CURRENT STUDIES IN BASIC SCIENCES, ENGINEERING AND TECHNOLOGY 2022

EDITORS

SABRI KOCER

OZGUR DUNDAR

ISRES Publishing

Current Studies in Basic Sciences, Engineering and Technology

Edited by

Sabri KOCER

Professor, Necmettin Erbakan University
Faculty of Engineering
Department of Computer Engineering, Konya, Türkiye

Ozgur DUNDAR

Asistant Professor, Necmettin Erbakan University
Faculty of Aeronautics and Astronautics
Department of Astronautical Engineering, Konya, Türkiye

ISRES Publishing

Current Studies in Basic Sciences, Engineering and Technology

Editors

Sabri KOCER
Ozgur DUNDAR

This book was typeset in 10/12 pt. Times New Roman, Italic, Bold and Bold Italic.

Cover image by Freepik

Current Studies in Basic Sciences, Engineering and Technology

Published by ISRES Publishing, International Society for Research in Education and Science (ISRES).

Includes bibliographical references and index.

ISBN

978-605-81654-2-7

Date of Issue

December, 2022

E-mail

isrespublishing@gmail.com

Address

Istanbul C. Cengaver S. No 2 Karatay/Konya/TÜRKİYE

www.isres.org

PREFACE

Current Studies in Basic Sciences, Engineering and Technology 2022 is published annually from selected articles invited by editors. This edition contains 16 articles in Technology, Engineering, Basic Sciences.

All submissions are reviewed by at least two international referees. The aim of the book is to provide readers with a scientific peer-reviewed publication in the field of basic sciences, engineering and technology. Current Studies in Basic Sciences, Engineering and Technology 2022, Published by ISRES Publishing This book contains new ideas in Electrical Electronics, Communications, Mechatronics, Software, Artificial Intelligence, Aviation and IOT engineering and technology.

We hope that the book will arouse curiosity about science and technology, be useful to new scientists, science readers and anyone who wants to learn the mystery of science, and contribute to the literature.

December, 2022

Sabri KOCER
Professor, Necmettin Erbakan University
Faculty of Engineering
Department of Computer Engineering, Konya, Türkiye
Email: skocer@erbakan.edu.tr

Ozgur DUNDAR
Asistant Professor, Necmettin Erbakan University
Faculty of Aeronautics and Astronautics
Department of Astronautical Engineering, Konya, Türkiye
Email: ozdundar@erbakan.edu.tr

IN THIS BOOK

In Chapter 1, Its power was recognized by everyone shortly after deep learning emerged. For robots to work in harmony with many different problems, it has required them to face real-world problems. Deep learning has become an indispensable part of robots in tasks such as object recognition and tracking and rapid decision-making in the real world. This study explains the place of deep learning in robotic vision.

In Chapter 2, Natural language processing is a field of science and engineering that deals with the design and implementation of computer systems that analyze, understand, interpret and produce a natural language. This subject, which was seen as a sub-field of artificial intelligence in the 1960s, is now accepted as a subject of computer science as a result of the successes of researchers and applications. Natural language processing is a wide field that can be studied in a wide variety of subjects such as engineering, economics, politics, sales and marketing, customer satisfaction, customer recommendation systems, disease detection in the health sector, and analysis of court decisions in the field of law. This study explains how natural language processing can be used in engineering.

In Chapter 3, Microstrip antenna structures are often prefered from lots of advantages in biomedical research fields. Especially, ease of design and cost of them increase the popularity in this field. The chapter prepared contains the aims of use for different cancer disease types and heart diseases. With this aim, the chapter content ise enriched by benefiting several references in the literature. The contributions of microstrip antenna structures to the biomedical fields are quite high. This multiplicity increases day by day and paves the way for new studies in the future. It is seen that the design richness of microstrip antenna structures is a mine that continues to be processed for the biomedical fields.

In Chapter 4, The main objectivity of this work is to provide a summary of the different machine learning techniques, applications and challenges along with the history of machine learning so far. This paper provides a description of the history of machine learning, including its origins and progression through the ages. In this book chapter: Machine learning and data analysis challenges, General machine learning model, Machine learning concepts and Machine learning algorithm are explained.

In Chapter 5, The power systems can be cited as the most important subsystem in spacecraft. Failure of the power subsystem can cause the space mission to fail, and is thus the cause of the majority of mission failures on early satellites. Because the most important factor that determines the life of a spacecraft is the active power systems they have. As systems developed and became more complex, the power requirement in spacecraft increased. While batteries and fuel cells were generally used in the first spacecraft to meet the energy needs, solar cells and nuclear power systems are used in today's space platforms in line with the needs of the mission area and mission platform. Today, with the emergence of deep space mission concepts such as space tourism and space mining, the need to develop new primary power systems that are more efficient and that will use different energy sources other than solar energy has emerged. For this reason, the field of space power systems has become a focus for researchers

In Chapter 6, Nowadays, the use of autonomous vehicles is becoming more and more widespread. Considering the autonomous levels, autonomous vehicles at certain levels are frequently seen in most vehicles in traffic. With the development of technology, it is predicted that there will be an upward increase in the number of autonomous vehicles in traffic in the future, as well as in autonomous driving levels. In this study, besides the historical developments of autonomous vehicles, predictions were made about the vehicles that will be found in traffic in the future. In addition, autonomous driving levels and the basic sensors used are mentioned.

In Chapter 7, The field of nonlinear optics offers many exciting opportunities for both fundamental research and technological application. As in other high-tech areas, such as microelectronics and genetic engineering, science and technology can be expected to share a vital interplay where advances on one front enable advances and present new challenges on the other. Improvement of nanotechnology and creation of new nano-materials with specific properties lead to a growing need to study the nonlinearity of the optical characteristics of the newly created materials. For example, optical devices based on nonlinear optical effects are becoming more and more widely used in practice today.

In Chapter 8, Images are one of the most significant channels that can be utilized to pass on a wide range of data. In general, images hold highly sensitive and private information that play an essential for a variety of applications, such as military, remote sensing, and medical imaging. Subsequently, as different kinds of information, the vast majority of us use the internet as an essential means of transferring image information from one point to another. It is important to protect sensitive and confidential data against unauthorized users in order to access and modify it in this context. Due to the necessity of security and the resource constraints, image encryption, and decryption techniques have been invented. The internet, interactive media, medical imaging, and military imaging frameworks all utilize image encryption. Each kind of media information has its interesting properties, like high pixel connection and redundancy. In recent years, researchers have developed various image encryption techniques to enhance security. In this study, it was conducted a comparative study of five encryption algorithms using color images obtained in the BSD500 database. The main purpose is to compare the performances of these algorithms by using different methods of image quality evaluation. The algorithms in the scope of the study include AES, DES, TDES, RSA and ECC. The first three algorithms are considered symmetric methods, while the other two algorithms are considered asymmetric encryption methods. In the experiment, some color images from BSD were used to test the performance of the algorithms. The efficiency of encoding was measured using some image evaluation methods and the results were presented in the study.

In Chapter 9, It leads to the efficient technical requirement for storing and transmitting data in the digital world. Therefore, it is important to minimize the size of the data stored or transmitted. Data compression ensures efficient use of available storage space and communication bandwidth. It is necessary to study data compression techniques and applications. The features, basic concepts and limitations of data compression techniques are explored in this book chapter.

In Chapter 10, Ancillary services in the electricity distribution grid is briefly explained and a concise review of the literature on ancillary services is provided. This chapter explains the current state of ancillary services, providing a list of suggestions for new roles of distribution SOs at the end.

In Chapter 11, With the introduction of the Internet into people's lives, it has become possible for a device anywhere in the world to communicate with another device. Internet technology has had a great impact especially on embedded system technologies (smartphones, watches, glasses and other low-power wearable and wearable devices). By making use of the sensors integrated into these embedded systems, it is possible to obtain many different data from the environment of the individual. These obtained data were collected, analyzed and analyzed in a center via wired or wireless internet. In this way, various information about the person who owns the device or the environment in which the device is located has been reached in a short time and the communication of objects with each other (IoT) has emerged. One of the areas where IoT is used the most is the field of health services. In the correct diagnosis, treatment and follow-up of diseases, the data to be obtained from the daily life of the patient outside the hospital is of great importance. The best way to obtain this data is to use IoT wearable or wearable healthcare devices.

In Chapter 12, In this section, the swarm concept and swarm UAVs are discussed. Swarm behavior; It is the behavior that causes living things with a colony structure such as birds, fish and some bacteria to clump together and move or move with the members of the swarm. The ability of living things to solve problems that they cannot solve individually, as a swarm, is called swarm intelligence. Swarm robots; It is a system that works for a certain purpose by using swarm intelligence software and swarm communication methods in coordination with robots that can work alone. In this section, the swarm concept and swarm UAVs are discussed. The article describes the communication and coordinated operation of swarm UAVs, and finally the use of UAVs in battlefields.

In Chapter 13, The Internet of Things (IoT) is without a doubt one of the most revolutionary technologies that has changed the way people live by the end of the twentieth century. Over the last three decades, IoT has expanded and become deeply embedded in a variety of sectors ranging from consumer applications to military applications, health, education, information technology, automobiles, and so on. Because traditional India is primarily based on agriculture, this section focuses on the impact and application of the Internet of Things in transforming farming methods into smart and indigenous farming. It obviously outlines the contribution of IoT to the modernization of agricultural techniques through the use of automated equipment, keeping abreast of monitoring techniques, and addressing issues of environmental disasters and calamities that will assist farmers and agriculturists in combating and being empowered.

In Chapter 14, Today, wearable technologies are used in many areas. They are frequently used in applications such as monitoring the physical activities of people, patients, and monitoring the life signals of elderly people. In this chapter, a review about one of the most important components of a wearable device i.e. antennas used to send and receive signals is performed. After a brief history and their evolution in time, commonly used fabrication techniques and materials are presented. Textile and polymer-based antennas are particularly shown to be widely used due to their high flexibility and strength. A more recent approach based on the laser direct structuring technique is also presented for manufacturing wearable antennas. Recent works related to wearable antennas and particular applications are explored in this chapter.

In Chapter 15, Along with the developing technology, digitalization of all segments of society, and the change of business models, large amounts of digital data are produced regularly or irregularly. In this study, information and methods on HDD and optical disk media where data is stored and how to recover inaccessible data in these media are examined in practice and various solutions are presented. In addition, the steps to be taken to avoid data loss and the steps to be followed in case of data loss are explained in detail.

In Chapter 16, With the digitalization of all segments of society with the developing technology and the change of business models, a large amount of digital data is produced regularly or irregularly. In this study, information and methods on how to recover inaccessible data on devices using Flash memory as a data recording unit and how to recover inaccessible data on these devices have been studied in practice and various solutions have been presented. In addition, the actions to be taken to avoid data loss and the steps to be followed in case of data loss are explained in detail.

Managing Editors

Sabri KOCER, He is a professor and senior researcher in the Department of Computer Engineering at Konya Necmettin Erbakan University. He completed his doctorate at the Department of Electronic Computer Education at Gazi University. He has held academic positions at different Universities in Turkey and has made great contributions to the on-going technological academic development in the region. His main research interests are biomedical engineering, artificial intelligence, optimization, and machine learning and microcontroller applications. In addition, he has many academic works such as books, articles, essays, posters published in many national and international journals.
E-mail: skocer@erbakan.edu.tr, **ORCID:** 0000-0002-4849-747X

Ozgur DUNDAR, PhD, works at Necmettin Erbakan University, Department of Astronautical Engineering. He graduated from the Electrical and Electronics Engineering Department of Selcuk University. He worked as an Automation Engineer for a while. His master's and doctorate degrees are from Selçuk University, Institute of Science and Technology, Department of Electrical and Electronics Engineering. Special fields of study are Automation, Robotic, Communication, Electromagnetic and Micro Strip Patch Antenna designs.
Email: ozdundar@erbakan.edu.tr , **ORCID:** 0000-0002-4142-4446

CONTENTS

CONTRIBUTORS

Kursad UCAR
Department of Electrical & Electronics Engineering,
Selcuk University, Konya, Türkiye

Hasan Erdinç KOCER
Department of Electrical & Electronics Engineering,
Selcuk University, Konya, Türkiye

Seda YILDIRIM
Department of Computer Programming,
KTO Karatay University, Konya, Türkiye

Rabia TOPRAK
Department of Electrical - Electronics Engineering,
Karamanoglu Mehmetbey University, Karaman, Türkiye

Seemant TIWARI
Department of Electrical Engineering,
Southern Taiwan University of Science and Technology

Omer Faruk KESER
Department of Astronautical Engineering,
Necmettin Erbakan University, Konya, Türkiye

Bayram BULAT
Staff Officer, Operations Plans & CIS Division,
NATO NAEW&C Force Headquarters Geilenkirchen, Germany

Tarik UNLER
Department of Aviation Electrics and Electronics,
Necmettin Erbakan University, Konya, Türkiye

Ferit ARTKIN
Department of Machinery and Metal Technologies,
Kocaeli University, Kocaeli, Türkiye

Mehmet HACIBEYOĞLU
Department of Computer Engineering,
Necmettin Erbakan University, Konya, Türkiye

Bashir Sheikh Abdullahi JAMA
Department of Computer Engineering,
Necmettin Erbakan University, Konya, Türkiye

Sabri KOCER
Department of Computer Engineering,
Necmettin Erbakan University, Konya, Türkiye

Bashir Sheikh Abdullahi JAMA
Department Of Computer Science And Technology,
Simad University, Mogadishu, Somali

Yasin ER
System Administrator, Department of Meteorological Data Processing,
Turkish State Meteorological Service, Ankara, Türkiye

Beyza DUNDAR
Deparment of Computer Engineering (Undergraduate)
KTO Karatay University, Konya, Türkiye

Mert BAYER
Meram Electricity Distribution Company, Konya, Türkiye

Harun KOROGLU
Meram Electricity Distribution Company, Konya, Türkiye

Mustafa Emre SAFAK
Electricity Distribution Company, Sakarya, Türkiye

Burak ALTUN
Meram Electricity Distribution Company, Konya, Türkiye

Hüseyin Oktay ALTUN
Department of Electrical and Electronics Engineering,
KTO Karatay University, Konya, Türkiye

Hakan OCAL
Department of Computer Engineering,
Bartin University, Bartin, Türkiye

Ozgur DUNDAR
Department of Astronautical Engineering,
Necmettin Erbakan University, Konya, Türkiye

Hasan Berkant DUNDAR
Department of Electrical and Electronics Engineering,
Istanbul Medipol University

Suleyman KAPLAN
Department of Astronautical Engineering,
Necmettin Erbakan University, Konya, Türkiye

B Ravi CHANDRA
Department of Electronics and Communication,
G Pullaiah college of Engineering and Technology, Kurnool,
Research Scholar in Lovely Professional University, Punjab, India

I Sharath CHANDRA
Department of Electronics and Communication
Matrusri Engineering College, Hyderabad, India

Krishan KUMAR
Department of Electronics and Communication Engineering,
Lovely Professional University, Punjab, India

Hulusi ACIKGOZ
Department of Electrical and Electronics Engineering,
KTO Karatay University, Konya, Türkiye

Ilayda CAN
Department of Electrical and Electronics Engineering,
KTO Karatay University, Konya, Türkiye

Abdulkadir ÖZDEMİR
Cyber Crime Branch Office, Konya, Türkiye

Şaban GÜLCÜ
Department of Computer Engineering,
Necmettin Erbakan University, Konya, Türkiye

Robotic Vision with Deep Learning

Kursad UCAR
Selcuk University

Hasan Erdinc KOCER
Selcuk University

Introduction

A robot is a mechanical device that can automatically perform a complex sequence of actions and is usually programmable by a computer. Internal or external controllers can control the robots. Robots can be created to resemble humans or animals, but most robots are designed to perform a task regardless of their appearance (Ruiz-del-Solar et al., 2018). As the same time robots are controlled fully or semi-autonomously in terms of being controlled.

Robotics is the field of technology that deals with the design, manufacture, operation, and application of robots (Samuel E. Anthony, 2016). Robots are machines that can replace humans in hazardous environments or production processes. For this reason, they are also more similar to humans in terms of appearance, behavior, or cognition. In addition to humans, there are also robots inspired by nature, such as spiders and worm robots. With these robots, a new robotic field was born under the name of soft robotics. Robotics researchers are trying to use robots in many different tasks. Some tasks can be summarized as acting like a human, garbage sorting, automatic direction finding and navigation, collaborative automation, debris search and rescue, and drawing pictures. Deep learning will play an active role in all these tasks. The difficulties in this and similar tasks can be briefly summarized as follows.

1) Advanced manipulation. Grasping of an object may differ according to the object. It is very hard to grasp things with complex shapes, which can break or change shape when squeezed. Likewise, tasks required by the grasping process, such as painting, opening a door, or using a tool, are of interest to researchers.
2) Learning dynamics: It has always been a problem for robots, as changes force them to step outside of the algorithm that controls them. Monitoring and capturing changes requires human expertise and is time-consuming. It is significant to adapt to new dynamics in order to solve problems such as grasping objects that are different from the ones taught, and moving in a new environment.
3) Sensor fusion & dimensionality reduction. Robots need to be fast to be used in real-time applications. For this reason, obtaining meaningful data from low-quality images also plays a critical role. Thanks to deep learning, the data can be represented and used with data of lower dimensions by cleaning it from data that

will not be used.

4) Interpreting and anticipating human actions. Robots can also interact with humans. In emergency rescues, robots that are put into the service of people, such as elderly / sick care, should work in a way that does not harm people. In such tasks, robots must also be able to predict human movements.
5) Advanced object recognition. Robots deal with a wide variety of objects. Object recognition and classification are some of the things that deep learning does best. Thanks to deep learning, it has become possible to recognize objects and obtain their properties, to make necessary path calculations, to obtain critical properties (slippery, sharpness, etc.) and surface characteristics of objects.

A robot perceives the world it is in (simulation or real) with its different sensors, creates a model of the world and updates this model over time when necessary. However, in order for the robot to perform a task, it is essential to make decisions, plan the necessary actions and implement these actions. While performing these actions, it is highly adhered to the program written by the human. Keeping the robots working without being affected by environmental or parameter changes is still not an insurmountable problem. Machine learning is a useful tool to overcome the mentioned and similar challenges. In particular, deep learning has become well known for its outstanding success in many fields.

Robot Kinematics

A robot is structurally created with joints that perform translational (prismatic-P) and rotational (revolute-R) motions, as shown in Figure 1, and joints that connect these joints to each other. The displacement due to the rotational movement and the translation due to the displacement between the joints are called the joint angle and the joint offset, respectively. Robots are positioned by controlling the joints separately. Thus, the positioning of the endpoint is ensured. Movement in Cartesian coordinates is achieved by obtaining the joint coordinates that will provide the current position using kinematic equations.

Kinematics is a science of motion that deals with the motion of objects. The masses of the objects and the forces acting on the object are not taken into account in the movements discussed. In robot kinematics, a relationship is defined between the end-effector and the joints that provide the movements.

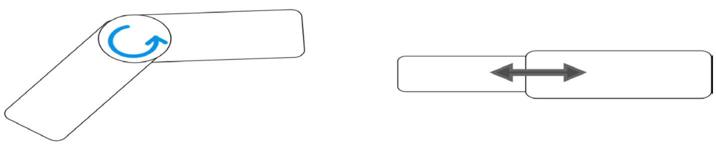


Figure 1. Prismatic and revolute joints

The position and orientation of each joint of a robot is determined relative to the previous or the next joint. This consecutive relationship is called an open kinematic chain.

Transformation matrices are used to create an open kinematic chain. The transformation matrix contains the robot's position and orientation information. A homogeneous transformation matrix is created for each joint. The degree of freedom (DOF) of the robot determines the number of these matrices created. Six degrees of freedom are sufficient to reach any point in three-dimensional space with any orientation. For this reason, industrial robots are generally produced with 6 DOF. Of course, there are robots with a DOF of less than six which can be used for simpler tasks. On the other hand, redundancy occurs in robots with more than six DOF. In this case, more than one joint is scanning the same area. This phenomenon is seen in special-purpose robots designed for special conditions.

One of the most important stages of robot studies is the kinematic model of robots. Many methods have been developed for this purpose. The solution to kinematic problems can be realized in two different spaces, three and four-dimensional. Since the kinematic model is extracted using transformation operators such as matrices or vectors in three-dimensional space, this method is called point transformation, and since linear vectors and transformation operators are used in four-dimensional space, this method is called linear transformation method (Aspragathos & Dimitros, 1998).

Robot Operating System

Robot Operating System is software that allows controlling robots. Although it is called an operating system, it is not an operating system. ROS is open-source interface software that creates human-robot communication.

ROS is widely used and preferred software in robots. The robot evaluates the data it receives from the environment via sensors and sends it as a command to the robot as a result. With the help of sensors in the controlled robot, data collected from the outside world such as sound and images are transferred to the computer via ROS. The data is processed according to the algorithm and usage area in the computer. The result of the processed data is sent back to the robot as a command. While performing these operations, it works in publishing/subscriber logic. The topic and messages are used to provide communication between the computer and the robot. In addition to being open-source, it is possible to use it with more than one programming language. In other words, it allows using different languages (Java, Lisp, C++, Python) on a robot.

With ROS, it is aimed for the user to develop software that fulfills the desired task by making minimum changes to the code and using it according to the needs. In addition, the user can quickly run the previously developed algorithms in the download-use format and thus make performance comparisons and impressions.

Robotics is a difficult and costly field to implement. For this reason, it may be important to test the application in a simulation environment before it is implemented in the real world. In the light of this importance, there are simulation environments that are fully compatible with ROS. The most common of these simulations are Gazebo and Rviz environments. With the libraries and ready-made 2D and 3D objects in Gazebo, it offers a wide working area to the developers. Data packaged in ROS can be easily simulated and examined in 3D in the Gazebo environment.

While developing robot software, it can be used in interfaces such as Orocos, Carmen, Microsoft Robotics Studio, and YARP apart from ROS. The advantages of ROS over others can be summarized as follows.

1) It publishes a different issue (publish) in each transaction. Thus, error management can be made easier.
2) It is not dependent on a programming language. Among the languages it supports, there are many languages such as Java, C++, Python, and Lisp, which are among the most widely used programming languages. It is also platform-independent. Even a ROS node can be programmed in different languages according to need or teamwork. It comes to the fore in the realization of steps such as teamwork, remote work, or performance optimization by working integrated during compilation and study.
3) The software is open source. Being open-source code does not remain in a limited framework; it allows communicating with many people through the community. With the tools available in the community, studies can be developed in the desired direction.
4) Its modular structure is another advantage. While a problem in other robot control software causes the main node to stop, the node structure used here prevents this.

ROS is created from nodes as a system. These nodes communicate with each other and use the publish-subscribe messaging model. Units that can operate here are called nodes. For example, while one node performs the data collection process with the sensor from the outside world, another node performs the task of processing this data. Another node creates a result. The communication and operation of all these nodes with each other are carried out synchronously. To provide this, a server called ROS-Master creates the connection between the nodes. This Master server is an XML-RPC-based application.

Robotic Vision

With the emergence of deep learning, pattern recognition has become powerful tool that raises the direction of studies in the fields of machine learning and computer vision. Thanks to deep learning, many problems have been solved or reached better solution (Ucar & Kocer, 2022).. It has also become a research area and improvements have been made in various algorithms, data sets, and computational power. Such a powerful tool has attracted attention in all areas. Undoubtedly, robotic vision is among them. The deep learning algorithms used in robotic vision were transferred from computer vision. Robotic vision is the addition of action to the computer vision result. As shown in Figure 2, if the output of the image inputs is data, it is computer vision. However, if the image inputs are transformed into action and a job is performed, robotic vision is realized. However, it is not possible to transfer every algorithm or method. Robotic vision and computer vision are separated because of the two main differs from each other on several issues.

1) Computer vision can generally be operated with hardware such as Graphical Processing Unit (GPU) that offers high computing power, while robotic vision is difficult to perform with such high hardware.
2) Robotic vision is run in real-time. For this reason, computer vision algorithms and methods can be transferred, which can meet the requirements of robotic vision such as real-time and limited computing power.

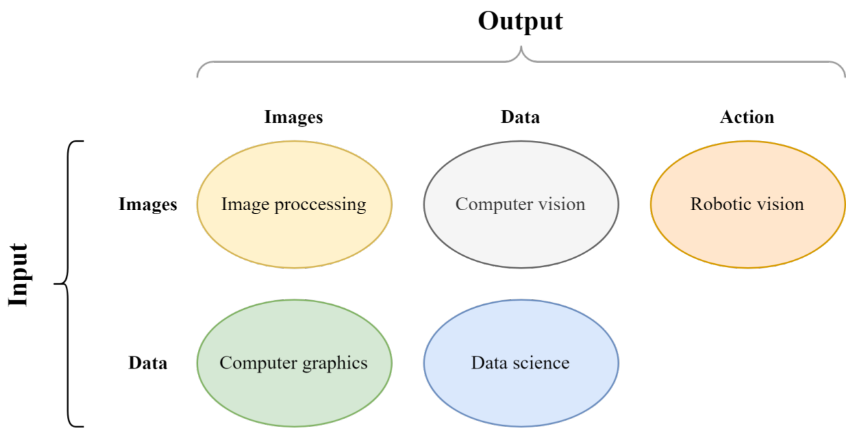


Figure 2. Position of robotic vision according to computer vision

An example of a robotic vision application is shown in Figure 3. General view of robotic vision and grasping mechanism is shown in Figure 3a. When the objects on the conveyor belt (Figure 3b) will be separated with the robot arm, the objects are viewed with a camera as in Figure 3c. By running the deep network on the displayed objects, necessary information for gripping such as detection, location, and speed of the objects is obtained. According to the information obtained, the robot arm positions itself as in Figure 3d. Then, the robot arm waits for the objects to come to the grip area as in Figure 3e. When the robot arm realizes the grip, it separates the desired object (Figure 3f). Since all these processes have to be carried out in a short time, the need for fast-running and moving systems is huge.

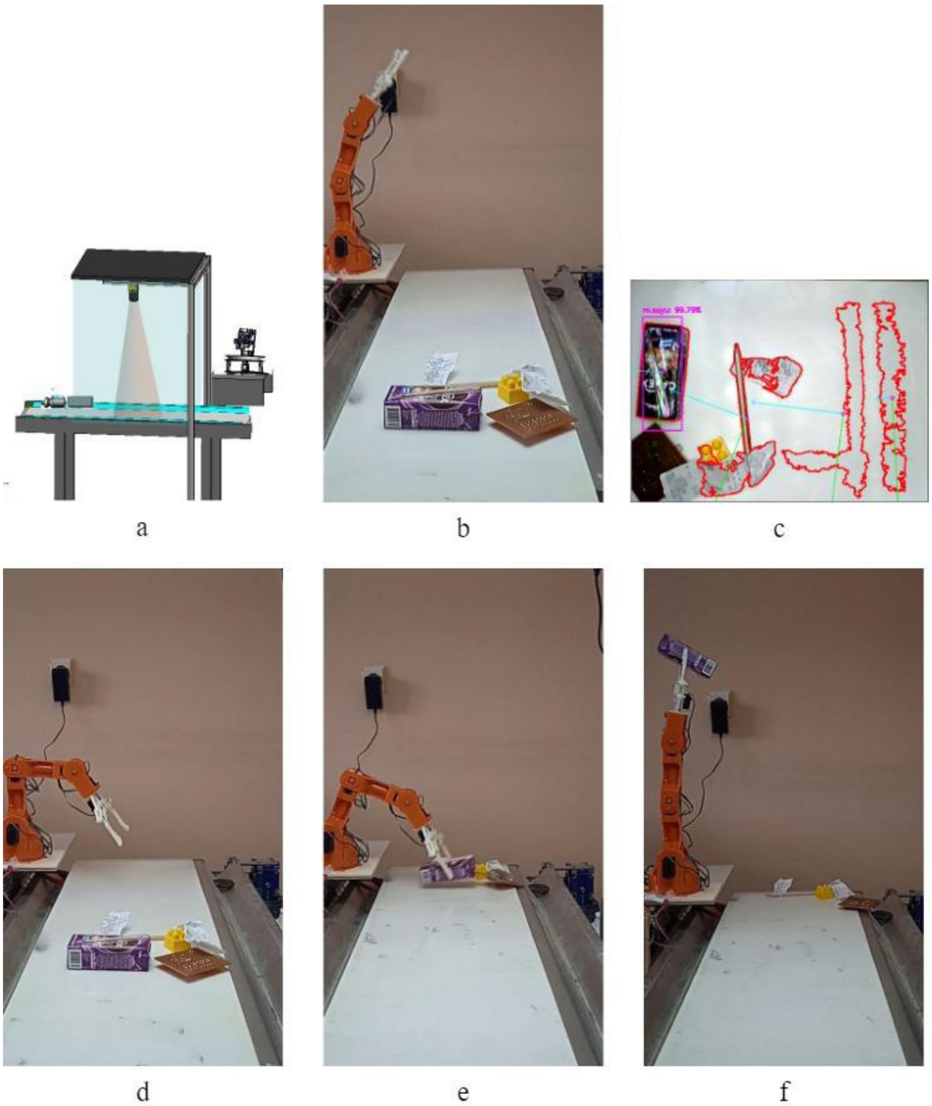


Figure 3. a) General view of robotic vision & grasping mechanism, b) Robot arm and objects to be separated, c) Viewing objects and obtaining deep network related data, d) Positioning of the robot arm according to the deep learning result, e) The robot arm grasps the desired object, f) Separation of the grasped object

Deep Learning

Deep learning seeks solutions using a more complex architecture for problems where artificial intelligence algorithms are insufficient. For this reason, it needs higher computational power and data set. In the early years, it took some time to realize its importance and power, since it was not possible to meet these needs. Many different architectures have been developed as it has become easier to train and use.

Convolutional Neural Networks

CNN is an architecture of DNN. This structure, inspired by the visual system of mammals, has local connectivity, pooling, shared weight, and the use of many layers (Bengio & LeCun, 2007). The large numbers of hidden layers that make up CNN are used in 3 different tasks. Convolutional layers with linear filtering are the layers where the features are obtained. The non-linearity of the network is gained with the activation function applied after the convolution layer. A pooling layer is used between the convolution layers. Pooling layers serve to reduce the data size to reduce the processing load. The use of the pooling layer should be careful, as there will be data loss along with the data size reduction. Maximum and average pooling are the most commonly used pooling layers. After multiple layers of convolution and pooling, an input-based decision is required. Decision-making is done in the fully connected layer. The operations of these layers are shown in Figure 4.

The convolution, pooling, and fully connected layers that make up a CNN, the number and size of filters vary according to the application. However, since there must be at least one connection for each data in each layer, it needs quite high computational power. However, since most operations are done locally in CNN, it is suitable to work with GPUs.

Fully connected layers are an important layer for converting 2D data to 1D. This layer is similar to the classical ANN. Also here the final classification is carried out.

CNNs can be created in different structures by using the different number of layers, filter size, and number, dropout, pooling, and activation functions. Many architectures have been created according to these differences and attract attention with their performance (AlexNet (Krizhevsky et al., 2012); VGG (Simonyan & Zisserman, 2014); GoogleNet (Szegedy et al., 2015); ResNet (He et al., 2016); DenseNet (Huang et al., 2017)).

Although CNN is successful in classification, it cannot produce an output for object tracking, recognition, and tracking tasks. In such tasks, the target usually covers a small area in the images and this area needs to be detected. The window scrolling method can be used for this task, but its slow operation makes it impossible to use for robotic tasks. To eliminate this disadvantage, it is suggested to use region proposals containing objects. Fast RCNN (Girshick, 2015) is one of these methods. In the light of later studies, SSD (Liu et al., 2016) and YOLO (Redmon et al., 2016) emerged. YOLO has been an algorithm that has been studied a lot. It is a very effective method as it can perform object classification and location detection with a single CNN. It also allows trade-offs between accuracy and speed. The most striking feature is that it offers high speed and accuracy even with low hardware, thanks to different architectures and developments.

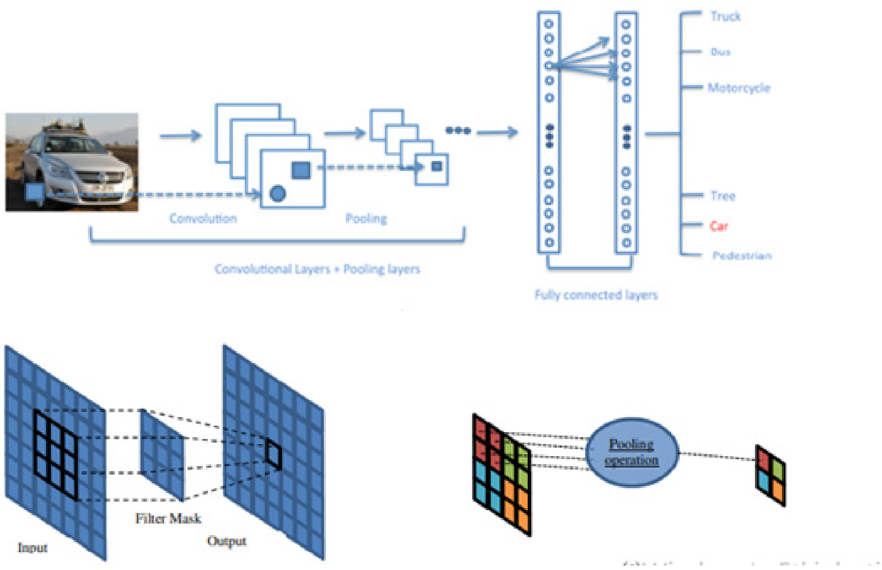


Figure 4. The operations of CNN layers (Guo et al., 2016)

Fully Convolutional Networks

The difference between CNN is that there is no fully connected layer in Fully CN. For this reason, the output in the last convolution layer is the network output. In these networks, the input and output also consist of 2D or 3D tensors. Among the most common uses are dimensionality reductions, regression, or semantic segmentation.

One type in which FCNs are used is Auto Encoders (Hinton & Salakhutdinov, 2006). An encoder and decoder are created using two FCNs. The encoder reduces the size of the input image and creates a tensor output. In the decoder, this tensor regenerates the image by performing inverse operations. Auto encoder's training is also carried out for this purpose. For this reason, it can be used in noise removal and segmentation tasks.

Deep Reinforcement Learning

Reinforcement Learning consists of a set of techniques that learn by rewarding agents cumulatively for their success in their tasks. Here, agents interact with their environment to take an action and receive rewards (Li, 2017).

Deep Learning in Robotic Vision

Robotic vision studies using deep learning are divided into three different categories (Ruiz-del-Solar et al., 2018): object detection and categorization, object grasping and manipulation, scene representation, and classification. Although many robotic vision studies have been carried out in these areas, it is not known whether they can meet the real robotic vision requirements as some of them are carried out in simulation environments.

- Object Detection and Categorization: One of the most fundamental steps of robotic applications is object recognition and classification. Thus, robots can interact with the real world. As a result of the interaction, the robots can detect the surrounding targets and, if necessary, avoid obstacles such as walls and foreign objects that will prevent them from working. Of course, this contribution of deep learning to robots becomes limited by the real-time requirement. As a result, applications are carried out around task-based methods. With task-based applications, it is aimed to work in real-time by using the most useful, fast, and few features of algorithms with a limited number of objects. In addition, some applications have increased the speed by using smaller deep networks.

- Scene Representation and Classification: In recent years, mobile robots have become quite common. With these robots, which have many different tasks, each scene is different and can be important in terms of their duties. Even outside the task, robots can understand the movement route and location information through images. For this, the features that were previously extracted manually have become easy with the help of deep learning.
- Object Grasping and Manipulation: Grasping is a very important task in robotic applications. Both gripper and object detection are areas where deep learning is frequently used. In this context, tasks such as grasping, manipulation, and visual servoing are performed.
- Although the task of grasping objects is very easy for humans, the situation for robots is quite complex. With the help of deep learning, robots can gain the ability to grasp objects in a fixed shape. However, as the shapes, sizes, positions, and directions of objects change, robots face a difficult situation. In such cases, robots undergo long-term training with trial-and-error methods.

Deep Learning in Robotic Vision

Most jobs involving robots also involve object grasping. For this reason, a gripper is mounted on the end of the robots. Grippers are mechanical, pneumatic, or electrical equipment used to carry or hold a part. The selection of this critical part can be chosen according to the task to be performed by the robot. The most widely used robotic grippers are; 1) Parallel Motion Two-Jaw Gripper, 2) Three-Jaw Gripper, 3) Bellows Gripper, 4) Collet and Expanding Mandrel Grippers, 5) O-Ring Grippers, 6) Needle Grippers, 7) Multi-Finger and Adaptive Grippers, 8) Vacuum Cup or Vacuum Cup Arrays, Electromagnets, and Electrostatic Force Grippers.

Deep Reinforcement Learning

A frequently preferred area of deep learning is grasping. Grasp can be defined as the partial or complete closure of an object with an end effector. A gripping problem can be solved by correctly positioning the robot. For this, it is necessary to know the features of the object to be grasped, such as shape, size, and location. According to these features, possible gripper configurations should also be investigated (Bohg et al., 2013; Kumra & Kanan, 2017).

Robots need as much information as possible about the targets they need to grasp.

Presenting this information by people who are experts in the subject is both a time-consuming and a very difficult approach.

For humans, grasping is an unthinking and automatic action. However, the same is not true for robots. Comprehension of objects is a very challenging problem due to factors such as an unlimited variety of objects, different physical properties of objects, and different viewing angles. All these difficulties must be overcome to achieve successful grasping. When we want to grasp an object, we see the object and determine its position. Then we reach out and hold the object. The same procedures apply to robots. However, since robots are not as smart as humans, they cannot practice grasping that much. In addition, robotic grippers can even have only two fingers. After detecting the target on the images, the robots obtain their position in the image. It should estimate the location in the real world based on the location in the image. It must position itself and calculate grip points by comparing its position with the position in the real world. Then he should reach and grasp the object without being hindered by the surrounding obstacles.

References

Aspragathos, N. A. & Dimitros, J. K. (1998). A comparative study of three methods for robot kinematics. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, *28*(2), 135–145.

Bengio, Y. & LeCun, Y. (2007). Scaling learning algorithms towards AI. *Large-Scale Kernel Machines*, *34*(5), 1–41.

Bohg, J., Morales, A., Asfour, T. & Kragic, D. (2013). Data-driven grasp synthesis—a survey. *IEEE Transactions on Robotics*, *30*(2), 289–309.

Girshick, R. (2015). Fast r-cnn. *Proceedings of the IEEE International Conference on Computer Vision*, 1440–1448.

Guo, Y., Liu, Y., Oerlemans, A., Lao, S., Wu, S. & Lew, M. S. (2016). Deep learning for visual understanding: A review. *Neurocomputing*, *187*, 27–48.

He, K., Zhang, X., Ren, S. & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770–778.

Hinton, G. E. & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science*, *313*(5786), 504–507.

Huang, G., Liu, Z., Van Der Maaten, L. & Weinberger, K. Q. (2017). Densely connected convolutional networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 4700–4708.

Krizhevsky, A., Sutskever, I. & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, *25*.

Kumra, S. & Kanan, C. (2017). Robotic grasp detection using deep convolutional neural networks. *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 769–776.

Li, Y. (2017). Deep reinforcement learning: An overview. *ArXiv Preprint ArXiv:1701.07274*.

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y. & Berg, A. C. (2016). Ssd: Single shot multibox detector. *European Conference on Computer Vision*, 21–37.

Redmon, J., Divvala, S., Girshick, R. & Farhadi, A. (2016). You only look once: Unified, real-time object detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 779–788.

Ruiz-del-Solar, J., Loncomilla, P. & Soto, N. (2018). *A Survey on Deep Learning Methods for Robot Vision*. 1–43. http://arxiv.org/abs/1803.10862

Simonyan, K. & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *ArXiv Preprint ArXiv:1409.1556*.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V. & Rabinovich, A. (2015). Going deeper with convolutions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1–9.

Ucr, K., & Kocer, H. E. (2022). Determination of Angular Status and Dimensional Properties of Objects for Grasping with Robot Arm. IEEE Latin America Transactions, 100(XXX). Retrieved from https://latamt.ieeer9.org/index.php/transactions/article/view/6999

About the Authors

Kursad UCAR received the Master degree from the Department of Electric–Electronic Engineering, Institute of Science, Selcuk University, Konya, Türkiye (2018). He continues his Ph. at the same university. He has been a research assistant at Selçuk University since 2016. His research interests include image processing, robotics, automation and control, and machine learning.

E-mail: kucar@selcuk.edu.tr, **ORCID:** 0000-0001-5521-2447

Hasan Erdinc KOCER received the Ph.D. degree from the Department of Electric–Electronic Engineering, Institute of Science, Selcuk University, Konya, Türkiye (2007), where he has been an associate professor since 2018. His research interests are in machine vision and pattern recognition, industrial automation on camera, medical image processing, and automation and control.

E-mail: ekocer@selcuk.edu.tr, **ORCID:** 0000-0002-0799-2140

Similarity Index

The similarity index obtained from the plagiarism software for this book chapter is 3%

To Cite This Chapter:

Ucar K, Kocer H. E. (2022). Robotic Vision with Deep Learning. In S. Kocer & O. Dundar (Eds .), *Current Studies in Basic Sciences Engineering and Technology* (pp. 1–11). ISRES Publishing.

Natural Language Processing Applications in Engineering

Seda YILDIRIM
KTO Karatay University

Introduction

Natural language processing (NLP) is a subfield of linguistics, computer science, and artificial intelligence. NLP related to the interactions between computers and human language, how to program computers to process and analyze big amounts of natural language data ("Natural Language Processing", 2022).

NLP brings together theories, methods and technologies developed in many different fields for example artificial intelligence, formal language theory, theoretical linguistics, and computer-aided linguistics. The purpose of NLP is to examine problems in automatically generating and understanding natural languages (Young et al., 2018). NLP helps ensure human-computer interaction by processing human-generated sounds and texts.

NLP are theoretically motivated computational techniques for automated analysis and representation of human language (Cambria et al., 2014). In an environment where people of all ages can access social media, the amount of data generated continues to increase day by day. Naturally generated data by humans is not in a condition to be processed directly. Therefore, many areas need to work together to make sense of data and use it efficiently to provide human-machine communication.

The main objectives of research in the field of NLP are as follows:

- Better understanding of the function and structure of natural languages
- Using natural language as an interface between computers and humans and facilitating communication between computer and human.
- Making language translation with computer

To understand the structure of natural languages, it is necessary to make a detailed analysis of natural language and to extract its mathematics. Natural Language Processing aims to understand or reproduce the canonical structure of natural languages by analyzing them.

History of NLP

The first natural language processing studies started in 1950 with the article Computer Machines and Intelligence published by Alan Turing. In this article, he made a definition that includes automatic interpretation and natural language generation known as the Turing Test ("Natural Language Processing", 2022).

NLP history consists of 3 stages:

- Symbolic NLP (1950s-early 1990s): Rule-based emulation of Natural Language understanding and Generation Labeling operations made according to grammatical rules. They work according to the grammatical structures of languages. For example, ‘x’ is an unknown word in English. If there is a sentence in the form of Adverb + x + Noun, the word x is an adjective.

- Statistical NLP (1990s–2010s): Statistics-based methods like bag-of-words and n-grams became popular, also thanks to increased data availability from the Internet.
- Neural NLP (present): Due to the successes achieved in language modeling and parsing studies (Goldberg, 2016, Goodfellow et al., 2016, Jozefowicz et al., 2016, Choe & Charniak, 2016, Vinyals et al., 2014), machine learning techniques such as representational learning and deep neural learning have become widespread in natural language processing since 2010.

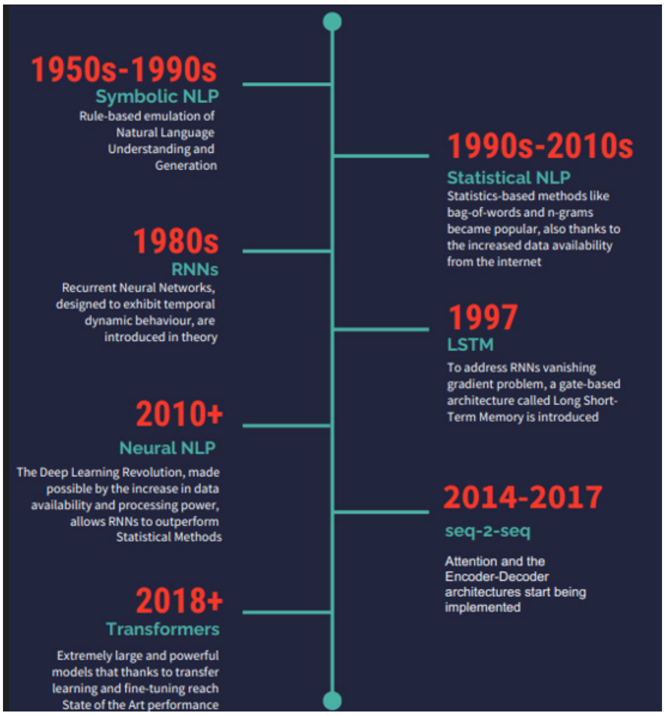


Figure 1. NLP brief history (“Natural Language Processing History”, 2022)

Methods

Statistical-Based Methods

The first of the commonly used methods in natural language processing is *statistical-based methods*. Language modeling is the job of finding statistical distributions that capture patterns in natural language. A statistical language model is built on the weights or probabilities of terms. Word representation methods are divided into two as *Frequency Based Representation* and *Prediction Based Representation*. Frequency-based word representations, which are defined as more traditional methods, are focused on the principle of detecting the words in the documents and the frequencies of these words.

Bag of Words Method

The most preferred method developed based on *frequency-based word representation* is the Bag of Words method. According to this method, each sentence in the document is divided into unique words and converted into a unique word-sized matrix. While the columns of the matrix consist of the words in the document (N), the rows consist of the number of the document (D). As a result, the entire corpus is represented as a DxN matrix.

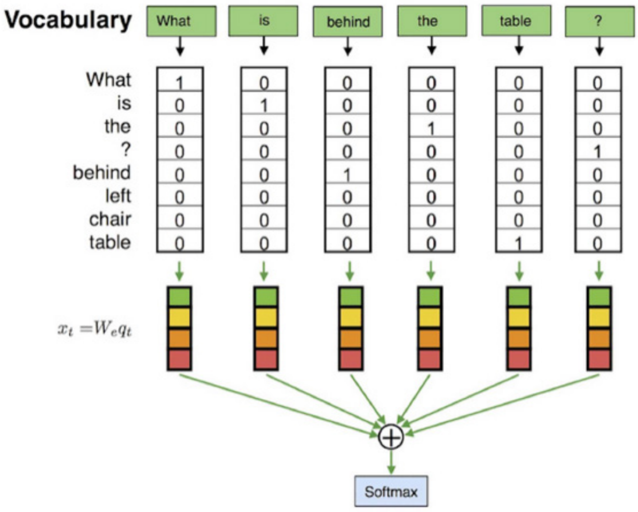


Figure 2. Bag of words example (Malinowski & Fritz, 2016)

Frequency-based methods have two main disadvantages (Karcı & Aydoğan, 2019)

- First, because there are unique words in the rows and/or columns according to the method used, matrices in the size of the word count occur, and since the value of most of them will be 0, a sparse matrix is formed. This causes the matrix to take up space in memory.
- Another important disadvantage is the inability to detect semantic closeness between words. In other words, if a word always comes after a word in a text, the link between it cannot be determined. For example, let's examine the sentence "Ali came to school". When the words "Ali" and "to school" come, the future of the words "came" or "went" should be deduced from the text.

Word2Vec Method

One of the word representation methods, the *prediction-based word representation method*, is the Word2Vec model, which was developed in 2013 and is based on the principle of training words with an artificial neural network. The Word2Vec model consists of two different algorithms named *CBOW (Continuous Bag of Words)* and *Skip-gram.*

The main differences of these two models are based on the methods of taking inputs and outputs. Word2Vec method, takes unique words in a large corpus as input and a matrix with a specified representation vector size is created. This method scans a sentence in the document each time by scrolling through a structure called a window and produces a vector consisting of many dimensions according to the target word in the window (Karci & Aydogan, 2019).

CBOW Model (Continuous Bag of Words)

In the CBOW model, words that are not in the window size center are taken as input and the words in the center are tried to be estimated as output. This process continues until the end of the sentence. This situation is tried to be shown in Figure 3. The value indicated by $w_{(t}$ is the output value in the center of the sentence and desired to be estimated, while the values indicated by $w_{(t-2)}$…..$w_{(t+2)}$ are off-center output values according to the preferred window size (window size).

Example Sentence: "Artificial intelligence is the new electricity."

The CBOW model, which takes this sentence as input and window size = 1, works as follows. First, the word "artificial" is placed in the center of the window, then the word "artificial" in the center of the window is tried to be predicted with the neural network model by taking the 1 word on the right and the left as separate input (window size=1). Then window 1 is shifted to the right, this time with the word "intelligence" in the center of the window. Table 1 shows the step-by-step operation of the sentence according to the CBOW model. The words written in red represent the output, and the words in blue represent the input (Karcı & Aydoğan, 2019).

Table 1. How the CBOW model works (Karcı & Aydoğan, 2019)

	Sentence (window size=1)	Input	Output
1	Artificial intelligence is the new electricity.	(intelligence)	(artificial)
2	Artificial intelligence is the new electricity.	(artificial) (new)	(Intelligence) (Intelligence)
3	Artificial intelligence is the new electricity.	(Intelligence) (electricity)	(new) (new)
4	Artificial intelligence is the new electricity.	(new)	(electricity)

Skip- Gram Model

In the Skip-Gram Model, it has a reverse operation of the CBOW model. The word in the center is taken as input and the words that are not in the center are tried to be guessed as output. This process continues until the sentence ends. This situation is shown in Figure 4 (Karci & Aydogan, 2019).

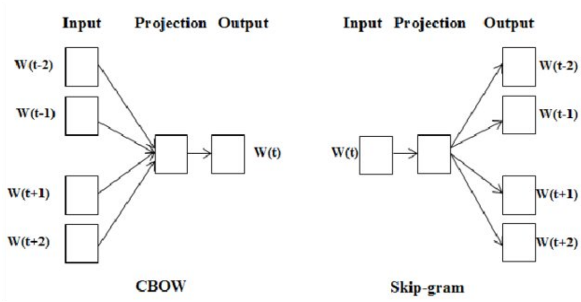


Figure 3. Structure of Skip – Gram and CBOW (Suleiman & Awajan, 2019)

Skip-gram model uses a neural network to create word representations. A common way to represent a word in machine language is to encode a word into an array of characters or a string. However, an array of characters does not carry any meaning. The Skip-gram model will create a vector, an element of vector space to represent each word, called a word vector. A word vector is a position relative to the origin in a graph. Words with similar meaning will have their word vectors closer together whereas the words with different meaning will be far apart. Interestingly, the word vectors encode the semantic relations through linear translations.

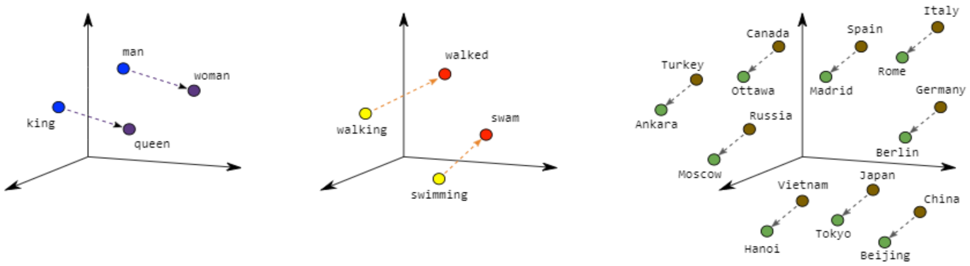


Figure 4. Word Embeddings Representation (Embeddings, 2022)

Artificial intelligence is the new electricity.

We will create (center, context) pairs of words to train our model. The context is the window of words to the left and to the right of a center word. Using a window size of 2,

- Center word, “artificial” will have context words, “intelligence” and “is”.
- Center word, “intelligence” will have context words, “artificial”, “is” and “the”.
- Center word “is” will have context words, “artificial”, “intelligence”, “the” and “new”.
- Center word, “the” will have context words, “intelligence”, “is”, “new” and “electricity”.

Training dataset of (center, context) pairs. The center words will be the input to the neural network whereas context words are the target.

(artificial, intelligence), (artificial, is), (intelligence, artificial), (intelligence, is), (intelligence, the), (is, artificial), (is, intelligence), (is, the), (is, new), etc.

Neural NLP

A major disadvantage of statistical methods is the manual determination of features. Since 2015, statistical methods have been abandoned for natural language processing and neural networks have been used for machine learning (Socher, 2020) Among the different methods of deep learning, especially *recurrent neural networks* and *convolutional neural networks* give successful results in solving natural language processing problems. For example, the use of convolutional neural networks to solve very different problems of natural language processing has been demonstrated in a study that introduced a unified architecture for multitasking learning (Collobert & Weston, 2008). With these two methods, it has been stated in different studies in the literature that other deep learning methods can be used in various problems of natural language processing and successful systems can be developed based on these methods (Socher et al., 2012).

Deep learning, which is one of the machine learning methods, predicts the results with the given data set and consists of multiple layers. Deep learning, machine learning; machine learning is a sub-branch of artificial intelligence. Deep learning (also deep structured learning, hierarchical learning, or deep machine learning) is a field that includes artificial neural networks and machine learning algorithms with one or more hidden layers. Deep learning can be conducted supervised, semi-supervised or unsupervised. Deep artificial neural networks have also given successful results with the reinforcement learning approach. In order to define a model with classical machine learning techniques, the feature vector must be extracted first. In order to extract the feature vector, the data must be preprocessed. Deep Learning has eliminated the data preprocessing problem that machine learning workers have been dealing with for many years. Because, unlike traditional machine learning and image processing techniques, deep networks perform the learning process on raw data. While processing the raw data, it obtains the necessary information with the representations it has created in different layers.

For deep learning to learn distinctive features by itself, the number of data must be large. The more data there is for the learning process, the more successful the system will be. The data passes through multiple layers, revealing the details on the data.

Three main types of deep web models:

- Convolutional Neural Networks
- Recurrent Neural Networks
- Multilayer Perceptron (Multilayer Perceptron)

Convolutional Neural Networks

The CNN algorithm is very successful in problems such as image classification, object identification and image segmentation. CNNs are enhanced versions of ANNs. CNN is the network that gets deeper as a result of the increase in the number of hidden layers in ANNs. This depth in CNN is achieved using 2D filters. CNNs realize learning in a hierarchical structure (Kucuk & Arici, 2018).

CNN uses convolution and pooling operators. A CNN has three basic types of layers:

- Convolutional layer

- Pooling layer
- Fully connected layer

Multiple convolution + pooling can be done in succession. Then there are several fully connected layers. In multi-label classification problems, there is the SoftMax layer at the end. The fully connected layer receives the three-dimensional input by reducing it to one dimension and obtains a class label. The SoftMax layer calculates the probability tribution of the output classes.

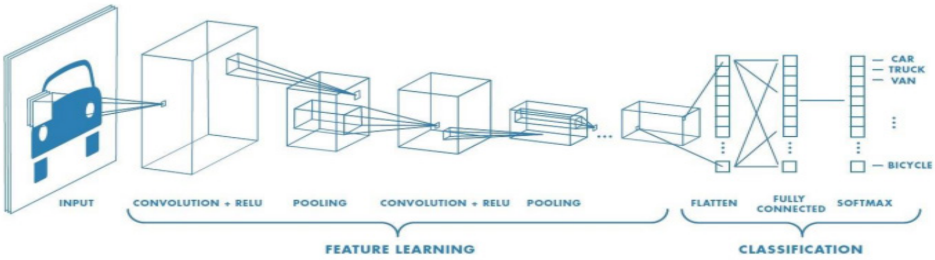


Figure 5. CNN Layers (Akcayol, 2016)

Recurrent Neural Networks

Recurrent neural networks (RNN) are a family of artificial neural networks specialized for processing sequential data. Convolutional neural networks can be scalable on data of large width and height, or they can process data of varying sizes. Similarly, iterative neural networks can scale and process sequential arrays for row-based data larger than non-specialized networks can handle. Recursive networks are based on the idea of sharing parameters between different parts of the model so that the model can be extended and applied to different form instances and generalized (Goodfellow et al., 2018). Although the difference between standard neural networks and iterative neural networks may seem insignificant, the implications of sequential learning of iterative neural networks are far-reaching. A neural network model can only map from input vectors to output vectors, while iterative neural networks can map to any output from the entire history of previous inputs. The equivalent consequence of the universal approximation theory for standard neural networks is that an iterative neural network with enough hidden units can approximate any measurable sequence-to-sequence mapping (Hammer, 2000).

The key here is that recurring connections allow a “memory” of previous inputs to remain in the internal state of the network and thus affect network output. That is, recurrent neural networks have a dynamic nature and backward neurons have backward connections. The output of the network depends on the input values of the network, the previous input values or the parameters such as the previous output values (Geron, 2017).

In summary, there is a loop fed into the RNN. The RNN structure is as follows.

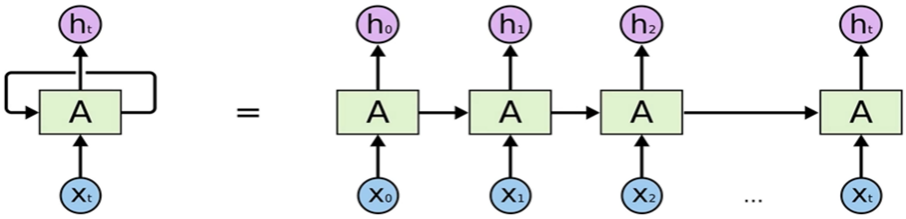


Figure 6. Recurrent Neural Network (An Introduction to Recurrent Neural Networks, 2022)

As a standard, an input comes through the neural network and an output is produced. At the same time, neurons connect with each other over time. In this way, the neural network working now can receive information from the network that worked at the previous time in this way, it can provide a memory flow. For example; The network takes a word and produces any output. When the same network moves to the next word and takes that word as input, the neural network that worked at the previous time also returns a vector, so the neural network can also access information about the previous word. In RNN, the neural network has a memory and information flows over time. Since the network can receive information from networks that have worked in the previous steps, it can interpret the sequential data in a healthier way.

Different RNN Architectures

One to one: It is the standard neural network architecture. It takes an input of a certain size and returns an output of a certain size. For example; image classification

One to many: It is the classical RNN architecture. For example; In automatic image captioning, an image is given as input, and words describing what is happening in this image are taken as output. It could also be the other way around.

Many to one: Give more than one input and get a single output. For example; In Sentiment Analysis, a text is given to the neural network and it is determined whether this text is positive or negative.

Many to many: Different numbers of inputs are given, and different numbers of outputs are received. For example; machine translation and Chatbots

An output is produced for each processed input that differs from the other. In the former, the input is processed first and then the output is produced. For example; The input can be a video and it is desirable to make a video in which every frame of the video is tagged.

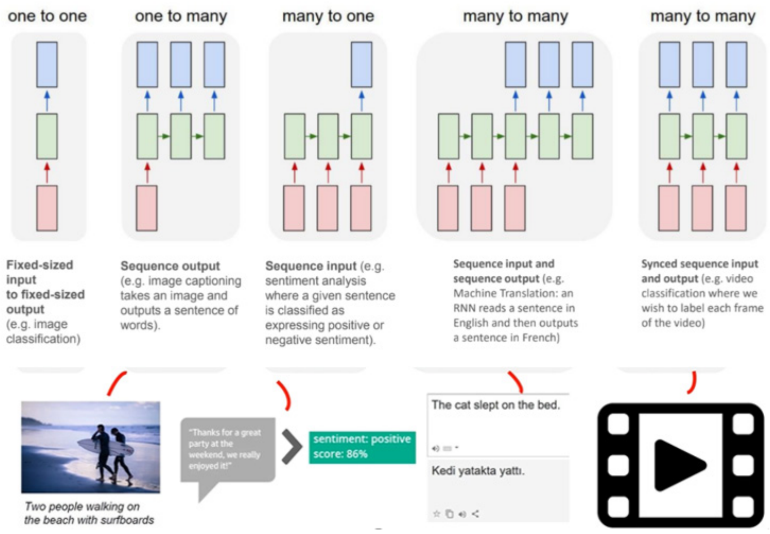


Figure 7. Different RNN architectures (Seq2Seq models,2019)

NLP Studies

The studies in the field of Natural Language Processing and the Deep Learning methods used in these studies are as follows.

Table 2. Natural Language Processing Problems and Deep Learning Methods Used (Socher et al., 2012)

Natural Language Processing Problem	**Deep learning methods used**
Text Classification	Convolutional Neural Networks Recurrent Convolutional Neural Networks Long Short-Term Memory
Text Parsing	Convolutional Neural Networks
Sentiment Analysis	Deep Auto Encoders Convolutional Neural Networks
Information Extraction	Deep Neural Networks
Asset Name Recognition	Convolutional Neural Networks
Temporal Relationship Inference	Convolutional Neural Networks
Event Inference	Convolutional Neural Networks
Vocabulary Labeling	Deep Neural Networks Long Short-Term Memory
Text Sorting	Convolutional Neural Networks
Automatic Transliteration	Deep Belief Networks
Automatic Question Answering	Convolutional Neural Networks Long Short-Term Memory

(*) LSTM is an artificial iterative neural network (RNN) architecture used in the field of deep learning. Unlike standard feed-forward neural networks, LSTM has feedback connections.

Text Classification: One of the long-studied and important application areas of natural language processing is text classification. Deep learning methods are also used for text classification and there are various current studies on this subject. In a sample study, very successful results were obtained with a simple single-level convolutional neural network approach for the sentence classification problem (such as positive/negative customer reviews) on different datasets (Kim, 2014). In another case study, an iterative convolutional neural network infrastructure is proposed for text classification. In the proposed approach; Iterative neural network is used for contextual information and convolutional neural network is used for text representation (Lai et al., 2015)

Text Parsing: In text parsing, it is aimed to reveal the grammatical structure of a given text. In a study on this subject, a fast method based on convolutional neural networks has been proposed for text parsing. In this method, convolutional neural networks are combined with label extraction structured in graphs, and thus the graph converter network method, which was previously available in the literature, was used (Collobert & Weston, 2008).

Sentiment Analysis: Sentiment analysis is defined as the automatic determination of emotion, opinion and subjectivity in each text (Chen et al., 2015). Commonly, using the text given in sentiment analysis studies, it is aimed to classify this text as positive, negative or neutral. Today, the widespread use of the Internet to enter all areas of life provides people with virtual environments such as social media, forums, blogs, e-commerce sites, where they can share their ideas, feelings and opinions. In these virtual environments, there is a huge amount of data in which people's opinions are expressed. These data attract the attention of many people and researchers, especially the manufacturers and sellers of these products and services. Idea mining and sentiment analysis aims to reveal the feelings, ideas and thoughts hidden in the texts in which people express their views on topics such as products, services, organizations, events, political thoughts in virtual environments.

Information Extraction: Information extraction from texts in natural language; Many important pieces of information such as entities, events, date and time expressions, concepts are automatically extracted. *Entity name recognition* is a sub-problem of the information extraction field of natural language processing and is a long-studied research topic. Entity name recognition systems, in general, automatically extract and classify the names of entities such as people, places and institutions in each text. *Inference of temporal relations* is also a sub-topic of information extraction and thus an important research area of natural language processing. Temporal relations include the following and similar expressions: "before, after, just before, immediately after, during (during)". Automatic extraction of events mentioned in natural language generated texts is also a sub-field of information extraction and thus a problem of natural language processing (Kucuk & Arici, 2018). In a recent study in this field, an approach using convolutional neural networks is introduced for the extraction of sentence-level features for automatic event extraction. In this study, a different word representation approach was adopted for the used word-level attributes. This study, which uses deep learning methods for event extraction, has been shown to give successful results in event extraction experiments.

Natural Language Processing and Infrastructure

NLP studies first go through a preprocessing that extracts features from audio or text to understand the text.

Preprocessing steps in natural language processing can be examined under 4 main headings: (Delibaş, 2008)

- Phonology: It examines the sounds of letters and how they are used in language. All languages have an alphabet, and each letter sounds different from the others. The aim of phonology is to translate spoken language into written language. Sounds are tried to be made into words.
- Morphology: At this stage, the words are handled individually and the structure of the word is examined in accordance with the rules of the language. At the end of this study, every part of each word is analyzed. Suffixes, roots, rules about them and classification of these structures are handled within the scope of morpheme. Morphology is to define the types of elements that make up the form in the language and to classify the formal elements called grammatical rules. The word structure of Turkish is realized by adding derivation and inflection suffixes to roots as suffixes.
- Syntax: Examines how words must be arranged to form sentences. At this stage, the words whose analysis has been completed are combined to form sentences and texts, which are larger elements of the language.
- Semantic: Understanding the sentence structure and taking action as a result happens at this stage. The basic function is to examine the meanings that the sequence of words in the language give to the sentences and to give meaning in this way.

Basic Elements of Natural Language Processing

Since the purpose of natural language processing is to communicate with the computer in natural language, the computer needs to learn natural language rules. For this, the computer needs a general dictionary and various algorithms to use this dictionary. In addition to the general information about the language, the computer also needs a field or task-specific knowledge base that it needs and must be perceived independently of the general structure of the language. There are generally five basic elements in a natural language processing system. These are *parser, dictionary (lexicon), understander, knowledge base* and *generator*. (Delibas, 2008)

Parser: The parser syntactically analyzes the given sentence and builds the parser tree. One of the most widely recognized approaches in parsing is phrase-structure grammars.

This approach is based on Chomsky's generative transformational grammar theory. It aims to group sentences by dividing them into groups. According to this approach, the basic and constitutive unit of the language is the sentence. The sentence consists of two basic structures, the noun phrase and the verb phrase. These clusters are also divided into smaller clusters within themselves. After the parsing process, the words whose tasks are determined are subjected to semantic analysis and an output sentence is formed according to the input sentence.

Dictionary (Corpora): It is a structure that contains all the words required to be recognized by the program. The parser works by doing syntactic analysis with the dictionary. The dictionary contains the root and meanings of each word that are required to be recognized by the natural language processing system. It is a work that gives the vocabulary of a language or languages with the way they are spoken and written and shows the words and meanings they form with other elements, and their different uses, based on the root of the word.

Understander: Tries to determine what the sentence means with the knowledge base.

Knowledge base: Conceptually, it consists of two sub-components: general knowledge base and task dependent knowledge base. The main task of the understander is to find the equivalent of the generated parser tree in the knowledge base. The interpreter prepares the appropriate answer for the entered sentence.

Generator: It is the most basic system used in the field of natural language processing and is the display of certain stored patterns for certain words and sentences to the user (Delibaş, 2008).

Natural Language Processing Libraries

Natural Language Processing is a field that produces solutions to many different problems in many different fields. That's why there are different libraries for working in the field of NLP. Python language is commonly used for natural language processing operations. Many reasons such as simple syntax, transparent semantics, open source code, community support are enough to choose Python language. The reason why it is preferred in the field of Natural Language Processing is that it has useful tools for operations such as machine learning and deep learning (Ozen, 2021). These libraries are; *Language Toolkit (NLTK), GENSIM, TORCHTEXT*, *TEXTBLOP*, *CORENLP, ZEMBEREK, ITU Turkish Natural Language Processing Software Chain, SentiTurkNet*

ZEMBEREK: Zemberek is an open source Turkish Natural language processing library and OpenOffice is a LibreOffice extension. Developed entirely in Java, the library has functions such as spell checking, suggestion for incorrect words, hyphenation, and faulty coding cleaning.

ITU Turkish Natural Language Processing Software Chain: It is a work that is not open source for all-natural language processing but supports natural language processing projects with API support. This study is a platform consisting of tools such as Turkish character converter (deasciifier), separation into sentence elements, spell checker, entity name recognition. Since this platform has both a web interface and an application interface, researchers at different levels can benefit from this platform. There is a preview where we can apply natural language processing methods (Eryigit, 2014).

SentiTurkNet: It is an open source Turkish natural language processing library developed especially for sentiment analysis method. In this library, which was developed by Dehkharghani et al., 2015, a dictionary indicating the emotion poles of Turkish words was created. The classification of the text in this dictionary is made according to six emotions called anger, hate, fear, happiness, sadness and surprise.

Machine learning libraries are also used in natural language processing, which is a part of the world of artificial intelligence and data science. For example, while *Scikit Learn* is frequently used in natural language processing projects, deep learning libraries such as *Tensorflow*, *Keras*, and *Pytorch* are among the libraries used to take natural language processing to the next level with the development of methods and possibilities. It is also used to visualize output with visualization libraries such as *Matplotlib*, *Seaborn* and *Bokeh*. The following table lists commonly used NLP libraries for natural language processing problems.

Table 3. Natural Language Processing Method and Libraries (Yılmaz & Yumuşak, 2021)

Natural Language Processing Methods	Libraries
Named Entity Recognition	NLTK, Spacy, AllenNLP, Stanford-NLP, Zemberek
Text Classification	Zemberek, Flair, PyTorch
Language Identification	TextBlob, Zemberek
Vectorization	Gensim, Scikit-Learn
Separating sentence into elements, finding the root of words, finding the head word	Zemberek, Turkish NLP, Turkish Stemmer, Turkish Lemmatizer, Zemberek Parser
Morphology	Zemberek, ITU Turkish NLP Pipe-line
Normalization	Zemberek, Fast.ai, Pyspellchecker
Sentiment Analysis	SentiTurkNet, TextBlob

Natural Language Processing Steps

Segmentation of Sentence (Tokenization): It is the process of breaking down sentences in the text into meaningful small unit particles. As a result of the decomposition of the sentence, too many unit particle tokens are formed. Removal of ineffective words (stop word), removal of misspelled words, stemming and lemmatization can be performed as methods to reduce these particles. Tokens are meaningful small units, symbols, words, phrases can be given as examples of tokens.

Word Tokenizer: Splits the sentence into words and extracts the punctuation marks, also separates the possessive “apostrophe s” in English together.
Sentence Tokenizer: Splits the paragraph into sentences.
Treebank Word Tokenizer: Separates words in sentences according to spaces and punctuation marks.
Word Punct Tokenizer: Extracts punctuation marks from the sentence.
Extraction of ineffective words (stop word): It is one of the methods used to preprocess texts. Stop words are generally unnecessary words. Words that are used to improve the flow of sentences but do not make sense when analyzing data are called stop words. They are words like “a, an, the, and, I, me, myself” in English.
Words like these mean nothing. They are important words for understanding sentences, but they mean nothing to the machine when analyzing data. If the maximum number of words in a text is calculated, the words in the first 10 are made up of stop words. Therefore, these words in the text unnecessarily can be removed or a threshold value can be determined so that it does not affect the text to be learned too much.
Finding the root of words: There are 2 methods to find the roots of words: *Stemming* and *Lemmatization*. Both are methods that try to find the root of the words by discarding the suffixes. The method of finding the root of words (stemming) is simpler than finding the root word (lemmatization), and a strong grammatical knowledge is required for algorithms to find the root word (Yilmaz & Yumusak, 2021). The words highlight, highway and high have the same roots and are high.

Table 4. Stemming and Lemmatization difference

Words	Stem roots	Lemma roots
Drive	Drive	drive
Driving	Drive	driving
Driver	Driver	driver
Drives	Drive	drive
Drove	Drove	drove
Cats	Cat	cat
Children	Children	child

Part of Speech Tagging:It is the tagging of words in texts with their elements as syntax. It is the process of putting the class as a label, whichever class belongs to the word, such as noun, verb, adjective, conjunction.

1) Normalization: Text normalization is the name given to the methods used to reduce clutter in the analyzed text. Case adjustments, unnecessary spacing or character removal, number/text corrections, abbreviation corrections, etc. includes regulations. Statistical methods and the use of some distance measuring methods are common in this field. The most well-known method is

the Levenshtein distance method (Yılmaz & Yumusak, 2021).

2) Morphology: It is the process of classifying sentence components that are parsed as tokens according to grammar.
3) Named Entity Recognition: It is the process of defining predefined categories such as person, place, organization, institution through text documents.
4) Vectorization: The first step for the computer to understand the language is to understand the words. For words to be understood, they must be symbolized in a way that the computer can understand, that is, they must be represented numerically. When words are represented numerically, mathematical operations can be performed on words. Vectors are used to represent words numerically. The conversion of texts into numerical expressions is called word embedding.

Methods that allow words, sentences, or documents to be expressed as vectors are called vectorization methods. Since machine learning algorithms or mathematical models work with numbers, textual expressions need such a transformation.

Developed methods for determining grammatical features in written texts are basically grouped into 2 groups:

- Rule-based methods
- Methods based on statistics / probability

These methods are described in the Methods section.

5) Text Classification: Natural language processing applications that predict which subject/class the content of a text belongs to are called text classification applications. There are also various applications that help us decide who is the author of a text or whether an email is spam based on its content.

Sentiment Analysis

Sentiment analysis is defined as the automatic determination of emotion, opinion, and subjectivity in each text (Chen et al., 2015). Commonly, using the text given in sentiment analysis studies, it is aimed to classify this text as positive, negative, or neutral. Idea mining and sentiment analysis aims to reveal the feelings, ideas and thoughts hidden in the texts in which people express their views on topics such as products, services, organizations, events, political thoughts in virtual environments.

Since the early 2000s, sentiment analysis has become a very active field of study as a sub-branch of natural language processing. Sentiment analysis studies are carried out in the fields of data mining, web mining and text mining as well as natural language processing (Liu, 2012). The term sentiment analysis was first used by Tetsuya and Jeonghee in 2003, and the term opinion mining was first used by Kushal et al., in 2003. Although these statements emerged in 2003, studies on this subject started in previous years. The studies of Vasileios and Janyce in 2000, Tong et al., in 2001, Turney in 2002, Pang et al., in 2002 can be cited as examples of the first studies in this area.

Sentiment analysis and opinion mining are the studies that reveal the feelings, ideas and thoughts hidden in the texts in which people express their views on different topics such as products, services, organizations, events. In sentiment analysis studies, things such as products, services, events, people, about which opinions are expressed, are called assets. Expression of emotion may be about the entity itself, or it may be about an aspect or feature of the entity (Ozyurt & Akcayol, 2018).

(v, ö, d, n) → v: existence, ö: property, d: emotion, n: subject

The main tasks in sentiment analysis are to extract some or all of these four components of emotion expression from texts according to the scope and level of the study.

Existence: The task of extracting entities from the text is not performed in all sentiment analysis studies. For example, in product reviews or product comments on e-commerce sites, an opinion is expressed about a certain entity, so there is no need for entity extraction from the text. On the other hand, in a column about politics, the author can express his opinion about different political parties and politicians. Asset extraction is important in the idea of mining work to be done in such texts. Entity extraction is a rule-based entity extraction task. There are studies in Turkish in this field by Dalkilic et al. in 2010, Seker and Eryigit in 2012, Kucuk and Yazici in 2009.

Property: Whether feature extraction is performed depends on the level of sentiment analysis.

Emotion: This is also called emotional polarity. Emotion polarity is positive or negative. Although some studies have also been classified as neutral, it is not very common. In addition to the positive/negative binary classification, emotion rating studies are also carried out to determine the positivity/negative level of the emotion. Removing emotion polarity is the main task in sentiment analysis and opinion mining.

Subject: The subject can be a person as well as a legal person. In forums, blogs and e-commerce sites, the subject is the user who writes the message, except in very exceptional cases. In texts such as newspaper news in which the ideas and thoughts of people such as administrators and politicians are conveyed, the subject can be third parties instead of the author himself (Ozyurt & Akcayol, 2018).

References

Akcayol, M. Ali, (2016), *Deep Learning* [PowerPoint slides]. SlideShare. https://w3.gazi.edu.tr/~akcayol/files/DL_L6_CNNs.pdf

An Introduction to Recurrent Neural Networks, (2019), https://datascience.eu/machine learning/an-introduction-to-recurrent-neural-networks/.

Cambria, E., White, B., 2014, Jumping NLP Curves: A Review of Natural Language Processing Research, IEEE Computational Intelligence Magazine, 9(2), 48-57.

Chen, Y., Xu, L., Lıu, K., Zeng, D., Zhao, J., 2015, Event extraction via dynamic multi-poolin convolutional neural networks, Annual Meeting of the Association for Computational

Chen, Y., Xu, L., Lıu, K., Zeng, D., Zhao, J., 2015, Event extraction via dynamic multi-pooling convolutional neural networks, *Annual Meeting of the Association for Computational Linguistics and International Joint Conference on Natural Language Processin*g, 167-176.

Choe, D. K., Charniak, E., (2016), Parsing as Language Modeling. classification, AAAI Conference on Artificial Intelligence, 2267-2273.

Collobert, R., Weston, J., 2008, A unified architecture for natural language processing: Deep neural networks with multitask learning, *International Conference on Machine Learning (ICML),* 160-167.

Dalkilic, F.E., Gelisli, S., Diri B., 2010, Named Entity Recognition from Turkish texts, IEEE 18.Signal Processing and Communication Applications Congress, Diyarbakır, Türkiye, ss. 918-920.

Dehkharghani, Yucel, R.S., Yanikoglu, B., Oflazer, K., 2015, SentiTurkNet: a Turkish polarity lexicon for sentiment analysis, Language Resources and Evaluation. 50. 10.1007/s10579-015-9307-6.

Delibas, A., 2008, Turkish Spell Check With Natural Language Processing, (Master Thesis, Istanbul Technical University, Institute of Natural and Applied Sciences), İstanbul.

Embeddings: Translating to a Lower-Dimensional Space, (2022, August 12), https://developers.google.com/machine-learning/crash-course/embeddings/translating-to-a-lower-dimensional-space

Eryigit, G., 2014, ITU Turkish NLP web service, *In Proceedings of the Demonstrations at the 14th Conference of the European Chapter of the Association for Computational Linguistics* (pp. 1-4).

Geron, A., 2017, Hands-On Machine Learning with Scikit-Learn and TensorFlow, O'Reilly Media.

Goldberg, Y., (2016), A Primer on Neural Network Models for Natural Language Processing, Journal of Artificial Intelligence Research 57: 345–420.

Goodfellow, I. Bengio, Y., Courville, A., 2018, Derin Öğrenme, Buzdağı Yayınevi.

Goodfellow, I., Bengio, Y., Courville, Aaro (2016), Deep Learning MIT Press.

Hammer, B., 2000, On the approximation capability of recurrent neural networks, Neurocomputing, 31 (1-4), 107-123, (2000).

Jozefowicz, R., Vinyals, O., Schuster, M., Shazeer, N., Yongh, W., (2016), Exploring the Limits of Language Modeling.

Karci, A., Aydogan, M., (2019), Investigation of Word Similarities with Word Embedding Methods, Cukurova University Journal of the Faculty of Engineering and Architecture, 34(2),. 181-195.

Kim, Y., 2014, Convolutional neural networks for sentence classification.

Kucuk, D., Arici, N., 2018, A Literature Study On Deep Learning Applications In Natural Language Processing, International Journal of Management Information Systems and Computer Science, 2(2):76-86.

Kucuk, D., Yazici, A., 2009, Named Entity Recognition Experiments on Turkish Texts, In Proceedings of FQAS-2009, 8th International Conference on Flexible Query Answering Systems*, Roskilde*, Danimarka, ss. 524-535.

Kushal, D., Steve, L., Pennock, D.M., 2003, Mining the Peanut Gallery: Opinion Extraction and Semantic Classification of Product Reviews, In Proceedings of WWW'03, 12th International Conference on World Wide Web, Budapest Congress Centre, Macaristan, ss. 519-528, 2003.

Lai, S., Xu, L., Liu, K., Zhao, J., 2015, Recurrent convolutional neural networks for text Linguistics and International Joint Conference on Natural Language Processing, 167-176.

Liu, B., 2012, Sentiment Analysis and Opinion Mining, Morgan & Claypool Publishers.

Malinowski, M., Fritz, M., (2016), Tutorial on Answering Questions about Images with Deep Learning.

Natural Language Processing History, (2022, August 12), https://towardsai.net/p/nlp/conditional-story-generation%E2%80%8A-%E2%80%8Apart1

Natural Language Processing, (2022, August 12), https://en.wikipedia.org/wiki/Natural_language_processing

Ozen, M.S., (2021), Python Doğal Dil İşleme Kütüphaneleri, https://www.datasciencearth.com/dogal-dil-isleme-1-4-python-dogal-dil-isleme kutuphaneleri/

Ozyurt, B, Akcayol, M.A., 2018, A Survey On Sentiment Analysis And Opinion Mining, Methods And Approaches, Selcuk Univ. J. Eng. Sci. Tech., v.6, n.4, pp. 668-693, 2018.

Pang, B., Lee, L., Vaithyanathan, S., 2002, Sentiment Classification Using Machine Learning Techniques, In Proceedings of EMNLP-2002, Conference on Empirical Methods in Natural Language Processing, University of Pennsylvania, PA, ABD, ss. 79-86.

Seker, G.A., Eryigit, G., 2012, Initial Explorations on using CRFs for Turkish Named Entity Recognition, In Proceedings of COLING 2012, *24th International Conference on Computational Linguistics*, IIT, Bombay, Hindistan, 2459–2474.

Seq2Seq models and the Attention mechanism, (2019), https://mett29.github.io/posts/2019/12/seq2seq_and_attention/.

Socher, R., (2020), Deep Learning For NLP-ACL 2012 Tutorial.

Socher, R., Bengio, Y., Manning, C. D., 2012, Deep learning for NLP (without magic), Tutorial Abstracts of ACL, 5.

Süleiman, D., Awajan, A., 2019, Comparative Study of Word Embeddings Models and Their Usage in Arabic Language Applications.

Tetsuya, N., Jeonghee, Y., 2003, Sentiment Analysis: Capturing Favorability Using Natural Language Processing, In Proceedings of KCAP-03, *2nd International Conference on KnowledgeCapture*, Sanibel Island, FL, ABD, ss. 70-77.

Tong, R.M., 2001, An Operational System for Detecting and Tracking Opinions in On-Line Discussion, In Proceedings of SIGIR 2001 Workshop on Operational Text Classification, New Orleans, Louisiana, ABD.

Turney P.D., 2002, Thumbs up or Thumbs down: Semantic Orientation Applied to Unsupervised Classification of Reviews, In Proceedings of ACL'02, *40th Annual Meeting of the Association for Computational Linguistics*, Pennsylvania, ABD, ss. 417-424.

Vasileios, H., Janyce, M.W., 2000, Effects of Adjective Orientation and Gradability on Sentence Subjectivity, In Proceedings of COLING-2000, *18th Interntional Conference on Computational Linguistics,* Saarbrücken, Almanya, ss. 299-305.

Vinyals, O., Kaiser, L., Koo, T., Petrov, S., Sutskever, I., Hinton, G., (2014). Grammar as a Foreign Language, NIPS, http://arxiv.org/abs/1412.7449

Yilmaz, H., Yumusak, S., 2021, Open Source Natural Language Processing Libraries, Istanbul Sabahattin Zaim University Journal of the Institute of Science and Technology,3 (1): 81-85.

Young, T., Hazarika, D., Poria, S., Cambria, E., (2018), Recent Trends in Deep Learning Based Natural Language Processing. IEEE Computational Intelligence Magazine, 55-75.

About the Authors

Seda YILDIRIM graduated from Selcuk University, Faculty of Engineering and Architecture, Computer Engineering in 2015. She received the master's degree from the Department of Computer Engineering, Institute of Science, Konya Food and Agriculture University, Konya, Turkey (2019). She continues his Ph. at Konya Technical university. She has been Lecturer at KTO Karatay University since 2020. His research interests include image processing, natural language processing, artificial intelligence, machine learning and deep learning.

E-mail: seda.yildirim@karatay.edu.tr, **ORCID:** 0000-0003-2944-6826

Similarity Index

The similarity index obtained from the plagiarism software for this book chapter is 19%.

To Cite This Chapter:

Yıldırım S. (2022). Natural Language Processing Applications in Engineering. In S. Kocer & O. Dundar (Eds .), *Current Studies in Basic Sciences Engineering and Technology* (pp. 12–30). ISRES Publishing.

Microstrip Antennas in Biomedical Fields

Rabia TOPRAK
Karamanoglu Mehmetbey University

Introduction

Antennas are extremely useful equipments in communication applications. Receiving information taking from a supply and transmiting to another supply is provided by the antenna structures. With these features, antennas are indispensable for telecommunication applications. There are lots of types of the antenna structures. They are horn, microstrip, monopol, dipol, Vivaldi, loop, Bow-Tie, log-periodic and etc. In the designs of antenna, there are important parameters. They are operating frequency, bandwidth, radiation pattern, gain, directivity, impedance, return loss, efficiency, VSWR (Voltage Standing Wave Ratio). These parameters explain the performance of the antenna structures. A radiation pattern defines the variation of the power radiated by an antenna structure as a function of the direction away from the antenna structure. Therefore, gain, directivity and field graphs are generally explain with the radiation pattern. Another parameter is the efficiency of the antenna structure. Efficiency is an important paramater for an antenna structure because of a ratio of the power delivered to the antenna structure relative to the power radiated from the antenna structure (Balanis, 2013).

Microstrip antennas (MAs) consist of a dielectric layer called as subtrate between two conduction layer. Copper is generally used as conduction material. Substrate is a material has different dielectric constant and tangent loss. The thickness and dielectric constant value of substrate can be designed differently by different companies. This provides easily the ability of MAs design. MAs are also be able to call as microstrip patch antennas, too. So, MAs are prefered in large search areas (B. J. Kwaha, O. N. Inyang, 2011; Bahrami et al. 2021; Dey & Mittra, 1996; Kanj & Popovic, 2005; Nalam et al., 2014; Peixeiro, 2011; Rabbani & Ghafouri-Shiraz, 2017; Soontornpipit et al., 2004; Top et al., 2020a; Top et al., 2019; Top et al. 2017, 2020b; Uyanik et al., 2020a; S. Yang et al., 2008). The reason of this preference is that they have lots of advantages. One of the advantages is that could directly be printed onto a circuit board. The other is low production cost. Easy design, lightness, large design facility and etc. are possible to sort as another advantages. Disadvantages of MAs are low gain and efficiency values. According to their designs, the efficiency, gain and directivity of microstrip antenna structures are able to increase by applying different methods. Therefore, MAs become attractive for different research applications in the worldwide (Singh & Tripathi, 2011).

Every research related to human life is valuable. So, Biomedical/Medical applications are interested in healt issues. Sometimes they aim to diagnose and sometimes to treat. Especially as far as human health is concerned, all of biomedical applications get importance. Biomedical researches and developments are made for purposes that are very different. For examples; Wireless Body Area Network (WBAN), Implantable Medical Devices (IMDs), Wireless Local Area Network and Wearable Medical Devices (WMDs) (Kolokotsa et al., 2002; Paracha et al., 2019; Samal et al., 2019; Simorangkir et al., 2018).

Microstrip Antenna Applications on Cancer Researches

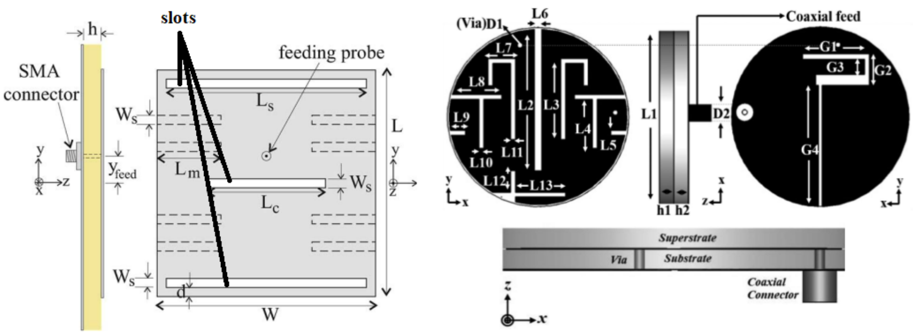
MAs are highly used in cancer researches. It is possible to say that cancer researches are fairly important because of the effect of death ratios. All over the world people suffer from diseases suppled cancer. Once the statistics are examined, it is seen one of the biggest death ratios about 19% are from diseases suppled cancer (Institution, 2020). Therefore, to diagnose or treat the cancer diseases become quite important. With this aim, in the literature there are several studies. Specially, to solve such health problems, microstrip antenna structures are explained in this book chapter. MAs are especially chosen to produce solution easily cancer diseases. Because these antenna structures have lots of advantages mentioned above. Also, there are a lot of the types of cancer. They are sortable as breast, brain, pancreatic, kidney, colon cancer and etc. Each type of cancer is a necessary and different research topic.

Such works are generally related to detect cancer diseases. Because imaging methods are MR (Magnetic Resonance), ultrasound or CT (Computing Tomography) (Alsharif & Kurnaz, 2018; Caliskan et al., 2015; Top et al., 2019). These methods are contained very high frequency bands. For example, X-ray. X-ray is able to affect the atom of human tissue. Although it has low cost, it has low image resolution and the effect of ionizing radiation. MR has high image resolution and nonionizing radiation. But, it is expensive, bulky, noisy and time consuming. Frequency region is very important because of the effect of radiation. The radiation intensity increases at higher frequencies (McKeen, 2016). Also, CT is expensive, bulky and has the effect of ionizing radiation in human body (Alqadami et al., 2019). This could harm the human body at the frequent measurements to detect cancer diseases. So, MAs have proposed for microwave imaging solutions. Microwave frequency region has no effect on human body. It is called as Microwave Tomography. Medical diagnostic, identification via wireless communication, wireless sensors and etc. applications are continually researching and developing with this aim. Although the proposed antenna types differ, MAs are mostly preferred.

Microstrip Antenna Applications for the Type of Breast Cancer

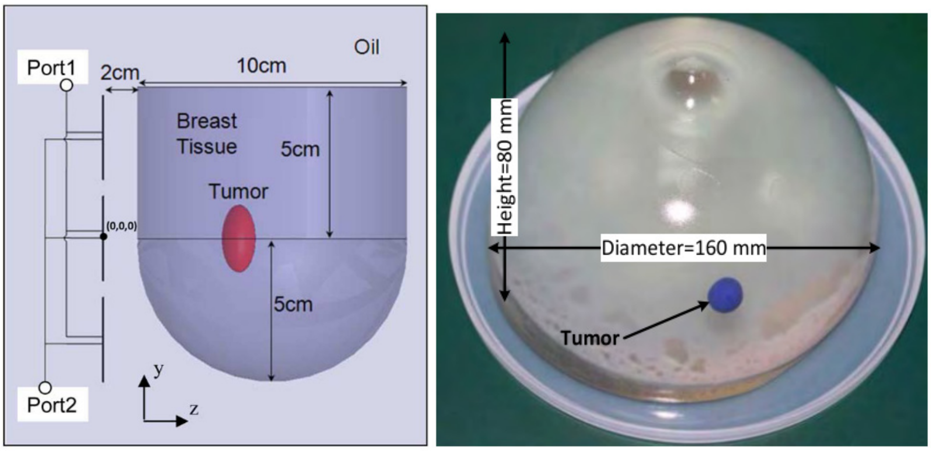
All over the world women suffer from breast cancer. One in 8 women gets breast cancer. In men, this rate is 1 in 833 (“Breast Cancer Facts and Statistics,” 2022; World Health Orgabization, 2021). Breast cancer is that has the biggest death rates between women. Early diagnosis is very important and beneficial to treat it. For early diagnosis the most important part is detection to tumorous tissue in breast. Namely, imaging method is almost vital in women life. Because, every woman above 30-years old need to make check-up twice a year with an imaging method. With this aim, the importance of imaging methods increases day by day. The less radiation intensity used in imaging method is, the more useful women who are measured is. Therefore, Microwave Tomography applications are searched with an increasing demand. Both its application is simple and its cost is low. MAs are developing as an alternative imaging method that has non-ionizing radiation intensity against other imaging methods that has high ionizing radiation intensity level.

Operating frequency and band area are topics examining in antenna structures. Dual-band frequency operation is applied by handling the fundamental resonant mode of the MAs and one of its higher order modes. Dual-band (multi frequency band) operation provides solving the inverse scattering problems. Also, the coupling effect of antenna in an immersion medium is decreased with dual-band operation. Slotting is a method to obtain dual-band in patch of MAs (Al-Joumayly et al., 2010). Short-circuit pin and superstrate methods are another methods to get dual-band operation (Blauert & Kiourti, 2014; Ganeshwaran et al., 2020; Woten & El-Shenawee, 2008). Figure 1 shows some example of MAs slotted and added via in the patch. Figure 1. (a) shows an example of the rectangular patch with three slots and Figure 1. (b) shows an example of the circular patch with slots and via. Superstrate method is used to increase the gain of MAs and to get dual-band operation frequency regions (Faisal & Yoo, 2019a). It is also used to prevent direct contact of the patch with the human tissue. Slotting in the patch and ground plane are methods to increase the gain of MAs (Çalışkan et al., 2015; Toprak et al., 2021). Proposed antennas are used to simulate and measure several 3D measurement set-ups. Some examples are presented in Figure 2. Different breast models are presented at different studies (AbdulSadda et al., 2010; Li et al., 2015).



(a) (b)

Figure 1. (a) Slotted example of MAs, (b) Slotted and added via example of MAs in the patch (Al-Joumayly et al., 2010; Ganeshwaran et al., 2020)



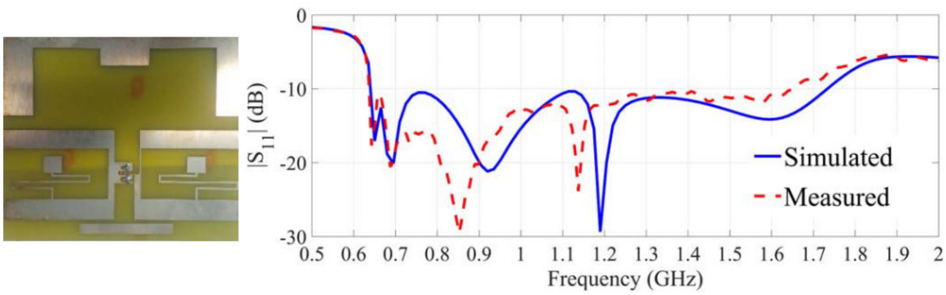
(a) (b)

Figure 2. 3D tumorous breast model (a) for simulation, (b) for measurement (Mahmud et al., 2018; Woten & El-Shenawee, 2008)

Another technique to detect breast cancer is to get ultra wide-band (UWB). As differing slot number in the patch, it is possible to make UWB. UWB ensures better return loss and VSWR (Voltage Standing Wave Ratio) value in the measurements. These values are important parameters for evaluating antenna performance (Ul Haq & Khany, 2014).

Microstrip Antenna Applications for the Type of Laryngeal and Trachea/Bronchus/Lung Cancer

In 2019, Türkiye, according to Türkiye Statistical Institute the Larynx and trachea/bronchus/lung tumors caused the most deaths from tumors (Institution, 2020). When deaths from benign and malignant tumors were analyzed according to sub-causes of death, it was seen that 30.2% of the deaths were caused by malignant tumors of the larynx and trachea/bronchus/lung. For imagine these cancer types, MAs are used as microwave imaging method. A study has used MAs as dipole antenna structure. The proposed antenna technique and set-up are quite fast in scanning, computation, and image creation because it needs no additional devices for accurate boundary estimation. Array elements are utilized in MAs structure because of obtaining the dipole antenna structure (Zamani et al., 2018). To do this, by taking the feeding point as the center, a co-design has been made in the patch. So, dipole antenna structure is obtained. Figure 3 shows the antenna structure and its return loss graphs.



(a) (b)

Figure 3. (a) Dipole antenna from MAs, (b) Its return loss (Zamani et al., 2018)

MAs are designed as resonators for imaging lung cancer type. There are different resonator types such as square, circular and etc. Measurements with resonator method are more accurate than other characterization techniques. Resonator method also provides to miniature the proposed technique. Therefore, while smaller resonator structure is got, higher accuracy results are gained (Hardinata et al., 2017).

MAs with array elements are utilized for electromagnetic imaging. Metamaterial is also used as meta surface element. Cause, backscattering problem is minimized with this method (Zamani et al., 2019). Advantage of this study is better imaging reconstruction. While there is no touch on human torso, imaging method could be easily applied.

A prototype structure, microstrip impedance transformer and antenna have been designed and simulated. In order to deliver the microwave energy into lung tissue at the end of the cable, an antenna has designed. The design must maintain access to the hollow channel at the end to not block the field of view of the CMOS (Complementary metal–oxide–semiconductor) imaging and illumination (Jones et al., 2019). This study is thinkable as lung ablation sensor that designed by using MA structure.

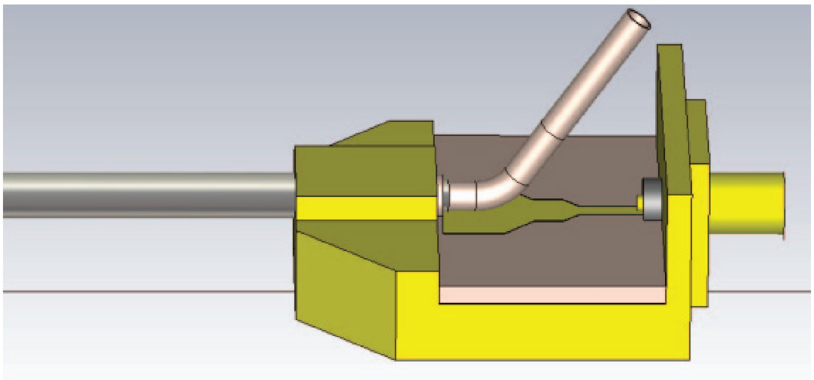


Figure 4. Modelled microstrip and ground structure tray (Jones et al., 2019)

Microstrip Antenna Applications for the Type of Stomach and Colon Cancer

After breast, laryngeal,trachea/bronchus/lung cancer, stomach and colon cancer types have the most death ratios. Both stomach and colon cancer types have about 8% of the death ratio in death ratios supplied benign/malignant tumour in Türkiye (Institution, 2020).

At the detection of colon cancer, colonoscopy is vital. For microwave colonoscopy system, a compact cylindrical U-slot antenna is applied. This antenna with MA structure has 16 L-shaped microstrip lines with two rows. One row is for transmission and other is for reception. L-shaped microstrip feed point is designed to feed U-slot effectively. To make easily colonoscopy, antenna structures wrapped around metal cylinder like a camera. Of course this style is cheaper than a system with a camera. The shape of the antenna provides better input impedance matching when there ise human tissue. Antenna performance is tested in liquid phantom (Guardiola et al., 2019).

To measure the reaction of different organ tissue, two layered printed compact dual polarized antenna is designed. Both horizontal and vertical polarization is maintained with double layer (Sabban, 2011). To prefer double or multiple layer on the design of antenna structure is quite effective method. Therewithal, multiple layer assures the compactness of the MAs.

There are wireless capsule endoscopy systems with MAs. In this system, antennas are the most important element. Flexible substrate is used because of wrapping around the capsule. Proposed antenna type is a loop antenna. Also it is suitable for paediatric application (Miah, Icheln et al.,2018). Another capsule antenna structure is presented by differing the shape of patch (Suzan Miah et al., 2019). In this work, antenna structure is an ultra-wideband conformal loop antenna for the in-body transmission.

Microstrip Antenna Applications on Heart Failure Diseases

Circulatory system diseases are any conditions that affect your heart or blood vessels. Circulatory system is also called your cardiovascular system. It keeps blood moving in your body. Diseases originating from the circulatory system have the highest mortality rate in the world today. The biggest ratio in this rate is from heart failure diseases (Institution, 2020; World Health Organization, 2020). Especially coronary artery occlusion is vital in the development and progress of heart failures. At the early stage, the detection of cardiac occlusion could be able to save lots of people life.

MAs in the detection of heart failure diseases are researching and development. A microstrip patch antenna has been designed at ISM (Industrial, Scientific, Medical) frequency band. A 3D heart model is developed by using the human tissue phantom receipt. This study focuses on the different patch shapes. Therefore, the performance of the proposed antenna is increased to detect coronary artery occlusion (Top, 2017; Top, Gultekin, & Uzer, 2017; Uyanik et al., 2020b). To detect heart failure at early stage, a compact antenna array with wideband performance is designed and tested in a study. The performance of the proposed antenna is tested in an artificial torso phantom. In the study, the proposed antenna with microstrip structure is designed as a loop antenna structure. So, wideband operation is provided (S.Ahdi Rezaeieh et al., 2015). Figure 5 shows the proposed antenna structure. Wideband frequency operation ensures the fine resolution of reconstruction. Another study suggests a leadless pacemaker by proposing a prototype rectenna. They propose a dipole antenna that is planar. Proposed antenna has a fractal geometry that is inductively doubled with rectangular strips. Fractal geometry is for minimization of the antenna structure. The performance of the proposed antenna is tested and measured with an ovine which is alive. The results of measurements are successful to pace a live heart using wearable conformal antenna array (Asif et al., 2019). Another leadless pacemaker by using MAs is suggested by differing dimensions of patch and ground plane (Abbas et al., 2022).

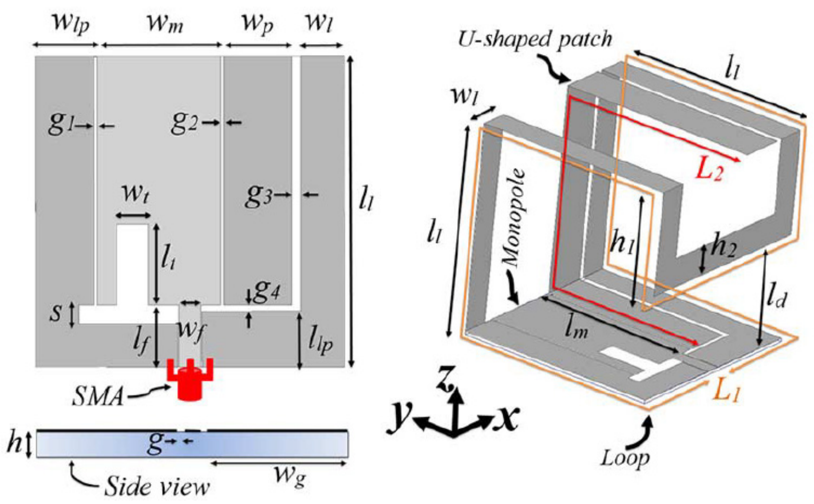


Figure 5. The proposed antenna structure for detection cardiac occlusion at early stage (S.Ahdi Rezaeieh et al., 2015)

A unidirectional and wideband antenna is proposed to image thorax. Reflector structure is utilized for resonance-based. Patch part is designed for reflection. So dimension of the proposed antenna is minimized. Multiple antenna structures could be located adjacent to each other with no experiencing any impedance mismatch or frequency shift. Figure 6 shows the proposed antenna structure. It is seen from the picture that antenna structure provides easily transfer electromagnetic energy to thorax (Sasan Ahdi Rezaeieh et al., 2018).

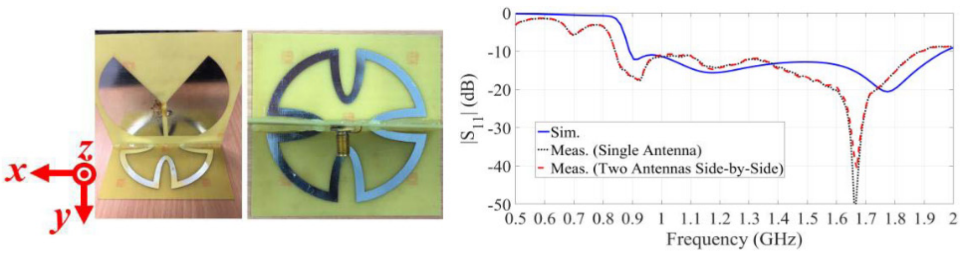


Figure 6. The proposed anten structure (left) and its return loss graph (right) (Sasan Ahdi Rezaeieh et al., 2018)

MAs are examined for cancer and heart failure diseases above. There are different MAs in the literature to produce different solution for biomedical areas. For example, wearable microstrip antenna is designed for detection brain cancer. Proposed antenna structure has small dimensions and compatible for measuring in the brain structure. Patch part is from pentagon shaped and ground plane is modified. This design supplies the increase of current flow (Raihan, Alam Bhuiyan, Hasan, Chowdhury, & Farhin, 2017). MAs is very preferable to detect brain cancer type (Mahmoud & Montaser, 2021; Rodrigues et al., 2021).

Telemetry applications are very searchable subject in medical. So MAs are used frequently to apply telemetry applications (Basir & Yoo, 2019; Faisal & Yoo, 2019b; Wireless, 2018; Xia et al., 2020). MAs are popular antenna types for other biomedical applications (Alkhaibari et al., 2017; Hasan et al., 2018; Iqbal et al.,2019; Lin et al., 2018; Nalam et al., 2014; Nesasudha & Fairy, 2018; Rahaman & Delwar Hossain, 2019;

Saha et al., 2018; Shadid et al., 2018; Xia et al., 2020; Z. J. Yang & Xiao, 2018a, 2018b). It is possible to rise MA examples in biomedical areas.

Conclusion

When the examples are examined, the importance of MAs is come out. Ease of design of microstrip antennas, which was also emphasized above, provide a huge research facility/possibility. The variety of substrate material used in designs provides design richness. According to the aim of design, it is usable both robust and flexible materials. Thickness and dimensions of it are able to adjusted as desired. There are many companies that produce PCBs. On account of the situation, it could be found in desired thickness.

The patch of MAs is able to design desire shapes. A flower, symbol, geometric shapes and letters are able to use. Moreover, one or more slot could make a slit in the patch structure. The shape of slot may differ. These differences are used for various aims. For example, to minimize the dimensions of antenna, to obtain wideband, ultra wideband, to increase the gain of antenna, to determine the directivity of antenna, to increase the efficiency of the antenna, to adapt the impedance matching and etc.The via used for antenna designs especially provides to decrease the dimensions of antenna structure. To minimize the antenna structure, generally via is utilized. But there is a disadvantage of it. Impedance matching may be a problem because of solder on ports. The ground plane of MAs has different shapes. Modified ground plane generally uses to obtain wideband or ultra wideband operations. Shape of it could be a geometric shape, letter or any other design. Using multiple layer called as superstrate is a different design method. Same or different substrate materials are designed by superimposing. It is generally used to increase the gain of antenna. Gain value is important because of effecting the directivity of antenna. Moreover, it is able to prefer for minimization and wideband operation.

MAs are possible to transform a loop or reflector antennas after designs. Similarly, the surface of patch is able to cover another different material for the aim of hypothesis. Metamaterial is another design and research subject in this area. There are several successful studies with metamaterial used. All of the examples and explanations are shown that MAs are very charming candidates for biomedical applications. Despite all the work, MAs continues to be developed at the moment. MAs are sometimes a biosensor, sometimes an implantable device and sometimes a wearable device. The importance of them is vital in biomedical applications.

References

Abbas, N., Basir, A., Iqbal, A., Yousaf, M., Akram, A., & Yoo, H. (2022). Ultra-Miniaturized Antenna for Deeply Implanted Biomedical Devices. *IEEE Access*, *10*, 54563–54571. https://doi.org/10.1109/ACCESS.2022.3176720.

AbdulSadda, A., Bouaynaya, N., & Iqbal, K. (2010). Noninvasive breast tumor localization based on ultrawideband microwave backscatter. *2010 International Conference on Signal Processing and Communications, SPCOM 2010*. https://doi.

org/10.1109/SPCOM.2010.5560490.

Al-Joumayly, M. A., Aguilar, S. M., Behdad, N., & Hagness, S. C. (2010). Dual-Band Miniaturized Patch Antennas for Microwave Breast Imaging. *IEEE Antennas and Wireless Propagation Letters*, *9*. Retrieved from https://ieeexplore.ieee.org/document/5433057/.

Alkhaibari, A., Sheta, A. F., & Elshafiey, I. (2017). Notched anti-podal Vivaldi antenna for biomedical applications. *2017 7th International Conference on Modeling, Simulation, and Applied Optimization, ICMSAO 2017*. https://doi.org/10.1109/ICMSAO.2017.7934874.

Alqadami, A. S. M., Bialkowski, K. S., Mobashsher, A. T., & Abbosh, A. M. (2019). Wearable Electromagnetic Head Imaging System Using Flexible Wideband Antenna Array Based on Polymer Technology for Brain Stroke Diagnosis. *IEEE Transactions on Biomedical Circuits and Systems*, *13*(1), 124–134. https://doi.org/10.1109/TBCAS.2018.2878057.

Alsharif, F., & Kurnaz, C. (2018). Wearable Microstrip Patch Ultra Wide Band Antenna for Breast Cancer Detection. *2018 41st International Conference on Telecommunications and Signal Processing, TSP 2018*. https://doi.org/10.1109/TSP.2018.8441335.

Asif, S. M., Iftikhar, A., Hansen, J. W., Khan, M. S., Ewert, D. L., & Braaten, B. D. (2019). A Novel RF-Powered Wireless Pacing via a Rectenna-Based Pacemaker and a Wearable Transmit-Antenna Array. *IEEE Access*, *7*, 1139–1148. https://doi.org/10.1109/ACCESS.2018.2885620.

B. J. Kwaha, O. N. Inyang, P. A. (2011). The circular microstrip patch antenna-design and implementation. *International Journal of Recent Research and Applied Studies (IJRRAS)*, *8*(1), 86–95.

Bahrami, S., Moloudian, G., Miri-Rostami, S. R., & Bjorninen, T. (2021). Compact Microstrip Antennas with Enhanced Bandwidth for the Implanted and External Subsystems of a Wireless Retinal Prosthesi. *IEEE Transactions on Antennas and Propagation*, *69*(5), 2969–2974. https://doi.org/10.1109/TAP.2020.3025245.

Balanis, C. A. (2013). *Anten teorisi : Analiz ve Tasarım*. Nobel Akademik Yayıncılık.

Basir, A., & Yoo, H. (2019). A Stable Impedance-Matched Ultrawideband Antenna System Mitigating Detuning Effects for Multiple Biotelemetric Applications. *IEEE Transactions on Antennas and Propagation*, *67*(5), 3416–3421. https://doi.org/10.1109/TAP.2019.2905891.

Blauert, J., & Kiourti, A. (2014). *Dual-Band (2.4 / 4.8 GHz) Implantable Antenna for Biomedical Telemetry Applications* (Vol. 1).

Breast Cancer Facts and Statistics. (2022). Retrieved August 23, 2022, from https://www.breastcancer.org/facts-statistics#section-us-breast-cancer-statistics.

Caliskan, R., Gultekin, S. S., Uzer, D., & Dundar, O, (2015). A microstrip patch antenna design for breast cancer detection. *Procedia-Social and Behavioral Sciences*, *195*, 2905–2911.

Dey, S., & Mittra, R. (1996). Compact microstrip patch antenna. *Microwave and Optical Technology Letters*, *13*(1), 12–14. https://doi.org/10.1002/(sici)1098-2760(199609)13:1<12::aid-mop4>3.0.co;2-q.

Faisal, F., & Yoo, H. (2019a). A miniaturized novel-shape dual-band antenna for implantable applications. *IEEE Transactions on Antennas and Propagation*, *67*(2), 774–783. https://doi.org/10.1109/TAP.2018.2880046.

Faisal, F., & Yoo, H. (2019b). A miniaturized novel-shape dual-band antenna for implantable applications. *IEEE Transactions on Antennas and Propagation*, *67*(2), 774–783. https://doi.org/10.1109/TAP.2018.2880046.

Ganeshwaran, N., Jeyaprakash, J. K., Alsath, M. G. N., & Sathyanarayanan, V. (2020). Design of a Dual-Band Circular Implantable Antenna for Biomedical Applications. *IEEE Antennas and Wireless Propagation Letters*, *19*(1), 119–123. https://doi.org/10.1109/LAWP.2019.2955140.

Guardiola, M., Djafri, K., Challal, M., Ballester, M. A. G., Fernandez-Esparrach, G., Camara, O., & Romeu, J. (2019). Design and Evaluation of an Antenna Applicator for a Microwave Colonoscopy System. *IEEE Transactions on Antennas and Propagation*, *67*(8). Retrieved from https://ieeexplore.ieee.org/document/8631178/.

Hardinata, S., Deshours, F., Alquié, G., Kokabi, H., & Koskas, F. (2017). Biosensor miniaturization for non-invasive measurements of materials and biological tissues. *Proceedings - 2017 International Seminar on Sensor, Instrumentation, Measurement and Metrology: Innovation for the Advancement and Competitiveness of the Nation, ISSIMM 2017*, *2017-Janua*, 137–140. https://doi.org/10.1109/ISSIMM.2017.8124278.

Hasan, R. R., Shanto, M. A. H., Howlader, S., & Jahan, S. (2018). A novel design and miniaturization of a scalp implantable circular patch antenna at ISM band for biomedical application. *2017 Intelligent Systems Conference, IntelliSys 2017*, 166–169. https://doi.org/10.1109/IntelliSys.2017.8324286.

Institution, T. S. (2020). Death Causes Statistics. *TUIK*, p. 33710. Retrieved from https://data.tuik.gov.tr/Bulten/Index?p=Olum-ve-Olum-Nedeni-Istatistikleri-2019-33710

Iqbal, A., Alazemi, A. J., & Mallat, N. K. (2019). Slot-DRA-based independent dual-band hybrid antenna for wearable biomedical devices. *IEEE Access*, *7*, 184029–184037. https://doi.org/10.1109/ACCESS.2019.2960443.

Jones, A., Preston, S., White, M., & Hancock, C. (2019). A prototype system for optical visualisation and microwave ablation of soft tissue. *Asia-Pacific Microwave Conference Proceedings, APMC*, *2019-Decem*, 1411–1413. https://doi.org/10.1109/APMC46564.2019.9038787.

Kanj, H., & Popovic, M. (2005). Miniaturized microstrip-fed “Dark Eyes” antenna for near-field microwave sensing. *IEEE Antennas and Wireless Propagation Letters*, *4*. Retrieved from https://ieeexplore.ieee.org/document/1532173/.

Kolokotsa, D., Stavrakakis, G. S., Kalaitzakis, K., & Agoris, D. (2002). Genetic algorithms optimized fuzzy controller for the indoor environmental management in buildings implemented using PLC and local operating networks. *Engineering Applications of Artificial Intelligence*, *15*(5), 417–428. https://doi.org/10.1016/S0952-1976(02)00090-8.

Li, Q., Xiao, X., Wang, L., Song, H., Kono, H., Liu, P., … Kikkawa, T. (2015). Direct Extraction of Tumor Response Based on Ensemble Empirical Mode Decomposition for Image Reconstruction of Early Breast Cancer Detection by UWB. *IEEE Transactions on Biomedical Circuits and Systems*, *9*(5), 710–724. https://doi.

org/10.1109/TBCAS.2015.2481940.

Lin, J. Y., Chen, H. C., & Yen, M. Y. (2018). Sensor/Antenna Interface IC for Implantable Biomedical Monitoring System. *IEEE Transactions on Microwave Theory and Techniques*, *66*(3), 1660–1667. https://doi.org/10.1109/TMTT.2017.2755647.

Mahmoud, K. R., & Montaser, A. M. (2021). Design of Hyperthermia Applicator to Heat Multi-Brain Tumors Simultaneously Based on Adaptive Beamforming Technique. *IEEE Journal of Electromagnetics, RF and Microwaves in Medicine and Biology*, *5*(2). Retrieved from https://ieeexplore.ieee.org/document/9163268/.

Mahmud, M. Z., Islam, M. T., Misran, N., Kibria, S., & Samsuzzaman, M. (2018). Microwave imaging for breast tumor detection using uniplanar AMC Based CPW-fed microstrip antenna. *IEEE Access*, *6*, 44763–44775. https://doi.org/10.1109/ACCESS.2018.2859434.

McKeen, L. W. (2016). Fluoropolymer Coating Processing Technology. *Fluorinated Coatings and Finishes Handbook*, 211–226. https://doi.org/10.1016/B978-0-323-37126-1.00012-6.

Miah, M. S., Icheln, C., Islam, M. M., & Haneda, K. (2018). An Ultrawideband Conformal Antenna at 433 MHz for Wireless Capsule Endoscope of Pediatric Patients. *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC*, 350–355. https://doi.org/10.1109/PIMRC.2018.8580890.

Nalam, M., Rani, N., & Mohan, A. (2014). Biomedical Application of Microstrip Patch Antenna. *International Journal of Innovative Science and Modern Engineering (IJISME)*, *2*(6), 6–8.

Nesasudha, M., & Fairy, J. J. (2018). Low profile antenna design for biomedical applications. *Proceedings of IEEE International Conference on Signal Processing and Communication, ICSPC 2017*, 139–142. https://doi.org/10.1109/CSPC.2017.8305825.

Paracha, K. N., Rahim, S. K. A., Soh, P. J., Kamarudin, M. R., Tan, K. G., Lo, Y. C., & Islam, M. T. (2019). A Low Profile, Dual-band, Dual Polarized Antenna for Indoor/Outdoor Wearable Application. *IEEE Access*, *7*, 33277–33288. https://doi.org/10.1109/ACCESS.2019.2894330.

Peixeiro, C. (2011). Microstrip patch antennas: An historical perspective of the development. *Undefined*, 684–688. https://doi.org/10.1109/IMOC.2011.6169224.

Rabbani, M. S., & Ghafouri-Shiraz, H. (2017). Liquid Crystalline Polymer Substrate-Based THz Microstrip Antenna Arrays for Medical Applications. *IEEE Antennas and Wireless Propagation Letters*, *16*. Retrieved from https://ieeexplore.ieee.org/document/7809072/.

Rahaman, M. A., & Delwar Hossain, Q. (2019). Design and overall performance analysis of an open end slot feed miniature microstrip antenna for on-body biomedical applications. *1st International Conference on Robotics, Electrical and Signal Processing Techniques, ICREST 2019*, 200–204. https://doi.org/10.1109/ICREST.2019.8644334.

Raihan, R., Alam Bhuiyan, M. S., Hasan, R. R., Chowdhury, T., & Farhin, R. (2017). Awearable microstrip patch antenna for detecting brain cancer. *2017 IEEE 2nd International Conference on Signal and Image Processing, ICSIP 2017*, 432–436.

https://doi.org/10.1109/SIPROCESS.2017.8124578.

Rezaeieh, S.Ahdi, Zamani, A., Bialkowski, K. S., & Abbosh, A. M. (2015). Foam Embedded Wideband Antenna Array for Early Congestive Heart Failure Detection with Tests Using Artificial Phantom with Animal Organs. *IEEE Transactions on Antennas and Propagation*, *63*(11), 5138–5143. https://doi.org/10.1109/TAP.2015.2475606.

Rezaeieh, Sasan Ahdi, Antoniades, M. A., & Abbosh, A. M. (2018). Compact and unidirectional resonance-based reflector antenna for wideband electromagnetic imaging. *IEEE Transactions on Antennas and Propagation*, *66*(11), 5773–5782. https://doi.org/10.1109/TAP.2018.2866516.

Rodrigues, D. B., Ellsworth, J., & Turner, P. (2021). Feasibility of Heating Brain Tumors Using a 915 MHz Annular Phased-Array. *IEEE Antennas and Wireless Propagation Letters*, *20*(4). Retrieved from https://ieeexplore.ieee.org/document/9317736/.

Sabban, A. (2011). New compact wideband printed antennas for medical applications. *IEEE Antennas and Propagation Society, AP-S International Symposium (Digest)*, 251–254. https://doi.org/10.1109/APS.2011.5996689.

Saha, P., Mitra, D., & Parui, S. K. (2018). A circularly polarized implantable monopole antenna for biomedical applications. *Progress In Electromagnetics Research C*, *85*, 167–175. https://doi.org/10.2528/pierc18051807.

Saleeb, D. A., Helmy, R. M., Areed, N. F. F., Marey, M., Abdulkawi, W. M., & Elkorany, A. S. (2021). A Technique for the Early Detection of Brain Cancer Using Circularly Polarized Reconfigurable Antenna Array. *IEEE Access*, *9*. Retrieved from https://ieeexplore.ieee.org/document/9548093/.

Samal, P. B., Jack Soh, P., & Zakaria, Z. (2019). Compact and Wearable Microstrip-based Textile Antenna with Full Ground Plane Designed for WBAN-UWB 802.15.6 Application. *13th European Conference on Antennas and Propagation (EuCAP)*, 1–4.

Shadid, R., Haerinia, M., Roy, S., & Noghanian, S. (2018). Hybrid inductive power transfer and wireless antenna system for biomedical implanted devices. *Progress In Electromagnetics Research C*, *88*, 77–88. https://doi.org/10.2528/PIERC18061604.

Simorangkir, R. B. V. B., Yang, Y., Hashmi, R. M., Bjorninen, T., Esselle, K. P., & Ukkonen, L. (2018). Polydimethylsiloxane-Embedded Conductive Fabric: Characterization and Application for Realization of Robust Passive and Active Flexible Wearable Antennas. *IEEE Access*, *6*, 48102–48112. https://doi.org/10.1109/ACCESS.2018.2867696.

Singh, I., & Tripathi, V. S. (2011). Micro strip Patch Antenna and its Applications: a Survey. In *Article in International Journal of Computer Applications in Technology*. Retrieved from https://www.researchgate.net/publication/232318276.

Soontornpipit, P., Furse, C. M., & Chung, Y. C. (2004). Design of implantable microstrip antenna for communication with medical implants. *IEEE Transactions on Microwave Theory and Techniques*, *52*(8), 1944–1951. https://doi.org/10.1109/TMTT.2004.831976.

Suzan Miah, M., Khan, A. N., Icheln, C., Haneda, K., & Takizawa, K. I. (2019). Antenna system design for improved wireless capsule endoscope links at 433 MHz. *IEEE*

Transactions on Antennas and Propagation, *67*(4), 2687–2699. https://doi.org/10.1109/TAP.2019.2900389.

Top, R. (2017). *A Transmitter Microstrip Antenna Design and Application Towards the Detection of Heart Disease Parameters*. Selcuk University.

Top, R., Gultekin, S. S., & Uzer, D. (2017). Modeling congestion of vessel on rectangular microstrip antenna and evaluating electromagnetic signals. *2017 25th Signal Processing and Communications Applications Conference, SIU 2017*. https://doi.org/10.1109/SIU.2017.7960624.

Top, R., Gultekin, S. S., & Uzer, D. (2017). A Microstrip Patch Antenna Design to Determine the Level of Occlusion in Heart Vein. *5th International Conference on Advanced Technology & Sciences (ICAT'17)*, 644–647.

Top, R., Gültekin, S. S., & Uzer, D. (2020a). Evaluation and Comparison of Electromagnetic and Scattering Parameters Data of Two Microstrip Patch Antennas Operating in ISM Band for Cancer Detection. *1st International Conference on Computer, Electrical and Electronic Sciences*. Konya.

Top, R., Gultekin, S. S., & Uzer, D. (2020b). Evaluation and Comparison of Electromagnetic and Scattering Parameters Data of Two Microstrip Patch Antennas Operating in ISM Band for Cancer Detection. *Avrupa Bilim ve Teknoloji Dergisi*, 237–244. https://doi.org/10.31590/EJOSAT.804330.

Top, R., Unlu, Y., Gultekin, S. S., & Uzer, D. (2019). Microstrip antenna design with circular patch for skin cancer detection. *Advanced Electromagnetics*, *8*(2), 71–76. https://doi.org/10.7716/aem.v8i2.996.

Toprak, R., Sinan Gultekin, S., & Uzer, D. (2021). Investigation of Normalization Values of S-Parameters in Microstrip Antenna Structure in Pathological Tissue Samples. *Avrupa Bilim ve Teknoloji Dergisi*, *28*(28), 1366–1371. https://doi.org/10.31590/EJOSAT.1017084.

Ul Haq, M. A., & Khany, M. A. (2014). A multiple ring slots ultra wide-band antenna (MRS-UWB) for biomedical applications. *17th IEEE International Multi Topic Conference: Collaborative and Sustainable Development of Technologies, IEEE INMIC 2014 - Proceedings*, (Mi), 56–60. https://doi.org/10.1109/INMIC.2014.7096911.

Uyanik, H., Uzer, D., Top, R., & Gultekin, S. S. (2020a). Determination of Cardiovascular Occlusion with Microstrip Antennas. *4th International Symposium on Multidisciplinary Studies and Innovative Technologies, ISMSIT 2020 - Proceedings*. https://doi.org/10.1109/ISMSIT50672.2020.9254528.

Uyanik, H., Uzer, D., Top, R., & Gultekin, S. S. (2020b). Determination of Cardiovascular Occlusion with Microstrip Antennas. *4th International Symposium on Multidisciplinary Studies and Innovative Technologies, ISMSIT 2020 - Proceedings*. https://doi.org/10.1109/ISMSIT50672.2020.9254528.

Wireless, M. I. F. (2018). Circularly Polarized Implantable Antenna for Power Transmission. *IEEE Antennas and Wireless Propagation Letters*, *17*(3), 373–376.

World Health Orgabization. (2021, March 26). Breast cancer. Retrieved August 23, 2022, from https://www.who.int/news-room/fact-sheets/detail/breast-cancer.

World Health Organization. (2020). Mortality and global health estimates. Retrieved

August 24, 2022, from https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates.

Woten, D. A., & El-Shenawee, M. (2008). Broadband Dual Linear Polarized Antenna for Statistical Detection of Breast Cancer. *IEEE Transactions on Antennas and Propagation*, *56*(11). Retrieved from https://ieeexplore.ieee.org/document/4685899/.

Xia, Z., Li, H., Lee, Z., Xiao, S., Shao, W., Ding, X., & Yang, X. (2020). A Wideband Circularly Polarized Implantable Patch Antenna for ISM Band Biomedical Applications. *IEEE Transactions on Antennas and Propagation*, *68*(3), 2399–2404. https://doi.org/10.1109/TAP.2019.2944538.

Yang, S., Kishk, A., Wireless, K. L.-I. A. and, & 2008, undefined. (n.d.). Frequency reconfigurable U-slot microstrip patch antenna. *Ieeexplore.Ieee.Org*. Retrieved from https://ieeexplore.ieee.org/abstract/document/4473039/?casa_token=M2HcpwXTq4AAAAAA:zVuv_8-0kazFzy7LT-wJyjyLd72kz1NPPZEJiEuvbpJbvOoqw0D5Ud3p883_6BPdlaO5c6f3Y4k.

Yang, Z. J., & Xiao, S. (2018a). A wideband implantable antenna for 2.4 GHz ISM band biomedical application. *2018 IEEE International Workshop on Antenna Technology, IWAT2018 - Proceedings*, 1–3. https://doi.org/10.1109/IWAT.2018.8379168.

Yang, Z. J., & Xiao, S. (2018b). A wideband implantable antenna for 2.4 GHz ISM band biomedical application. *2018 IEEE International Workshop on Antenna Technology, IWAT2018 - Proceedings*, 1–3. https://doi.org/10.1109/IWAT.2018.8379168.

Zamani, A., Ahdi Rezaeieh, S., Bialkowski, K. S., & Abbosh, A. M. (2018). Boundary Estimation of Imaged Object in Microwave Medical Imaging Using Antenna Resonant Frequency Shift. *IEEE Transactions on Antennas and Propagation*, *66*(2), 927–936. https://doi.org/10.1109/TAP.2017.2780898.

Zamani, A., Darvazehban, A., Rezaeieh, S. A., & Abbosh, A. (2019). Three-Dimensional Electromagnetic Torso Imaging Using Reconfigurable Antennas. *13th European Conference on Antennas and Propagation, EuCAP 2019*, (EuCAP), 2–4.

About the Author

Rabia TOPRAK, PhD, is a Research Assistant of Electrical and Electronics Engineering at Karamanoglu Mehmetbey University in Karaman, Türkiye. She holds a PhD in Electrical and Electronics Engineering from Konya Technical University and a MSc in Electrical and Electronics Engineering from Selcuk University. Her main areas of interest are antenna structures, microstrip antennas, biomedical applications of antenna structures and electromagnetic fields.

E-mail: rabiatop@kmu.edu.tr, **ORCID**: 0000-0002-3306-1163.

Similarity Index

The similarity index obtained from the plagiarism software for this book chapter is 11%.

To Cite This Chapter:

Toprak, R. (2022). Microstrip Antennas in Biomedical Fields. In S. Kocer & O. Dundar (Eds.), *Current Studies in Basic Sciences, Engineering and Technology* (pp. 31-44). ISRES Publishing.

Concepts and Strategies for Machine Learning

Seemant Tiwari
Southern Taiwan University of Science and Technology

Introduction

The current technological movement is paving the way for a tomorrow wherein smart devices, interconnected operations, and big data coexist. Machine learning helps the system mimic and modify people's behavior. Utilizing machine learning, every connection and activity produces anything the computer could understand and use as knowledge every subsequent time. This digital world has created a massive amount of data, hastening the development of applying machine learning and methods. This paper introduces the data science strategy that allows machines to understand and do what people instinctively do: learn from the involvement. It covers the fundamentals of computer vision, including the concept, terminology, and implementations that explain when, why, and where it works. Machine - learning is an additional tool for overcoming issues that standard statistical models are inadequate. The technological upcoming of supervised machine learning and unsupervised machine learning is to evaluate and validate their commercial and industrial application possibilities. This effort's fundamental goal is to show why machine learning is the future technology.

Machine learning is the study of creating computer programs that recognize data patterns, respond to changes, and improve over time. Since technology is predictable to grip more difficult matters as they become embedded into our everyday lifestyle, machine learning has evolved from becoming one of the original goals of computer programming to becoming essential. The broad definition of learning is the act of gaining fresh or altering current habits, attitudes, information, abilities, or inclinations. Institutional theory, functionalism, and sociological all strongly emphasize education and the person being; they provide knowledge on how to advance oneself or others. Unlike what individuals instinctively do, computers utilize data to learn through involvement. At its maximum elementary level, machine learning (ML) is a subset of artificial intelligence that allows systems to reason and acquire knowledge by themselves.

The cognitive processer perspective, commonly referred to as machine learning theory, seeks to comprehend the core ideas behind training as just a computer program. This subject aims to understand at a mathematically rigorous degree whatever talents and knowledge are compulsory for learning different activities efficiently and effectively and comprehend the essential computational concepts related to teaching algorithms using data and improving quality via input (Shai Shalev-shwartz, 2014). The problems covered by machine learning theory, which incorporates concepts through both statistical as well as computational theory, include:

- Constructing statistical models representing essential characteristics of machine learning thus allows one to assess the intrinsic simplicity or complexity of specific learning tasks.
- We are designing machine learning methods that may be proven to match necessary standards and providing assurances for techniques.
- Mathematical model examining fundamental problems like: "When can we be convinced regarding predicted from small dataset?" "How much strength may full engagement over observation add to learning?" and "What types of approaches can discover despite the presence of huge amounts of deflecting data?"

Scholars in published papers have properly defined ML. Arthur Samuel first used the phrase in 1967, significant machine learning (ML) as a discipline of investigation that permits machines to study without even being computer vision (A. L. Samuel, 1967). A broad range of study areas supports the multidisciplinary area of machine learning. Algorithmic Statistical, which primary goal is to concentrate on generating forecasts using computing, has a substantial relationship to the modeling of Ml algorithms. Additionally connected to the analytics discipline are mathematics optimization's concepts, methods, and tools. Researchers are using machine learning in many different computing domains to create and scheme expressive methodologies with excellent performance outcomes, such as phishing emails filtration, social media network fraud protection, internet equities, encounter, and form sense, clinical issue, road prediction, and wind speed prediction for wind power generation, among many others.

The principal impartial of this study is to provide an outline of machine learning's history up to this moment, as well as different machine learning techniques, implementations, and difficulties. An account of machine learning's history, including its inception and advancement through the ages, is provided in this study. The research is structured as follows: Challenges with machine learning and data analysis are described in part 2. The general machine learning model is explained in part 3. Machine learning concepts are described in part 4. Explain the machine learning algorithm in part 5. Implementation of machine learning is presented in part 6. Upcoming machine learning is defined in part 7. Discussion is explained in part 8. Conclusion explained in part 9.

Challenges with Machine Learning and Data Analysis

Upon that foundation of studying and steadily gaining experience to comprehend the nature of the situation and the requirement for adaptation, machine learning is necessary for the computer to complete the job without the mediation of the user required proficiently.

- *Towards Basic Capability Activities:* One such group of activities that can be carried out by machine learning must be done effectively to analyze vast and complicated data sources, such as those used in remotely sensed, meteorology, wind power forecasting, and internet research. It is challenging for humans to forecast relevant data when there is a lot of data.

- *Humans undertake the following activities:* Humans execute many different activities daily; however, the fundamental goal is to complete each one accurately and follow a predetermined plan. As: Voice Detection

The ability of machine learning to automatically resolve data science issues is well established (Chikio Hayashi, 1998). "The dictionary defines data science as a notion of uniting statistical, data processing, machine learning, and their technical assistance to comprehend and evaluate actual processes" with data." The issue should be characterized appropriately before being solved to use the most effective machine learning method. A suitable machine learning method may be employed based on the class of problems. The following list of divisions is clarified:

- *Difficulty with Anomaly Detection:* This group includes issues that examine a trend and find modifications or irregularities in the movement. Bank card businesses, for instance, use anomaly-based algorithms to recognize variations from their customers' typical transactional activity and issue notifications anytime there is an odd purchase. These issues center on identifying the extremes.
- *The Regression Issue:* Techniques for regression serve to resolve issues that include multiple and numerical information. These would be typically applied to problems with inquiries like "what more" or "many."
- *Reinforcement Issue:* Whenever a choice must be made according to previous learning opportunities, reinforcement learning techniques have been used. Computer intelligence develops new behaviors through investigation interactions with a constantly shifting environment. It offers a method of programming machines utilizing the incentives and consequences without defining how the assignment will be completed. Thermostat procedures are a few well-known instances of reinforcement learning applications.
- *A classification issue:* A classification technique is a circumstance in which only a limited amount of outcome classifications, such as True/False, are specified prior. Contingent on the amount of O/P layers, the issue might be binary before the multi-class cataloging model.
- *Clustering Issue:* Grouping is a type of technique used in unsupervised classification. Such techniques attempt to recognize patterns in the information and create groups based on similarities in the information's architecture. The various groups or categories are therefore given names. When such a technique is developed, it inserts previously unknown data into

particular clusters.

The General Machine Learning Model

ML is utilized to address various issues that call for learning algorithms. A learning issue has three characteristics.

- Task categories
- To enhance performance measurement
- The method of expertise acquisition

Six elements comprise the entire machine learning model, irrespective of the technique used. Every model element has a professional task, as outlined in the following subsection.

- *Data Gathering and Preparedness:* The primary responsibility of the machine learning method is to gather and arrange data in a manner that may be used as algorithmic inputs. Every issue might well be addressed with a significant sum. Internet data is frequently unorganized and noisy, such as duplicated and useless data. As a result, the information must be cleansed and prepared in an organized fashion.
- *Featured Choice:* Various factors of the information collected from the previous stage might not all be pertinent to the learning experience. After these characteristics are eliminated, a selection of the crucial traits must be determined.
- *Algorithms of Selection:* Not every problem can be solved by every machine learning technique. As mentioned in the subsection above, some methods are better suited to a specific problem. Choosing the ideal machine learning technique for such a task is crucial for obtaining the best outcomes.
- *Design and parameter estimation:* Most machine learning techniques involve early operator interaction to establish more suitable settings for various variables.
- *Training:* The models must be developed using a subset of the database as training data after choosing the proper technique and model parameters.
- *Performance Assessment:* In Earlier real-time applications of the organization, the method must be verified against hidden data to assess how much has been learned using numerous presentation parameters similar to accuracy, precision, and recall.

Machine Learning Concept

Machine learning approaches can be categorized into numerous groups based on the method used to develop an algorithm and the accessibility of the result during training. They consist of supervised and unsupervised learning (Omar Y. Al-Jarrah et al., 2015; Shai Shalev-shwartz, 2014).

Supervised Learning:

Forecasting problems are supervised learning since the objective is to foresee or categorize the desired result. To formulate critical metrics that forecast the beginning and progression of various psychological diseases and suicide conduct, supervised learning has been used to massive data architectures that include demographics, medical, and social factors (Jaimiel L. Gradus, 2020; Anthony J. Rosellini et al., 2020). Additionally, prognosis ratings estimating a person's likelihood of reacting to a specific treatment for mental illness have been created using supervised learning (Webb C. A. & Cohen Z. D., 2020).

It may be beneficial to use diagnostic and prognosis assessments to pinpoint those requiring preventative measures or specialized/intensive care. Whereas most research use predictors evaluated at a one-time point to produce hazard and prognosis ratings, several techniques incorporate recurring assessments within supervised learning algorithms (implementing a data format for people and time).

Supervised machine learning creates a system that brands forecasts founded on information in the context of indecision. A supervised learning technique develops a network to produce correct estimates for reaction to new data using a current set of I/P information and previous replies to the information (O/P). Opportunities for supervised learning can be found in last data-based forecasting. The basic supervised learning techniques are Deep Learning, Neural Network, Classification, and Regression.

- CNN (convolutional neural network), RNN (recurrent neural network), Auto-Encoder, and DBN (deep belief network) are components of Deep Learning
- MLP (multi-layer perceptron), and PNN (probabilistic neural network) are components of the Neural Network
- Decision Trees, Logistic Regression, and SVM (support vector machine) are components of the Classification
- Linear Regression and SVR (support vector regression) are components of Regression

Unsupervised Learning

Data patterns or intrinsic properties are discovered through unsupervised learning. It is employed to infer conclusions from large datasets having inputs but no marked answers. Clustering is a much more popular unsupervised learning method (Haosen Yang & Robert C. Qiu, 2020; R. Zheng & J. Gu, 2019).

Unsupervised learning relates to the challenge of identifying hidden patterns in large datasets and is considered a sample group learning method. The decision of sampling or objects to the group, the feature selection to be utilized in the grouping, the conclusion of a measure of similarity for comparing specimens, and the decision of a technique to apply are conditions for executing unsupervised machine learning. The basic unsupervised learning techniques are Neural Network, Clustering, and Dimensionality Reduction.

- Neural Network includes the algorithms: RBM (restricted Boltzmann machines) and Auto-Encoder

- Clustering includes the algorithms: K-Means, Fuzzy C-Means, Hierarchical Clustering, and DBSCAN (density-based spatial clustering of applications with noise)
- Dimensionality Reduction includes the algorithms: PCA (principal component analysis), LDA (linear discriminant analysis), GDA (Gaussian discriminant analysis), NMF (non-negative matrix factorization)

Machine Learning Algorithm

This section discusses the best machine learning algorithms, their benefits and drawbacks, and application scenarios (Sandhya N. dhage & C. K. Raina, 2017; Shai Shalev-shwartz, 2014). A couple of such processes are extensively explained in Table 1 below.

Table 1. A Couple of Such Processes

No	Algorithm	Definition	Benefits	Drawbacks	Applications
1	Logistic Regression	A straightforward formula for simulating the linear relationship between I/P and category O/P	Easily understood and explicable	Able to generalize tiny, high dimensional data	Forecast for customer defection
2	Linear Regression	A straightforward formula for simulating the linear relation among an I/P and a continuous target O/P variable	More quickly trained than alternative machine learning techniques	Observant of anomalies	Share Price Forecast
3	Decision Tree	Decision Tree methods provide estimates by applying judgment rules to the structures. It can be used for regression or cataloging.	Accepts incomplete data	Observant of anomalies	Sickness prognosis
4	Random Forests	a technique for ensemble methods that integrates the results of various decision trees	Greater precision in comparison to other methods	Not very understandable	Estimating home values trends
5	Gradient Boosting Regression	Gradient Boosting Regression uses raising to create forecasting methods from a group of ineffective forecasters.	Superior efficiency over competing regression methods	High intricacy and costly processing	Estimating vehicle emissions
6	Hierarchical Clustering	A “lowest part” strategy in which each data item is regarded as a separate cluster and the nearest two groups are repeatedly combined.	Informing is the generated dendrogram.	It May not always produce the optimal clustering	Identification of scam
7	K-Means	The most common grouping method is K-Means, which establishes the K network created on euclidean distances.	Produces compact clusters	Demands the anticipated number of clusters from the start	Segmenting consumers

8	Gaussian Mixture Models	A stochastic approach for depicting clusters that are evenly spaced out in a database	Superior outcomes to those of K-means	Needs intricate setup	Networks of suggestions

Implementation of Machine Learning:

Table 2 includes a few well-known ML real-world implementations (Shai Shalev-shwartz, 2014).

Table 2. ML Real-World Implementations

No	Implementation	Definition
1	Language Detection	Nowadays, the more advanced voice recognition programs use various machine learning techniques. For continuously modifying voices, dictionaries, noises, etc., different Neural Network learning approaches for understanding hidden Markov Models are pretty efficient.
2	Forecasting	Wind speed forecasting for wind power generation (Seemant Tiwari, 2022). Also, various machine learning techniques are used for wind energy generation (Tiwari S., 2021).
3	Email Scam	Spam emails could be filtered using machine learning. The machine learning-founded algorithm will merely remember all the messages marked as junk by the client. The machine learning-based system would explore, analyze, and decide on the previous spam messages whenever a second email enters the mailbox. Incoming mail will be flagged as junk if it fits any of these; otherwise, it will be forwarded to the person's mailbox.
4	On social networks	Internet companies, including Google and Yahoo, group websites by similarities using K-means grouping. Facebook evaluates updates conveying both positive and negative feelings using Nave Bayesian'.
5	Artificial Intelligence	Machine learning is viewed as an enhanced method of resolving issues. Learning may be enhanced using basic and training information with machine learning techniques, which will advance robots and A.I. to the next stage.
6	Healthcare Sector	Numerous other medicinal balances used to assess the harshness of a patient have been advanced by consuming logistic regression

Upcoming Machine Learning

Addressing long-standing fundamental issues and creating new structures to reflect the requirements of novel machine learning techniques make up machine learning scientific discourse. Although it is difficult to predict where the following innovations will occur, there are a few subjects that one can anticipate tomorrow to hold:

- Knowing the optimal way for a machine learning algorithm to use additional data, including unsupervised learning, human cues, or taught history tasks, can help it become more adept at acquiring new knowledge. Machine Learning Theory has concentrated primarily on issues with performing a function from labeled data. Furthermore, sort of information is frequently accessible. Large amounts of unsupervised learning could be available to one and could perhaps include valuable information.

- It is strengthening the links to modern economics. "Strategic" difficulties become more significant as machine learning-based computer products are utilized in the market environment. The majority of techniques and theories up to this point have concentrated on the scenario of a single supervised learning functioning in a setting that, although it could change, lacks its own goals and tactics. Even so, if ML algorithms are to function in environments where other existing algorithms predominately behave in the best interests of their users, like when bidding on goods or engaging in different types of agreements, then we truly have a fusion of computer programming and economic systems. Many of the underlying problems are still up for debate when these factors are combined.
- They are creating learning techniques intending to incorporate learning into a more extensive network. Many machine learning techniques consider learning a different process and emphasize predictive performance as a metric. Furthermore, other concerns could surface when a machine learning technique is integrated into a more extensive network. Systems that include both how to acquire something and the procedure of choosing what to study are desirable. Such concerns have received significant conceptual attention, although a great deal is undoubtedly accomplished.

Discussion

Despite the broad range of applications of ML, there are still several difficulties. Here are a few of them:

- Machine learning techniques need a lot of data presently unavailable to academics. Because they have accessibility to such vast amounts of data, tech behemoths like Facebook and Microsoft are at the front of artificial intelligence. In industries with scant digital information, such as finance and hospital, obtaining this information becomes even more challenging, making it harder to create reliable forecasts.
- Scam identification: Provided an email account, the currently available expert systems still cannot accurately identify spam messages. In conclusion, trash is sent to the inbox while non-spam emails are sent to a trash archive.
- Machine learning systems cannot yet identify things and photos. For machine learning approaches, this area remains accessible.

Machine learning techniques are constantly being improved and will undoubtedly expand in the years ahead. They are helpful in a wide range of applications, and substantial resources are invested in expanding the frontiers of development (Bhatia M. P. S., 2008). Several significant applications that are still open are:

- For example, big data analytics and data gathering are for forecasting commercial market dynamics.

- We are analyzing natural speech, such as in search results.
- Machine learning in the health industry (diagnostic imaging, dealing with medical studies, and thinking clearly of genetic information from enormous communities).
- The development of future A.I. frameworks with technology advancements from businesses like AMD and Intel.

Conclusion

Machine Learning Theory is both a fundamental principle with numerous significant and captivating founding questionnaires and a subject matter of actual relevance that contributes to the advancement of application state-of-the-art by offering numerical structures for creating new algorithms for machine learning. Several brand-new linkages are also being made and investigated in the discipline, while new machine learning implementations pose interesting new modeling and research issues. It is reasonable to claim that the theoretical and practical applications of machine learning go well beyond what we can now imagine. This study presents a thorough analysis of the machine learning method and techniques. The goal is to comprehend the function, benefits, and limitations of machine learning as a technological solution.

References

A. L. Samuel. 1967. "Some Studies in Machine Learning Using the Game of Checkers. II—Recent Progress," in *I.B.M. Journal of Research and Development*, vol. 11, no. 6, pp. 601-617.

Anthony J. Rosellini, Siyu Liu, Grace N. Anderson, Sophia Sbi, Esther S. Tung & Knyazhanskaya, E. 2020. "Developing Algorithms to Predict Adult Onset Internalizing Disorders: An Ensemble Learning Approach," *Journal of Psychiatric Research,* vol. 121, pp. 189–196.

Bhatia M. P. S., Khalid A. K. 2008. "Information Retrieval and Machine Learning: Supporting Technologies for Web Mining Research and Practice," *Webology,* vol. 5, no. 2.

Chikio Hayashi. 1998. "What is Data Science? Fundamental Concepts and a Heuristic Example," *Data Science, Classification, and Related Methods,* Studies in Classification, Data Analysis, and Knowledge Organization, pp. 40–51.

Haosen Yang, Robert C. Qiu, Xin Shi, Xing He. 2020. "Unsupervised Feature Learning for Online Voltage Stability Evaluation and Monitoring Based on Variational Autoencoder," *Electric Power Systems Research,* vol. 182.

Jaimie L. Gradus, Anthony J. Rosellini, Erzsebet Horvath-Puho. 2020. "Prediction of Sex-Specific Suicide Risk Using Machine Learning and Single-Payer Health Care Registry Data From Denmark," *JAMA Psychiatry,* vol. 77, no. 1, pp. 25 –34.

Omar Y. Al-Jarrah, Paul D. Yoo, Sami Muhaidat, Karagiannidis GK, Taha K. 2015. "Efficient machine learning for big data: A review," *Big Data Research,* vol. 2, no. 3, pp. 87-93.

R. Zheng, J. Gu. 2019. "Anomaly Detection for Power System Forecasting Under Data Corruption Based on Variational Auto-Encoder," *I.E.T. Digital Library*.

Sandhya N. dhage, Charanjeet Kaur Raina. 2017. "A Review on Machine Learning Techniques," *International Journal on Recent and Innovation Trends in Computing and Communication,* vol. 4, no. 3, pp. 395 – 399.

Seemant Tiwari. 2022. "Wind Speed Forecasting Methods for Wind Energy Generation," *International Conference on Informatics,* pp. 143-147.

Shai Shalev-shwartz and Shai ben-david. 2014. "Understanding Machine Learning: From Theory to Algorithms," *Cambridge University Press,* https://www.cs.huji.ac.il/~shais/UnderstandingMachineLearning/

Tiwari S., J. M. Ling. 2021. "A Review of Wind Energy Forecasting Techniques," *International Conference on Technology and Policy in Energy and Electric Power,* 2021, pp. 213-218.

Webb C. A., Cohen Z. D., Beard C., Forgeard M., Peckham A. D., Bj. T. 2020. "Personalized Prognostic Prediction of Treatment Outcome for Depressed Patients in a Naturalistic Psychiatric Hospital Setting: A Comparison of Machine Learning Approaches," *Journal of Consulting and Clinical Psychology,* vol. 88, no. 1, pp. 25–38.

About The Author

Seemant Tiwari graduated with a B. Tech in Electrical Engineering from AKTU in Lucknow, India. He earned his M. Tech. in Power Electronics in May 2013 from Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology in Chennai, India. He obtained his Post Graduate Certificate Program in November 2013 from the Fluid Control Research Institute in Palakkad, India. He has been working on a Ph.D. since September 2019 at the Department of Electrical at Southern Taiwan University of Science and Technology in Tainan City, Taiwan. Before joining STUST in Taiwan, he was a lecturer in the Electrical Department at RBS-Polytechnic in Agra, India.

E-mail: da82b203@stust.edu.tw, **ORCID:** 0000-0002-6886-5508

Similarity Index

The similarity index obtained from the plagiarism software for this book chapter is 12%.

To Cite This Chapter:

T. Seemant (2022). Concepts and Strategies for Machine Learning, In S. Kocer, O. Dundar (Eds .), *Current Studies in Basic Sciences Engineering and Technology*, (pp. 45–54). ISRES Publishing

SPACECRAFT POWER SYSTEM TECHNOLOGIES

Omer Faruk Keser
Necmettin Erbakan University

Bayram Bulat
Staff Officer, Operations Plans & CIS Division, NATO NAEW&C Force Headquarters, Geilenkirchen, Germany.

Introduction

Space platforms are configured based on three basic systems. The first of these is the ground segment, in which the ground station, command and control center and operating units are located, and the other is the launch center and launch vehicle, and finally the space segment in which the space platform or satellite is located.

The most important and decisive factor in a space platform is the payload and its configuration. Payloads differ, sometimes according to the mission of examining the physical elements of space or observing space objects, sometimes according to the mission of communication to transmit information. The function of a communications satellite is to receive signals, amplify and retransmit the received signal. All of these processes are carried out by the satellite's subsystems.

Although there may be differences in the subsystems it has depending on factors such as mission type, mission region, mission duration, structure, in the most general sense, the subsystems of a space platform; It can be counted as communication, propulsion, telemetry, tracking, control, and monitoring (TTC&M), power, attitude control, structure, thermal control, and command and data handling subsystems, especially payload.

Given the vital role that satellites play in contemporary society, it is crucial to maximize their technological efficiency and, consequently, the rate of return for investors. The power system, which produces the energy required for the satellite to operate, determines the satellite dimensions, operating life, mission efficiency and cost depending on the useful payload of the satellite. In this process, reducing weight and ensuring sustainability are considered very important in improving the performance of power systems for the satellite (Keser, et al., 2017; Idare, B., 2018).

When the Sputnik-1 satellite, which is considered to be the beginning of the space age, was sent into space on October 4, 1957, the silver-zinc battery on it was the only energy source. The battery provided a total of 1 W of power to the two signal transmitters on the satellite. At the end of 3 weeks, signal reception from the transmitters ended with the depletion of the battery (Hyder et al., 2003). Today, the International Space Station, with 32800 solar cells and a total power generation capacity of 248 kW, continues its activities in the space environment as a science center where scientists from many nations conduct space studies. While the amount of power needed in the first satellite in 1957 was 1 watt, the power requirement of the space platforms designed today has increased to hundreds of kW, enabling the design of different power systems in solar, thermal, nuclear and chemical structures that can meet the energy needs of these amounts.

While the power systems are being designed, besides the energy needed by the space platform, the power system through which this energy will be provided is equally important, and the main factor determining this is the mission area of the space platform. Because the primary power source for a spacecraft designed for deep space missions will be a nuclear power-based power system, while the primary power source for a space platform in earth orbits is likely to be photovoltaic systems.

The system that manages the process from the production of the energy determined for the payload and carrier system of the satellite to the distribution in a controlled manner is the satellite power system.

Spacecraft Power Subsystem

A spacecraft power subsystem usually consists of 3 basic components; primary and secondary power sources and a power control/distribution unit. These items are shown schematically in Figure 1.

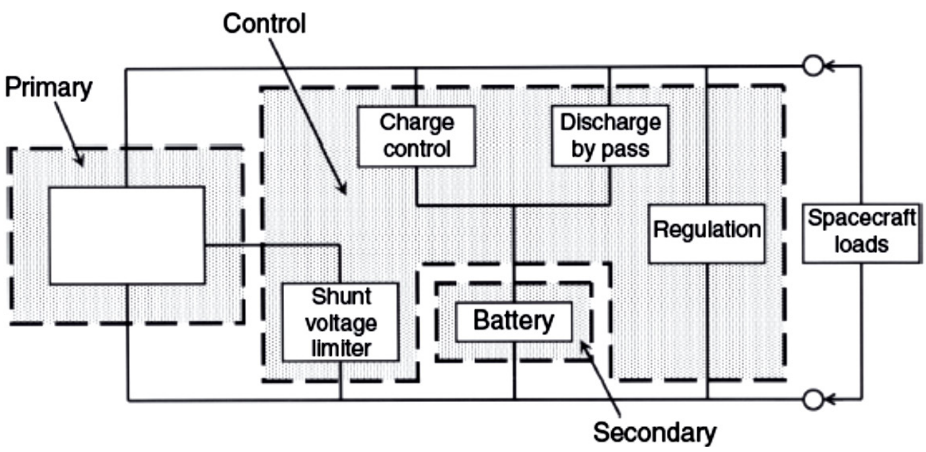


Figure 1. Schematic representation of the power system elements of a spacecraft (Fortescue et al., 2011)

The primary power system converts a fuel into electrical energy. In the first spacecrafts, the electrical energy was provided by batteries. Batteries doesn't convert a fuel into electrical energy, only use stored energy. In today's spacecrafts, solar cells are often used as the primary power system. While fuel cells are used as the primary power source in short-term manned space missions, fuel cell-solar cell hybrid systems are used if the mission period is longer. Hydrogen and oxygen molecules are combined to produce heat and water as the outcome of the reaction in the fuel cells. While the heat can be used to generate electricity, the water can be used for the needs of astronauts in spacecraft during the space mission. Nuclear powered systems use either a radioactive decay process (for RTGs) or a nuclear fusion process as their energy source. While RTGs benefit from the thermoelectric effect, fusion reactors operate in a similar way to nuclear power plants on the world (Bennett, 1995). Figure 2 shows the cycle of the energy produced by the primary power systems and its transmission to the related units with the power control/ distribution network.

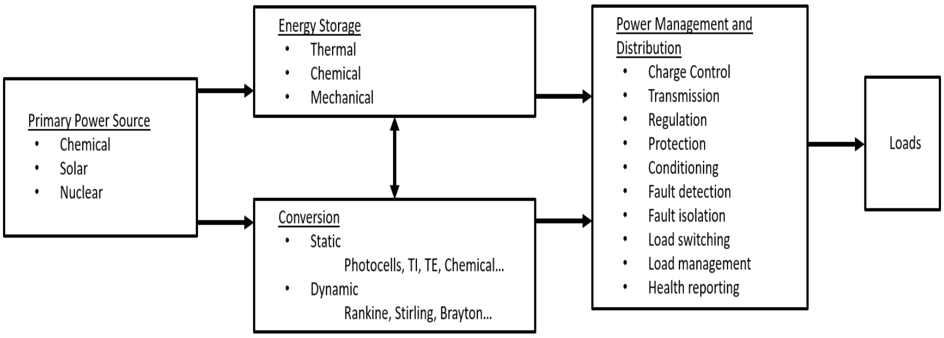


Figure 2.. Power generation-distribution cycle (Hyder et al., 2003).

The secondary energy source is used to store the converted electrical energy and distribute the energy to the spacecraft subsystems and its payloads in the absence of a primary energy source. For instance, secondary batteries are used when a solar powered spacecraft/satellite stays in the dark side of its period and primary power systems are not functional. Table 1. shows the specifications of some battery types used as secondary power systems.The power control and distribution system regulates the load-voltage values and distribute electrical energy to all spacecraft subsystems.

Table1. Secondary Batteries And Their Specifications (Hyder et al., 2003)

	Silver Zinc	Nickel Cadmium	Nickel Hydrogen
Energy Density (W.h/kg)	90	35	75
Energy Density (W.h/dm^3)	245	90	60
Oper Temp (^{o}C)	0 – 20	0 – 20	0 - 40
Storage Temp (^{o}C)	0 – 30	0 – 30	0 – 30
Dry Storage Life	5 years	5 years	5 years
Wet Storage Life	30 – 90 days	2 years	2 years
Max cycle life	200	20.000	20.000
Open circuit (V/cell)	1,9	1,35	1,55
Discharge (V/ cell)	1,8-1,5	1,25	1,25
Charge (V/cell)	2,0	1,45	1,50
Manufacturers	Eagle-Pitc-her, Yardney Te-chnical Prod	Eagle-Pitcher, Gates Aerospace Batteries	Eagle-Pitcher, Yardney, Gates, Hughes

Primary Power Systems

The primary energy system converts a fuel into electrical energy. In the first spacecrafts, only batteries were used as the power source and thus there were no conversion process. Batteries doesn't convert any energy source to electrical energy, only use the stored energy inside. Later, with the developing technology, the use of fuel cells, photovoltaic and nuclear power systems in space missions had become widespread. Different primary power systems are shown in Figure 3. As clearly seen on the figure, photovoltaic and nuclear power systems are the most suitable power systems for spacecraft and satellites.

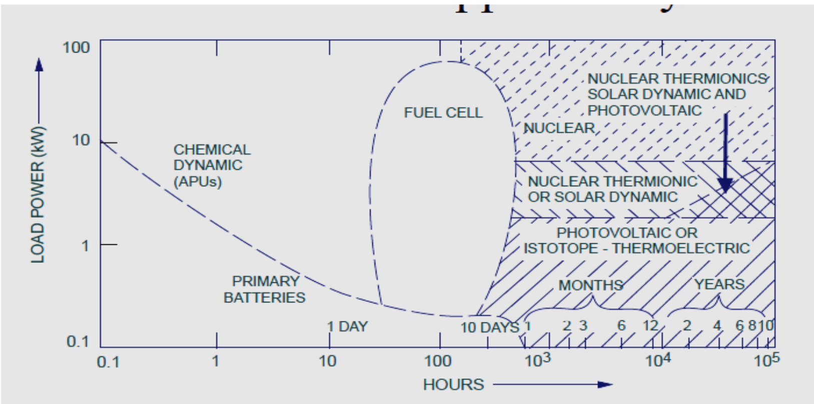


Figure 3. Primary power systems and their lifetime (Miller et al.)

Although they are named differently in different references, in the most basic sense, primary power systems can be listed as in Figure 4.

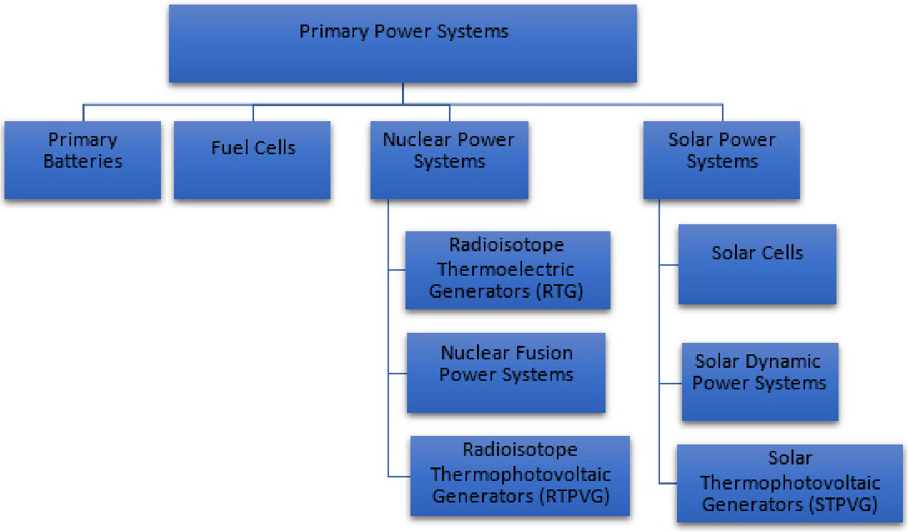


Figure 4. Primary power systems

Primary Batteries

Primary batteries convert chemically stored energy into electrical energy. The difference between primary and secondary batteries is that while the secondary batteries are used to store the electrical energy produced in the primary power systems, primary batteries are the only power generation system in the spacesraft and the energy capacity of the spacecraft is limited with the energy stored in the primary battery. Battery types used as primary batteries and their properties are shown in Table 2.

Table 2. Primary Batteries (Miller et al.)

	Silver zinc	Lithium sul-fur dioxide	Lithium carbon monofluoride	Lithium thionyl chloride
Energy Density (W.h/kg)	130	220	210	275
Energy Density (W.h/dm^3)	360	300	320	340
Op Temp (^{o}C)	0 – 40	-50 – 75	? - 82	-40 - 70
Storage Temp (^{o}C)	0 – 30	0 – 50	0 – 10	0 - 30
Storage Life	30-90 days wet, 5 years dry	10 years	2 years	5 years
Open circuit voltage (V/cell)	1,6	3,0	3,0	3,6
Discharge voltage (V/cell)	1,5	2,7	2,5	3,2
Manufacturers	Eagle-Pitc-her, Yardley	Honeywell, Power Conver	Eagle-Pitcher	Duracell, Altus, ITT

The first use of primary batteries dates back to Sputnik-1, which was sent into space by the Soviet Union in 1956. The functional life of Sputnik-1 ended with the discharge of the Silver-Zinc type primary batteries, since there was no other power system to support these batteries in the spacecraft. In the early stages of space studies, Silver-Zinc batteries were preferred, Nickel-Cadmium batteries were used for the next 20 years, and Nickel-Hydrogen and Lithium-Ion batteries have been used since the 1980s (Hyder et al., 2003).

Fuel Cells

The fuel cells shown in Figure 5 have been used as the primary power system, especially in space shuttles and manned space missions. Hydrogen and oxygen molecules are combined in electrochemicals in fuel cells resulting in water and electrical energy as the outcome of the reaction. It is especially designed for manned space missions, and hybrid power systems with photovoltaic-fuel cells are used in these missions. The water resulting from the electrochemical reaction can be used to meet the water needs of astronauts. Due to the necessity of carrying fuel, a fuel cell has limited duty time, heavy weight and needs a large storage area. Contrary to solar cells, a fuel cell is not affected by the external environment, since it is a closed system, It provides opportunity to produce electrical energy both in the sun-exposed period and in the dark period of the orbit.

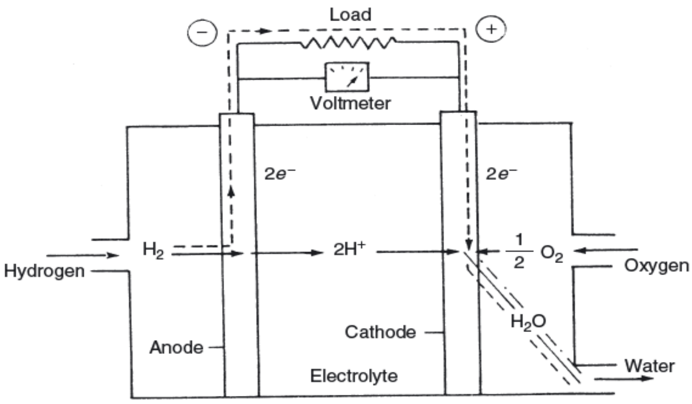


Figure 5. Schematic representation of a fuel cell (Fortescue et al., 2011)

Nuclear Power Systems

There are many nuclear power systems that have been used effectively in space studies for both power generation and heating since 1956. The basic principle of nuclear power systems is to generate electrical energy with different methods by making use of the heat emitted by the radioisotope material.

Radioisotope Thermoelectric Generators (RTG)

Space and terrestrial power applications using radioisotopes have been in use since 1956. RTGs are the most effective power generation devices, especially for deep space and planetary missions (Keser et al., 2019).

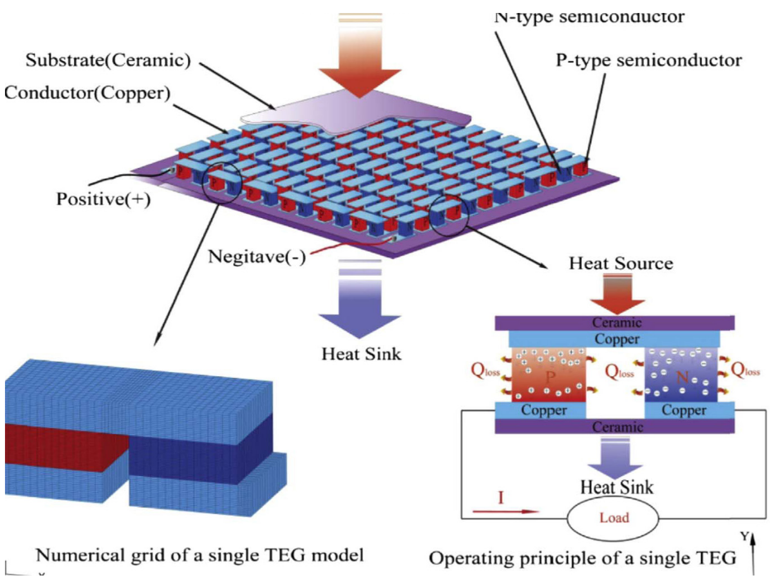


Figure 6. Working principle of a thermoelectric generator module (Twaha et al., 2016)

The Radioisotope Thermoelectric Generator is basically a device that produces electrical energy from the heat emitted by the radioactive material inside, based on the Seebeck effect method. Figure 6 shows the working principle of a thermoelectric generator modüle. NASA uses Plutonium 238 as the radioactive heat source in most of the RTGs it produces. While the heat inside RTGs are around 1000˚C, heat transfer elements are used for cooling and removing the heat. Silicon-Germanium (SiGe) is the most suitable material for use in RTGs, which operate at such high temperatures with a relatively low efficiency of 6.6% (Fagas et al., 2016). Figure 7. shows the structure of an RTG.

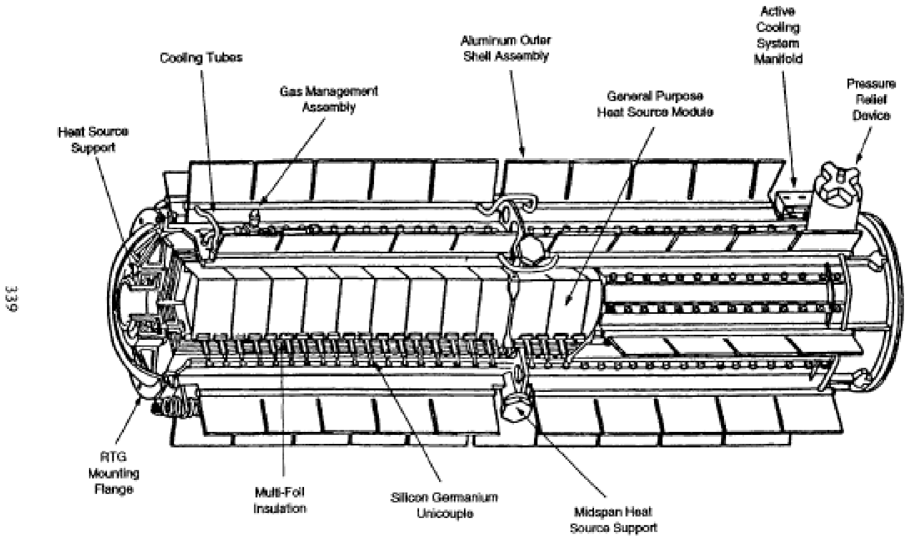


Figure 7. Schematic representation of a Radioisotope Thermoelectric Generator (RTG) (Pisacane, 2005)

Space and terrestrial power applications using radioisotopes have been in use since 1956. SNAP-3, one of the first devices developed in this context, was first introduced in 1959. SNAP-3 was the size of a bunch of grapes, weighed 1.8 kg, had a power of 2.5 W and was capable of producing 11.6 kW/h of energy in a 280-day period. A Nickel-Cadmium battery that can provide the same amount of energy should weigh 315 kg. SNAP-3 used $_{210}$Polonium alpha emitter isotope as its heat energy source (Raghep, 2012). A total of 45 RTGs in the USA have been used in 26 spacecraft designed for NASA's LEO orbit, missions on the Moon and Mars surface, and research-exploration missions on near-far planets for different missions of NASA.

RTGs are often used in deep space exploration and interplanetary missions. However, it is not cost-effective compared to solar cells and has many problems. RTGs operate at very high temperatures (about 1000°C) resulting in material problems, limiting the operational life of the spacecraft. Another important problem is that RTGs use radioactive materials as fuel, which is a very expensive, rare and dangerous material. These problems result in complex security restrictions and costly ground handling procedures. These factors make the costs of RTGs tens of times more expensive than solar cells, and this limits the use of RTGs to areas where solar cells cannot be used (Pisacane, 2005).

RTGs are not only used to generate energy, but also to protect the electronics, opto-electronics and other systems on the spacecraft from extremely low temperatures of deep space. By directing the waste heat emitted by the nuclear fuels in the RTGs to the related devices, this problem can be solved with the waste heat energy of the RTGs without the need to use extra energy to heat these devices. RTGs are available in miniature size versions manufactured for heating purposes only. These miniature devices called LW-RHU (Low-Weight Radioisotope Heater Unit) have been used many times in spacecrafts designed especially for deep space missions by the USA (Summerer, 2006).

Nuclear Fusion Power Systems

It works on the same principle as the nuclear fusion systems used on earth. It has been used in spacecraft and deep space missions where high power requirements are required. Uranium-235 is used as a radioactive material. They were mostly used by Soviet Russia and are not used today.

Radioisotope Thermophotovoltaic Power Systems

Although RTGs have been used in many space missions by the USA and Russia, their efficiency is very low compared to photovoltaic and other primary power systems. The main reason for the low efficiency of RTGs is that the efficiency of the SiGe thermoelectric modules inside RTGs are between 5-10%. Today, studies are carried out on different designs to solve the low efficiency problem of RTGs. One of the most striking one of these studies is the Radioisotope Thermophotovoltaic Generator (RTPVG), which is seen in Figure 8. While the efficiency of thermoelectric modules that determine the efficiency of RTGs is between 5-10%, the efficiency of thermophotovoltaic (TPV) modules that determine the efficiency of RTPVGs is 20-25% (Wernsman et al., 2004) and has the potential to increase up to 40% (Datas, 2015). It has been seen as a potential primary power system to replace RTGs.

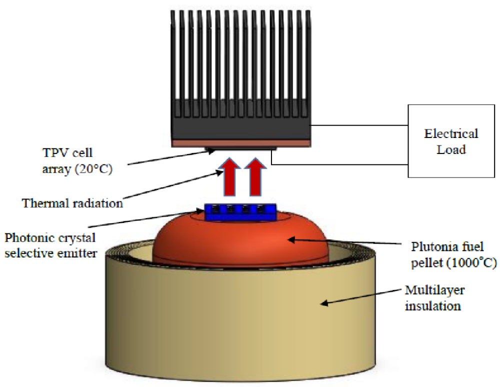


Figure 8. Radioisotope Thermo-Photovoltaic Generator (RTPVG) demonstration (Wang et al., 2015)

Radiozotope Thermophotovoltaic Generator (RTPVG) power systems were first introduced in the mid-1990s as an alternative system to RTGs and as a power system that could reach a power/mass ratio of 15 W/kg. This ratio means that it has a 3-5 times higher power/mass ratio than RTG and its derivatives (MMRTG(2.8 W/kg), GHPS-RTG(5.1 W/kg)). In line with this data, NASA started the first RTPVG project in 2003, 25 series-connected InGaAs TPV cells were used in the first prototype, and 50 watts of power was planned to be generated from a 250 watt heat source. In this context, the targeted efficiency was around 18-20% and the power/mass ratio was 17 W/kg, but the data obtained were far from these values. With the developing technology and production techniques over time, InGaAs TPV cells which has 20% efficiency have been produced (Datas et al., 2017).

Solar Power Systems

Solar power systems generate electrical energy by utilizing the thermal or photovoltaic properties of solar energy. While photovoltaic energy conversion is available in solar cells, solar dynamic and solar thermal systems intensify solar energy and take advantage of its thermal properties.

Air Mass Coefficient (AM) and Solar Spectrum

The air mass coefficient is used to determine the solar radiation spectrum at the targeted point after the sun's rays move in and out of the atmosphere (in space). The most common use of the air mass coefficient is to provide standardizations in order to determine the efficiency of photovoltaic cells. It is usually symbolized by the abbreviation AM followed by numbering formats. While determining the efficiency of photovoltaic cells produced for use on earth, the value AM1.5 is the most used air mass coefficient. The AM0 value is always used when determining the efficiency of photovoltaic cells designed for use in artificial satellites and other space platforms. As shown in Figure 9., the determining factor in determining AM values is the incidence angle of the sun's rays to the earth. Accordingly, the AM value is determined as follows:

$$AM = 1/\cos\theta \quad (1.1)$$

As shown in Figure 9, the angle ϴ is the angle between the point reached by the sun's rays and the surface normal (zenith). The AM value is calculated using the ϴ angle, and according to the AM value, the solar radiation intensity (W/m^2) and solar radiation spectrum can be reached in the region where the sun's rays fall, as shown in Table 3.

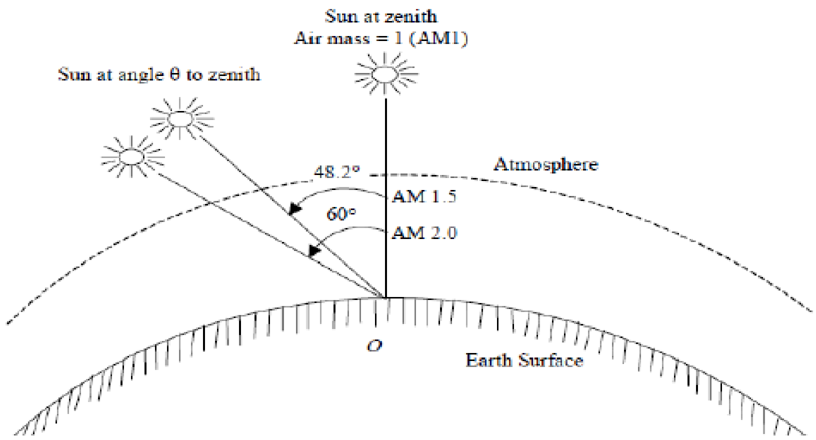


Figure 9. Relationship between AM values and incidence angle of sun rays

Outside the atmosphere, the solar radiation value in outer space is expressed by AM0 and is 1366 W/m^2. AM values according to the angles of reaching the earth and radiation values are shown in Table 3.

Table 3. AM values table

Incidence angle	AM	Standard	Solar radiation intensity (W/m^2)
-	0	ASTM E-490	1366,1
0°	1		1040
48.2°	1.5	ASTM G-173-03	1000,4
60°	2		840
75°	3.8		620
90°	38		20

As shown in Table 3, while the solar radiation intensity is 1000 W/m^2 in AM1.5, which is accepted as the common value for the sun rays reaching the earth, the intensity of sunlight emitted in space (AM0) is approximately 1.36 times of this value. Figure 10 shows the solar spectrum of solar radiation at different AM values. As seen in the solar spectrum, 5% of the sun's rays are in the ultraviolet (0-400 nm), 43% in the visible light region (400-700 nm), and 52% in the infrared region (700-2600nm) (Kruse et al., 2006).

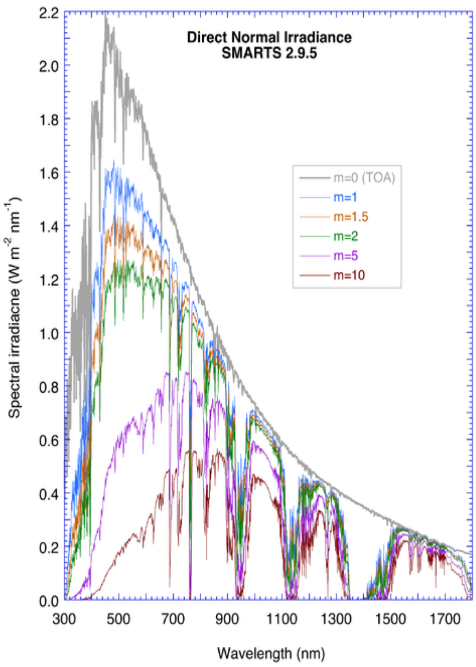


Figure 10. Solar spectrum of solar radiation values for different AM values (Gueymard, 2004)

Solar Cells (Photovoltaics)

Photovoltaic energy systems are energy conversion systems in which electromagnetic waves in the part of the visible light spectrum of solar spectrum are converted into DC current with semiconductor materials having photovoltaic effect. Photovoltaic energy conversion is explained and illustrated in Figure 11 (Apostoleris et al., 2018).

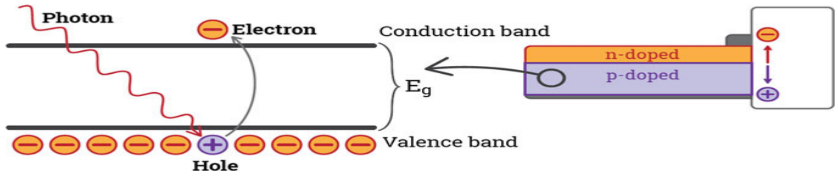


Figure 11.1. The operating principle of a photovoltaic cell: valence-band electrons in a semiconductor are excited to the conduction band by the energy of incident photons. Conduction band electrons and valence band holes are separated by a PN junction, causing current to flow through the external circuit (Apostoleris et al., 2018)

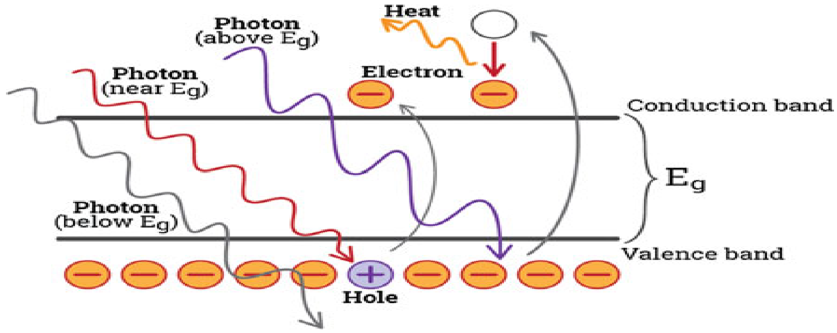


Figure 11.2. Interaction of photons of different energies with a PV cell: photons with energy below the band gap are transmitted; those with energy above the band gap create a carrier pair by promoting an electron, but the difference between the photon energy and the band gap becomes heat (Apostoleris et al., 2018)

In space studies, solar cells were first used in 1958 on the Vanguard and Sputnik-3 spacecrafts. The first solar cells used in space studies were of monocrystalline silicon structure and had an efficiency of approximately 10% (Kalogirou, 2017). Although the efficiency of these cells increased to 14% and above with the developments in silicon solar cell technology in the 1960s-1970s, Ga-As solar cells which were developed in the early 1960s and had lower efficiencies (11%) were mostly preferred in space studies due to their stability at high temperature and radiation.

Today, in parallel with the studies carried out in the field of photovoltaics and space, great developments have also been achieved in the field of space-qualified solar cells. Multilayer space-qualified solar cells with 28-32% efficiency have been used in many space platforms.

CPV (Concentrated PV) cells which have been produced within the scope of studies on multilayer solar cells also offer high efficiency energy conversion in lower PV cell areas compared to solar cells. As shown in Figure 12., solar energy concentrated up to 1000x by optical systems focuses on CPV cells, providing the opportunity to produce electrical energy at high power density. The efficiency values of CPV cells reached 44%. It is considered that they have a high potential for use in space platforms due to their low cost and high efficiency power generation in lower cell areas (Keser et al., 2019).

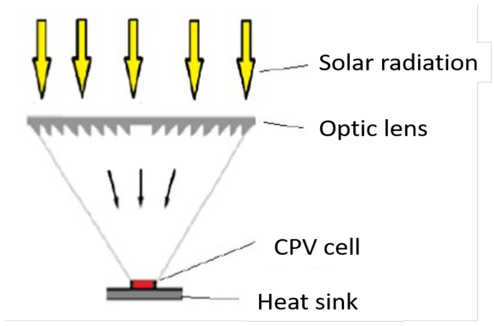


Figure 12. CPV cell power generation (Selimoğlu, 2013)

Solar Dynamic Power Systems

In these systems, high heat is obtained by concentrating sunlight on the receiver. Electric energy is obtained by circulating the gas/liquid heated by the heat obtained on the circuit. Studies have shown that it is 25% more efficient than photovoltaic solar cells. With such systems, it is possible to produce electrical energy up to 25 kW. Although studies on this subject have been carried out by NASA, it has not yet been used in a real space mission.

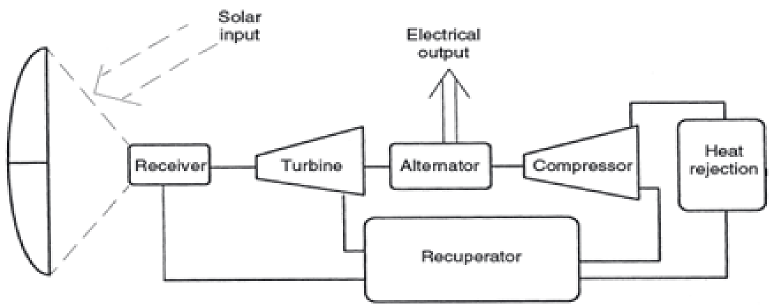


Figure 13. Solar dynamic Brayton cycle (Fortescue et al., 2011)

Figure 13 shows a solar dynamic Bryton cycle designed by NASA for use on the International Space Station. While the International Space Station was being built, it was designed to replace the 75 kW output solar cells currently in use with two 25 kW solar dynamic Bryton cycle power systems. Helium and Xenon gasses would be used in the cycle, and a temperature of 1042ºK would be achieved there by focusing the sunlight on the receiver with the concentrators. However, the project could not be implemented due to high cost and budget constraints (Fortescue et al., 2011).

Solar Thermophotovoltaic Power Systems

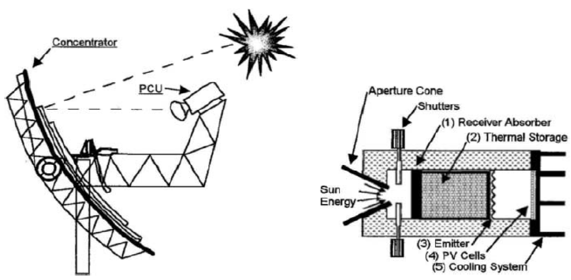


Figure 14. Solar thermophotovoltaic power system diagram (Hyder et al., 2003)

As shown in Figure 14, the sun rays are concentrated on the receiving system, resulting in a temperature of over 1000˚C in the thermal storage. With the temperature obtained, the emitter material emits infrared radiation, and the infrared rays emitted reach onto the photovoltaic cells selected in accordance with the emitted wavelength, and electrical energy is obtained from there.

Although solar thermophotovoltaic power systems (STPVG) have not been used in a real space mission, it has shown that it can be among the future space power systems with its high efficiency and low weight. A typical photovoltaic power system consists of solar cells and a battery pack. A satellite in LEO orbit spends approximately 30 minutes of each orbital period in the dark side. While in the dark part of the period, the satellite uses energy stored in secondary batteries. For this reason, instantaneous energy production at least one and a half times the instantaneous need is mandatory in power systems working with solar cells. The most striking feature of thermophotovoltaic power systems is that they continue to produce energy in the dark part of the orbit. In this way, they do not need extra energy production or a secondary battery group to store the energy produced, as in solar cells. Thanks to these advantages, it is possible to design power systems that are up to 20% more efficient than solar cells (Hyder et al., 2003).

While factors such as high efficiency, low cost and high production and supply network in market conditions can be shown as the main reason for the preference of solar cells in all space platforms designed for orbital missions, factors such as being affected by harsh space conditions, low operating temperature range, and high volume-mass ratio. There are also negative aspects. On the other hand, the only reason why RTGs are preferred in planetary and deep space missions despite their safety problems and very low efficiency values is that they are the only primary power system that can actively produce power in the harsh conditions of the region where they are operating. However, due to the high potential of TPVGs, which have not yet been used in a real space mission, to be the primary power system of the future, there are findings in the relevant literature that innovative power systems should be designed for application (Wernsman et al., 2004; Datas, 2015; Wang et al., 2015).

In this context, in a study carried out to investigate the designability of a TPVG-based power system for space applications that can eliminate the negative aspects of existing primary power systems, studies to develop the designs of TPVG power systems, which are considered to be the primary power system of the future for space platforms, and the availability of these power systems in space platforms has been researched.

As seen in the references, the information presented in this section was obtained from the previous studies by the authors and other literature, and with a general evaluation, the introduction part of a postgraduate study in this field (Bulat, B; 2021) was used.

References

Apostoleris H., Stefancich M., Chiesa M., 2018, Concentrating photovoltaics (CPV): the path ahead, Springer.

Bulat B., 2021, Designing a thermophotovoltaic power system for space platforms, Master Thesis in Electrical and Electronics Engineering, The Graduate School of Natural And Applied Science, Necmettin Erbakan University, Konya, Türkiye.

Bennett G., 1995, Summary of the US use of space nuclear power, ESA SP 369, 163–169.

Datas A., 2015, Optimum semiconductor bandgaps in single junction and multijunction thermophotovoltaic converters, *Sol. Energy Mater. Sol. Cells,* 134 275–290.

Datas A., Marti A., 2017, Thermophotovoltaic energy in space applications: review and future potential, *Solar Energy Materials & Solar Cells,* 161 285–296.

Fagas G., Gammaitoni L., Paul D., Berini G.A., 2014, "ICT –energy concepts towards zero–power information and communication technology", pp 211-212.

Fortescue P., Swinerd G., Stark J., 2011, Spacecraft Systems Engineering (Aerospace Series), Wiley.

Gueymard C.A., 2004, The sun's total and spectral irradiance for solar energy applications and solar radiation models, Solar Energy, vol. 76, 423-453.

Hyder A. K., Wiley R. L., Halpert G., Flood D. J. and Sabripour S., 2003, Spacecraft power technologies, Imperial College Press, London.

Kalogirou S., 2017, McEvoy's handbook of photovoltaics, Third Edition, Fundamentals and Applications, Elsevier Science Academic Press

Keser, O. F., Yenisoy, A., Idare, B., 2017. Development of Space Qualified Microlens Arrays for Solar Cells Used on Satellite Power Systems. Sakarya University Journal of Science, 21(4), pp. 627-635.

Keser O. F., Bulat B., Idare B., Okan A., 2019, The Usability of PV-TEG Hybrid Systems on Space Platforms, IEEE, 22 July 2019, DOI: 10.1109/RAST.2019.8767785, https://ieeexplore.ieee.org/document/8767785.

Keser O. F. and Idare B., 2019, Concentrated Photovoltaic Solar Power Investigation in Space Application, 978-1-7281-3789-6/19, IEEE.

Keser, O. F., Idare, B., "Designing Anti-Reflective Microlens Arrays for Space Solar Cells" Solar Energy Materials and Solar Cells, ISSN:0927-0248, 200, 15 Sept. (2019) 110003

Miller D.W. and Keesee J., 2012, Spacecraft Power Systems.

Pisacane, V.L., 2005, Fundamentals of space systems' (2nd Edition), Oxford University Press.

Raghep M., 2012, Radioisotope power generation, NPE Chapter 3.

Selimoğlu O., 2013, Design and realization of a new concentrating photovoltaic solar energy module based on lossless horizontally staggered light guide, PhD Thesis, Physics Dept. Middle East Technical University.

Summerer L., 2006, Technical aspects of space nuclear power sources.

Twaha, S.,Zhu, J., Yan, Y., Li, B., 2016, A comprehensive review of thermoelectric technology: materials, applications, modelling and performance improvement, Journal of Renewable and Sustainable Energy Reviews.

Wang X., Chan W.R., Stelmakh V., Soljacic M., Joannopoulos C.D., Celanovic and

Fisher P.H., 2015, Prototype of radioisotope thermophotovoltaic system using photonic crystal spectral control, Journal of Physics: Conference Series 660 (2015) 012034.

Wernsman B., Siergiej R.R., Link S.D., Mahorter R.G., Palmisiano M.N., Wehrer R.J., Schultz R.W., Schmuck G.P., Messham R.L., Murray S., Murray S.C., Newman F., Taylor D., DePoy D.M., Rahmlow T., 2004, Greater than 20% radiant heat conversion efficiency of a thermophotovoltaic radiator/module system using reflective spectral control, IEEE Trans. Electron Devices 51 (3) 512–515.

About the Authors

Omer Faruk KESER, He is a professor, Head of Department and senior researcher at Konya Necmettin Erbakan University, Department of Aerospace Engineering. He held academic positions at different universities in Turkey and made great contributions to the ongoing technological academic development in the region as the Regional Coordinator of TUBITAK - Turkish Scientific and Technological Research Council. He was the Coordinator of the OPMER Precision Optical Systems Development Laboratory, which was established to develop Satellite Camera Optics at the TUBITAK Space Technologies Research Institute, where the communication satellite Turksat 6A, which will be launched in 2023, and the observation satellite IMECE were developed locally and nationally. His main research areas are satellite power systems, wireless power transmission technologies, electromagnetism, physics, space environment, satellite technologies and space applications. He also has academic studies published in various national and international respected journals.

E-mail: ofkeser@erbakan.edu.tr, **ORCID**: 0000-0003-1138-4815

Bayram BULAT, is a staff officer representing Türkiye at NATO AEW&C Force Headquarters in Germany. He graduated from the Air Force Academy and started to work as an air traffic controller in Ankara. In August 2020, he was promoted to captain rank and assigned to NATO AEW&C Force Headquarters in Germany as an operations planning officer. In December 2021 he completed his master thesis in the field of spacecraft power technologies and had the degree of master of science in electrical and electronics engineering. His main research interests are spacecraft power systems, semiconductors and energy harvesting.

E-mail: bayram.bulat@gmail.com, **ORCID:** 0000-0002-5938-2942

Similarity Index

The similarity index obtained from the plagiarism software for this book chapter is 8%.

To Cite This Chapter:

Keser O. F., Bulat B.(2022). Spacecraft Power System Technologies, In S. Kocer, O. Dundar (Eds .), *Current Studies in Basic Sciences Engineering and Technology* (pp. 55–69). ISRES Publishing.

Vehicles of the Near Future: Driverless Autonomous Cars

Tarık UNLER
Necmettin Erbakan University

Introduction

With the development of technology, many institutions and organizations from automobile manufacturers, electronics, software industry and universities are working very seriously on autonomous driving systems, which are seen as the technology of the future. These institutions and organizations continue to make highly automated vehicle prototypes and road tests, whose capacities are rapidly developing due to the development of advanced sensor processing technologies, adaptive algorithms, and high-definition mapping systems. Autonomous vehicles perform tests in the United States almost exclusively in autonomous driving mode. These vehicles make long-distance highway and arterial trips in Europe and Japan (Vanderblit, 2012). In addition, the effects caused by reasons such as safety concerns in traffic, the high number of property damage and fatal accidents, environmental effects, technological developments and the increase in comfort expectations lead to a revolutionary change in personal transportation vehicles, which have an important place in our daily lives (Tastan & Kaymaz 2017).. These reasons also play a major role in the development of autonomous vehicles. Autopilots are used today to perform very different autonomous driving operations for aircraft, since air traffic is much more convenient than ground traffic in aircraft. These systems, which make almost less errors than a pilot in terms of error rate, continue to be developed very widely. As in the weather example, the production of vehicles capable of autonomous driving is of great importance in order to prevent accidents, especially by preventing driver errors. The heavy ground traffic compared to air traffic has many problems such as physical obstacles and pedestrians. Especially as the processing capacity and speed of the controllers to be used in this regard increased, the problems that were seen as difficult before started to be solved one by one. Being able to determine the traffic signs provides a serious convenience in the decision-making mechanisms of the estimation of the object approaching the vehicle (pedestrian, animal, vehicle, etc.). Developing applications in image processing methods and developing platforms at the same time indicate that autonomous vehicles will be very common in our lives in the near future. Especially thanks to real-time image processing, it allows to make instant and fast decisions. In addition, with the developing LIDAR and RADAR technologies, approaching or moving objects can be detected. Thus, the danger is detected in advance and necessary precautions are taken. It is predicted that autonomous vehicles will have an important place in our lives in the near future due to the advancement of technology and the comfort and safety concerns of people.

History of Autonomous Vehicles

Since at least the 1920s, automatic driving systems (ADS) have been the subject of tests, with actual implementation starting in the 1950s. Japan's Tsukuba Mechanical Engineering Laboratory created the first fully automated automobile in 1977. The car used an analog computer for signal processing and two cameras to understand the white street signs. The vehicle, which was supported by an elevated track, could travel up to 30 km/h (19 mph) (McAleer, 2017). The Navlab and ALV projects at Carnegie Mellon University, funded by DARPA, began in 1984, and the EUREKA Prometheus Project at Mercedes-Benz and the Bundeswehr University Munich, funded by EUREKA, began in 1987. These projects marked the beginning of the emergence of autonomous prototype cars in the 1980s. The ALV demonstrated off-road driving in both daylight and darkness in 1987, as well as two-lane highways that were 31 kilometers (19 mph) long and obstacle avoidance in 1986. Since the 1960s and up to the second DARPA Grand Challenge in 2005, automated vehicles have been on American roads. All three of DARPA's funding partners the US Army, US Navy, and US Department of Defense have helped advance research on controls, speed-boosting sensors, and performance in increasingly demanding, quick-changing environments. Research organizations and system firms have produced prototype. In 1991, the United States allotted $650 million for research on automating by combining embedded automation and automation, as well as research on the National Highway System, which uses automobiles on highways to show how automation technology works with highway infrastructure and highway driving. The initiative concluded in a successful demonstration in 1997, but it lacked the clear guidance and funding necessary to put the system into practice on a bigger scale. The Carnegie Mellon University Navlab traveled 4,584 km (2,848 mi), 4,501 km (2,797 mi), or 98% autonomously throughout the United States in 1995. This work was partially funded by the National Automated Highway System and DARPA. During Navlab's 2015 visit to 15 states, when Delphi drove an Audi powered by Delphi and created a 5,000 km record, its accomplishments were 99% unmatched. (Hawkins, 2017). The newest A8 will be automated utilizing "Audi AI" at speeds up to 60 kph, according to a 2017 statement from Audi (37 mph). Instead of regularly grasping the steering wheel, the driver wouldn't need to do safety checks. In addition to using cameras and ultrasonic sensors, the Audi A8 is thought to be the first production vehicle to reach level 3 automatic driving. Audi would also be the first automaker to deploy laser scanners. Waymo said in November 2017 that it has started testing a driver who was unreliable when not in the driver's seat, but there was still an employee in the car. Waymo stated in October 2018 that its test vehicles have covered 10,000,000 miles (16,000,000 km) in automated mode. (Balzer, 2002).

With the development of sensor technologies in recent years, the development and commercialization of vehicles has accelerated. Technically speaking, different autonomous features are available in vehicles produced in the market in recent years. Vehicles with many autonomous driving features such as cruise control, lane tracking system or following the vehicle in front are frequently seen in traffic. It is clearly seen that the vehicles with only one of these features or a few different ones will be seen frequently in the traffic in the coming days.

Main Sensors Used in Autonomous Driving

The effectiveness of sensor fusion and the use of cutting-edge applications directly relates to the success of autonomous driving applications. The employment of algorithms in sensor fusion will significantly lower the error rate. Internal sensors that monitor things like location, speed, acceleration, and motor torque gather data about the actual device. Cameras, range sensors (such as infrared, laser, and ultrasonic), proximity sensors (such as photo diode detectors and touch sensors), and force sensors are examples of external sensors that gather environmental data and transfer it to the decision unit (CPUC, 2020). It is a sensor that is employed. Figure 1 depicts the general architecture for autonomous driving in autonomous vehicles.

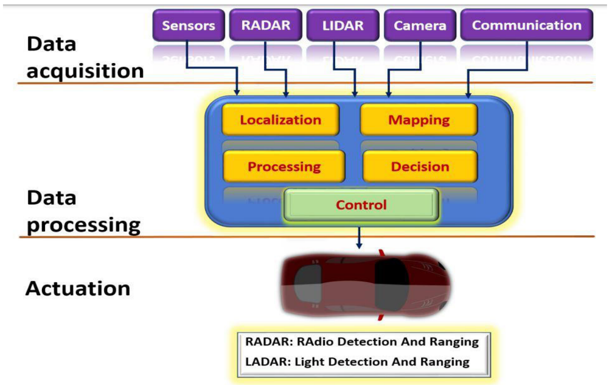


Figure 1. Basic Block Diagram for Autonomous Driving (Hussain &. Zeadally,2018)

The basic sensors required for autonomous driving, their measuring ranges and their positions on the vehicle are shown in Figure 2.

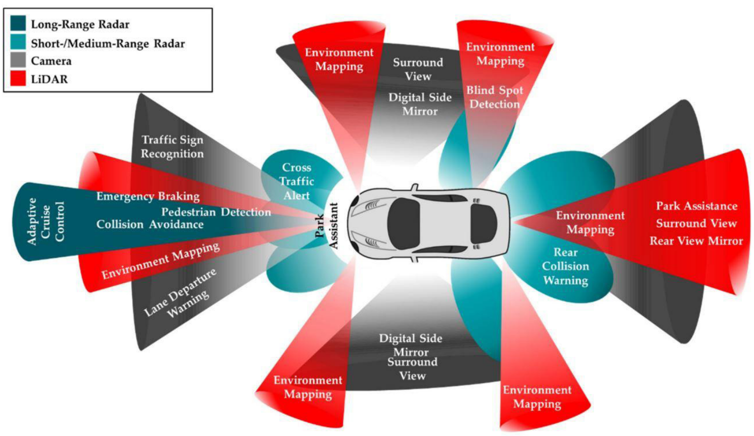


Figure 2. Basic Sensors Required for Autonomous Driving and Their Placement on the Vehicle (Hussain &. Zeadally,2018)

The following five sensors are basically used in an autonomous driving:

- Camera
- RADAR
- LIDAR
- IMU((Inertial measurement unit)
- GPS(Global Positoning System)

Camera

For manufacturers that implement many autonomous driving applications, the camera is an indispensable part. In fact, some manufacturers only use cameras in their applications. When choosing cameras, basically three important comparison features should be considered. (Trommer,2017) The first property, resolution, is the value that specifies the number of dots or pixels in an image. Resolution is the most basic way of determining the quality of the image. (Dawson,2017) The higher the resolution, the higher the image quality will be. Figure 3 shows a high resolution camera.

Figure 3. High Resulotion Camera

Field of view is what the eyes see both directly in front of and around them when viewed from a fixed point. This feature can be changed with lens selection and zoom. Another important metric is dynamic range, expressed as the difference in light between the brightest and darkest areas in an image. The wide dynamic range makes it easier to perceive the image for autonomous vehicles, especially in different lighting conditions encountered during night driving. There is an important ratio between the choice of cameras and lenses between the field of view and the choice of resolution. When it is desired to increase the wide field of view, it is necessary to increase the resolution in order to perceive the same quality due to the fewer pixels that absorb light from a certain object (Grush, 2016). The above factors should be taken into account when choosing the camera, which is the most important part of autonomous driving.

RADAR (RAdio Detection and Ranging)

With the development of technology, RADAR (RAdio Detection And Ranging) is used in many areas besides military fields today. (Ünler & Seyfi, 2022). Radars are generally used for distance measurements for autonomous vehicles. RADAR uses radio waves to measure distance. Radio waves travel at the speed of light and have the lowest frequency of the electromagnetic spectrum. RADAR covers long distance measurements and can detect large objects in the environment well. Since it uses magnetic waves, it is not affected much by adverse weather such as precipitation and snow. RADAR is selected based on key features such as detection range, field of view, position and velocity measurement accuracy. (Lawson,2018). Figure 4. show a 24 Ghz 100m range RADAR.

Figure 4. 24 Ghz RADAR

Unlike ultrasonic sensors, since they have the ability to measure at much longer distances, they can also ensure that the danger is detected from a distance. It has also made good use of RADAR systems and vehicle-to-vehicle communication in the autonomous vehicle industry, leading to advances such as adaptive cruise control and emergency braking.(Trommer,2017)

LIDAR, which has the same measurements as radar, can make erroneous measurements, especially in bad weather conditions. Effects such as attenuation and reflection in camera or LIDAR data need to be considered. These make detection methods erroneous in estimating object locations (Montemerlo, 2008). Considering the characteristics of RADAR and LIDAR sensors, RADARs are of great importance for autonomous vehicles in order to reduce erroneous measurements in bad weather conditions. In addition, RADARs provide great advantages when long distance measurements are needed.

LIDAR (Laser Imaging Detection and Ranging)

LIDAR (Laser Imaging Detection and Ranging) sensors are devices for measuring distance by emitting Laser light. In terms of distance measurement, they work with the same logic as RADARs that emit EM waves. Instead of the radio waves used in radar sensors, LIDAR makes distance measurements using laser pulses. Data collected from sensors is used to create highly detailed 2D and 3D maps of the environments around us. LIDAR sensors send between 50,000 and 200,000 pulses per second to cover an area and compile the returning signals into a 3D point cloud. The more points collected per second, the more detailed the 3D point cloud can be. The rapid update of 3D point clouds is directly proportional to the rotation speed of the LIDAR sensor. The power of the laser light used in the LIDAR sensor determines the measurement range (Lawso,2018). Figure 5 shows a HDL-64E model LIDAR sensor produced by google and used on autonomous vehicles.

Figure 5. HDL-64E LIDAR sensor

While the traditional mapping method is used in fields such as surveying, agriculture and mining, today, with the development of technology, LIDAR technology is preferred in many areas in order to make more detailed mapping. With the help of complex algorithms, LIDAR sensors are used to solve the Simultaneous Localization and Mapping (SLAM) problem for use in autonomous navigation in military applications, autonomous vehicles and urban search and rescue missions. Various SLAM algorithms are available for air, land and sea platforms. SLAM is preferred because it has many uses and can make much more precise measurements than GPS (Topal & Yigit,2021).

IMU (Internal Measurement Unit)

The IMU consists of two types of sensors to measure the 3-axis acceleration and 3-axis rotational force on an object in motion in space. The IMU measures the acceleration, angular acceleration and rotations of the system and converts these data into instantaneous position data of the system. Figure 6 shows an IMU sensor manufactured by Bosch and used in the automotive industry.

Figure 6. AN IMU sensor produced by BOSCH

The IMU, called the inertial measurement unit, has processors that interpret the information received by the sensors from the environment like the human brain and make decisions. It is related to the inertia measurement unit INS, GPS and LIDAR, which provides safe driving by monitoring the dynamic changes in the movement of the vehicle (Gao et al., 2015)

The two-sensor models produced with previous technology basically contain two separate sensors as an accelerometer and a gyroscope placed on the x-y-z axes. With the developing technology, the magnetometer in the three-sensor model has been added to the IMUs as the third sensor. The magnetometer measures the magnetic direction of the spindle bearing, so the reading of the gyroscope is improved relative to the 2 axes. (Hazry et al. ,2013)

GPS (Global Positioning system)

They are positioning systems with the help of certain mathematical calculations using GNSS and GPS satellite systems. Position estimation is made when the receiver sees at least 4 satellites. As the number of satellites seen by the receiver increases, the position sensitivity also increases. Figure 7 shows a standard GPS receiver.

Figure 7. A GPS receiver

Satellite-based positioning and navigation systems are preferred in real-time applications. The US navy TRANSIT navigation satellite system is known as the first system that uses the Doppler Effect in positioning and navigation applications of moving objects. Today, GPS is the most widely used satellite-based positioning system. Interruption of satellite signals in urban, rural and indoor areas creates problems in GPS applications. In order to prevent such problems, within the scope of GNSS-based positioning technologies, GPS is supported with additional equipment called inertial measurement equipment (IMU). (Chaiang, 2004; Ozcelik, 2009). Inertial navigation systems are systems used in addition to GPS equipment in positioning, target tracking and navigation applications of air, land, sea and space vehicles. The system's inerial sensors (gyroscope and accelerometers) measure the acceleration, angular acceleration and rotations of the system and convert these data into instantaneous position data of the system (Ozcelik, 2009; Walcho, 2002; Bayraktar, 2004).

Accumulation of errors increases with time in INS measurements. Errors in GPS measurements are independent of time. Integration of INS, which gives effective results in short time intervals, and GPS data, which gives effective results in long time intervals, is used more (Walcho, 2002; Bayraktar, 2004). Therefore, since GPS/INS integration provides reliable and more accurate location and navigation information, it is preferred in applications where small-scale precision is required (Bayraktar, 2004).

The GPS satellite orbits are about 20200 kilometers away from the planet. Two frequency values are present in GPS satellite signals: and. To mimic the delays in the atmosphere, distinct L1 and L2 wavelengths are utilized. These signals are modulated using one or both of the C/A code and P code PRN (Pseudo Random Noise) codes. L2 carrier phase is exclusively modulated with P code, whereas L1 carrier phase is modulated with C/A and P code. Every millisecond, the C/A code is broadcast at a frequency that is one-tenth that of the GPS signal (10.23MHz). Every 267 days, the P code is delivered at its original frequency. GPS receivers measure the apparent transit time of the satellite signal between the satellite and the user, called pseudorange. Pseudorange includes receiver clock deviations and signal delays (Chiang, 2004; Ozcelik, 2009).

GPS is not a self-directed or automated system. To perform position determination, the sensors must see enough satellites. Satellite visibility may decrease due to structures such as buildings, bridges, tunnels. In kinematic applications, electronic interference or other interference may cause errors/deficiencies in the detector's detection of carrier phase waves. The data frequency of many sensors is usually 1 Hz. is Most of the errors encountered in positioning with GPS are atmospheric effects that are difficult to predict (Kocaman, 2003).

Classification of Driving Automation Systems

Driving automation systems are classified in 6 different levels according to autonomous driving situations. (SAE, 2016). These levels are determined according to the rate of the driver's intervention at the time of travel. These levels are shown in Figure 8 clearly.

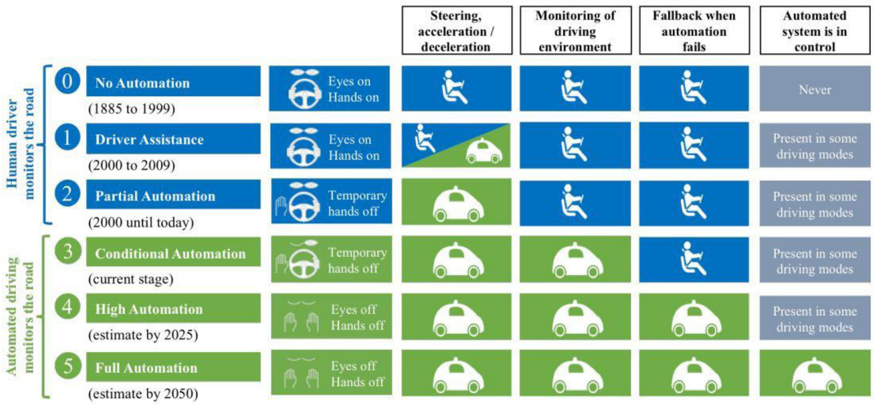


Figure 8. 6 Levels of Autonomous Driving (Cavazza et al.,2019)

Level 0 (No Automation)

Level zero vehicles are those without any automation and where the dynamic driving work is totally carried out by the user. These vehicles include technological equipment that can only function as a consequence of human interaction, which we refer to as manual usage at this level. In these vehicles, the human driver controls the vehicle's horizontal and vertical motion under all traffic circumstances. Any normal car without automation may be used as an example of this level, and cars at this level don't need to be taken into account for the study's problem. (SAE, 2016).

b. Level 1 (Driver's Assistant)

This level of automation requires the user to continuously watch their driving while a few controls are only partially taken over by the system. One, sometimes known as driver helper, is this level. It is the driving level where an automation system at this level cannot simultaneously take over the vehicle's vertical and horizontal controls, just the steering or simply the speed (Cavazza et al.,2019). The driver unquestionably has either horizontal or vertical supremacy, it should be stressed. This category is shown by cars with features like adaptive cruise control or parking assistance.

c. Level 2 (Partial Automation)

Contrary to the driver assistant, the level of automation in which both vertical and horizontal control of the dynamic driving task continues is described as level two, or partial automation. The only difference between partial automation and driver assistance is that the automation system has the technology to handle both horizontal and vertical control of the dynamic driving task (Cavazza et al., 2019). For example, it is possible to consider vehicles with both adaptive cruise control system and lane alignment system in this category.

d. Level 3 (Conditional Automation)

The first and third level automation vehicles are the vehicles in which no human driver interferes with the driving activity and the driving is controlled by the system. At this level of driving automation, the entire driving task is performed by the system within a limited driving area, and under normal conditions, the user does not need to participate in the driving task. However, in case of any malfunction in the automation system, when the system calls the user to intervene, it is the driving level that the user must respond to this call (Cavazza et al.,2019).“Traffic jam assist systems”, which can perform the entire driving task in a stop-and-go traffic situation at low speed, are shown as an example of third-level driving automation.

e. Level 4 (High Automation)

Driving automation is defined as level four, which can move within a limited driving space as in level three automation systems, but unlike conditional automation, where the user does not have to respond to the intervention call sent by the system and the entire dynamic driving task can be carried out by the system without driver intervention (SAE, 2016). For this reason, it is possible to say that in a vehicle with level four automation, the operational design area can be realized by the driving automation system with the expectation that the user will take over the rest of the control, and the user is only in the vehicle as a passenger, as long as the user stays within the limits of the control task (Cavazza et al.,2019). Today, highly automated level four vehicles are still at the prototype level, but autonomous taxis and buses, which serve in a limited area and do not have human drivers, can be given as examples of these vehicles.

f. Level 5 (Full Automation)

Level five, also known as full automation, is the driving automation where it is possible to control the dynamic driving task by the system in an unlimited operational design area, in all conditions suitable for movement, without the need for any user intervention (SAE, 2016). In the event of a malfunction or unexpected road and weather conditions during the dynamic driving task in a fully automated vehicle, the system can ex officio bring the vehicle to a minimum risk condition without the need for user intervention (Cavazza et al., 2019). The biggest difference that distinguishes full automation from high automation is that the operational design space is unlimited in vehicles with this automation system, which means that there are no restrictions on weather, time or geographical conditions as to where and when the system can operate the vehicle (

Cavazza et al.,2019). Today, there is no tool that has this level of automation. In the doctrine, it is stated that the main condition for the transition to fully autonomous driving is that the vehicles are programmed completely safely, taking into account all real-life scenarios, and that there are no possible scenarios for which the system does not know the solution (SAE, 2016) . In this respect, it should be stated that it is a difficult possibility to see fully autonomous vehicles on highways in the near future and that these vehicles do not constitute a current need in terms of criminal law for now.

The Future of Autonomous Vehicles

The increase in accidents due to human errors in traffic and the situation in environmental factors is seen as a need to develop autonomous vehicles. In addition, the increase in comfort expectations increases the demand for these vehicles and the studies on these vehicles. Considering the defined levels of autonomous vehicles, it is clearly seen that they are in our lives. Especially in level 1, level 2 and level 3 traffic, it is one of the most common autonomous vehicle classes. However, although there are companies such as Tesla, which produces at level 4 and level 5, where the driver is less involved, it is not used very intensively yet. Intensive laboratory studies are continuing, especially for level 5. In the future, it is predicted that autonomous vehicles will be seen more frequently in traffic, especially with the development of sensor technology. It is thought that developments in cameras will accelerate this process even more. In order to perform real-time image processing, which is of great importance for autonomous driving, computers must have high-speed processing capability. With the developing technology, both the reduction of the size of the computers that process the image data quickly and the increase in the processing levels will accelerate this process even more. Since this situation is very necessary especially in object recognition, developments in these areas allow autonomous vehicles to develop very quickly. On the other hand, increasing the number of sensors will increase the detection accuracy of the objects around while driving. In addition, the use of sensor fusion in the interpretation of the data of the sensors used will increase the quality of autonomous driving. In addition, with the development of artificial intelligence algorithms used in these situations, the system's self-decision mechanism will improve significantly. Thus, it is thought that the obstacles in front of level 4 and level 5 will be removed.

References

Balzer, D. (2002). Safety enhancement by lane observation of Advanced Driver Assistance Systems (ADAS), VDA, Technical Congress 2002, Germany.

Bayraktar L.H., Kundak N., (2004). "Düşük Maliyetli Alıcılarla, İnsansız Hava Araçları için Geliştirilen Ataletsel Navigasyon ve GPS Entegrasyonu Uygulaması" HITEK 2004 Havacılıkta İleri Teknolojiler Conference Proceedding Book Volume¬2, Air Force Academy, 9 December 2004, Yeşilyurt, Istanbul

Cavazza, B. H., Gandia, R. M., Antonialli, F., Zambalde, A. L., Nicolaï, I., Sugano, J. Y., & Neto, A. D. M. (2019). Management and business of autonomous vehicles: a systematic integrative bibliographic review. International Journal of Automotive Technology and Management, 19(1-2), 31-54.

Chiang, K.W., (2004). "INS/GPS Integration Using Neural Networks for Land Vehicular Navigation Application", Ph.D. Thesis, University of Calgary, Alberta, Canada.

CPUC (2020). Autonomous Vehicle Pilot Permits Issued, California Public Utilities Commission (www.cpuc.ca.gov/avcissued).

Dawson C. (2017). WSJ ,Why Your Next Car May Look Like a Living Room, Wall Street Journal, www.wsj.com, https://www.wsj.com/articles/your-next-car-may-be-a-living-room-on-wheels-1497881613 Accessed Date:20.10.2022

Dickey, S. (2008). Development of hardware in the loop simulation and paramics/VS-PLUS integration. Technical report UCB-ITS-PRR-2008- 29, California PATH research report

Gao, Y., Liu, S., Atia, M. M., & Noureldin, A. (2015). INS/GPS/LiDAR integrated navigation system for urban and indoor environments using hybrid scan matching algorithm. Sensors (Switzerland), 15(9), 23286–23302. https://doi.org/10.3390/s150923286

Grush, B. (2016), Driverless Cars Ahead: Ontario Must Prepare for Vehicle Automation. Residential and Civil Construction Alliance of Ontario (RCCAO).

Hawkins, Andrew J. "The Federal Government Just Granted Its First Driverless CarExemption", The Verge, Accessed Date: 16.10.2022 <https://www.theverge.com/2020/2/6/21125358/nuro-self-driving-deliveryrobot-r2-fmvss-exemption> (Date of Access: 20.10.2022).

Hazry, D., Sofian, M., and Azfar, A. Z. (2009). "Study Of Inertial Measurement Unit Sensor," No. October, Pp. 11–13, [11] Ahmad, N., Ghazilla, R. A. R., Khairi, N. M., & Kasi, V. (2013). Reviews on various inertial measurement unit (IMU) sensor applications. International Journal of Signal Processing Systems, 1(2), 256-262.

Kocaman, S., (2003). "GPS and INS Integration with Kalman Filtering for Direct Georeferencing of Airborne Imagery", Geodetic Seminar Report, ETH Hönggerberg, Zurich.

Lawson, S., (2018). Tackling the Transition to Automated Vehicles, Roads that Cars Can Read Report III, European Road Assessment Association.

Montemerlo, M. (2008). Junior: the Stanford entry in the urban challenge. J. Field Robot. 25,569–597.

Ozcelik A. E., (2009). "GPS/INS data fusion with kalman filtering methods", Msc.

Thesis, Erciyes University, Kayseri)

R. Hussain and S. Zeadally, (2019). "Autonomous Cars: Research Results, Issues, and Future Challenges," in IEEE Communications Surveys & Tutorials, vol. 21, no. 2, pp. 1275-1313, doi: 10.1109/COMST.2018.2869360.

SAE International. (2016). Surface vehicle recommended practice: (R) Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. USA.

Tastan Y. & Kaymaz, H. (2021). Difficulties in the Proliferation of Autonomous Vehicles. International Journal of Advances in Engineering and Pure Sciences, 33 (2) , 195-209 . DOI: 10.7240/jeps.741594

Topal A., &Yigit T., (2021). "A 3d Environment Mapping System for Unmanned Ground Vehicles Usıng Lidar Technology"International Journal of 3D Printing Technologies and Digital Industry 5.2: 171-186.

Trommer S. (2016)., Autonomous Driving: The Impact of Vehicle Automation on Mobility Behaviour, Institute of Transport Research www.ifmo.de.

Unler, T., & Seyfi, L. (2022). Experimental Investigation of Range Measurement Error for Radar under Different Weather Conditions. In Proceedings of the Bulgarian Academy of Sciences (Vol. 75, No. 9, pp. 1325-1333).

Vanderblit, T. (2012). Autonomous Cars through the Ages.

Walchko, K.J.,(2002). "Low Cost Inertial Navigation: Learning to Integrate Noise and Find Your Way", Msc. Thesis, University of Florida, Florida.

About the Authors

Tarik UNLER, PhD, is an Assistant Professor of Aviation Electrical and Electronic at Necmettin Erbakan University in Konya, Türkiye. He holds a PhD in Electric-Electronic Engineering from Konya Technical University. His main areas of interest are embedded systems, autonomous systems, RADAR and aviation applications.

E-mail: tunler@erbakan.edu.tr, **ORCID:** 0000-0002-2658-1902.

Similarity Index

The similarity index obtained from the plagiarism software for this book chapter is 20%

To Cite This Chapter:

Unler T. , (2022). Vehicles of the Near Future: Driverless Autonomous Car, In S. Kocer, O. Dundar (Eds .). *Current Studies in Basic Sciences Engineering and Technology* (pp. 70–81). ISRES Publishing.

Engineering Applications of Nonlinear Optics

Ferit ARTKIN
Kocaeli University

Introduction

In the early 1980s nonlinear optical materials was extended with a new production of semiconductor materials. For example, multi-quantum well is produced by the amplification of two different semiconductor layers (GaAs and AlGaAs). That let the formation of consecutive hills and valleys of variable potential, for electrons and holes. It is regarded that the carriers are within these potential walls. Hence, they display a semi two-dimensional behavior and large optical nonlinearity.

It can be an useful instrument for network communication for Photonic information and communication systems. It has developed very fast in last years. However, using only fiber optical systems has not ability to meet the necessity on carrying the information flow in future of communication systems. Fast optical switches which can verify (process) large scale of data should be developed so as to remove the gateways caused by optoelectronic return throughout the network. Large and fast nonlinear optical (NLO) materials are appropriate tool for photo-dynamic therapy to treat serious diseases like cancer, and other biomedical applications in the near future soon (L. W. Tutt et al., 1993) for optic controller, optic correction and optic power limiting practices (T. D. Mody et al., 2001) (Wung, C. J. et al., 1993).

Some of the advances in science and technology is crucial for the development of third-order non-linear optical materials are summarized below:

Theoretical Modeling: To increase the magnitude, especially for third-order nonlinearities, predictive capabilities need to be improved. Progress at this point might enable the prediction of polymer structural and other functionalization requirements that may be necessary.

Molecular Engineering through Polymer Design and Chemical Synthesis: Progress in systhetic and polymer chemistry and chemical engineering is significant for sustainable progress in producing systematically varied structures where the impact of important electronic or structural features can be studied. The rigorous processing requirements should be on center of synthetic strategies associated with waveguide fabrication and performance issues. Examples of such issues are control of domain structures that can lead to refractive index in homogeneities, environmental stability, photo stability, mechanical compatibility with substrates, and adhesion (Artkin F., 2014).

Material Processing: Processing conditions are vital in figuring the ultimate quality of fibers and films of organic and polymeric materials. One promising avenue for optimization of electronic properties, meanwhile the variety of needed ancillary properties, is done by of polymer blends and composites. To show a sample of this approach is a current improvement of a composite between sol-gel processed silica and polyparaphenylene vinylene (Sutherland, R. L. 2003). This silica gel is an excellent linear optical medium with extremely low optical losses, yet it has rather low third-order nonlinear coefficient. Special processing techniques has been improved to eradicate general tendency for phase separation which occurs at higher concentrations of polymer. The resulting films have excellent optical quality and a nonlinear coefficient is consistent with the polymer intensity in the glassy medium. Device Structures: There are considerable opportunities in the process of development and production of device structures. A number of preliminary device structures have been found out, as properties of materials are improved and their intrinsic advantages and limitations realized, refinements in devices which can be benignant in terms of their advantages will undoubtedly be pursued (B. G. Maiya et al., 1989).

Nowadays, characterization is particularly important for materials showing suitable properties for nonlinear optical devices. Various techniques are applicable for these materials' characterization. In this study, Z-Scan technique is used which provides advantages in many aspects.

Methods for Studying of Optical Nonlinearity of Engineering Applications

It is highly popular nowadays to find materials with large but fast non-linear properties, database enlargement becomes necessary. Methods for determining non-linear parameters are discussed throughout this book chapter, search which creates curiosity, for materials purely optical controlling and detector protection ingredients, it's about nonlinear absorption (NLA) and nonlinear refraction (NLR). Database for optic features matters, especially natural, some situations are not sufficient for metric determination to guide synthesis efforts.

Degenerate Four Wave Mixing (DFWM): Operated for measuring third order non-linear reaction and magnitude this reaction.

Third Harmonic Generation: Used only for measuring third order nonlinearity.

Electro-absorption technique: Third order nonlinearity dispersion, this technique is applied.

Time-Resolved Optical Kerr Effect and Transient Absorption Techniques: Used presenting photophysical processes that determine nonlinearity.

Z-Scan: Used for measuring magnitude of third order nonlinearity and its sign.

In this regard, the reader sould be aware that the results for using the z-scan method, active non-linear processes of any reference are frequently a matter of debate. The Z-scan technique is a type of method that quickly measures both NLA and NLR in solids, liquids and liquid solutions. The non-linear optics community has swiftly adopted the Z-scan approach as a standard method for identifying Non-linear variations inindex and absorption must be considered independently. It should be emphasized, however, that the approach is susceptible to any non-linear optical phenomena that cause variations in the refractive index and absorption coefficient, making it difficult, in most cases, to infer the underlying physical mechanisms from a Z-scan (Paras N. Prasad et al., 1991).

Degenerate Four Wave Mixing (DFWM)

A physical description degenerate four-wave mixture, principle, treated similarly to harmonic generation (or frequency mixing), except for three spatially distinguishable field waves in this process, except for E1(ω) ,t), E2(ω, t), E3(ω,t) interacts with the same frequency (ω), creating the fourth wave E4(ω,t) in the same order, but with a different propagation direction.

Figure 1. can be solved under the assumption of slowly varying amplitude approximation, which assumes that it wave density variation throughout the propagation distance negligible on scale of wavelength of light, and negligible pump beam depletion to obtain the expression for field E_4 as a function of propagation distance z in non-linear media.

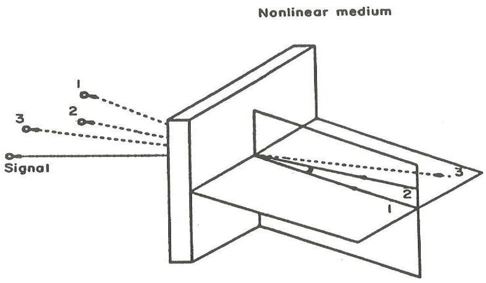


Figure 1. Four beams are available in scheme in the forward-wave arrangements for degenerate four-wave mixing.

As in the example of harmonic generation, the wave equation also yields the requirement to match phases $k_1 + k_2 + k_3 + k_4 = 0$.

Many different beam geometries have been used for degenerate four-wave mixing (DFWM). The two common geometries are shown in Figures 1. and 2. The Figure 1 shows the forward wave geometry arrangement with all incident beams (E_1, E_2, and E_3) are propagating from right to left by following the same way. It also shows the phase-matched direction of the output beam (signal) E_4. Figure 2. shows a backward-wave geometry in which two beams E_1 and E_2, called the forward and reverse beams E_f and E_b, are counterpropagating and third beam E_3, also called the probe beam E_p. The process of DFWM synonymous with optical phase conjugation in backwave geometry (Fisher, R.A., 1983). From this perspective, it is considered the interaction of two counterpropagating pump beams I_f and I_b, which set up a phase conjugate mirror. A probe beam I_p incident on it is reflected to retrace itself. However, it can swap energy with the other beams. I_s is used to refer the conjugate of the probe beam I_p. Therefore, any phase distortion suffered by the probe beam I_p is reconstructed in beam I_s. The process is utilizable for high-contrast dynamic holography.

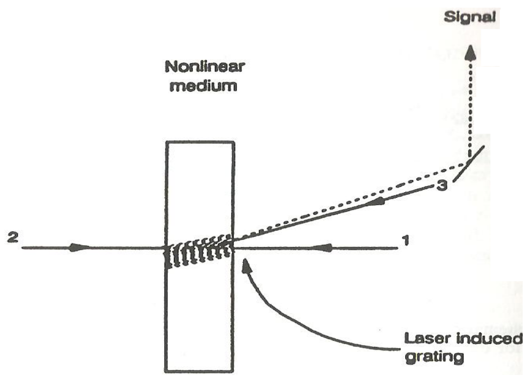


Figure 2. Schematics of the four rays in the reverse-wave arrangement degenerate four-wave mixing.

The DFWM process can also be intuitively comprehended by imagining a laser-induced grating (Eichler, H. et al., 1986). At this level, the interaction of two harmonic beams in a material causes intensity modulation. Because the refractive index of nonlinear media is intensity dependent, the refractive index modulates, resulting in a diffraction grating. If the refractive index of the medium is complicated, the refractive index's real component changing opens the way for what is known as a phase grating (Eichler, H. et al., 1986) (Fayer, M. D. et al., 1982). An amplitude grating is created when the refractive index's imagined portion modulates. In the case of reverse DFWM geometry, each rays pairings generate three separate grids, but only two of these grids yield phase-matched conjugate signals (Shen Y.R. et al., 1984).

Third Harmonic Generation (THG)

However, the it component is included in the wave equation for the propagation of a plane wave electric field in a nonlinear medium. Under low signal or low conversion conditions, the solution of the consequent connected amplitude equations. This circumstance occurs when the phase promptness of the primary and third harmonic waves is equal. However, dispersion refractive indices (which are often different) results in varying phase promptness for the main and third harmonic waves.

Therefore, for finite values of phase mismatch, decreases from its value at and undergoes damped oscillations.

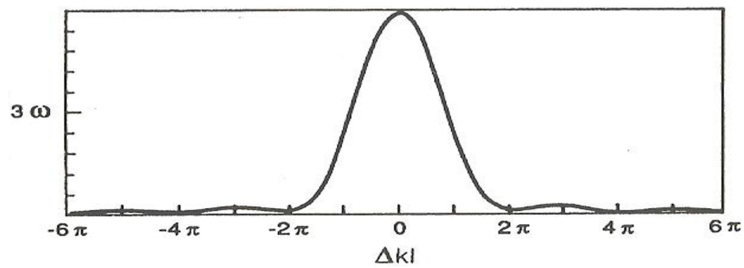


Figure 3. Third-harmonic intensity in random parts, marked as a function of *l*, for a finite wave-vector mismatch

Natural birefringence to phase match may be favorable in anisotropic media, such as birefringent crystals, since it is useful for efficient second-harmonic production. For instance , it is possible to choose a direction of propagation for which

$$n_e(3\omega) = n_0(\omega) \qquad (1)$$

The letters e and o stand for extraordinary and ordinary rays, respectively. However, due to the significantly divergent frequency values of fundamental and third harmonic waves and inadequate birefringence, phase matching for third harmonic production in a material such as crystal is difficult to achieve in practice.

Due to dispersion effects on the refractive index, phase matching cannot be met in pure isotropic media such as gases, liquids, or amorphous polymers. However, phase matching may be achieved in theory by mixing two isotropic media A and B (acting same two gases) (Shen, Y.R. et al., 1984).

$$n_A(\omega) + n_B(\omega) = n_A(3\omega) + n_B(3\omega) \qquad (2)$$

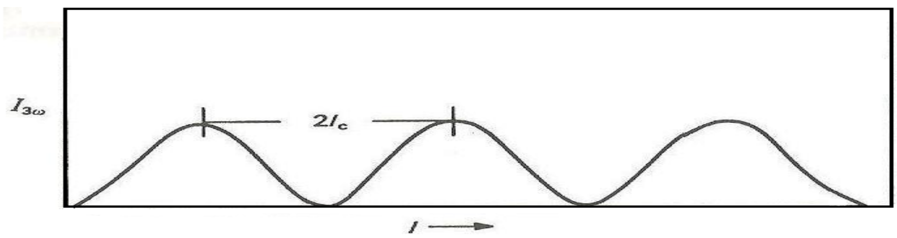


Figure 4. Wedge fringe pattern for third-harmonic signal obtained as a sample is translated to change the path length *l*.

When the interaction length is changed, the oscillatory behavior of is described by a fringing curve. In the case of a wedge-shaped sample, the change of interaction length by translation of the sample produces a wedge fringe pattern, demonstrated in Figure 4.

Harmonic generation (or frequency mixing in general) is a coherent process by taking account the perspective out wave generated at the new frequency has a definite relationship with the phase of the input waves. It is derived from non-linear interactions that appear instantaneously. Such consistent non-linear states are only obtained from electronic interactions that have no dependence on the excited state population. We shall see later that dynamic nonlinearities that involve electronic or molecular excitations are incoherent non-linear responses of a medium and depend on the population density of excitation in the medium.

Electric Field-Induced Second-Harmonic Generation (EFISH)

It has contributions both from second- and third-order nonlinearities. However, for centrosymmetric structures or structures with small second-order nonlinearity, the EFISH generation is derived from the third-order susceptibility, γ. The EFISH process can be depicted as unusual example of four-wave mixing where one area is in bottom in terms of frequency or dc-electric area for which crosschecked with optical field's sequence.Thus it can be regarded , a particular situation of sum-frequency generation.

For rigid solid structures where alignment of molecular bipolarity corresponding with the electric field is energetically improbable, the EFISH process is derived from γ even though the molecular structure might permit a large. This is the case with isotropic polymers below their glass transition temperatures. As for third-harmonic generation, the EFISH process probes only purely electronic third-order nonlinearity, which has an instantaneous response (Paras N. Prasad et al., 1991).

Z-Scan Technique

A z-scan measurement is operated in nonlinear optics to estimate the nonlinear index n2 Kerr nonlinearity and nonlinear absorption coefficient using open and closed approaches, respectively. Because non-linear absorption might alter the measurement of the non-linear index, the explicit approach frequently operated to rectify computed value of the implicit technique.

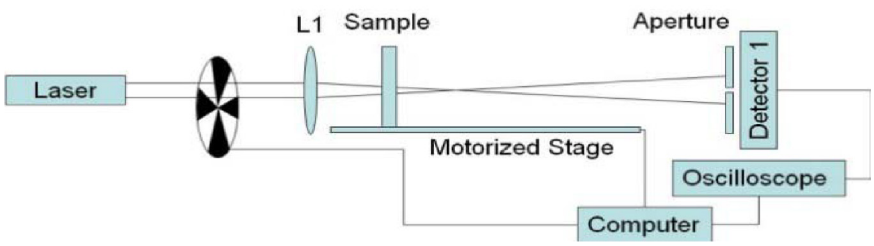


Figure 5. Experimental Z-scan setup.

An aperture is constructed in this experimental configuration to prevent part light from passing through detector is set up as shown in the diagram. A laser is focused by a lens. Some extent before the beam naturally drifts out of focus. An aperture detector is placed behind it after a considerable distance. The aperture allows just light cone's field center to reach the detector. It is frequently within the normalized permeability range

$$0,1 < S < 0,5 \quad (3)$$

The detector is sensitive to any focussing or defocusing caused by the sample. Sample is normally positioned at the focal point of lens and then passed across the z axis, the distance between which is provided by the Rayleigh length zo;

$$z_0 = \frac{\pi W_0^2}{\lambda} \qquad (4)$$

According to the thin sample approximation, sample L' density must be less than Rayleigh length L < z_0.

Z-Scan Measurements

Definition: A method for determining the strength of Kerr nonlinearity in a material that relies on self-focusing.

The z-scan measuring technique (M. Sheik-Bahae et al., 1989) (M. Sheik-Bahae et al et al., 1990) is frequently used to calculate an optical material's kerr nonlinearity force (ie the magnitude of the non-linear index n2). To demonstrate the functional location of the sample, A sample of the material under investigation is moved along the laser beam's focus, and the beam radius (or on-axis intensity) is measured at one point while holding it behind the focus zone. These figures are influenced by the self-focus effect. If the non-linear index is positive and the sample is positioned below the focus, self-focusing reduces beam deflection and hence increases detector signal (as shown in Figure 5). When the sample is moved to the left of the focal point, the focus shifts to the left and the detector signal decreases. The observed reliance of the detector signal on the sample position may be used to calculate the magnitude of the non-linear index. Non-linear absorption, such as two-photon absorption, can also have an effect on the measured signal. This, however, may be monitored individually while keeping the power of the total transmitted beam constant. Nonlinearity measurements can be adjusted using these data.

Advantages and Disadvantages of Z-Scan Technique

Advantages:

- There are no complicated situations except focusing the beam on the slut Magnitude and sign of nonlinearity can be measured simultaneously.
- Refractivity and absorptivity parts of nonlinearity can be isolated. Yet, emplying DFWM technique cannot be helpful.
- Except for a few exceptions, data analysis is quick and easy.
- Intensity is high. If sample has high optical quality, even a phase distortion of λ/300 can easily be decomposed.
- A close similarity exists between Z-scan and geometry that limits optical power.

Disadvantages:

- The requirement for a high quality Gaussian TEM00 beam for several experiments is a drawback because to the difficulty in obtaining it.
- Analysis is completely different for non-Gaussian beams.
- Beam is infractible because of irregularities of the sample.

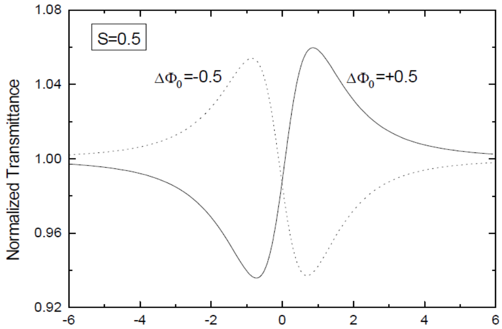


Figure 6. Typical Z-scan for positive data (solid line) and negative (dashed line) data nonlinear refraction of third order.

A typical closed aperture Z-scan result for a thin sample with non-linear refraction is shown in Figure (6). (solid line). As the example moves away from the lens, a self-focusing nonlinearity, n > 0, generates a peak followed by a trough in normalized transmittance (increasing Z).

One of the enticing aspects of the Z-scan technique is the ease and simplicity with which non-linear optical coefficients may be created with approximate accuracy.

However, as is the case with the vast majority of non-linear optical measurement techniques. The accurate estimate of nonlinear coefficients, such as n2 or, is hence dependent on competing nonlinearities. It becomes critical as reliance on the temporal and spatial profiles of the laser source, power or energy content, as well as stability Z scan may be thoroughly modeled for every beam form and sample thickness by solving the appropriate Maxwell equations under certain nonlinearity conditions (for example, a very rapid (3) response). Many appropriate assumptions and forecasts, on the other hand, will result in simple analytical expressions and make data analysis easy yet exact. Outside of the usual SVEA (slow changing envelope approximation), neither diffraction nor non-linear refraction within the non-linear sample will alter the beam profile, as long as the non-linear sample is narrow.

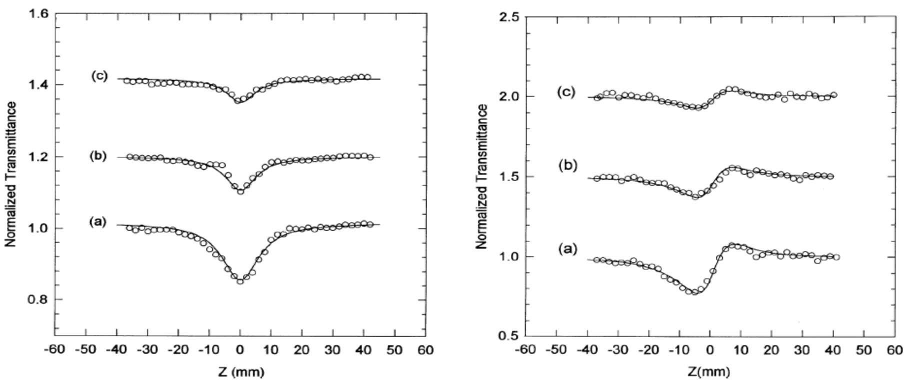


Figure 7. Z-scan measurements for various laser intensities on lithium niobate (Heping Li et al., 1997): (a) 22 GW/cm2, (b) 12 GW/cm2, (c) 6 GW/cm2 experimental data (circles) and theoretical fits (straight lines). Left: non-linear absorption through open aperture geometry; right: non-linear refraction via closed aperture geometry. For graphing, the downs for (b) and (c) have been adjusted vertically.

References

Artkin F., (2014), PhD Thesis, Computer Aided Z-Scan Testing Apparatus Integrated into Servo Motor and PLC for Investigation of Non-Linear Materials' Permeability, Technical University of Sofia.

B. G. Maiya, A. Harriman, J. L. Sessler, G. Hemmi, T. Murai, and T. E. Mallouk, J.Phys, Chem. 93, 8111 (1989).

Eichler, H., P. Gunther, and D. W. Pohl, Laser Induced Dynamic Gratings, Springer-Verlag, Berlin, (1986).

Fayer, M. D., Annu. Rev. Phys. Chem. 33, 63 (1982).

Fisher, R.A. (Ed.), Optical Phase Conjugation, Academic, New York, (1983).

Heping Li, Feng Zhou, Xuejun Zhang, Wei Ji. Picosecond Z-scan study of bound electronic Kerr effect in LiNbO3 crystal associated with two-photon absorption. Appl. Phys. B 64, 659–662 (1997).

L. W. Tutt and T. F. Boggess, Prog. Quantum Electronics 17, 299, (1993).

M. Sheik-Bahae et al., "High-sensitivity, single-beam n2 measurements", Opt. Lett. 14 (17), 955 (1989).

M. Sheik-Bahae et al., "Sensitive measurement of optical nonlinearities using a single beam", IEEE J. Quantum Electron. 26, 760, (1990).

Paras N. Prasad, David J. Williams, Introduction to Nonlinear Optical Effects in Molecules and Polymers John Willey & Sons Inc., U.S., 1991., ISBN: 978-0-471-51562-3.

Paras N. Prasad, David J. Williams, Introduction to Nonlinear Optical Effects in Molecules and Polymers John Willey & Sons Inc., U.S., 1991, Chap. 8.3., ISBN: 978-0-471-51562-3.

Shen, Y.R., The Principles of Nonlinear Optics, Wiley, New York, 1984, Chap. 2.

Sutherland, R. L. 2003, Handbook of Nonlinear Optics, Marcel Dekker Inc., 971 p., USA

T. D. Mody and J. L.Sessler, J. Porphyrins Phthalocyanines 5, 134 (2001).

Wung, C. J., Y. Pang, P. N. Prasad, F. E. Karasz, Polymer, Elsevier, (1993).

About the Author

Ferit ARTKIN, PhD, is a Lecturer of Hereke Vocational School at Kocaeli University. Author graduated from the Technical University of Sofia with a PhD in Mechanical Engineering. Optical and Laser Devices and Methods, Nonlinear Optics, Machine Design and Machine Equipment, Manufacturing Technologies, Product Design are his key areas of expertise. He is the Vice Head of the Vocational School of Hereke Asim Kocabiyik, Department of Machinery and Metal Technologies, Kocaeli University,

E-mail: artkinf@kocaeli.edu.tr, **ORCID:** 0000-0002-8543-6334

Similarity Index

This book chapter has a similarity index of 16 % according to plagiarism software.

To Cite This Chapter:

Artkin F., (2022). Engineering Applications of Nonlinear Optics Learning. In S. Kocer, O. Dundar (Eds .), *Current Studies in Basic Sciences Engineering and Technology* (pp. 82–91). ISRES Publishing.

Performance Analysis of Symmetric and Asymmetric Encryption Algorithms in Digital Images

Mehmet HACIBEYOGLU
Necmettin Erbakan University

Bashir Sheikh Abdullahi JAMA
Necmettin Erbakan University

Introduction

With the progression of innovation and the internet, the main issue about information security and protection has arisen (Furht, Socek, & Eskicioglu, 2004). Many individuals these days are worried about their data being taken while looking for data on the internet. We carefully peruse the internet for data or share significant data with associates, as though somebody was gazing at us out the window. A thought known as information encryption has been introduced as an answer for this issue. Information encryption is the demonstration of diverting information from a plain message or clear, design into a format that the right sender and planned recipient can comprehend. Encoded information ordinarily shows up as a long line of arbitrary letters and numbers. The expected beneficiary of the information has the key or an exceptional means to change over the information from encoded to the decoded structure (Davis, 1978).

Recently days, the attackers is utilized an approach known as a brute force which endeavor arbitrary keys until they obtain the fitting one which is the most basic method of attack on encryption. The length of the key impacts the number of keys that can be utilized and the persuasiveness of this attack. It's memorable's crucial that while encryption strength is straightforwardly connected with key size, the quantity of resources needed to do the computation raises as well (Bhanot & Hans, 2015). Data encryption ensures the security of advanced information as it is put away on computer network frameworks and conveyed across the internet or other network systems (Davis, 1978). Encryption should be appropriately intended to try not to dial back a framework's exhibition. Equipment, such as chips or hard drives, operating frameworks, like MS Windows or Linux, and encryption programs from NIST-ensured providers are the three levels of encryption, recorded from best to most exceedingly terrible execution. As indicated by merchant benchmarks for every one of the three degrees of encryption, frameworks will scarcely lose a small portion of a percent of their speed, and the end client won't see the distinction. To check these statements, more benchmarks might be required. This data might be accessible from staff who give specialized help to different projects at your foundation; if not, the CDC's NPCR can work with NPCR projects to build up a rundown of benchmarks (Furht et al., 2004).

Image encryption has been the subject of broad exploration of research. A picture encryption plot is introduced in which the places of picture pixels are rearranged and the dim upsides of image pixels are changed to befuddle the connection between the encrypted picture and the plain picture (Guan et al., 2005). The Arnold feline guide is utilized to rearrange the places of image pixels in the spatial space in this procedure. From that point onward, Chen's turbulent framework's discrete result signal is preprocessed to make it satisfactory for grayscale picture encryption. An astounding picture encryption algorithm the usage of a plane little bit of a source image changed into delivered as bitplane bit safety key to scramble pics to work on protection picture encryption strategies primarily based on bitplane bit decay (Zhou et al., 2014). To scramble the photographs, Ye et al. (2010) utilized the strategic guide. The stage is done at the piece level in this technique. This technique is impervious to an assortment of security attacks. Zhu et al. (2011) suggested a piece-level change and dissemination algorithm for picture encryption. Change and dissemination activities are performed utilizing the Arnold map and strategic guide in this strategy.

Kanso et al. (2012) used a three-layered Arnold change to create secret keys to encode the pictures. This approach is broken into three stages rearranging, scrambling, and veiling. This method is very touchy to the picture utilized as a piece of information. It secures against known plaintext and picked plaintext assaults all the more viably. Khan and Shah (2014) worked on the mathematical and factual elements of the S-confine by carrying out it in a turbulent way using relative and Lorenz changes. S-confine is for the most part utilized square encryption to make pixels become confounded. It works on the disarray and dispersion of the information picture to shield it from security attacks.

Material and Methods

Encryption methods are arranged by three standards. The first is the number of keys utilized by the encryption process (zero, one, and two keys), the second is the operation type of the encryption algorithm (substitution, displacements, and production), and the third is the plain text handling structure utilized by the encryption calculations (block-based and stream-based). Besides, the encryption methods are extensively ordered into fundamental classes: Symmetric and asymmetric encryption (Sharif et al., 2010).

Symmetric encryption utilizes a solitary, private key for encryption and unscrambling. It is a quicker strategy than unbalanced encryption and is best utilized by people or inside shut frameworks. Involving symmetric strategies with numerous clients in open frameworks, for example, over an organization, requires the transmission of the key and sets out freedom for robbery. The most regularly utilized kind of symmetric encryption is AES (Sharif et al., 2010).

Asymmetric encryption utilizes a couple of numerically connected public and private keys that must be utilized together. Encryption should be possible with one or the other key, yet unscrambling should be finished with the related key. Since the public key might be uninhibitedly traded without gambling information burglary, awry encryption is used by different clients and over open organizations, like the Internet. ElGamal, RSA, DSA, and PKCS are the most routinely utilized sorts of hilter kilter encryption (Lai et al.,2010).

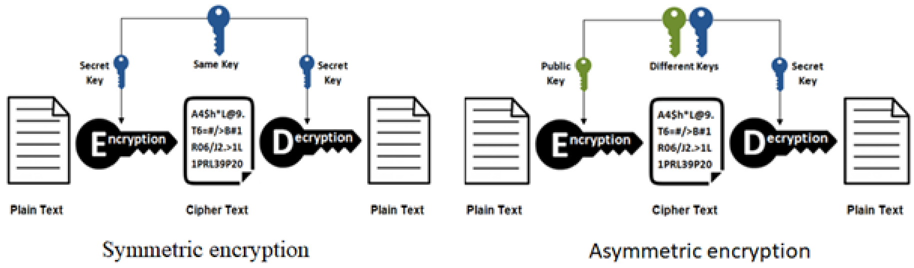


Figure 1. Symmetric encryption and asymmetric encryption processes ("Symmetric vs. Asymmetric Encryption," 2022)

Data Encryption Standard Algorithm

The Data Encryption Standard is a block-based symmetric-key methodology for the encryption of different forms of information in 64-digit blocks. This suggests that 64 pieces of plain text are taken care of into DES, which produces 64 pieces of ciphertext. Encryption and unscrambling utilize a similar calculation and key, with slight varieties. The DES algorithm has a 56-bit key length. Since the DES approach has been demonstrated to be helpless against amazingly strong assaults, its prevalence has been on disappearing (Kansal & Mittal, 2015).

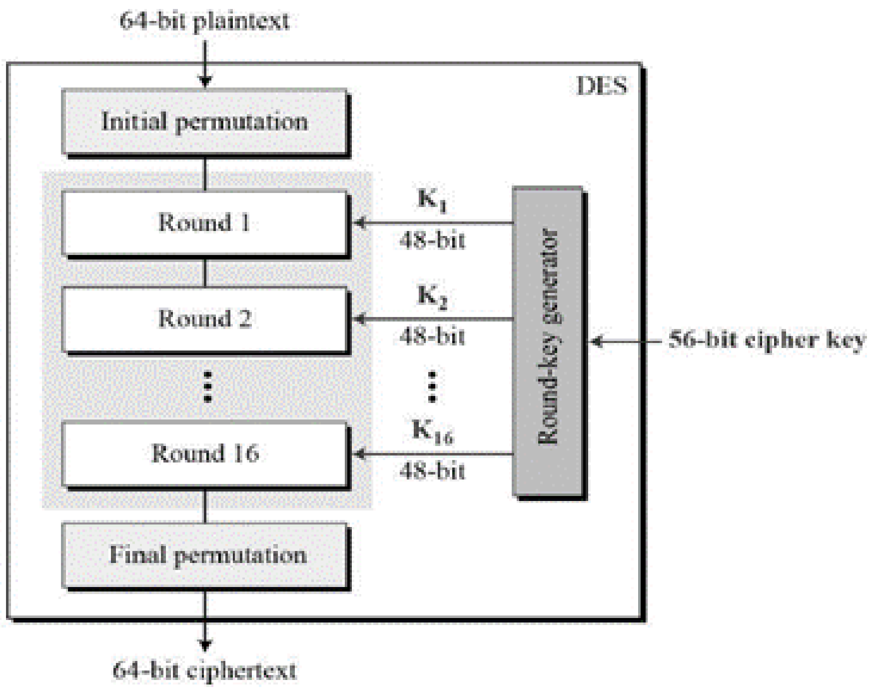


Figure 2. DES algorithm operation procedure ("Data Encryption Standard", 2022)

DES is a block-based approach depending on the Feistel Cipher, the method should determine three fundamental things needed for the encryption operation, in particular the Round work, the Key table, and the First and last permutation. The first and only changes are straight Permutation boxes (P-boxes) that are inverses of one another. They have no cryptography importance in DES (Kansal & Mittal, 2015). The underlying and last changes are displayed in fig (3).

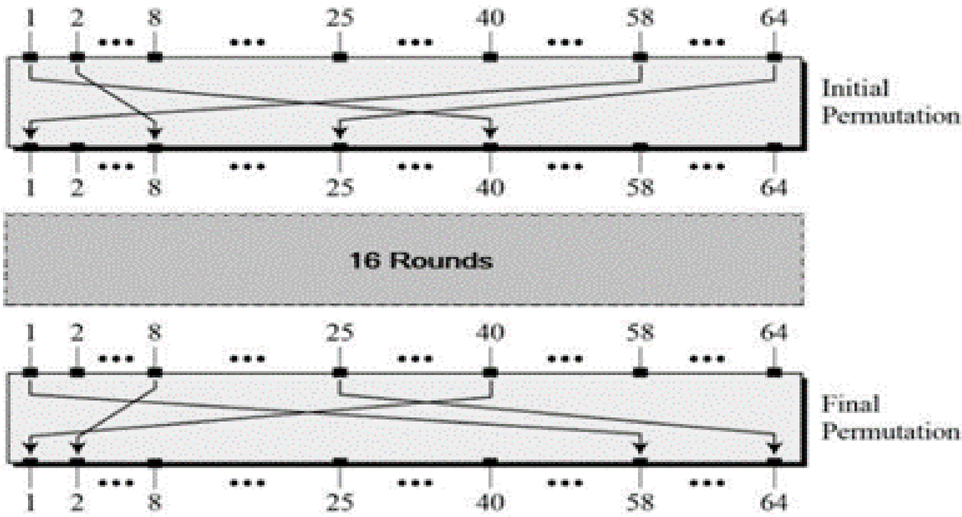


Figure 3. DES algorithm operation procedure ("Data Encryption Standard ", 2022)

After performing the first and last permutation, the round function *f* was used, which applied a 48-bit key to the rightmost 32 bits to produce a 32-bit output.

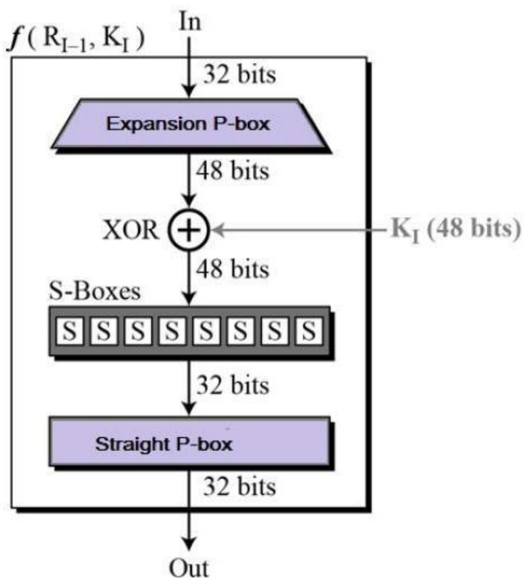


Figure 4. DES algorithm operation procedure ("Data Encryption Standard", 2022)

Since the round key is 48 parts and the correct facts (data) are 32 bits, it's miles expected to first decorate the suitable contribution to forty eight components in the DES path, that's the execution of the improvement exchange field.(Davis, 1978).

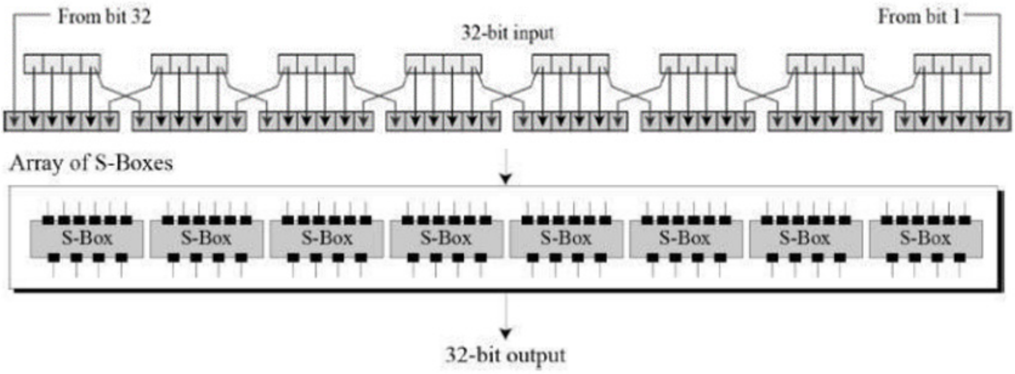


Figure 5. Demonstrates key expansion and XOR of DES algorithm ("Data Encryption Standard", 2022)

Key creation is the penultimate stage in the DES method's encryption cycle, which creates sixteen 48-piece keys from a 56-digit figure key. Figure 6 portrays the key creation process.

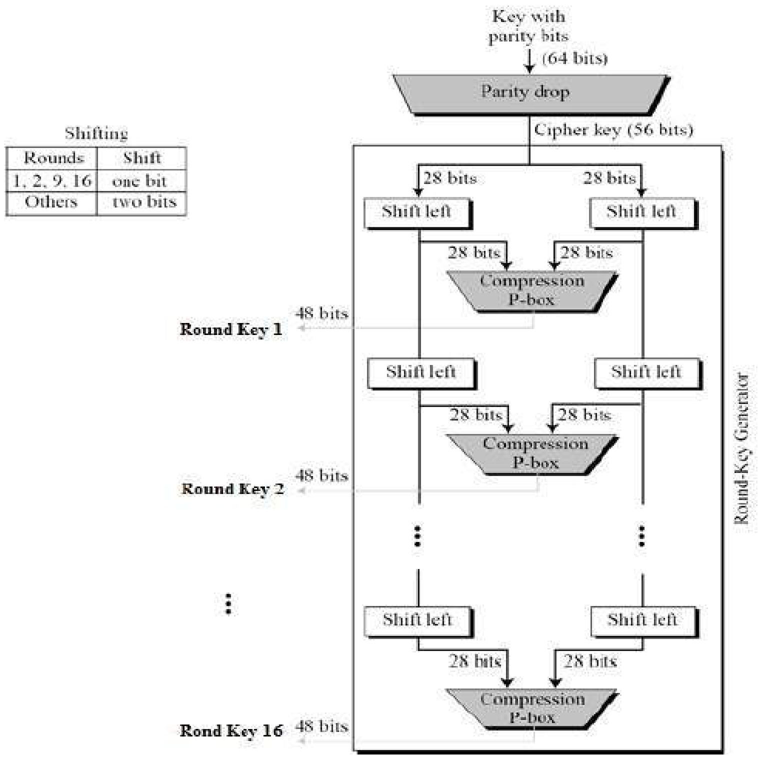


Figure 6. Key generation of DES algorithm

Triple DES

The DES cipher is repeated three times to produce the Triple DES block cipher algorithm. The Triple DES algorithm is called TDES or TDEA (Triple Data Encryption Algorithm). (Barker & Mouha, 2017). TDES was selected as a straightforward option to expand the key space without switching to a novel algorithm after it was discovered that a 56-bit key DES was insufficient to prevent brute force assaults. To keep away from meet-in-the-center attacks, which might be a hit with double DES encryption, 3 strategies are required. 3DES is a frequently utilized encryption technique in the financial industry. It's to up amazing the preceding block cipher DES. It has some of advantages, which include ease of implementation in each hardware and software program, and wide approval among cryptographic libraries as well as protocols. 3DES improves the security of DES by using three 56-bit keys DES instead of two keys. (Cardarilli, Di Nunzio, Fazzolari, & Re, 2015). The following is a description of the 3DES encryption method, where K1, K2, and K3 are the three 56-bit keys.

3DES Encryption Algorithm:

- Begin the encryption process of plaintext with the use of the K1 and the DES method.
- In the second step, decryption is executed on the encrypted output by using the DES algorithm and K2.
- The decrypted output should be encrypted another time using the DES algorithm and K3.

This kind of encryption technique produces the output of the 3DES cipher text.

3DES Decryption Algorithm:

- The cipher picture can be decrypted using the DES and K3.
- The decrypted output should be encrypted using the DES and K2.
- To covert the encrypted data that has been decrypted, use the DES and K1.

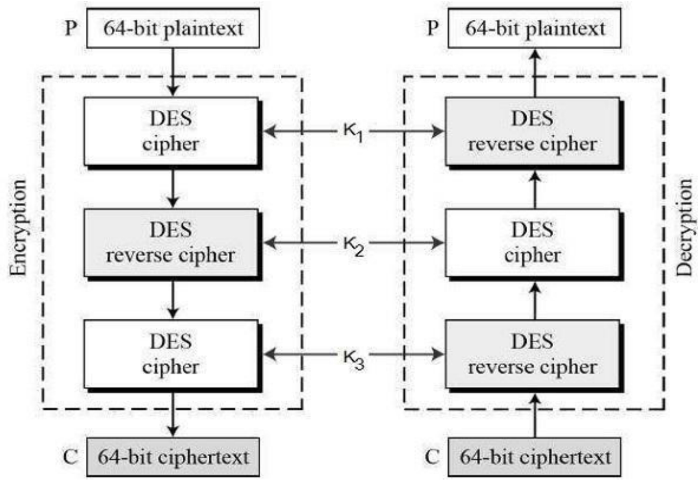


Figure 7. 3DES steps

Advanced Encryption Standard Algorithm

The most well-known and often preferred symmetric encryption (SE) method is the Advanced Encryption Standard (AES). It is regarded as being considerably faster than triple DES. The excessively small key size of DES necessitated a trade. With growing registering power, it was perceived as being defenseless versus all-out key inquiry attacks. Triple DES was designed to overcome this drawback, but it was thought to be slow. The element of AES is a Symmetric-key square code algorithm that utilizes 128-digit information with 128/192/256-bit keys. It provides complete determination and setup details and is faster and more stable than Triple-DES (Kaur & Kumar, 2020).

AES is an iterative rather than stream cipher. It is carried out by using the concept of replacement permutation technique which determines it. It includes a series of linked processes, some of which replace inputs with clear results (replacements), while others involve moving parts about (changes). It's odd that AES does all of its calculation on bytes rather than bits. From this point forward, AES interprets a plaintext block's 128 pieces as 16 bytes. For handling as a lattice, these 16 bytes are divided into four pieces and four lines. In contrast to DES, the number of rounds in AES varies and depends on the size of the key. AES uses 10 rounds for keys with 128 digits, 12 rounds for keys with 192 pieces, and 14 rounds for keys with 256 cycles. These rounds each use a different 128-cycle round key that is derived from the initial AES key (Xiang et al., 2006).

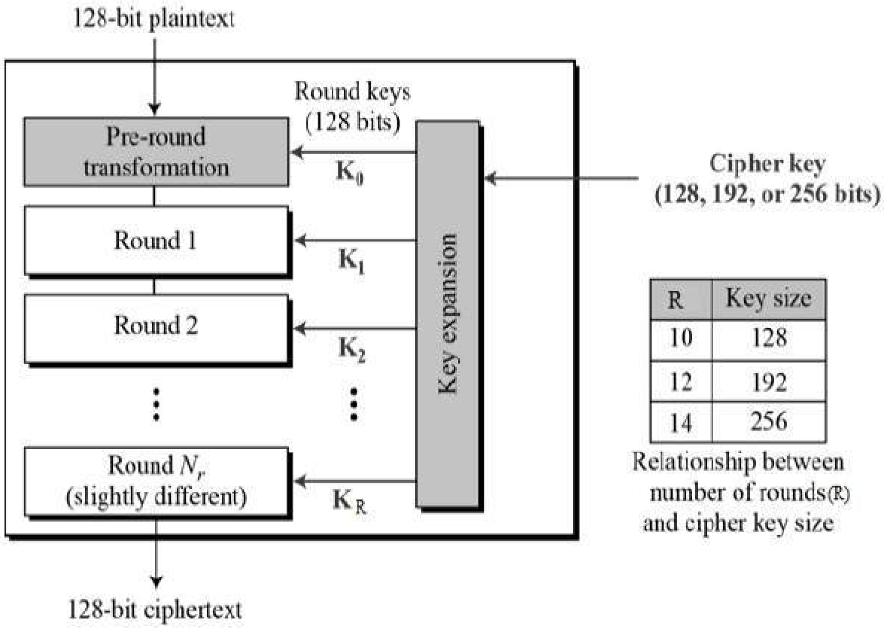


Figure 8. Demonstrates structure of AES algorithm

The main information to be scrambled is separated into a progression of the block. each block (16 input bytes) is supplanted by alluding to a proper table (S-box) given in the plan, and every one of the four columns of the matrix is moved. On the left, every four-byte segment is presently changed over utilizing an exceptional numerical capacity. This capacity takes four bytes of a segment as info and results from four new bytes that supplant the first segment, and 16 bytes of the network are presently viewed as 128 pieces and XORed to 128 pieces of the round key(Nadeem, & Javed, 2005,).

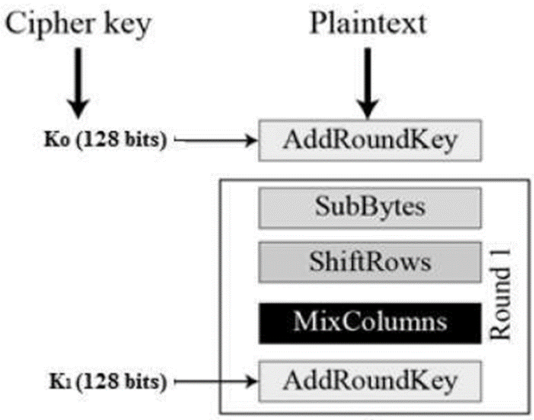


Figure 9. Demonstrates steps of the AES encryption process

RSA Encryption Algorithm

The RSA encryption is a largely utilized asymmetric encryption (SE) method in a range of products and services. Hence to carry out encryption and decryption in a specific data, the SE makes use of a couple of keys which are mathematically consolidated. The set of key pairs are formed by using a private and a public key, in which the general public key is to be had to anyone, and the personal secret is a secret known only to the creator of the important thing pair. With RSA, facts may be encrypted with either the private or public key, at the same time as the alternative key decrypts it. This is one of the motives why RSA is the maximum widely used asymmetric encryption approach (Dhakar, Ravi Shankar, Amit Kumar Gupta, 2012).

The functionality to scramble with either the non-public or public key continues a massive range of services to RSA users. When the public one used for encryption, the data have to be decrypted with a personal or private key. This is ideal for transferring sensitive information over a connection or the Internet, wherein the information receiver affords its public key to the sender. After the use of the receiver's public key to encrypt the sensitive facts, the sender of the data transmits it to the recipient. Because the public key has encrypted the touchy facts, most effective the proprietor of the non-public key can decrypt it. Only the intended recipient of the records may decrypt it, even though the records is intercepted within the transmission (Liu & Yang, 2018). Using a personal key to encrypt a message is another approach to making use of RSA for asymmetric encryption. In this illustration, the sender of the records uses its private key to encrypt the facts before sending both the encrypted facts and its public key to the recipient of the information. Using the sender's public key to decode the data, the receiver can then affirm that the sender is who they claim to be. Although this approach enables statistics to be stolen and tested in transit, its true aim is to establish the sender's identification. The recipient will be aware that the records was altered if the information had been stolen and modified within the direction because the general public key cannot decrypt this message. The recipient may be aware that the information became altered at the same time as in transit if the statistics are stolen and updated in transit when you consider that the general public secret is not able to decipher these messages (Aufa et al., 2018). The RSA algorithm makes sure the keys are as safe as possible. The following steps will show you how it works:

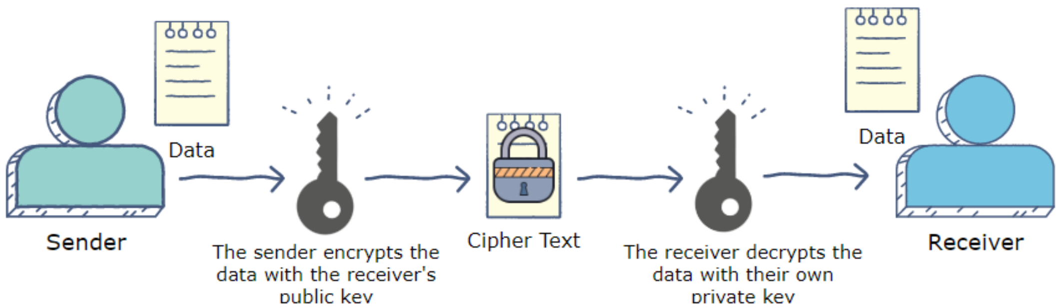


Figure 10. RSA

The RSA algorithm makes sure the keys are as safe as possible. The following steps will show you how it works:

Generating the keys

a) Choose any two big prime numbers, x and y. The prime numbers need to be big so that they shall be hard for anyone to guess it.
b) Compute n=x×y
c) Compute the totient function same as Ø(n)=(x-1)(y-1).
d) Choose an integer e, in such way that e is co-prime to Ø(n) and 1<e<Ø(n). The pair of numbers (n,e) sum up the public key.
e) Compute d such that e.d=1 mode Ø(n). In this cas4 d can be calculated using the extended Euclidean algorithm. The pair (n, d) sum up the private key.

2. Encryption

a) Let plaintext, represented as a number, the ciphertext is calculated as: .

3. Decryption

a. While using the private key, the plaintext can be computed using: .

Elliptic Curve Cryptography (ECC)

Another efficient method for data encoding and encoding is elliptic curve cryptography (ECC) which is encryption generation based on keys. The ECC specializes in pairs of public and private keys for decrypting and encrypting website visitors. A better generation of advanced cryptography than RSA is ECC algorithm. To generate safety among key pairs for public key encryption, it uses elliptic curve arithmetic (Zhang & Wang, 2018).

ECC has lately gained popularity because of its smaller key size and capability to keep security, while RSA calls for excessive top numbers as opposed to ECC to reach similar final results. The desire for devices to stay relaxed grows as keys get longer, placing pressure on the limited cell resources, therefore it is very likely that this trend will continue. For this reason, understanding elliptic curve cryptography is essential (Dawahdeh et al., 2018; Bhanot & Hans, 2015).

ECC's public key encryption techniques, in contrast to RSA's, are based on the algebraic structure of elliptic curves over finite fields. As a result, ECC generates keys that are more difficult to decode mathematically. ECC is therefore regarded as the next generation of public key cryptography and is much more secure than RSA. Applying ECC also makes sense to maintain high levels of safety and overall performance. This is owing to the fact that ECC is being employed more frequently by websites as a means of advancing online data protection while also increasing mobile optimization. The call for a concise guide to ECC grows as greater internet websites use ECC to encrypt information (Bhanot&Hans, 2015). The latest ECC utilizes EC which is a plane curve over a finite field composed of points that satisfy the equation: $y^2=x^3 + ax + b$.

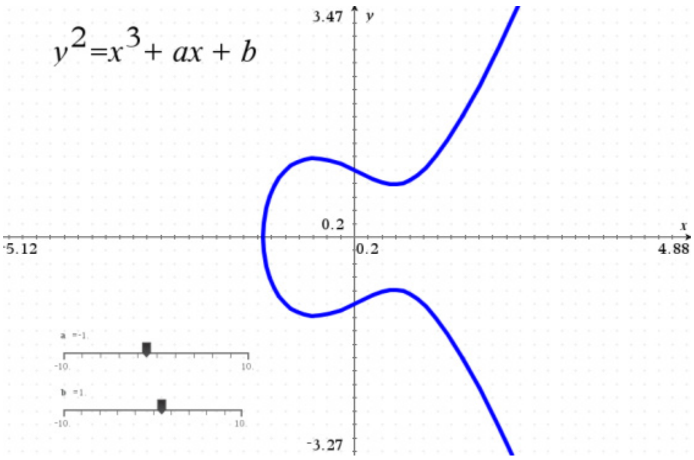


Figure 11. ECC

In this elliptic curve cryptography example, any point on the curve may be reflected over the x-axis and the curve will not alter. Any non-vertical line that crosses the curve fewer than 3 instances can be appeared as such. Elliptic curve encryption is far smaller than different kinds of encryption, which offers smaller, cellular devices greater energy. For two keys of the equal size, RSA's factoring encryption is extra susceptible considering factoring is less difficult and requires less effort than calculating for an elliptic curve discrete logarithm. With ECC, you could use smaller keys to acquire the equal stage of protection. When compared to RSA, ECC offers advanced security with quicker, shorter keys considering cell devices will need to do increasingly more cryptography with reducing computational power inside the destiny(Zhang & Wang, 2018). One of the most famous strategies for imposing digital signatures in cryptocurrencies is ECC. Both Bitcoin and Ethereum use the Elliptic Curve Digital Signature Algorithm to signal transactions (ECDSA). ECC, however, is used outdoors of cryptocurrency as properly. The majority of online apps will finally put into effect this encryption fashionable because of its shorter key period and effectiveness (Wang et al., 2020).

Performance Metrics

Differential Analysis

Many performance matrices are used to study differential attack, but the most common are: the Number of Pixel Change Rate (NPCR) and Unified Average Changing Intensity (UACI) parameters. The differential attack is used to test the encryption algorithm's sensitivity to even smallest changes in the plain picture. Attackers frequently make slight improvements to the plain image. Then, using the same mystery key, encrypt both the plain and modified images. After that, it tries to find a link between encrypted plain and modified image (Kaur & Kumar, 2020).

The Number of Pixel Change Rate is characterized as an action used to compute the proportion of various pixel numbers between two encoded pictures whose ordinary images vary by a solitary pixel. A set of policies is extra resistant to differential assaults if it has an excessive NPCR score (Kaur & Kumar, 2020). The following formula can be used to calculate NPCR (Kaur & Kumar, 2020).

$$NPRC = \frac{\sum_{i,j} D(i,j)}{W * H} * 100 \quad (1)$$

$$D(i,j) = 0 \; if \; E(i,j) = E'(i,j) 1 if \; E(i,j) \neq E'(i,j) \quad (2)$$

W and H stand for the image's width and height, respectively. The variation in comparable pixels between the simple image and the scrambled image shown as E' (i, j) and the scrambled image of the modified image E (i, j) is shown as D (i, j). The range of NPCR € [0, 100]. Also the NPCR value of a scrambled image should be close to 100 (Kaur & Kumar, 2020).

The UACI (Unified Average Changing Intensity) is a matrix for calculating the average intensity of differences between the plain and encrypted image (Dawahdeh et al., 2018). It can be defined in the following way (Kaur & Kumar, 2020).

$$UACI = \frac{(\sum_{i,j} E(i,j) - E'(i,j))}{(255 * W * H)} * 100 \quad (3)$$

Here E(i, j) and E'(I, j) are the encrypted picture of plain and changed images, respectively. The values of NPCR and UACI supposed to be maximized.

Statistical Analysis

Statistical analysis of a scrambled image can also be used to crack encryption systems. The HA (Histogram analysis) as well as CC (correlation coefficient) are image evaluation metrics used to analyze histograms and the correlation between neighboring pixels of an encoded, or original one to confirm the resistance of an coding system to statistical attacks (Kaur & Kumar, 2020).

Histogram Analysis (HA) displays the pixel value distribution of the image. An original image's histogram ought to look entirely different from a scrambled image. Normal photos' histograms are by nature not uniform. The histograms of photos that have been scrambled, however, should be uniform (Kaur & Kumar, 2020). This indicates that the spatial distribution of all pixels is uniform.

Correlation Coefficient (CC) is used to compare identical pixels between an image that is encrypted and one that is not. In the original image, the horizontal, diagonal, and vertical values of adjacent pixels are closely connected. The most effective picture encryption method reduces this relationship in the ciphered picture (Somaraj, 2015). Correlation coefficient can be computed as follows (Bhanot & Hans, 2015):

$$r_{x,y} = \frac{C(x,y)}{\sqrt{D(x)} * \sqrt{D(y)}} \tag{4}$$

$$C(x,y) = \frac{(\sum_{(i=1)}^{K} (x_i - E(x)) * (y_i - E(y)))}{K} \tag{5}$$

$$D(x) = \frac{1}{K} \sum_{(i=1)}^{K} (x_i - E(x))^2 \tag{6}$$

$$D(x) = \frac{1}{K} \sum_{(i=1)}^{K} (y_i - E(y))^2 \tag{7}$$

Here, C(x,y) represents the co-variance between the x and y patterns. Furthermore, x and y are the 2 coordinates of an image. K is the number of pixel couples (x_i, y_i). D(x) and D(y) are the deviation of x and y, simultaneously. E(x) is the mean of xi-pixel values. The range of CC should be a value between [1, 1]. The CC value of an encoded image must be close to zero.

Information Entropy (IE)

It determines the typical amount of data included in each picture bit. It offers all the details that are contained in an image. Every pixel has a different value. So each pixel in an encrypted image has a uniform probability distribution, according to the entropy of the picture (Bhanot & Hans, 2015). It can be calculated as follows (Belazi, Abd El-Latif, & Belghith, 2016):

$$H(s) = -\sum_{s} (P(s_i) * \log_2 P(s_i)) \tag{8}$$

Here H(S) represents the entropy of message source (S). P(s_i) demonstrates the likelihood of occurrence of s_i. The value of IE [0; 8]. It must be close to 8 for 8-bit picture. Bit Correct Ratio (BCR) is used to determine how an encrypted image differs from its original. It checks to make sure the decrypted picture is accurate. (Belazi et al., 2016). It may be solved using the formula (Belazi et al., 2016):

$$BCR = (1 - \frac{(\sum_{(x,y)}^{(M*N)} O(x,y) \oplus R(x,y))}{M*N}) \tag{9}$$

Here, x and y represent the M x N pixel pictures' pixel coordinates. O is a plain image and R is the decrypted image. represents the XOR operation. The range of BCR [0, 1].

Mean Squared Error (MSE)

The MSE is a technique for comparing an original image's "true" pixel values to a decoded image. The discrepancy between an image's original values and its decrypted values is the mistake.(Kaur & Kumar, 2020). Below is a definition of MSE:

$$MSE = \frac{1}{MN} \sum_{(x=1)}^{M} \sum_{(y=1)}^{N} [O(x,y) - R(x,y)]^2 \tag{10}$$

Here, x and y represent the MxN-pixel pictures' pixel coordinates. The encoded and plain pictures are O and R, respectively. The value of MSE [0, ∞]. The value of MSE between plain and decoded must be minimum.

Peak Signal to Noise Ratio (PSNR)

PSNR is used to measure the quality between the plain and decoded images (Kaur & Kumar, 2020). PSNR can be calculated mathematically as follows:

$$PSNR = 10 \log_{10} \frac{(2^n - 1)^2}{MSE} \tag{11}$$

Here n is the number of bits per pixel. PSNR is measured in decibel (dB). The value of PSNR [0, 1]. The value of PSNR must be maximum between plain and decoded images.

Structural Similarity Index (SSIM)

The SSIM shows how comparable the plain and decrypted images are. It is a parameter for evaluating the quality of decrypted images. It is calculated between many windows of images of the same size (Kaur & Kumar, 2020). Below you will find a definition:

$$SSIM = \frac{(2\mu_1\mu_D + C_1)*(2\sigma_{ID} + C_1)}{(\mu_1{}^2 + \mu_D{}^2 + C_1)*(\sigma_1{}^2 + \sigma_D{}^2 + C_2)} \tag{12}$$

Here μ_i and μ_D are the average of an input (I) and decoded (D) image, respectively. Sign2i Sign2d D denote the variance of I and D, respectively. is the covariance of I and D. C_1 and C_2 are the regularization parameters with values $(0.01P)^2$ and $(0.03P)^2$, respectively, where P is the specified value for the dynamic range. The range of SSIM € [-1, 1].

Root Mean Squared Error (RMSE)

The RMSE evaluates the root of the MSE and delivers more exact and accurate statistics. (Cao, Wei, Guo, & Wang, 2017). It can be calculated analytically as:

$$RMSE = \sqrt{\frac{\sum_{x=1}^{M} \sum_{y=1}^{N} [O(x,y) - D(X,Y)]^2}{M*N}} \tag{13}$$

Where the coordinates for each pixel in an image with a size of MxN are x and y. O and D stands for the original and decrypted images, respectively. The range of RMSE € [0,∞].

Mean Absolute Error (MAE)

The MAE calculates the variation between plain and encrypted pictures (Kaur & Kumar, 2020). It can be calculated as:

$$MAE = \frac{1}{MN} / \sum_{x=1}^{M} \sum_{y=1}^{N} |O(x,y) - E(x,y)| \tag{14}$$

Here O(x, y) and E(x, y) mean plain and scrambled images, respectively. x and y demonstrate the pixel coordinates of picture with size of M x N. The range of MAE Execution Time (ET) [0, $2^n - 1$], where n represents the number of bits per pixel. The MAE between the original and encrypted photos should be at maximum.

Execution Time (ET)

The execution time (ET) of a specific picture encryption algorithm is calculated. It is the sum of the compile and run times. ET should be kept to a minimum to make image encryption practical. Seconds, milliseconds, and minutes are the most common units of measurement (Kaur & Kumar, 2020).

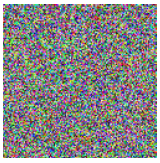
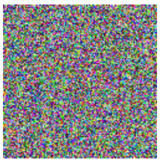
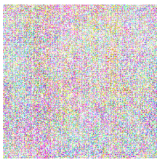
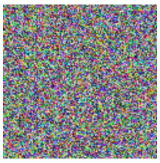
Dataset

The BSD dataset is abbreviated as the Berkeley Segmentation Dataset, which is a dataset designed to conduct different scientific research in the fields of image and computer vision (Arbeláez, Maire, Fowlkes, & Malik, 2011). This database is frequently used for image noise removal, segmentation, enhancement, and many other image processing operations. In this article, some sample color images are taken from BSD to implement the encryption algorithms used in the study. The goal is to test how AES, DES, 3DES, RSA and ECC algorithms can encrypt and decrypt images to protect sensitive digital pictures from attackers and unauthorized users.

Experimental Results

The performance of the different encryption algorithm are examined in this study was using some color images taken from the BSDS500 database. The applications within the scope of the study were developed using MATLAB which effective software environment tool that enables image analysts to perform some image-related activities such as analysis, demonstration, and enhancement, and a comprehensive set of image processing applications and standard algorithms for image processing. The visual image result of the study were given in below:

Table 1. Visual result

Algorithm	Input image	Encrypted image	Decrypted image
AES			
DES			
TDES			
RSA			
ECC			

In the figure above, the encryption and decryption results of the five algorithms in the study. As can be seen, the results of these algorithms are not visually separated but can be distinguished using image evaluation and performance matrices.

In the study, NPCR and UACI analysis was performed, and Table 1 shows all the average values of the different algorithms in the study for the images in the experiment. According to the results of the experiments in Table 1, the NPCR and UACI values for all eight test images are more than 99 and 33%, respectively. These average values are very high, which is due to the diffusion stage (DS) included in these algorithms. The DS causes a large change in the encrypted image when a particular pixel in the plain image is altered, i.e., the DI make the process very susceptible to beginning pixel setting.

Table 2. NPCR and UACI values for one-bit change in key value

	Lena (1)		**28083**		**78019**		**113044**	
Algorithms	**NPCR**	**UACI**	**NPCR**	**UACI**	**NPCR**	**UACI**	**NPCR**	**UACI**
AES	**99.6414**	33.4250	99.7256	**33.8467**	99.6938	33.2103	99.7846	33.2481
DES	94.0019	32.7270	99.0042	33.0121	98.6037	33.0541	98.8054	33.4627
3DES	98.0932	33.0110	**99.8702**	33.7601	99.0454	33.1471	99.4048	**32.8031**

RSA	99.0782	33.0456	99.0166	34.4017	**99.9238**	**33.9975**	**99.89228**	33.7459
ECC	99.0836	**33.9701**	98.0723	33.6403	99.0521	33.4343	99.0461	33.2043
	135069		**157055**		**179084**		**124084**	
Algorithms	**NPCR**	**UACI**	**NPCR**	**UACI**	**NPCR**	**UACI**	**NPCR**	**UACI**
AES	99.8743	33.2947	99.8941	33.3145	**99.7985**	33.0071	**99.5419**	33.9807
DES	97.2654	32.9514	98.7519	33.2974	98.1564	33.0175	96.2126	32.7942
3DES	**99.8785**	33.7534	99.3497	**33.7614**	99.7161	33.4201	98.1317	33.5475
RSA	99.8502	**33.9997**	**99.9734**	34.0014	99.1469	**33.9976**	99.4149	**33.9992**
ECC	98.9985	33.2891	97.3157	33.1764	96.4782	32.1403	99.4712	33.1037

An examination on the correlation analysis (CC) and information entropy (IE) of the mixed and unadulterated pictures. Every one of the pixel sets contains a haphazardly chosen pixel and a neighboring pixel. It tends to be seen that the RSA gave extremely low correlation coefficient values for every one of the eight test pictures, mirroring the high opposition of the technique to statistical attacks. There is a correlation in the flat image, and the pixel values are only occasionally irregular, so the entropy value is largely more modest than the best value of 8. The entropy can be the extreme ideal value of 8 when all pixel values are indiscriminately entangled. Table 2 displays the entropy values of the encoded images. From the results obtained, it can be seen that the entropy of the encrypted images is extremely close to the best value of 8. The data leakage in the cryptographic algorithms is insignificant and is safe from entropy-based attacks.

This demonstrates their decreased resistance to the statistical attacks. However, these qualities are still relatively lesser than the first CC, henceforth guaranteeing protection from statistical attacks dependent upon some degree. An overall pattern of lower connection can likewise be viewed as the picture size increments which support the histogram perception of expansion in thus can't be viewed as effective on visual evaluation grounds. At long last, every one of the plans give the unscrambled pictures like the first pictures which guarantee the unwavering quality of decoded picture on the off chance that security is guaranteed. The CC value is between - 1 and +1, where the limits shows an ideal negative or positive straight correlation respectively. The zero value of CC indicates that there is no linear relationship between consecutive pixel values.

Table 3. CC and IE values for one-bit change in key value

	Lena (1)		**28083**		**78019**		**113044**	
Algorithms	**CC**	**IE**	**CC**	**IE**	**CC**	**IE**	**CC**	**IE**
AES	0.0692	7.8015	0.07193	7.6705	0.0532	7.6126	0.057936	7.9294
DES	0.08345	7.9086	0.06789	7.8785	0.05641	7.6818	0.07414	7.7891
3DES	0.07693	7.8578	0.08657	7.0621	0.06934	7.6718	0.07571	7.4106
RSA	-0.00685	7.6442	0.02311	7.8273	0.24604	7.6079	-0.02828	7.9134

ECC	0.06979	7.7949	0.08291	7.8201	0.02359	7.6418	0.056202	7.4772
	135069		**157055**		**179084**		**124084**	
Algorithms	**CC**	**IE**	**CC**	**IE**	**CC**	**IE**	**CC**	**IE**
AES	0.061627	7.5876	0.08146	7.9604	0.07508	7.6544	0.06709	7.8578
DES	0.095478	7.9898	0.09969	7.5438	0.09351	7.3492	0.08601	7.4141
3DES	0.07810	7.7882	0.09678	7.7141	0.08746	7.7565	0.079605	7.6795
RSA	-0.04217	7.4939	0.05997	7.7459	0.02464	7.8352	0.023765	7.8581
ECC	0.092505	7.6379	0.08934	7.5457	0.07846	7.5102	0.087168	7.7545

In the analysis, BCR measurement was also tested to determine the difference between the plain image and the decoded image. The metric usually checks the correctness of the encoded image. The value of this metric should be a number in the range of zero to one. The MSE was also used to compare the original image's true pixel values to the decrypted image. The error is the difference between the values of an original picture and the values of the decrypted image. The MSE of the encrypted image should be maximized to ensure that attackers cannot find the correct pixels of the original image. The result of both BCR and MSE is shown in the table below.

Table 4. BCR and MSE values for one-bit change in key value

	Lena (1)		28083		78019		113044	
Algo-rithms	BCR	MSE	BCR	MSE	BCR	MSE	BCR	MSE
AES	0.05484	0.60024	0.07468	0.71062	0.01351	0.54721	0.062402	0.31233
DES	0.068878	0.49174	0.08678	0.33645	0.06878	0.38704	0.069878	0.09331
3DES	0.05946	1.7196	0.06964	0.42857	0.43874	0.51065	0.079464	0.27469
RSA	0.041107	0.3617	0.04585	0.62874	0.02107	0.76454	0.061107	0.84787
ECC	0.052402	2.09834	0.06738	0.62242	0.05914	0.55626	0.058465	0.70116
	135069		157055		179084		124084	
Algo-rithms	BCR	MSE	BCR	MSE	BCR	MSE	BCR	MSE
AES	0.055439	0.40979	0.06402	0.63092	0.08242	0.65434	0.072402	0.47971
DES	0.068878	0.64464	0.0887	0.54234	0.07978	0.6986	0.08878	0.7219
3DES	0.089464	0.63371	0.07946	0.37046	0.08064	0.82154	0.079464	0.4959
RSA	0.043107	0.09751	0.04117	0.95421	0.03107	0.08741	0.051107	0.7883
ECC	0.050914	0.60256	0.05914	0.46545	0.04814	0.84029	0.060914	0.4857

The PSNR metric of the encoded image was tested in the analysis. The PSNR value range should start at zero and go to infinity. Normally PSNR should be maximized as much as possible. The Structural Similarity Index (SSIM) has also been tested, and the encrypted SSIM value should be minimized so that attackers cannot fully locate the original image. The PSNR and SSIM results are shown in the table below.

Table 5. Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index (SSIM)

	Lena (1)		28083		78019		113044	
Algorithms	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
AES	39.5111	0.00841	38.5956	0.02680	36.9558	0.00602	35.5718	0.00475
DES	39.0102	0.00852	40.4546	0.00529	38.3998	0.00598	37.4035	0.00483
3DES	40.4991	0.00653	41.1322	0.00498	37.6827	0.00560	37.5261	0.00491
RSA	41.0848	0.00504	18.5969	0.01003	9.02328	0.00116	13.2627	0.01008
ECC	36.5244	0.00527	37.3107	0.00937	39.5754	0.00676	37.7165	0.00633
	135069		157055		179084		124084	
Algorithms	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
AES	36.3375	0.00627	36.0703	0.00674	33.6552	0.00706	37.041	0.00836
DES	40.8517	0.00454	37.333	0.00730	39.7253	0.00752	41.1306	0.00906
3DES	40.7503	0.00414	37.4737	0.00698	38.9231	0.00629	41.0213	0.00832
RSA	9.7663	0.00324	10.0839	0.00177	7.71929	0.00535	6.61093	0.00804
ECC	39.3381	0.00435	38.1078	0.00439	36.8471	0.00632	39.2697	0.00842

The RMSE metric of the encoded image was tested in the analysis. The RMSE range of values should start at zero and go to infinity. Normally PSNR should be minimized as much as possible. The RMSE results are shown in the table below.

Table 6. Root Mean Squared Error (RMSE)

	Lena (1)	**28083**	**78019**	**113044**
Algorithms	**RMSE**	**RMSE**	**RMSE**	**RMSE**
AES	0.90103	0.89722	0.9246	0.88303
DES	0.89493	0.87993	0.8719	0.87493
3DES	0.91068	0.91068	0.9168	0.92166
RSA	0.95906	0.97176	0.8916	0.85906
ECC	0.88703	0.89418	0.8835	0.88292
	135069	**157055**	**179084**	**124084**
Algorithms	**RMSE**	**RMSE**	**RMSE**	**RMSE**
AES	0.91183	0.88702	0.8873	0.8832
DES	0.87493	0.87493	0.8749	0.8794
3DES	0.91268	0.93168	0.9168	0.9166
RSA	0.89906	0.89596	0.8996	0.8956
ECC	0.88535	0.88135	0.8535	0.8813

In the analysis, the Mean Absolute Error (MAE) and Execution Time (ET) matrices of the encoded image were tested. The MAE value range should start at zero and go to infinity. Normally MAE should be maximized as much as possible to prevent attacks from decrypting the encrypted image. On the other hand, ET was tested to measure the time required to execute a given image encryption algorithm. The ET should be kept to minimum to show the performance of the encryption algorithm. The ET and MAE results are shown in the table below.

Table 7. Mean Absolute Error (MAE)and Execution Time (ET)

	Lena (1)		28083		78019		113044	
Algorithms	MAE	ET	MAE	ET(sec)	MAE	ET	MAE	ET
AES	86.4335	1.29688	82.1031	1.73438	80.3625	1.2338	87.03155	1.4844
DES	84.773	2.3594	77.0587	2.60938	79.1609	2.9838	83.05234	2.48438
3DES	85.0148	5.4375	81.0904	4.85938	81.1789	3.4065	84.73543	5.4375
RSA	87.0656	5.23438	86.2739	5.95313	86.456	4.935	87.3679	5.6875
ECC	88.0638	1.95313	84.2384	2.04587	85.1615	1.52	87.03624	2.04688
	135069		157055		179084		124084	
Algorithms	MAE	ET	MAE	ET	MAE	ET	MAE	ET
AES	87.07656	1.65625	88.1904	1.4375	85.0267	1.625	89.47442	1.77656
DES	85.14562	1.9375	84.4277	2.10938	83.4852	2.4025	85.8698	2.71875
3DES	86.34568	3.28125	85.7598	4.29688	86.0183	4.4375	87.07948	5.92188
RSA	88.4913	5.625	89.4882	5.48438	86.8384	5.4013	91.8328	5.78125
ECC	87.76373	2.125	82.0598	1.9375	83.6637	1.7663	86.17841	1.82813

Conclusion

Digital images are assuming an inexorably significant part in our lives, because of the huge advances in imaging gadgets and the innovations of these technologies. This makes image security one of the most interesting research fields in image processing and computer vision. Attackers always take advantage of weak or unprotected data transmitted and stored over the internet. They try to sit in the middle of the network to control the data transfer over the network. Therefore, they focus on critical data shared between the ten senders and receivers. In this study, we conducted a comparative study of five encryption algorithms using color images obtained in the BSD300 database. The main purpose is to compare the performances of these algorithms using different methods of image quality evaluation. The algorithms in the scope of the study include AES, DES, TDES, RSA and ECC. The first three algorithms are considered symmetric methods, while other algorithms are considered asymmetric encryption methods. In the experiment, some color images from BSD were used to test the performance of the algorithms. The efficiency of encoding was measured using some image evaluation methods and the results were presented in the study.

In the future, more encryption algorithms such as Blowfish, RC4, RC5, RC6, and Chaotic Map can be added to the work. In addition, performance evaluation criteria can be expanded by adding a feature similarity measurement index (FSIM), key spacing, pixel-based analysis, and many other image evaluation criteria.

References

Arbeláez, P., Maire, M., Fowlkes, C., & Malik, J. (2011). Contour detection and hierarchical image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 33(5), 898–916. https://doi.org/10.1109/TPAMI.2010.161.

Aufa, F. J., Endroyono, & Affandi, A. (2018). Security System Analysis in Combination Method: RSA Encryption and Digital Signature Algorithm. Proceedings - 2018 4th International Conference on Science and Technology, ICST 2018. https://doi.org/10.1109/ICSTC.2018.8528584.

Barker, E., & Mouha, N. (2017). Recommendation for the triple data encryption algorithm (TDEA) block cipher.

Belazi, A., Abd El-Latif, A. A., & Belghith, S. (2016). A novel image encryption scheme based on substitution-permutation network and chaos. Signal Processing, 128, 155–170. https://doi.org/10.1016/J.SIGPRO.2016.03.021.

Bhanot, R., & Hans, R. (2015). A Review and Comparative Analysis of Various Encryption Algorithms. International Journal of Security and Its Applications, 9(4), 289–306.

Cao, X., Wei, X., Guo, R., & Wang, C. (2017). No embedding: A novel image cryptosystem for meaningful encryption. Journal of Visual Communication and Image Representation, 44, 236–249. https://doi.org/10.1016/J.JVCIR.2016.08.003.

Cardarilli, G. C., Di Nunzio, L., Fazzolari, R., & Re, M. (2015). TDES cryptography algorithm acceleration using a reconfigurable functional unit. 2014 21st IEEE International Conference on Electronics, Circuits and Systems, ICECS 2014, 419–422. https://doi.org/10.1109/ICECS.2014.7050011.

Data Encryption Standard. (2022). Retrieved July 22, 2022, from https://www.tutorialspoint.com/cryptography/data_encryption_standard.htm.

Davis, R. M. (1978). The Data Encryption Standard in Perspective. IEEE Communications Society Magazine, 16(6), 5–9. https://doi.org/10.1109/MCOM.1978.1089771.

Dawahdeh, Z. E., Yaakob, S. N., & Razif bin Othman, R. (2018). A new image encryption technique combining Elliptic Curve Cryptosystem with Hill Cipher. Journal of King Saud University - Computer and Information Sciences, 30(3), 349–355. https://doi.org/10.1016/J.JKSUCI.2017.06.004.

Dhakar, Ravi Shankar, Amit Kumar Gupta and P. S. (2012). Modified RSA encryption algorithm (MREA). Second International Conference on Advanced Computing & Communication Technologies, 426–429.

Furht, B., Socek, D., & Eskicioglu, A. M. (2004). Fundamentals of Multimedia Encryption Techniques. Multimedia Security Handbook, 95–132. https://doi.org/10.1201/9781420038262-3.

Guan, Z. H., Huang, F., & Guan, W. (2005). Chaos-based image encryption algorithm. Physics Letters A, 346(1–3), 153–157. https://doi.org/10.1016/J.PHYSLETA.2005.08.006

Kansal, S., & Mittal, M. (2015). Performance evaluation of various symmetric encryption algorithms. Proceedings of 2014 3rd International Conference on Parallel, Distributed and Grid Computing, PDGC 2014, 105–109. https://doi.org/10.1109/PDGC.2014.7030724.

Kanso, A., & Ghebleh, M. (2012). A novel image encryption algorithm based on a 3D chaotic map. Communications in Nonlinear Science and Numerical Simulation, 17(7), 2943–2959. https://doi.org/10.1016/J.CNSNS.2011.11.030.

Kaur, M., & Kumar, V. (2020). A Comprehensive Review on Image Encryption Techniques. Archives of Computational Methods in Engineering, 27(1), 15–43. https://doi.org/10.1007/S11831-018-9298-8.

Khan, M., & Shah, T. (2014). A Novel Statistical Analysis of Chaotic S-box in Image Encryption. 3DResearch 2014 5:3, 5(3), 1–8. https://doi.org/10.1007/S13319-014-0016-5.

Lai, X. J., Lu, M. X., Qin, L., Han, J. S., & Fang, X. W. (2010). Asymmetric encryption and signature method with DNA technology. Science China Information Sciences 2010 53:3, 53(3), 506–514. https://doi.org/10.1007/S11432-010-0063-3.

Liu, Yang, et al. (2018). Secure and robust digital image watermarking scheme using logistic and RSA encryption. Elsevier, 95–105.

Nadeem, A., & Javed, M. Y. (2005, August). A performance comparison of data encryption algorithms. In 2005 international Conference on information and communication technologies (pp. 84-89). IEEE. https://doi.org/10.1109/ICICT.2005.1598556.

Sharif, S. O., Kuncheva, L. I., & Mansoor, S. P. (2010). Classifying encryption algorithms using pattern recognition techniques. Proceedings 2010 IEEE International Conference on Information Theory and Information Security, ICITIS 2010, 1168–1172. https://doi.org/10.1109/ICITIS.2010.5689769.

Somaraj, S., and, M. H.-I. J. of S., & 2015, undefined. (2015). Performance and security analysis for image encryption using key image. Researchgate.Net. https://doi.org/10.17485/ijst/2015/v8i35/73141.

Symmetric vs. Asymmetric Encryption. (2022). Retrieved July 22, 2022, from https://www.ssl2buy.com/wiki/symmetric-vs-asymmetric-encryption-what-are-differences.

Wang, H., He, D., & Ji, Y. (2020). Designated-verifier proof of assets for bitcoin exchange using elliptic curve cryptography. Future Generation Computer Systems, 107, 854–862. https://doi.org/10.1016/J.FUTURE.2017.06.028.

Xiang, T., Liao, X., Tang, G., Chen, Y., & Wong, K. W. (2006). A novel block cryptosystem based on iterating a chaotic map. Physics Letters, Section A: General, Atomic and Solid State Physics, 349(1–4), 109–115. https://doi.org/10.1016/J.PHYSLETA.2005.02.083.

Ye, G. (2010). Image scrambling encryption algorithm of pixel bit based on chaos map. Pattern Recognition Letters, 31(5), 347–354. https://doi.org/10.1016/J.PATREC.2009.11.008.

Zhang, X., & Wang, X. (2018). Digital Image Encryption Algorithm Based on Elliptic Curve Public Cryptosystem. IEEE Access, 6, 70025–70034. https://doi.org/10.1109/ACCESS.2018.2879844.

Zhou, Y., Cao, W., & Philip Chen, C. L. (2014). Image encryption using binary bitplane. Signal Processing, 100, 197–207. https://doi.org/10.1016/J.SIGPRO.2014.01.020.

Zhu, Z. L., Zhang, W., Wong, K. W., & Yu, H. (2011). A chaos-based symmetric image encryption scheme using a bit-level permutation. Information Sciences, 181(6), 1171–1186. https://doi.org/10.1016/J.INS.2010.11.009.

About the Authors

Mehmet HACIBEYOGLU, PhD, is an Associate Professor of Computer Engineering at Necmettin Erbakan University in Konya, Türkiye. He holds a PhD in Computer Engineering from Selcuk University. His main areas of interest are artificial intelligence, data mining and cyber security applications.
E-mail: hacibeyoglu@erbakan.edu.tr, **ORCID:** 0000-0003-1830-8516.

Bashir Sheikh Abdullahi JAMA, is a computer engineer and junior researcher in the field of computer science and engineering, He is interested in conducting research that improves human life and helps to adapt to the modern technological world. His research interests include artificial intelligence, machine learning, computer vision digital image processing. Currently, he is a PhD student in the Department of Computer Engineering at Necmettin Erbakan University.
E-mail: bashiir204@gmail.com, **ORCID:** 0000-0002-2417-9223

Similarity Index

The similarity index obtained from the plagiarism software for this book chapter is 20%.

To Cite This Chapter:

Hacibeyoglu, M., & Jama, B.S.A (2022). Performance Analysis of Symmetric and Asymmetric Encryption Algorithms in Digital Images, In S. Kocer, O. Dundar (Eds .), *Current Studies in Basic Sciences Engineering and Technology*, (pp. 92–113). ISRES Publishing.

Analysing Trends and Applications in Data Compression Techniques

Sabri KOCER
Necmettin Erbakan University

Bashir Sheikh Abdullahi JAMA
Necmettin Erbakan University

Yasin ER
Necmettin Erbakan University

Beyza DUNDAR
KTO Karatay University

Introduction

Information storage and transmission are frequently significant issues for businesses, states, and different associations. These organizations can deal with additional information by packing it, which likewise decreases the space and cost expected to store it. It could be valuable to appreciate how compress capabilities and the advantages it might propose to you and your association on the off chance that you save or send information as an aspect of your responsibilities. Data compression (DC) makes documents occupy less space on a hard drive and gets some margin time to transport or download. Cost decreases may be huge because of this decrease in existence. For example, organizations and medical care offices that keep a ton of information could lessen their capacity costs by utilizing compression to store more records on a given amount of space. Furthermore, since compressed documents transfer more quickly over the web or internet, these businesses and organizations don't need to spend as much money on pricey bandwidth upgrades. Compression empowers a few different organizations to offer the most ideal help at the most advantageous time. Consider how much sound and video information is dealt with by telecoms organizations. They can serve much more customers because of compression, which doesn't altogether bring down the sound or video quality.

In this chapter, we present a brief definition of the terminologies used in data compression, examine their significance, outline different compression techniques, explain the metrics used to compression performance and give an overview of how to apply compression to various types of data.

A Brief History of Data Compression

The history of DC dates back to the 18th century, when Morse code was invented, which was used for broadcast use in 1838. it dates back to the 18. century, is an exemplary illustration of the DC that depends on using more limited codewords for English alphabetic letters like "e" and "t" that are more predominant. With the appearance of data hypothesis in the last part of the 1940s, current DC research started off (Wolfram, 2003). Claude Shannon and Robert Fano created a methodical methodology for allocating codewords based on block probability In 1949. David Huffman later discovered an ideal approach for doing this in 1951. Early implementations were frequently carried out in hardware, and certain codeword selections were made as trade-offs between compression and error correction(Blelloch, 2013). A methodical methodology for allocating codewords based on block probabilities was developed in 1949 by Claude Shannon and Robert Fano. Then, in 1951, David Huffman discovered the best technique for doing this. Early implementations were primarily hardware-based, with particular codeword selections made as trade-offs between compression and error-correction(Thanki & Kothari, 2019b). With the rise of the Internet in the 1970s, the transfer speed of the File increased significantly. The first studies on universal lossless compression algorithms were on Lempel-Ziv-Welch (LZW) in the mid-1980s. (Badshah et al., 2016). With LZW, the first widely used data compression algorithm introduced in computers, a text file could be reduced to about half. (Thanki & Kothari, 2019b). The Morse code, created by Samuel Morse in the middle of the 19th century, is an early instance of data compression. Dots and dashes are used to encode leers conveyed by telegraph. Morse observed that some letters appeared more frequently than others (Salomon, 2008). He assigned shorter sequences to leers that happen more frequently in order to decrease the average amount of time needed to transmit a message like e (·) and a(· −), and longer successions to sneers that happen less oftentimes, like q (− − · −) and j (· − − −). This thought of involving more limited codes for all the more every now and again happening characters is utilized in Huffman coding, which we will portray in the chapter.

How Data Compression Works

He gave leers that occur more frequently shorter sequences in order to decrease the average amount of time needed to transmit a message (Sayood, 2017). Every DC method has its own set of guidelines. As an illustration, when text compression is started, the computer will give each of the text gaps a single byte. It will then pack the byte into a string that instructs the decoder where to place everything back (Jovanovic, 2022).

Image compression operates same way as text. Depending on the algorithm, you can get a file that is significantly smaller but with noticeably worse image quality, or something that is about the same size and virtually similar to the original. By either eliminating extraneous data or collecting identical or related bytes and assigning them a new value, compression enables the computer to reassemble the original data (Jamil et al., 2022; Otair et al., 2022).

The same information is repeated throughout most computer file types, making them fairly redundant. Programs for compressing files merely eliminate redundant information. A file-compression program presents a piece of information only once and then references it anytime it appears in the original program, as opposed to listing the same information repeatedly. Consider John F. Kennedy's 1961 inaugural address, where he gave the following well-known quote: **"Ask not what your country can do for you -- ask what you can do for your country."** The quote consists of 17 words, encompassed 61 letters, 16 spaces, 1 dash, and 1 period make up the quotation. The entire file size is 79 units if each letter, space, and punctuation mark eats up one unit of memory. We must look for duplications in order to reduce the file size.

Word	Number of occurrences
Ask	2
What	2
Your	2
Country	2
Can	2
Do	2
For	2
You	2

About half of the phrase is unnecessary if the distinction between capital and lower-case letters is ignored. We have practically all the information we require for the complete quote in nine words. We just point to the words in the first half of the statement and add the necessary punctuation and spaces to complete the second half.

Classification of Data Compression Techniques

The data compression (DC) approach has recently been created for a variety of data kinds, making it challenging to choose the right strategy for a given application. The quality of the data, different coding methods, different types of data, and applications can all be used to classify data compression approaches (Jayasankar et al., 2021). The following subsections provide more information on various DC techniques.

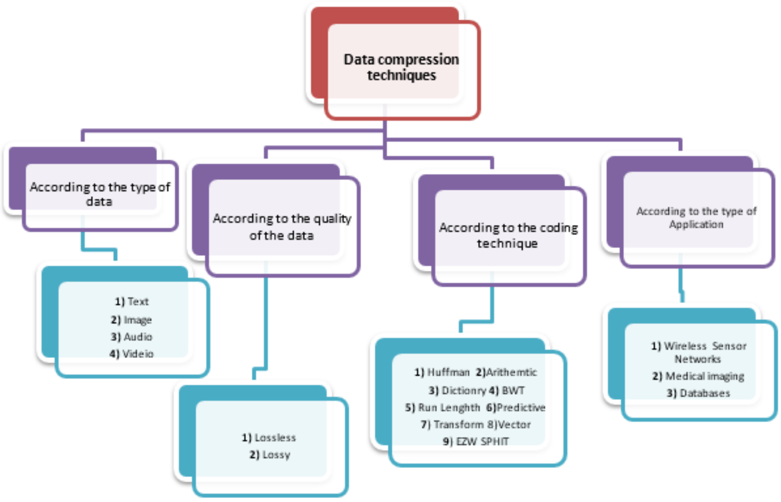


Figure 1. Classification of the DC methods

Classification of DC by data type

Techniques for DC are frequently employed for text, image, video, and audio data (Thanki & Kothari, 2019a). This section covers information on various DC methods and their features for each type of data.

Text Data

Lossless compression methods are frequently employed for text data. In order to compress textual data, Abel and Teahan created a number of preprocessing methods in 2005, including word substitution, phrase replacement, transliteration (from capital to tiny), coding of line endings, and alphabet recording (Navvabi et al., 2022). The drawback of these techniques is that data compression requires longer processing times. Plato presented a BWT-based compression method for repeating text files in about 2008 (Díaz-Domínguez & Navarro, 2022). The BWT approach investigated binary text data's Boolean reductions to produce compressed text data. Here, compressed text data was encoded using BWT coding (Platoš et al., 2008). Numerous real-world applications, including Facebook and texting, use this concept. Reversible transformation-based pre-processing method for converting text data to different forms was created by Roberts and Nadarajan (Sayood, 2018b, 2018a).

Image Data

The idea of picture compression was developed as a result of efforts to optimize website and media data over the network. By using picture compression, the image size can be decreased without significantly affecting the image's quality. You'll be able to do this to cut down on transfer and storage costs as well as time (Nuha et al., 2022).

The process of compressing image files can be done in a variety of ways. The JPEG format and the GIF format are the two most popular compressed graphic image formats for use on the Internet. While the GIF approach is frequently used for line art and other images with relatively simple geometric features, the JPEG methodology is more frequently utilized for photography. Wavelets and fractals are two other methods of image reduction(Javed et al., 2018). As of the time of this writing, these techniques have not become widely accepted for use on the Internet. Nevertheless, both techniques show potential as they provide higher compression ratios for particular types of images than the JPEG or GIF techniques (Larobina & Murino, 2014). The PNG format is another emerging technique that might eventually take the place of the GIF format. There is a limit to how much a text file or software may be compressed without introducing mistakes. Lossless compression is the term for this. Errors are introduced after this. Lossless compression is essential for text and program files because a single error can significantly alter a text file's meaning or prevent a program from running. A slight quality reduction during image compression is typically undetectable. There is no "critical point" beyond which compression is impractical but up to which it functions perfectly(Javed et al., 2018). The compression factor can be higher when there is some loss tolerance than it can be when there is none. Graphical pictures can therefore be compressed more than text or program files.

You can gain from image compression by having your photos load more quickly and having your webpages load more quickly (Brahimi et al., 2022). Additionally, image compression might provide you with extra advantages like:

- **Reduction of the File Size**: One of the main advantages of image compression is this. You can reduce the image until it is the right size, regardless of the file size you are working with. This suggests that, until you alter the image, it occupies less space on the webpage and stays the same size physically.
- **Create Websites with less Bandwidth**: Webmasters can build websites with numerous images without using a lot of storage or bandwidth if they choose reduced image sizes.
- **Faster Loading on Mobiles**: People access websites on a variety of devices, including tablets, iPads, and much more, in addition to mobile phones. Your website will load more quickly on all devices if the picture files on it are small.
- **Maintain Image Quality**: If the quality of your photographs is an issue for you, you can conduct lossless image compression without suffering any image quality loss.
- **Easy Storage**: You can save more photographs on your hard disk since lower image sizes use less storage space than larger image sizes.

Audio Data

In order to drastically reduce file sizes and make it easier to share music and recordings over the network, audio data compression is a crucial and essential technique. Simply put, the current audio streaming environment would not exist without audio compression. Recently, authors created the linear prediction coding-based audio compression standard IEEE 1857.2 (LPC) (RANDHAWA, 2022). This standard offers higher-quality, lossless compression for audio signals with fast processing. In this method, pre-processing and an entropy coding block are done after the audio stream has already been encoded using LPC. When applied to voice and audio signals, this approach is effective. This method allows compressed original audio signals to be recreated. The development of a scalable variable bit rate encoding-based audio compression method (Hoang Thi Hang, 2022) It can be used for any variation in bandwidth.

Video Data

Real-time video compression is a challenging and significant subject that has sparked a lot of study. This corpus of knowledge has been heavily included into the H.261/H263 and MPEG motion video specifications (Ma et al., 2020). However, certain crucial issues have not yet been addressed. This chapter describes a potential replacement for these standards that has superior compression properties and requires far less processing power to fully implement. (Duan et al., 2020).

Moving digital video pictures have been incorporated into applications since around 1989. The enormous bandwidth needed for the encoding of video data is what makes it challenging to integrate moving digital video. A quarter-screen image (320 x 240 pixels), for instance, needs to be stored and sent at a rate of 6.9 million bytes per second in order to play at full speed on an RGB video screen at 30 frames per second (FPS). This data rate is just unaffordable, so finding ways to compress digital video so that it may be played back in real-time is a must for the mainstream adoption of digital motion video applications(Chowdhery & Chiang, 2018). There are numerous digital video compression techniques that have been created and used. These algorithms' compression ratios (CR) change depending on the subjectively accepted degree of error, how the term "compression" is used, and who is making the claim. The digital video often has a lot of information for both transmission and storage (Javed et al., 2018). Compression is therefore required for the storage and transmission of any video stream. Researchers created a number of methods to effectively compress video streams. The Moving Pictures Experts Body (MPEG) is an ISO working group that creates numerous video stream compression standards (Duan et al., 2020).

Classification of DC by data quality

Generally, the employment of a data compression technique in a particular application has an impact on the quality of the data. The quality of compressed data is not given much consideration when compression techniques are utilized in applications that pertain to communication. However, the data loss is intolerable in applications involving text data compression. Small changes in image pixels are also undesirable when compressing satellite images or medical images. As a result, the type of data or application involved has a significant impact on the quality of the data in the data compression technique. Data compression techniques can be classified into lossless compression and lossy compression based on the quality of the compressed data that is required (Thanki & Kothari, 2019a).

Reconstructed data from compressed data in lossless compression resembles the original data. Data cannot be increased by the noise. Thus, it is referred as lossless. Most applications where information loss is unacceptable, such as those involving medical images, text, satellite imagery, etc., use lossless data compression. Reconstructed data in lossy compression is not an exact replica of the original data. This kind of approach involves a certain amount of data or information loss. In comparison to lossless compression techniques, lossy compression approaches can obtain a larger compression ratio(Blelloch, 2013; Rao & Yip, 2018).

Classification of DC by coding approach

This section categorizes data compression methods based on well-known coding methods like dictionary-based predictive coding (PC), transform coding (TC), fractal compression (FC), wavelet transform (WT), and vector quantization (VQ). Also these methods include arithmetic, run-length encoding (RLE), Burrows-Wheeler Transform (BWT), Huffman coding (HC), and run-length encoding (RLE).

Huffman Coding

The fundamental concept of file compression is Huffman coding, a data compression algorithm. The method functions by building a binary tree with nodes. A node can either be an internal node or a leaf node. All nodes are initially leaf nodes, which include the character's weight (frequency of presence), as well as the character itself. Character weight and linkages to two child nodes are both found in internal nodes. Following the left child is denoted by bit 0 and following the right child by bit 1, respectively. n leaf nodes and n-1 interior nodes make to a finished tree. For the most ideal code lengths, it is advised that Huffman Tree throw away any unneeded characters in the text (Sayood, 2018a). The Huffman Tree can be built using a priority sequence as a basic illustration of the HC method, with the node with the lowest frequency having the highest priority. Below are all of the steps:

1. For every character, make a leaf node and add them to the priority queue.

2. Although there are multiple nodes in the queue:

- Remove the two nodes of the highest priority (the lowest frequency) from the queue.
- Add these two nodes as children of a new internal node with a frequency equal to the sum of both nodes' frequencies.
- Add the new node to the priority queue.

3. The tree is finished when just the root node remains.

Arithmetic Coding

Important coding for the creation of compressed data utilising variable length codes was created by Langdon in 1984 (Kumar et al., 2021). Arithmetic coding is the name of this coding method. This method is more popular in practical applications and is superior to human coding. Lossless data compression employs a type of entropy encoding known as arithmetic coding (AC). As with the ASCII code, a string of characters is often represented using a fixed number of bits per character. An algorithm for arithmetic coding converts a whole file's worth of symbols into a single decimal integer (Guo et al., 2021). Each iteration processes a single input symbol. The code word for the complete string of symbols is chosen using the interval that is generated at the end of this division operation. Two categories of coding exist, including binary arithmetic coding (Klein & Shapira, 2021) and adaptive arithmetic coding (Dudhagara & Patel, 2017). When dealing with sources with limited alphabets, including binary sources, and alphabets with significantly skewed probability, the AC is very helpful. When the modelling and coding components of lossless compression need to be kept apart for a variety of reasons, it is also a very helpful method.

Dictionary Coding

In order to produce greater compression, dictionary compression algorithms use a variety of strategies that take into account the data's structure (Cinque et al., 2009). The idea is to avoid keeping redundant strings for words and phrases that repeatedly appear in the text stream. The most frequently used words or phrases are kept in a dictionary, which is used by the coder to generate tokens for output. Ideally, the words or phrases are repeated throughout the page and the tokens are significantly shorter than the actual words or phrases (Thanki & Kothari, 2019b).

When reading an input string, an encoder looks for recurrent words and outputs those words' dictionary indices. A new word is produced in its uncompressed form and added as a new entry to the dictionary. The primary operations are efficient encoding, dictionary upkeep, and string comparison. Both compressors and decompressors keep a lexicon on their own. Generally speaking, dictionary-based algorithms are quicker than entropy-based ones (Cinque et al., 2009; Navvabi et al., 2022). Instead of treating the input as a stream of bits, they treat it as a series of characters. The compression method receives a stream of symbols as input, and the output is a combination of tokens and words in their original form. The coding system can be categorized as working in variable-to-fixed fashion when it outputs tokens since, in its simplest version, each string to be encoded is of a varied length but the code words, or the dictionary's indexes, are of the same length. Due to the fact that the dictionary is changed during the compression and decompression processes, dictionary-based techniques are adaptable (Blelloch, 2013). The dictionary's contents change depending on the order in which the compressed text is entered. Dictionary techniques rely on finding recurring patterns rather than any statistical model. As a result, neither the statistical model's quality nor the source's entropy have any bearing on the compression effect. As a result, it can frequently obtain a better compression ratio than techniques that use statistical models (Cinque et al., 2009).

Run Length Encoding (RLE)

A straightforward method of lossless data compression called run-length encoding (RLE) applies to sequences where the same value appears repeatedly. In order to store just one value and its count, the sequence is encoded (Birajdar et al., 2019). For instance, if the sample string is "mmmmaaadexxxxxx", then the process should return "m4a3d1e1x6". To solve this issue, adhere to the methods listed below::

- Choose the first letter of the source string.
- Add the chosen character to the string's final destination.
- Count how many times the chosen character appears after that and add that number to the final string.

- If the string's end isn't reached, choose the next character and repeat steps 2, 3, and 4.

Predictive and Transform

Predictive coding and transform coding are the bases for another classification of data compression techniques. Predictive coding uses input data that is already accessible to anticipate compressed data, and it performs encoding depending on the discrepancy between the predicate data and the actual input data (Chowdhery & Chiang, 2018). This method is the simplest, clearest, and most useful for compressing texture images. When using transform coding, the input data is first converted into its transformed coefficients, and then these coefficients are encoded. Here, different transformations like the discrete Fourier transform (DFT), discrete cosine transform (DCT), and discrete wavelet transform (DWT) are widely used (Rao & Yip, 2018). The transform coding method, which is referred to as the better compression method, offers superior compression to data utilising a small number of coefficients.

Vector-based Compression

In vector quantization, blocks or vectors are formed from the source output. As an illustration, we can think of [X] successive speech samples as the elements of a [x-dimensional] vector. Alternately, we can treat each pixel value in a block of L pixels from an image as a component of a vector of size or dimension [X]. The vector quantizer receives this vector of source outputs as input. We have a collection of [X-dimensional] vectors at the vector quantizer's encoder and decoder known as the vector quantizer's codebook(Blelloch, 2013; Faria et al., 2012). The "code-vectors" in this codebook are chosen to be typical of the vectors we produce from the source output. A binary index is assigned to each code-vector. The input vector is compared to each code-vector at the encoder to determine which code-vector is most similar to the input vector (Blelloch, 2013). The quantized values of the source output make up the codevector's elements. We broadcast or store the binary index of the code-vector in order to provide the decoder with information about which code-vector was determined to be the most similar to the input vector. Given its binary index, the decoder can obtain the code-vector because it uses the exact codebook.

EZW SPHIT

The Embedded Zero-tree Wavelet is known as EZW. On many test photos, it provides good results for compression. J. Shapiro provided the introduction. It uses adaptive arithmetic coding and is a lossless compression technology. It is a useful algorithm for compressing images (Dudhagara & Patel, 2017). For picture coding, it generates a fully embedded bit stream. This method uses DWT, critical information is missing, entropy coding, and lossless compression using arithmetic coding. Most coefficients produced by a subband transform will be 0 or very close to zero when the bit rate is low. Since low frequency information is included in nearly all standard test photos, it is close to zero. High frequency information is sent using high quality schemes. This is crucial to how people judge the quality of images (Dudhagara & Patel, 2017). Set Partitioning in a Hierarchical Tree is referred to as SPIHT. This EZW technique has been changed. It is an image compression coder based on wavelets. It transforms the image into a wavelet. The best image quality and high PSNR values were produced by this method, which is what makes it so essential. It provides a fully progressive bit stream and can be used for lossless image compression. This algorithm's key benefit is that it is entirely progressive. Wavelet transform coefficients are coded and sent using this technique. Therefore, it is beneficial to gradually regenerate the original copy of the image. This technique for compressing images is quite effective. It produces a compressed bit-stream and has fewer complexes, which makes it easier to decode at various transmission rates (Dudhagara & Patel, 2017).

Classification of DC by type of Application

Depending on the application, the DC techniques can be used in a variety of apps. It is challenging to categorise each method according to where it can be used. In this part, various application areas of DC methods are briefly explored along with their features for diverse purposes. A data compression strategy can be used to limit the quantity of information being transferred in a network, resulting in a reduction in power consumption, which is a major power consumer in wireless sensor networks (WSN). This implies that a higher data compression ratio results in a higher percentage of power savings (Shin & Kim, 2016). Image compression significantly reduces the amount of storage space and transmission bandwidth needed, which not only supports the broad use of medical imaging but also serves as one of the most crucial foundational technologies for telemedicine applications. This is crucial in endoscopy as well, which is the process of gathering and saving a lot of data for more accurate diagnosis. Different DC approaches are utilised for different kinds of data, including databases, signals, data from remote sensing, and DNA sequences. As a result, the DC has a wide range of applications, which demonstrates its vital role in current technological trends (Thanki & Kothari, 2019b).

Evaluation metrics for data compression

There are several ways to assess a compression algorithm (AA Azeez, 2016; Baker et al., 2014). We may evaluate the algorithm's relative complexity, the memory needed to implement it, the speed at which it runs on a specific machine, the level of compression, and how closely the reconstruction resembles the original (Faria et al., 2012). The final two criteria will be our main focus in this chapter. Let's examine each one individually. Looking at the ratio of the number of bits needed to represent the data before compression to the number of bits needed to represent the data after compression is a pretty natural way to gauge how well a compression algorithm compresses a given batch of data. The compression ratio is the name of this ratio. Let's say that 65,536 bytes are needed to store a 256x256 pixel square image. 16,384 bytes are needed to store the compressed version of the image. The compression ratio, in our estimation, is 4:1. The reduction in the quantity of data required as a percentage of the size of the original data is another way to express the compression ratio. The compression ratio determined using this method in this specific case would be 75%.

The average number of bits used to represent a single sample is another approach to report compression performance. Typically, this is referred to as the rate. For instance, in the above-mentioned compressed image, the compressed representation's average bit rate per pixel is 2. We could therefore state that the rate is 2 bits per pixel.

The reconstruction in a lossy compression is different from the original data. As a result, we need a method for measuring the difference in order to establish how effective a compression algorithm is. The distortion is the term used to describe the variation between the original and the reconstruction. Lossy compression methods are typically employed for data that started off as analog signals, such speech and video(AA Azeez, 2016). Human beings are the ultimate arbiters of quality in speech and video compression. The quality of the reconstructed waveforms is assessed using a variety of approximations of distortion because it is challenging to mathematically represent human responses.

Conclusion

Data compression is the technology term utilized in current computing technology that spotlights on the decrease in the quantity of pieces expected to represent data. Compression can increase file transfer speed, preserve storage space, and lower the cost of storage hardware and network bandwidth. In this chapter, the basic concepts and importance of advanced technological DC methods used in many real-life applications are discussed. The role and importance of data compression technology are described in detail in the first chapters, and in the second, the categories and classification of DC methods are presented. The main methods of the DC are described in detail in this section. In the section, the evaluation criteria of DC, which are more important for evaluating the performance of DC techniques, are presented. In the last part of the chapter, general and special purpose applications such as WSN, medical imaging, DNA data, signal data, databases and remote detection using DC methods are discussed.

References

AA Azeez, A. L. (2016). Empirical and statistical evaluation of the effectiveness of four lossless data compression algorithms. *Igerian Journal of Technological*, *13*(2). https://www.ajol.info/index.php/njtd/article/view/152843.

Badshah, G., Liew, S. C., Zain, J. M., & Ali, M. (2016). Watermark Compression in Medical Image Watermarking Using Lempel-Ziv-Welch (LZW) Lossless Compression Technique. *Journal of Digital Imaging*, *29*(2), 216–225. https://doi.org/10.1007/S10278-015-9822-4.

Baker, A. H., Levy, M. N., Xu, H., Nychka, D., Dennis, J. M., & Mickelson, S. A. (2014). A methodology for evaluating the impact of data compression on climate simulation data. *HPDC 2014 - Proceedings of the 23rd International Symposium on High-Performance Parallel and Distributed Computing*, 203–214. https://doi.org/10.1145/2600212.2600217.

Birajdar, A., Agarwal, H., Bolia, M., & Gupte, V. (2019). Image Compression using Run Length Encoding and its Optimisation. *2019 Global Conference for Advancement in Technology, GCAT 2019*. https://doi.org/10.1109/GCAT47503.2019.8978464.

Blelloch, G. (2013). Introduction to data compression. *Orion.Towson.Edu*. http://orion.towson.edu/~mzimand/ainternetalg/Blelloch-Compression.pdf

Brahimi, T., Khelifi, F., Laouir, F., & Kacha, A. (2022). A new, enhanced EZW image codec with subband classification. *Multimedia Systems*, *28*(1). https://doi.org/10.1007/S00530-021-00781-X.

Chowdhery, A., & Chiang, M. (2018). Model Predictive Compression for Drone Video Analytics. *2018 IEEE International Conference on Sensing, Communication and Networking, SECON Workshops 2018*, 1–5. https://doi.org/10.1109/SECONW.2018.8396351.

Cinque, L., De Agostino, S., & Lombardi, L. (2009). Practical parallel algorithms for dictionary data compression. *Proceedings - 2009 Data Compression Conference, DCC 2009*, 441. https://doi.org/10.1109/DCC.2009.84.

Díaz-Domínguez, D., & Navarro, G. (2022). *Efficient Construction of the BWT for Repetitive Text Using String Compression*. http://arxiv.org/abs/2204.05969.

Duan, L., Liu, J., Yang, W., Huang, T., & Gao, W. (2020). Video Coding for Machines: A Paradigm of Collaborative Compression and Intelligent Analytics. *IEEE Transactions on Image Processing*, *29*, 8680–8695. https://doi.org/10.1109/TIP.2020.3016485.

Dudhagara, C. R., & Patel, M. M. (2017). *A Comparative Study and Analysis of EZW and SPIHT methods for Wavelet based Image Compression*. https://doi.org/10.13005/ojcst/10.03.17.

Faria, L. N., Fonseca, L. M. G., & Costa, M. H. M. (2012). Performance evaluation of data compression systems applied to satellite imagery. *Journal of Electrical and Computer Engineering*. https://doi.org/10.1155/2012/471857.

Guo, Z., Fu, J., Feng, R., & Chen, Z. (2021). Accelerate neural image compression with channel-adaptive arithmetic coding. *Proceedings - IEEE International Symposium on Circuits and Systems*, *2021-May*. https://doi.org/10.1109/ISCAS51556.2021.9401277.

Hoang Thi Hang, N. V. V. (2022). *APPLY QMF FILTER TAPE IN DIGITAL FILTER BAND TO COMPRESS DIGITAL AUDIO SIGNALS*. UTEHY Journal of Science and Technology. http://tapchi.utehy.edu.vn/index.php/jst/article/view/495.

Jamil, S., Jalil Piran, M., & Rahman, M. (2022). *Learning-Driven Lossy Image Compression; A Comprehensive Survey*. http://arxiv.org/abs/2201.09240.

Javed, M., Nagabhushan, P., & Chaudhuri, B. B. (2018). A review on document image analysis techniques directly in the compressed domain. *Artificial Intelligence Review*, *50*(4), 539–568. https://doi.org/10.1007/S10462-017-9551-9.

Jayasankar, U., Thirumal, V., & Ponnurangam, D. (2021). A survey on data compression techniques: From the perspective of data quality, coding schemes, data type and applications. *Journal of King Saud University - Computer and Information Sciences*, *33*(2), 119–140. https://doi.org/10.1016/J.JKSUCI.2018.05.006.

Jovanovic, B. (2022). *Data Compression Explained*. Dataport. https://dataprot.net/guides/data-compression/.

Klein, S. T., & Shapira, D. (2021). Integrated encryption in dynamic arithmetic compression. *Information and Computation*, *279*, 104617. https://doi.org/10.1016/J.IC.2020.104617.

Kumar, G., Processing, R. K.-G. and S., & 2021, undefined. (2021). Analysis of Arithmetic and Huffman Compression Techniques by Using DWT-DCT. *Mecs-Press.Org*, *4*, 63–70. https://doi.org/10.5815/ijigsp.2021.04.05.

Larobina, M., & Murino, L. (2014). Medical image file formats. *Journal of Digital Imaging*, *27*(2), 200–206. https://doi.org/10.1007/S10278-013-9657-9.

Ma, S., Zhang, X., Jia, C., Zhao, Z., Wang, S., & Wang, S. (2020). Image and Video Compression with Neural Networks: A Review. *IEEE Transactions on Circuits and Systems for Video Technology*, *30*(6), 1683–1698. https://doi.org/10.1109/TCSVT.2019.2910119.

Navvabi, S., Najafi Oshnari, M., & Novikov, B. (2022). Lossless Genome Data Compression Using V-Gram. *Springer Proceedings in Physics*, *268*, 141–149. https://doi.org/10.1007/978-3-030-81119-8_15.

Nuha, H., Mohandes, M., Liu, B., & Al-Shaikhi, A. (2022). Seismic Data Compression: A Survey. *Advances in Science, Technology and Innovation*, 253–255. https://doi.org/10.1007/978-3-030-73026-0_58.

Otair, M., Abualigah, L., & Qawaqzeh, M. K. (2022). Improved near-lossless technique using the Huffman coding for enhancing the quality of image compression. *Multimedia Tools and Applications*. https://doi.org/10.1007/S11042-022-12846-8.

Platoš, J., Snášel, V., & El-Qawasmeh, E. (2008). Compression of small text files. *Advanced Engineering Informatics*, *22*(3), 410–417. https://doi.org/10.1016/J.AEI.2008.05.001.

RANDHAWA, S. S. (2022). *EEE 1857.2 LOSSLESS AUDIO COMPRESSION : A LITERATURE REVIEW*. Researchgate. https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&as_ylo=2018&q=IEEE+1857.2++linear+prediction+coding+&btnG=.

Rao, K., & Yip, P. (2018). *The transform and data compression handbook*. https://www.taylorfrancis.com/books/mono/10.1201/9781315220529/transform-data-compression-handbook-kamisetty-ramam-rao-patrick-yip.

Salomon, D. (2008). *A Concise Introduction to Data Compression*. https://doi.org/10.1007/978-1-84800-072-8.

Sayood, K. (2017). *Introduction to data compression*. https://books.google.com/books?hl=en&lr=&id=3DFHDgAAQBAJ&oi=fnd&pg=PP1&dq=how+data+compression+works&ots=gGMmmxOh74&sig=dv3DcxUHkYiElz-kpjpBxLj69Gk.

Sayood, K. (2018a). Huffman Coding. *Introduction to Data Compression*, 41–88. https://

doi.org/10.1016/B978-0-12-809474-7.00003-3.

Sayood, K. (2018b). *Text Compression - an overview* |. ScienceDirect. https://www.sciencedirect.com/topics/computer-science/text-compression.

Shin, D., & Kim, C. (2016). Data Compression Method for Reducing Sensor Data Loss and Error in Wireless Sensor Networks. *Journal of Korea Multimedia Society*, *19*(2), 360–374. https://doi.org/10.9717/KMMS.2016.19.2.360.

Thanki, R. M., & Kothari, A. (2019a). Classification in Data Compression. *Hybrid and Advanced Compression Techniques for Medical Images*, 17–29. https://doi.org/10.1007/978-3-030-12575-2_2.

Thanki, R. M., & Kothari, A. (2019b). Hybrid and advanced compression techniques for medical images. *Hybrid and Advanced Compression Techniques for Medical Images*, 1–95. https://doi.org/10.1007/978-3-030-12575-2/COVER

Wolfram, S. (2003). *Historical Notes from Stephen Wolfram's A New Kind of Science.*

About the Authors

Sabri KOCER, He is a professor and senior researcher in the Department of Computer Engineering at Konya Necmettin Erbakan University. He completed his doctorate at the Department of Electronic Computer Education at Gazi University. He has held academic positions at different Universities in Turkey and has made great contributions to the on-going technological academic development in the region. His main research interests are biomedical engineering, artificial intelligence, optimization, and machine learning and microcontroller applications. In addition, he has many academic works such as books, articles, essays, posters published in many national and international journals.

E-mail: skocer@erbakan.edu.tr, **ORCID**: 0000-0002-4849-747X

Bashir Sheikh Abdullahi JAMA, He is a computer engineer and junior researcher in the field of computer science and engineering. He is interested in conducting research that improves human life and helps to adapt to the modern technological world. His research interests include artificial intelligence, machine learning, computer vision and digital image processing. Currently, he is a PhD student in the Department of Computer Engineering at Konya Necmettin Erbakan University.

E-mail: bashiir204@gmail.com, **ORCID**: 0000-0002-2417-9223

Yasin ER, he works as a computer engineer at the Turkish State Meteorological Service. He has a master's degree from Hoca Ahmet Yesevi University. His studies are meteorological communication, enterprise-specific software development, and Windows/Linux-based system management. He is a member of various expert groups and working groups at the World Meteorological Organization, currently, He is studying for a doctorate at the Institute of Pure and Applied Science of Necmettin Erbakan University.

E-mail: yasineryasiner@gmail.com, **ORCID**: 0000-0001-6729-5278

Beyza DUNDAR, She is a senior student in Computer Engineering at KTO Karatay University. System has studied data science, artificial intelligence, image processing and natural language processing. She works on smart cities on the Capsule Technology Platform.
E-mail: dundarbeyza777@gmail.com, **ORCID**: 0000-0002-2843-1043

Similarity Index
The similarity index obtained from the plagiarism software for this book chapter is 16%.

To Cite This Chapter:
Kocer, S. , Jama, B.S.A ., Er Y. , Dundar, B. (2022). Analysing Trends and Applications in Data Compression Techniques, In S. Kocer, O. Dundar (Eds .), *Current Studies in Basic Sciences Engineering and Technology* (pp. 114–128). ISRES Publishing.

PROVIDING ANCILLARY SERVICES FROM DISTRIBUTED ENERGY RESOURCES TO ELECTRICITY DISTRIBUTION GRID

Harun KOROGLU
MEDAŞ, Konya, Turkey

Mustafa Emre SAFAK
MEDAŞ, Konya, Turkey

Mert BAYER
MEDAŞ, Konya, Turkey

Burak ALTUN
MEDAŞ, Konya, Turkey

Huseyin Oktay ALTUN
KTO Karatay University, Konya, Turkey

Ancillary Services

Ancillary services (AS) are services related to the security and maintenance of distribution systems such as frequency response stability, providing fast backup, providing reactive power, system restoration, black start etc. In conventional electric power systems where the power flows unilaterally (i.e. from transmission systems towards distribution systems), AS are usually provided by production plants connected to the transmission system. Various institutions emphasize the transmission system operator (TSO) as the acting agent when defining AS. In this context, European Network of Transmission System Operators for Electricity (ENTSO-E) defines AS as the application of assorted functions to ensure system security of the TSO (ENTOS-E, 2017).

EURELECTRIC defines AS as the services required for transmission and distribution system operators to provide the power quality, integrity and stability of the electrical power system at the specified level (EURELECTRIC, 2011). AS; supplied to its production, transmission and control systems, ensure that energy produced is transported to consumers safely and continuously (Luo, 2015).

The Federal Energy Regulatory Commission (FERC) of the United States defined AS as; "Given the obligations of the control areas, the services required to support the transmission of electrical energy from the seller to the buyer and to maintain reliable operations of the interconnected transmission system in these control areas. AS provided with production include load tracking, reactive power voltage regulation, system protection services, loss compensation services, system control, load dispatch services and energy imbalance services." It is also stated that traditionally AS are provided by generators, but they require a change in the equipment that can be used to offer variable services, to accommodate variable renewable energy resources production and to develop smart grid technologies (FERC, 2018: https://www.ferc.gov/market-oversight/guide/glossary.asp). In the Electricity Market Ancillary Services Regulation published by Republic of Turkey Energy Market Regulatory Authority (EMRA), it is defined as "the services to be provided by the relevant real/legal persons connected to the transmission system or distribution system, defined to ensure the reliable operation of the transmission or distribution system and the provision of electricity to the service under the necessary quality conditions". Supply of AS varies by country. In some countries, services are provided with an ancillary service market according to the free market conditions, while in other countries, they are compulsory and in the form of services based on legal regulations.

Types of AS

From the production of electrical energy to the end consumer, energy must be handled at the quality determined by standards and regulations. In conventional systems where the power flow has only one way which is from the transmission system to the distribution systems, AS are provided from the production facilities under the control of the TSOs in order to keep the voltage and frequency parameters in certain specified limits. AS can be classified roughly as in Figure 1.

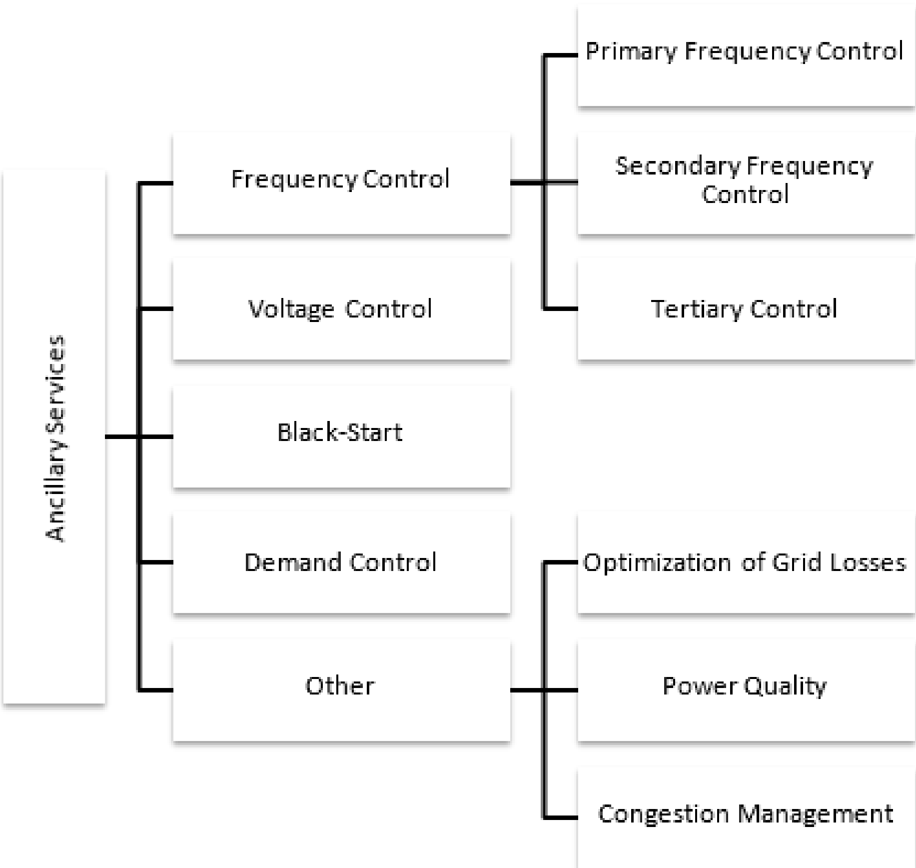


Figure 1. Classification of AS

Frequency Control:

The frequency is controlled by TSOs. TSOs are in charge of maintaining a precise balance between electricity production and consumption, which is necessary for reliable grid operation. By regulating the quantity of active power produced and/or consumed, the supply-demand balance of electric energy should be maintained in the electrical power system's 200 mHz band. A fixed percentage of the active power reserve is used for frequency regulation. In cases where the system frequency drops below the determined value, frequency control backup is used. Frequency control is divided into three categories:

- Primary frequency control,
- Secondary frequency control,
- Tertiary frequency control.

With the oscillation in the frequency of the power system, primary frequency controls are activated first. Primary control reserves in the frequency swing ± 200 mHz band are active and keep the oscillations intact. Then, secondary frequency control is activated, bringing the system frequency to nominal value. Tertiary frequency control generally provides production and consumption balance with load throw-on and load throw-off instructions. Similarly to secondary control, it supports secondary reserves with tertiary frequency control and indirectly makes primary control reserves suitable (OpHB-Team, 2009).

The response time of the system frequency resources used for the restoration of European countries and in Turkey is as follows:

- Interaction of primary frequency control sources within 30 seconds and maintaining at least 15 minutes,
- The secondary frequency control reserves should be activated within 30 seconds and increase the frequency value to the nominal level within 15 minutes,
- Tertiary frequency reserves are also expected to increase to maximum power within 15 minutes.

Voltage Control

It is a mandatory service applied to keep the voltage of the power system within the certain operating voltage limits. Voltage control in the electrical power system is largely carried out by reactive power control. In order for the reactive power demanded by the grid to be transmitted over long distances, the locations of the reactive power generation resources throughout the grid should be set and compensation investments should be made at appropriate locations. The regulation of the voltage between the determined limit values in the entire electrical power system is carried out with a three-level hierarchical system:

- Primary voltage control,

- Secondary voltage control,
- Tertiary tension control.

Primary voltage control is carried out with controllable devices such as Automatic Voltage Regulators (AVR) and Static VAR compensators. The primary side of voltage control is an automatic, local control that brings the voltage value in a given feeder to the set value. Secondary side of voltage control is a central, and again, automatic control system used by regulatory agencies to manage reactive power flows locally. Tertiary voltage control optimizes the reactive power flow in the grid manually. Based on the relationship between voltage and reactive power, equipment capable of producing and consuming reactive power is included in the tertiary control. Equipment capable of producing and consuming reactive power are included in the tertiary control. (2022, URL-2).

The reactive power needed for voltage control is provided locally at the points of requirement. Reactive power; generators from dynamic sources are provided from systems such as synchronous compensators (STATCOM). In addition, capacitor banks from static sources are provided from static voltage controllers and flexible alternating current transmission systems (FACTS) (Hung, 2011).

Black – Start

Black-start refers to re-energization processes in case the electric power system collapses in a wide area or completely after the disruptive effect. After re-energization actions are ensured by system stability, the customers are fed with the production facilities that can be commissioned without the need of an external energy source, the power system is energized and the other production facilities are commissioned. Restoration of the power system which means re-energization from system collapse is carried out with frequency stabilization and re-synchronization stages.

Demand Control

Agreements between the system operator (SO) and consumers are based on participation in frequency control by reducing the demand for electrical energy. Demand control is carried out by separating some of the burdens of consumers from the grid. In cases where the energy supply is insufficient or the demand is high, instantaneous demands are created that can be used in frequency control. It aims to participate in demand control with a large strong industrial facility.

In the United Kingdom, a maximum period of 30 minutes has been determined in demand control agreements and this period has been determined as 15 minutes in the Electricity Market AS Regulation published by EMRA. Similarly, in the UK, the demand control reserve must be at least 3 MW. In Turkey, while there is no lower limit value, it is stated that the offers of the consumers who will participate in instant demand control should be in the form of 1 MW and its multiples.

Instant demand control is realized with low frequency relays. The setting value of the low frequency relay is based on the automatic reduction of the demand from 10% to 20% if the system frequency drops to 49.0 Hz. It is suitable for participating in the instant demand control of the cement and iron-steel industries. Consumers' demand for electrical energy can be reduced to another time period or the demanded energy can be supplied from the generator or embedded production facilities owned by the facility behind the point of measurement.

In coming years, as the renewable-based distributed resources integration in electrical power systems increases, demand control seems to be increased consumption as well.

In the statement made by European Power Exchange (EPEX), which is one of spot market managers in Europe, it was reported that the prices of electrical energy units in Germany were negative on 24 and 26 December 2017 (2022, URL-1). It is likely that excess production will occur due to seasonal and climatic characteristics, especially during long periods such as night, weekend and holidays (EURELECTRIC, 2004).

Other Ancillary Service Types: In the above titles, the types of AS that have been adopted in almost all energy markets are given. Other types of AS in different world applications are as follows:

Optimization of Grid Losses

In electrical power systems, some energy is lost due to the heating of the conductors and transformer windings. Countries like Austria and Belgium expect these losses to be compensated by the SO. Due to that, special market mechanisms have been created to dispose of the linked loss.

Power Quality

The electric power system is subject to several disruptive effects in terms of power quality. Power quality issues are caused by variations in the primary dimensions of the power system—current, voltage, and frequency—that are outside the permissible limit values. These variations lead to consumer electronics failure or breakdown. The provision of reactive power services, frequency control, voltage control, and limit values established by voltage and frequency standards would all improve power quality characteristics within the ambit of AS in electrical power systems.(Delkhooni, 2019).

Congestion Management

With the rising prevalence of renewable energy sources, AS will be provided in the long term, from the distribution level to grid connected production. In order to provide AS at the distribution level, studies have been carried out on a project basis against the lack of a definite framework on how to coordinate between transmission and distribution SOs (TSO and DSO). In order to benefit from the flexibility that can be provided by the generation facilities in the power system, it is important to establish a legislative infrastructure that will ensure the collaboration between TSO and DSO. Thus, congestion situations can be overcome with strong communication and cooperation between SOs and by making more use of system flexibility (Zipf, 2016).

Resources in Distribution Grid for AS

With the spread of Distributed Energy Resources (DER), the current roles of transmission and distribution systems, Transmission System Operator (TSO) and Distribution System Operator (DSO) are changing day by day. In conventional electrical systems;

- The transmission grid is a system with high voltage (154 kV, 380 kV, etc.), interconnected with parallel lines and bidirectional energy flows, where large production facilities and large loads are directly connected.
- Distribution grids are characterized by one-way energy flow, medium and low voltage (31.5 kV, 15.8 kV, etc.), radial, from connection point in transmission grid to loads, almost any production facility (very limited production facilities compared to all production facilities) is not directly connected and loads are only a statistically predictable system.
- Only large power plants were allowed to participate in wholesale markets, and these provided frequency and voltage control AS to the system (Rebours, 2007).

Renewable production sources with inverter systems such as solar and wind connected to the distribution grid, storage systems, electric vehicles, smart meter systems, etc. new technologies change the architecture and operation of the power system.

With the increase of renewable-based distributed production facilities connected on the distribution level, utilization of AS in the distribution grid comes to the agenda. The increase in the prevalence of distributed generation facilities in the distribution grid has significant effects on the concepts of operation, control, protection and reliability in existing power systems. Active and reactive power flows cease to be unidirectional. The energy source non continuance (cloud transitions, sudden changes in wind speed) in RES based production facilities such as wind and solar causes sudden changes in power systems. In addition, the operation of the distributed generation facilities with a constant power factor independent of the grid operating conditions causes a voltage increase at the connection point. In cases where the demand for electrical energy changes, constant active power supply to the distribution system causes increased line losses and reverse power flows from the distribution system to the transmission. Therefore, it is aimed at reshaping the ancillary service markets and providing AS from the distributed production sources of both transmission and distribution system operators; studies for determining pre-qualification, procurement, activation and settlement processes and TSO-DSO coordination are ongoing.

Below, the sources in the distribution that can be used for the AS in the distribution network are examined.

Distributed Energy Resources

Despite the advantages of conventional electrical energy systems, the trend towards distributed production resources is increasing for the following reasons:

- To meet the increasing demand for electrical energy, more fossil fuels are needed to meet the demand in conventional energy systems.

- To meet the increasing demand, adopt the concept of consume where you produce (prosumer) due to problems such as the existing systems are not in the desired reliability and cause technical and economic losses encountered in the transmission of the produced energy to the consumed region.
- Many countries prefer renewable energy sources by reducing the use of fossil fuels that cause environmental pollution according to the Kyoto Protocol. Thus, it is desired to prevent climate change and to prevent global warming problems by reducing greenhouse gas emissions

Wind, hydraulic, solar, biomass, geothermal and cogeneration energy systems which are among renewable energy sources are preferred due to some of the mentioned problems of conventional production systems. In the coming period, it is expected that distributed generation plants based on renewable energy sources such as solar, river hydro, geothermal and wind will become more widespread. With the widespread use of distributed generation plants in Turkey it is expected to reduce the investment costs for the transport of electricity generated from east to west. In addition, it is predicted that technical losses caused by long lines will decrease. The following table shows the ratio of installed power distributed generation plants connected to the distribution grid in Turkey, according to the February 2020 report shows EMRA sector.

Table 1. Distributed generation installed power plants in Turkey

Resource Type	Installed Power (MW)	Rate (%)
Solar	5,883.41	92.36
Natural Gas	331.18	5.20
Biomass	75.67	1.19
Wind	70.83	1.11
Hydraulic	8.65	0.14
The Overall Total	**6,369.74**	**100.00**

In addition, with the developing technologies, micro grid systems are expected to be developed by using storage systems, with the spread of distributed production plants and cost reduction. Rural areas supply with micro-grid, especially long feeders in distribution regions;

- Failure to be not subjected to long-term interruptions in failures (increasing energy supply continuity)
- It will be beneficial to reduce the distribution grid investment and operating costs to be made in the region.

Reactive power support can be obtained from storage systems using solar inverters, solar power plants, and some wind power plants.

With the increase in grid integration of renewable production sources, it will have the potential to provide AS to the grid. Many of the AS that are currently given by synchronous generators and voltage regulators might potentially be provided by inverters placed with distributed generation systems and rooftop solar systems. These services include harmonic reduction, voltage regulation, flicker control, reactive power compensation, and active power filtering. Wind turbines and variable speed generators have the ability to provide the grid more.

Energy Storage Systems

Energy storage systems are an important component in providing flexibility and supporting renewable energy integration in the energy system. It can balance central and distributed electricity production, as well as contribute to supply security. It will have a complementary effect on energy storage, demand-side management, flexible production and grid development. Energy storage can also contribute to the de-carbonization of other economic sectors and support the increased share of the integration of variable renewable energy in transport, building or industry. The role of the energy storage area with electricity, which technologies and innovative solutions seem more suitable for different purposes and possible policy approaches should be evaluated within the framework of clean energy (Kling, 2000).

Storage technologies have started to appear as more cost effective, wide-ranging and very profitable grid participants, considering today's technologies. Below are some energy storage systems.

- Chemical Energy Storage Systems
    - Lithium-ion (li-ion) batteries
        - Lithium cobalt oxide
        - Lithium manganese oxide
        - Lithium iron phosphate
        - Lithium titanate
    - Redox flow batteries
        - Vanadium redox batteries
        - Zinc bromine batteries
    - Lead batteries
    - Nickel cadmium batteries
    - Sodium sulfide batteries
    - Sodium nickel chloride batteries
    - Metal-air batteries
- Ultra / Super Capacitors
- Compressed Air Energy Storage
- Pumped Hydro Storage
- Flywheel

Storage systems can perform many functions such as frequency control, voltage control, power balancing, commercial and improving grid power quality at the same time. This makes it difficult to answer the question of how to define storage systems in the grid. While defining energy storage systems;

- Firstly, the target application should be determined (AS, investment deferral, energy arbitrage etc.),
- Optimum power value and optimum capacity value should be determined according to the application determined. Optimum size and location can be determined with algorithms that will work according to the final purpose.
- After correct sizing, the connection level to the grid should be selected. The usage factor should be determined according to the battery technology chosen.
- The application type should be determined according to the primary, secondary and tertiary applications according to the application priority. For example, primary application is primary frequency control, secondary application is energy arbitrage, etc. priorities can be determined.

Unlike other equipment used in the electricity grid, energy storage systems have a wide variety of technologies and capabilities. Unlike other equipment used in the power grid, energy storage systems have a wide variety of technologies and capabilities. For this reason, storage technology is selected according to the application and location of the application. Battery features seem to vary according to the battery type to be selected. Battery features:

Energy capacity: The energy capacity of a battery is the total energy obtained by a full discharge, expressed in an ampere-hour (Ah) or watt-hour (Wh). Energy available from a battery is a function of discharge rate and temperature. Therefore, the capacity of the battery is usually indicated on the label next to these conditions. For many battery technologies, the energy capacity decreases over time and usage. Also, while high temperatures accelerate capacity reduction, the average charging status is also effective for capacity reduction for some technologies.

Energy density: Energy density (or specific energy) defines how much energy a battery stores (at a certain temperature and discharge rate) per mass or volume unit.

Power capacity: Power is the speed of energy transfer. The ability of a battery to charge or discharge power generally varies depending on the state of charge (SOC) and temperature. Charge rate (C-Rate) values are used to compare the charge and discharge power capabilities of the batteries. C-Rate is the normalization of the power output / input according to the energy capacity of the battery. For example, a 10kWh battery with 5 kW charge / discharge is known to charge / discharge at C2 or 0.5C. A 10kWh battery with 20kW charge / discharge is known to charge / discharge at C0.5 or 2C. Batteries with high C-Rate can charge or discharge their capacity quickly.

Power density: Power density (or specific power) defines how much power a battery can deliver or supply (at specific temperature and SOC level) per unit of mass or volume.

Response time: The time it takes for a battery system to respond to commands.
Ramp rate: It is the rate of increase or decrease of the output power of the battery system.
Cycle life: Cycle life defines the ability of a battery to withstand repeated charge and discharge.The cycle life definition is usually provided by lead acid battery manufacturers based on a specific temperature and discharge depth (DoD). For example, a 10 kWh battery with 3500 cycle life at 40% DoD is expected to have 1400 cycle life at 100% DoD. However, it should be noted that the conversion rates rarely occur in applications.
Shelf life: Shelf life indicates the time to reach the end of the battery's life if stored under ideal conditions. Therefore, this value represents the upper limit of the lifetime that can be achieved in the field. It should be noted that the ideal conditions for batteries to charge are significantly different between technologies.
Round-trip efficiency: It is the ratio of energies entering and exiting the battery during a complete charge-discharge cycle. It is a useful parameter to learn the expected loss during energy transfer.
Self-discharge rate: Self-discharge rate defines the discharge rate of the battery in the no-load state. It is usually expressed in x% per month. For example, a 10kWh battery with a 5% self-discharge rate is expected to lose 5% SOC or 0.5 kWh for a month while in standby. Self-discharge typically increases with increasing ambient temperature.
Temperature range: Batteries can be damaged or become unsafe when certain temperatures are reached. Therefore, a temperature range in which the battery has to be operated is usually determined by the manufacturer.
Voltage range: Some battery technologies are sensitive to overcharge or over discharge. As a result of this situation, manufacturers often determine a voltage range in which the battery should be present.

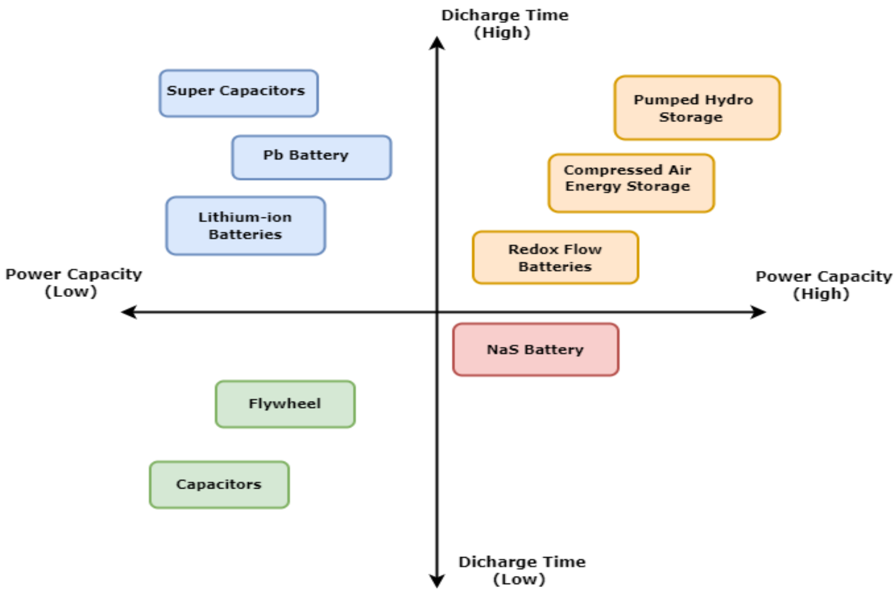


Figure 2. Energy storage systems technologies and characteristics

Participation in AS Capabilities of Storage Systems

Some storage technologies are excellent frequency and voltage regulation providers because the zero net energy source matches the zero net energy source. The fast response and sensitive control offered by storage outperform the control capabilities of many conventional production plants. Technologies that can provide repeated high-speed storage capacity without degradation in their performance are the most suitable frequency and voltage regulation providers.

In the EU countries, the need for fast and flexible storage systems will increase further in the near future, as base-loaded power plants (e.g. coal-fired power plants or nuclear power plants) are increasingly replaced by renewable production sources (e.g. wind or photovoltaic). For example, WWF Solar GmbH, located in Eberswalde, Germany, which integrates renewable energy sources into the grid, started to build a storage system to provide secondary balancing capacity using Vanadium Redox Flow storage systems (2018, The Independent, http://www.independent.co.uk/news/business/news/germany-energy-consumer-christmas-paidsupply- demand-outstrip-renewables-wind-solar-a8129716.html). In addition, a multi-purpose power station is being built in AK Berlin-Schönefel to provide AS to a 3,800 kWp photovoltaic system, the TSO of the German utility area in the high-voltage grid. For this purpose, 25 Vanadium Redox Flow Storage Systems with 12,500 kW power and 50,000 kWh capacity are used.

Another energy storage application was used in the black start application. The 33MW / 20MWh lithium-ion battery energy storage system (BESS), which provides grid stability in daily use and helps smooth out output from local renewable power supplies, has been used to launch 44 MW combined cycle natural gas at El Centro Production Station in Valley, in California (Divya, 2009).

Energy storage systems provide the following services in terms of AS participation in electric power systems:

- Load following,
- Regulation,
- Frequency response,
- Reserve capacity,
- Voltage support,
- Black start.

Electric Vehicles

The use of electric vehicles by the day as well as throughout the world is increasing in Turkey. In addition to adding a new load to distribution and transmission grids, the widespread use of electric vehicles can be evaluated in the context of grid reliability and load management and energy quality as they are used as storage sources (Saygın, 2019). The Electric Vehicles' (EV) batteries are designed in both directions so that EVs can support AS in the grid. In this way, it will be possible to transfer energy from vehicle to grid (V2G) or from grid to vehicle (G2V).

In order to use electric vehicles in AS, the distribution grid manager must keep the number of EVs connected to the grid under control. In the event that the number of vehicles connected to the grid and the occupancy rate of the vehicle batteries is sufficient, the signal to be supplied will be used to start the reference part (this signal is an example of the sinusoidal signal in the grid) and to provide AS to the grid, since the voltage signal will include magnitude and phase shift.

Prevalence of Electric Vehicles

A robust charging infrastructure must be built in order to expand EV use. Since EVs represent the mobile demand for electricity, they theoretically have the ability to draw power from the grid at any time of the day, whether at home or at work, at low charge rates or by using the public charging infrastructure for fast charging. The amount, location, and timing of the electrical energy that electric vehicles (EVs) receive from the grid can vary greatly and are influenced by both driving habits and charging-related business models.

It will be necessary to coordinate the charging models with digital technologies (smart charging). With smart charging, it can be encouraged to charge vehicles with low cost, renewable electricity generation over price and control signals or when there is no congestion in the grid. The value of smart charging applications will increase if it provides support to the grid to supply power quality and flexibility. Beyond smart charging, more flexibility can also be carried out through vehicle-to-grid (V2G) technologies that enable two-way charging.

Where and when the EVs are charged is also critical for distribution electricity companies. EV's can be charged through "behind the meter" systems. Thus, the additional load is provided by rooftop PV systems and/or batteries, and charging is optimized to take place during the day at home or at work.

The value of smart charging applications will increase if it provides support to the grid to supply power quality and flexibility. Beyond smart charging, more flexibility can also be implemented through vehicle-to-grid (V2G) technologies that enable two-way charging (Sortomme, 2012).

Shura "Transformation of Turkey's Transportation Sector: Effects on Turkey Distribution Grid of Electric Vehicle" as stated in the report two different electric vehicle market scenarios were investigated. According to the "High Growth" scenario, electric vehicles will constitute 10% of the total passenger vehicles by 2030 and the number of electric vehicles is expected to be 2.5 million. The "Medium Growth" scenario, on the other hand, predicts that there will be around 1 million electric vehicles in the same year. When the average distance traveled per vehicle is 10,000 km per year and 1 kWh charge is equivalent to 6 km travel, the impact of the total country electricity demand in 2030 will be 4.2 TWh per year according to the high growth scenario and 1.7 TWh according to the medium growth scenario (Deutsche Energie-Agentur GmbH, 2018).

The main effect of electric vehicles will be on the load management on the LV side of the distribution grid. Electricity distribution companies are expected to apply smart grid elements correctly, manage demand at micro level, and eliminate the effect of electric vehicles increasing transformer points with the help of rooftop solar panels and storage modules. In addition, electricity distribution companies; in the "prosumer" category to be formed, 1 million meters / customers, LV grid control, smart meter, big data management, communication and information infrastructure, cyber security will have to adapt their infrastructures accordingly, in other words, they will have to plan and implement the smart grid elements (Gustavsson, 2013).

Vehicle to Grid (V2G) Concept

Electric vehicles could support the grid, particularly for load balancing and spinning reserves AS. Electric cars have the capacity to discharge power back into the grid in both directions, serving as distributed energy storage. It can deliver power quickly, which can be utilized as rotating backups to maintain grid stability as intermittent production sources like electric vehicles, wind turbines, and solar power are increasingly utilized. Although they are not yet commonly used, technologies that use electric vehicles. In addition to taking off energy from the grid while charging, electric vehicles can also be considered as a mobile storage component that can power the grid. Thus, it may provide different balancing services or AS to the grid operator. In addition, it can act as a storage element that meets the demand for housing when necessary. Although the scale of benefits provided by the flexibility provided for the grid is high, technological constraints still remain. EV battery life will decrease very quickly due to the widespread use of V2G applications and the number of completed cycles associated with charge-discharg

Demand Side Management

Demand side participation is one of the increasingly important methods in developed markets. In generally, it can be defined as the reaction of customers, who can be flexible to the price. However, with the passing of the distribution grid to bidirectional power flow, the increase in the prevalence of the facilities producing energy based on renewable energy sources increases the importance of demand side management. Situations where electricity supply will be insufficient are more likely to occur than conventional grid structure. For this reason, it is expressed as demand side management that the demand for electrical energy is shifted to a different timeframe or disposed of completely against the reduction of the production amount to be provided from RES based production facilities (Delfanti, 2014). Demand side management is carried out with the following functions:

- Peak Demand Decrease
- Low Demand Time Filling

- Load Shifting
- Demand Savings
- Strategic Demand Growth
- Flexible Load Shaping

With the demand-side management, the expansion of power plants based on renewable energy sources (wind, solar, etc.) and the change in electricity consumption habits brought the necessity of making the electricity grids more and more smart. Energy storage systems, ventilation, air conditioning, lighting equipment, renewable production sources, and DSOs on how to use them all will play a major role. Demand side management should provide information such as when electricity is used and producer-consumer incentives for the spread of management. Electricity tariffs should provide information such as when electricity is used and producer-consumer incentives for the prevelance of demand side management. As a result, both producers and consumers will want to participate in demand-side management with incentives.

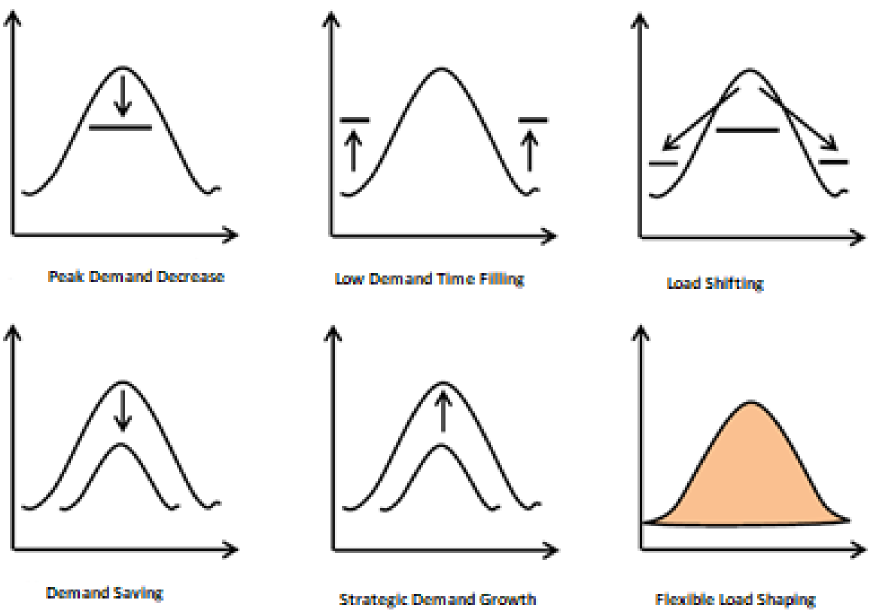


Figure 3: Demand Side Management Functions

Demand side management is carried out directly and indirectly in two different ways. The methods indicated in the figure below are detailed (Yumak, 2016).

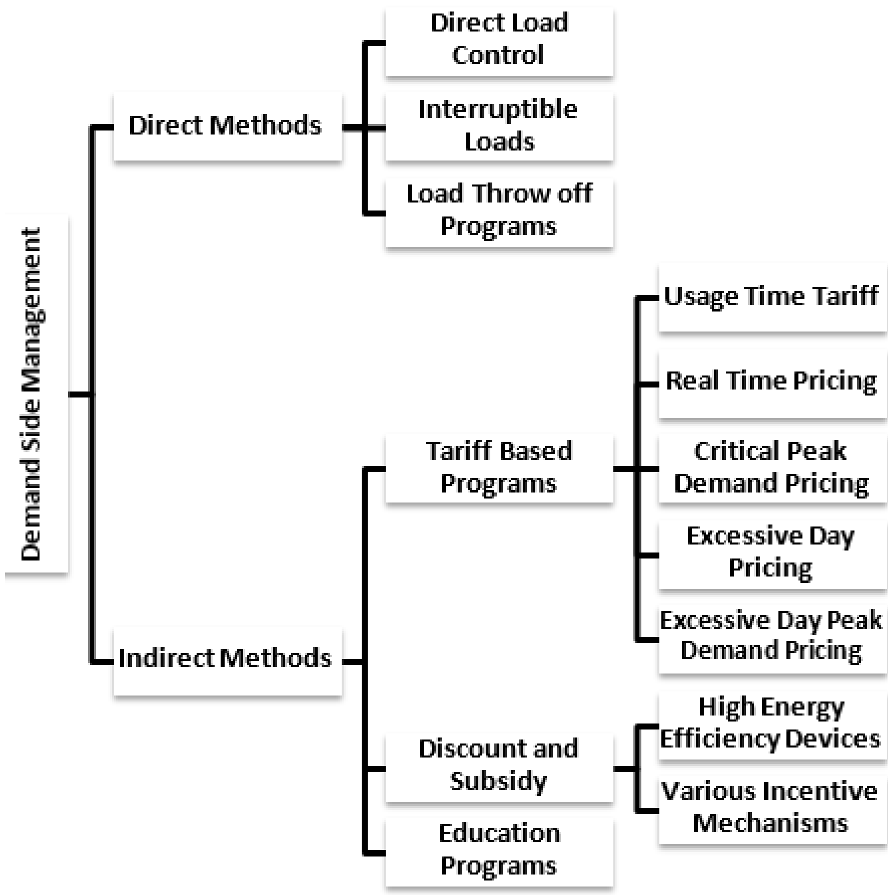


Figure 4. Demand Side Management Methods

Direct load control is the remote control of end consumer equipment during peak demand times and critical operating situations with the development of the smart grid concept. Water heater, pool pump, air conditioning systems used by residential consumers are the sources used in direct load control. Interruptible loads are reduced by the SO with demand limiters and signals in exchange for a certain pricing. In load throw off programs, consumers are offered payment or discounted tariffs to reduce peak demand for a certain period of time. It is usually carried out with large industrial consumers (ENTSO-e, 2015).

Indirect methods are evaluated under the sub-headings of tariff-based, discounts, subsidies and education programs. Tariff-based programs are briefly explained below:

- Usage Time Tariff: It is a fixed price tariff structure in different time periods, months or seasons.
- Real Time Pricing: Energy prices change on an hourly basis and reflect energy supply costs more accurately.
- Critical Peak Demand Pricing: Pre-determined higher energy price application for a few days or hours per year, taking into account time-of-use tariff and real-time pricing mechanisms.
- Excessive Day Peak Demand Pricing: A situation where higher prices are valid

for 24 hours, similar to critical peak pricing.

- Excessive Day Critical Peak Demand Pricing: Critical peak demand pricing conditions are valid throughout the day.

By enabling consumers to follow instant consumption values, they can shift their consumption to more favorable tariff times and / or prefer devices that operate more efficiently, thus making them pay less. In addition, under favour of the load management practices that voluntarily participate in the consumers with the incentives, the investments budgets can be reduced by using the distribution grid more efficiently. When evaluated from the perspective of the SO, demand side management is important to prevent line loadings and control technical loss parameters in congestion situations.

Evaluation of AS in Distribution Grid

AS at the transmission system level are used to balance the system and manage energy fluctuations. Services with similar features can be used for the management of local system difficulties at the distribution system level with the increase of distributed production resources connected to the distribution grid. At the TSO level, these services can be classified as frequency control, voltage control and other AS.

The table below shows that TSOs and DSOs can make AS contracts in Austria, Belgium, Denmark, Finland, Norway, Spain, Turkey (Labato Miguelez, 2008).

Table 2. Current Status Of AS Contracting At Distribution System Level

Contracting AS from DER directly connected to the DSO-grid	Austria	Belgium	Denmark	Finland	Spain	Norway	Turkey
TSO contracting AS from DER directly connected to the DSO grid	Yes	Yes	Yes	Yes	Yes	Yes	No
DSO contracting local services (for own purposes) from DER directly connected to the DSO grid	No	No	No	No	No	No	No
DSO contracting AS on TSO's behalf	No	No	No	No	No	No	No

In general, there are several restrictions on the participation of DER (distributed energy resource) in AS. In many countries today, the diversity of DER participation in AS markets is large, but the contracted volume is still marginal. It is clear, in addition to that DER has increased participation in AS markets.

Local needs for AS in distribution systems must be combined with system needs in balancing and congestion management. Sources in distribution systems such as demand management, storage systems and distributed production are expected to participate in AS for competitive AS markets for both the local and the entire power system (Singh, 1999).

Frequency Control

Primary frequency control (PFK) is the local automatic control method, it is based on adjusting the active power output of generators, especially in cases where a large power generation facility or load is cut out. Keeping the frequency at a stable point is inevitable for the power system. In generators working synchronously with each other in the power system, primary frequency control is performed automatically with speed regulators. Secondary frequency control (SFK) is the central automatic control method. It is carried out by sending automatic generation set value (set-point) to the plants through Turkey Electricity Transmission System (TEIAS-TSO) National Load Dispatch Center (NLDC) automatic generation control (AGC) software.

Increasing asynchronous resources included in the grid with new technologies will cause the synchronous conventional generators to be economically disabled. Asynchronous plants are examples of inverter-controlled renewable energy sources, energy storage technologies.

High-performance, swiftly responding inverter-controlled facilities are anticipated to have their own service category. Comparing these technologies to conventional synchronous production reveals a wide For example;

- Systems for battery storage and inverter-based production do not add much,
- Power output from solar and wind energy generation is unpredictable and only available intermittently.

However, the frequency response from a battery storage system, solar power or wind power plant is controlled by the inverter and has the potential to react faster than conventional generators. Increasing asynchronous resources included in the grid with new technologies will cause the synchronous conventional generators to be economically disabled. Asynchronous sources are connected to the system via the inverter and the system causes the Inertial Response capability potential to decrease. The current frequency control AS philosophy set is not specifically designed to eliminate this challenge (Aubut, 2013).

Other inverter-based sources including wind turbine generators (WTG) may also be able to inject active power into the system following a malfunction caused by the control system (e.g. generator failure). Many WTG producers already have access to this so-called synthetic inertia. The response of a wind power plant in Hydro Quebec with a synthetic inertia capacity is depicted in the image below (Figure 5). Without wind energy reduction before failure, this feature is possible. However, a wind turbine will give a frequency improvement period as illustrated in the figure below (Lalor, 2005).

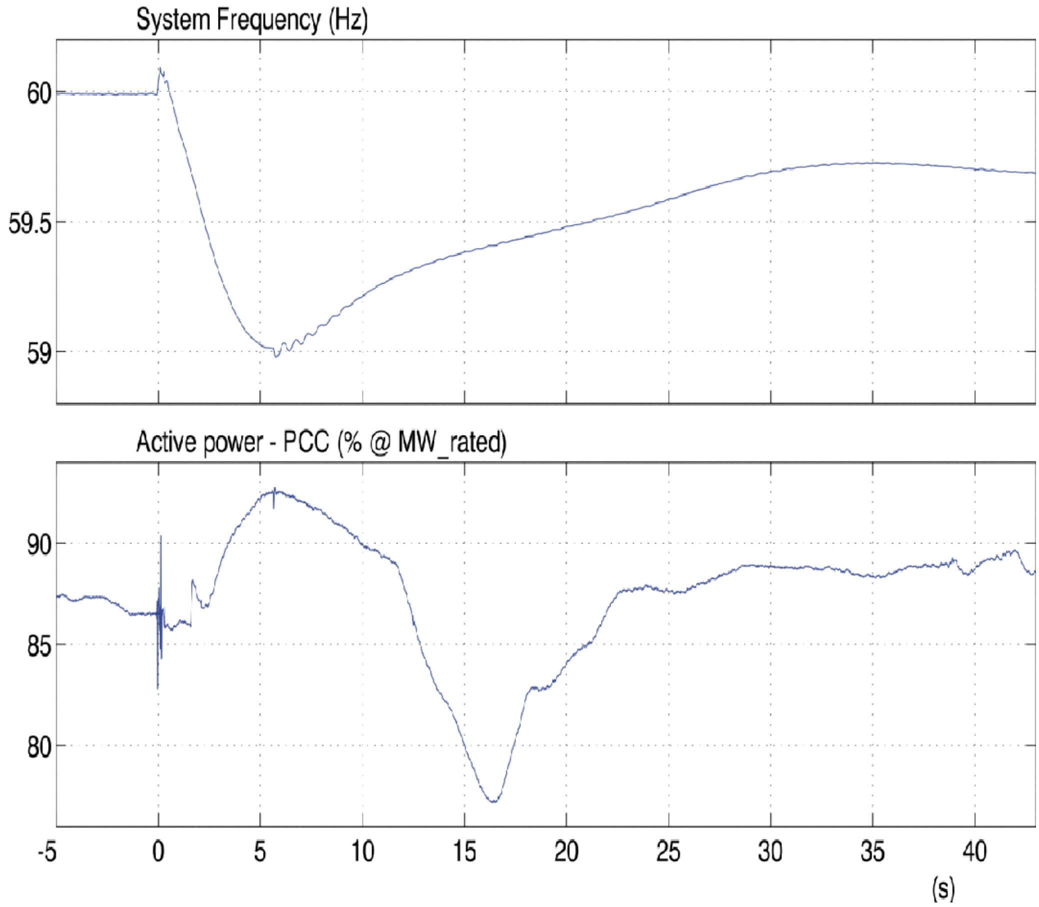


Figure 5. Wind Power Plant Synthetic Inertia Response

All synchronous machines in the circuit perform Synchronous Inertial Response after power system failures (SIR). The ability for synthetic inertia necessitates control action in response to a decrease in system frequency. Because of this, synchronous machines do not have the same inertia or natural response as synchronous machines in the event of generator failure. Synthetic inertia, on the other hand, will enhance the Rate of Change of Frequency (RoCoF) and aid in preventing system deterioration. The performance of synthetic inertia needs to be examined and explored.

Performance requirements related to primary frequency control ancillary service are defined in the Electricity Market Regulatory Authority Electricity Grid Regulation Annex 17, AS performance test procedures, Annex 17.A primary frequency control performance test procedures. The minimum performance criteria defined in the primary frequency control test procedure are summarized in the table below.

Table 3. Primary Frequency Control Performance Requirements

Requirements	Value
Reserve Capacity	Minimum ± 1 MW
Sensitivity	Minimum ±10 mHz
Response Time	Maximum 2 seconds

Ramp Rate	Minimum Reserve / 30 seconds
Duration	Minimum 15 minutes
Tolerance	± 1% of Tolerance Installed Power
Success Rate	Minimum 90%

When the requirements stated in the table above are evaluated in terms of storage facilities, other requirements (reserve capacity, sensitivity etc.) are related to the inverter, except for the duration. Depending on the duration, the charge / discharge capacity of the storage facilities is determined.

In addition, according to the current regulation, storage facilities should have at least ± 1 MW and their multiples (such as ± 2 MW, ± 3 MW, not ± 1.5 MW, ± 2.5 MW). As the maintenance time is at least 15 minutes, the charge / discharge capacity should be at least 250 kWh. As a sensitivity, inverters should be able to react to changes in mains frequency at least 10 mHz and start to react within a maximum of 2 seconds. The responses should be provided within the tolerance limits with a success rate of at least 90%. For storage facilities, tolerance limits must be redefined depending on reserve capacity. For example, it can be defined as ± 10% of reserve capacity.

Performance requirements related to secondary frequency control ancillary service are defined in the Energy Market Regulatory Authority Electricity Grid Regulation Annex 17 Ancillary Services Performance Test Procedures ANNEX 17.B Secondary Frequency Control Performance Test Procedures. The minimum performance criteria defined in the secondary frequency control test procedure are summarized in the table below.

Table 4. Secondary Frequency Control Performance Requirements

Requirements	Value
Reserve Capacity	Minimum ± 5 MW
Sensitivity	-
Response Time	Maximum 30 seconds
Ramp Rate	Minimum Reserve / 5 minutes
Duration	Minimum 1 hour
Tolerance	± 1% of Tolerance Installed Power
Success Rate	Minimum 90%

49 MW lithium-ion energy storage systems have been installed at coal and gas fired power plants in Roosecote, England (Energy UK, 2017). The purpose of installing the storage system is to maintain fluctuations in electricity demand, power frequency and grid stability. The system responds to fluctuations in less than a second (UK, 2017).

German utility WEMAG has decided to increase the energy storage capacity in Schwerin from 5MW to 10MW power output and from 5MWh to 14.5MWh capacity. Currently, there is no storage facility installed in size and capacity that can participate in frequency control in Turkey. However, within the scope of the R&D project accepted by EMRA and carried out by Meram Electricity Distribution Company will establish a 1 MW chemical energy energy storage facility for primary frequency control ancillary service and congestion management applications in the coming days.

The power value, capacity value, storage size and location related to the design of storage facilities in MV grid should be determined. The location focuses mainly on two different architectures: central energy storage architecture and distributed energy storage architecture. High-capacity storage systems with centralized power storage architecture Turkey Electricity Transmission Company (TEIAS) is connected to the first distribution centers on feeders coming from substations. With the distributed energy storage architecture, multiple low capacity energy storage facilities are preferred over high capacity storage facilities. Low capacity energy storage facilities are connected to such as distribution centers, circuit breaker measurement cabinets within the grid and transformer centers (Koohi-Kamali, 2013).

Voltage Control

At all voltage levels in the electrical power system, the voltage value must remain within the specified operating range. Over-active power generation / consumption at low and medium voltage levels may cause voltage increase / decrease and technical loss rates increase.

If there is no production in the local area, the voltage decreases as the distance increases until the electrical energy is delivered to the consumer. When the amount of distributed production is high at the end of a line, and there is no load to use the entire production, power flows occur in the opposite direction (from distribution to transmission). As a result, the voltage can rise to very high levels or lower the voltage limits. DSOs try to minimize voltage drops (<1 pu) with compensation investments. However, it does not affect the voltage increase (> 1 pu) too much.

When the number and capacity of distributed energy resources increase, voltage change problems will increase in the distribution grid. Thus, it is necessary to control the voltage to compensate the effects of resources on the distribution grid, first of all, the reactive power properties of inverter controlled power plants should be used and then other options should be checked. With the control of reactive power, compensation investments made annually will decrease significantly. In addition, technical losses will be reduced and energy efficiency will be increased by controlling the voltage.

For the reactive power characteristics of the plants are used, the active power value at the exit point of the feeder can approach zero due to the active power generation of the plant. In such cases, legal regulations should not constitute any penal situation. Furthermore, for the distribution grid operating conditions have changed, the active powers of the plants must also be controlled by the distribution grid operator.

Table 5. Reactive Power Control Performance Requirements

Requirements	Value
Reserve Capacity	± 0.9 power factor
Sensitivity	Minimum 0,01 pu
Response Time	Maximum 1 seconds
Ramp Rate	Minimum Reserve / 30 seconds
Duration	Continuous
Tolerance	± 10% of the reserve
Success Rate	Minimum 90%

As a reserve capacity requirement, it is recommended that the reactive capacity, which is determined by taking into account the generation facilities, inverter installed powers and ± 0.9 power factor, is determined as the minimum mandatory requirement. For the 0.8 power factor, fixed payment business model can be developed. As a sensitivity, it can be expected to react to at least 1% changes in the grid voltage (0.01 pu). As a response time, it can be expected to start reacting within 1 second at the latest and provide the entire reserve within 30 seconds. If the reactive power control ancillary service is to be provided autonomously, these times can be reduced to a few seconds depending on the inverter capacity. As of the end of 2019, there are 1114 solar power plants connected to the Meram Electricity Distribution Company grid. In the table below, the reactive power control potential for Meram Electricity Distribution Company is shared. In inverter controlled sources, it is assumed that facility power is equal to inverter power.

Table 6. Meram Electricity Distribution Company Reactive Power Potential

Resource Type	Number of Plants (Number)	Plant Power (MW)	Power Factor	Reactive Power Potential (MVAR)	Power Factor	Reactive Power Potential (MVAr)
Solar Power System	1114	920,043	±0,80	690,032	±0,90	445,597
Other	8	6,598	±0,80	4,948	±0,90	3,196

When the field type solar power plant systems in the Meram Electricity Distribution Company region is examined, 1250 kVA transformers and inverters with a total power of 1 MW are used in the typical installed power of 1 MW. While solar power plant systems are producing at full load; although transformer power is sufficient to provide reactive power control ancillary service with a power factor of ± 0,8, inverter powers are insufficient. Reactive power control ancillary service can be obtained from the existing solar power plant systems if they do not produce at full load.

In addition, the number of central and series inverters in the provinces in the Meram Electricity Distribution Company region can be seen in the table below. Central inverters offer remote communication and communication options. Central inverters serve smoothly as infrastructure. A full range of data communication methods are available, including remote monitoring.

Table 7. Number of PV Inverter Types

PROVINCE	SERIES (pieces)	CENTER (pieces)	The Overall Total
Aksaray	69	31	100
Karaman	60	25	85
Kırşehir	35	23	58
Konya	375	183	558
Nevşehir	144	42	186
Niğde	104	23	127
The Overall Total	**787**	**327**	**1114**

Energy storage systems, which are anticipated to become commonplace in the future, will be connected to the distribution grid in addition to producing facilities. It is important to define the power value, capacity value, storage size, and location for storage facilities reactive power control auxiliary service design. According to the distributed or central connection architecture, the location determination should be connected to the distribution (Evans, 2012).

One of the difficulties of increasing the renewable energy sources is the voltage change problems instantly due to the reasons of the sun, the wind speed etc. Voltage changes cause serious difficulties to consumers and distribution grid operator. The following figure shows the production profile for a random 2-day solar power plant selected from the Meram Electricity Distribution Company region. As can be seen in the figure, production increases at noon, but production may change instantaneously at 13.00 due to the cloud coming in front of the sun or for any other reason.

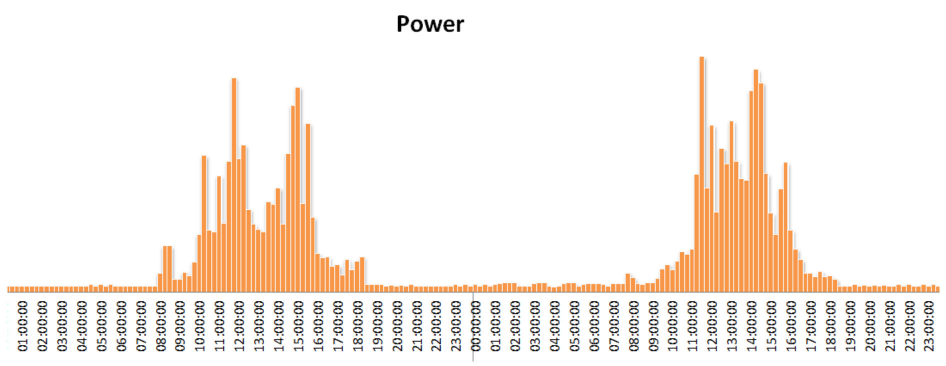


Figure 6. Production Profile For 2 Days Randomly Selected

Instant power changes and reactive power flow in the grid should be monitored effectively and communication infrastructure should be established. RTU is installed for field type solar power plants in Meram Electricity Distribution Company region. By connecting RTUs placed on the field to Meram Electricity Distribution Company's SCADA control center, an effective monitoring can be made, and even inverters can be controlled according to reactive need.

Congestion Management

Congestion management ancillary service should be used to eliminate both production-related and consumption-related congestions that occur or expected to occur in distribution grids.

In optimum distribution grid planning, distributed production sources should be close to the loads. Thus, it should be contributed to the power quality, the reduction of transmission and distribution peak loads, the need for long-distance transmission, the prevention of grid over-capacity, the deferral of grid investments and the reduction of grid losses, so that the congestion caused by production and consumption in the grid can be prevented. However, in reality, distributed production resources cannot always be close to loads and cannot be controlled. Therefore, distributed production resources do not always meet local demand.

Regarding the congestions caused by production; the increase in the prevalence of DERs in the distribution grid, the spread of rooftop type solar electricity generation in the regions where residential consumers are concentrated, congestion occurs in the transformers and lines at times when consumption in the grid is low. Especially in feeders with low load and high energy production with distributed generation facilities, problems such as overload, reverse power flow and voltage increase occur. As a result of reverse power and overload, it will cause various problems such as damage to conductors, aging of conductors and increased losses (Joos, 2000).

Also, the preveaance of EV's and the including of charge stations in the grid, charge EV batteries without any control can overload transformers and cables during peak hours, where EV's penetration is relatively high.

In order to overcome the problems mentioned in the conventional distribution grid operation, new investments are made to increase the cross-sections of conductors, or the prevalence of distributed generation resources in the relevant feeders is limited. However, in order to both investments deferral and to expand renewable-based energy production, it is necessary to monitor, dynamically control of distributed energy resources and supply AS from these sources.

Distribution grids, which traditionally have a one-way load flow, need to evolve into a grid structure suitable for bidirectional load flow and establish the infrastructure, depending on the renewable energy sources that are increasing today. At this point, it is an important requirement for distribution grid operators to provide congestion management ancillary service with demand side participation, both for production resources and for high consumption loads such as agricultural irrigation in Meram Electricity Distribution Company region, such as industry, commercial.

With the congestion management defined as ancillary service, technical losses can be reduced by ensuring that the investments are defer in the distribution grids, the energy supply-demand is balanced and the balanced loading of the grid.

In order to provide congestion management ancillary service from production sources, primary sources of production facilities must be stored or available if needed. For this reason, storage facilities should be used in congestion management for both production and consumption.

Performance requirements regarding congestion management ancillary service should be defined separately for production and consumption facilities. Production facilities that will provide congestion management ancillary service and consumers are required to sign a congestion management ancillary service with distribution companies.

Midwest Energy provides electricity distribution to 48.000 subscribers in Kansas. In order to solve the problems experienced due to the increase in demand, Enernoc has developed an agricultural load management program. Within the scope of the program, the irrigation pumps of the subscribers are automatically closed on weekdays of the summer between 14.00 and 19.00. Subscribers who are subject to the agreement are paid by Midwest Energy. It is an agricultural subscriber using 2000 of the subscribers over 30 kW. This corresponds to a peak load of 45 MW. The electricity distribution company pays 20 dollars for each kW that the subscriber does not use. The project, which started in 2009, tripled the number of subscribers in 2011 (2022, EnerNOC, Agricultural Load Management Project, http://www.enernoc.com/our-resources/case-studies/547-midwest-energy-grows-new-energy-supply-with-enernoc-agricultural-demand-response).

The company PG&E charges an additional fee of 1 dollar per kW when its subscribers consume energy during peak time. This method which is applied is very expensive for farmers and the wages even increase up to 10 times of their normal tariffs.Farm owners agreed with the authorized company and installed their automatic shut-off systems on water pumps. With this method applied, the farmers are asked whether they want to take part in load management at the time of the peak and if the farmers agree, the electricity of the irrigation pumps is interrupted for 2 - 4 hours in line with their agreement.

On the other hand, the prevalance of grids where periodic loads (agricultural, etc.) are dominant, where the load factor is quite low, also offers a significant potential for congestion management in terms of load management needs.

24% of total consumption in Meram Electricity Distribution Company region is agricultural irrigation consumers. In the table 8, there is provincial information based on agricultural irrigation cooperatives in Meram DSO region. In certain periods of the year, there is congestion in transformers and lines for agricultural irrigation purposes. In order to solve this problem, the demand side participation project is carried out with the agricultural irrigation consumer subscribers within the scope of the Ancillary Services Pilot Application project, which is an R&D project adopted by EMRA in July 2019 EnerNOC, 2022).

Table 8. Agricultural Irrigation Cooperatives Transformer Numbers And İnstalled Powers

PROVINCE	Number of Agricultural Irrigation Transformers (pieces)	Agricultural Irrigation Installed Power (KVA)
Konya	722	213.502
Karaman	84	41.706
Niğde	328	28.252
Aksaray	72	13.168
Kırşehir	18	1975
Nevşehir	40	13.061

With the installation of a storage system or systems in the grid, distributed energy storage architecture should be preferred. Because when the distributed energy storage architecture is compared with the central energy storage architecture; especially, it is expected to provide maximum benefit from congestion management and reactive power control AS (to deferral investment, to minimize compensation investments, to decrease grid losses and to increase grid efficiency).

The necessary communication infrastructure must be established in order to monitor and to send warning signals the producer / consumer points to participate in the congestion management ancillary service.

Within the scope of the congestion management ancillary service, components such as smart meters and sensors are applied at the producer / consumer points. The size of the data transferred from these components to the center increases with the spread of the program. Therefore, big data needs to be used / managed effectively due to increasing data size. As a result, it can be ensured that parameters such as price and required capacity amount can be estimated in real time and accurately by making use of big data management techniques.

Micro Grid Applications

Recently, there has been a lot of discussion on the controlled and reliable integration of DER into big power grids and microgrids. Due to its flexibility, controllability, and energy management skills, the microgrid concept—which essentially entails the coordinated functioning of a group of loads, distributed generators, and energy storage systems—is highly alluring. Grid-connected and stand-alone modes are the two ways that microgrids can operate to supply loads with consistent power. A microgrid will typically function in a grid connected mode and be connected to the mains. The microgrid is separated from the distribution grid in independent mode (Jiang, 2013).

Microgrids are required to function in both distribution grid-connected and stand-alone modes, and economically fulfill the demand on an instantaneous basis, in order to offer uninterruptible power supply to the loads. The issue of managing resources in a microgrid as efficiently as possible is a topic of intense research, and recent studies have suggested applying both centralized and distributed control schemes employing multi-agent systems, heuristic techniques, and optimization algorithms (URL-3, 2022). The highest priority for microgrids is to provide customers with a reliable power supply instead of economic benefits. Therefore, goals and energy management strategies are different in both modes. Microgrids are expected to be a part of the next electrical energy system development, not only in rural and remote areas, but also in urban areas. Planning processes should be directed to economic feasibility as a guarantee of long-term stability, as microgrids are expected to coexist with conventional power grids. Microgrids are required to function in both distribution grid-connected and stand-alone modes, and economically fulfill the demand on an instantaneous basis, in order to offer uninterruptible power supply to the loads. The issue of managing resources in a microgrid as efficiently as possible is a topic of intense research, and recent studies have suggested applying both centralized and distributed control schemes employing multi-agent systems, heuristic techniques, and optimization algorithms. (Gamarra, 2015) (URL-3, 2022).

In the case of isolation, the only structure that will control the quality of energy is the production systems that remain within the island. Thus, the production plants in the island (microgrid) must control the voltage and frequency. Therefore, dangerous energy oscillations are likely to occur in the microgrid structure. Thus, the Energy Management System (EMS) is critical for managing power and energy between loads and resources. Active and reactive power references can be checked with EMS (Zehir, 2013 Several functions are used to examine microgrid management, and they are categorized according to a temporal scale.

Long-term energy management includes the following:

- Forecasting RES output on an hourly basis, accounting for production costs and environmental effects,
- Management of controllable loads not connected to the grid according to the inspection requirements,
- Offering a suitable level of power reserve capacity in accordance with the electrical market and the predicted load demand,
- Maintenance frequencies.

Short-term energy management includes the following;

- Real-time power distribution, rms voltage regulation and primary frequency control via DERs,
- Real-time power distribution via DERs. (Gustavsson, 2013)

Improving Power Quality

Any issue with voltage, current, or frequency deviation that causes a customer's electrical equipment to malfunction can be categorized as a power quality issue. Power plants can currently only manage voltage; they have no control over the current a given load draws. As a result, they make an effort to maintain power quality standards within a given range.

While various definitions of power quality are not consistent with each other in the literature and standards, generally everyone's perception is the same. On the other hand, modern industrial processes working with new electrical equipment are affected by the lower power quality more than older equipment. Considering these situations, higher quality power is needed every day within the scope of our industry and individual life.

In power systems, harmonic voltage levels and flicker effect mentioned by specified in part of Annex 1 of the Grid Regulation in Turkey. It is not known exactly how much of the deterioration of the power quality is from the consumer and how much from the grid. Negative effects due to harmonics are generally not detected. The deterioration in power quality is also gradually increasing. Filters are used to eliminate current harmonics in power systems. Filters are classified under two main titles as passive and active harmonic filters.

Black-Start

According to the definition in the Electricity Market Ancillary Services Regulation of the Energy Market Regulatory Authority (EMRA); It is defined as energizing the transmission system through production facilities that can be commissioned without the need for an external energy source, supplying electrical energy to customers and re-commissioning other production facilities if the transmission system is partially or fully black-out (Turkey Energy Market Regulatary Authority-EMRA, 2017.)

Re-commissioning the system after the system is black-out is another task of the SO. For the majority of the producers need to get electricity from the grid to energize themselves and start production, they cannot facilitate reboot after failure. Therefore, black-start services are required from production facilities that can be commissioned without the need for an external energy source. When we consider that the supply of start-up services from energy sources with black-start capability is a purely public good, it is clear that this task is given to the SO, a regulated actor. In principle, both TSO and DSO are eligible for this service.

The use of storage systems in black-start is in trial phase, on May 10 2017, an electricity company in Southern California successfully demonstrated the use of a black-start by firing a deactivated combined cycle gas turbine with a battery energy storage system.

By using the black-start capability of the said battery energy storage system, an electricity generator and grid were energized, respectively, without the need for external transmission lines, thus the grid power functions were restored. 33MW / 20MWh lithium-ion battery energy storage system (BESS), which provides grid stability in daily use and helps smooth out output from local renewable power supplies, was used to operate the 44 MW combined cycle natural gas turbine at the Imperial Valley El Centro Production Station in California.

Optimization of Grid Losses

International regulations are specifying that distributed productions should not cause to an increase in losses. It is also specifying in the EMRA Electricity Market Unlicensed Electricity Generation Regulation that one of the purpose of distributed energy resources is to reduce losses in the.

Power plants to be connected to the distribution grid have the purpose of reducing losses in the distribution grid. However, integrations without any calculations can increase distribution grid losses. For example, a large group of power plants to be connected to the end of a relatively low feeder load will create an inverse power of greater power than the power previously carried in the related feeder, since only the load condition is considered in the respective feeder design, resulting in increased loss in the related line. The integration of distributed energy resourse into the distribution system will reduce the load of transmission transformers and reduce losses. However, the direction of losses decrease and increase in the distribution grid area which the distribution grid operator is responsible will directly depend on the locations of the distributed energy resources, the load distribution in the related region and the line / cable cross-section carrying the related power. For this reason, it is necessary to make loss analysis in production integrations and determine the connection points and line cross-sections that will prevent loss increase.

Regional Capacity Rental

Regional Capacity Rental Ancillary Service is a mechanism that enabling the required capacity to be established in the regions needed in the market structure in order to increase the supply continuity and quality and to fulfill the economic needs of the regional system, if the distribution system does not have sufficient capacity to ensure supply reliability.

Meram Electricity Distribution Company’s rural distribution grid realizes unforeseen long-term power interruptions due to reasons such as pole demolition from ice load or wind force, conductor breaks, especially in winter when there is heavy snowfall. In terms of economics of scale, there are no system strengthening investments in order to feed these scattered small loads supply with long feeders safely. For this reason, in order to meet the electrical energy needs of small residential units in the feeder failures that may occur in distribution grid, demand management can be provided in a hybrid micro grid with renewable energy source so that the energy need of the load can be meet continuously. Technology, DERs and consumption diversity play an important role in the development and adoption of micro grids.

Regional capacity rental ancillary service can also be used for micro grids in order to increase supply continuity and quality and meet the possible regional system needs economically if there is not enough capacity in ensuring distribution system supply reliability (Holttinen, 2012).

The Electricity Market Ancillary Services Regulation is regulated on the basis of the fact that only the TSO provides regional capacity leasing ancillary service and it does this in order to overcome the transmission system congestions and ensure supply security against transmission system congestions. The regulation needs to be reorganized by the DSOs, in particular for regional capacity leasing items, in order to ensure the continuity of supply in the distribution system and to overcome grid congestion, so that they can receive this ancillary service for microgrids and include storage systems.

TSO-DSO Coordination

The conventional power grid is a unidirectional power flow, and the electricity produced in conventional power plants is designed to deliver electricity to consumers connected to high, medium and low voltage levels. The task of the TSO in the conventional electricity grid can be summarized as the safe operation of the power system, the balancing of the power system in the balancing area, the measurement of electricity demand in the HV grid, ancillary service supply and the maintenance of the HV grid. Some of the task of DSO is to operate and maintain the MV and LV grid, deliver electricity to consumers, manage the grid capacity to minimize grid capacity costs and measure the electricity demand in the MV / LV grid.

However, in recent years, large amounts of RES have been connected to both the transmission and distribution grids, and it has become difficult for production to be programmable at the transmission level. In addition, power flows are now bidirectional in distribution grids, and a relatively high volume of electricity is generated at the system's MV and LV levels. The logic of making a strict distinction between transmission and distribution systems has disappeared, and TSO and DSO roles must be redesigned as a result.

While TSOs generally serve hundreds of large industrial consumers and producerss, DSOs generally have millions of residential customers as well as thousands of medium-sized commercial customers. In addition, today there are thousands of medium-sized producers (eg solar PV, wind, biomass, etc.) at the distribution level. In the future, hundreds of thousands of unlicensed small producers (eg rooftop PV panels, etc.) are expected to be active in the distribution grid at medium and low voltage levels (Kanchev, 2011).

It seems that DSOs will need AS as well as grid reinforcement investments in order to cope with this new situation and difficulties in grid operation. However, a stronger cooperation between TSOs and DSOs will be required to guarantee the current level of supply security.

Due to the situations mentioned above, the need for coordination between TSO and DSO becomes more and more important. SmartNet basically offers five models for TSO and DSO coordination.

Coordination Models

Centralized AS Market Model

In this scenario, the DER owner that is directly connected to the DSO grid for AS is contracted by the TSO. Although the DSO can supply and deploy resources to address local grid concerns, the procurement process differs from that of the centralized AS market in terms of timing. Without the widespread participation of DSO, TSO runs a market for both sources tied to the level of transmission and distribution. The most resembling example of the conventional structure is this one.

The prequalification process specified in the DSO role can be divided into two separate processes;

- **Technical Prequalification:** The TSO controls and approves the technical requirements of a unit that wants to join the AS market.
- **System Prequalification:** Approves the participation of distributed energy resources in AS market, provided that it complies with the operating conditions in the distribution grid (URL-4, 2022).

The role of DSO according to the centralized AS market model is limited by its role in the system pre-qualification process. Consequently, according to this model, DSO must provide the TSO with the necessary data, or the TSO must provide full observability of the DSO grid. Full observability means that real-time / near-real-time telemetry data of DERs are transmitted to TSO infrastructures instantly (SCADA etc.).

Local AS Market Model

The main principle of the local ancillary market model is the establishment of a local market and its management by DSO. After DSO has aggregated these resources and transferred them to the TSO AS market, TSO can sign a contract with DER indirectly from the local markets.

DSO organizes a local market for DSO-connected resources and, after resolving local grid restrictions, collects the remaining offers and makes them available to the TSO.

The local AS market model deviates from the centralized AS market model by promoting a local market. The implementation of such a market model shifts the priorities to DSO. In case of compliance with the distribution grid restrictions, all AS that are not needed / procured in the local market are included in the central market under the control of the TSO in bulk. AS offered to the central market by DSO collectively provide an advantage to TSO in terms of manageability.

Shared Balancing Responsibility Model

The TSO passes the "balancing" duty of the distribution grid to the DSO in accordance with the shared balancing responsibility paradigm. To carry out the DSO balancing duty, it executes DER's Daily Production Programs in accordance with a "predefined program" that takes into consideration demand and grid loss estimates for the territory under its jurisdiction. To carry out its balancing duties, DSO adheres to this predetermined program and makes use of DEKs connected to the distribution network via the regional AS market (Gubina, 2015).

The predetermined program is based on the fact that DERs are proposed (for the entire DSO region) in accordance with daily production plans, their offer to the market (i.e., merit order) as balancing units, and historical data projections at each TSO-DSO interconnection point. The predetermined program is currently only expressed in terms of the outcomes of the energy markets, and TSO and DSOs do not alter this program in order to control congestion. This predefined program means that it is determined at the level of the entire DSO area, not at the TSO-DSO interconnection point, because candidates are not specifically made to enter Merit Order for every TSO-DSO interconnection. Alternatively, TSOs and DSOs could determine the pre-defined schedule, using historical forecasts for each TSO-DSO interconnection point, together with constraints for both the transmission and distribution grid. In this second option, the pre-defined schedule is determined for each individual TSO-DSO interconnection point (URL-4, 2022).

In summary, TSO and DSO balancing responsibilities are fulfilled individually, using resources available on each grid. While DSO organizes a local market taking into account a balance program agreed with TSO, TSO does not have access to resources connected to the distribution grid.

The only coordination model in which TSO does not have access to connected resources in the distribution grid is the shared balancing responsibility model. Only the DSO is permitted to use the distribution grid's flexibility in order to carry out its duties related to balancing and addressing local grid restrictions.

Common TSO-DSO AS Market Model

The common TSO-DSO AS market model develops a common flexibility market for SOs. The main purpose of this coordination model is to minimize the total procurement costs of flexibility. This model requires incentives to improve the system as a whole, rather than focusing on minimizing the costs of TSO and DSO separately, regarding the revenue requirement of TSO and DSO.

The cost of the resources TSO and DSO require must be reduced. This shared objective can be accomplished by the coordinated efforts of a common market, the dynamic integration of a local market run by DSO, a central market run by TSO, or both (Enerquire, 2017).

The common TSO-DSO AS market model can be seen as an extension of both the local AS market model and the centralized AS market model. By combining the resources connected to the transmission grid and distribution grid, the optimization is still set up in the centralized variation. However, in this model, DSO grid restrictions and prospective local flexibility requirements are also included in the common markets. In contrast to the decentralized variant, the local AS market model does not give priority to exploiting flexible resources in its distribution grid. Based on a combined optimization of the needs for flexibility at both the distribution level and the transmission level, the DSO will decide which resources to deploy to meet local restrictions. (Atzeni, 2012).

Integrated Flexibility Market Model

The integrated flexibility market model encourages a regulated market (TSO and DSO) and commercial market parties to provide flexibility in a common market. This market is open to both TSO and DSOs and balancing units, aggregators etc., and requires the use of an independent market operator to ensure neutrality. As a result, the boundaries between balancing markets and AS can disappear.

The integrated flexibility market concept puts forth a market mechanism wherein SOs and commercial market players can both purchase available flexibility under the same terms. Actors that are controlled and those that are liberalized are interchangeable. The distribution of flexibility is decided by market forces. But this distribution will account for grid constraints at all voltage levels.

Evaluation of Coordination Models

All these models have different strengths and weaknesses. There is a lot of disagreement about which model to choose under which conditions. The advantage-disadvantage information about TSO and DSO coordination models described above are summarized in the table below.

Table 9. Comparison of Coordination Models

Coordination Model	Advantages	Disadvantages
Centralized AS Market Model	• Only suitable for TSO to purchase AS • Supports standard operations with low operating costs due to a single market • Compatible with many of the existing regulations	• DSO does not have real participation in the market • DSO grid restrictions are not always taken into account
Local AS Market Model	• DSO takes priority in the use of local resources • DSO actively manages the procurement process • Less barriers to small-scale DER participation in the market	• TSO and DSO market are balanced in order • Local market may have low liquidity • Comprehensive communication may be needed between the TSO market and the DSO market

Shared Balancing Responsibility Model	• TSO will require less AS in the future. • Less barriers to small-scale DER participation in the market • System operating limits are clear between TSO and DSO	• Total amount of As to be supplied by TSO and DSO will be higher • Balancing units may face higher balancing costs • Small local markets may not have enough liquidity to provide adequate resources to DSOs • The predefined program methodology adopted by both TSO and DSO can be difficult to define
Common TSO-DSO AS Market Model	• Total system costs of AS are minimum • TSO and DSO collaborate closely to take advantage of the most out of available resources	• TSO and DSO may have higher individual costs than other models • Costs between TSO and DSO can be difficult to allocate
Integrated Flexibility Market Model	• Opportunities for balancing units to address their portfolio imbalances are expanding. • There are many buyers and sellers, which leads to high liquidity and price competition	• Independent operator is required to operate the market platform • A detrimental effect on the growth and liquidity of intraday markets • TSO and DSO need to share data with independent market operator

As a result, it is recommended that each distribution grid operator should act as an aggregator in the flexibility of its region and use it in the distribution grid according to need or offer it to the transmission AS market, as the distribution regions differ (in terms of DERs, in terms of needed AS, size of distribution grid etc.) in each other. As a first step, it is considered appropriate to prefer the local AS market model, and then to pass to the common TSO-DSO market model.

New Roles of Distribution SOs

In the future, the new roles of DSOs will need to be defined in the legislation as follows;

1. DSOs, as neutral market facilitators, should be able to control, use and coordinate the impact of AS operations on their grids. DSOs can allow a distribution grid user to be activated by a TSO or a market player. However, it should ensure that control architectures as part of DSO's active system management responsibilities and, among other tasks, to provide for prior notification, supervising, evaluation and blocking potentially harmful control signals.
2. Cooperation and coordination between TSOs and DSOs is an important factor when DERs play an important role in providing real-time AS. It is especially important that the services are offered to the TSO and that all systems, as well as the local distribution networks to which they are connected, affect safe operation. Although it is appropriate for TSOs to be responsible for providing balancing services, they may still need to share some of this responsibility with DSOs, when it is considered that the share of the distribution-related organizations will increase in importance (Migliavacca, 2019).

3. TSOs and DSOs should work together to realise the efficient and non-discriminatory utilization of the capabilities of distribution connected generators and demand resources (wherever its connection point) to provide system services (e.g. voltage, frequency, inertia, etc.). As a first goal, relevant responsibilities for each system service should be stated.
4. DSOs should be allowed to use AS to manage their grids and optimize the capacities they can operate. DSOs should be allowed access and use of AS by DSOs such as technical solutions of their own assets, connection agreement, grid tariffs and market-based supply. Regardless of the AS model used, it should be financially viable for all interested parties.
5. The analysis made in the distribution grid of Meram Elektrik Distribution Company, with the spread of DERs, revealed possible problems in the distribution system; it should be able to provide AS from DERs in order to eliminate the effects of voltage rise, grid congestion and to deferral grid investments.
6. Because distributed resources are having an increasingly significant impact on the system's overall operation and planning, TSOs should collaborate with DSOs and regulators to determine the requirements for observability and active power management of distributed generation and demand side response.
7. Since DSOs need more tools to operate their grids (e.g. agreement with flexible resources) segmentation of markets should be avoided, there should be a single market for flexibility and balancing. TSOs, DSOs, regulators and market players should collaborate. A unique set of market rules should be developed to support the active participation of flexible resources of both TSOs and DSOs and to ensure their efficient supply. This situation will enable DSOs to provide flexible resources connected to their grid to manage local congestion and provide voltage control.
8. DSOs must be enabled to select the best and most cost-effective technology to operate the distribution system. Legislation should not limit the choice of technologies DSOs can use to fulfill the legal obligations of DSOs. Nevertheless, DSOs should ensure that the use of these technologies does not cause discomfort in the market. As long as it is efficient, a market-based solution should be preferred. Especially for grid-scale storage, DSOs should be allowed to own and operate such devices to secure the technical operations of the grid within the approved requirements.
9. The development of new distribution network tariff structures based on more capacity for efficient use of distribution system capacity should be supported.
10. In addition to the TSO-DSO coordination model, the types, features and arrangements for the provision of AS will need to improve to enable more participants to participate and use these new business opportunities. Therefore,

since potential flexibility providers are mostly small DERs, the regulation will need to take full account of the features of potential flexibility providers connected to the DSO side. In particular, the importance of market design for AS should not be ignored. However, if the architecture of real-time markets can fully take into account the characteristics of potential flexibility providers connected to distribution grids, it can offer its parties the opportunity to gain significant participation.

11. When making decisions on the TSO side, side effects on the DSO side should be taken into account to avoid lack of resources for alternative purposes or stimulated grid problems in another grid.
12. TSOs and DSOs should get enough information from each other. TSOs and DSOs must receive sufficient data (eg grid user data) to efficiently monitor and operate their grids. When considering for additional observability, level of detail and transparency demands expected by policymakers, data collection requires improvement.
13. DSOs should be encouraged to use AS for congestion management (cost effective for this). DSOs should be able to decide on the best solution to overcome specific challenges, either through ancillary solutions or grid improvement / empowerment.
14. Battery type storage systems respond to frequency faster than conventional production facilities and the success rate in providing frequency control service is close to 100% without deviation. Distribution grids should offer the ancillary service provided from storage systems to TSO as a different frequency control service category.
15. Storage systems can play a very important role in ensuring the production-consumption balance in microgrids.
16. All barriers to aggregators must be removed to ensure that the demand-side participation is placed on the market. This means that consumers can aggregate regardless of their connection points and will increase the gathering potential of consumers for each DSO marketplace.
17. The function played by the collector is crucial: Aggregators must be able to create a market interface that is streamlined and hide the intricate intricacies of a single flexibility provider’s characteristics. Aggregators should provide competitive pricing to entice flexibility providers to participate.
18. All market actors, including DERs, aggregators, and other customers, must have access to coordination models. This is anticipated to potentially involve additional rules, such as the development of the appropriate incentive schemes, as needed.
19. To lower the risk of clearing assets, grid planning will need to maximize the use

of RESs while minimizing infrastructure investments or deferral investments.

20. Finally, technical optimism should be supported by a full cost-benefit analysis, and necessary arrangements should be made in the electricity grid regulation, distribution grid regulation, and balancing and settlement regulation, in particular the AS regulation.

References

Atzeni, I., Ordóñez, L. G., Scutari, G., Palomar, D. P., & Fonollosa, J. R. (2012). Demand-side management via distributed energy generation and storage optimization. *IEEE Transactions on Smart Grid*, 4(2), 866-876.

Aubut, N. (2013). Wind power status and experience with inertial response from wind power plants. In *Proc. 2nd Workshop Active Power Control from Wind Power*.

Deutsche Energie-Agentur GmbH (dena) (2018). “Ancillary Services Innovation Report”. Berlin, Germany.

Divya, K. C., & Østergaard, J. (2009). Battery energy storage technology for power systems—An overview. *Electric power systems research*, 79(4), 511-520.

Delfanti, M., Galliani, A., & Olivieri, V. (2014, June). The new role of DSOs: Ancillary Services from RES towards a local dispatch. In *Cired Workshop*, Rome.

EnerNOC, Agricultural Load Management Project, http://www.enernoc.com/our-resources/case-studies/547-midwest-energy-grows-new-energy-supply-with-enernoc-agricultural-demand-response

EnerNOC, Agricultural Load Management ve Equipment Control, http://www.enernoc.com/our-resources/case-studies/farm-managers-harvest-energy-savings-and-avoid-surcharges

ENTSO-e. (2015). Towards smarter grids: developing TSO and DSO roles and interactions for the benefit of consumers. Availabe: https://eepublicdownloads. entsoe. eu/clean-documents/Publications/Position% 20papers% 20and% 20reports/150303 ENTSOE Position Paper TSO-DSO interaction. pdf.

Energy UK (April 2017), “Ancillary Services Report”, UK.

Enerquire, TSO & DSO coordination key for the integration of renewables, 2017

ENTOS-E: Balancing and Ancillary Services Markets, Retrived: 08.01.2017.

EURELECTRIC, 2011, 2012 ve 2013a.

EURELECTRIC (February 2004). “Ancillary Services - Unbundling Electricity Products – an Emerging Market” (Ref: 2003-150-0007). Brussels, Belgium. Available: http://pierrepinson.com/31761/Literature/Eurelectric2004-ancillaryservices.pdf

Hung, D. Q., & Mithulananthan, N. (2011, September). Community energy storage and capacitor allocation in distribution systems. In AUPEC 2011 (pp. 1-6). *IEEE*.

Evans, A., Strezov, V., & Evans, T. J. (2012). Assessment of utility energy storage options for increased renewable energy penetration. *Renewable and Sustainable Energy Reviews*, 16(6), 4141-4147.

FERC - Federal Energy Regulatory Commission, Guide to Market Oversight: Glossary, Retrived: 08.01.2018. https://www.ferc.gov/market-oversight/guide/glossary.asp).

Holttinen, H., Kiviluoma, J., Cutululis, N., Gubina, A., Keane, A., & Van Julle, F. (2012).

Ancillary services: technical specifications, system needs and costs: REserviceS Deliverable D2. 2.

Gamarra, C., & Guerrero, J. M. (2015). *Computational optimization techniques applied to microgrids planning: A review. Renewable and Sustainable Energy Reviews*, 48, 413-424.

Gubina, A. (2015). Ancillary services in the distribution network: Where are the opportunities?. no. December, T309.

Gustavsson, R., Hussain, S., & Saleem, A. (2013, June). Ancillary services for smart grids—Power quality markets. *In 2013 IEEE Grenoble Conference* (pp. 1-6). IEEE.

Jiang, Q., Xue, M., & Geng, G. (2013). Energy management of microgrid in grid-connected and stand-alone modes. *IEEE transactions on power systems*, 28(3), 3380-3389.

Joos, G., Ooi, B. T., McGillis, D., Galiana, F. D., & Marceau, R. (2000, July). The potential of distributed generation to provide ancillary services. *In 2000 power engineering society summer meeting* (cat. no. 00ch37134) (Vol. 3, pp. 1762-1767). IEEE.

Kling, W. L., Scarpellini, P., Alonso-Llorente, J. F., Auer, W., Beune, R. J., Braun, J., ... & Delabre, J. Y. (2000). Connection rules for generation and management of ancillary services. *Union of the Electricity Industry-EURELECTRIC, Tech. Rep*.

Kanchev, H., Lu, D., Colas, F., Lazarov, V., & Francois, B. (2011). Energy management and operational planning of a microgrid with a PV-based active generator for smart grid applications. *IEEE transactions on industrial electronics*, 58(10), 4583-4592.

Koohi-Kamali, S., Tyagi, V. V., Rahim, N. A., Panwar, N. L., & Mokhlis, H. (2013). Emergence of energy storage technologies as the solution for reliable operation of smart power systems: A review. *Renewable and Sustainable Energy Reviews*, 25, 135-165.

Lobato Miguelez, E., Egido Cortes, I., Rouco Rodriguez, L., & Lopez Camino, G. (2008). An overview of ancillary services in Spain. *Electric Power Systems Research*, 78.

Lalor, G., Mullane, A., & O'Malley, M. (2005). Frequency control and wind turbine technologies. *IEEE Transactions on power systems*, 20(4), 1905-1913.

Luo, X., Wang, J., Dooner, M., & Clarke, J. (2015). Overview of current development in electrical energy storage technologies and the application potential in power system operation. *Applied energy*, 137, 511-536.

Migliavacca, G. (Ed.). (2019). TSO-DSO Interactions and Ancillary Services in Electricity Transmission and Distribution Networks: Modeling, Analysis and Case-Studies. *Springer Nature*.

OpHB-Team, U. C. T. E. (2009). UCTE Operation Handbook Policy 1: Load-Frequency Control and Performance. Available on: http://www. entsoe. eu.

Rebours, Y. G., Kirschen, D. S., Trotignon, M., & Rossignol, S. (2007). A survey of frequency and voltage control ancillary services—Part I: Technical features. *IEEE Transactions on power systems*, 22(1), 350-357.

Saygın, D., Tör, O., Teimourzadeh, S., Koç, M., Hildermeier, J., & Kolokathis, C. (2019). Türkiye ulaştırma sektörünün dönüşümü: Elektrikli araçların Türkiye dağıtım şebekesine etkileri. *SHURA Enerji Dönüşümü Merkezi*. Retrieved from www. shura. org. tr.

Singh, H., & Papalexopoulos, A. (1999). Competitive procurement of ancillary services by an independent system operator. *IEEE Transactions on Power Systems*, 14(2), 498-504.

Sortomme, E., & El-Sharkawi, M. A. (2011). Optimal scheduling of vehicle-to-grid energy and ancillary services. *IEEE Transactions on Smart Grid*, 3(1), 351-359.

The Independent, Germany Energy Consumers Paid To Use Power Over Christmas As Supply Outstrips Demand, http://www.independent.co.uk/news/business/news/germany-energy-consumer-christmas-paidsupply- demand-outstrip-renewables-wind-solar-a8129716.html, Date: 02.02.2018.

Turkey Energy Market Regulatary Authority-EMRA, Electricity Market Ancillary Services Regulation, 12.10.2017.

UCTE Operation Handbook – Glossary, Final v2.2 E, 24.06.2004.

UK utility picks Younicos for 49 MW storage project". Utility Dive. Retrieved 2017-01-30.

URL-1: http://www.energystorageexchange.org/projects

URL-2: http://www.epdk.org.tr/TR/DokumanDetay/Elektrik/Mevzuat/Yonetmelikler/Sebeke

URL-3: ieeexplore.ieee.org

URL-4: https://smartnet-project.eu/wp-content/uploads/2016/12/D1.3_20161202_V1.0.pdf

Woods, R., Remøe, K. S., Hestnes, A. G., & Gustavsen, A. (2019). *Annual report* 2018.

Yumak, K., Tosun, G., Varlik, B., & Bağriyanık, M. (2016, November). Load curve classification for the evaluation of demand side management programs. *In IOP Conference Series: Materials Science and Engineering* (Vol. 161, No. 1, p. 012111). IOP Publishing.

Zehir, M. A., Batman, A., & Bagriyanik, M. (2013, June). An event-driven energy management system for planned control of thermostatic loads. In 2013 *IEEE Grenoble Conference* (pp. 1-6). IEEE.

Zipf, M., & Möst, D. (2016, June). Cooperation of TSO and DSO to provide ancillary services. In 2016 *13th International conference on the European energy market (EEM)* (pp. 1-6). IEEE.

About the Authors

Mert BAYER works as a network operations manager at Meram Electricity Distribution Company in Konya, Turkey. He has an undergraduate degree in electrical engineering and a master's degree in business administration. His main areas of interest are power systems, power distribution planning and analysis, power system reliability, ancillary services, distribution network master planning, smart grids, professional coaching, executive coaching, and leadership training program.

E-mail: mert.bayer@meramedas.com.tr, **ORCID:** 0000-0003-2221-0714

Harun KOROGLU is the R&D Chief Engineer at Meram Electricity Distribution Company in Konya, Turkey. He has a degree in electrical engineering from Eskişehir

Osmangazi University. His main areas of interest are electrical distribution system, protection system, electrical grid planning, electrical fault-repair, electrical storage systems, and R&D project management.
E-mail: harun.koroglu@meramedas.com.tr, **ORCID:** 0000-0001-9936-598X

Mustafa Emre SAFAK is a performance improvement coordinator at Sakarya Electricity Distribution Company in Sakarya, Turkey. He has an undergraduate degree in electrical engineering from Middle East Technical University and master's degree in business administration from Bahçeşehir University. His main areas of interest are electrical distribution system, power distribution, renewable energy, electrical fault repair, electrical storage systems, and project management.
E-mail: emresafak@gmail.com, **ORCID:** 0000-0002-8213-8892

Burak ALTUN is the purchasing and contract manager at Meram Electricity Distribution Company in Konya, Turkey. He has a degree in electrical engineering from Kocaeli University. His main areas of interest are electrical distribution system and cryptology.
E-mail: Burak.Altun@meramedas.com.tr, **ORCID:** 0000-0002-9446-2841

Mert BAYER works as an assistant professor network operations manager at Meram Electricity Distribution Company in Konya, Turkey. He has an undergraduate degree in electrical engineering and a master's degree in business administration. His main areas of interest are power systems, power distribution planning and analysis, power system reliability, ancillary services, distribution network master planning, smart grids, professional coaching, executive coaching, and leadership training program.
E-mail: mert.bayer@meramedas.com.tr, **ORCID:**0000-0003-2221-0714

Huseyin Oktay ALTUN works as an assistant professor at KTO Karatay University in Konya, Turkey. He has an undergraduate degree in electrical engineering from Bilkent University. He obtained his MSc and PhD degrees from University of Rochester. His main research areas of interest are reinforcement learning, convex optimization and energy systems design.
E-mail: oktay.altun@karatay.edu.tr, **ORCID:** 0000-0002-8875-6878

Similarity Index

The similarity index obtained from the plagiarism software for this book chapter is %3.

To Cite This Chapter:

Koroglu H. ,Safak M. E., Bayer M. , Altun B., & Altun H.O., (2022). Providing Ancillary Services From Distributed Energy Resources Connected To Electricity Distribution Grid,   In S. Kocer, O. Dundar (Eds .), *Current Studies in Basic Sciences Engineering and Technology* (pp. 129–167). ISRES Publishing.

Use of IoT and Wearable Technology Design Fundamentals in Healthcare Industry

Sabri KOCER
Necmettin Erbakan University

Hakan OCAL
Department of Computer Engineering,
Bartin University

Ozgur DUNDAR
Necmettin Erbakan University

Hasan Berkant DUNDAR
Istanbul Medipol University

Introduction

Today, IoT and Wearable Technology are among the fastest growing technologies. It can be shown as an innovative feature that allows remote access, processing and control, which can be connected to the Internet. Wearables technologies are used in smart homes, between machines (M2M), transportation between vehicles (V2V), healthcare, electrical networks and many other industrial applications.

The world's population is growing older and older due to low birth rates in many parts of the world and the desire to live longer with advancing technology. It is estimated that from 2025 to 2050, the population over the age of 65 in the world will reach 1.6 billion. With this increase in the elderly population around the world, it is inevitable that problems related to providing various health services (stroke, dementia, heart attack, etc.) will arise. In addition, it has been expanded to cover various health problems (actuators, GPS devices, mobile devices and various sensors, etc.) from all age groups in order to address more objects.

The IoT also has the potential to analyze captured data by accurately tracking people, equipment, assay samples, supplies, and test animals. For example, sensors can detect deficiencies in the patient's exercise activities by comparing excessive cardiological or walking activities with normal walking pace. The data obtained in these two examples can be collected and processed in a center with the help of wired or wireless communication technologies, providing very valuable information to healthcare providers.

In addition, these data can be reviewed by insurance companies, healthcare providers and pharmaceutical companies and used for future studies. In addition, the fitness movements of the person can be detected by sensors and collected with the help of wired or wireless network technologies and analyzed in a center. Thanks to the analyzed data, it can be ensured that the person performs more accurate and regular exercise movements in order to keep his body fitter in intense work tempo. Diet status can be controlled simultaneously with a wearable or wearable device (glasses, wearable devices placed on the teeth or devices that detect mouth movements, etc.) to be placed on the body while performing the fitness movements of the individual.

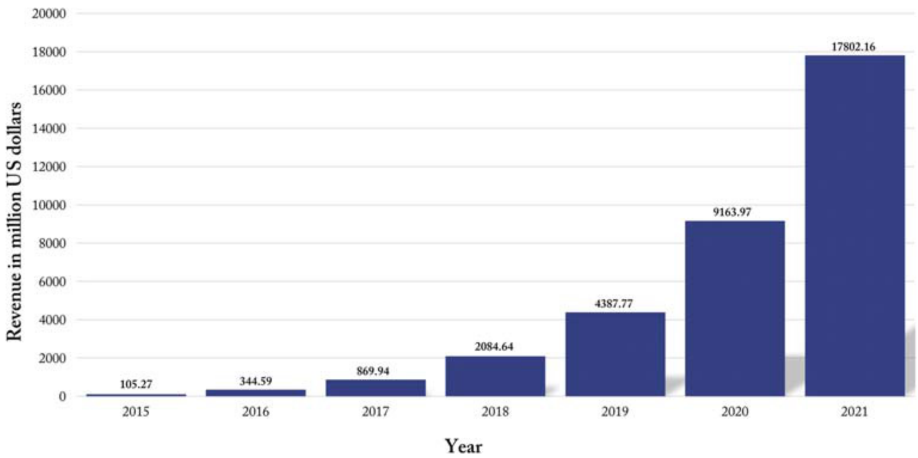


Figure 1. Projected global market size for wearables in healthcare from 2015 to 2021. (Junata,2018)

Wearable technology itself is proliferating in development in both the commercial and research worlds. Figure 1 shows the gradual increase of the projected global market for wearable devices in healthcare from 2015 to 2021. In particular, wearable technology will be a large part of healthcare and has a large market share within the pharmaceutical industry in the future.

With an increasing number of global aging populations, wearable technologies paired with home health and rehabilitation services will be extremely beneficial. We can think of mechanical watches as the ancestor of wearable theologies. Likewise, glasses are examples of wearable technologies. We use crutches for clothes to keep us warm in winter and cold days or to move around when a leg is broken. Humans have a long history of developing technology to extend the functionality of the human body. In recent years, with the development of technology for the defense industry, health services and professional sports, many wearable devices have been developed. Today, our smart watches, smart glasses and different kinds of smart clothes are produced. If we take a brief look at the history of wearable technology in Figure 2;

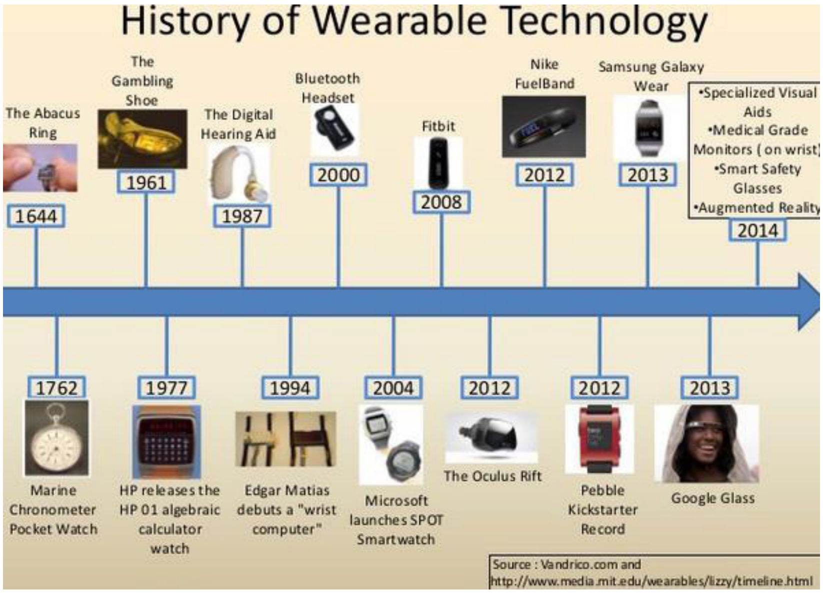


Figure 2. A Brief Look into the History of Wearable Technology (URL 1,2016)

The first wearable computer was invented in the 1960s by mathematics professor Edward Created by Thorp.

The first algebraic calculator wristwatch went public in 1977. The Sony Walkman was released four years later. In the 1987s, digital hearing aids, which are the ancestors of today, were introduced.

Steve Mann, a Canadian researcher, invented the wearable wireless webcam in the 1990s. In 1994, the wrist computer was developed. Such devices have made it easier to use future IoT technologies. Fairs and academic studies on wearable technology gained popularity in the 90s. Wearable technology gained popularity in 2000 with Bluetooth headphones.

In 2011, wearable technology gained a lot of attention when Google developed the first prototype of smart glasses, today called Google Glass.

In the 2020s, the use of AR and VR glasses, especially in game software on computers and smartphones, has increased a great interest in wearable technologies.

It offers various health-related areas such as patient disease symptom monitoring for guidance and rehabilitation monitoring in patients' daily activities and operating rooms. Smart healthcare has spurred the development of smart wearable mobile devices. Some of the smart wearable mobile devices are smart phones, smart watches, etc. to monitor the heart rate, sleep status, blood pressure and other activities. This opens new facets in the medical field by sensing data with wearable devices to analyze patients' health, behavior and recommendations to further improve health. It also helps patients to be controlled by doctors remotely and in real time. Wearable smart devices should be convenient, cost-effective, safer and easily usable by the user. In the future, cloud computing can be further enhanced with big data and deep neural networks.

Wearable technologies provide continuous monitoring of physiological and biochemical parameters as well as human physical activity and behavior throughout daily life. Vital signs such as heart rate, blood pressure and body temperature can be observed through wearable devices. It can also measure electrocardiogram (ECG), ballistocardiogram (BCG), and oxygen saturation. Wearable devices can be attached to glasses, shoes, gloves, earrings, clothing and watches. Wearable devices can also be mounted on the skin.

The sensors can be placed in environments such as car seats and chairs. Advanced smartphone technologies, on the other hand, can transmit this information to a remote server for storage and analysis. Wrist activity monitors and mobile phone applications can be used to analyze gait with accelerometers and gyroscopes. It is also possible to do data analysis with software. In this way, wearable technologies can be innovative solutions to health problems..

What Is The Internet of Things (IoT) and How Does It Work?

IoT refers to all systems that can transmit data over a network, information received on interrelated computing devices, mechanical and digital machines, objects, animals. It includes advantages such as working more efficiently with IoT technology, improving customer service and improving decision making.

Figure 3 shows a picture of all devices connected to the internet through databases, data aggregators or relays, display services or other devices.

IOT devices send the sensor data they collect to the cloud for analysis and share it with the desired devices when necessary. Sometimes these devices communicate with other related devices and can make decisions based on this information. The best example of an advanced IoT device is the human body. It stores data from our sense organs and can mimic most sensory abilities when needed. With IoT we can use it to draw conclusions about analog or digital values from sensors.

Figure 3. All devices connected to the internet through databases(URL 2,2022)

For example; With the smart watch on your wrist, we can monitor all the activities of the day, the distance you have covered, the steps you take and how your heart beats. This data collected here is analyzed and we can reach important information about our health.

Basic structure of IoT solutions

An IoT solution consists of many layers, from a physical device to the end user, by combining many different technologies into a single product. These layers;

- Device hardware: Every IoT project usually includes hardware including Microcontroller Unit (MCU) and sensors/actuators. It is also very important to choose the most suitable communication medium for the IoT device.
- Device firmware: According to the content of the project, it is necessary to write the device firmware to run on the MCU. This is where we collect data and transfer it to other components as needed.
- Communication: There should be a common language for obtaining and sharing data from the physical environment. The protocols that provide the connection between these devices enable the communication of data both to the physical environment and to the application layer.
- Backend system: In this part, all data is collected in the back-end system and the management, monitoring and integration of the product is provided. Depending on the project requirements, data can be sent to cloud providers. In addition, big data analytics can be applied on the data from the Sensors or artificial intelligence algorithms can be used when necessary.
- End-user applications: In this part, a modern interface can be designed to control and monitor the necessary information.

The Figure 4 diagram below shows the general structure of IoT solutions:

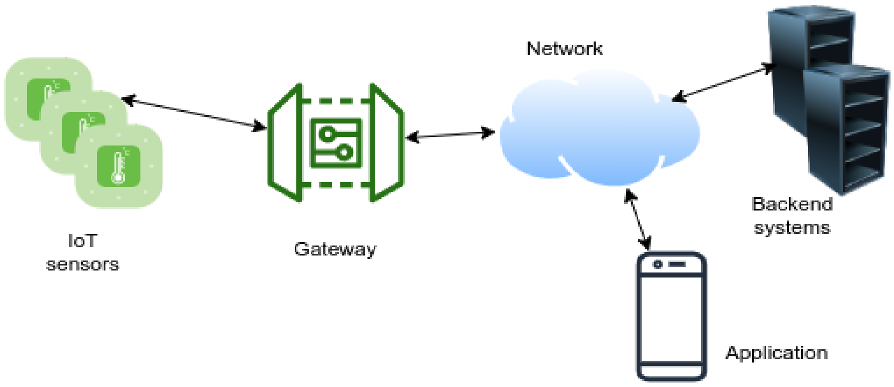


Figure 4. General structure of IoT solutions (URL 3,2021)

The Wearable is based on 3 layers: First Layer, this layer is sensors. The sensors are placed closest to the body. Sensors monitor items such as temperature, movement, and heart rate. Layer Two, this layer is the link and control layer. The Bluetooth Low Energy (BLE) protocol is the most widely used protocol for connecting wearable devices to a smartphone or home network. Layer Three, this layer is the cloud from which the wearable provides and reads data.

Different Types of Wearable Technology

Different types of wearable technology, commonly known as wearable technologies, refer to smart electronic devices that can be incorporated into clothing or attached to the body as implants or accessories. Smart watches, fitness trackers, shiny jewellery, smart clothing, body wearables, and head-worn displays are examples of standard wearables in Figure 5.

Wearables have gained great momentum with customers over the past decade. Smartwatches are probably the best-known wearables. In addition, products such as glasses, head-mounted displays (HMDs), Socks and gloves are examples of modern wearable technology on the market.

Figure 5. Different types of wearable technologies (URL 4,2022)

Since these devices are in contact with the user's body, they offer services such as calorie tracking and sleep monitoring thanks to embedded systems.

Some of the different types of wearable technology include;

1. Smart Watches

Smart Watches can easily access the desired information both on their own screen and by connecting to the smart phone. It allows to take EKG heart rhythm, SpO2 or body temperature through sensors on the watch.

2. Audible

Hearing aids allow you to hear what is happening around you. It also has audio input and output. The earbuds are specially designed to hear ambient noise. These devices have microphones on them.

3. Fitness Tracker

With fitness trackers, heart rate, SpO2, steps taken, calories burned and other fitness indicators are all tracked. Fitness-oriented products can also send notifications and alert messages to smartphones when necessary.

4. Smart Shoes

Smart Shoes have also gained importance in recent years. With these devices, the number of steps and speed and step length, distance, walking time and other data are collected with smart shoes.

5. Smart Clothing

Smart suits can provide deeper insights than other examples of modern wearable technology, as they come into contact with a more important part of the body, enabling advanced monitoring for medical care and lifestyle improvement.

Figure 6. Smart sock (URL 5,2022)

Each smart sock is infused with three thin, soft textile pressure sensors. Attached to each smart sock, a magnetic Bluetooth smart electronic anklet collects data and transmits it wirelessly to Sensoria's mobile app and web dashboard to provide runners with real-time visual and audio feedback, as well as post-workout analysis.

It includes sensor calibration and motion capture algorithms with smart pants and combines it with e-skin smart clothes that can be used in daily life. This allows for a variety of applications in rehabilitation, sports, occupational safety and wherever detailed and robust motion capture is required.

6. Smart Glasses

Smart glasses record what the user sees and send the data to a cloud system when necessary. These glasses can do anything like our smartphones and tablets. They have internal and external sensors that can collect data from computers, smartphones and other electronic devices.

7. Wearable Cameras

Wearable cameras capture images from a first-person perspective. They are often attached to clothing such as shirts, hats or helmets. In addition, wearable cameras capture events such as daily life, parties, performances, sports and other events.

8. Implantable body sensors

They are implantable, communicating from the inside with the wearer's body rather than the skin. Some of such wearables are available both in fitness monitoring units and medically, such as blood glucose, blood pressure, electrocardiography, pulse rate, respiratory rate, body temperature, etc. It is intended to measure body information such as (Figure 7).

Pill-like sensors have been produced that can monitor blood pressure. After swallowing such drugs, the user can easily monitor the data collected from the body.

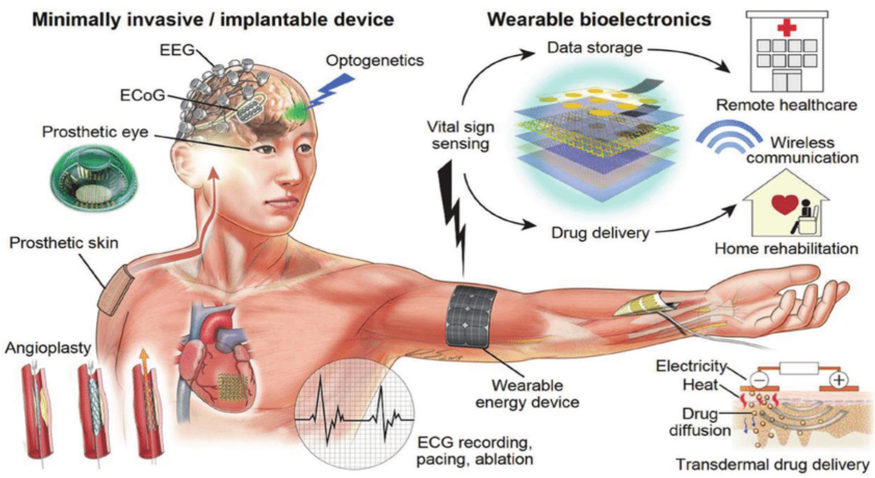


Figure 7. Body sensors (URL 6,2022)

Body sensors can be implanted on or inside the skin's surface. They move with the skin, collecting data and transmitting it to computers or mobile phones connected to them. Electrocardiogram (ECG) and electroencephalogram (EEG) readings can be made with body sensors, and also when the patient takes the sensor pill can be checked. In addition, body sensors monitor physical activity and sleep patterns.

9. Life Belt

The Life Belt is a transkarin wearable device that monitors the physical health of pregnant women. Obstetricians and doctors do not want the patient to visit them often. Thus, they monitor their health remotely using a life belt. It sends alerts to patients and authorities if any complications occur. So they can take quick action to save both the mother and the fetus.

10. Head Mounted Displays

Head-mounted display (HMD) is a device that provides virtual reality and augmented reality for the user. It is worn on the head or as part of a helmet.[1] In front of it is a small screen optics (monocular or binocular HMD). There is also an optical head mounted display (OHMD) type. This can mirror the projected images and allow the user to see it. A typical HMD has one or two small screens, lenses and translucent mirrors built into the glasses. It is used as augmented reality (AR) virtual reality (VR) glasses with HMDs.

11. Smart Jewelry

Another product used in monitoring health data is smart jewelry. These smart jewels can be in the form of necklaces, bracelets, watches and rings. They can track your heart rate, sleep, stress and advise you on how to stay healthy.

12. Location Tracker

With Location Tracker, it usually receives GPS data and sends the location of the product to the connected device. GPS trackers for kids are extremely useful. Especially parents can use such devices to find out what is going on around their children and their location.

13. Exoskeletons

It uses a combination of electric motors, pneumatics, levers, hydraulics and tentacles on exoskeletons, also known as outer suits.

Also, Active and passive exoskeletons are the two main types of exoskeletons. The user's movements are strengthened by actuators on active exoskeletons. They can be used especially in medical, military and civil settings.

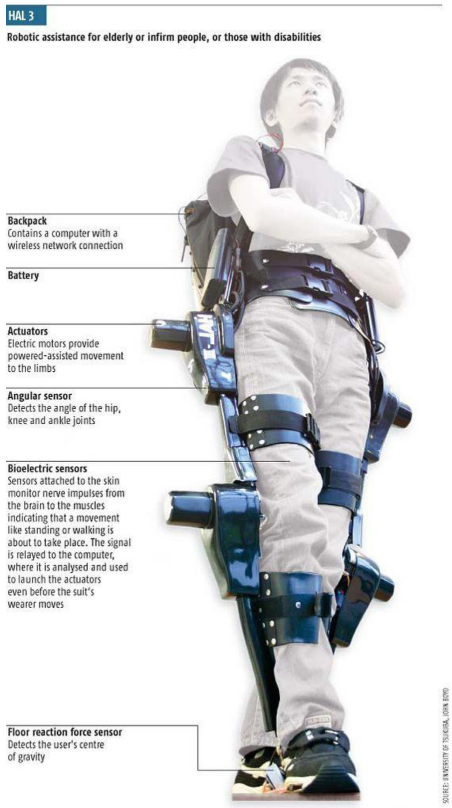


Figure 8. Robot asistance for elderly or infirm people (URL 7,2022)

14. Smart Jacket

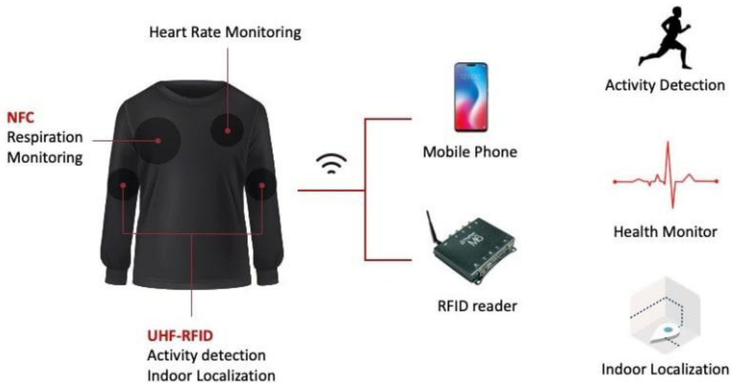


Figure 9. Smart jacket (URL 8,2022)

This Smart jacket can collect information with human health and identify where people are in Figure 9. It can wirelessly transmit the data on the sensors on the jacket to the cloud. At the same time, the jacket can harvest energy to provide a self-contained power source. The life jacket is one of the most widely used wearable devices in the medical field. Thanks to these devices, we can control a person's heart rate or blood pressure.

When necessary, this information can be transferred to the hospital environment.

Results And Discussion

The environment in which doctors or other healthcare professionals can best monitor the patient for chronically ongoing diseases is the patient's natural life. Therefore, various portable monitoring devices are needed to monitor patients in their natural life. These devices can send the data they receive from the patients to the relevant health service providers (hospitals, insurance companies, etc.), enabling them to take various decisions more efficiently for that person and for the future in health services.

Today, wearable health products come up with devices with much more advanced capabilities. For example, it is used to collect vital values such as heart rate with sensors attached to the skin surface and to transfer data to doctors on their devices. With portable Holter-like devices, the patient's electrocardiogram can be recorded and monitored when necessary. This is done via a wireless electrode that must be in contact with the patient. It is usually affixed to the patient's finger or chest. Smart gloves are designed for patients who cannot control their hand movements. The device helps the patient regain hand movement with repetitive arm and hand movements. In the academic community, many articles and projects have been made in the field of wearable technologies in the Health sector.

Modern smartwatches help detect cardiovascular diseases by capturing systolic and diastolic pressure. It also offers advice on oxygen and blood pressure levels, sleep patterns, diet and activity levels.

The Future of Wearable Technology in Healthcare

Many health-related reasons will affect wearable technology. First, technological advances make it easier to design and build non-intrusive technologies that monitor a person's health. Second, technological evolution will solve some pressing technical challenges, such as the short battery life and size of these devices. Implantable devices, which already have an important role in the field of health (pacemakers), will come to the fore with the miniaturization of this technology. Wearable health monitoring devices and any technology with similar monitoring functions will be easy to install, use and manage, effectively improving quality of life and health delivery.

Technology has helped make significant advances in both patients and healthcare. Thanks to its telehealth solutions, patients can now consult digitally and take the necessary medicines on time instead of going to hospitals for consultation.

The increase in the use of smart devices has made people want to follow all kinds of data on these devices. Businesses in the competitive market, on the other hand, have taken technology to different dimensions by not ignoring this desire. One of these dimensions and a technology that has attracted attention with its usage rate in recent days has been smart products. With this technology, people can keep track of their daily activities and evaluate their statistical results. Human beings can obtain and record many data in different ways, but wearable products help in interpreting this data. These products, which can be used for all animate and inanimate beings apart from being used only for humans, have now become a necessity. The rapid progress of technology makes human life much easier, and it can also reveal some vital problems. Before having a technology, it is necessary to know the most effective use and protection methods.

References

Arun, C. S., & Alexander, A. (2017, April). Mobile ECG monitoring device using wearable non contact armband. In 2017 International Conference on Circuit, Power and Computing Technologies (ICCPCT) (pp. 1-4). IEEE.

Benadda, B., Beldjilali, B., Mankouri, A., & Taleb, O. (2018). Secure IoT solution for wearable health care applications, case study Electric Imp development platform. International Journal of Communication Systems, 31(5),

Brophy, K., Davies, S., Olenik, S., Çotur, Y., Ming, D., Van Zalk, N., ... & Yetisen, A. K. (2021). The future of wearable technologies. Imperial College London: London, UK.

Chen, B. R., Patel, S., Buckley, T., Rednic, R., McClure, D. J., Shih, L., ... & Bonato, P. (2010). A web-based system for home monitoring of patients with Parkinson's disease using wearable sensors. IEEE transactions on biomedical engineering, 58(3), 831-836.

Choi, S., Lee, H., Ghaffari, R., Hyeon, T., & Kim, D. H. (2016). Recent advances in flexible and stretchable bio-electronic devices integrated with nanomaterials. Advanced materials, 28(22), 4203-4218.

Yotha, D., Pidthalek, C., Yimman, S., & Niramitmahapanya, S. (2016, December). Design and construction of the hypoglycemia monito wireless system for diabetic. In 2016 9th Biomedical Engineering International Conference (BMEiCON) (pp. 1-4). IEEE.

Dey, N., Ashour, A. S., Fong, S. J., & Bhatt, C. (Eds.). (2019). Wearable and Implantable Medical Devices: Applications and Challenges.

Gay, V., & Leijdekkers, P. (2007). A health monitoring system using smart phones and wearable sensors. International journal of ARM, 8(2), 29-35.

Geng, H. (Ed.). (2017). Internet of things and data analytics handbook. John Wiley & Sons.

He, W., Goodkind, D., & Kowal, P. R. (2016). An aging world: 2015.

Westwood, J. (2011). Online predictive tools for intervention in mental illness: the OPTIMI project. Med. Meets Virtual Real. 18 NextMed, 163, 86.

Liu, Q., Poon, C. C. Y., & Zhang, Y. T. (2012). Wearable technologies for neonatal monitoring. In Neonatal Monitoring Technologies: Design for Integrated Solutions (pp. 12-40). IGI Global.

Huang, C. H., & Cheng, K. W. (2014). RFID technology combined with IoT application in medical nursing system. Bulletin of Networking, Computing, Systems, and Software, 3(1), 20-24.

J. Varma, S., Sambath Kumar, K., Seal, S., Rajaraman, S., & Thomas, J. (2018). Fiber-type solar cells, nanogenerators, batteries, and supercapacitors for wearable applications. Advanced Science, 5(9), 1800340.

Junata, M., & Tong, R. K. Y. (2018). Wearable Technology in Medicine and Health Care: Introduction. Wearable Technology in Medicine and Health Care, 1.

Kang, J. J., Luan, T. H., & Larkin, H. (2016, November). Inference system of body sensors for health and internet of things networks. In Proceedings of the 14th International Conference on Advances in Mobile Computing and Multi Media (pp. 94-98).

Kaplan, M., Berk, T. N., Çemrek, B., Şahin, S., & Fidan, U. (2017, October). Mobile physiological signal monitoring system for family medicine. In 2017 Medical Technologies National Congress (TIPTEKNO) (pp. 1-4). IEEE.

Khan, S. F. (2017, March). Health care monitoring system in Internet of Things (IoT) by using RFID. In 2017 6th International conference on industrial technology and management (ICITM) (pp. 198-204). IEEE.

Kurniawan, A. (2019). Internet of Things Projects with ESP32: Build exciting and powerful IoT projects using the all-new Espressif ESP32. Packt Publishing Ltd.

Kuusik, A., Alam, M. M., Kask, T., & Gross-Paju, K. (2018, February). Wearable m-assessment system for neurological disease patients. In 2018 IEEE 4th World Forum on Internet of Things (WF-IoT) (pp. 201-206). IEEE.

Labrador, M. A., & Yejas, O. D. L. (2013). Human activity recognition: Using wearable sensors and smartphones. CRC Press.

Laplante, P. A., & Laplante, N. (2016). The internet of things in healthcare: Potential applications and challenges. It Professional, 18(3), 2-4.

Lay-Ekuakille, A., & Mukhopadhyay, S. C. (2010). Wearable and autonomous biomedical devices and systems for smart environment. Heidelberg, Germany:: Springer.

Lebepe, F., Niezen, G., Hancke, G. P., & Ramotsoela, T. D. (2016, July). Wearable stress monitoring system using multiple sensors. In 2016 IEEE 14th International Conference on Industrial Informatics (INDIN) (pp. 895-898). IEEE.

Lee, J., Wheeler, K., & James, D. A. (2019). Wearable sensors in sport: a practical guide to usage and implementation. Springer Singapore.

Lukowicz P., Anliker U., Ward J., Troster G., Hirt E., (2002), Neufelt C., "Amon: A wearable medical computer for high risk patients", 6th International Symposium on, Seattle, 133-134.

Majoe D., Bonhof P., Kaegi-Trachsel T., Gutknecht J. and Widmer L., , (2010). "Stress and Sleep Quality Estimation from a Smart Wearable Sensor", 5th International Conference on Pervasive Computing and Applications, Maribor, 14-19

Mills, S., Lucas, S., Irakliotis, L., Rappa, M., Carlson, T., & Perlowitz, B. (2012). Demystifying big data: a practical guide to transforming the business of government. TechAmerica Foundation, Washington.

Mukhopadhyay, S. C. (Ed.). (2015). Wearable electronics sensors: For safe and healthy living (Vol. 15). Springer.

Mukhopadhyay, S. C., Suryadevara, N. K., & Nag, A. (2021). Wearable Sensors and

Systems in the IoT. Sensors, 21(23), 7880.

Neagu G., Preda Ş. and Stanciu A., (2017) “A Cloud-IoT Based Sensing Service for Health Monitoring”, The 6th IEEE International Conference on E-Health and Bioengineering – EHB, Sinaia, 53 – 56,.

Otto, C., Milenković, A., Sanders, C., & Jovanov, E. (2006). System architecture of a wireless body area sensor network for ubiquitous health monitoring. Journal of mobile multimedia, 307-326.

Qi, J., Yang, P., Min, G., Amft, O., Dong, F., & Xu, L. (2017). Advanced internet of things for personalised healthcare systems: A survey. Pervasive and Mobile Computing, 41, 132-149.

Rahimi, F., Bee, C., Debicki, D., Roberts, A. C., Bapat, P., & Jog, M. (2013). Effectiveness of BoNT A in Parkinson’s disease upper limb tremor management. Canadian journal of neurological sciences, 40(5), 663-669.

Rahimi, F., Duval, C., Jog, M., Bee, C., South, A., Jog, M., ... & Boissy, P. (2011, August). Capturing whole-body mobility of patients with Parkinson disease using inertial motion sensors: expected challenges and rewards. In 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (pp. 5833-5838). IEEE.

Sabban, A. (2017). Novel wearable antennas for communication and medical systems. CRC Press.

Salarian, A., Russmann, H., Wider, C., Burkhard, P. R., Vingerhoets, F. J., & Aminian, K. (2007). Quantification of tremor and bradykinesia in Parkinson’s disease using a novel ambulatory monitoring system. IEEE Transactions on biomedical engineering, 54(2), 313-322.

Santhi, V., Ramya, K., Tarana, A. P. J., & Vinitha, G. (2017). IOT based wearable health monitoring system for pregnant ladies using cc3200. International Journal of Advanced Research Methodology in Engineering & Technology, 1(3), 56-59.

Sazonov, E. (Ed.). (2020). Wearable Sensors: Fundamentals, implementation and applications. Academic Press.

Szczęsna, A., Nowak, A., Grabiec, P., Paszkuta, M., Tajstra, M., & Wojciechowska, M. (2017). Survey of Wearable Multi-modal Vital Parameters Measurement Systems. In Innovations in Biomedical Engineering (pp. 323-329). Springer, Cham.

URL 1: A brief history of wearable computing.” http://media.mit.edu/wearables/lizzy/timeline.html. [Accessed: 21-Mar-2016].

URL 2:MIT Professional Education announces new online course on the Internet of Things, 2016 https:/news.mit.edu/2016/mit-professional-education-announces-new-online-course-internet-things-0408

URL 3: Wearable technology applications in healthcare: A literature review. (2021, April 2). HIMSS. https://www.himss.org/resources/wearable-technology-applications-healthcare-literature-review

URL 4: Uwadiare, J. (2022, October 3). 14 different types of wearable technology. The Techtian. https://Thetechtian.com/different-types-of-wearable-technology

URL5: Sensoria Home Page (2022). https://www.sensoriafitness.com

URL 6: Wearable Devices, 2022 https://doi.org/10.1002/advs.201870057

https://onlinelibrary.wiley.com/…adma201504150-fig-0001-m.jpg (1333×758) (wiley.com)

URL 7: Exoskeletons. (n.d.), 2022. https://www.oobject.com/category/exoskeletons

URL 8:Liu, Y. (n.d.). "Smart jacket" for future urban life. Keysight IoT Innovation Challenge, 2022. https://www.iotchallengekeysight.com/2019/entries/smart-land/292-0515-145528-smart-jacket-for-future-urban-life

About the Authors

Sabri KOCER, He is a professor and senior researcher in the Department of Computer Engineering at Konya Necmettin Erbakan University. He completed his doctorate at the Department of Electronic Computer Education at Gazi University. He has held academic positions at different Universities in Turkey and has made great contributions to the on-going technological academic development in the region. His main research interests are biomedical engineering, artificial intelligence, optimization, and machine learning and microcontroller applications. In addition, he has many academic works such as books, articles, essays, posters published in many national and international journals.

E-mail: skocer@erbakan.edu.tr, **ORCID**: 0000-0002-4849-747X

Hakan OCAL, He is a PhD and A class Occuppational Safety Specialst in the department of Occupational Health and Safety at Republic of Türkiye Ministry of National Education. He completed his doctorate at the Department of Computer Engineering at Gazi University. His main research interests are biomedical engineering, artificial intelligence, and machine learning and microcontroller applications.

E-mail: hknocal@gmail.com, **ORCID**: 0000-0002-8061-8059

Ozgur DUNDAR, PhD, works at Necmettin Erbakan University, Department of Astronautical Engineering. He graduated from the Electrical and Electronics Engineering Department of Selcuk University. He worked as an Automation Engineer for a while. His master's and doctorate degrees are from Selcuk University, Institute of Science and Technology, Department of Electrical and Electronics Engineering. Special fields of study are Automation, Robotic, Communication, Electromagnetic and Micro Strip Patch Antenna designs.

Email: ozdundar@erbakan.edu.tr , **ORCID:** 0000-0002-4142-4446

Hasan Berkant DUNDAR, He is a final year student of Electrical and Electronics Engineering at Istanbul Medipol University. He has worked on computer vision, machine learning, rf and microwave. He took part in international unmanned aerial vehicle races organized by TÜBİTAK.

E-mail: hbdundar42@gmail.com, **ORCID**: 0000-0002-8404-5059

Similarity Index

The similarity index obtained from the plagiarism software for this book chapter is 9%.

To Cite This Chapter:

Kocer S., Ocal H., & Dundar H. B., (2022). Use of IoT and Wearable Technology Design Fundamentals in Healthcare Industry, In S. Kocer, O. Dundar (Eds .), *Current Studies in Basic Sciences Engineering and Technology* (pp. 168–182). ISRES Publishing.

Swarm Concept and Swarm UAV Systems

Ozgur DUNDAR
Necmettin Erbakan University

Suleyman KAPLAN
Necmettin Erbakan University

Introduction

Swarm behavior; It is the behavior that causes living things with a colony structure such as birds, fish and some bacteria to clump together and move or move with the members of the swarm. The swarm behavior was first tested by computer graphics specialist Craig W. Reynolds in 1986 by examining bird flocks in a computer environment with a simulation program called BOIDS. The simulation allowed the swarm to act with simple factors according to the basic rules (Reynolds Engineering & Design). The ability of living things to solve problems that they cannot solve individually, as a swarm, is called swarm intelligence.

In swarm behavior, the movements are determined by a common intelligence, which indicates swarm intelligence. Swarm intelligence; If it is adapted to herd technologies, maximum efficiency and sustainability are obtained from the herd by dividing the same task among multiple herd members (Kaplan, 2021).

History of Swarm Technologies

In this part of the article, the general swarm systems in the field of robotics will be addressed as “swarm robot systems”. In the other subjects of the article, the member elements of the swarm robot systems (UAV, IDA and IKA) will be explained in detail and system distinction will be made. Swarm robot technologies are a field that examines swarm behaviors in natural life and tries to adapt them to robot systems. The field of swarm technologies is trying to adapt the swarm functions of living things such as ants, birds and fish in natural life, such as organizing among themselves, having an autonomous structure, cooperation and coordination, to robot systems as a flexible, scalable and robust structure. Swarm robot systems; it generally operates without a central entity to control the swarm members. Communication between swarm members is robot-to-robot or robot-to-environment (roboticsbiz.com, 2022). An example of herd life in nature is given in figure 1.

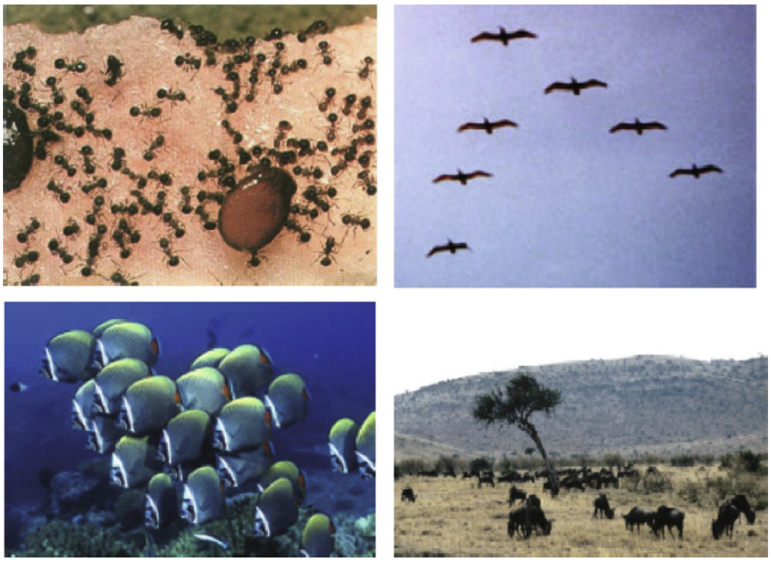


Figure 1. Swarm life in nature (Song et al., 2013).

In the field of robotics, the subject of "swarm" was used in G. Beni's work "The concept of cellular robotic system" in 1988, and in "Approach to the dynamically reconfigurable robotic system" by T. Fukuda and S. Nakagawa. As G. Beni has stated, cellular robotics operates autonomously in an n-dimensional environment and is coordinated with limited communication without a central entity. According to G. Beni, these robots cooperate in a coordinated way to achieve a common goal (Cheraghi et al., 2021).

T. Fukuda, on the other hand, stated that as a result of his studies, a group of robots that act with the logic of a swarm can come together like cells in the human body and do complex tasks. In 1989, G. Beni and J. Wang, using the term swarm intelligence at the seventh Japan Robot Society Meeting, claimed that robot systems could display intelligent behavior by coordinating them (Eberhart et al., 2001). Ronald Kube and Hong Zohng; In 1993, they built a multi-robot system (swarm robot) inspired by the behavior of natural swarms. Gregory Dudek and his teammates for the swarm robots in the same year; He worked on the concepts of swarm size, communication range and spacing, communication topology between swarm elements, communication bandwidth, swarm organization, swarm heterogeneity or homogeneity, which describe the abilities of swarm elements. In these studies, which can be considered the ancestors of swarm systems, the concepts such as swarm robot, multi-robot system or swarm systems are synonymous, but the use of these concepts in different names is the reason why they did not have a clear terminology at that time.

Beni, who studies swarm systems, suggested in 2004 that swarm members should have a simple, identical, scalable and local communication system. The concepts put forward by Beni form the basis of swarm systems even today (Cheraghi et al., 2021). A number of system examples are given in Figure 2.

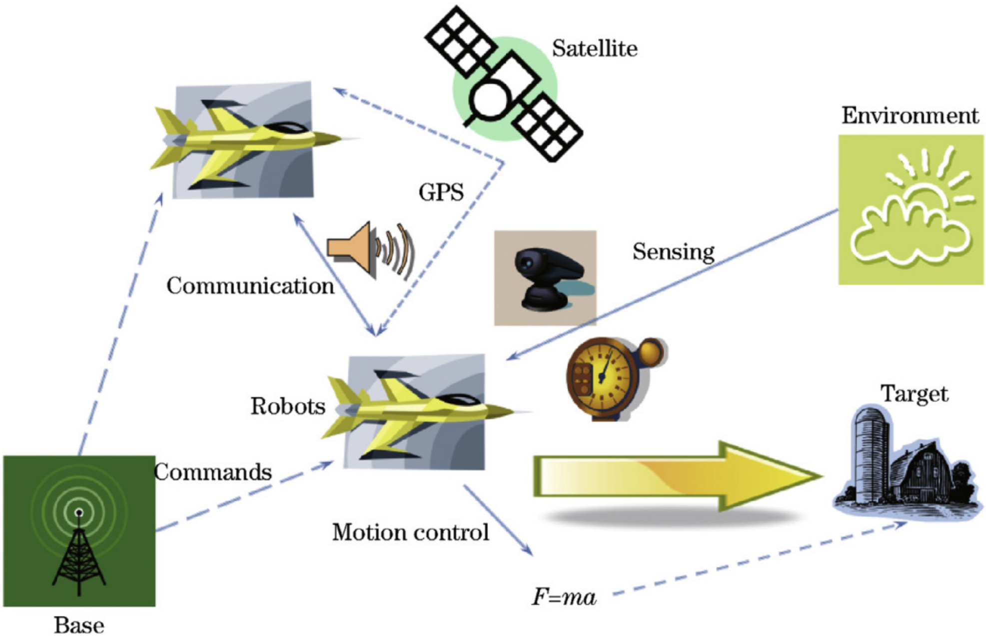


Figure 2. An example of a swarm system (Tan & Zheng, 2013).

Swarm Intelligence and Robotics

Swarm robots; It is a system that works for a certain purpose by using swarm intelligence software and swarm communication methods in coordination with robots that can work alone. For the efficient and successful implementation of swarm robots in the field, technology that can organize quickly and take on different formations is needed. This need has found itself in light, small, cost-effective and sacrificable robots (UAV, UGA and IDA) (Alemdar, 2021).

Swarm robots; they communicate with each other by swarm algorithms receiving and giving various data to communication units of swarm elements using swarm communication. In general, in swarm algorithms, swarm elements act according to local rules and the general behavior emerges from the interaction of swarm elements with each other or with the environment. According to the “Swarm Robotic Behaviors and Current Applications” study, by converting individual robots to the field of swarm robotics; It exhibits behavior based on a local rule set, which can range from simple reactive mapping between sensor inputs and actuator outputs to elaborate local algorithms. The behaviors in question include the interaction of swarm robots among themselves or with the environment (Schranz et al., 2020). Figure 3 shows an example array of swarm robots.

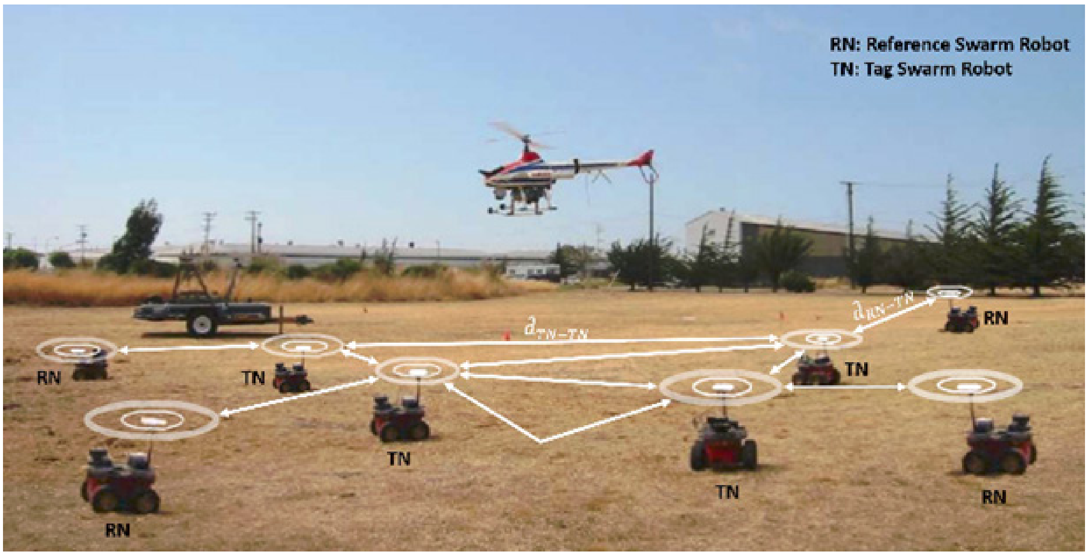


Figure 3. Swarm robots (Hanet al., 2015)

Interactions; It includes reading the data from the sensors of the swarm member, processing the swarm member's task computer or the computer of the relevant unit, and transforming the processed data into a response by the actuators (Floreano & Mattiussi, 2008). By repeating this interaction, the basic behavior of the herd is defined. Before entering the usage areas or working principles of herd robots, it is necessary to examine the basic principles of "herd intelligence", which scientists create by observing animal herds in nature and abstracting their behavior patterns.

Awareness

Awareness, which is one of the factors that form the basis of herd intelligence, is important for a good application of herd intelligence. In its interaction with the environment, the swarm member needs to instantly transmit critical elements or situations to the swarm of which he is a member, by scanning his surroundings in general threat/opportunity or friend/enemy dilemma. It provides awareness by detecting the threat/opportunity data individually at the right time and transmitting the relevant data to the other members of the swarm. Awareness determines the fate of the herd (Alemdar, 2021).

Autonomy and Self-government

The biggest difference of swarm intelligence from typical robotic hierarchy processes is that each element in the swarm can collect data on its own with its sensors and make decisions with the relevant data, and react with the decision taken by the microcontroller by using the actuators on the robot, which is the swarm member. The swarm algorithms of the swarm robots were previously developed with common procedures according to the swarm task. Although the roles of the herd members in the herd are similar, they may differ according to the task design of the herd. In swarm systems, in order to minimize the need for the swarm's communication system, it is generally requested that the swarm decide autonomously according to the instantaneous state of the swarm member without receiving a command from the system (Alemdar, 2021).

Solidarity

One of the situations in which the swarm-like communities and swarm intelligence give priority to their basic behaviors is the survival of the swarm. For this reason, every member of the swarm should give their basic reactions to the outside in solidarity with their swarm. Here, the survival of the individual herd is not important, but the survival of the herd itself. Another issue in solidarity is that the swarm immediately starts another task after performing the relevant task, and this is cyclical (Alemdar, 2021).

Figure 4. Circular swarm robots (Canciani et al., 2019)

Resiliency

Although the concept of Flexibility, which means the swarm's rapid recovery after the relevant situation, is referred to as "Resiliency" in the English literature, there is no exact equivalent for its use in swarm systems. In any disaster that may happen to the herd, the herd's realization of the current losses and the rapid improvement of the herd explains this concept within the herd system. No matter how heavy the loss or losses of the herd members are, the herd intelligence should analyze this situation and instead make a new role distribution within the other herd members or the herd. It should also adapt quickly to this new situation (Alemdar, 2021).

Swarm UAV: Robotic State of Bees

Swarm UAVs, a robotic copying of the behavior of bees between natural swarm systems; It consists of a large number of UAVs that perform the task in a coordinated way within the herd to perform a task with the operator or the decision mechanism within the herd (After et al., 2016)

The sensors of the swarm UAVs can be the same or different depending on the tasks, the autonomy levels and the platform types. In this case, we can use the homogeneous and heterogeneous classification that we use in swarm systems for swarm UAVs. Although the level of system complexity is low in homogeneous swarm UAV systems, sensor type, information processing capacity or flexibility capabilities according to tasks are more limited than heterogeneous systems. In heterogeneous swarm UAV systems, the diversity in the UAV mission capability and the high sensor capability in the payloads carried separately by the swarm elements are very beneficial to the swarm, but technical complexity is beneficial compared to homogeneous systems (Oskoei, 2014).

Basically, two control architectures are used in swarm UAV operations, "centralized" and "decentralized". The central control architecture of the UAV system in the swarm has a low level of autonomy and autonomy. There is no mutual communication structure between the UAVs in the swarm. The operator or operators of the system provide the coordination of the herd by getting information from each UAV system separately. For this reason, task assignment for the UAV system is made in advance. The advantages of low autonomy in the central control architecture are that the system is simple and easy to optimize, but the herd is less durable in case of system redundancy or communication problems.

On the other hand, the decentralized, that is, distributed control architecture, in the UAV system, the herd needs to have high autonomy and mutual communication. UAVs with this structure should be able to receive and share information by communicating in order to transfer sensor data among themselves. It should be able to make the necessary decisions in response to incoming information. In this case, the role of the operator in the herd moves to a higher level of management.

Distributed control architecture has a more complex structure than centralized architecture because they require a high level of autonomy. In a distributed or decentralized control architecture, since the tasks and information are distributed among the UAVs, the system is more durable, but it is more important because of its flexibility and redundancy. This makes it easier for UAVs to adapt to dynamic environments. The scattered herd; It is more resistant to damage that may occur in UAVs by cooperating within the system.

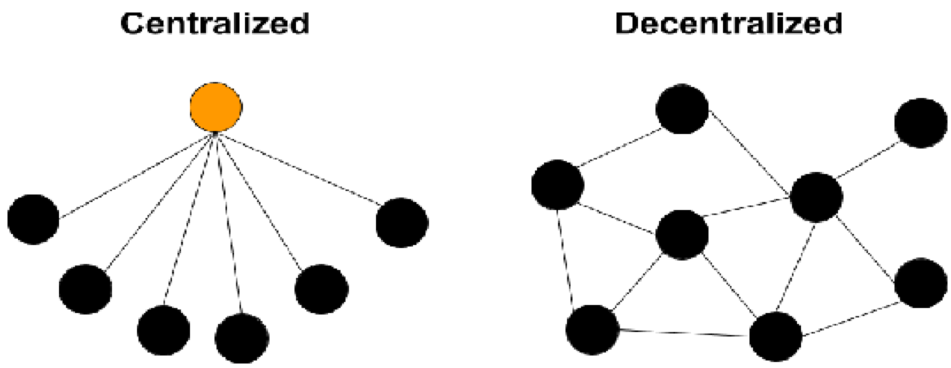


Figure 5. Central and Decentralized structures (Madey, 2013)

Formation Deployments in Herd Systems

Another issue in swarm systems is their deployment depending on the location and task. There are 3 deployment tubes, namely "Static Swarm", "Dynamic Swarm" and "Hybrid Swarm". In static swarm formation, swarm members are selected before the task by establishing trust in communication, trust and cooperation among swarm members by the system that controls the flight on the ground. In the static herd, participation in the herd is prevented from the beginning of the task, and new members are not included in the herd. In dynamic herd formation, new members are allowed to join or leave the herd at any time, either before or after the task. The addition of new robots from the same control mechanism or group to the herd is called Closed-Dynamic Swarm Formation. In the Open-Dynamic Swarm Formation, any third-party robot can be included in the herd and the relevant assignment can be made. In dynamic herd formation, difficulties such as secure communication and cooperation among herd members come to the fore (Akram et al., 2017). Hybrid flock formation allows static flock formation and dynamic flock formation to work together as a single body. In this formation, there is a flock with a static herd behavior in the center of the herd. The static swarm in the center allows other swarm members (UAV) to join the flock.

Communication and Interaction in Swarm UAV Systems

In this section, Swarm Systems and the internal and external interactions of Swarm UAVs, which are a structure of Swarm Systems, and communication ways will be examined. Coordination among herd members is among the important conditions for successfully performing tasks in herd systems. The transfer of the information received from the sensors of the swarm elements and the processing of this information and its distribution to the swarm elements are important for swarm communication. UAV systems included in swarm robots establish swarm interaction with wireless connection due to their fast and very flexible movement and wide task areas (Zhu, , 2015).

Current Swarm UAV systems include a ground control system that provides command and control. Here, the relevant data is transferred to the ground control system, providing information to the operator and the commands transmitted by the operator are delivered to the swarm elements. UAVs transmit location information (GPS, if available), speed and data from payloads to the ground control system as telemetry. In these systems, receivers and transmitters use unlicensed radio frequencies such as 900 MHz to transfer data (Campion et al., 2018). The general representation of swarm communication is given in Figure 6.

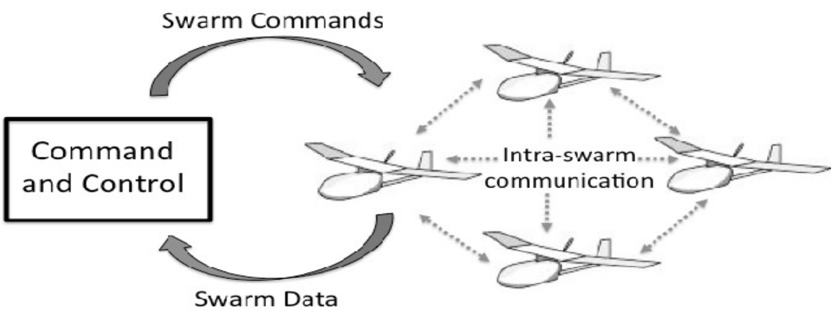


Figure 6. General display of swarm communication (Madey, 2013)

Swarm communication is divided into two as “Direct Communication” and “Indirect Communication” when we look at the communication and coverage area within the swarm. Direct Communication is used when swarm systems are from a single center and have limited coverage for the respective task. In this communication method, they communicate directly between the swarm elements and the ground control system or station without a base station. Communication modules such as ZigBee, which are used for direct communication, have low power consumption, low delay communication and low cost.

Indirect communication, on the other hand, provides communication between the UAVs in the herd and between the ground control station without distance restrictions. In this communication method, there are fixed base stations or mobile base stations between the herd and the ground control station, and the communication between the systems is ensured. Communication units such as GSM or satellite communication are used in indirect communication that enables long-distance data transfer. This situation increases the power consumed, the cost spent for communication and the delay between units in direct proportion to the increase in the coverage area.

The data transfer within the swarm UAVs is divided into three as Broadcasting, Inquiry and Synchronization. This distinction is outlined in the STM Thinktech February 2019 report titled "HOT UAV SYSTEMS: A Future Forecast for Modern Warfare". Broadcasting; It realizes the current state of the swarm element by sharing it with other swarm members. In Inquiry, the swarm element sends a message containing the relevant data from other swarm elements. Synchronization sends a message to the herd leader of the pack containing the synchronization request. Relevant swarm elements that receive the request broadcast the data to other swarm elements.

Swarm UAV and Battlefield

UAV systems first begin their historical process with the logic of remote control. It is accepted as the first use of UAV for military attack by Austria on August 22, 1849, by putting time-controlled fuse bombs in 200 balloons and sending them to Venice, Italy. "Ruston Proctor Aerial Target", the first unmanned aerial vehicle or aircraft, was designed in 1916. In 1917, the gyroscope-controlled automatic aircraft became the first official unmanned aircraft of the US Army (Kahveci & Nazlı, 2017).

Unmanned aerial vehicles, which are modern and comply with today's definition of UAV, have been developed since the 1960s. Ryan Model 147, an unmanned aerial vehicle developed by the USA, was used in the Vietnam War in the 1970s. UAVs, which send videos and pictures to the operator console where the UAV is controlled, were first used in the Gulf War in 1991 (Ekmekcioğlu & Yıldız, 2018). Turkey's UAV development process started with the agreement signed by the Undersecretariat for Defense Industries with TAI in 2004. The first automatic flight test of the successful Bayraktar Blok A, performed by Baykar Defense in 2009, made significant contributions to Turkey's national UAV development adventure (Aydın, 2017). Bayraktar TB2, which made its first flight on April 29, 2014; With its armed configuration, it was actively used as an armed force along with reconnaissance, intelligence and surveillance in Operation

Euphrates Shield, Operation Olive Branch, Operation Claw, Kıran and Peace Spring (URL-1).

Clustered use of drones was first seen in 2018 at the Russian-controlled Khmeimim Air Base in western Syria. The attack on the base was carried out by an anti-regime group. Russian personnel detected 13 UAVs that were preparing to attack at low altitudes and prevented them with EW & SHORAD systems (URL-2). On 14 September 2019, a clustered drone attack was carried out on Aramco's oil facilities in Abqaiq and Khurais in Saudi Arabia. The attack was carried out with 25 UAVs in the form of two sequential waves. Analysis after the attacks revealed that the facility in Abqaiq received 19 hits. The Saudi Arabian air defense layer, which consists of the MIM-104 Patriot and Crotale Short Range Air Defense Systems developed by France, could not stop the UAV attacks (Joshi, 2021).

Turkish Armed Forces; targeted the Syrian Armed Forces in the countryside of Idlib in 2020 with the use of clustered (Herd) armed UAVs. Using UAV systems in droves, the Turkish Armed Forces destroyed more than 200 targets, 5 helicopters, 23 tanks, 23 guns and the Russian-made BUK and Pantsir Air Defense systems in a short time (Urcosta, 2020).

References

After A. Kolling, et al .(2016). "Human Interaction with Robot Swarms," IEEE Transactions on Human-Machine Systems, Vol. 46.

Akram, R. N., Markantonakis, K., Mayes, K., Habachi, O., Sauveron, D., Steyven, A., & Chaumette, S. (2017, September). Security, privacy and safety evaluation of dynamic and static fleets of drones. In *2017 IEEE/AIAA 36th Digital Avionics Systems Conference (DASC)* (pp. 1-12). IEEE.

Alemdar, A. (2021, January 15). Suru Dronelar 'Askeri Doktrinleri' nasıl etkileyecek? DefenceTurk. https://www.defenceturk.net/suru-dronelar-askeri-doktrinleri-nasil-etkileyecek

Alemdar, A. (2021, January 15). Suru Dronelar 'Askeri Doktrinleri' nasıl etkileyecek? DefenceTurk. https://www.defenceturk.net/suru-dronelar-askeri-doktrinleri-nasil-etkileyecek

Aydın, F. (2017, December 20). Birçok alanda kullanılan İHA'ların Türkiye serüveni. GZT. https://www.gzt.com/jurnalist/bircok-alanda-kullanilan-ihalarin-turkiye-seruveni-2911729

Boids (Flocks, herds, and schools: A distributed behavioral model). (n.d.). Reynolds Engineering & Design. https://www.red3d.com/cwr/boids/

Campion, M., Ranganathan, P., & Faruque, S. (2018). UAV swarm communication and control architectures: a review. Journal of Unmanned Vehicle Systems, 7(2), 93-106.

Canciani, F., Talamali, M. S., Marshall, J. A., Bose, T., & Reina, A. (2019, May). Keep calm and vote on: Swarm resiliency in collective decision making. In *Proceedings of workshop resilient robot teams of the 2019 ieee international conference on robotics and automation (ICRA 2019)* (p. 4).

Cheraghi, A. R., Shahzad, S., & Graffi, K. (2021, September). Past, present, and future of swarm robotics. In Proceedings of SAI Intelligent Systems Conference (pp. 190-233). Springer, Cham.

Eberhart, R. C., Shi, Y., & Kennedy, J. (2001). Swarm intelligence. Elsevier.

Editorial. (2022, September 6). Advantages and history of swarm robotics (SR) explained. RoboticsBiz. https://roboticsbiz.com/advantages-and-history-of-swarm-robotics-sr-explained/

Ekmekciolu A., Yildiz M. (2018). The Use of Unmanned Aerial Vehicles in Military and Civil Fields : USA and Turkey Models and Sone Policy Offers. Türk İdare Dergisi. Sayı 486, ss. 169-227

Floreano, D., & Mattiussi, C. (2008). Bio-inspired artificial intelligence: theories, methods, and technologies. MIT press.

Han, Z., He, J., Geng, Y., Xu, C., Xu, L., & Duan, S. (2015, August). CRLB for TOA based near-ground swarm robotic localization. In 2015 IEEE 12th Intl Conf on Ubiquitous Intelligence and Computing and 2015 IEEE 12th Intl Conf on Autonomic and Trusted Computing and 2015 IEEE 15th Intl Conf on Scalable Computing and Communications and Its Associated Workshops (UIC-ATC-ScalCom) (pp. 1853-1858). IEEE.

Joshi, S. (2021). 'Drone swarms' are coming, and they are the future of wars in the air. The Print, 2.

Kahveci, M., & Nazlı, C. A. N. (2017). Unmanned Aerıal Vehıcles: Hıstory, Descrıptıon, Legal Status In The World And In Turkey. Selcuk University Journal of Engineering, Science and Technology, 5(4), 511-535.

Kaplan S., (2022, November 6). Takım Uydular: 21 Yüzyıl Uzay Oyuncuları. Dergiler. DefenceTurk. https://www.defenceturk.net/dergiler#dflip-df_75428/1/

Madey, A.G., and Madey G. R., 2013. "Design and evaluation of UAV swarm command and control strategies." Proceedings of the Agent-Directed Simulation Symposium. Society for Computer Simulation International. USA.

Oskoei E (2014). Swarm of UAVs: Search & Rescue Operation in Chaotic Ship Wakes, Master Thesis by, KTH Royal Institute of Technology.

Schranz, M., Umlauft, M., Sende, M. ve Elmenreich, W. (2020).Swarm robotic behaviors and current applications. Frontiers in Robotics and Artificial Intelligence. 7 , 36.

Song, Y., Hao, Q., & Zhang, K. (2013). Review of the modeling, simulation and implement of intra-body communication. Defence Technology, 9(1), 10-17.1

Srivastava, D., Pakkar, R., Langrehr, A., & Yamane, C. (2019, March). Adaptable UAV swarm autonomy and formation platform. In 2019 IEEE Aerospace Conference (pp. 1-6). IEEE.

Tan, Y., & Zheng, Z. Y. (2013). Research advance in swarm robotics. Defence Technology, 9(1), 18-39.

Urcosta, R. B. (2020). The revolution in drone warfare: The lessons from the Idlib de-escalation zone. Air Force Journal of European, Middle Eastern & African affairs, 3(2), 50-65.

URL-1: Türkiye'nin gökyüzündeki başarı imzası: Yerli İHA ve SİHA'lar. (2020, March

3). TRT Haber: Son Dakika Haberler ile Türkiye'nin Haber Kaynağı. https://www.trthaber.com/haber/gundem/turkiyenin-gokyuzundeki-basari-imzasi-yerli-iha-ve-sihalar-464449.html

URL-2: https://harkul.com.tr/blog/siha-suru-teknigi-ve-savasin-degisimi/

Zhu, X., Liu, Z., & Yang, J. (2015). Model of collaborative UAV swarm toward coordination and control mechanisms study. Procedia Computer Science, 51, 493-502.

About The Authors

Ozgur Dundar, PhD, works at Necmettin Erbakan University, Department of Astronautical Engineering. He graduated from the Electrical and Electronics Engineering Department of Selcuk University. He worked as an Automation Engineer for a while. His master's and doctorate degrees are from Selçuk University, Institute of Science and Technology, Department of Electrical and Electronics Engineering. Special fields of study are Automation, Robotic, Communication, Electromagnetic and Micro Strip Patch Antenna designs.

Email: ozdundar@erbakan.edu.tr , **ORCID:** 0000-0002-4142-4446

Suleyman Kaplan, I'm moving on with my education in Astronautical Engineering as a student. I graduated from the ACCESS Project, which was organized by the US Embassy in Ankara and TED College. For a short time, I was an R&D intern on the Technofest Rocket team. I wrote a magazine article on team satellites, launch systems, and space telescopes as a senior editor for Defence Turk. I graduated from the T3 Initiative Center Entrepreneurship Program with the designation of an entrepreneur. In my initiative, CarryVibe Logistics Technologies, I am a co-founder and team leader in business development, design, planning, and building connections. I'm presently developing myself in data science, machine learning, and an MBA.

Email : suleymankaplan7254@gmail.com , 0000-0002-3161-1977

Similarity Index

The similarity index obtained from the plagiarism software for this book chapter is 4%.

To Cite This Chapter:

Dundar, O. & Kaplan, S. (2022). Swarm Concept and Swarm UAV Systems. In S. Kocer. & O. Dundar (Eds .), *Current Studies in Basic Seciences, Engineering and Tecnology* (pp. 183–193). ISRES Publishing.

Role of Internet of Things in Modern Day Farming

Ravi Chandra
Research Scholar in Lovely Professional University

Krishan Kumar
Professor, lovely Professional University

I Sharath Chandra
Associate Professor, Matrusri Engineering College

Introduction

With the world's population growing by the day, according to UN Food and Agricultural Organization sources, the world needs to increase food production by 70% by 2050, compared to current statistics. Howver, with the shrinking of agricultural land and changes in the climate that support farming due to dwindling natural resources, the agricultural sector is experiencing a downtrend, with many traditional farmers looking to other sectors for sustainability, which is a concerning situation.

To address workforce issues, many of them began automating agricultural processes through the use of machinery and technology. One such technology is Internet of Things (IoT), which is used to fill gaps in the supply-demand chain. This technology paved the way for farmers to increase crop yield and profits while also addressing environmental concerns. IoT Technology in agriculture has resulted in advancements in the development of smart machinery, wireless connectivity, and the use of IT services.

Intelligent farming based on IoT technology assists farmers in increasing productivity by taking into account critical parameters such as keeping track of soil testing, understanding environmental conditions, the amount of fertilizer used, the number of times agricultural vehicles were used, and the use of resources such as water and electricity.

Assessing these parameters assists them in improving their business and increasing their efficiency. The use of sensor technology in conjunction with IoT enabled them to gain knowledge on crop health and monitor the condition of the field from any location. Because it has addressed many issues in traditional farming procedures, smart farming has become an obvious choice for many of them. In the following sections, we will look at how IoT has changed the agriculture sector in various ways.

Agriculture Sensors

This section focuses on various sensors used in agriculture, also known as agriculture sensors.Based on their usage and application, these sensors can be classified into several types. The table 1 below shows a variety of sensors and their applications.

Table 1. A Variety of Sensors and Their Applications [URL-1].

Agriculture Sensors	Usage and Application
Placement Sensors	The latitude, longitude, and angle of any location within the required area are determined by these sensors. For this, they use GPS satellite help and support.
Laser Sensors	These sensors measure the soil's characteristics using light. To measure the amount of clay, organic matter, and moisture in the soil, they are mounted on satellites, drones, or robots.
Sensors based on electro-chemical	These sensors aid in the collection of chemical data from soils by accurately identifying charged particles in the soil. They supply data in the form of pH and soil nutrient concentration.
Machine-based Sensors	These sensors are used to gauge mechanical resistance or soil compaction.
Detectors of Dielectric Soil Moisture	These sensors gauge soil's dielectric constant to determine moisture levels.
Sensors for Air Flow	Air permeability is measured using these sensors. They can be utilised in mobile or fixed positions.

Sensors of light

Optical sensors, as its name suggests, use light to detect the characteristics of the soil. They are able to ascertain the soil's organic matter, clay content, and moisture content. Satellites, robots, and drones all use this sensor. There are a variety of commercially accessible optical sensors with slightly different traits and structures.

Sensors based on electrochemistry

Two of the most important soil factors to examine are the pH levels and nutrient content. These sensors use electrodes to measure the voltage difference between two locations in order to calculate the maximum ion concentration like H+, K+, NO3-, and others. In order for this method to function, the electrode needs to be in contact with the soil sample. Chemical soil analysis is the customary process, which is time-consuming and expensive. ISFET stands for ion selective field efficient transistors and chemically modified electrochemically active area efficient transistors. ISE is for ion-selective electrode.

Sensors for location

Using GPS, location or inertial measurement units are utilised to properly map farms. Farmers use these tracking devices to determine where and how much pesticides and fertiliser to apply. This can be used to identify irregular landscapes, uneven ground, and leveling issues that cause water logging, among other things.

Sensors for airflow

Sensors that measure airflow are used to gauge soil air permeability (Figure 1). How well soil prevents air escaping all over it is determined by its permittivity. The kind, shape, and water content content of the soil can all be determined using this component. In these sensors, silicon chips measure temperature and heating. They adjust their direction and speed in response to wind and gases moving thru the soil.

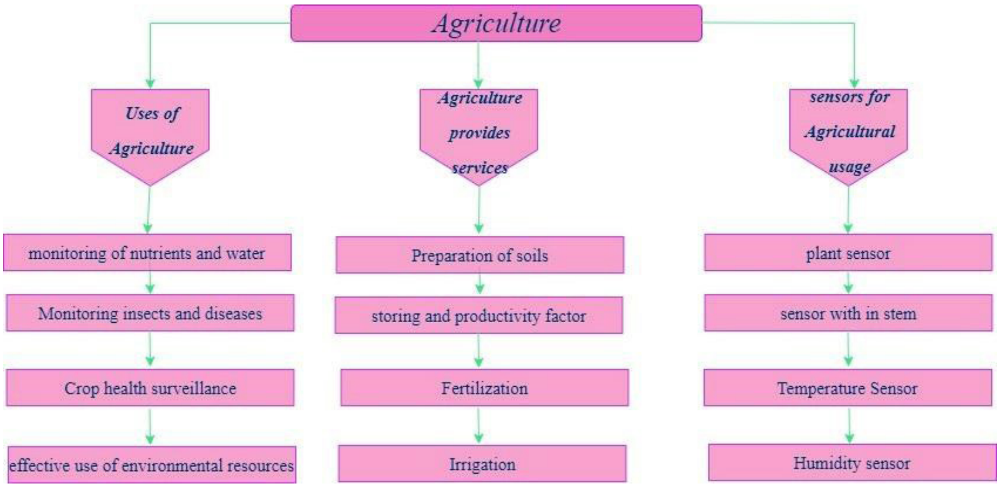


Figure.1. Agriculture related sensors, services, and applications

These sensors help with conserving and field mapping, which leads to higher productivity. They are easier to use and less expensive. Farmers can use these agricultural sensors to take an innovative and data-driven approach to each and every stage of the crop's life span, from seed production to cultivation. Field and IoT sensor data can be linked and incorporated into the centre console of any responsive system that delivers real-time farm analysis..

Drones for Agriculture

The recent agricultural revolution has been fueled by precision farming. Agronomists, agricultural engineers, and farmers are turning to unmanned aerial vehicles (UAVs) to acquire extra successful crop insights and more precisely plan and manage their operations, and crop monitoring from the air with farmland drones appears to be the next big thing. Global positioning system (GPS) technology and geographic information system (GIS) tools are heavily used in precision farming techniques to empower fine scale monitoring and mapping of yield and crop parameter data within fields. These provide more intensive and effective farming techniques, enabling farmers to modify fertilizer application recommendations or spot crop diseases early on. With more data at their disposal, farm owners can make choices based on both economic and environmental factors. For instance, increase the project and environmental savings can be actually realized by maximizing fertilizer care and attempting to implement only the proper level at the right time (URL-2).

There isn't a single agricultural mapping drone that works for everyone. Thinking about the particular use case is the first step in selecting the best drone for agricultural use. When choosing a drone for a particular application, it's crucial to take into account its cruise speed, form factor, payload capacity, physical range, and flight time with sensor payload. From emergency and early growth assessments to yield forecasting before harvest, durable drone technology adds value throughout the growing season. Drones typically fly at a height of 150 meters above the ground while taking pictures with cameras. The image resolution on the ground with this flight height and camera is approximately 30 cm per pixel. Then, for large fields, this data is revised down to one metre per pixel in order to shorten delivery time. The drones map crops at a rate of about three hectares per minute, depending on the size of the field.

Crop non-visible data are gathered by a camera with multi-spectral sensors in four different portions of the spectrum: green, red, red-edge, and near infrared. Agronomists are then able to provide a precise crop evaluation for each square meter of a farmer's field using the highly accurate maps created from this data. The amount of dry matter in the field and nitrogen absorption at crucial stages of crop development are determined using data from the drone in Figure 2.

Figure 2. Dusting/Spraying with Drones

When contrasted to farms studied utilising conventional, nondrone methods, drone utilisation led to an average higher yield of 10%. By 2024, 200,000 agricultural drones will have been shipped, and the market will be worth over $1 billion, predicts Global Market Insights. According to GMI, farmers' increased awareness of the advantages and disadvantages of employing drones in agriculture will contribute to the growth through 2024.

The following agricultural drones are offered for sale commercially:

- Precision Hawk's DJI Matrice 200 v2: Since working with DJI in 2016, Precision Hawk has prospered in the industry and is a well-liked drone option for farmers.
- The M200, according to Precision Hawk, is perfect for "the hardest farming settings" since it has sophisticated obstacle detection and can operate in below-freezing temperatures.
- SenseFly's eBee SQ is marketed as a "sophisticated agriculture drone," and the company's own emotion software, which makes developing a flight plan simple, is its key selling feature.
- In a single trip, the drone can record footage covering hundreds of acres.
- Sentera PHX Complete System: The PHX can gather data from 700 acres and features a lengthy unidirectional communication link.

Farmers can increase operational effectiveness and carry out plant health analytics thanks to SenetraUnmanned aerial vehicles (UAVs) are quickly becoming acknowledged as a more precise and costeffective alternative to satellites and manned aircraft, which have historically been employed to monitor farmland. Studies show that even on overcast days, drone imagery has a greater rate of accuracy and resolution. Accurate crop health checks utilising UAVs can be done all year round, but traditional terrestrial approaches to data collecting in difficult weather conditions could possibly delay projects for days.

Livestock Monitoring

Farmers can increase operational effectiveness and carry out plant health analytics thanks to SenetraUnmanned aerial vehicles (UAVs) are quickly becoming acknowledged as a more precise and costeffective alternative to satellites and manned aircraft, which have historically been employed to monitor farmland. Studies show that even on overcast days, drone imagery has a greater rate of accuracy and resolution. Accurate crop health checks utilising UAVs can be done all year round, but traditional terrestrial approaches to data collecting in difficult weather conditions could possibly delay projects for days.

Figure 3. Monitor and manage huge livestock

Common issues in the livestock sector include animals ingesting harmful materials like plastic and getting sick, a high animal fatality rate, no record of livestock if there are a lot of them, and if a farm animal goes missing there is no way to know where they went and therefore no record.A cuttingedge solution called a livestock monitoring system is created and constructed utilising sensors, GPS, and other technologies that are all combined with a protocol stack for communication.

Using this surveillance device, farmers can keep an eye on their livestock from a distance. The tracker gadget, which can be fastened to collars, can keep an eye on the wellbeing of the animals as well as their chewing habits, geographic information, and pasture management. The Internet of Things-enabled cattle monitoring system could be a boon to livestock farming. It aids in the administration of farm animals as well as other farm equipment.

IoT-enabled herd management solutions provide data on numerous facets of cattle health.

Animals' position, temperature, pulse rate, and heart rate are tracked via a wearing collar or tag with sensors, which wirelessly communicates the information to farmers' devices in almost real-time. Each of the farm animals has a movable tracker on them. To avoid sticking out from the animal's body, the tracker gadget is made to be compact and low in weight. The device includes a GPS tracker and sensors that enable round-the-clock monitoring of the animals' movements and well-being. The technology is designed to monitor the behaviour of livestock.Farmers can use the system's behavioural tracking capability to see what their cattle are consuming. In addition to the equipment, there are mobile apps, online apps, and IoT Dashboards. Farmers may safeguard areas where cattle can wander freely by using the smartphone's associated application to establish virtual borders with geofencing. The tracker notifies the mobile application if the cattle departs from the predetermined bounds.

If the livestock ingest something harmful or unsuitable, it will also transmit alarms. In this way, farmers mayprevent disease in their cattle. The system can run on LPWAN, which is widely available and guarantees proper connectivity. LPWAN protocols come in a variety, and the one picked depends on the application.

Farms today have a lot of animals, making direct connection between personnel and all of the animals unfeasible. It can be used to track farmed equipment to avoid theft. IoTenabled livestock monitoring makes it possible to remotely monitor all animals and offers reliable data on each one on the farm.

GPS Tracking of Livestock Movement

These days, it's crucial to observe animal behavior and their movements. The whereabouts of cattle may be followed around-the-clock, every day of the week, thanks to the livestock monitoring system. A lost animal can also be found and saved right away. The solution makes it simple to set up grazing borders for cows and other farm animals by integrating a GPS tracker with a geofencing capability.

Monitor Livestock Health

In order for us to obtain nutritious goods from cattle, the animals must be in good health. The cattle remote monitoring assists in keeping track of the wellbeing of the cattle. To make smarter choices about cattle weight, timing of meals, and other issues, the feeding pattern can be observed. It is possible to monitor the time and behavior of the medication as well as the birthing, lactation, and other events if cattle become ill, such as lameness. If livestock are properly fed and digested, they can produce milk that is more nutrient-rich. Animals can receive better care if feverish symptoms are caught early thanks to temperature sensors.

Lower the Livestock Mortality Rate

Farmers may force their animals to consume antibiotics or steroids in order to obtain additional products from farm animals. These antibiotics can harm livestock in ways that can result in the animal's death. With the livestock monitoring system, you can track your livestock's eating habits and provide them with nutritious feed to produce better results. In addition, if farm animals become ill, the illness can be detected and treated before it is too late. In this way, the monitoring system contributes to lower livestock mortality rates.

Livestock Identification and Better Security

With livestock monitoring systems, animal identification becomes easier. You can easily track birth information, breed, identification marks, and so on. The livestock's production history can be tracked. The tags make it easier to identify the livestock and improve their security. This can reduce the likelihood of theft and lost cases (URL-3).

Pasture Management Improvement

The livestock monitoring system can help to improve the pasture or grass that animals eat. Farmers can track the movement of the animals as well as the area in which they graze. Animals will eat better grass as a result, and we will get better products from them. It will increase the profit earned by farmers from livestock

GEO Fencing

Crop destruction caused by wild animal intrusion into farm fields has recently become a common challenge for farmers all over the world. Wild animals straying into farm fields and squashing crops are becoming more common, posing an economic challenge for farmers. Animal intrusions cause crop damage; therefore, farmers must be constantly vigilant to protect their crops. Farmers suffer a significant loss of crops and resources required for crop expansion as a result of this. As a result, preventing animal aggression is critical for crop protection. This problem can be addressed with the aid of a WSN (Wireless sensor network). Together with the IoT (Internet of Things), the sensor network lets the farmers stay connected to their farms anytime, anywhere. Emerging IoT technology with IoT devices can be used to capture wild animal detection and environmental information for human observations through the smartphone application. The IoT enhances things to be detected and/or remotely controlled through the current network infrastructure, generating opportunities for more direct interaction with the physical environment through computer-based systems and contributing towards precision, performance, and economic benefits.

Smart Green Houses

The world's agriculture sector is under intense strain as a result of climate change, diminishing resources, and rising population. It should come as no surprise that growers are utilizing cutting-edge techniques to enhance crop resiliency and operational efficiencies as the degree of unpredictability rises. Smart greenhouses serve as a perfect example of how the Internet of Things (IoT) is much more commonplace nowadays in agriculture. A greenhouse offers a regulated environment that is tailored to the needs of the vegetation being grown within. Microclimate and agronomic factors have historically been recorded somewhat manually and unevenly. What can be quantified has a limit, and farming operations follow a predetermined, speculative schedule. On the other hand, the climate in the greenhouse is constantly influenced by changing weather patterns and "invisible" factors like doors open or early-stage infections that could harm crops. Smart greenhouses, which are outfitted with contemporary sensor and communications technology, automatically record and disseminate information on the environment and crop around-the-clock. In order to find bottlenecks and anomalies, collected data is sent into an Iot ecosystem where analytical algorithms transform it into useful insight. As a result, on-demand control over HVAC, lighting, and irrigation as well as spraying operations is possible. The creation of predictive models to evaluate the risks of agricultural disease and infection is made easier by continuous data monitoring. A smart greenhouse enables growers to optimize yield rates while reducing labor requirements and increasing resource and chemicals use efficiency in Figure 4 (URL-4).

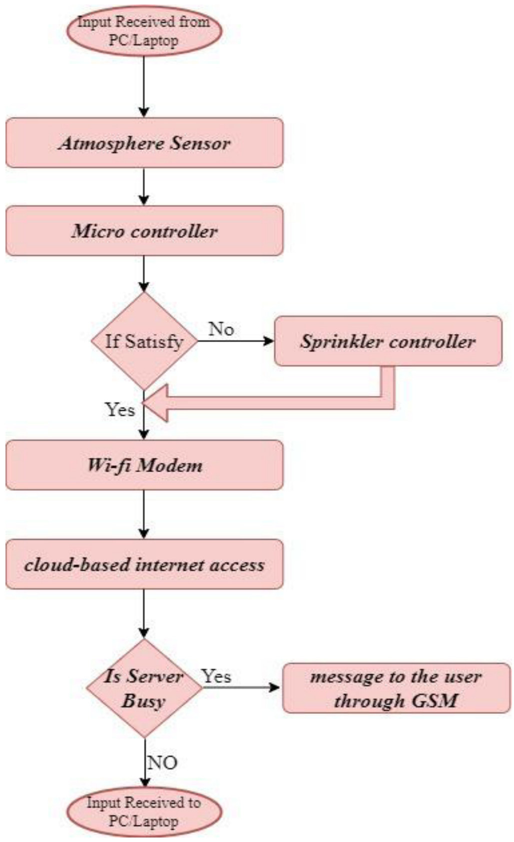


Figure 4. Flowchart of Smart GreenHouse

Keep optimum microclimatic conditions.

Farmers may gather multiple data points with previously unheard-of granularity thanks to IoT sensors. In the greenhouse, they offer real-time data on important climate variables like temperature, humidity, light exposure, and carbon dioxide. The ideal conditions for plant growth are maintained while promoting energy conservation thanks to the data that prompts pertinent adjustments to HVAC and lighting settings. To maintain a strictly controlled atmosphere, motion/acceleration sensors assist in identifying doors that are mistakenly left open.

Improve Fertilization and Irrigation Techniques

Farmers can monitor agricultural status as well as environmental measurements thanks to intelligent conservatories. This ensures that fertilization and irrigation practices are in accordance with the actual requirements of grown plants for best yields. Soil volumetric water content measurements, for instance, can indicate that crops are under moisture stress. Similar to that, measurements of soil salinity assist determine whether fertilization is necessary. With the least minimum of user input, watering and sprinkling systems can be programmed to respond automatically to the group's needs in real time.

Infection control and disease outbreak prevention

Harvest disease is a recurring problem for farmers, with every epidemic having a negative impact on crop margins. Although farmers have access to agrochemical treatments, they frequently don't know when to use them. A failure to utilize treatments could result in harmful disease outbreaks, whereas applications performed too frequently create ecological, safety, and financial problems. Data on greenhouse settings, outside weather, and soil properties are combined with a machine learning platform to gain significant insights into current insect and fungal concerns. By using this knowledge, farmers may use the least amount of chemicals while still providing a healthy crop.

Improve Security and Prevent Theft

Crops grown in greenhouses with high value are a tempting target for burglars. Many farmers lack a reliable security system since installing typical CCTV surveillance networks is expensive. In this situation, Wearables in smart greenhouses offer a cost-effective infrastructure to track door status and spot shady activity. When a safety issue occurs, they immediately alert growers because they are connected to an automatic alarm system.

Components of an Intelligent Gardener

What therefore is required to put a smart greenhouse in place? There are a few crucial aspects you need to consider:

a) Battery-powered, low-power sensors that can record different meteorological, agronomic, and safety datasets

b) Trustworthy, affordable wireless communication to send information from cross-greenhouse detectors to a distant gateway

c) A diagnostic deep learning platform that extracts information from sensor data and presents it on preferred UIs to enable decision-makers to make well-informed decisions about farming activities. The platform can be quickly integrated with conservatory control systems that have already been put in place to get HVAC, illumination, watering, and sprayer networks working on their own.

Because they are spread out over large geographic areas, advertising greenhouses facilities need protracted wireless connection with excellent penetration capabilities. By using this method, a gateway can be placed closer to the power source, saving money on trenching for wires while assuring dependable data delivery. Scalability, which minimizes the number of ports necessary to reduce the cost of hardware, installation, and operation, is another factor to consider.

Predictive Analytics

The agricultural value chain provides a wide range of use cases and specific applications for analytics. Analytics solutions make use of Big Data, IoT, Cloud, GPS, and other technologies to produce pertinent data from which actionable insights can be derived. Farmers and financial institutions may now make much better information decisions thanks to this.Basic data from Information Management Systems should be requisitioned to unleash the potential of analytics. In order to reap the benefits, this will ultimately encourage smart farming and better decision-making. The benefits of analytics-enabled smart farming are outlined in the framework shown in Figure 5.

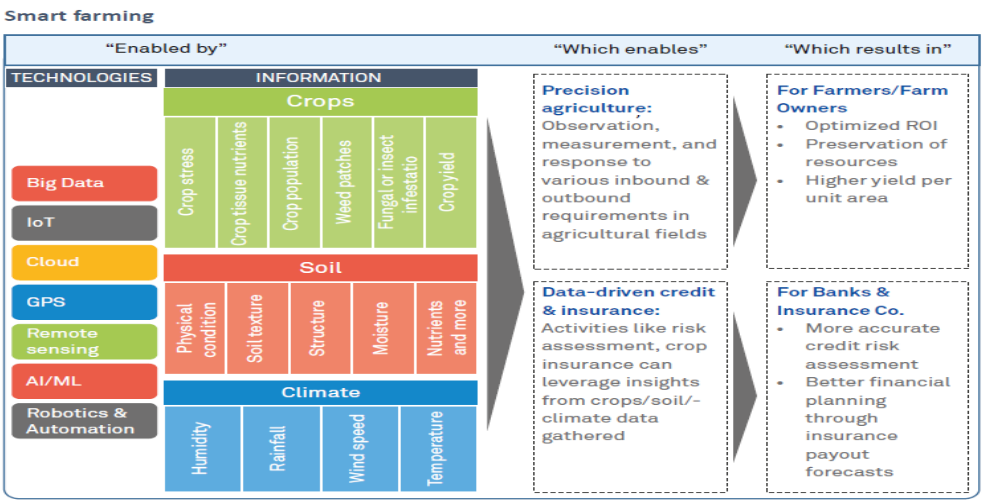


Figure 5. Information of Management system

Challenges in Building The IOT Platform

IoT technology and smart farming both demand constant internet access. In many other parts of the world as well as underdeveloped nations, this is not accessible. Market speculation holds that consumers aren't always eager to acquire the newest Internet of Things gadgets outfitted with agricultural sensors. There aren't always the necessary pieces of infrastructure, such smart grids, transportation systems, and cell towersThis makes it more difficult for its use to increase. Above all, the link needs to be very energy-efficient so that sensors can be left in place for years with little need for upkeep.

Conclusion

According to FAO forecasts, the globe would need to produce 70% extra foodstuff in 2050 as the human population grows exponentially. As a result, increasing crop yields becomes crucial. As a result, IoT may play a significant role in this process.Farmers and ranchers can now practise smart farming thanks to the use of IoT. Smart farming has two advantages: farmers can spend less time in the fields while increasing crop yields.

In a nation like India, in which farming is the foundation of the business and the industry is founded on it, this contemporary farming strategy has the ability to usher in a new era of agricultural revolution and pave the way for progress. Additionally, the utilization of various sensor types in IoT is illustrated using general scientific principles. This study's objective was to introduce the Internet of Things and discuss how it might be applied to agricultural practices. The main objective was to provide a more thorough explanation of the various agricultural sensors utilized to gather data. The advantages of many frequently made sensors were briefly discussed. The report also covered the uses, advantages, disadvantages, and restrictions of agricultural sensors.

References

Sai Xu, Xiuxiu Sun, Huazhong Lu, Hui Yang, Qingsong Ruan, Hao Huang and Minglin Chen, (2018) Detecting and Monitoring the Flavor of Tomato (Solanum lycopersicum) under the Impact of Postharvest Handlings by Physicochemical Parameters and Electronic Nose, MDPI,.

A. Bseiso, B. Abele, S. Ferguson, P. Lusch, and K. Mehta, (2015) "A decision support tool for greenhouse farmers in low-resource settings," in Global Humanitarian Technology Conference (GHTC), 2015 IEEE, pp. 292–297.

K. A. Czyzyk, S. T. Bement, W. F. Dawson, and K. Mehta, (2014) "Quantifying water savings with greenhouse farming," in Global Humanitarian Technology Conference (GHTC), 2014 IEEE, pp. 325–332.

Z. Xiaoyan, Z. Xiangyang, D. Chen, C. Zhaohui, S. Shangming, and Z. Zhaohui, (2013)"The design and implementation of the greenhouse monitoring system based on gsm and rf technologies," in Computational Problem-solving (ICCP), 2013 International Conference on,, pp. 32–35.

M. Mahdavian, M. B. Poudeh, and N. Wattanapongsakorn, (2013) "Greenhouse lighting optimization for tomato cultivation considering realtime pricing (rtp) of electricity in the smart grid," in Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 2013 10th International Conference on,, pp. 1–6.

Demmers, T.G., Gauss, S., Wathes, C., Cao, Y., Parsons, D., et al., (2012). Simultaneous monitoring and control of pig growth and ammonia emissions. In: The Ninth International Livestock Environment Symposium (ILES IX). International Conference of Agricultural Engineering-CIGR-AgEng 2012: Agriculture and Engineering for a Healthier Life, Valencia, Spain, 8-12 July 2012. CIGR-EurAgEng.

Dr. K. Gayathri Devi, N.Sowmiya, R.K.Yasoda, Dr.K.Muthulakshmi, B.Kishore, (2020) Review On Application Of Drones For Crop Health Monitoring And Spraying Pesticides And Fertilizer, Journal of Critical Reviews, ISSN- 2394-5125, Vol 7,Issue 6,.

S.R.Kurkute, B. D. Deore, Payal Kasar, Megha Bhamare, Mayuri Sahane, (2018) Drones for Smart Agriculture: A Technical Report, International Journal for Research in Applied Science & Engineering Technology (IJRASET), ISSN: 2321-9653, Volume 6 Issue IV,.

[URL-1]: https://www.analyticssteps.com/blogs/5-applications-iot-agriculture

[URL-2]: https://tropogo.com/blogs/application-of-drones-in-agriculture-in-india

[URL-3]: https://www.digitalmatter.com/applications/livestock-tracking/

[URL-4]: https://www.community.ruggedboard.com/post/iot-smart-agriculture

About the Authors

B Ravi Chandra, (Ph.D), He is Research Scholar in Lovely Professional University, Punjab. He is Postgraduate in Digital Electronics and communication System from Narayana Engineering College, Gudur Nellore. He completed his graduate. Currently working as an Assistant Professor in G Pullaiah College of Engineering Technology, Kurnool.
Email: ravichandraece@gpcet.ac.in, **ORCID:** a000-0002-7812-931X

Krishan Kumar, Dr Krishan Kumar received B.Tech degree and M.Tech Degree in ECE from Government institute. Ph.D. degree From Thapar University Punjab India. His research interests include Wireless communication.
E-mail: krishan.22397@lpu.co.in,**ORCID:** 0000-0003-1680-820X

I Sharath Chandra, Ph.D. degree From GITAM University Visakhapatnam, currently working as associate professor in Matrusri Engineering College in Hyderabad, Telangana
E-mail: drsharath@matrusri.edu.in ,**ORCID:** 0000-0002-3077-0060

Similarity Index

The similarity index obtained from the plagiarism software for this book chapter is 11%.

To Cite This Chapter:

B Ravi Chandra, Krishan Kumar, & I Sharath Chandra, (2022). Role Of Internet Of Things In Modern Day Farming, In S. Kocer, O. Dundar (Eds .), *Current Studies in Basic Sciences Engineering and Technology* (pp. 194–205). ISRES Publishing.

Current Trends on Wearable Antennas, Fabrication Techniques and Materials

Hulusi ACIKGOZ
KTO Karatay University

Ilayda CAN
KTO Karatay University

Introduction

We are witnessing a rapid increase of Information and Communication Technologies (ICT). The tremendous growth of new services and demand has led to the design and development of new devices which can satisfy various demands in terms of shape, size and functionality. Undoubtedly, the Internet of Things (IoT), generic term to describe the ability of physical objects equipped with sensors and processing units to connect and exchange data with other devices, has been decisive for the deployment of those equipment. On the other hand, with the deployment of 5G, a new wireless mobile network standard, everything can be connected with higher data rate and speed, lower latency and more reliability. The spread of 5G networks is essential for the full usage of IoT and virtually connects everything (or everyone) to everything. According to (Galov, 2022), the number of connected devices will be around 29.4 billion by 2030.

Besides, academic studies on antennas, which are indispensable elements of wireless communication and internet of things applications, continue intensively. Along with the developing technology, it is important that antennas are portable on the user in terms of not interrupting the communication, as well as being able to work with high gain in the newly allocated frequency bands. Here, portable means that the devices are part of the user body, clothes, or accessories, rather than being carried in their hands or in pockets like mobile phones or tablets. In fact, the word *wearable* is more often preferred to the word *portable* and it is used for any electronic device that is worn as an accessory or embroidered into clothing or implanted under the skin or tattooed on the skin. Wearable antennas can be integrated into many portable devices worn on the user, as shown in Figure 1. These devices can be, for example, smart glasses, smart belt, as well as smart shirts and trousers (Rodrigues et al., 2018).

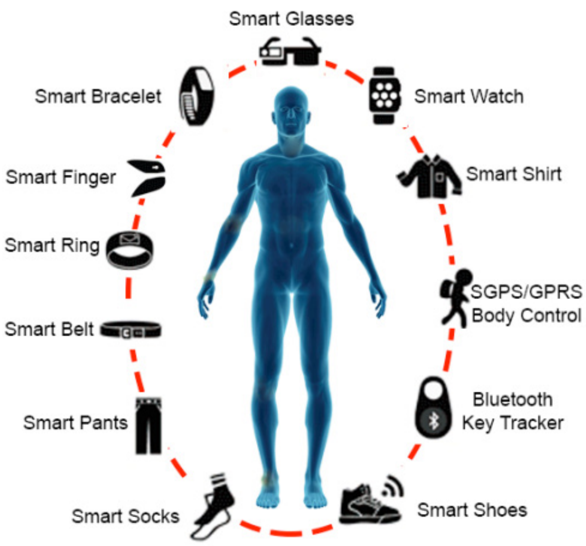


Figure 1. Wearable technologies that can be worn on the user (Rodrigues et al., 2018)

Wearable technologies are used in many areas today. They are frequently used in applications such as monitoring the physical activities of people (Tokucoglu, 2018) and patients (al Bassam et al., 2021), and monitoring the life signals of elderly people (Baig et al., 2013) etc. For instance, measuring body health signals such as cardiac rhythms, blood pressure, temperature and transferring them to a central hospital instantly is vital in the follow-up of patients (al Bassam et al., 2021; Dias & Cunha, 2018; Roudjane et al., 2018). On the other hand, in sportive activities, wearable technologies are used to obtain location and direction information of athletes, monitor their performance and track targets, as well as monitor their health status (Tokucoglu, 2018). Likewise, in the defense industry, wearable technologies are designed and developed for purposes such as monitoring the health status of soldiers, navigation and communication (Sharma et al., 2020). In the next five years, the military wearable technologies market is expected to reach 3.4 billion USD (Marketsandmarkets, n.d.). In mining, wearable technologies are of vital importance to instantaneously locate workers with wireless communication and, if necessary, to evacuate them from dangerous areas.

One of the most important components of a wearable device is the antenna part used to send and receive signals to be processed later. Depending on the application and the type of wearable device, the antenna needs to be tailored to extract the most useful information from the sent/received signal with the lowest degradation. If the wearable device is embedded in the clothes of the user, the antenna should be made from flexible materials for achieving a highly embedded device that is not obtrusive. However, the flexibility of the antenna that has to be considered when designing an antenna is also a factor for degradation of the signal. In parallel, the antenna should be low profile for compactness purposes. Additionally, the antenna designer should pay attention to the proximity of the antenna to the body that is likely to degrade the performance of the device.

This book chapter is dedicated to reviewing current trends on wearable antennas. The second part of this review will be on a brief history and applications of wearable devices. Then, the antenna, an essential element of wearable devices will be reviewed. In particular, some of the different fabrication processes existing in the literature and materials used for the antenna fabrication will be explored.

Brief History and Applications of Wearable Antennas

Although some consider the introduction of wearable devices back to the Middle Ages, to the invention of the first glasses, modern wearable technologies reached their current usage with the development of computers in the middle of the 20th century (Ometov et al., 2021). The rapid development of portable computers has brought different applications in many other fields. The emergence of the first hearing aids in the 80s with the increase in the functionality of wrist watches in parallel to technological developments can be shown as the first examples of wearable technologies. The 70's and 80's were also the years when wristwatches with calculators were introduced to the market. Examples are Hamilton Pulsar (1975), HP Algebraic Calculator (1979) and Seiko Data-2000 (1983).

The advances of wearable devices have been boosted by the introduction of the Bluetooth Technology (BT), a protocol for connecting different devices wirelessly (Ullah, 2009). This enables remote control and wireless transfer of information between any two or more wearable devices equipped with the BT. This highly popular technology stands out because it is cheap, easy to use and does not have to be on the line of sight of communication (Verma et al., 2015). In 1999, the first Bluetooth headset was released as one of the first wearable devices using the Bluetooth technology to communicate with a cell phone (Seneviratne et al., 2017).

In parallel, we witnessed the development of a new type of wearable device which can be mounted on the head next to the eye of the user and displays images as if it were seen by the eye. One of the early initiatives, the so called EyeTap project (Mann et al., 2005), is considered as the father of the first prototype of Google glasses which were available to the public in 2014 (Wikipedia contributors, 2022). The ambition of Google was to build a head-worn computer that will help to "explore and share the world". These new technologies contributed to the widespread use of Augmented Reality (AR) technologies (Carmigniani et al., 2011).

The existing commercial wearable devices may be classified as: accessories, e-textiles, and e-patches. The above-mentioned devices are considered as accessories because they are not part of the clothes nor the body of the user. These are accessory devices in the form of wristbands, wristwatches and glasses. They can be worn and removed at will. On the contrary, e-textiles are the product of the integration of electronics with the traditional textiles (Komolafe et al., 2021). Thus, sensing physical activities, communicating with other electronic devices can be easily performed. For delivering drugs to the patient and for monitoring biometric parameters such as body temperature, heart rate etc., e-patches are employed (Vijayalakshmi et al., 2018). These devices are placed directly on the skin of the user.

Although one the first e-textile devices (electrically heated gloves) was released in 1910, most of them have been developed after 2000, triggered by the rapid rise of modern technology (Hughes-Riley et al., 2018). Technologically more advanced e-textile devices consisting of electronic devices incorporated into pockets of the clothing have been proposed (Heilman & Porges, 2007). Heart rate monitoring sport bra (Navalta et al., 2020), ECG sensors in garments (Uz Zaman et al., 2020; HealthWatch contributors, n.d.), smart socks equipped with textile sensors to count footsteps, distances etc. (Sensoria contributors, n.d.) are some of the many examples that can be found in the literature. For more applications of wearable e-textiles, readers can refer to (Uz Zaman et al., 2022; Stoppa & Chiolerio, 2014; Komolafe et al., 2021; Dacuna & Pous, 2009).

Wearable Antenna Fabrication Techniques and Materials

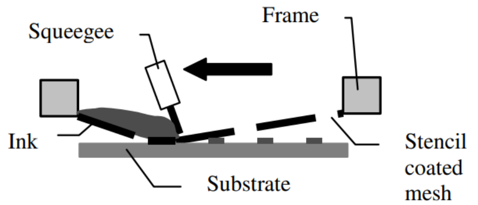
An antenna is a key component of a wearable device. It allows the device to communicate with other surrounding equipment and to get rid of the cable bundle if the communication must be made with other on-body devices. Although, theoretically an antenna can be in various forms, it should obey certain criteria that will qualify it as wearable. For instance, to be unobtrusive, its overall size should be as small as possible. Thus, as opposed to bulky antennas, like horn antennas, its thickness should be the lowest possible. These antennas are said to be planar or low profile antennas. Low profile antennas are widely used in applications such as RFID localization (Dacuna & Pous, 2009), communication (An et al., 2018), biomedical (Islam et al., 2022) etc. These low-profile antennas are generally made of a substrate under which a metallic ground plane exists and a metallic patch to radiate the electromagnetic wave to the surrounding medium. Depending on the material of the substrate and the fabrication process, these antennas can be characterized as flexible, conformable etc. Below, various wearable antennas will be presented. They will be examined within two categories: the process of fabrication and materials used for their manufacture.

Fabrication Techniques

Various techniques have been considered for the fabrication of wearable antennas. Among other techniques, screen printing (Blayo & Pineaux, 2005), inkjet printing (H. R. Khaleel et al., 2013; Le et al., 2013) and embroidery (Tsolis et al., 2014) methods are probably the most popular ones.

The antenna can be easily and economically printed employing the screen printing technique (Figure 2a). It consists of transferring the ink to the substrate through a stencil. As shown in the figure, a squeegee blade is used to force the ink downward. Despite the advantages underlined above, this process has some disadvantages. Because of the nature of the process, it leads to printed patterns with low resolution and the thickness of the layer deposited can not be controlled. This can be a problem for antenna applications where precise thickness is needed (el Gharbi et al., 2020).

The inkjet printing method is an alternative printing technology (Figure 2b). Small amount of droplets, in the order of 1 pL, are deposited on the surface of a substrate either using the continuous inkjet or the drop-on-demand technique. The amount of liquid deposited to the substrate is controlled by a computer and directly deposited to the position where the antenna pattern is. Thus, highly precise patterns can be printed without the need for a mask. This makes the technique very efficient and economic when compared to other techniques. In the drop on demand process, a single drop of the ink can be ejected by increasing the pressure within the reservoir either by a piezoelectric actuator or a resistive heater. Most electronic boards manufacturers use the drop on demand technique. Although many other types of metal inks such as copper or gold exist, silver nano-particle ink is prefered for printing conductive circuits because of its relatively low curing temperature and high electrical conductivity (Kim, 2020).



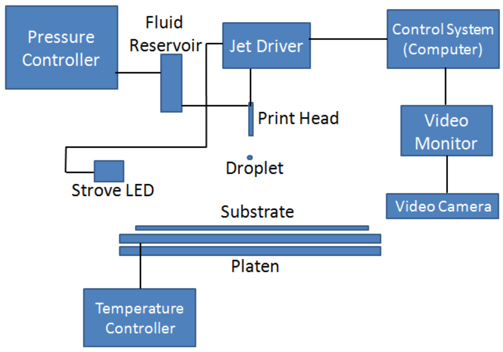


Figure 2. Printing processes, a) screen printing (Blayo & Pineaux, 2005), inkjet printing (H. R. Khaleel et al., 2013; Le et al., 2013)

In recent years, embroidering has become popular among scientists and engineers for fabricating wearable antennas (Tsolis et al., 2014). It uses conductive threads and specialized computer assisted machines to sew the metallic part of the antenna on a fabric substrate. The process of embroidering is shown in Figure 3. First, the antenna geometry is drawn on a computer, then after the digitization step, it is sent to the computerized embroidering machine that is purposely used to embroider the antenna on a textile. This technique is the prefered one when it comes to the fabrication of textile antennas since there is no need to use adhesive material to glue the patch to the substrate. In fact, as will be discussed below, using a third material between the patch and the substrate may change the antenna physical properties and deteriore the antenna performance. There are several criteria that have to be considered when using embroidering technique to fabricate textile antennas. First, the conductivity of the thread has to be as high as possible. For this reason, silver threads are mostly used for embroidering the antenna conductive parts on the substrate. Second, the threads should be strong enough but flexible at the same time so as not to split under hard working conditions (Ali et al., 2020).

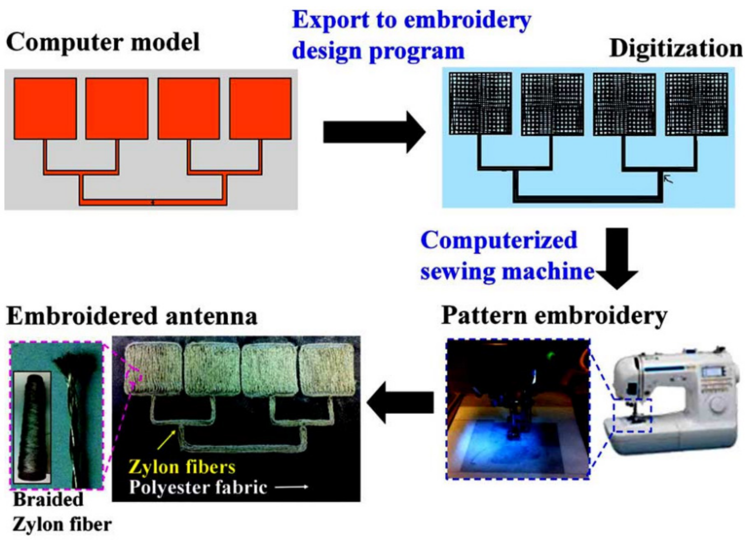


Figure 3 Process of fabrication of the embroidered antenna (Wang et al., 2012)

Recently, Laser Direct Structuring (LDS) method is gaining attention for the fabrication of conformable and wearable antennas. LDS is an innovative technology for fabricating antenna structure on thermoplastic materials specifically manufactured for that purpose. As shown in Figure 4, the fabrication process is implemented following 4 distinct operations. First, the part on which the metallic part will be patterned is fabricated by the injection molding technique. Commercially available thermoplastic materials provided with LDS additives (thermoplastic doped with a metal-organic compound) or any other plastics, metals, glass, FR4 can be used for the molded part. In the latter case, the material should be coated with some special painting material to prepare the surface for the LDS process. The LDS ready material is a special plastic doped with a mixed metal oxide that can be activated with a laser beam. The second step is the activation of the surface where the antenna pattern will be metallized. The laser patterning of the conductive traces is essential for the copper to adhere firmly to the surface of the part. Once the surface of the part (antenna) is activated by treating it with the laser beam, the part can be metalized (deposition of the copper ions to the activated surface) by the electroless plating process. Subsequently, other surface finishing techniques such as reduction of the oxidation can be employed (Yang et al., 2017; Stafford, 2014; Kumar et al., 2020).

Although the LDS technique is well-known for structuring electronic circuits for molded interconnect devices, to the best of the author's knowledge, it has not been fully exploited for flexible and wearable antenna fabrications. For many years, it has been considered mainly to manufacture antennas for consumer devices such as smartphones, laptops and tablets (Friedrich et al., 2016). Only recently, other devices such as smart watches and wristbands have been shown as potential applications for the LDS method (Kumar et al., 2020; Berkelmann et al., 2018). Future direction of research should definitely include other wearable antenna applications such as flexible and textile ones. The appropriateness of textile materials for LDS technique should be studied and examined.

In (Kumar et al., 2020), LDS technique has been proposed for the fabrication of an antenna for a wristwatch wireless sensor (Figure 5a). In the aforementioned work, the antenna structure is printed on an ASA thermoplastic enclosure. In order to allow metallization of the selected parts, the part is then coated with an LDS epoxy layer. After having implemented a π-type matching network, it has been shown that the measured results exhibit acceptable impedance and radiation characteristics.

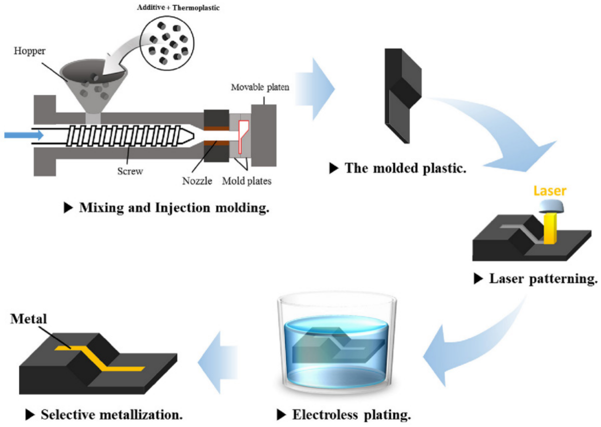
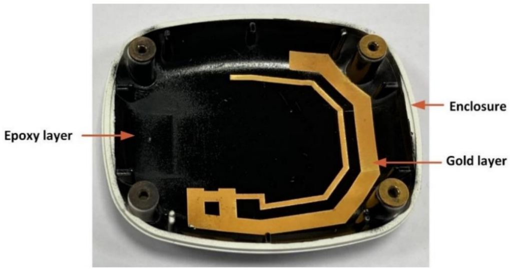
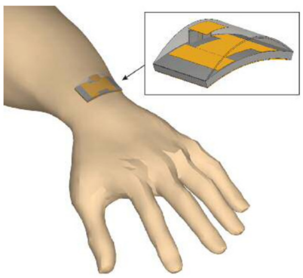


Figure 4. Laser Direct Structuring Process (Yang et al., 2017)

In (Berkelmann et al., 2018) authors proposed the LDS technology for the fabrication of off and on-body communication antennas to fit the 3D surface of a human body. A slotted patch antenna and an inverted-F antenna are designed to operate at the same frequency band of 2.4 GHz for Wireless Body Area Networks (Figure 5b). The Vectra 840i LDS, a Liquid Crystal Polymer (LCP) is used as a substrate for structuring the antennas. It has been shown that by utilizing the LDS technology, the desired excitation of the electromagnetic wave components for both off and on-body communication is possible. Moreover, this technology allows a compact and conformal integration of the antenna system to the human wrist.



a)



b)

Figure 5. a) Wristwatch device showing the antenna unit (Kumar et al., 2020), b) Wristband on a 3D human wrist model (Berkelmann et al., 2018)

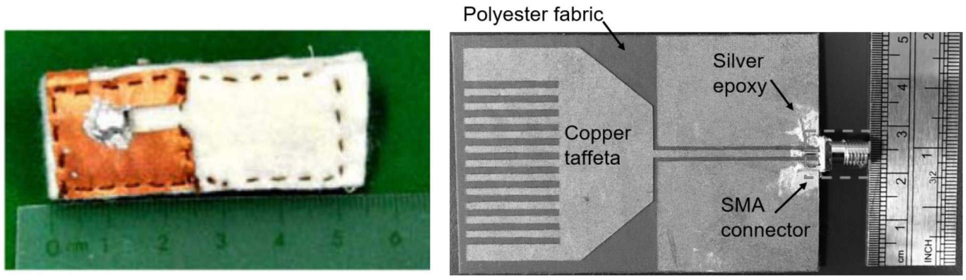
Antenna Materials

The choice of the material is utmost important for the fabrication of the wearable antenna. Both of the dielectric (substrate) and the conductive (ground and radiating patch) parts of the antenna should be chosen properly. In fact, the nature of the material of the antenna will directly affect its performance as well as its integration to the garment. An appropriate selection of the material will contribute to the flexibility and conformability of the antenna. In this work, although other materials exist, only textile and polymer based antennas and related works will be presented and discussed.

Textile Based Antennas

Early textile antennas were based on rigide conductive copper foils directly glued to the textile materials. It is known that one of the first studies on antennas on fabric materials was conducted in 2001 by (Salonen et al., 2001). In the aforementioned study, a flexible PIFA antenna working as single band and dual band in Bluetooth and UMTS bands was designed and manufactured. It has been observed that the working performance of the antenna structure, which is claimed to have been the first commercial thin and flexible antenna, is satisfactory. The same author, in another study, examines the effects of different fabric substrates on a GPS antenna (Salonen et al., 2004). In this study, it is mentioned that fiber-reinforced synthetic fabric, whose trade name is Cordura, can be preferred because of its resistance to mechanical changes and the presence of water. In (Salvado et al., 2012), the authors compare many fabric materials used in antenna manufacture. As expected, Cordura® + copper foil gives good results.

According to the authors, the first all-fabric PIFA antenna was manufactured using conductive copper polyester taffeta, in 2012 (Soh et al., 2012). As seen in Figure 6a, the antenna is completely made of fabric, and the conductive copper taffeta fabric is sewn directly onto the felt fabric with thread. More recently, a low-profile ultra-wideband antenna capable of operating between 1.198–4.055 GHz for microwave medical imaging has been developed by applying conductive copper taffeta fabric on a polyester fabric (Lin et al., 2020). By using conductive copper taffeta fabric instead of traditional copper foil, a 100% textile wearable antenna was developed and the flexibility, which is a very important parameter in wearable antennas, was increased (Fig. 6b). In this study, copper taffeta was attached to the polyester fabric with a very thin thermal film and iron steam. In addition to providing a solid adhesion, it has been reported in (Loss et al., 2016) that the addition of the adhesive thermal film may experience shifts in conductivity and dielectric properties of the material, which will affect the performance of the antenna. Since the copper taffeta fabric is a highly conductive and corrosion resistant material, the decrease in antenna performance due to the weakening of the conductivity caused by corrosion can be prevented.



a) b)

Figure 6. a) PIFA antenna (Soh et al., 2012) and b) UWB antenna (Lin et al., 2020), made of copper taffeta on polyester fabric.

In (Loss et al., 2016), it is aimed to harvest energy at 900 MHz and 1800 MHz frequencies by integrating the antenna directly on the fabric made of Cordura ® fabric.

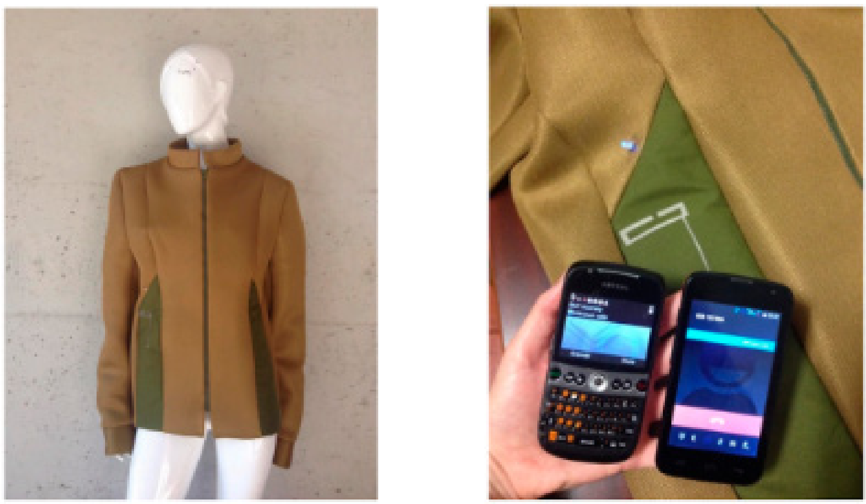


Figure 7. Smart coat with integrated dual band antenna for energy harvesting (Loss et al., 2016)

In this study, instead of copper foil as a conductive material, the structure forming the antenna is embroidered on the fabric with a conductive thread (Figure 7). It is thought that in this way, the antenna will be more invisible and can be easily hidden in the clothing. In addition, the use of a conductive thread instead of a copper foil, which is a harder and solid material, is important for the ease of use of the clothes. Processing antennas directly on fabrics with embroidery machines provides advantages for mass production. Especially with the digitalization of embroidery machines, the transfer of the antenna pattern directly to the computer-controlled machine facilitates the processing of antennas in different colors and complex forms (Kiourti & Volakis, 2015). Embroidered antennas are easy to apply and mechanically robust, which makes them advantageous over other fabric antennas. However, it is necessary to pay attention to aspects such as the conductivity, durability and flexibility of the thread used, and the direction and density of embroideries (Chauraya et al., 2012). The direction and the density of the embroidery (Zhang et al., 2012) not interfering with the flow of the electric current will

positively affect the antenna performance.
One of the antennas embroidered directly on the fabric is given in (Moradi et al., 2013). In this study, RFID tag performances operating at UHF frequencies were examined according to embroidery direction and patterns, and the reading distance of the antenna was tested on the body. In 2005, the authors developed another approach by integrating the antenna directly into the embroidered logos on the garment (Soh et al., 2012). A similar study can be found in (Atanasova et al., 2022). Thanks to the high sensitivity of the embroidery machine, more resistant, flexible, lightweight and colorful logo antennas have been developed. More recently, an RFID tag antenna that can operate in the 870-990 MHz range has been produced with X-static (©Kwontex, Korea) conductive yarn formed by twisting many silver-plated thin elastic nylon fibers (Figure 8). In this study, the effects on antenna performance of parameters such as yarn density and the antenna deformation were investigated. It is concluded that increasing the yarn density would also increase the resonance frequency (Truong et al., 2021). In another study, a compact textile antenna, knitted with Zari silver conductive material, was embroidered on a cotton substrate in order to monitor the heart rate. Besides the humidity, deformation, and wrinkles, the effect of the human body on the antenna performance is investigated (Anbalagan et al., 2020).

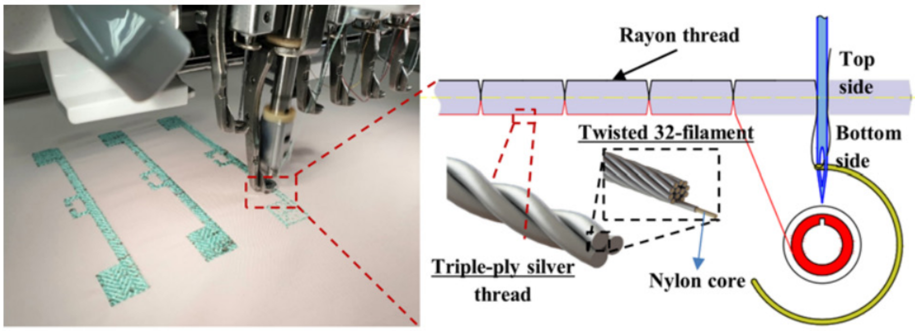


Figure 8. Embroidering RFID tag antenna using X-statik conductor thread (Truong et al., 2021)

In (el Gharbi et al., 2022), the authors embroidered a curved dipole antenna on a cotton t-shirt to monitor real-time breathing (Figure 9). The conductive parts of the antenna are made of a Shieldex® brand conductive thread. It has been shown that the antenna operating at the 2.4 GHz frequency consistently detects the person's breathing in different positions (standing and sitting) and in different breathing patterns. It is predicted that the developed antenna can be used to detect problems such as respiratory failure in newborn babies, sleep apnea and asthma in children and adults.

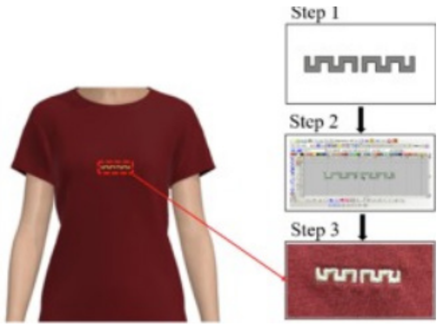


Figure 9. Embroidered antenna integration into a cotton T-Shirt for breath monitoring (el Gharbi et al., 2022)

Polymer Based Antennas

Polymers have the great advantage to be highly flexible and stretchable that make them suitable for wearable applications. As such, polymers are mostly used as substrates. Due to the proximity of the antenna to the human body, it must be robust against any deformation. For instance, if the antenna is mounted on a clothes, it must withstand any bendings and twistings of the fabric. Moreover, polymers are highly resistant to water and they can be used without altering their electrical properties. Among various polymer materials, the polyimide, the Polyethylene terephthalate, (PET), the polydimethylsiloxane (PDMS) materials are the most widely used ones (H. Khaleel, 2015).

The Kapton material, a polyimide film, is one the most commonly used substrates for wearable antenna applications. This material can withstand extreme temperatures while maintaining its physical properties (H. R. Khaleel et al., 2013). In (H. R. Khaleel et al., 2012), authors presented a compact elliptical UWB monopole wearable/flexible antenna (Figure 10a). The antenna is printed on a 50.8- m Kapton polyimide substrate. It has been shown that the impedance matching and the radiation pattern characteristics of the antenna has a very small change when the antenna is subjected to deformation. In a more recent study, an UWB antenna, where the Kapton material is used for the substrate and a conductive nanocomposite material based on a polymer (polyaniline: PANI) and charged by multiwalled carbon nanotubes (MWCNTs) are used for the ground plane and the radiating patch respectively, is developed (Hamouda et al., 2018). Flexible polymer antennas intended to be used in IoT applications and based on Polyethylene Terephthalate (PET), Polytetrafluoroethylene (PTFE) Teflon, and Polyvinylchloride (PVC) substrates operating at 2.45, 4.45, and 7.25 GHz frequencies were fabricated and their bending capabilities are evaluated in (Khan et al., 2021). Next, a highly flexible antenna based on the use of PDMS material was proposed in (Simorangkir et al., 2016; Simorangkir et al., 2018). A conductive silver fabric was embedded into the PDMS substrate which has the advantage of having flexibility as well as water and thermal resistance (Figure 10b).

Figure 10. a) Elliptical UWB antenna based on Kapton polyimide material (H. R. Khaleel et al., 2012), b) PDMS-embedded conductive fabric antenna (Simorangkir et al., 2018)

Conclusion

In this chapter, some of the most recent works on wearable antenna technologies have been reviewed. After a brief history and their evolution timeline, commonly used fabrication techniques and materials have been presented. Due to their wide use, textile and polymer-based antennas are considered in this work. Among textile materials, embroidered patches on fabrics have shown to provide much more strength and flexibility compared to patches made of copper foils. High flexibility, resistance to high temperatures and water make polymer-based antennas the preferred solution for applications where extreme conditions prevail. Regarding the fabrication process, the embroidering technique has become popular for integrating textile antennas directly on the clothes. Depending on the electrical and mechanical properties of the thread used for embroidering the metallic patch and the density of the stitches, high performances can be reached. A more recent approach based on laser direct structuring was also presented for manufacturing wearable antennas. Although this technique is promising for the development of more complex design, only few works exist. This technique should be proven on substrates such as textile and flexible polymers.

References

Al Bassam, N., Hussain, S. A., al Qaraghuli, A., Khan, J., Sumesh, E. P., & Lavanya, V. (2021). IoT Based Wearable Device To Monitor the Signs of Quarantined Remote Patients of COVID-19. *Informatics in Medicine Unlocked*, *24*. https://doi.org/10.1016/j.imu.2021.100588.

Ali, S. M., Sovuthy, C., Imran, M. A., Socheatra, S., Abbasi, Q. H., & Abidin, Z. Z. (2020). Recent Advances of Wearable Antennas in Materials, Fabrication Methods, Designs, and Their Applications: State-of-the-Art. *Micromachines*, *11*(10). https://doi.org/10.3390/mi11100888.

An, W., Li, Y., Fu, H., Ma, J., Chen, W., & Feng, B. (2018). Low-Profile and Wideband Microstrip Antenna with Stable Gain for 5G Wireless Applications. *IEEE Antennas and Wireless Propagation Letters*, *17*(4), 621–624. https://doi.org/10.1109/LAWP.2018.2806369.

Anbalagan, A., Sundarsingh, E. F., Ramalingam, V. S., Samdaria, A., Gurion, D. ben, & Balamurugan, K. (2020). Realization and Analysis of a Novel Low-Profile

Embroidered Textile Antenna for Real-time Pulse Monitoring. *IETE Journal of Research*, 1–8. https://doi.org/10.1080/03772063.2020.1787877.

Atanasova, G. L., Atanasov, B. N., & Atanasov, N. T. (2022). Fully Textile Dual-Band Logo Antenna for IoT Wearable Devices. *Sensors*, *22*(12), 1–19. https://doi.org/10.3390/s22124516.

Baig, M. M., Gholamhosseini, H., & Connolly, M. J. (2013). A Comprehensive Survey of Wearable and Wireless ECG Monitoring Systems for Older Adults. *Medical and Biological Engineering and Computing*, *51*(5), 485–495. https://doi.org/10.1007/s11517-012-1021-6.

Berkelmann, L., Martinelli, T., Friedrich, A., & Manteuffel, D. (2018). Design and Integration of a Wearable Antenna system for on and Off-Body Communication Based on 3D-MID Technology. *IET Conference Publications*, *2018*(CP741). https://doi.org/10.1049/CP.2018.0849.

Blayo, A., & Pineaux, B. (2005). Printing Processes and Their Potential for RFID Printing. *Proceedings of the 2005 Joint Conference on Smart Objects and Ambient Intelligence Innovative Context-Aware Services: Usages and Technologies - Joint SOc-EUSAI Conference*, 27–30. https://doi.org/10.1145/1107548.1107559.

Carmigniani, J., Furht, B., Anisetti, M., Ceravolo, P., Damiani, E., & Ivkovic, M. (2011). Augmented Reality Technologies, Systems and Applications. *Multimedia Tools and Applications*, *51*(1), 341–377. https://doi.org/10.1007/s11042-010-0660-6.

Chauraya, A., Zhang, S., Whittow, W., Acti, T., Seager, R., Dias, T., & Vardaxoglou, Y. C. (2012). Addressing the Challenges of Fabricating Microwave Antennas Using Conductive Threads. *Proceedings of 6th European Conference on Antennas and Propagation, EuCAP 2012*, 1365–1367. https://doi.org/10.1109/EuCAP.2012.6205910.

Dacuna, J., & Pous, R. (2009). Low-Profile Patch Atenna for RF Identification Applications. *IEEE Transactions on Microwave Theory and Techniques*, *57*(5), 1406–1410. https://doi.org/10.1109/TMTT.2009.2017322.

Dias, D., & Cunha, J. P. S. (2018). Wearable Health Devices-Vital Sign Monitoring, Systems and Technologies. *Sensors (Switzerland)*, *18*(8), 1–28. https://doi.org/10.3390/s18082414.

el Gharbi, M., Fernández-García, R., Ahyoud, S., & Gil, I. (2020). A Review of Flexible Wearable Antenna Sensors: Design, Fabrication Methods, and Applications. *Materials*, *13*(17). https://doi.org/10.3390/MA13173781.

el Gharbi, M., Fernández-García, R., & Gil, I. (2022). Embroidered Wearable Antenna-Based Sensor For Real-Time Breath Monitoring. *Measurement: Journal of the International Measurement Confederation*, *195*. https://doi.org/10.1016/j.measurement.2022.111080.

Friedrich, A., Geck, B., & Fengler, M. (2016, September). LDS Manufacturing Technology For Next Generation Radio Frequency Applications. *2016 12th International Congress Molded Interconnect Devices - Scientific Proceedings, MID 2016*. https://doi.org/10.1109/ICMID.2016.7738939.

Galov, N. (2022, August 12). *How Many IoT Devices Are There In 2022? [All You Need To Know]*. https://techjury.net/blog/how-many-iot-devices-are-there/#gref.

Hamouda, Z., Wojkiewicz, J. luc, Pud, A. A., Kone, L., Bergheul, S., & Lasri, T. (2018). Flexible UWB Organic Antenna for Wearable Technologies Application. *IET Microwaves, Antennas and Propagation*, *12*(2), 160–166. https://doi.org/10.1049/iet-map.2017.0189.

HealthWatch contributors. (n.d.). *A t-shirt That Monitors Your Heart - HealthWatch Ltd.| Healthcare Wearables*. Retrieved October 8, 2022, from https://healthwatchtech.com/a-t-shirt-that-monitors-your-heart/.

Heilman, K. J., & Porges, S. W. (2007). Accuracy of the LifeShirt® (Vivometrics) In the Detection of Cardiac Rhythms. *Biological Psychology*, *75*(3), 300–305. https://doi.org/10.1016/j.biopsycho.2007.04.001.

Hughes-Riley, T., Dias, T., & Cork, C. (2018). A Historical Review of the Development of Electronic Textiles. *Fibers*, *6*(2). https://doi.org/10.3390/fib6020034.

Islam, M. S., Kayser Azam, S. M., Zakir Hossain, A. K. M., Ibrahimy, M. I., & Motakabber, S. M. A. (2022). A Low-Profile Flexible Planar Monopole Antenna For Biomedical Applications. *Engineering Science and Technology, an International Journal*, *35*. https://doi.org/10.1016/j.jestch.2022.101112.

Khaleel, H. (Ed.). (2015). *Innovation In Wearable and Flexible Antennas* (Vol. 82). WIT Press.

Khaleel, H. R., Al-Rizzo, H. M., Abbosh, A. I., Khaleel, H. R., Al-Rizzo, H. M., & Abbosh, A. I. (2013). Design, Fabrication, and Testing of Flexible Antennas. *Advancement in Microstrip Antennas with Recent Applications*, 363–383. https://doi.org/10.5772/50841.

Khaleel, H. R., Al-Rizzo, H. M., Rucker, D. G., & Mohan, S. (2012). A Compact Polyimide-Based UWB Antenna For Flexible Electronics. *IEEE Antennas and Wireless Propagation Letters*, *11*, 564–567. https://doi.org/10.1109/LAWP.2012.2199956.

Khan, M. U. A., Raad, R., Tubbal, F., Theoharis, P. I., Liu, S., & Foroughi, J. (2021). Bending Analysis of Polymer-Based Flexible Antennas for Wearable, General Iot Applications: A Review. *Polymers*, *13*(3), 1–34. https://doi.org/10.3390/polym13030357.

Kim, S. (2020). Inkjet-Printed Electronics On Paper For RF Identification (RFID) and Sensing. *Electronics (Switzerland)*, *9*(10), 1–22. https://doi.org/10.3390/electronics9101636

Kiourti, A., & Volakis, J. L. (2015). Colorful Textile Antennas Integrated Into Embroidered Logos. *Journal of Sensor and Actuator Networks*, *4*(4), 371–377. https://doi.org/10.3390/jsan4040371.

Komolafe, A., Zaghari, B., Torah, R., Weddell, A. S., Khanbareh, H., Tsikriteas, Z. M., Vousden, M., Wagih, M., Jurado, U. T., Shi, J., Yong, S., Arumugam, S., Li, Y., Yang, K., Savelli, G., White, N. M., & Beeby, S. (2021). E-Textile Technology Review-From Materials To Application. *IEEE Access*, *9*, 97152–97179. https://doi.org/10.1109/ACCESS.2021.3094303.

Kumar, S., Buckley, J. L., Barton, J., Pigeon, M., Newberry, R., Rodencal, M., Hajzeraj, A., Hannon, T., Rogers, K., Casey, D., O'sullivan, D., & O'flynn, B. (2020). A Wristwatch-Based Wireless Sensor Platform for IoT Health Monitoring Applications. *Sensors (Switzerland)*, *20*(6), 1–25. https://doi.org/10.3390/s20061675.

Le, T., Lin, Z., Vyas, R., Lakafosis, V., Yang, L., Traille, A., Tentzeris, M. M., & Wong, C. P. (2013). Inkjet Printing of Radio Frequency Electronics: Design Methodologies and Application of Novel Nanotechnologies. *Journal of Electronic Packaging, Transactions of the ASME*, *135*(1). https://doi.org/10.1115/1.4023671.

Lin, X., Chen, Y., Gong, Z., Seet, B. C., Huang, L., & Lu, Y. (2020). Ultrawideband Textile Antenna For Wearable Microwave Medical Imaging Applications. *IEEE Transactions on Antennas and Propagation*, *68*(6), 4238–4249. https://doi.org/10.1109/TAP.2020.2970072.

Loss, C., Gonçalves, R., Lopes, C., Pinho, P., & Salvado, R. (2016). Smart Coat with a Fully-Embedded Textile Antenna for IoT Applications. *Sensors (Switzerland)*, *16*(6), 1–13. https://doi.org/10.3390/s16060938.

Mann, S., Fung, J., Aimone EyeTap Personal Imaging Lab, C., & Chen, D. (2005). Designing EyeTap Digital Eyeglasses for Continuous Lifelong Capture and Sharing of Personal Experiences. *Citeseer*. http://christinamann.com.

Marketsandmarkets. (n.d.). *Military Wearables Market Size, Growth, Trend and Forecast (2022-2027)*. Retrieved October 16, 2022, fromhttps://www.marketsandmarkets.com/Market-Reports/military-wearable-market-105612791.html.

Moradi, E., Koski, K., Ukkonen, L., Rahmat-Samii, Y., Björninen, T., & Sydänheimo, L. (2013). Embroidered RFID Tags In Body-Centric Communication. *2013 International Workshop on Antenna Technology, IWAT 2013*, 367–370. https://doi.org/10.1109/IWAT.2013.6518367.

Navalta, J. W., Ramirez, G. G., Maxwell, C., Radzak, K. N., & McGinnis, G. R. (2020). Validity and Reliability of Three Commercially Available Smart Sports Bras During Treadmill Walking and Running. *Scientific Reports*, *10*(1). https://doi.org/10.1038/s41598-020-64185-z.

Ometov, A., Shubina, V., Klus, L., Skibińska, J., Saafi, S., Pascacio, P., Flueratoru, L., Gaibor, D. Q., Chukhno, N., Chukhno, O., Ali, A., Channa, A., Svertoka, E., Qaim, W. bin, Casanova-Marqués, R., Holcer, S., Torres-Sospedra, J., Casteleyn, S., Ruggeri, G., … Lohan, E. S. (2021). A Survey On Wearable Technology: History, State-of-the-Art and Current Challenges. *Computer Networks*, *193*, 108074. https://doi.org/10.1016/J.COMNET.2021.108074.

Rodrigues, J. J. P. C., de Rezende Segundo, D. B., Junqueira, H. A., Sabino, M. H., Prince, R. M. I., Al-Muhtadi, J., & de Albuquerque, V. H. C. (2018). Enabling Technologies For the Internet of Health Things. *IEEE Access*, *6*, 13129–13141. https://doi.org/10.1109/ACCESS.2017.2789329.

Roudjane, M., Khalil, M., Miled, A., & Messaddeq, Y. (2018). New Generation Wearable Antenna Based on Multimaterial Fiber For Wireless Communication and Real-Time Breath Detection. *Photonics*, *5*(4), 1–20. https://doi.org/10.3390/photonics5040033.

Salonen, P., Keskilammi, M., Rantanen, J., & Sydänheimo, L. (2001). A Novel Bluetooth Antenna On Flexible Substrate For Smart Clothing. *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, *2*, 789–794. https://doi.org/10.1109/icsmc.2001.973011.

Salonen, P., Rahmat-Samii, Y., Schaffrath, M., & Kivikoski, M. (2004). Effect of Textile Materials On Wearable Antenna Performance: A Case Study of GPS Antennas.

IEEE Antennas and Propagation Society, AP-S International Symposium (Digest), *1*, 459–462. https://doi.org/10.1109/aps.2004.1329673.

Salvado, R., Loss, C., Gon, & Pinho, P. (2012). Textile Materials For the Design of Wearable Antennas: A survey. *Sensors (Switzerland)*, *12*(11), 15841–15857. https://doi.org/10.3390/s121115841.

Seneviratne, S., Hu, Y., Nguyen, T., Lan, G., Khalifa, S., Thilakarathna, K., Hassan, M., & Seneviratne, A. (2017). A Survey of Wearable Devices and Challenges. *IEEE Communications Surveys and Tutorials*, *19*(4), 2573–2620. https://doi.org/10.1109/COMST.2017.2731979.

Sensoria contributors. (n.d.). *Sensoria Home Page*. Retrieved October 8, 2022, from https://www.sensoriafitness.com/.

Sharma, P. K., Park, J., Park, J. H., & Cho, K. (2020). Wearable Computing For Defence Automation: Opportunities and Challenges In 5G Network. *IEEE Access*, *8*, 65993–66002. https://doi.org/10.1109/ACCESS.2020.2985313.

Simorangkir, R. B. V. B., Kiourti, A., & Esselle, K. P. (2018). UWB Wearable Antenna With a Full Ground Plane Based On PDMS-Embedded Conductive Fabric. *IEEE Antennas and Wireless Propagation Letters*, *17*(3), 493–496. https://doi.org/10.1109/LAWP.2018.2797251.

Simorangkir, R. B. V. B., Yang, Y., Matekovits, L., & Esselle, K. P. (2016). Dual-Band Dual-Mode Textile Antenna On PDMS Substrate For Body-Centric Communications. *IEEE Antennas and Wireless Propagation Letters*, *16*, 677–680. https://doi.org/10.1109/LAWP.2016.2598729.

Soh, P. J., Vandenbosch, G. A. E., Ooi, S. L., & Rais, N. H. M. (2012). Design of a Broadband All-Textile Slotted PIFA. *IEEE Transactions on Antennas and Propagation*, *60*(1), 379–384. https://doi.org/10.1109/TAP.2011.2167950.

Stafford, S. (2014). *What Are Molded Interconnect Devices?*. What-are-molded-interconnect-devices.pdf (lpkfusa.com).

Stoppa, M., & Chiolerio, A. (2014). Wearable Electronics and Smart Textiles: A Critical Review. *Sensors (Switzerland)*, *14*(7), 11957–11992. https://doi.org/10.3390/s140711957.

Tokucoglu, F. (2018). Monitoring Physical Activity with Wearable Technologies. *Archives of Neuropsychiatry*, 63–65. https://doi.org/10.29399/npa.23333.

Truong, T. T. N., Kim, J. S., & Kim, J. (2021). Design and Optimization of Embroidered Antennas On Textile Using Silver Conductive Thread for Wearable Applications. *Fibers and Polymers*, *22*(10), 2900–2909. https://doi.org/10.1007/s12221-021-0030-1.

Tsolis, A., Whittow, W. G., Alexandridis, A. A., & Vardaxoglou, J. Y. C. (2014). Embroidery and Related Manufacturing Techniques for Wearable Antennas: Challenges and Opportunities. *Electronics*, *3*(2), 314–338. https://doi.org/10.3390/electronics3020314.

Ullah, M. Z. (2009). *An Analysis of the Bluetooth Technology- Features, Challenges, Future and Security, Thesis no: MCS-2009-17 [Master Thesis Computer Science], School of Computing, Blekinge Institute of Technology*. www.bth.se/tek.

Uz Zaman, S., Tao, X., Cochrane, C., & Koncar, V. (2020). Understanding the Washing

Damage To Textile ECG Dry Skin Electrodes, Embroidered and Fabric-Based; Set Up of Equivalent Laboratory Tests. *Sensors (Switzerland)*, *20*(5). https://doi.org/10.3390/s20051272.

Uz Zaman, S., Tao, X., Cochrane, C., & Koncar, V. (2022). Smart E-Textile Systems: A Review For Healthcare Applications. *Electronics (Switzerland)*, *11*(1). https://doi.org/10.3390/electronics11010099.

Verma, M., Singh, S., & Kaur, B. (2015). An Overview of Bluetooth Technology and Its Communication Applications. *1588| International Journal of Current Engineering and Technology*, *5*(3). http://inpressco.com/category/ijcetGeneralArticle.

Vijayalakshmi, K., Uma, S., Bhuvanya, R., & Suresh, A. (2018). A Demand For Wearable Devices In Health Care. *International Journal of Engineering & Technology*, *7*(1), 1–4. https://doi.org/10.14419/IJET.V7I1.7.9377.

Wang, Z., Zhang, L., Bayram, Y., & Volakis, J. L. (2012). Embroidered Conductive Fibers On Polymer Composite for Conformal Antennas. *IEEE Transactions on Antennas and Propagation*, *60*(9), 4141–4147. https://doi.org/10.1109/TAP.2012.2207055.

Wikipedia contributors. (2022, September 14). *Google glass - Wikipedia*. https://en.wikipedia.org/wiki/Google_Glass.

Yang, J. uk, Cho, J. H., & Yoo, M. J. (2017). Selective Metallization On Copper Aluminate Composite Via Laser Direct Structuring Technology. *Composites Part B: Engineering*, *110*, 361–367. https://doi.org/10.1016/J.COMPOSITESB.2016.11.041.

Zhang, S., Chauraya, A., Whittow, W., Seager, R., Acti, T., Dias, T., & Vardaxoglou, Y. (2012). Embroidered Wearable Antennas Using Conductive Threads With Different Stitch Spacings. *LAPC 2012 - 2012 Loughborough Antennas and Propagation Conference*, 1–4. https://doi.org/10.1109/LAPC.2012.6403059.

About the Authors

Hulusi ACIKGOZ received his B.Sc. and M.Sc. degrees in Applied Physics from Paris-Est Marne La Vallée University, France in 2003 and 2005, respectively. His Ph.D. degree in Electrical Engineering is obtained from Pierre et Marie Curie University, France, in 2008. He was a post-doctoral researcher at the Laboratoire d'Électronique et Électromagnétisme (L2E), France, in 2010. In 2015, he spent a sabbatical year at Penn State University, USA, within the Electromagnetic Communication Lab, directed by Prof. Raj Mittra. From 2011 to 2020, he worked at KTO Karatay University, Turkey as an assistant professor in the electrical and electronics department. Since 2020, he is associate professor at the same university. Currently, he is also holding a technical leader position at the Smart, Technologies, Design, Development and Prototyping Center, funded by the European Union. His main research interests include computational electromagnetics, antennas modelling and analysis, RFID antennas, graphene antennas, statistical electromagnetics, development of antennas for microwave cancer ablation, inverse problems, neural networks, and electromagnetic compatibility. Dr. Acikgoz's is a member of the steering committee of URSI Turkey. He organized and served as the general chairman of URSI Turkey national congress, in 2018.

Email: hulusi.acikgoz@karatay.edu.tr, **ORCID:** 0000-0003-1648-2596

Ilayda CAN, M.Sc, is Research Assistant of Electrical and Electronics Engineering at KTO Karatay University in Konya, Turkey. She graduated from the Electrical and Electronics Engineering Department of KTO Karatay University (2020). She is continuning the master's degree in Electrical and Computer Engineering at KTO Karatay University. Her main areas of interest are electromagnetics, antennas, microwave, and electromagnetic compability.
Email: ilayda.can@karatay.edu.tr, **ORCID:** 0000-0002-4925-7339

Similarity Index

The similarity index obtained from the plagiarism software for this book chapter is 8%.

To Cite This Chapter:

Acikgoz H, Can I. (2022). Current Trends on Wearable Antennas, Fabrication Techniques and Materials. In S. Kocer & O. Dundar (Eds.), *Current Studies in Basic Sciences Engineering and Technology* (pp. 206–223). ISRES Publishing.

Scientific Data Recovery Methods and Solutions in Optics Discs and HDDs

Abdulkadir ÖZDEMİR
Konya Cyber Crime Branch Office

Şaban GÜLCÜ
Necmettin Erbakan University

Introduction

In the history of humanity from the earliest times to the recent past, the source of power; physical strength was evaluated on the basis of land size or numerical value. In today's conditions, power has become equivalent to knowledge. When we look at the world states which are called "superpowers" it is seen that one of the reasons why they are called by this name is that they have advanced technologies based on information. The concept of data is the general name given to any unprocessed raw data. Institutions, organizations, and individuals prefer digital media (HDD - SSD - flash memory, etc.) as a data storage medium, for reasons such as not occupying physical space, providing ease of access, and fast processing. Even today, it is observed that cloud storage methods are used more and more widely. As the amount of data created in daily life increases, the areas where this data will be stored will increase in direct proportion. The increase in data storage units and their widespread use will indirectly increase the possibility of digital data loss in different places, people, and time zones. The bit is the smallest unit used to store data in information systems. Bit expression also comes from the Binary digit. Data loss is mentioned when some or all of the data expected to be interpreted by the information system cannot be processed/understood. If the scaling of the bad effects of individual or corporate digital data loss is to be made; the losses of individuals due to digital data loss are generally higher in moral value, while the losses of institutions or organizations due to digital data loss are higher in monetary value. According to the IBM Security and Ponemon Institute, they have determined the costs incurred by data loss due to data breaches in the last 5 years as follows (IBM Security and Ponemon Institute, 2021):

- Average of 3.62 million USD for 419 companies in 13 countries in 2017,
- Average of 3.86 million USD for 447 companies in 15 countries in 2018,
- Average of 3.92 million USD for 507 companies in 16 countries in 2019,
- Average of 3.86 million USD for 524 companies in 17 countries in 2020,
- Average of 4.24 million USD for 537 companies in 17 countries in 2021.

When digital data loss events are examined, it is seen that data loss occurs mainly due to exposure to various effects. These various effects are divided into software, hardware, and semi-hardware problems. Software malfunctions are cases where the required information cannot be accessed even though the digital recording medium is intact and the device containing the recording medium is also in working condition. Hardware failures, on the other hand, are the types of failures that differ according to the environment in which the data is recorded, but basically, it is used to express the situations where the data recording environment does not work (Özdemir and Gülcü, 2021). Semi-hardware failures are failures caused by software (firmware) embedded in the hardware in data recording units. In this study, software and hardware-based failures will be examined.

Data Storage Units

Data is a general name given to unprocessed pieces of raw information stored in the form of signals or bit strings carried in digital media (Durak, 2011). Data are stored in information systems in two ways: permanent (non-volatile) and temporary (volatile). Volatile memories hold data when electrical current (power) is present, they cannot hold data if power is lost. Data storage units by access options it is divided into two according to the options to offer local and remote connections. What is meant by the storage media accessed by local connection, is the access of a user to devices containing optical devices, magnetic and flash memory that can be directly connected to the motherboard of the information system. By means of a storage medium accessed by remote connection, it is possible to connect to the storage unit in an information system, either wired or wirelessly, from the local network or over the internet. The information system, which is connected via a remote connection, offers a local storage environment, considering the storage environment it offers to the user. Meanwhile, as a result of the widespread use of the internet with portable devices such as mobile phones and tablets, it has increased the use of cloud storage very quickly, as it allows it to be accessed regardless of location and time (Keskin et al., 2020). Data storage units have different features from each other in terms of connection interface, packaging/ production format, and data recording technique. Digital data in storage units; magnetic, optical, and semiconductor (electronic) media are stored. Storage media containing data are classified according to certain criteria. This classification is presented in Figure 1 as an infographic. Detailed information about these storage units will be given under the relevant headings (Özdemir, 2022).

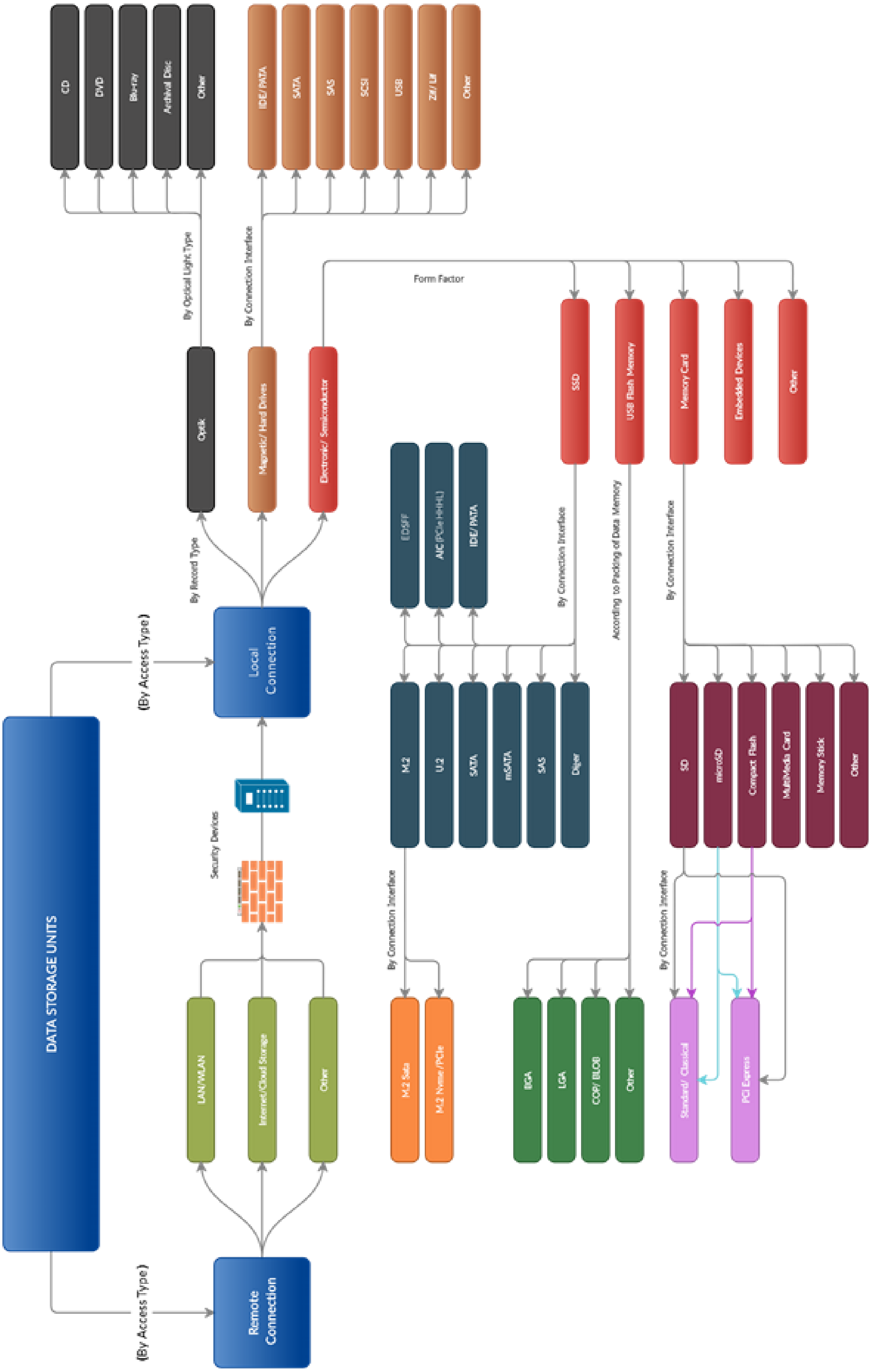


Figure 1. Data storage units and their classification

Data to Your Loss Causes of Causation and Data Recovery Methods

Data loss occurs when digital data cannot be accessed/used in normal ways for various reasons. It is also called data recovery when the reasons that cause data loss are determined and the necessary interventions are eliminated by using special hardware and software. These various reasons are generally; data hiding, deletion, corruption, encryption; breaking, burning, mechanical or electronic failure of data-hosting equipment; hardware (firmware) corruption, inability to be used by the device, etc. can be exemplified. These reasons can be natural, accidental, or sabotage. In Figure 2, the results of the research conducted by "Ontract Data" company in 2003 and "DeepSpar Data Recovery Systems" company obtained from 50 data recovery companies in 14 countries in 2007 regarding the causes of data loss are presented.

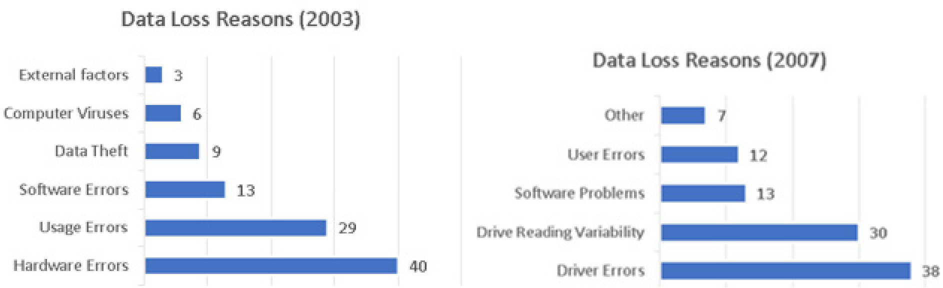


Figure 2. Causes of data loss (David, 2000; DeepSpar Data Recovery Systems, 2007))

Cloud storage or other structures (Das, Nas, San) or the problems caused by these are systemic failures, if these are ignored, data loss situations usually occur in data recording units. For example, the corruption of a disk in the device using Raid levels affects the entire structure. If the problem on this disk is resolved, the entire structure will be ready for use again. However, devices such as Das, Nas, and San offer the end user a logical space (volume) by using various parameters on the disks. If the device itself is corrupted, the data will not be accessible (except for Raid 1), since the logically presented structure will not be available. If the parameters used while creating this structure are created (simulated) software, only the data will be accessed.

Data recovery operations are subject to various classifications according to the method used or the type of digital material tampered with. The classification of data recovery methods in the literature is made according to the "cause of the failure" and "data recording technology" (Güllüce and Benzer, 2015). The main purpose during the data recovery process is to eliminate the reason why the user cannot access the data. The problems encountered in data recovery and how their solutions are carried out will be tried to be explained by using visual examples as much as possible. If the term in the solution methods is to be explained;

Raw Recovery: It is also called recovery or File carving. When the hexadecimal structure of the files in the data recording medium is examined, it is seen in Table 1 different types of files have different headers and some files have a footer in addition to the header.

Table 1. Headers and Footers of Files (Exterro Inc., 2021; Kessler, 2022)

File Extension	File Description	Signature (Header)	footer
($MFT)	Windows MFT Records	46 49 4c 45 30 00	
(FAT1)	FAT32 File Allocation table	F8 FF FF 0F FF FF FF 0F	
apk	Android Archive File	50 4B 03 04	
doc	MS Office Word document	EC A5 C1 00	
docx	MS Office 2007 documents	50 4B 03 04 14 00 06	
jpg/jpeg	JFIF/JPEG Image file	FF D8 FF E0 00 10 4A 46 49 46	FF D9
mp3	MPEG-1 Audio Layer 3 audio file	49 44 33	
mp4	MPEG-4 Video Files	00 00 00 18 66 74 79 70 33 67 70 35	
mov	Apple Quicktime movie file	00 00 00 20 66 74 79 70 4D 34 41 20	
pdf	Adobe Portable Document Format	25 50 44 46	0D 25 25 45 4F 46 0D
psd	Adobe Photoshop Image File	38 42 50 53	81 FF 81 FF 91 FF 81 FF F9 FF
rar	WinRAR **v3** Compressed Archive File	52 61 72 21 1A 07 **00 CF**	
	WinRAR **v5** Compressed Archive File	52 61 72 21 1A 07 **01 00**	
xls	MS Office Excel Spreadsheet	FD FF FF FF	45 78 63 65 6C 2E 53 68 65 65 74 2E 38 00 F4 39 B2 71
xlsx	MS Office 2007 documents	50 4B 03 04 14 00 06	
zip	WinZIP compressed archive	57 69 6E 5A 69 70	

This header and footer information, which enables files to be distinguished from each other, is called file signature. Data recovery and forensic software read/scan the data recording medium from beginning to end. During this scan, the process of recovering files or file system structures based on the signatures (header-footer) of the files and the clustering information (cluster) they use in the data recording medium is called raw recovery. If data recovery software can't access the clustering information and they didn't use sequential clustering from the files (in the fragment), the files will not be fully recovered. Therefore, the priority of the data recovery software used should be to recover the file system structures (MFT, FAT1, VBR, etc.) (Özdemir, 2022).

File-based recovery: Raw recovery is performed for one or a few files, not all files.

File repair: It is the process of removing the corruption or damage with software if the "recovered" files are corrupted or damaged as a result of raw recovery or file-based recovery.

In general, although it is tried to collect the malfunctions under certain headings, it should be known that the cause of each malfunction is different and the solutions will be different depending on this.

Software Data Loss and Solutions

Software data loss is defined as the situation where the requested files cannot be opened/viewed/used while accessing the normally working digital storage unit, where there is no physical failure.

Many users believe that their files are permanently deleted by simply moving their files to the recycle bin and then emptying the recycle bin. But there are two types of deletion. Figure 3 and we have two buttons named “button1” and “button2” on our form, after changing the text of “button1” to “Delete button2” Click Add event “button2.Visible= false;” Let’s write.

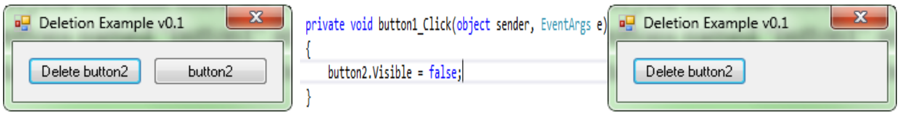


Figure 3. For the deletion example, form design, button typing is done with code and button

“Delete button2” in Figure 3 clicked, it will be seen that the button labeled “button2” has disappeared. The action done here is only to prevent “button2” from being displayed on the form. This is how the files are deleted from the computer. When a “Delete” command is given to the file system by the user, the file system provides the user with the illusion that the file has been deleted by simply making the file invisible (inaccessible) as shown above. As it can be understood, the button is not really deleted here, it is only hidden from the user, the first type of deletion is the case that it is not shown to the user while the data is still on the disk.

A question may come to mind: when files are deleted, the amount of free space on the disk increases, how would this amount of free space appear if there was no real deletion? Before answering this question, the concepts of allocated and unallocated space should be explained. In file systems, the disk areas where the user’s files are stored are called reserved space, and the areas marked as empty in the file system that the user can use to save new files are called unallocated space.

The second type of deletion is permanent deletion (wipe-data erasure) In addition to the deletion process described above, it is the process of writing FF or 00 or different hexadecimal values on the sectors where the data is located. In such cases, data recovery is not possible even if data recovery procedures are applied. Some of the software data recovery software are CDRoller, Disk Doctors, Easeus Data Recovery Wizard, File Scavenger, Flash Extractor, PC-3000, R- Studio, Recover MyFiles, Rusolut, Softcenter Image Explorer. This software can recover files that are consciously or unconsciously deleted by viruses, information system attacks or users, disk partitions damaged or deleted for any reason, file-based corruption, and data loss as a result of hardware or system crashes. It can also retrieve deleted data and configuration information from RAID disks.

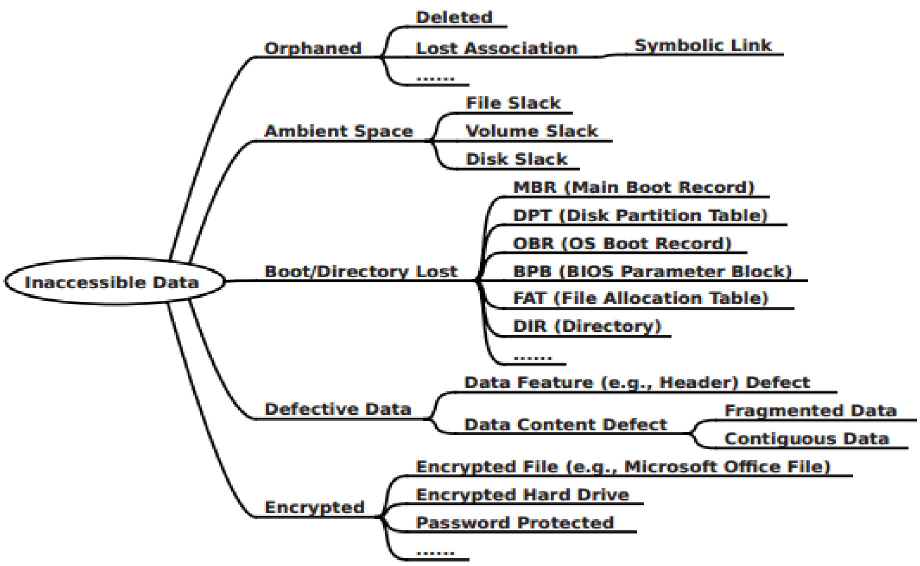


Figure 4. Classification of data according to inaccessibility (Guo and Slay, 2010)

If the software problems and their solutions are exemplified, the following examples can be given;

* Corruption of the file system, the deletion or corruption of the files of the filing system, and the inaccessibility of data and files as a result of the disk becoming Figure 5).

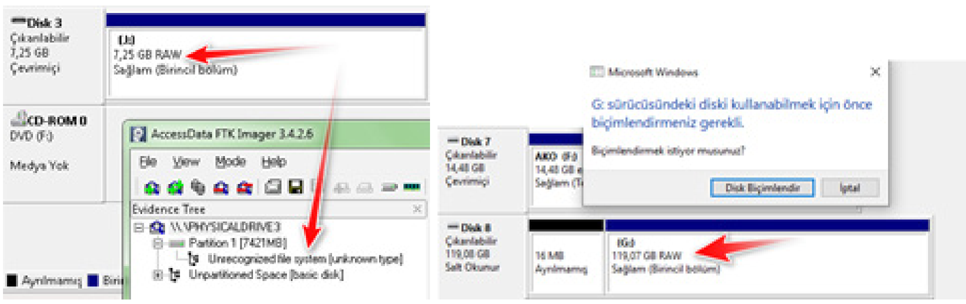


Figure 5. Device instances with a RAW file system

- Raw with data recovery software to access data on drives in this state Recovery or file type-based scanning is required. As an alternative method, try again to access the drive with a Unix-based operating system. The chkdsk commands should not be used without a forensic copy/image of the device.
- This problem is usually caused by static electricity, care should be taken when connecting data loggers to the computer system.

* If the drive letter cannot be assigned to the drive by the operating system, the operating system cannot logically show the drive to the user:

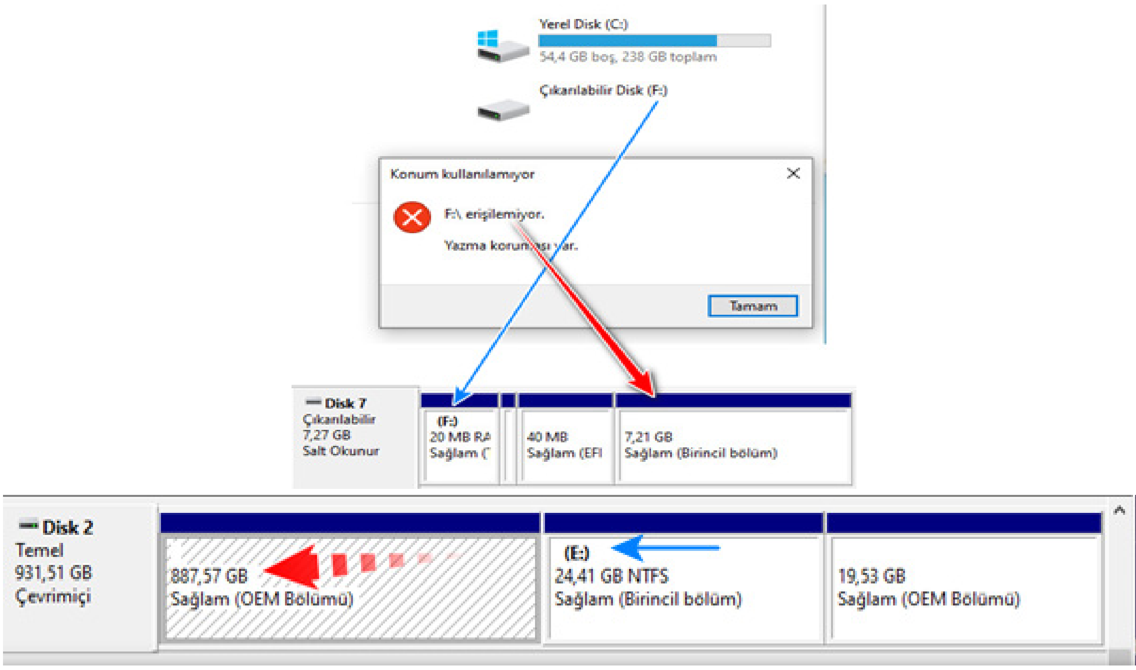


Figure 6. Examples of devices that cannot be assigned a drive letter

- The solution is provided by assigning drive letters to the drives in this situation from the disk management.

* Encryption of the data recording medium: It can be done deliberately by the user (BitLocker, FileVault, TrueCrypt, etc.) as well as the measures taken to protect the devices themselves from unauthorized access.

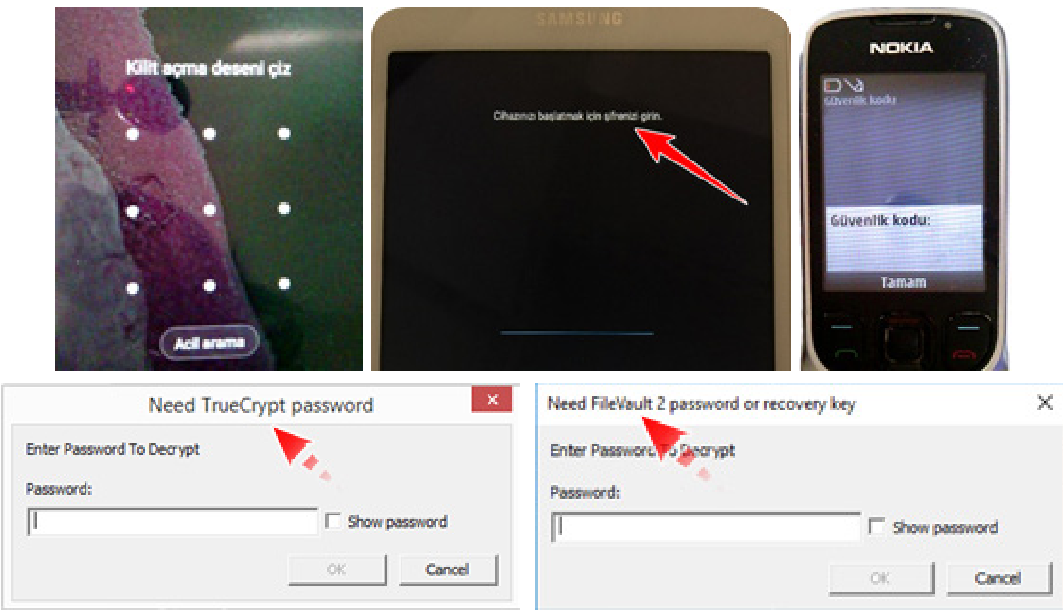


Figure 7. Encrypted device examples

* The surest way to recover data from devices in this situation is to learn the password. If it is not possible to learn the password;

- Manual password attempt
- Special software used (Passware Kit Mobile, PC-3000 Mobile, Cellebrite Ufed 4PC, Oxygen Forensics, Magnets By means of

Axiom, etc.) the pattern should be skipped/deactivated or closed physical forensic copy should be taken,

- Again, with such software, a brute force (brute-force) attack should be done.

* Crashing during updating of operating systems or for any other reason:

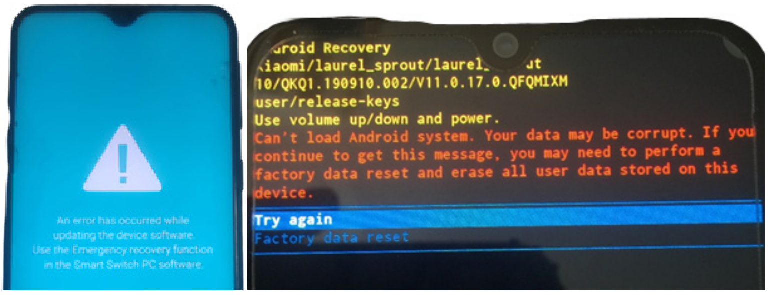


Figure 8. Examples of devices crashing during OS update

* If the device used by the hardware in this situation is a computer, the data can be safely recovered by inserting the disk in the device into a different computer system. There are passwords set in Windows user accounts that do not prevent data recovery or access, data can be accessed by connecting a storage unit to a different computer.

If the device is a mobile device;

- Stock rom loading,
- Taking a physical copy of the device with forensic software,
- Recovery options offered by forensic software can be used.

* Partition table corruption: These tables are corrupted due to sudden voltage changes, power cuts, etc.

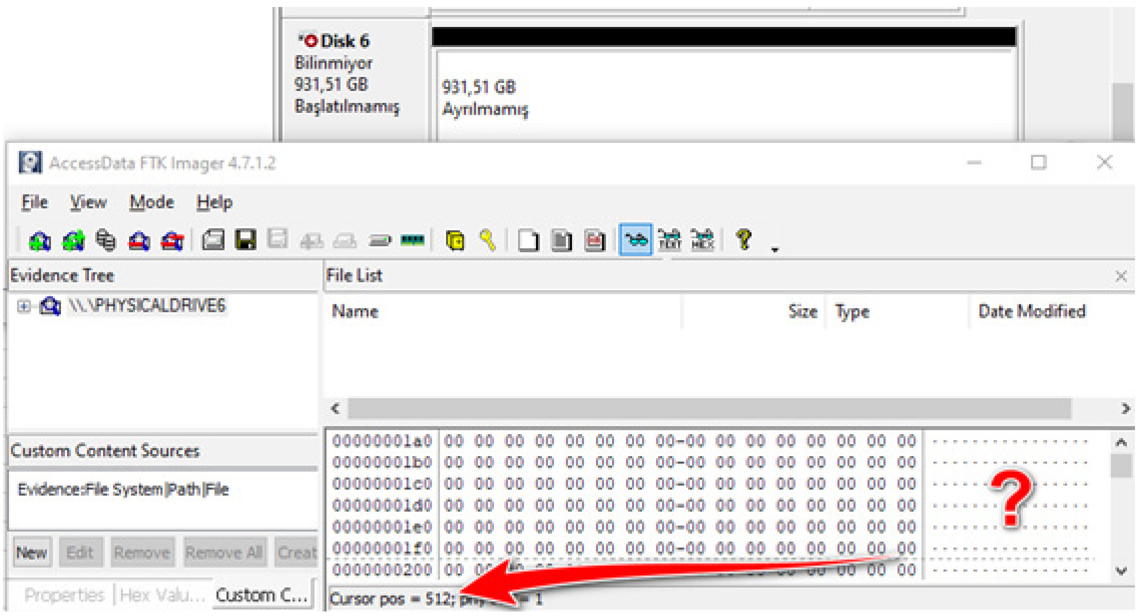


Figure 9. Partition table corrupted disk

- Raw with data recovery software to access data on drives in this state the file system structures should be found by recovering. Or file-based recovery should be done.

* During data recovery, the files do not come completely or not at all due to the scattered (fragment) records on the disk: As can be seen in Figure 10, it is seen that the first of the “recovered” files consists of 4 parts, and the other photo can be recovered in half.

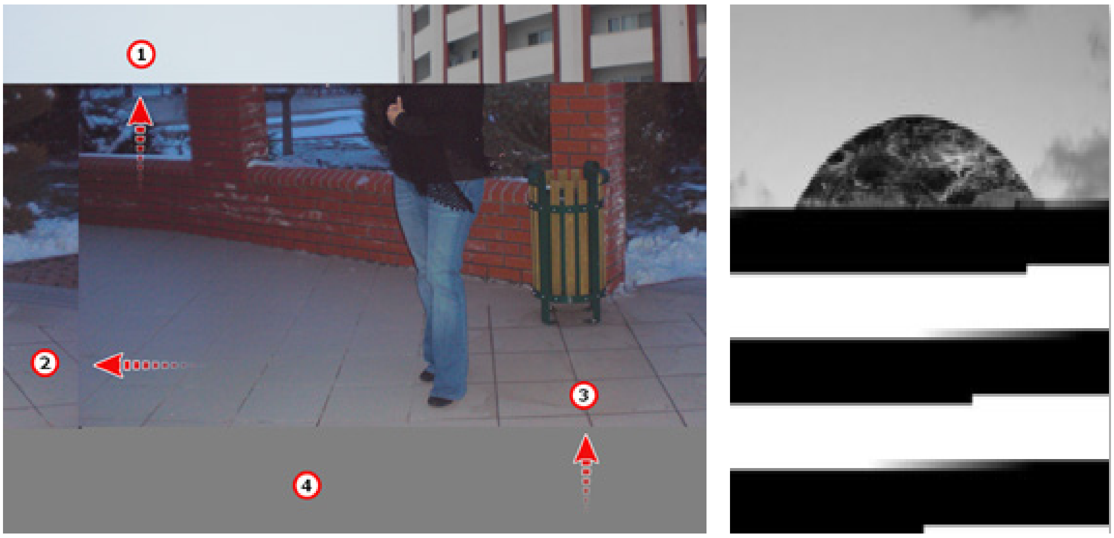


Figure 10. Fragmented files not coming in full

- That the software in which this process is performed recovers the file system structures and recovers the files on the disk hierarchically (tree structure) over it, or it is necessary to try alternative software.

* File corruptions: deliberate or accidental power cuts during file operations (move, copy, delete), not using the remove hardware option.

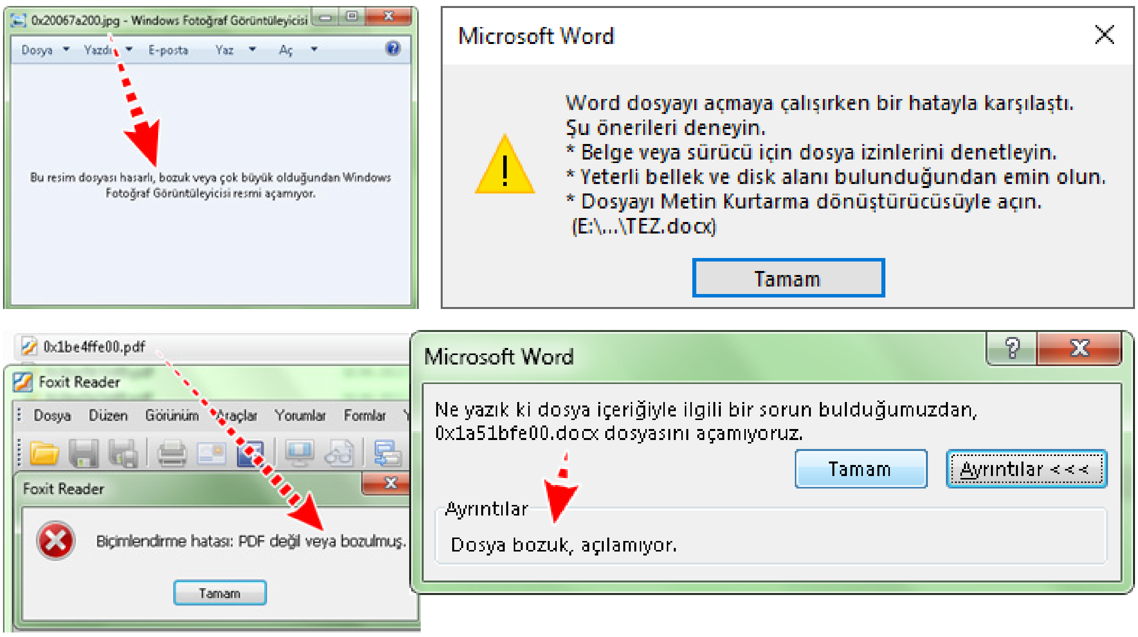


Figure 11. Examples of corrupt files

- If the methods described in Figure 10 cannot be obtained, a positive result can be obtained as a result of searching the temporary (.tmp) files created according to the file type (such as automatic save for office files) if an automatic backup is made while working with the relevant file on the computer. If the positive result is still not achieved, file repair (corrupt file repair) software (eg S2 Recovery Tools) should be tried on the recovered file. In Figure 12, pre- and post-repair images of a broken video part are presented.

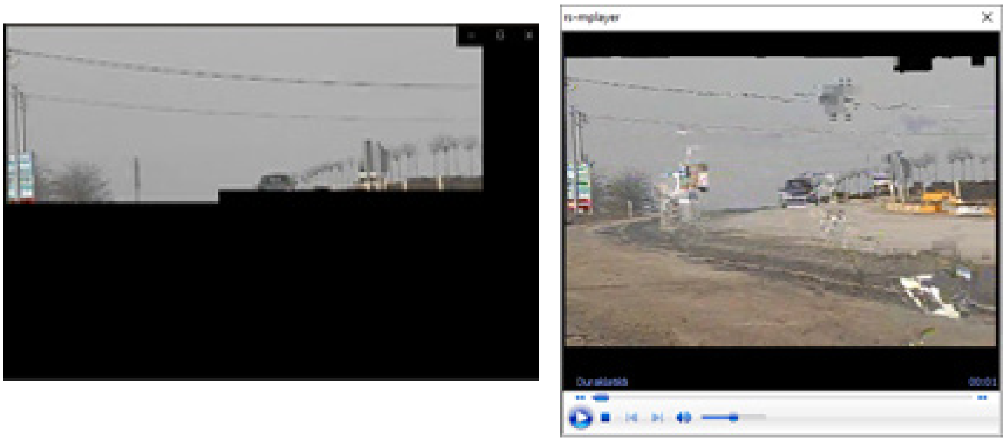


Figure 12. Repaired video example

* Intentional encryption of files or encryption by viruses (ransomware):

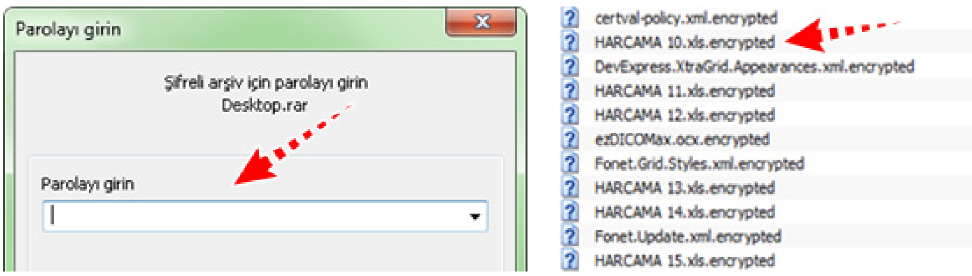


Figure 13. Encryption on files

- The surest way to recover data from files in this situation is to learn the password of the file. If it is not possible to learn the password, a brute force attack should be made by means of special software (Passware Kit Forensic, Hashcat, John the Ripper, etc.) If it is a Ransomware attack, developed software should be used to decrypt this Ransomware (MalwareHunterTeam, 2022; No More Ransom, 2022) or shadow copies (Shadow Copy - ShadowExplorer) should be used.

* To access data as a result of burning the motherboard of the Nas device:

Figure 14. Inability to access data due to corruption of Nas device

- After detecting the Raid level on the Nas device (Raid0), software that can create software Raid in the computer environment (PC-3000 Express RAID System, UFS Explorer, Raid Reconstructor, R - Studio, Softcenter the build must be rebuilt with Raid Explorer, etc.).

* Permanent wiping of data by users or viruses:

  * In such a case, data recovery is not possible.

Hardware Data Loss

In the previous titles, the software and hardware structures of data storage units were mentioned in detail. The mentioned components ensure the healthy operation of the data storage unit. If one or more of these components fail, the device will not work and data will be lost.

Hardware data losses; These refer to the situations in which the data recording unit does not work in any way as a result of exposure to natural disasters (earthquake, flood, landslide, fire), environmental factors (electrical fluctuation, static electricity, electromagnetic effect) or sabotage (throwing into acid, crushing, breaking, throwing into water). In short, it refers to all kinds of internal or external factors that will prevent the operation of the data recording medium. The aforementioned factors that the data logger is exposed to may have developed in accordance with the natural flow of life or may have been consciously carried out by a user or a cybercriminal. According to Figure 2, hardware data loss is seen as a 40% and 38% problem.

Data Loss in Optical Discs (Media) and Solutions

Since recording discs using optical write-reading technology do not have any casing or frame to protect them from external factors, they can be easily damaged.

Before moving on to the solutions, let's explain the terms to be used in the solutions:

Intervention With Electron Microscopy**:** It is to obtain a data stream from the images obtained from the scanned parts of an optical disc using an electron microscope, and then to interpret the data obtained with software specially developed for this work

(Jarolím, 2016).

Tagging Intervention: It is the fixation of the cracked part of an optical disc with a crack with a paper tape or label. The point to be considered here is that the adhesive property of the tape or label to be used should not be too strong. Otherwise, the coating of the pasted surface may be removed.

Filling Process: It is the use of specially developed optical disc cleaning and filling devices (Disc Body) to fill the scratches on the surface if it is deeply scratched on the optical disc.

Sanding Process: If there are not very deep scratches on the optical disc, these parts can be moistened and the data can be accessed by gently rubbing with very fine sandpaper (2000+) without causing new scratches.

Intervention With a Laser: The damaged part of the broken optic disc is cut with a laser from the broken places. Subsequently, a piece of a solid blank optical disc of the same size is laser cut and replaced with the slice removed from the previous optical disc. It should be adhered to the reflective face to reflect the light. It should be noted that there is no residue and level difference (Kara, 2013).

Heat Treatment Application: If the optic disc has taken a concave shape, a device (LCD Screen) capable of applying heat treatment and vacuuming an optical disc is placed on the separator. It is to wait for the optical disc to take its original shape by placing a suitable weight on the optical disc.

Cleaning Process: If the reading side of the optical disc is dirty, it is the process of first cleaning the solid dirt on it by holding it under water and then cleaning the sticky or water-removable dirt on it with isopropyl alcohol or cologne with the help of a cotton ball.

If the hardware problems specific to optical discs are to be exemplified;

* If the surface of the optical disc is dirty:

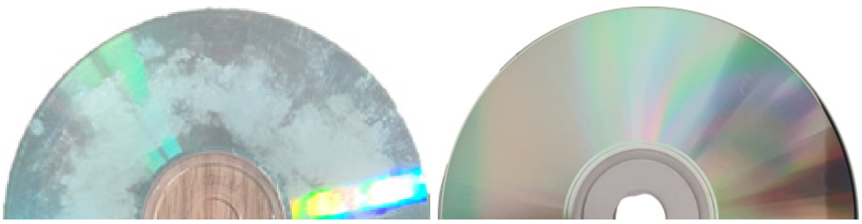


Figure 15. Dirty and cleaned optical disc surface

- The disc surface should be cleaned.

* Cracked optic disc:

Figure 16. A crack in the optic disc

- Cleaning and labeling processes should be applied to the disc surface.

* Fracture of the optic disc:

Figure 17. Fracture of the optic disc

- Disk surface cleaning and laser intervention should be done.

* Light scratches on the optical disc:

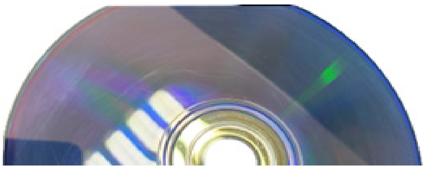


Figure 18. There are light scratches on the optical media

- The disc surface should be sanded and then cleaned.

* Having deep scratches on the optical disc

Figure 19. Having deep scratches on the optical disc

- The disc surface should be filled and cleaned afterward.

* Error during data recording to optical disc

Figure 20. Error during data recording to optical disc

- Data recovery should be attempted by trying alternative brands of optical disc readers.

* Wear of the reflective surface of the optical disc

Figure 21. Wear of the reflective surface of the optical disc

- The worn dimple should be covered with a material that will reflect light.

* Optical laser burning of the optical disc

Figure 22. Optical laser burning of the optical disc

- Data recovery should be attempted by trying alternative brands of optical disc readers.

* Distortion of the physical shape of the optical disc

Figure 23. Example of a physically deformed optical disc

- Heat treatment should be applied.

In cases of overwriting or deleting the unlogged optical media, software data recovery operations should be performed by making a logical copy of the optical media. In addition, in case the optical disc is not recognized by the optical disc reader, alternative brand optical disc readers should be tried and data recovery should be tried.

Data Loss and Solutions on Hard Drives

Data loss for hard drives; Printed circuit board failures, electronic components on the printed circuit board, motor malfunctions as a result of the disc being kept unused/ forgotten for a long time, deformation in the flex cable that provides the motor power connection, the pad available for the read/write heads on the printed circuit board to perform their necessary functions. Dirt/abrasion of the 's, scratches/scrapes on the surface (platter), not being able to read the modules as a result of firmware (firmware) failure, inability to function as a result of contamination of the read/write heads, sticking of the read/write heads to the disk plate (platter), breakage of the read/write heads, etc. occurs in such cases (Schroeder and Gibson, 2007).

Before moving on to the solutions, let's explain the terms to be used in the solutions:

Repairing The Device: It refers to the normal operation of the device by eliminating the software or physical malfunctions that prevent the data recording medium from working normally.

Donors: The basic components (printed circuit board, hard disk assembly) that make up a data recording environment and one or more of the parts that make up the subunits of these components (electronic components on the printed circuit board, read/write head, engine) are compatible and healthy working data. It is the process of taking the defective data from the recording unit to the recording unit. Data logger manufacturers (except Seagate) do not perform data recovery operations other than production. Therefore, there is no valid method/recipe to find a 100% compatible donor for the data registers. There is no publication or documentation of the manufacturers on this subject. Instead, documentation or blog posts published by data recovery companies and companies selling data recovery tools are used, but the criteria shown are not met. It is also stated that there is no 100% guarantee of finding a donor. The best donor device is exactly the same except for the serial number.

Donor Search for The Printed Circuit Board (PCB): If the power connection of the disc is made and no sound or movement is felt from the disc, it is considered as the source of the fault. Donor search should be done by finding the relevant number on the card and searching for the donor card according to that number (Donor Drives LLC, 2012). Example card numbers according to brands are presented in (HDDZone, 2010). As can be seen from the examples, PCB numbers are in different standards and locations according to brands. An alternative method is to search for donors by identifying the manufacturing family of the hard disk (Figure 25). In addition, when the donor of the defective printed circuit board is found, it is often not sufficient to replace the cards directly. The chip on the card must also be carried in the rom, which contains the hardware software (firmware) that ensures the healthy operation of the hard disk. If there is no physical defect on the PCB, no sound is heard from the disc and no movement is felt, the source of the fault is the PCB.

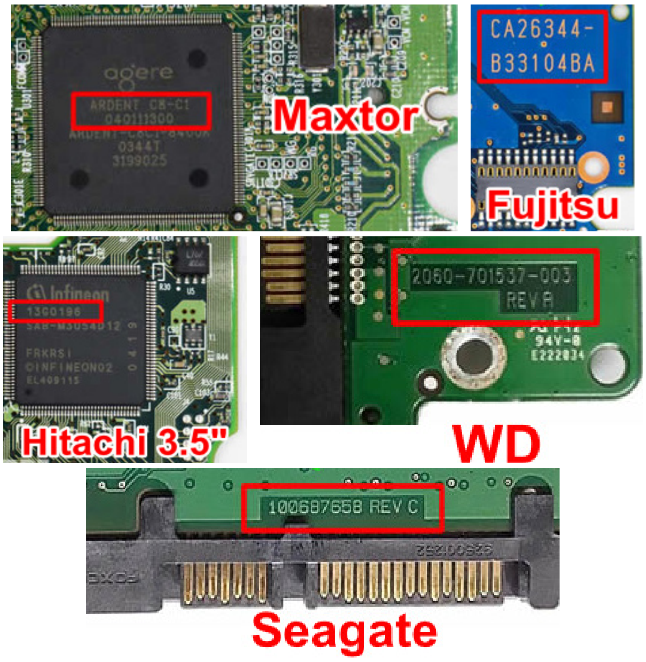


Figure 24. Criteria to be searched according to PCB brands

Western Digital		
Minimum PCB requirements	The middle characters in the second part of the model information. These characters indicate the family information to which the disc belongs. (Family ID)	WD6400AAKS-65 **A7** B0 WD10TMVW-11 **ZSM** S1
Seagate (Barracuda / F 3 Mimarisi)		
Minimum PCB requirements	Seagate brand drives produced before 2008 were produced with this architecture. There is no need to change ROM on these discs. Most likely a chip transfer will be required.	ST9640320AS ST31000528AS
Toshiba		
Minimum PCB requirements	Chip Transfer is required in most cases.	MK3265GSX MK1017GAP
Samsung		
Minimum PCB requirements	Chip transfer / swap or similar services may be required for new models.	HA250JC HM120JI SV0602H

Figure 25. Compatibility criteria for PCBs (Güllüce and Benzer, 2015)

Donor Search for Read/Write Heads and Motor: Although it is the easiest type of malfunction to detect, it is the most difficult compared to other types of malfunctions due to the highest level of operation sensitivity. This is because the read/write heads work very close to the platter. The main symptom is unusual noise from the disc. If there is such a malfunction in the disc, it is the sound of "squeaking", "click- knock", and "pop-pop" when the power is supplied to the disc. These sounds vary according to the brand and model of the disc and the type of malfunction (head sticking, head breaking, etc.). These sounds created by the discs are cataloged according to their brands and types, from the website of the Hdd Guru community (files.hddguru.com/download/Hard Disk Clicking Noise/) can be listened to. In cases where there is no sound but no data can be received, to test whether the heads are sound or not, they should be connected to advanced hardware that can send disc manufacturer-level signals regardless of the operating system, and through the software of these hardware (Atola, Deeppsar, Dolphin Data, MRT Lab, PC 3000, Salvation Data, etc.) should be tested with head test commands. Sometimes it may not come at all. In addition, sometimes these sounds can occur when an incompatible PCB is attached to the disc.

Donor for printed circuit board Similar to the criteria, a donor compatible with the defective disc is sought. It is shown by coloring on the photographs of the discs that the criteria are higher than on the printed circuit board. Donor Example discs with marked criteria are presented in Figure 26.

RED – Strictly required.
ORANGE – High priority.
YELLOW - Medium priority.
GREEN - Low priority.

Figure 26. Donor sample discs with marked (Donor Drives LLC, 2012)

Also, one of the best ways to find a good donor is to look at the preamp for Seagate and the microjog for WD. However, this is very difficult to do because it does not have the software and hardware to enable end users to see these values.

Preamp/microjog values match, there’s a good chance the heads are compatible. When you give the command “ctrl + l” (L) from the terminal for Seagate or “View Rom Info”

for WD disks, many values on the ROM will be displayed about the disk. One of them is the Preamtype value. For Seagate, for example, these values are "CC 16" or "B2 03" or "CA 05". To be a donor, at least the first two characters must match.
Head changes are in a dust-free clean room (clean room) cleaned with HEPA filters. It is performed by moving the head from the healthy disk to the defective disk in special areas called room). Special equipment is used for these changes.

Figure 27. Outer parking head replacement

To SATA Convert:

In hard disks with USB interface;

- The disc cannot be recognized by computers when power is on,
- Failure to fix the faults in the data paths with socket replacement,
- More commands can be run through the Sata interface,
- Since data recovery software (PC-3000, MRT, etc.) used with advanced data recovery hardware needs a Sata connection interface to send basic level commands such as ATA and PIO, disks must be converted to a Sata interface.

There are two ways to access USB disks with the Sata interface:
By looking at the production diagram (datasheet) of the controller, after determining which of the paths to the controller are Sata buses, the first method is to solder those paths to the SATA socket. Soldering operations should be done from the back of the capacitors at the input of the USB - SATA bridge circuit, according to the bus outputs, as shown in Figure 28 for the electrical connection, after the USB cable is attached to the PCB, the multimeter is brought to the diode mode (buzzer). After one end of the probe is touched from the USB cable end to the 5V pin, it is determined which point it corresponds to on the PCB. When the same process is done in the ground connection, the electrical pins on the PCB are also detected.

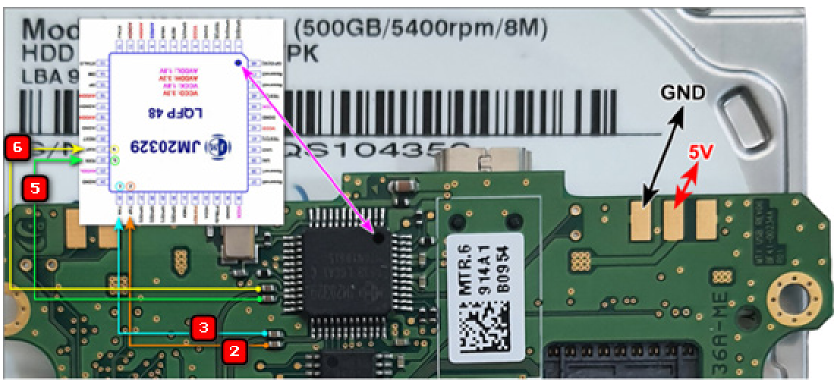


Figure 28. Soldering points when converting USB disk to Sata

To move the rom or replace it with a compatible PCB by writing the rom data. The list of compatible PCBs that can be used for Toshiba, Samsung, and WD brand USB disks is presented in Table 2.

Table 2. USB and SATA PCB Compatibility List (ACELab Europe s.r.o., 2014; Jared, 2016a; 2016b)

	USB PCB No	**Sata PCB No.**			**USB PCB No**	**Sata PCB No.**
SAMSUNG	100725482	100720903		**WESTERN DIGITAL**	701605	701499
	100760718	100720903			701615	701499
	100740633	100739392			701635	701572
	100765396	100767980			701650	701499
	BF41-00300A	BF41-00306A			701675	701672, 701609, 771672
	BF41-00365A	BF41-00354A			701737	771692, 771692
	BF41-00373A	BF41-00354A			771754	701692
	BF41-00325A	BF41-00315A			771754	701692, 771692
	BF41-00357A	BF41-00315A			771761	771692
	BF41-00282A	BF41-00249B			771801	771823
	BF41-00311A	BF41-00315A			771814	771820
	BF41-00231B	BF41-00157A			771817	771820
					771859	771852
TOSHIBA	G3448A	G003235B, G003235C			771939	771959
	G003309A	G003235B, G003235C			771949	771959, 771939
	G003296A	G003138A, G003235C			771961	771960, 771933, 771939, 771959
	G003250A	G003138A			771962	771931, 771959
	G003054A	G002825A			771964	771983
	G3959A	G3918A			800038	800025
	G003189A	G003138A			800041	800022
	G4330A	G4311A			800067	800065
	G0034A	G4311A			800069	800066
	G0039A	G4311A			810003	800065
	G003296A	G003138A			810012-000	800022-002
	G003309A	G003235B				
	G3711A	G3235C				

SSDs with M.2 or other interfaces to convert to Sata, special converter cards must be used as presented in Figure 29.

Figure 29. SSD conversion cards with (Acelab Europe s.r.o., 2022)

The following examples of hardware problems for the hard disk can be given;

* Contamination of pads to motor or read/write heads:

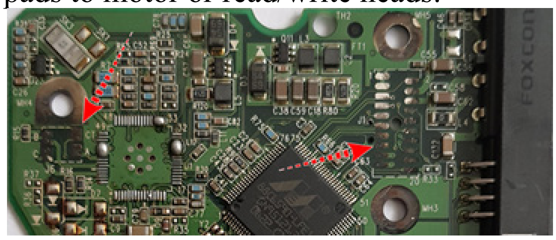


Figure 30. Contamination of pads going to a motor or read/write heads

- Related pads fiberglass should be cleaned with a pencil or eraser.

* Breaking the hard disk bus:

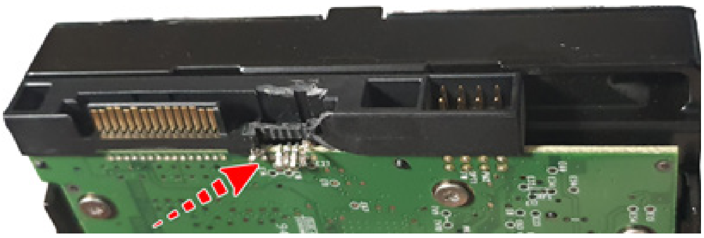


Figure 31. Breaking the data path

- The PCB can be changed (with the rom) as well as only the temporary connections to the buses can be soldered.

* Breaking of circuit elements from PCB:

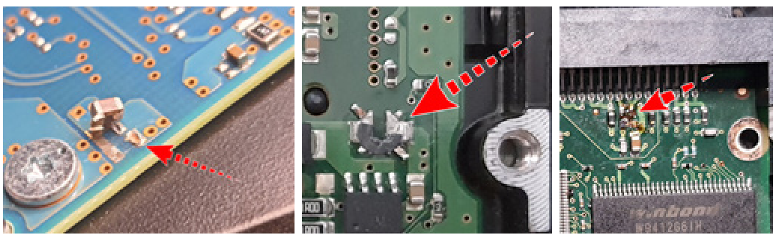


Figure 32. Breaking of circuit elements from PCB

- If the part can fulfill its function, it should be soldered, if it cannot fulfill its function, the donor should be provided.

* Loss of function of parts on the PCB:

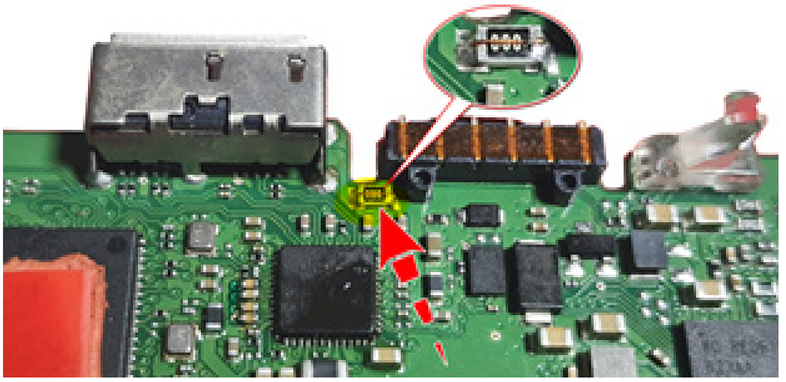


Figure 33. PCB with broken fuse resistor

- It is necessary to eliminate the malfunctions of the electronic parts on the PCB. Figure 33 how to bypass the fuse resistor using a jumper to repair it.

* Path break in PCB:

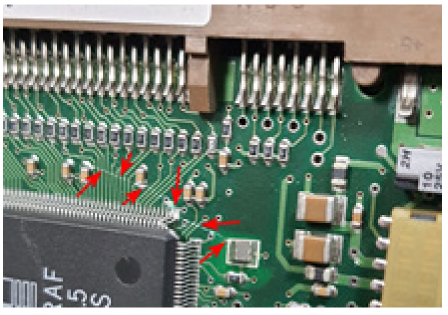


Figure 34. Example of path break in PCB

- If the PCB can be repaired, it should be performed using jumpers, if it is beyond repair, the PCB should be replaced by obtaining a donor.

* Crushing PCB or HDD metal case:

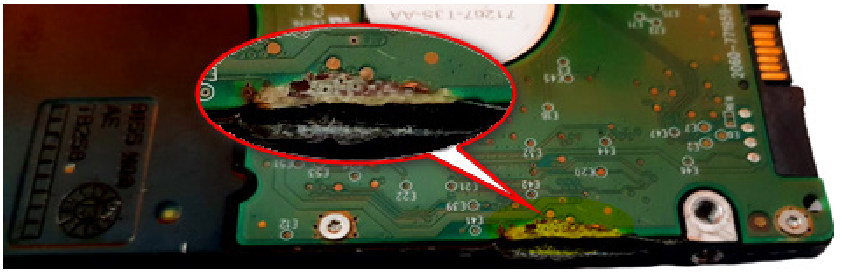


Figure 35. Crushing PCB or HDD metal case

- The donor should be obtained, the ROM should be moved and the PCB should be replaced.

* Combustion of parts on the PCB:

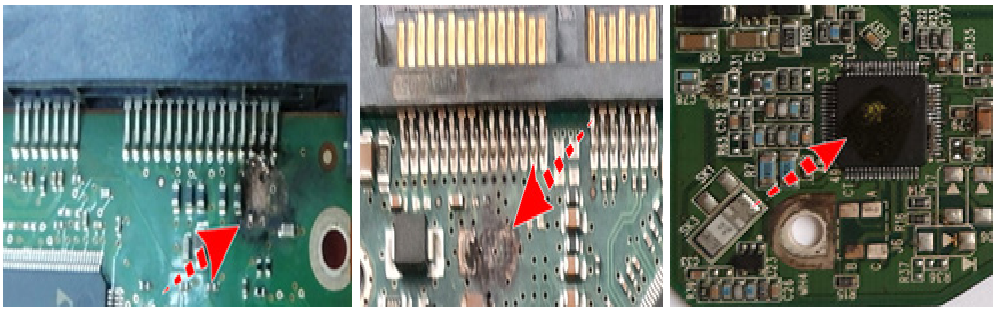


Figure 36. Part burn-in on PCB

- The burnt part(s) should be replaced by obtaining a donor. TVS diodes often burn out due to sudden voltage changes.

* Pin to the motor not touching the pad:

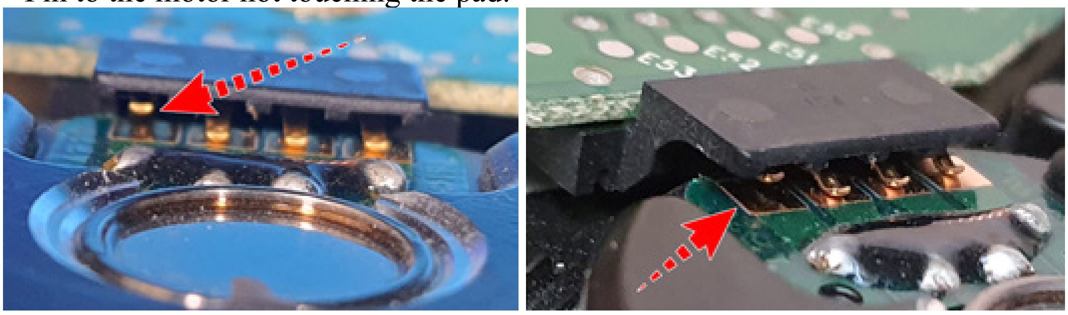


Figure 37. Pin to motor pad contact

- The pins pad should be touched.

* Failure of read/write heads:

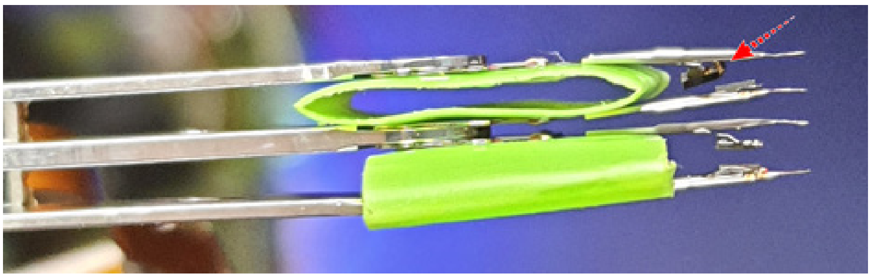


Figure 38. Failure of read/write heads

- A donor read/write head should be used.

* Formation of internal path break in PCB as a result of crushing:

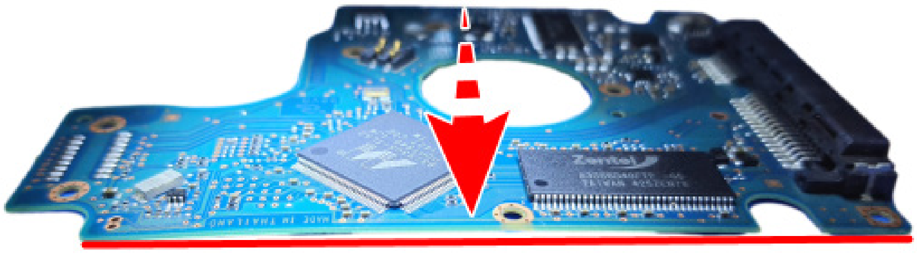


Figure 39. Internal path break in PCB as a result of bending

- The donor should be obtained, the ROM should be moved and the PCB should be replaced.

* Loss of read/write head:

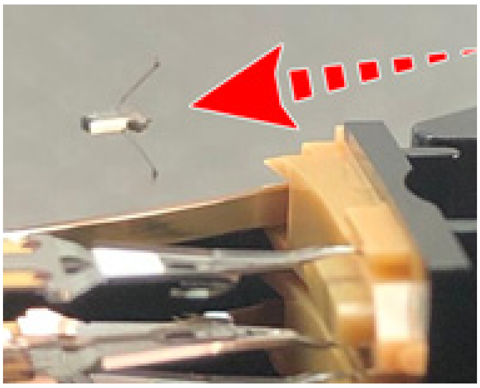


Figure 40. Break of read/write head to the platter

- A donor read/write head should be used.

* Hard disk USB socket clogged:

Figure 41. The hard disk USB socket clogged

- The socket must be cleaned.

* Sticking of read/write heads:

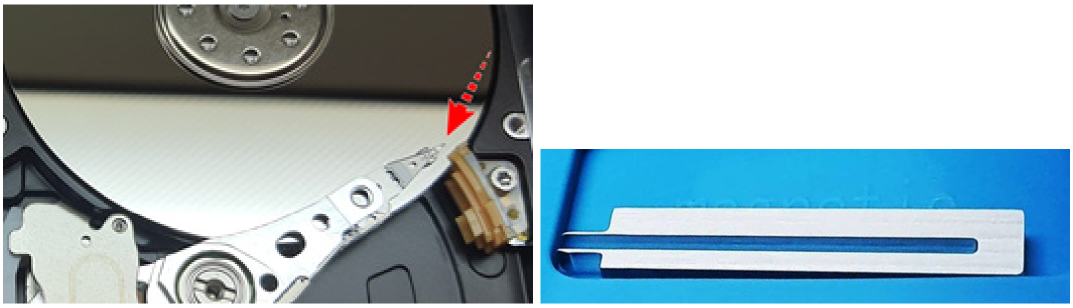


Figure 42. Adhesion of read/write heads and head spacer

- Head snapping/sticking is usually caused by the absence of airflow, or the motor not turning Head sticking also occurs when a circuit element on the PCB is electrically burned. Sticking heads should be transported to the park using the head spacer presented in Figure 42. An alternative and risky method is to pull the read/write heads to the park by turning the read/write head clockwise while the platter is turned counterclockwise.

*   Breaking of the power cables to the motor:

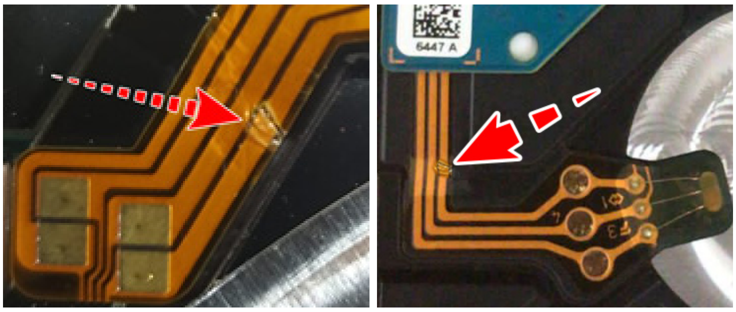


Figure 43. Breaking the power cable to the motor

- The power cable must be soldered or soldered with a jumper.

*   HDD with scratched plate and broken read/write head

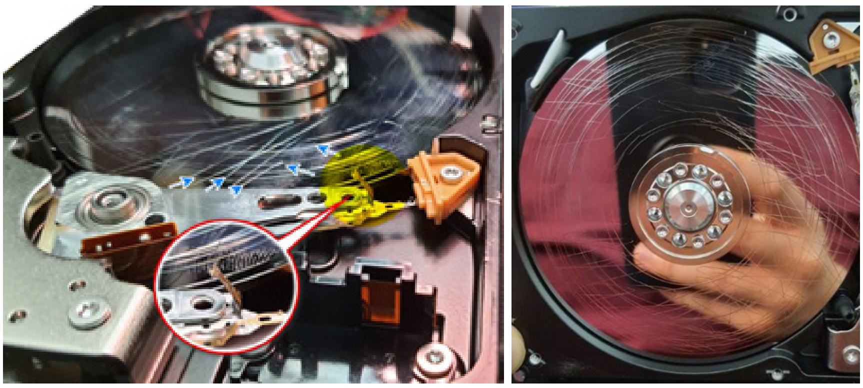


Figure 44. HDD with scratched plate and broken head

- Heads should be replaced by obtaining a donor, and data from the lower surface of the scratched plate and other plates should be tried.

*   Contamination of read/write heads

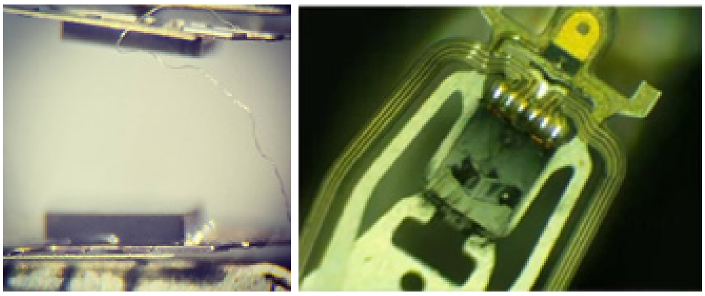


Figure 45. Contamination of read/write heads (Tri-State Data Recovery & Forensics LLC, 2020; ACELab Europe s.r.o., 2021)

- A donor read/write head should be used. As an alternative and risky procedure, it can be cleaned with isopropyl alcohol under a microscope.

* The disc is exposed to a natural disaster or sabotaged:

Figure 46. Hdd inside burnt DVR device

- If the heat did not affect the plates of the disc, the plates and rum should be transported by obtaining a donor.

* Unable to access the disk from the USB interface:

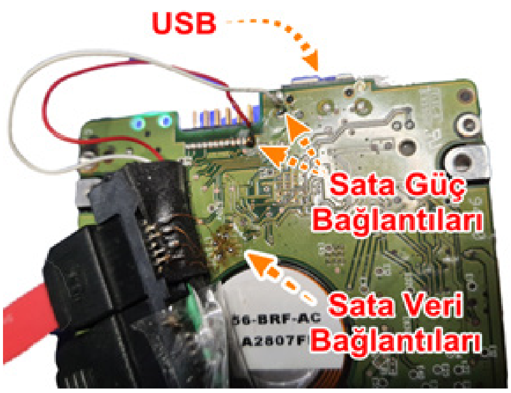


Figure 47. Example of SATA bus made to disk with bad USB socket

- Compatible Sata PCB must be used or converted to Sata.

* PCB that will block/redirect the electric current:

Figure 48. Presence of materials that will block the electric current in the PCB

- Relevant material needs to be removed from the PCB.

* Local scratches on disc platters:

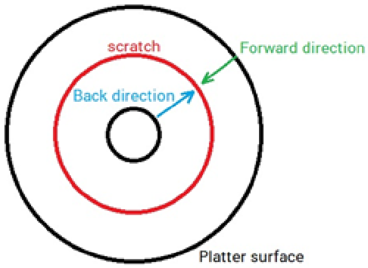


Figure 49. Presence of materials to prevent electric current in (Acelab Europe s.r.o., 2018)

- In the parts where there is regional scratching (red line), a forward reading should be started from the first sector of the disk via the data recovery software, and a backward reading should be made from the last sector of the disk when the parts with regional scratching come. In this way, the scratched region will be skipped and the heads will not be damaged during reading.

* Intervention by non-specialists:

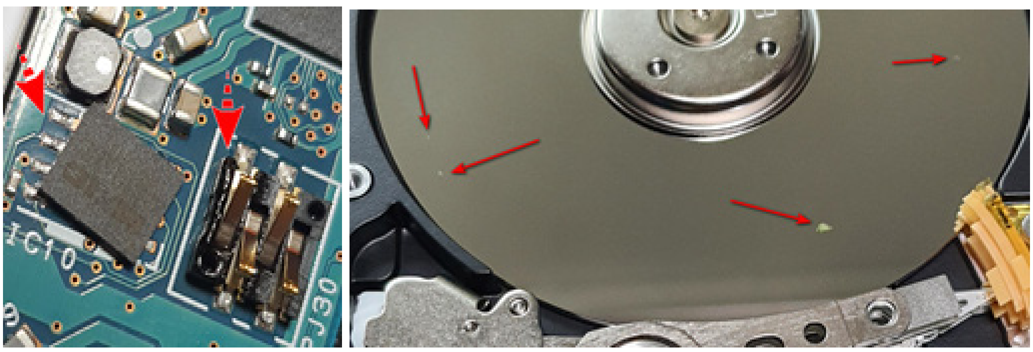


Figure 50. Intervention by non-experts

- If the intervened parts are still intact, the operation of the device is done correctly, if the parts do not work, a donor should be obtained.

- Intentional breakage of the printed circuit board:

Figure 51. Intentional breakage of the printed circuit board

- Donor printed circuit board must be provided, and the rom on the original board should be moved to the donor board.

* Physical damage to the disk

Figure 52. Physical damage to the disk (Recovery HDD, 2018)

- If the platters of the disc are not damaged, the platters should be moved to the donor disk.

- Failure of the motor to rotate the plates:

  - PCB of the disc and there is no audible sound, if the disc does not rotate despite the power supply, it is necessary to talk about a motor failure. It occurs when the engine oil in old-generation hard drives that have not been used for a long time loses its function. If the engine is heated with a hot air gun, the engine oil will regain its former functionality. Alternatively, the platters should be moved to a donor hard disk.

After the data recovery processes from 190 optical devices, which we performed on optical media and whose detailed information is given in Table 3, it was observed that data was recovered from 146 of them and data could not be recovered from 44 of them.

Table 3. Optical Disk Data Recovery Research Results

Problem	Type	Number of Optical Discs	Action Taken	Is data recovered?
Unable to copy files	CD	6	File-based forensic copying	Yes
	DVD	18		

Crack	CD	1	Labeling	Yes
Lightly scratched and stained	CD	22	Cleaning	Yes
	DVD	12		
Moderate scratch-es and stains	CD	29	Filling and Cleaning	Yes
	DVD	24		
Deep scratch	CD	5	Sanding and Cleaning	Yes
	DVD	8		
Deformity	CD	1	Heat treated	Yes
	DVD	2		
Unable to Read Data	DVD	18	Reading with a different reader	Yes
Broken	CD	5	Cleaning, Sanding, Fill-ing	No
	DVD	2		
	Blu-ray	4		
Reflective part removed	CD	6		
	DVD	4		
Very deep scratch	CD	1		
	DVD	6		
Coated with glue	CD	11		
	DVD	5		

The detailed information of 338 data recovery processes that we performed on hard disks and solid state disks is given in Table 4. 137 of them are hardware problems. 29 of them are hardware and semi-hardware problems. 23 of them are hardware and software problems. 86 of them are semi-hardware problems. 2 of them are semi-hardware and software problems. It has been observed that 61 of them are software. During the 338 data recovery processes we performed, when the methods described above were applied, it was observed that 284 data were recovered.

Table 4. Types Of Failures Encountered In Data Recovery Operations

Fault Type	**MOQ**
Hardware	137
Hardware and semi-hardware	29
Hardware and software	23
Semi-hardware	86
Semi-hardware and software	2
Software	61
Total	338

Table 5. General Data Recovery Status

Situation	MOQ
Data recovery done	284
Data recovery not possible	54
Total	338

Table 6. Data Recovery Only in Hardware Failures

Situation	MOQ
Data recovery done	79
Data recovery not possible	54
Total	133

Table 7. Types of Hardware Failures by Brands

Brand/ Fault Type	HGST	Hita-chi	Ki-oxia	Sam-sung	Sea-gate	Sk Hy-nix	Toshi-ba	WD	To-tal
Head failure	2	4		13	29		5	21	71
Head weakening				4	9		3	5	21
engine failure				1	5		4	1	11
PCB failure	1		1	6	9	3	2	5	23
Platter failure							2		1
USB socket corruption								1	1
Burnt								1	1
Total	3	4	1	23	52	1	12	33	129

Table 8. Data Recovery Only in Case of PCB Failures

Situation	MOQ
Data recovery done	18
Data recovery not possible	7
Total	23

Table 9. Data Recovery in Head Failures

Situation	MOQ
Data recovery done	39
Data recovery not possible	37
Total	76

A converter board has been developed to ensure that HDDs or M.2 SSDs with failed USB interfaces can be accessed via the Sata interface. The image of this developed card is presented in Figure 53.

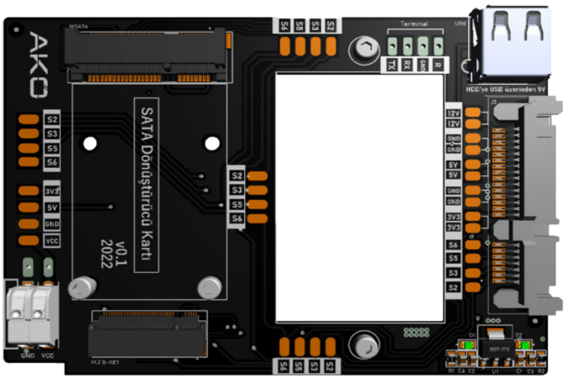


Figure 53. The card we developed for access from the Sata interface

Thanks to this card, all of the operations listed below can be performed without the need for any other hardware.

- To enable msata conversion,
- To enable M.2 – B Key SSD conversion,
- To ensure the safety of soldered card connection cables,
- To provide terminal connection to disks when necessary,
- Apart from these, converting SSDs that can be converted to Sata,
- To facilitate the soldering process and facilitate the transport of the soldered board,
- A converter board has been developed to provide socket connections for 1.8”, 2.5”, and 3.5” hard disks with broken power or data sockets.

The root cause of problems with optical discs is problems with the optical disc surface. If these problems are listed;

* The surface of the optic disc is dirty,
* The optical laser burns the optical disc,
* Fracture or cracking of the optic disc,
* Distortion of the physical shape of the optical disc,
* An error occurred during data recording to the optical disc,
* It can be listed as light or deep scratches on the optical disc.

The main causes of problems in hard disks are the electronic circuit elements on the PCB, the motor, and the read/write heads.
The main ones are listed;

* All kinds of reasons that prevent electrical conductivity, such as deterioration or fragmentation of circuit elements on the PCB,
* All kinds of motor-related reasons prevent the rotation of the platters and the platters,
* It can be summarized as any reason that prevents the read/write heads from functioning.

Process Steps to Follow During Data Recovery

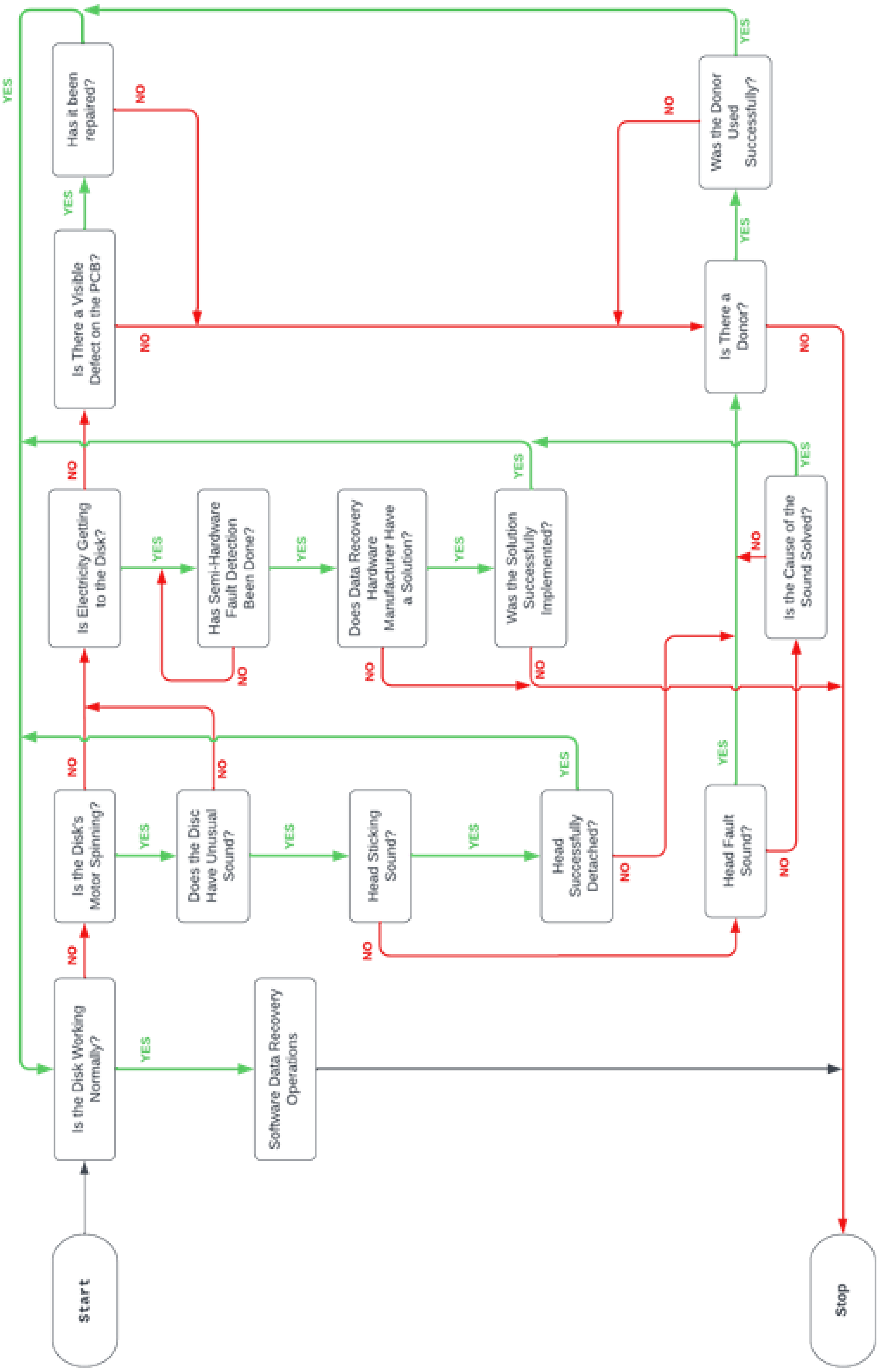


Figure 54. Data Recovery Steps on Hard Drives

References

ACELab Europe s.r.o., 2014, List of compatible SATA PCB for 2,5" USB drives, https://forum.acelab.eu.com/viewtopic.php?f=155&t=8376:

Acelab Europe s.r.o., 2018, PC-3000 Data Extractor. Practical advices and Settings Menu Dive In, https://blog.acelab.eu.com/pc-3000-data-extractor-practical-advices-settings-menu.html:

ACELab Europe s.r.o., 2021, ACE Lab Webinar on the Modern Challenges in Data Recovery & Digital Forensics, https://www.youtube.com/watch?v=47eXzxZk-0E:

Acelab Europe s.r.o., 2022, PC-3000 Product Catalogue for Digital Forensics.

David, M. S., 2000, The cost of lost data, *Storage Manag. Solut.*, 4 (6), 60–61.

DeepSpar Data Recovery Systems, 2007, Data Loss and Hard Drive Failure: Understanding the Causes and Costs, http://www.deepspar.com/wp-data-loss.html:

Donor Drives LLC, 2012, Hard Drive Circuit Board Replacement Guide or How To Swap HDD PCB, www.donordrives.com/pcb-replacement-guide:

Durak, R., 2011, Araştırma Yöntem Ve Teknikleri, https://atauni.edu.tr/yuklemeler/e6d32be6d3b1ad4dae5f6cc23165e2da.pdf:

Exterro Inc., 2021, Custom Carvers, https://support.exterro.com/support/solutions/articles/69000765593-custom-carvers:

Guo, Y. and Slay, J., 2010, Data Recovery Function Testing for Digital Forensic Tools, *Advances in Digital Forensics VI*, Berlin, Heidelberg, 297-311.

Güllüce, Y. Z. and Benzer, R., 2015, Hard Disk Failure And Data Recovery Methods In Computer Forensic, International Journal of Human Sciences, 12 (1), 206-225

HDDZone, 2010, Hard Drive PCB Replacement Guide, www.hddzone.com/conditions.html:

IBM Security and Ponemon Institute, 2021, Cost of a Data Breach Report.

Jared, 2016a, Samsung USB to SATA PCB Compatability, www.data-medics.com/forum/threads/samsung-usb-to-sata-pcb-compatability.1438:

Jared, 2016b, Toshiba USB to SATA Guide, www.data-medics.com/forum/threads/toshiba-usb-to-sata-guide.1305/:

Jarolím, J., 2016, Reconstruction Of Data From Damaged Cd/Dvd, Master Thesis, Brno University Of Technology.

Kara, Ş., 2013, Evaluation Of The Achievements Of Data Recovery Methods, Master Thesis, Firat University.

Keskin, N., Kıran, A. N., Eğdemir, F. K. and Eren, T., 2020, Multicriteria Decision Making Methods and Service Provider Selection to Cloud Information Security Requirements, International Journal of Information Security Engineering, 6 (1), 45-60.

Kessler, G., 2022, Gck's File Signatures Table, https://www.garykessler.net/library/file_sigs.html:

MalwareHunterTeam, 2022, ID Ransomware, https://id-ransomware.malwarehunterteam.com/:

No More Ransom, 2022, Decryption Tools, www.nomoreransom.org/en/decryption-tools.html:

Özdemir, A. and Gülcü, Ş., 2021, Causes Of Digital Data Loss And Data Recovery Methods, *2nd International 5 Ocak Congress on Applied Sciences*.

Özdemir, A., 2022, Investigation of data recovery methods and development of auxiliary equipment for data recovery, *Necmettin Erbakan University*, The graduate school of natural and applied science (Master thesis).

Recovery HDD, 2018, Atropelado sim senhor... www.instagram.com/p/BoxTBRPlKbV/:

Schroeder, B. and Gibson, G., 2007, Disk Failures in the Real World: What Does an MTTF of 1, 000, 000 Hours Mean to You?, p.

Tri-State Data Recovery & Forensics LLC, 2020, www.instagram.com/p/CGGW131jvQQ/: [22.06.2022].

About the Authors

Abdulkadir ÖZDEMİR is a police officer who has been working in Cyber Crime Branch Offices in different provinces since 2012. He has a master's degree in Computer Engineering from Necmettin Erbakan University. The main areas of work are digital data recovery, installation, and management of digital evidence inspection systems, malware analysis, and digital data inspection. He continues to work on various hardware and software developments to be used in digital evidence examination and data recovery processes.

E-mail: by.akozdemir@gmail.com, **ORCID**: 0000-0002-6992-5170.

Şaban GÜLCÜ graduated from the Software Engineering Department of Vienna University of Technology with B.S. and M.S. degrees in 2005 and 2006, respectively. He received a Ph.D. degree in Computer Engineering Department from Selcuk University in 2017. He is an Assistant Professor at the Computer Engineering Department at Necmettin Erbakan University. His main research interests are neural networks, optimization algorithms, and software engineering.

E-mail: sgulcu@erbakan.edu.tr, **ORCID**: 0000-0001-7714-8861.

Similarity Index

The similarity index obtained from the plagiarism software for this book chapter is 7%.

To Cite This Chapter:

Özdemir, A. & Gülcü, Ş. (2022).Scientific Data Recovery Methods and Solutions in Optics Discs and HDDs. In S. Kocer & O. Dundar (Eds.), *Current Studies in Basic Sciences, Engineering and Technology* (pp. 224-257). ISRES Publishing.

Scientific Data Recovery Methods and Solutions in Devices using Flash Memory

Abdulkadir ÖZDEMİR
Konya Cyber Crime Branch Office

Şaban GÜLCÜ
Necmettin Erbakan University

Introduction

Since devices using flash memory are used extensively in daily life, falling, crushing, breaking, and static electricity, etc. they can be overexposed and damaged. Due to their complex structure, data recovery in case of failure is much more difficult than hard disks.

There are two basic data recovery ways:

- The first way is to access the data by fixing/repairing the problem that prevents access to the data.
- The second way is the methods that will be used to access the data if the first method failed or the first method is not applicable. It is to access via Jtag/Isp/ protocols (On- Chip) without removing the data register unit, while the data chip is on the device.

If the data could not be reached with "On-Chip",

- If the device is encrypted,
- If the device only has a data logger,
- The device has an interface different from the standard,
- The device's firmware (firmware) is corrupted,
- The data cannot be accessed (corrupted) via the interface of the device, it is accessed from the eMMC or Nand interface by removing the data recording unit (Chip-Off) from the device.

"On-Chip" and "Chip-Off" methods are available in the literature. It is expressed as direct access to (Rusolut Sp. z o.o., 2022).

Full Swap: It is believed that the device's controller and data chip (s) are intact, and if the donor is available, data can be accessed by moving the controller and data chip(s) to the donor device, this is called full swap.

Jtag/ISP (On-Chip): JTAG, Joint Test Action Group; ISP is short for In-System Programming. Industry standards are used to verify and test printed circuit boards after manufacture. Although the data recording medium works, if all data cannot be retrieved from the standard interface (physical copy) or if the device's malfunction cannot be resolved, it is aimed to retrieve data using Jtag/ISP standards. Special devices and sockets are used to enable this process to take place. During these operations, no part is removed from the printed circuit board of the data recording medium. It is performed by soldering the pinpoints on the Jtag/ISP device around the data chip and the technological test points (pinout) of the data recording medium. The read data of the mobile device is transferred to the computer environment via the USB connection (Johnson and Christie, 2009). Technological test points/pinout to be soldered using Jtag/ISP is provided by the device manufacturer or provided from open sources. Pin names to be soldered for Jtag/ISP are presented in Table 1 and a sample ISP-connected mobile phone printed circuit board is presented in Figure 1.

Table 1. Jtag, ISP, UFS Pin Names to Be Soldered

Standard	Pin Names
JTAG	Gnd, Vcc , Tdi , Tdo , Tclk , Tms , Rtck , Trst
ISP	Gnd, Vcc (2.8V), Vccq (1.8), Clk , Cmd , D0
UFS	Gnd, Vcc (2.8V), Vccq (1.8), Txon (Dino_C), Txop (Dino_T), Rxon (Douto_C), Rxop (Douto_T)

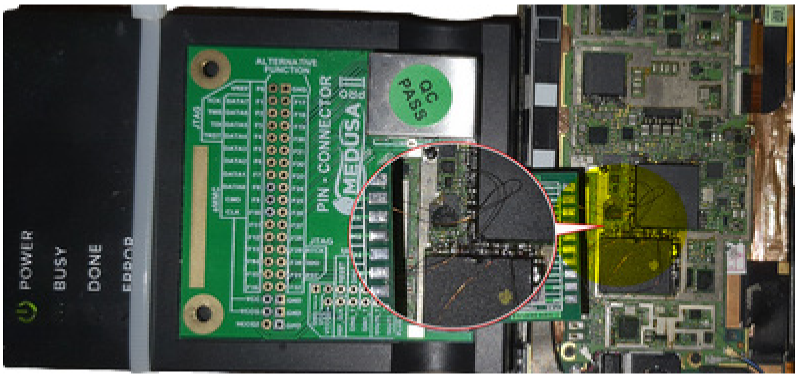


Figure 1. Example of a device with an ISP connection

Rooting, Jailbreak, Download mode for mobile devices, EDL mode or closed physical, file system, etc. offered by forensic software. Data can also be recovered using methods (Doğanay, 2019).

Chip-Off: The devices specially produced for these processes are called Box devices. Memory chips in a device are removed and placed with a special socket developed for these chips. After the socket is plugged into the Box device, the data read over the USB connection of the Box device is transferred to the computer environment. It is used as an alternative method in mobile devices using NAND memory if the data cannot be accessed by the previously mentioned methods. However, it should be used as a last resort (Özdemir and Gülcü, 2021). It is used on devices using the Android operating system and direct access to data can be provided up to the Android v6 version. On most devices after Android v6, the data is encrypted. In Figure 2, a chip-off phone (data chip is removed) and 3 different chip samples placed on the socket of the Box device, where the chips will be read, are presented.

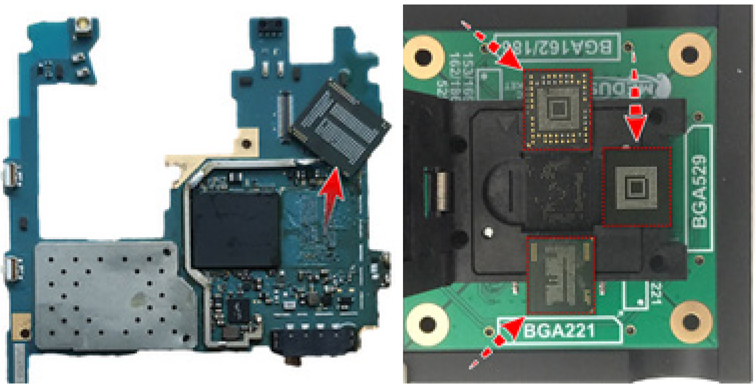
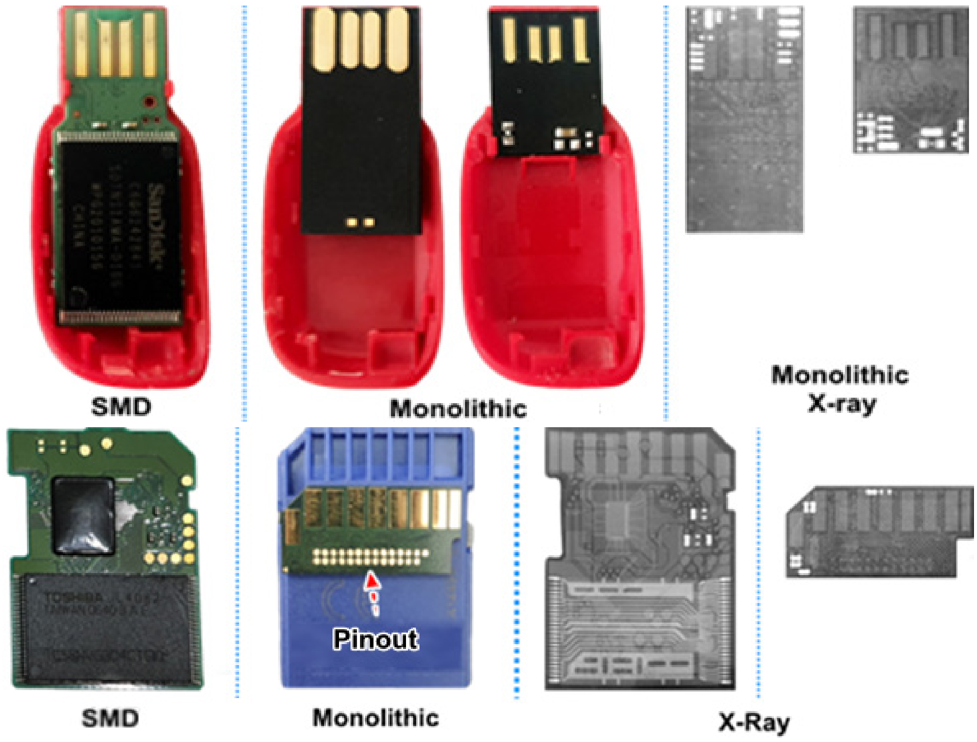


Figure 2. Three different data chips with a Chip-Off device

Nand Interface: According to whether the memory units are embedded or not, it is divided into memory using SMD components and memory with a monolithic structure. It is shown in Figure 3 through detailed examples.

Nand memory in the monolithic devices can be tested during the production phase whether it works correctly or not. These points are named pinout, technological pins, or technological contact points. Thanks to these points, memory manufacturers can test the robustness of the memories, while data recovery companies can perform data recovery operations using the same points. To recover data from the monolithic device, the signals specified in Table 2 must be soldered to the technological contact points/pinout (TTNs) in the data logger shown as an example in Figure 3 (Özdemir, 2022).



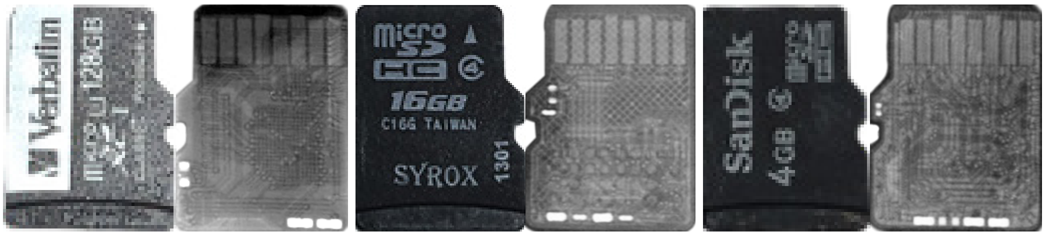


Figure 3. Examples of SMD and Solid type memory using NAND memory

Table 2. Nand Signals and Explanations (ONFI Workgroup, 2021)

Signal	Input/Output	Signal Description
ALE	I	Address Latch Enable is one of the signals used by the host to indicate the type of bus loop (command, address, data). When active, data can be written to an address.
CE_x	I	Chip Enable signal activates the device. The activity of the signal is that the square wave is in a low position. If the signal is not active, the device is in standby mode.
CLE	I	Command Enable signal is one of the signals used by the host to indicate the type of bus loop (command, address, data). This signal should be low when writing commands to the command register.
I/O 0-7	I/O	Inputs/Outputs is an 8-bit wide bidirectional port for receiving addresses, commands, and data to and from the device. (Data bus)
I/O 8-15	I/O	Nand memory is 16 bits, the first I/O 0-7 is used in addition to I/O8-15.
NC		No Connection
NU		Not Usable
R		Reserved for future use.
R/B_x	I/O	Ready/Busy signal shows the status of device operation. When low, it indicates that a write, erase, or read (LUN) operation is in progress. It returns to the high state when the process is complete.
RE	I	Read Enable signal controls the serial data output from the device. When active (low) data is output from the device.
VCC	I	Power
VDD		
GND	I	Ground
VSS		
VSP		Vendor Specific, the function of the signal is defined and specified by the NAND manufacturer.
WE	I	Write Enable signal controls the capture of commands, addresses, and input data.
WP	I	Write Protect signal disables write to memory and erase operations. It is generally recommended to connect to VCC, but the port can also be connected

Figure 4, the connection between the controller, Flash memory, and TTNs, where the data reading process passes from the flash memory to the controller and then to the computer environment via the USB interface.

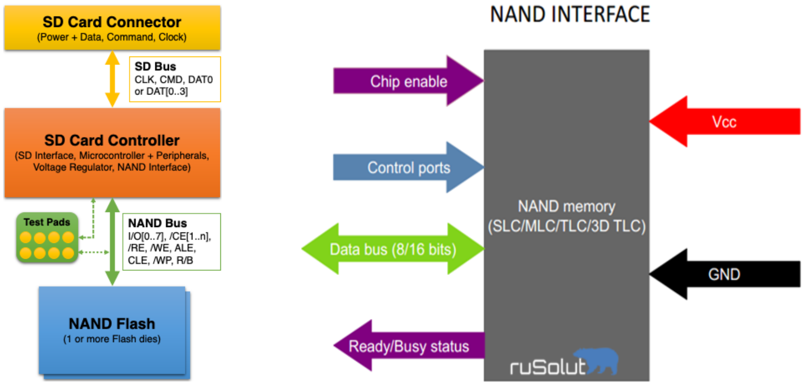


Figure 4. The connection between controller, Flash memory and TTNs, and I/O status (Gin, 2020; Rusolut Sp. z o.o., 2022)

In light of this information, it will be concluded that the flash memory of the devices can be accessed via the USB interface as well as via TTNs. TTNs must be used for data recovery as we cannot access solid devices through their standard interfaces (over the controller). The TTNs of monolithic devices are either covered with a protective layer or covered with tape. To access TTNs, the protective layer above the TTN needs to be scraped with a fiberglass pen or sandpaper first 1000+ then 2000+, or the protective tape removed. If sanding is done by moistening, it does not remove dust and scraping takes place faster because the sandpaper will not clog. Figure 5 shows a micro SD memory card with the adhesive tape removed and a solid USB flash memory (UFD) with an etched coating. Also, TTNs of an SD card are presented in Figure 5.

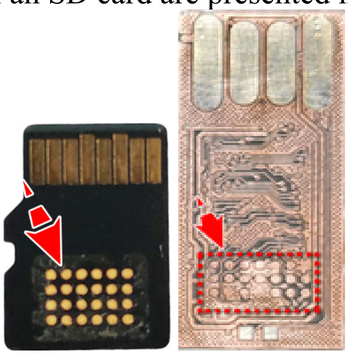


Figure 5. Tape - removed microSD and scraped UFD

This recovery process consists of 2 main stages:

First Stage: It is the process of making TTNs visible, soldering the signals on special equipment, or plugging them into sockets and reading this equipment as a dump (raw) to the computer over the USB connection.

Data writing is through the USB interface to the controller and then to the flash memory. Data is recorded using structures (Abrasion compensation, Bad block management, Erase before writing restriction, etc.) in the Flash memory conversion layer within the controller. In other words, the controller writes data to the (Özdemir and Gülcü, 2021). Figure 6 with Nand connections, microSD, and a Tsop48 data chip placed in the socket. To understand the sensitivity of the jumpers (blue arrowhead) used in the soldering processes, their photos were taken with 1TL.

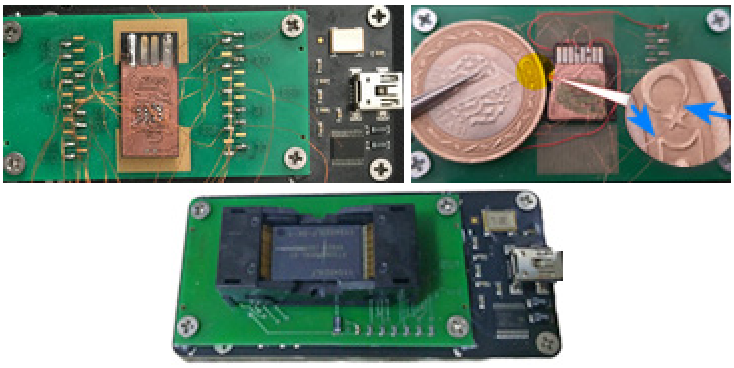


Figure 6. Examples of devices with Nand connections

All detachable circuit elements (including the controller) on the device must be removed since the USB flash drives whose packaging type is Cob/Blob shown in Figure 7 are not memory units that can be removed or scraped to reveal TTNs. After disassembly, the controller's manual (datasheet) should be found. After learning the output signals of the pads where the controller's legs sit from this manual, the actions described above as the first operation should be performed.

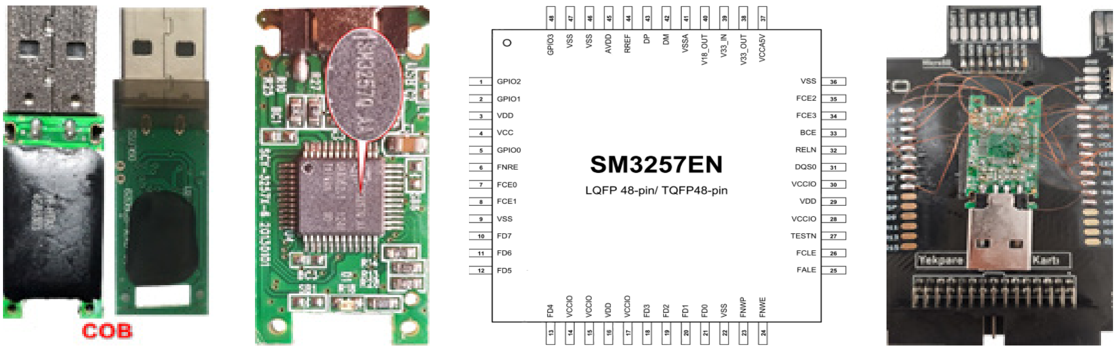


Figure 7. With COB-type flash memory and its soldered Nand connections

Nand Manufacturer brands according to their 3. According to the brands in this chart, Retry, power special reading settings such as mode must be performed. Power mod just Tsop is an enhanced option for reading chips (Özdemir, 2022).

Table 3. Nand Owned By Their Manufacturers ID

ID	Brand		ID	Brand
2C	Micron		98	Toshiba
45	Sandisk		AD	Hynix
89	Intel		EC	Samsung

Second Stage: The first step is to make sense of the raw data read through software (Flash Extractor, PC-3000 Flash, Rusolut) capable of data recovery from Nand devices. As an example, the first stage has been carried out and the imitation of the controller of a CF card with 4 Tsop48 data chips in Figure 8 with Flash Extractor software and the tree structure created after the imitation is presented.

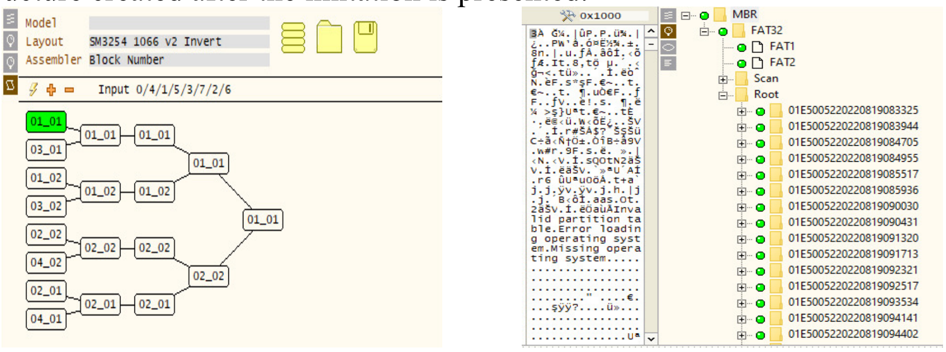


Figure 8. Imitation of the controller with Flash Extractor software and the tree structure formed as a result of this imitation

If hardware problems are to be exemplified;

- The legs of the data chip:

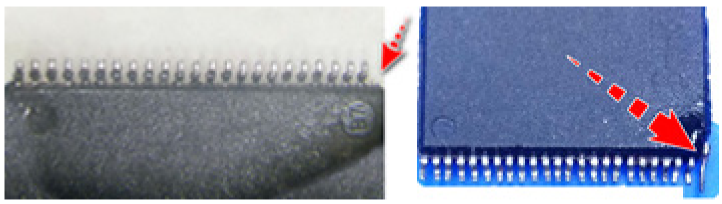


Figure 9. Example of tsop48 data chip legs breaking

  - It must be repaired using the jumper.

* Loss of USB busses:

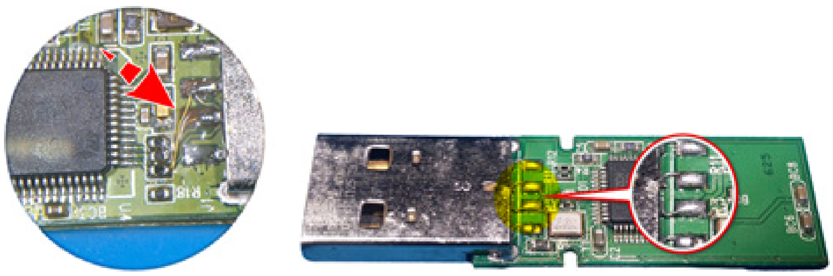


Figure 10. USB bus break

  * It must be repaired using the jumper.

* Breaking the USB PCB:

Figure 11. Breaking the USB PCB

  - Soldered to a solid USB socket.

* Breakdown of electronic circuit element:

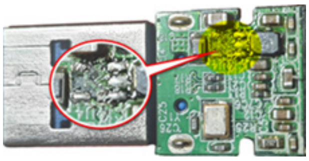


Figure 12. Fragmentation/loss of circuit element

- Donor or it should be completed by finding the diagram and determining the circuit element.

* USB socket was broken:

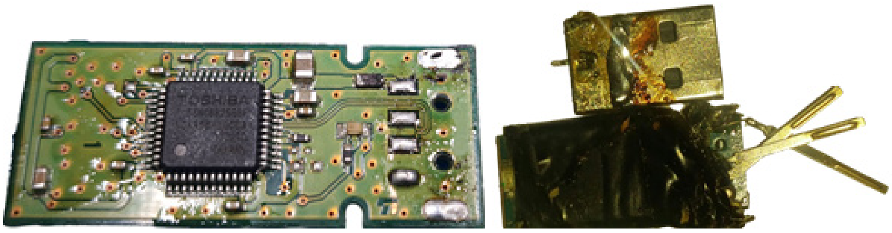


Figure 13. USB socket broken

- Soldered to a solid USB socket.

* USB socket pins not touching:

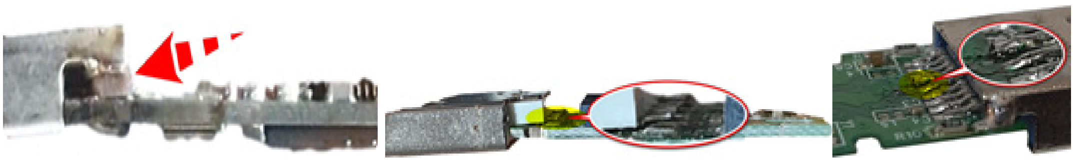


Figure 14. USB socket pins not touching

- Pins must be soldered.

* Deterioration of fuse resistor or capacitors:

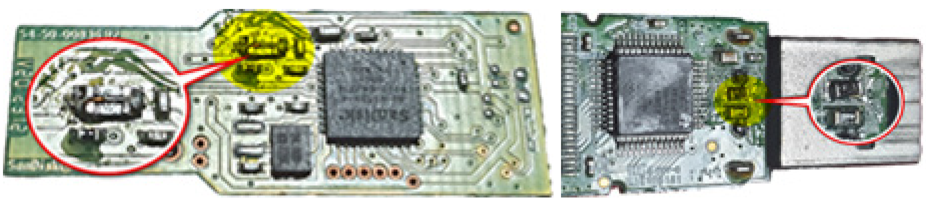


Figure 15. Bad fuse resistor

- The resistor must be changed or a jumper must be used. Capacitors must be replaced. Chip On or Chip If Repair Failed Off should be applied.

* Power (5V) not reaching memory or internal or surface bus being broken path:

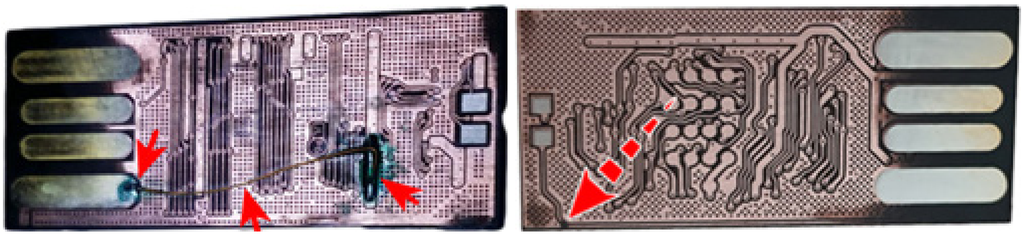


Figure 16. Internal and surface path break in USB flash memory

- Provided with the jumper or the broken path should be repaired.

- If the surface paths of the memory card are broken:

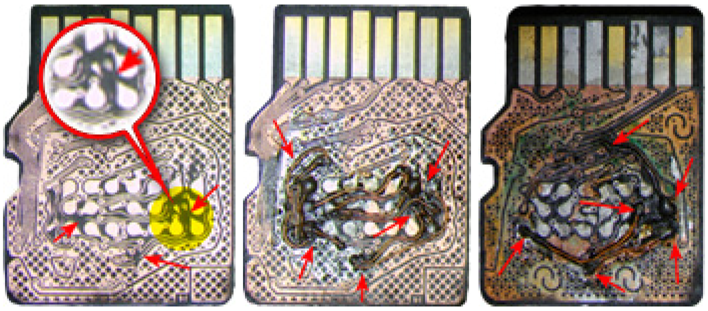


Figure 17. The paths of the memory card are broken

* Paths with jumpers must be repaired. Or it must be accessed from the Nand interface.

* Broken or broken memory card busses:

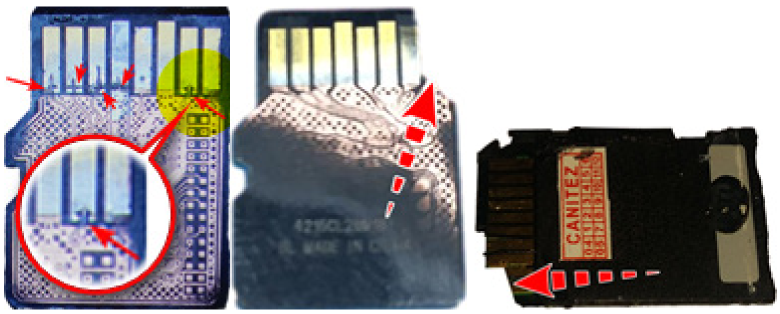


Figure 18. Broken or broken memory card busses

- Broken connections should be connected with a wire. Or by following the bus, the solder should be applied to the card reader by using the jumper from the last visible point.

* SSD data interface degradation and diode degradation:

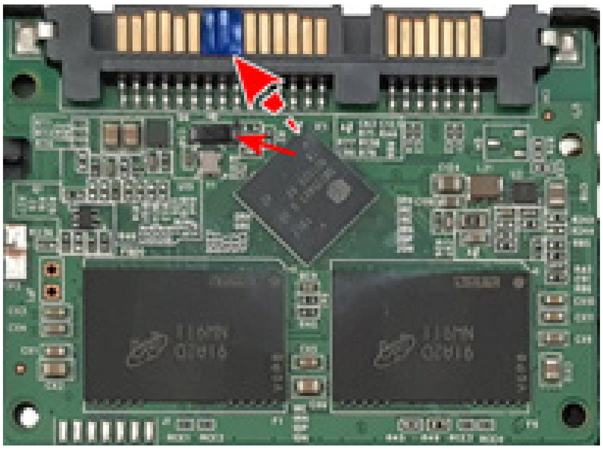


Figure 19. SSD data interface corruption

- The data interface should be repaired and the diode replaced or removed.

* Breakage of the circuit element on the SSD PCB:

Figure 20. Example of fragment breaking from SSD

- Soldered by detecting the circuit element by finding the donor or by finding the schematic.

- M.2 SSDs in the wrong socket:

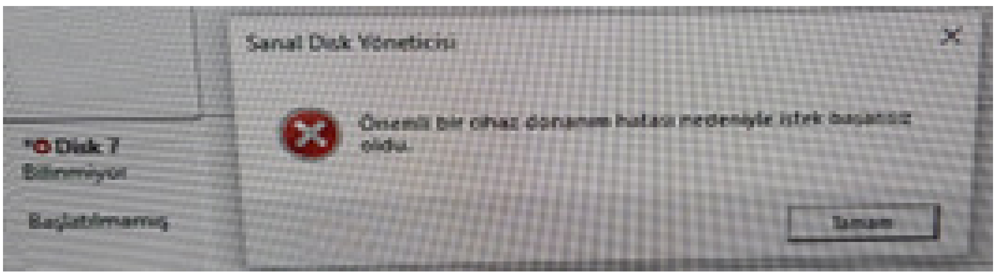


Figure 21. Installing M.2 SSDs in the wrong socket

- Socket details Must be plugged into a shape-matched socket or reader

* Deterioration of circuit elements on the memory card PCB:

Figure 22. Deterioration of circuit elements on the memory card

- The donor must be found and the circuit element must be replaced.

* USB socket not working:

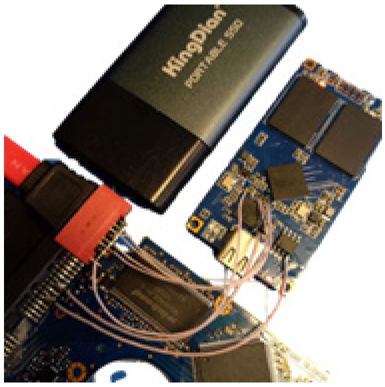


Figure 23. Example of SSD with USB socket not working (Total-Recovery, 2022)

- The data must be accessed by performing the conversion to the Sata interface.

* Failure of busses after liquid contact:

Figure 24. Bus not working after liquid contact with (ITHOPE s.r.o., 2021)

- M.2 disk that cannot be accessed from its default interface must be converted to SATA.

* Physical environmental conditions of the environment where the data logger is located:

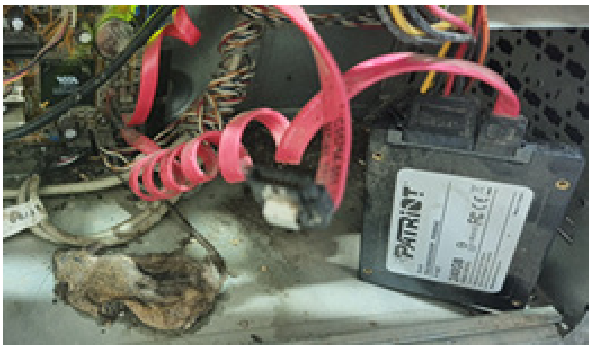


Figure 25. Example of a mouse biting the power cord

- The power cord must be replaced. Environmental conditions should be considered. From time to time, physical environments where data recording media are located should be checked.

* Breaking the data paths of mobile devices, breaking their sockets, breaking the socket pins, clogged the socket:

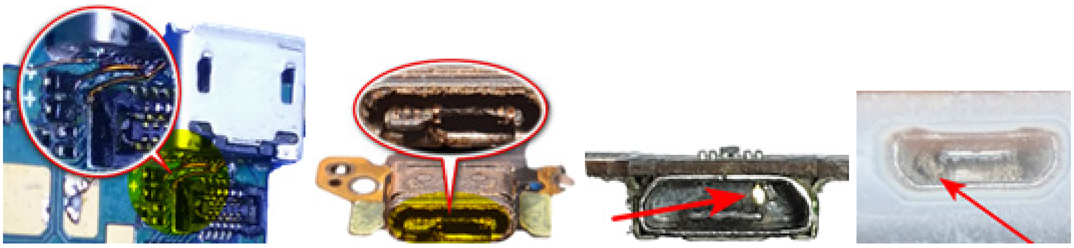


Figure 26. Socket failures in mobile devices

- Busses must be repaired with a jumper, breaks must be replaced with new sockets, and their blockages must be removed.

* Failure of circuit elements in the bus of mobile devices:

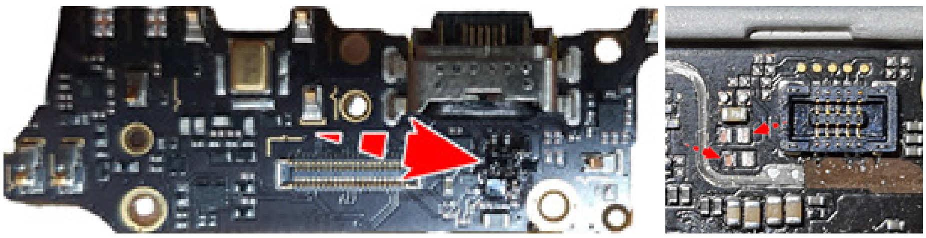


Figure 27. Malfunction of circuit elements in mobile devices

* The donor or the relevant part should be replaced by looking at its diagram.

- Battery deterioration of mobile devices:

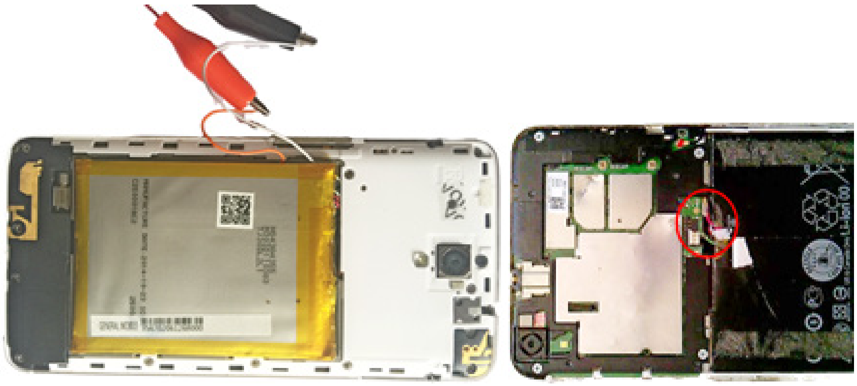


Figure 28. Battery deterioration of mobile devices

- External power must be supplied or a new battery must be installed.

* Cracked mobile screens:

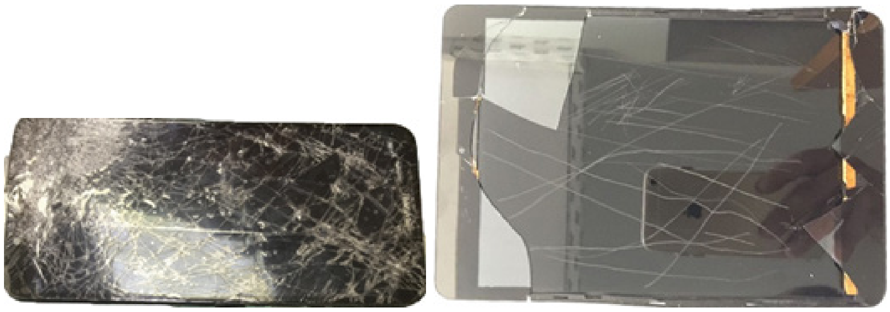


Figure 29. Cracked mobile screens

- The screen must be replaced.

- Malfunction of the socket of MP3 players:

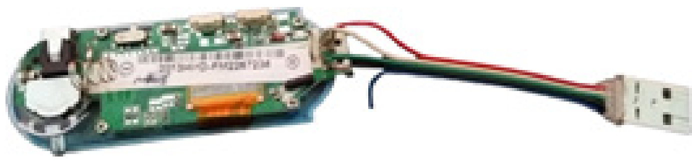
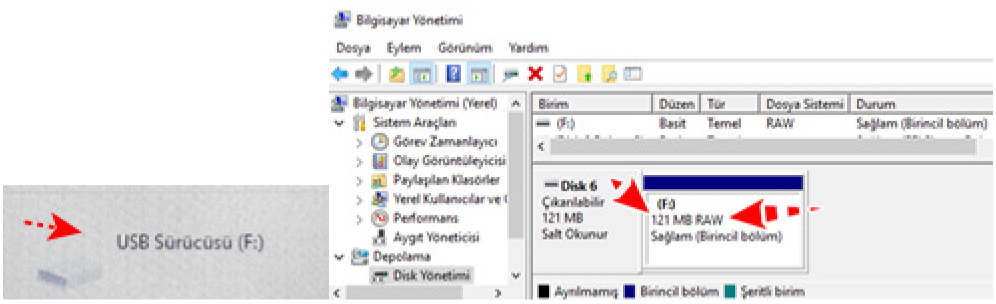


Figure 30. Malfunction of the socket of MP3 players

- The socket must be replaced.

* USB flash drives or memory cards appear below their normal capacity:



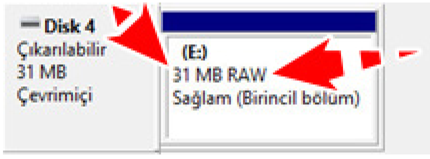


Figure 31. The appearance of the devices with hardware failures on the computer

- When there is a hardware failure in USB flash drives or memory cards or SSDs, as can be seen in Figure 31, they appear as an empty drive on my computer screen with a value below their normal capacity, such as “121MB” or “31MB”. It is necessary to access data on such devices using the Nand interface.

* USB flash drive and memory card with damaged controller:

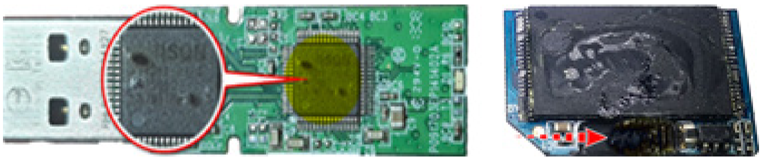


Figure 32. USB flash drive and memory card with damaged controller

- Requires access from the Nand interface.

* Flash drive oxidized by liquid contact:

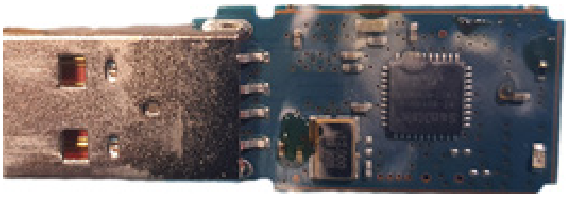


Figure 33. Oxidized USB flash drive

- Oxidation with isopropyl alcohol, if there is a faulty circuit element, it should be replaced. If the fault cannot be resolved, it must be accessed from the Nand interface.

* Burnt memory card and memory cards with broken PCB:

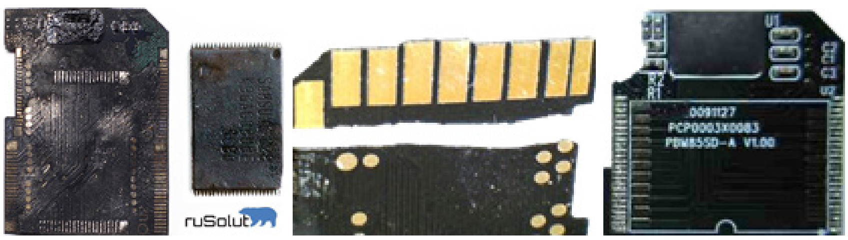


Figure 34. Burnt memory card (Rusolut Sp. z o.o., 2014) and memory cards with broken PCB

- Chip-Off should be tried

* Excessive scraping of memory card surface:

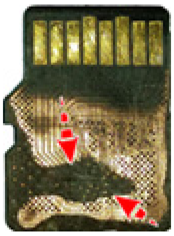


Figure 35. Excessive scraping of the memory card surface

- By being too scraped or physically damaged by the user to access TTNs. Access from the Nand interface should be attempted by taking X-rays and determining the exact locations of the TTNs.

* Memory card or data chip broken, cracked, or dented:

Figure 36. Examples of broken, cracked, and dented memory cards

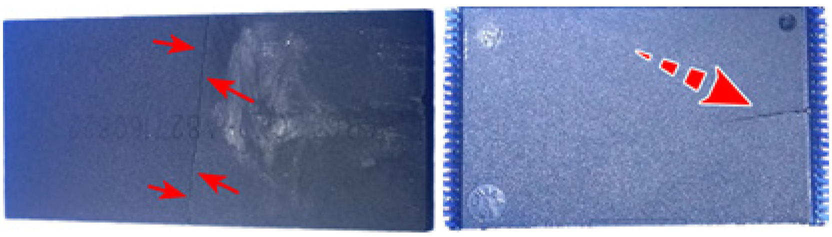


Figure 37. Broken UFD and Tsop48 chip examples

- An X-ray of the memory unit should be taken. If there is no damage to the Nand memory, try to get data from the Nand interface. Generally, data recovery is not possible in cases where the memory unit is broken or crushed.

* Combustion of parts on SSD PCB:

Figure 38. Burning parts on (DataSector, 2022)

- Data must be accessed from the Nand interface. As an alternative solution, a full replacement can be tried.

*   A mobile device whose physical failure cannot be fixed:

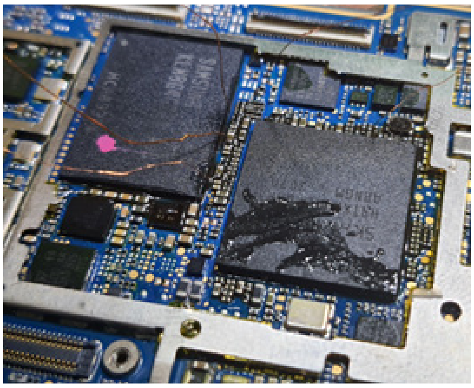


Figure 39. An example of ISP applied to a mobile device whose physical failure cannot be repaired

- Data must be accessed by making an ISP. If inaccessible, it must be accessed from the Nand interface.

*   Paths are broken or the protective coating has been removed on the PCB of the mobile device:

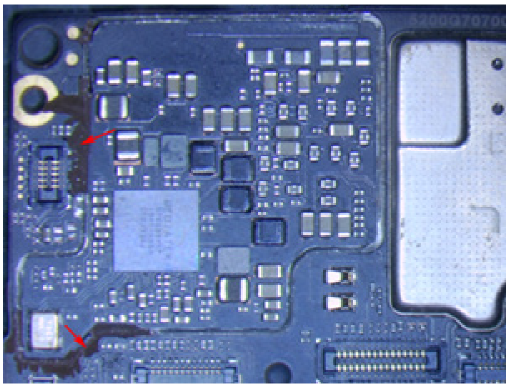


Figure 40. Remaining protective coating of mobile device PCB

- Jtag/ISP and then Chip-Off should be tried.

*   Burnt cell phone:

Figure 41. Burnt cell phone.

- Jtag/ISP and then Chip-Off should be tried.

* Mobile device PCB broken chip broken:

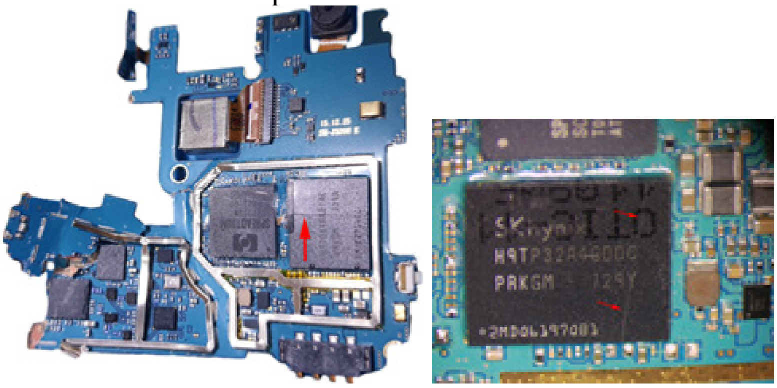


Figure 42. Mobile device PCB broken, chip broken

- An X-ray should be taken, if the Nand memory is not damaged, data should be obtained via Chip-Off or Nand interface.

* Breaking of data transmission legs of Usb Flash memory using a memory card as memory unit:

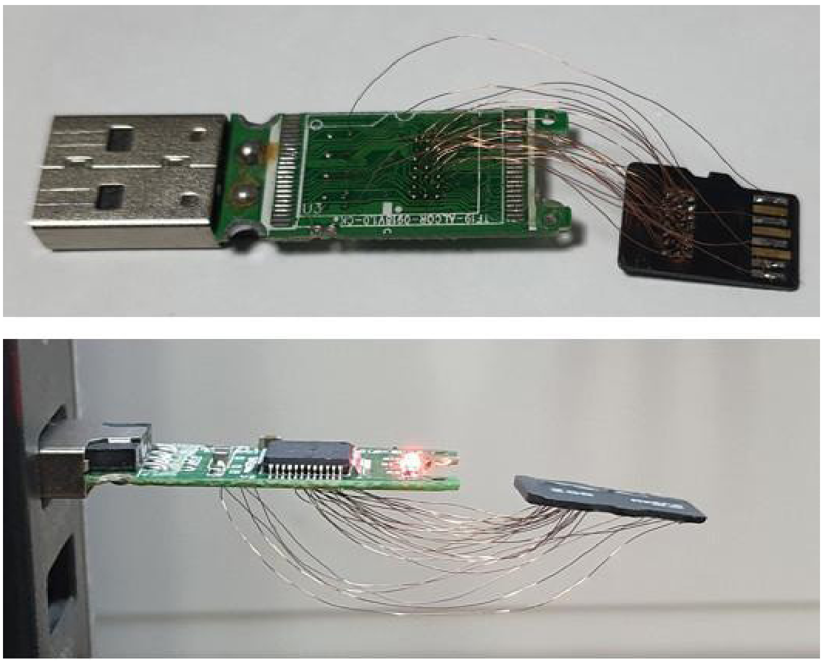


Figure 43. Breaking of data transmission paths

- If there are a few methods that can be used for data recovery processes;
  » Bus reconnection can be achieved by molding the TTNs of the memory card.

» The TTNs of the USB Flash memory and the TTNs of the memory card can be soldered with wires.
» The memory card can be accessed via the Nand interface (Chip-Off).

When TTNs are examined during data recovery operations on memory cards, species naming is shown in Table 4.

Table 4. Number of Encounters with TTN Types

TTN Type	Number of Encounters	Sample Photo
3x7	34	Figure 43– Number 1
4x6	52	Figure 43– Number 2
15x2	14	Figure 43– Number 3
Other	72	Figure 44
Total	172	

As can be seen from this chart, common TTN types are 3x7, 4x6, and 15x2. It was thought that it would be useful to make an adapter for the most common type of micro SD memory cards to be read quickly as a dump using the Nand interface. Previously, it has been explained how to perform data recovery over the Nand interface. To access the memory over the Nand interface, the signal equivalents of the TTNs to be used must be known. On Micro SD memory cards, detailed information of which is presented in Table 4, it was observed that 102 were hardware and 66 were software. When the methods described above were applied, it was observed that 97 data were recovered. When examined for data recovery processes for these memory cards, it was seen that some of the TTNs were the same (Figure 44) and some were completely different from each other (Figure 45). Pages belonging to data recovery companies (Arvika Data Recovery, 2013; GreyD Lab, 2015; Rflashdata.com, 2016; RecuperoDatos.com, 2020) When TTNs would be checked, it would be seen that there are hundreds of cards with different TTNs.

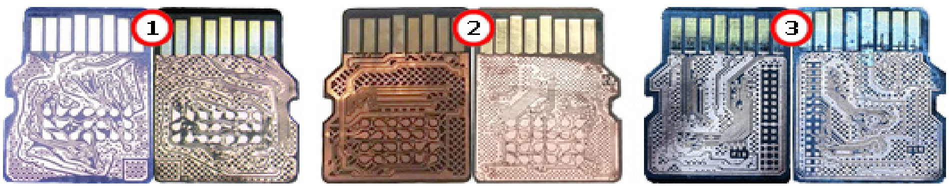


Figure 44. Same TTNs of different brands

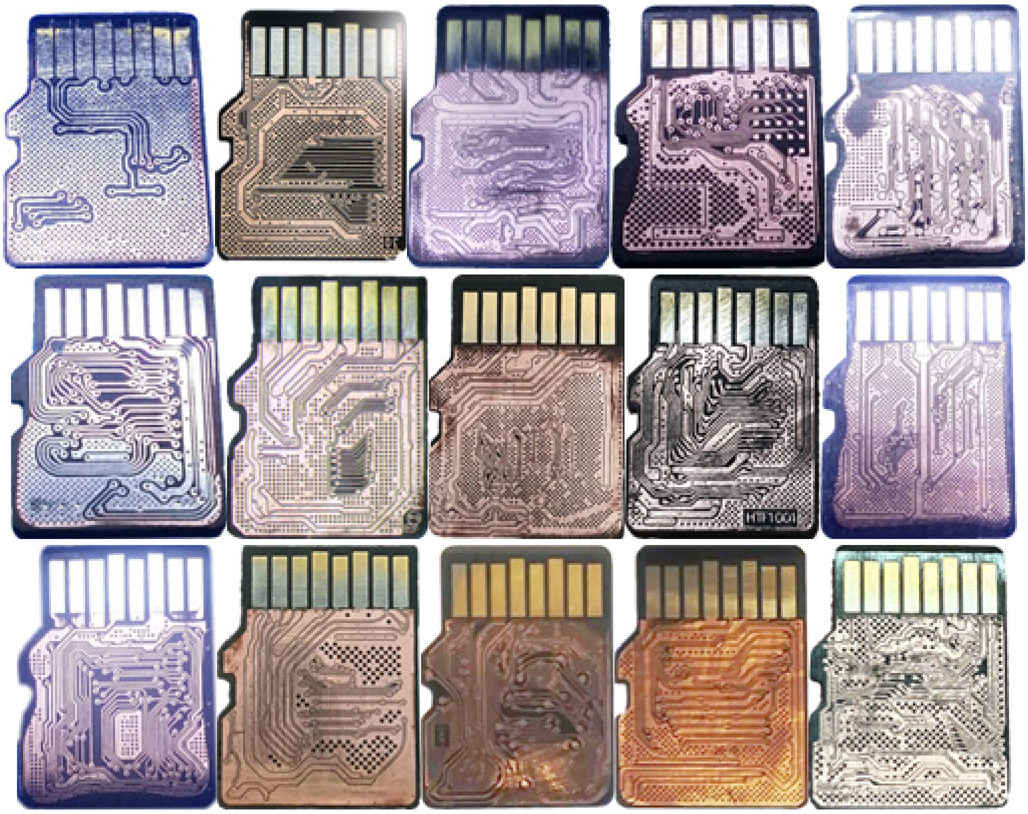


Figure 45. Different TTNs of different brands

The points called technological touch points cannot be seen directly from the outside and are used for tests. The TTN variety is developed according to the Nand memory type, circuit elements, and data buses used by the vendors during production, and the signal information used in TTNs cannot be disclosed due to copyrights. Memory cards vary in capacity, application performance, speed, and data paths. Multi-level cell technologies are also used to ensure this diversity. Due to this diversity, it is understood from the x-ray images that the internal data paths of the cards and the circuit elements used are also different.

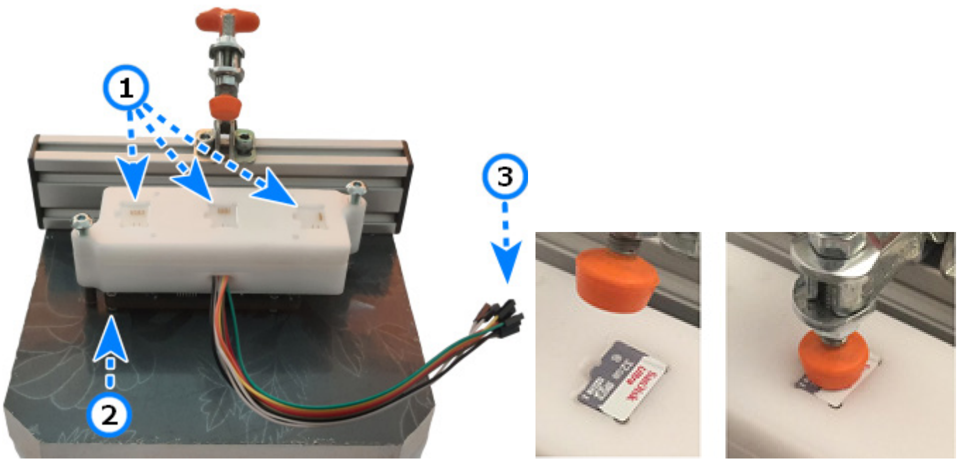


Figure 46. Completed adapter with all connections made

Thanks to this developed reader adapter;

* Common TTNs, without the need for soldering, the memory card Figure 46 reading

process is performed very quickly through the Nand Reader hardware belonging to the Soft-Center Company to the part indicated with number 2.

* Figure 46 is also performed for the devices of other companies that produce data recovery equipment.
* With this adapter produced, the loss of time spent on the soldering process will be eliminated.
* During soldering, problems such as incorrect soldering, solder, and wire breaks that repeat the soldering process and cause time loss will be eliminated.
* The loss of foreign currency spent on foreign companies that carry out similar transactions will be prevented.

Flash memories are mentioned as other and not frequently encountered in Table 4 Nand interface.

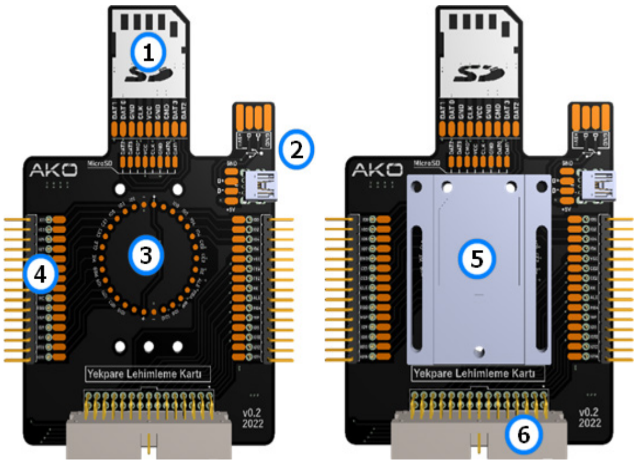


Figure 47. Soldering board developed for memory cards and USB sticks

Thanks to this card, all of the operations listed below can be performed without the need for any other hardware (Figure 47).

- MicroSD memory cards with a standard data interface broken or pin disconnected can be soldered to the area indicated by number 1 and data can be retrieved.
- USB flash drives with broken sockets can receive data via USB interface 2.
- 3 or 4 can be used crosswise/mixed for soldering memory cards, memory packaging type Cop/Blob or solid USB flash drives to TTNs. In this way, narrow soldering operations, which are formed by overlapping jumpers during the soldering of TTNs, are facilitated.
- To perform the soldering operations, it is necessary to fix the relevant memory to the soldering board, for this, bidirectional tapes are used. With the soldering holders we have developed, the formation of residues of the bidirectional tape on the memory card and soldering card is prevented.

The double-sided tape is prevented from losing its adhesive property due to the heat of the soldering iron.

- Multicom Sp. Zo.o. Since an output compatible with the output of the soldering card produced by the company is provided, this equipment can still be used if this equipment is already available.

In summary, the absence of mechanical parts in devices using flash memory, compared to hard disks, prevents the occurrence of mechanical problems such as head and motor failures, which are frequently encountered in hard disks. However, since devices using flash memory are constantly carried in daily life (mobile phones, USB flash memory, memory card, etc.), they are more likely to encounter physical effects. Therefore, physical failures occur more than software failures compared to other storage units.

These failures are;

- Fractures/cracks in the data chip,
- In addition, breakage/deterioration of components (such as screen, touchscreen) that enable the device to be used,
- can be listed as all kinds of reasons that prevent electrical conductivity, such as deterioration and disintegration of circuit elements in PCBs.

The closer the physical failures in data recording environments to the unit where the data is stored, the higher the sensitivity and difficulty level in the recovery process. For example;

1- A USB flash memory socket is broken

2- access Nand data chip of same USB flash memory

If a comparison is made, it will take 5 minutes for the problem to be resolved by replacing the socket, the fault in number 1, and to reach the data. However, it will only take ~5 minutes to remove the data chip to fix the fault in number 2. Afterward, the data chip will be read raw, and the read data will be tried to be made meaningful by imitating the controller. Here, however, data recovery will take days, maybe weeks.

Process Steps to Follow During Data Recovery

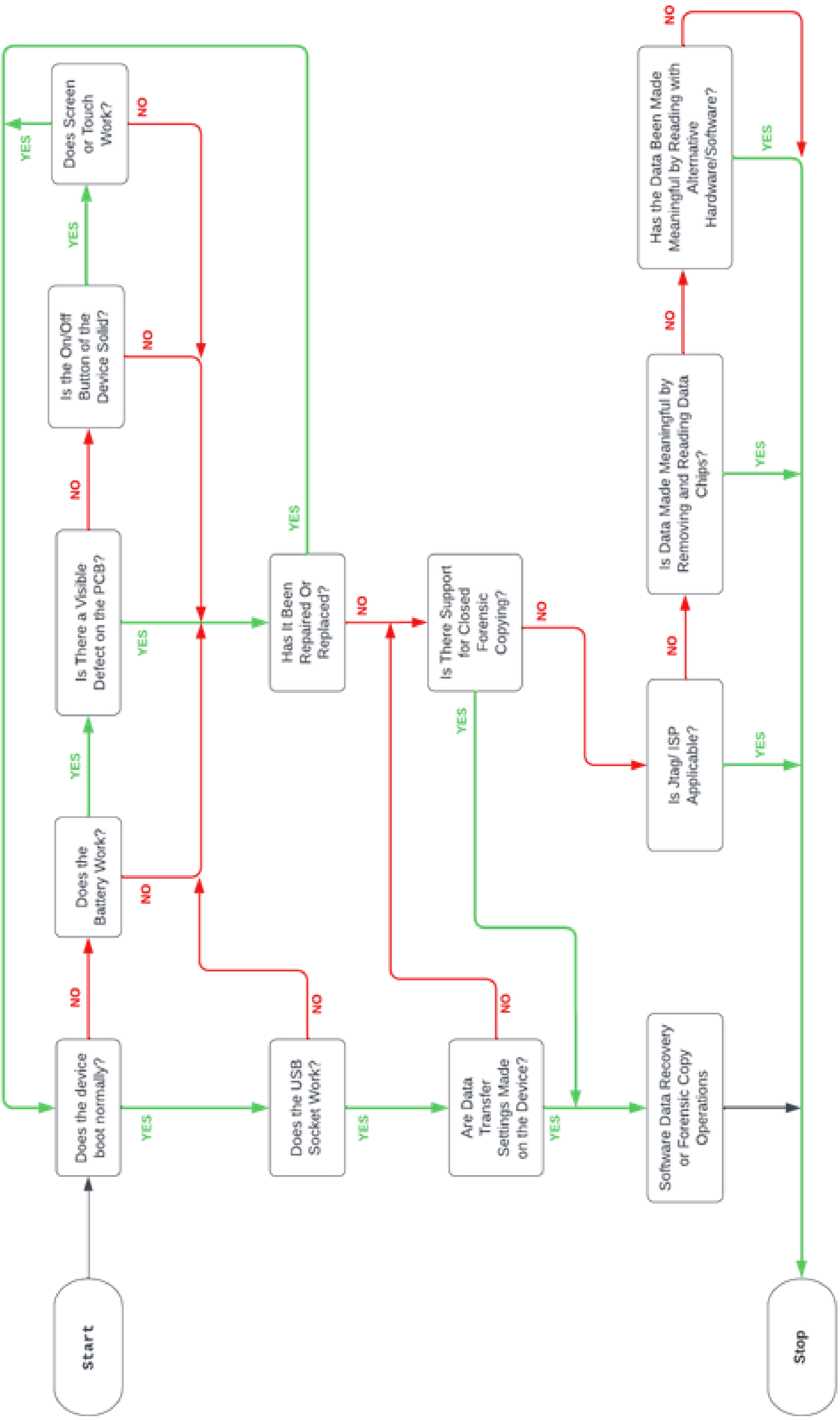


Figure 48. Data recovery steps from mobile devices

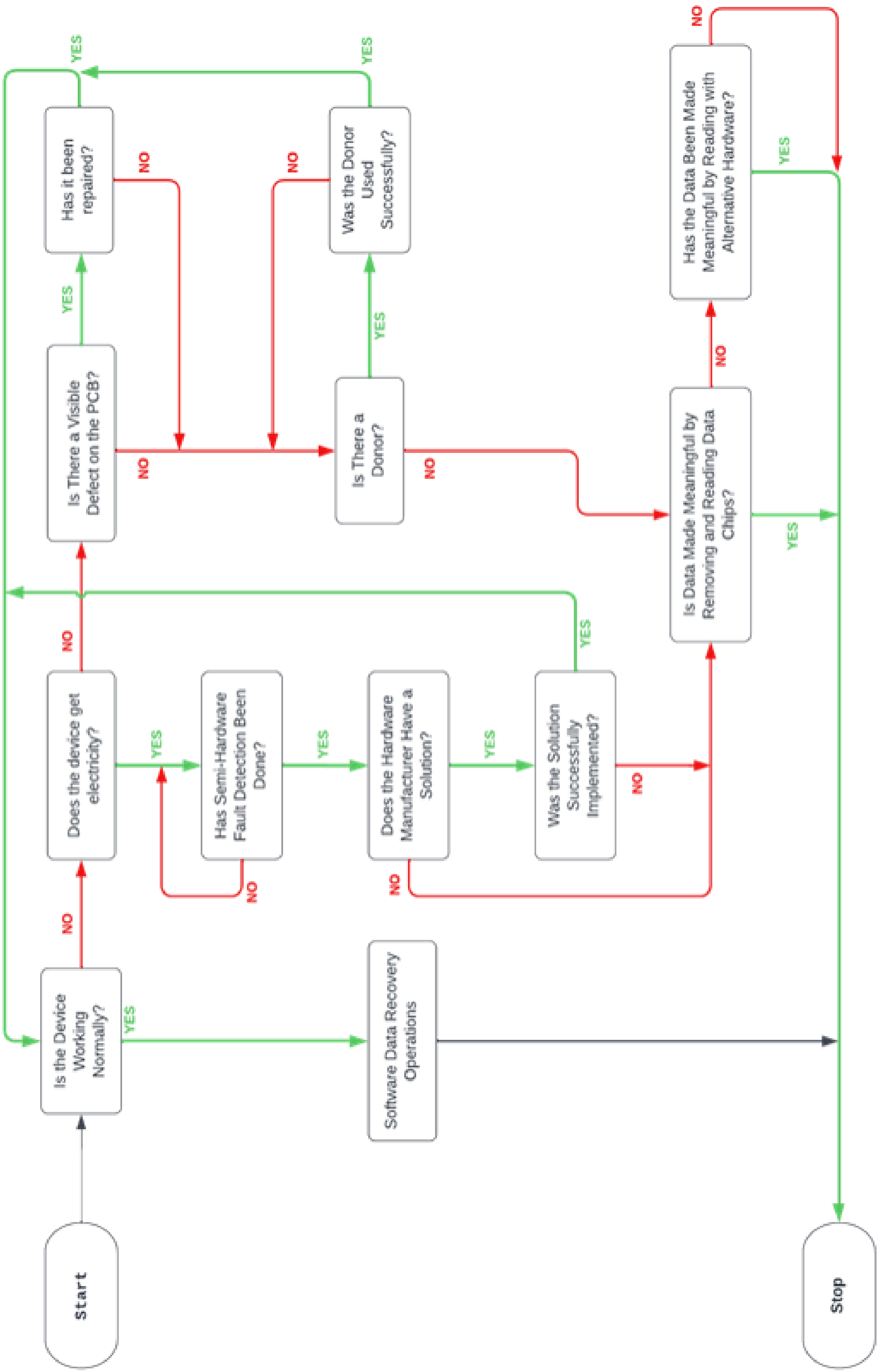


Figure 49. Data recovery steps from devices using flash memory

Precautions to be Taken to Avoid Data Loss

* To use a power supply for hardware protection of the device in sudden power cuts.
* To provide grounding to protect the device from static electricity.
* Using a case/bag to weaken the effects of physical impacts such as hitting or falling.
* Position data storage units with moving parts so that they cannot be moved as much as possible.
* Disconnecting from the computer using the “remove hardware” option to avoid corrupting files during data transfer to an external data storage or editing of data on it.
* Anti-virus etc. to prevent data encryption or corruption caused by a virus. using security software.
* To provide physical security to protect the data logger from forensic events such as theft.
* To use data storage units by taking appropriate measures to protect the data recording unit from natural disasters.
* Make at least two backups, one of which is cloud storage, to prevent data loss that may occur consciously or unconsciously.

References

Arvika Data Recovery, (2013), Monolith Database http://odzyskiwanie-danych.com.pl/monolith-database.html.

DataSector, (2022), Intel SSD, https://www.instagram.com/p/Ce0XJk6DVvH/?hl=tr.

Doğanay, H. A., 2019, Evaluation Of The Legal Process And The Term Chain Of Custody With Actual Difficulties Against Investigation Of New Generation Mobile Devices In The Context Of The Importance Of Mobile Device Forensics, Master Thesis, Ankara University, Department of Interdisciplinary Forensic Sciences.

Gin, J., (2020), Reverse-engineering and analysis of SanDisk High Endurance microSDXC card, https://ripitapart.com/2020/07/16/reverse-engineering-and-analysis-of-sandisk-high-endurance-microsdxc-card/.

GreyD Lab, (2015), Монолиты, http://greydcode.com/monolith.html.

ITHOPE s.r.o., (2021), SanDisk SSD U100, www.instagram.com/p/CLPYC6gj8t9/?hl=tr:

Johnson, R. and Christie, S., (2009). JTAG 101 IEEE 1149.x and Software Debug.

ONFI Workgroup, (2021), Open NAND Flash Interface Specification 5.0. Schmidt, K.-U.

Özdemir, A. and Gülcü, Ş., (2021), Causes of Digital Data Loss and Data Recovery Methods, 2nd International 5 Ocak Congress on Applied Sciences.

Ozdemir, A., (2022), Investigation of data recovery methods and development of auxiliary equipment for data recovery, Necmettin Erbakan University, The graduate school of natural and applied science (Master thesis).

RecuperoDatos.com, (2020), Monolith Memories Database (MicroSD, SD, Pendrive), https://datarecoveryhdd.com/monolith-database-microsd-sd-pendrive-cf.

Rflashdata.com, (2016), Monolith Flash Database, https://www.rflashdata.com/monolith-flash-database/.

Rusolut Sp. z o.o., (2014), BurntSD.jpg, https://rusolut.com/wp-content/uploads/2014/03/BurntSD.jpg.

Rusolut Sp. z o.o., (2022), Forensic analysis of flash memory using X-RAY and Logic Analyzer. PTXXI 2022 IV Konferencja Naukowa Przestępczość Teleinformatyczna XXI Wieku.

Total-Recovery, (2022), https://www.instagram.com/p/CaYlOyRjyW0/?hl.

About the Authors

Abdulkadir ÖZDEMİR is a police officer who has been working in Cyber Crime Branch Offices in different provinces since 2012. He has a master's degree in Computer Engineering from Necmettin Erbakan University. The main areas of work are digital data recovery, installation, and management of digital evidence inspection systems, malware analysis, and digital data inspection. He continues to work on various hardware and software developments to be used in digital evidence examination and data recovery processes.

E-mail: by.akozdemir@gmail.com, **ORCID**: 0000-0002-6992-5170.

Şaban GÜLCÜ graduated from the Software Engineering Department of Vienna University of Technology with B.S. and M.S. degrees in 2005 and 2006, respectively. He received a Ph.D. degree in Computer Engineering Department from Selcuk University in 2017. He is an Assistant Professor at the Computer Engineering Department at Necmettin Erbakan University. His main research interests are neural networks, optimization algorithms, and software engineering.

E-mail: sgulcu@erbakan.edu.tr, **ORCID**: 0000-0001-7714-8861.

Similarity Index

The similarity index obtained from the plagiarism software for this book chapter is 3%.

To Cite This Chapter:

Özdemir, A. and Gülcü, Ş. (2022). Scientific Data Recovery Methods and Solutions in Devices using Flash Memory. In S. Kocer & O. Dundar (Eds.), *Current Studies in Basic Sciences, Engineering and Technology* (pp. 258-286). ISRES Publishing.

Current Studies in Basic Sciences, Engineering and Technology
is published from the selected papers invited by the editors.

This edition includes 16 sections from the education, healt, industry etc of used in today's technology. All submissions are reviewed by at least two international reviewers.

The purpose of the book is to provide the readers with the opportunity of a scholarly refereed publication in the field of techonology, artificial intelligence and augmented reality.

Current Studies in Basic Sciences, Engineering and Technology
is published by ISRES Publishing.

ISRES Publishing