
T.C.

ALTINBAS UNIVERSITY

Graduate School of Science and Engineering

Information Technologies

# AN EFFICIENT MODEL FOR HUMAN ACTIVITY RECOGNITION USING CONVOLUTIONAL NEURAL NETWORK (CNN)

Hussein Riyadh Hussein AL-GBURI

Master of Science

Supervisor

Asst. Prof. Dr. Abdullahi Abdu IBRAHIM

Istanbul, 2021

# AN EFFICIENT MODEL FOR HUMAN ACTIVITY RECOGNITION USING CONVOLUTIONAL NEURAL NETWORK (CNN)

by

Hussein Riyadh Hussein AL-GBURI

Information Technologies

Submitted to the Graduate School of Science and Engineering

in partial fulfillment of the requirements for the degree of

Master of Science

ALTINBAS UNIVERSITY

2021

The thesis titled "AN EFFICIENT MODEL FOR HUMAN ACTIVITY RECOGNITION USING CONVOLUTIONAL NEURAL NETWORK (CNN)" prepared and presented by "Hussiın Riyadh Hussein AL-GBURI " was accepted as a Master of Science Thesis in Information Technologies.

Asst. Prof. Dr. Abdullahi Abdu IBRAHIM

Supervisor

Thesis Defense Jury Members:

| | | |
|---|---|---|
| Asst. Prof. Dr. Abdullahi Abdu IBRAHIM | School of Engineering and Natural Sciences, Altinbas University | _____ |
| Asst. Prof. Dr. Sefer KURNAZ | School of Engineering and Natural Sciences, Altinbas University | _____ |
| Assoc. Prof. Dr. Mohammed ABUBAKAR | Faculty of School Economics, Administrative and Social Sciences, Antalya Science University | _____ |

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

Approval Date of Institute of Graduate Studies: ____/____/___

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Hussein Riyadh Hussein AL-GBURI

# DEDICATION

First, I would like to thank Allah Almighty for the power of the mind, health, strength, guidance, knowledge, and skills to complete this study.

This thesis is wholeheartedly dedicated to my parents. There are no words to describe what you mean to me; there is nothing that I can repay for what you have done to me. I will continue to do my best to achieve your expectations.

Lastly, I dedicated this to the family, relatives, and friends who have been encouraging me during this study.

## ACKNOWLEDGEMENTS

I would like to thank my supervisor Asst. Prof. Dr. Abdullahi Abdu IBRAHIM for all support during my study. Its great pleasure to express my deepest gratitude to my friends who have shared with me best moments during my study for the Master degree.

# ABSTRACT

## AN EFFICIENT MODEL FOR HUMAN ACTIVITY RECOGNITION USING CONVOLUTIONAL NEURAL NETWORK (CNN)

AL-GBURI, Hussein Riyadh Hussein,

M.Sc., Information Technologies, Altınbaş University,

Supervisor: Asst. Prof. Dr. Abdullahi Abdu IBRAHIM

Date: 06/2021

Pages: 51

Information can be obtained from individuals through the systems of classifying and recognizing human activities at any time. These systems are used in different areas such as the detection of diseases, improvement of physical therapy stages, development of smart home projects. In this study, the data obtained from accelerometer and gyroscope sensors on smart phones are used. Most of the studies in the literature are unable to analyse higher-level features and their relationships based on machine learning and deep learning techniques. Convolutional Neural Network (CNN) model is a very suitable deep learning approach due to its ability to obtain high level and sensitive features. The deep learning-based approach that includes this background has been used in the classification of various human activities in the experiments in our study. In the experiments, the classification performance accuracy rate was measured by giving different input parameters, layer and network units to the relevant network models. As a result, it has been shown that six different classes are classified with high accuracy, achieving a classification performance of approximately 97.98%.

**Keywords:** CNN, human activity recognition, ELM, KNN, DT.

# ÖZET

## EVRİMSEL SİNİR AĞI (CNN) KULLANARAK İNSAN FAALİYETLERİNİN TANIMA İÇİN ETKİLİ BİR MODEL

AL-GBURI , Hussein Riyadh Hussein,

Yüksek Lisans, Bilgi Teknolojileri, Altınbaş Üniversitesi,

Danışman: Asst. Prof. Dr. Abdullahi Abdu IBRAHIM

Tarih: 06/2021

Sayfa Sayısı: 51

Bilgi, herhangi bir zamanda insan faaliyetlerini sınıflandırma ve tanıma sistemleri aracılığıyla bireylerden elde edilebilir. Bu sistemler hastalıkların tespiti, fizik tedavi aşamalarının iyileştirilmesi, akıllı ev projelerinin geliştirilmesi gibi farklı alanlarda kullanılmaktadır. Bu çalışmada akıllı telefonlarda ivmeölçer ve jiroskop sensörlerinden elde edilen veriler kullanılmıştır. Literatürdeki çalışmaların çoğu, üst düzey özellikleri ve bunların makine öğrenimi ve derin öğrenme tekniklerine dayalı ilişkilerini analiz edememektedir. Evrişimli Sinir Ağı (CNN) modeli, üst düzey ve hassas özellikler elde etme yeteneği nedeniyle çok uygun bir derin öğrenme yaklaşımıdır. Bu arka planı içeren derin öğrenmeye dayalı yaklaşım, çalışmamızdaki deneylerde çeşitli insan etkinliklerinin sınıflandırılmasında kullanılmıştır. Deneylerde, ilgili ağ modellerine farklı girdi parametreleri, katman ve ağ birimleri verilerek sınıflandırma performans doğruluk oranı ölçülmüştür. Sonuç olarak, altı farklı sınıfın yüksek doğrulukta sınıflandırıldığı ve yaklaşık% 97,98'lik bir sınıflandırma performansına ulaşıldığı gösterilmiştir..

**Anahtar Kelimeler:** CNN, insan aktivitesi tanıma, ELM, KNN, DT.

# TABLE OF CONTENTS

**Page**

# LIST OF TABLES

**Page**

# LIST OF FIGURES

**Page**

# LIST OF ABBREVIATIONS

| | | |
|---|---|---|
| SVM | : | Support vector machine |
| NN | : | Neural Network |
| PSO | : | Particle Swarm Optimization Algorithm |
| CNN | : | Convolutional Neural Network |
| KNN | : | k-nearest neighbours algorithm |

# 1. INTRODUCTION

Today, the rapid development of technology and as a result of this development, mobile device. The study conducted by the German statistics company Statistica supports this situation. The study, it was stated that the average number of mobile phones worldwide is approximately It is expected that this number will gradually increase and reach a total of 2.87 mobile phones by 2020. A similar situation is seen in wearable devices, except for mobile phones. According to the Visual Network Index Global Mobile Data Traffic Forecast Report, wearable devices in the world will reach from 109 million to 578 million in 2019, and 97% of mobile data traffic will be generated by mobile devices [2].

The rapid development of technology and the widespread use of mobile devices naturally mean that these devices adapt to this development in terms of hardware. Smartphones are equipped with better, more precise and diverse hardware every day. Today, in an average mobile phone, GPS, high resolution cameras and sensitive sensors that serve different purposes (such as accelerometer, gyroscope, light sensors) have become standard. The gradual decrease in the prices of mobile devices, the increase in their hardware efficiency, and the advantages of portability have been effective in these developments [3].

The hardware developments, lead applications in smart phones have become available for solving different and specific problems. One of these studies is the detection of human activities using smart phones. The definition of human activities is a field of study that is used in a wide variety of fields and is still difficult to apply. It is used in various fields such as analyzing the daily movements of people in the field of health, fall detection applications for elderly people, smart homes, and ensuring home security [4].

There are basically two approaches regarding the definition of human activity. The first of these approaches is the detection of the movements of the person or persons in the environment by image processing methods by means of various cameras placed in different positions in an environment [5]. However, these methods cause the cameras to be placed in a limited area, so that an application to be developed can be developed in a limited space. Although methods based on image processing are used in large areas, they need an infrastructure support. In addition, system setup costs are higher than the data mining methods to be mentioned [6].

Another method used in the field of human activity and becoming increasingly common is the use of data mining methods in the solution of the related problem. This method is generally based on the principle of solving this problem by processing the data obtained from sensors placed in different parts of the body in accordance with data mining methods while performing different types of movements and converting the related problem into a classification problem by considering human activities as different classes. It is tried to detect different types of motion by using different classification algorithms over the data set created over the raw data obtained from mobile sensors. Different algorithms such as Decision Tree (DT), Naive Bayes Classifier (NB), Multilayer Perceptron (MLP), Support Vector Machines (SVM) are frequently used in such studies.

Human activity is subdivided into gestures, atomic actions, person-to-object or person-to-person interactions, group actions, behavior and events depending on their complexity [7]. Also, a bottom-up approach then many researchers discover human activity. The main stages of these systems are: feature extraction, training and classification of actions, detection and segmentation of actions [8]. Although there has been a lot of research on this topic, it remains an active area of research with complex research topics such as the analysis of interventions to support living in smart homes [9].

Recognized activities can include daily activities such as walking and sitting, as well as sports activities such as jumping and running on a treadmill. More recent work on the automatic detection of daily activity focuses on machine learning algorithms that are based on the simultaneous input of different sensor modes, e.g. B. visual, inertial, acoustic, force, pressure, tension, physiological and kinetics, among others [10,11] ... Gathering information on user actions under comfortable living conditions in smart homes and recognizing abnormal behavior in order to help the elderly or people with special needs are complex research tasks.

Motion sensors make it possible to determine position and orientation information based on the measurement of physical quantities that are directly related to the movement of the body part in which they are positioned. HS are used to detect and track body movement in an almost unlimited field of study activity recognition,

Until today, it has only been used to track the movement of man-made vehicles, including aircraft, ships, submarines, cars, and more recently, wheeled and legged robots. But recent developments

in Microelectromechanical systems (MEMS); It has led to the development of a new generation of motion sensors.

Analysing movements and actions has a long history and interacts with various disciplines such as psychology, biology, and computer science. In recent years, sensor-based human HT has played an important role in the field of computer science due to its wide use in daily life. Evaluation of movement and posture and general for the kinematic analysis of behavior, progress in sensor technology and advanced methods in signal analysis are greatly utilized. Understanding the user's needs with the development of wearable devices with embedded sensors and it has become possible to collect data that can be analyzed to provide personalized services. Being able to recognize a person's situation has come to provide us with valuable information that can be used as input for other systems.

The general compounds of human activity recognition presented in the Figure 1.1.

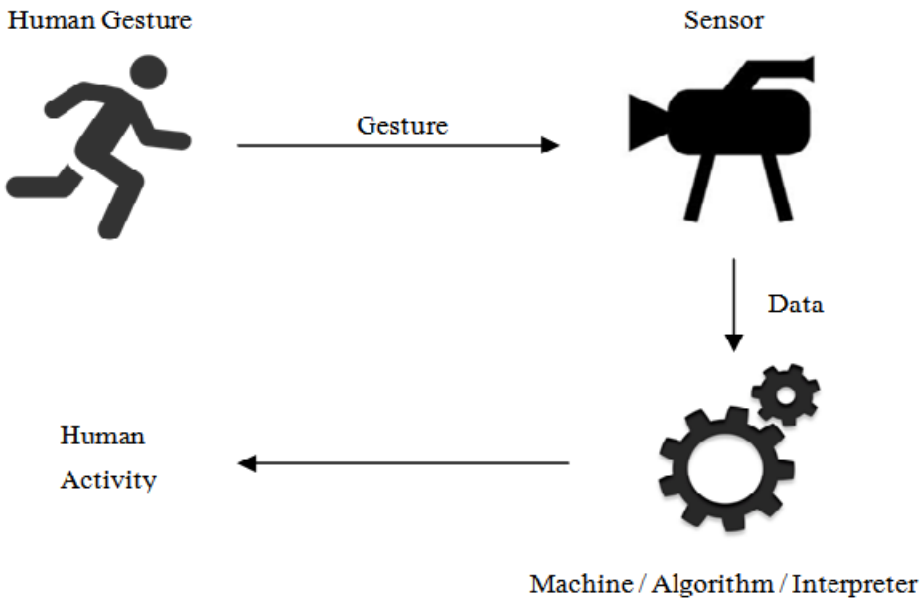


**Figure 1.1:** Human Activity Recognition [12]

**Sensors**: Sensors act as a bridge that connects the physical environment and industrial electrical / electronic devices. These devices have a wide range of uses such as control, protection and monitoring in the industrial process process.

One can mention hundreds of types of sensors produced today. Incredibly fast developments in microelectronics technology allow a new invention or a new application type to be developed every day.

Sensor or sensor is the name given to the sense organs of automatic control systems. Just as humans perceive what is happening in their environment with their sense organs, machines also perceive values such as temperature, pressure, speed, and the like, through their sensors. For example, a temperature sensor creates an electric potential difference (voltage) between its legs depending on the changing ambient temperature. When this information is transferred to a microcontroller, a closed loop temperature control unit is obtained.

The path designed to solve a problem or reach a determined goal, the steps of the process followed are called an algorithm. Algorithms are a set of operations that have a clearly defined beginning and an end. The solution ways and sequences to be processed to achieve the goal are determined and the algorithm follows this order to reach the most logical solution. The first algorithm is presented in Al-Khwarezmi's "Hisab-el Algebra and Al Muqabala" and the word algorithm comes from the name of Al-Khwarezmi. The term algorithm often appears in mathematics and computer science.
In this study, our goal is to developed new techniques based AI to developed new and effective algorithm to deal with this issue.

## 1.1 QUESTIONS OF THE STUDY

After analysing the previous studies and the field of the human activity recognitions, several questions are occurred and in this study we will answered these questions to enrichment the research field:

What is the human activity recognition? We tried to answer this question by presented several studies and techniques presented in this filed and also discuss the development of this filed after deploying the smart phones.

What is the best AI techniques applied to develop human activity recognition applications? several techniques presented in this field such as deep learning, machine learning, and data mining.

How we can developed effective AI based application for presenting remarkable results when compared with previous studies. The aim is to developed AI model based feature extraction techniques which lead to decrease the computation time. this lead also to reduce the hardware requirement for the application

## 1.2 THESIS STRUCTURE

**Chapter 1:** Introduction and brief Access to the problem field.

**Chapter 2:** related works and literature review

**Chapter 3:** material and methods

**Chapter 4:** results and discussions

**Chapter 5:** Conclusion and future works

# 2. OVERVIEW

## 2.1 LITERATURE REVIEW

In the literature, there are studies that consider the human activity definition problem as a classification problem and consider it as a solution to this classification problem. On the basis of these studies, it is based on the detection of these class tags with the attributes they obtain from sensor data, taking each human activity as different class tags.

Bao and Intille collected data from a total of 20 different subjects by placing 5 uniaxial accelerometers on different parts of the human body, such as the hips, wrists and ankles [13]. They trained 20 different human activities they collected from 20 different subjects with example-based learning, decision trees, Naive Bayes classifiers using Fourier Transform attributes and achieved a success close to 80%. Similarly, Banos et al. Tried to detect walking, running, standing and sitting movements by using accelerometers placed on different parts of the body [14].

Using a triaxial accelerometer, Niskham and Nikhil attempted to detect walking, running, standing, sitting, climbing stairs, climbing stairs, brushing teeth, and sweeping [15]. They concluded that majority voting classifiers were more successful than other classifiers. They also stated that it was impossible to distinguish between tooth brushing and stair climbing activities.

Kwapsiz et al. Detected an activity using three-axis accelerometer data obtained from an android device [16, 17]. They derived different properties such as average acceleration, standard deviation, and mean absolute deviation. They used the data set they obtained for the training of 3 different classifiers (Logistic Regression, Decision Tree and MLP). They stated that the models of the Decision Tree had higher performance than other models. In their results, it was observed that the performance was low due to the similarity between the stairs and climbing stairs.

In [18], 5 different human activities have been tried to be determined by using 13 different wave attributes in time and frequency space. The data set was collected from 7 different subjects, 2 of whom were women, aged 27-35. After obtaining 13 different time and frequency attributes from 69400 sensor data, they trained their systems with the C4.5 algorithm and Artificial Neural Networks. According to the results they obtained, they achieved a 94.13% success rate with the C4.5 algorithm.

Wu et al. Used gyroscope data together with accelerometer data for their human activity system [19]. According to their results, when used with accelerometer data, gyroscope data contributes to the establishment of more successful models. Similarly, in the study conducted in [20], a detection system for 7 different activities is presented by using accelerometer and gyroscope sensors together with GPS data. Like most studies in the literature, the Decision Tree algorithm has achieved the best result in the human activity detection system.

Lee and Cho [21] identified sitting, walking, staircase movements and running movements, which they first described as action, in their two-stage HMM structure. With a second HMM model based on the actions they have achieved, they tried to detect different human movements, which they call shopping, traveling by bus and moving state, which are their main objectives.

Kim et al. Used the Hidden Markov Model Ensemble (HMME) method [22]. They applied feature extraction on 3-axis accelerometer data. They obtained Decision Profiles (DP) from this data by using HMM models they created for different activities. They measured the similarities of DPs they obtained with Decision Templates (DTs) created for each activity and determined the type of an action. They compared their method with different classifiers and found that it was the most successful model after MLP.

Ranao and Cho presented a 2-stage activity detection system in their studies [23]. Apart from the acceleration data, they also used gyroscope data in the system they developed. In the first stage, which they call rough classification, they tried to determine whether a movement is a stationary or a moving action. According to whether the type of movement is stationary or mobile, they determined the real class of the activity by using one of two different HMMs in the second stage.

Han et al. Performed human activity identification with a versatile HMM that allows backward or skipped transitions [24]. They also took into account the gravity-induced bias and tried to compensate for this situation in the data. They implemented their systems over the obtained tilt compansated data. As a result, they concluded that inclination correction techniques are a powerful model against the rotation of the sensor and their versatile HMM model is more successful than a left-to-right unidirectional HMM.

In their study in [25], Mannini and Sabatini achieved a very high success with the Continuous Time Hidden Markov Models for 7 different basic human activities, using the data set they used in their work in [26] of Bao and Intille.

## 2.2 AI TECHNIQUES

### 2.1.1 Machine learning

Machine Learning) is a category of algorithms that makes software programs more accurate at predicting results without explicitly programming. The mainstay of machine learning is to create algorithms that can take input data and use statistical analysis to predict an output while updating output as new data emerges [27].

The processes involved in machine learning are similar to data mining and predictive modeling. Both require searching data to search for patterns and adjust program actions accordingly. Many people are familiar with machine learning from shopping online and posting ads for purchases. This is because recommendation engines use machine learning to personalize online ad delivery in near real time. Beyond personalized marketing, other common machine learning use cases include fraud detection, spam filtering, network security threat detection, predictive maintenance, and building news feeds [28].

### 2.1.2 KNN

KNN algorithm is one of the easy-to-implement supervised learning algorithms. Although it is used in the solution of both classification and regression problems, it is mostly used in the solution of classification problems in industry. KNN calculates the distances between neighbors, finds the closest neighbors and labels the data according to classes. KNN is one of the most popular machine learning algorithms due to its resistance to old, simple and noisy training data. However, it also has a disadvantage. For example, it requires a large amount of memory space when used for large data, since it stores all states while calculating distance. The main purpose of classification is the process of determining which class the objects belong to by looking at the properties of the objects [29].

Initial, discovery the KNN: scheming the Euclidean distance between the two facts or tuples. $X_1 = (x_{11}, x_{12}, \dots \dots x_{1n})$, $X_2 = (x_{21}, x_{22}, \dots x_{1n})$, the following formula in eq (2.1):

$$dist(x_{11}, x_{12}) = \sqrt{\sum_{i=1}^{n}(x_{i1} + x_{i2})^2}$$

(2.1)

In real operation, to avoid wide gap weight of different starting characteristics In the value domains, the value of the property is normalized as follows in eq (2.2):

$$v^{-} = \frac{v - min_A}{max_A - min_A}$$

(2.2)

Where the minimum and maximum values of the features represented by $min_A$ and $max_A$ respectively. The normalized and non-normalized eigenvalues represented by $v^{-}$ and $v$. Furthermore, the classification calculated according to the equation (2.3).

$$C_X = \arg max \sum_{y \in X_k} I(C_y = j)$$

(2.3)

Where $X_k$ is the KNN contain y, C represented the labe. The value 1 mean that the $I(.)$ is true. Or $(.) = 0$ .

### 2.1.3 Decision Tree

Decision trees are a classifier that starts from a single root and progresses to decision nodes and ends in tagged leaves. Decision trees are included in the scope of supervised machine learning [30]. Decision trees Easy to apply

A.It can be easily combined with database systems

B. Easy to interpret
C. Low cost
D. Algorithms applicable to all variables (qualitative, quantitative, continuous, discrete) to have

E. Frequently used for reasons such as reliable results It is one of the algorithms.

The algorithm consists of roots, decision nodes, branches and leaves. The place where the process begins is the root. Quality tests are performed on decision nodes. Test results can be seen on branches, and finally leaves represent classes. Decision trees divide big data into small groups using the method of induction [31].

### 2.1.4 Support Vector Machines

Supporter vector machines (DVM) is a machine learning method developed by Vladimir Vapnik and Alexey Chervonenkis in the late 1960's. This learning method is a controlled classification algorithm and is based on statistical learning theory [32].

DVM is a frequently used classification method in high dimensional data. It is applied by training on data. That's why it is within the scope of supervised machine learning.

DVM (Supporting Vector Machines) makes the separation of classes by ensuring that a line of appropriate value is found. The lines drawn must be drawn in such a way that they pass from the furthest point to the classes. It was used for the first time to model two-class linear data. SVM, which has developed over time, has also been used in the modeling of more than two-class and non-linear data. The working principle of DVM is to estimate the most appropriate decision function (hyperplane [33].

Supporting Vector Machines create the classification process by converting it to the squared optimization process. Thus, the quantity of transactions performed during learning decreases. In this way, faster solutions can be produced compared to other algorithms. Thanks to this skill, the volume of large data sets useful.

### 2.1.5 Neural Network

Artificial neural networks (ANN), which is a machine learning algorithm designed on the basis of biological neural networks, consists of neurons, cell bodies, dendrites and axons belonging to the biological nervous system. Axons are extensions that enable information transfer. Its main task is to communicate with other neuronal cells through junction points called synapses. Each of the

nerve cells with many nerve cells is in communication. It receives signals from many nerve cells at the same time. If these impulses are above a certain threshold level, they reach other neurons [34].

Artificial neural networks (artifical neural networks), the human nervous system It is an information processing system developed with inspiration from the networks that create it. They are computer programs made by analogy with biological neural networks. In other words, there are weighted links that connect parallel and distributed information processing structures, each consisting of processing units with its own memory.

Artificial neural networks collect information with examples, after making generalizations, when faced with new situations that they have not met before, they can make a judgment about new situations based on the information they have learned before. Thanks to these learning and generalization abilities of artificial neural networks

It has found application areas by spreading rapidly in many branches of science and has successfully found solutions to very complex problems [35].

### 2.1.6 AdaBoost

AdaBoost algorithm from English Adaptive Boosting words It is derived. Yoav Freund and Robert Schapire were the first to recommend AdaBoost. One of the important features that make AdaBoost the most popular machine learning algorithm is that it is superior to others in terms of speed, and the other is that it uses less memory. In addition, AdaBoost is preferred because of its applicability and strong theoretical basis.

AdaBoost is one of the community learning algorithms and can be used to determine classification or regression. In the initial case it starts with an even distribution for each sample. Then, according to his performance, the strongest remained weak finds the classifier. It then updates the weights and focuses on the misclassified sample. In this way, a strong classifier is created as the strongest weak classifiers are gathered together at the end of a certain number of iterations. This increases the classification success [36].

It has been applied to many pattern recognition problems such as activity and person recognition. AdaBoost aimed to increase the performance of algorithms mainly used for classification in order

to obtain a better accuracy. In this context, it tries to correct the attributes classified incorrectly by the classification algorithms [37].

# 3. MATERIAL AND METHODS

## 3.1 DEEP LEARNING

Artificial intelligence is systems or machines that perform various tasks similar to human intelligence and constantly improve themselves. Since they are systems that can learn from artificial intelligence errors that emerged in the 1950s, it constantly improves the system. Machine learning emerged in the 1980s and is to process a given data set and make predictions or classify. There are two types of learning in machine learning algorithms: supervised and unsupervised learning [38].

Supervised learning is learning from tagged data. Both the input and the desired result are defined. Accuracy feedback on the forecast is made by the individual.

Unsupervised learning is the process of learning from unlabeled observations. With the result data, there is no need for training. The algorithm itself derives from data to result. It is expected to discover by itself.

Machine learning, like artificial intelligence, has not shown a significant development for many years. In the 1990s, data mining increased in popularity, and in the early 2000s, deep learning began to develop. The reasons for the failure in the 90s, insufficient data sets, computers are very weak and insufficient in terms of CPU, and incorrect initialization Incorrect nonlinear activation functions Over time, the development of technology has enabled computers to increase their performance. Computers that were insufficient in artificial intelligence processes got stronger, data access became easier and the data sets needed by algorithms increased, and existing algorithms were improved or new algorithms were developed. Thus, development began to be observed in artificial intelligence applications.

Deep learning can be carried out with or without supervision. In deep learning, it learns the distinctive features by itself with lots of data entry. The more data input for the learning process, the more successful it will be. Data goes through multiple layers. Top layers are layers that add more detail. then, there are three main types of deep learning, -Multilayer Perceptron (Multilayer Perceptrons), Convolutional Neural Networks (CNN),and Recurrent Neural Networks [39]. Many applications developed using deep learning techniques such as:

A Face recognition system

B. In voice recognition systems
C. It is used in vehicles with autopilot feature or driverless vehicles.
D. In alarm systems, instead of constantly checking the camera records, technologies such as activating the alarm system only in case of unusual movements are possible thanks to deep learning.
E. It eliminates the loss of time in cancer researches in the health sector. Deep learning algorithms in which cancerous cell samples are introduced are both faster and more successful in diagnosing whether new cells are cancerous or not.
F. Image enhancement
G. In recommendation systems, to offer music and movie suggestions that can be liked
H. Deep learning methods can also be developed in cyber threat analysis. Apart from the examples given above, many examples can be given.

### 3.2 Convolutional Neural Network (CNN)

It is a particular type of advanced nutrition of the neural network that can draw inspiration from the biological processes that occur in the visual lobes (primarily biological brains). It is the solution to many artificial vision problems in artificial intelligence, such as image and video processing. CNN is a multilayer network (MLP) with a unique topology, which contains multiple hidden levels [40] It is differing from traditional MLPs by combining a number of locally connected layers that is used for feature extraction, followed by a number of fully connected(FC) layers which is used for classification [41] .The CNN network is mainly used for object recognition in image processing, handwriting recognition, and voice recognition, as they will automatically extract the distinctive characteristics of their levels from the initial input information without any proper normalization. This type of module is useful for data entry with internal structures (such as images) and where fixed functionality should be found. One of the main problems of using CNN is to avoid manually designed input functions, which may not be obtained by looking at general problems [42]. See figure 3.1.

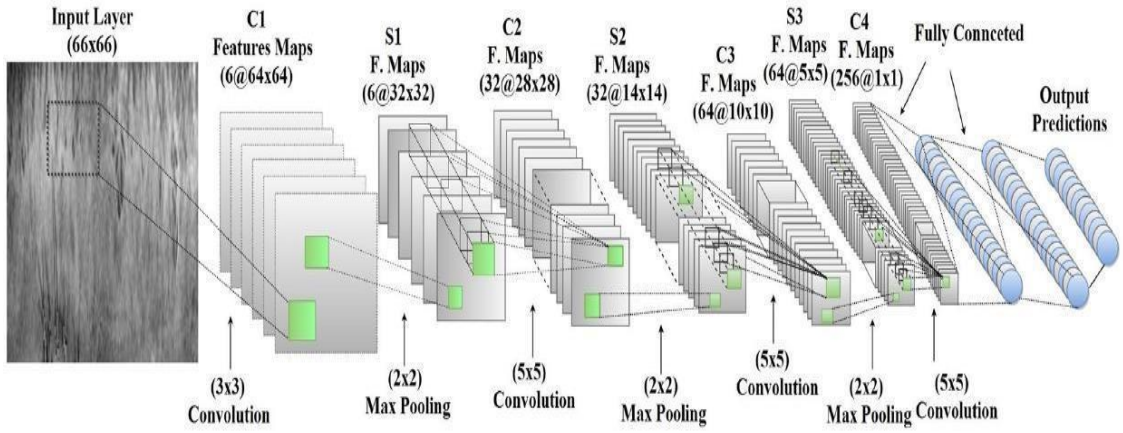


**Figure 3.1:** Classification Using CNN [43]**.**

Represents the feature map and the learning torsion core (filter) respectively. The cross line between the last two layers indicates the fully connected layer. Each pixel in the input image represents a neuron in the input feature map. On CNN, each layer contains multiple feature maps organized into a two-dimensional and two-dimensional network of neurons. Each core of the weight will be combined with all the feature maps in a specific layer to generate a feature map in the next layer. This process is responsible for extracting the local characteristics from the input image [44].

The concepts of local contact and universal weights are applied to convolutional layers, where each neuron receives information the same size as the torsion nucleus from a small area (receptor domain) of the previous layer, which makes the number of weights relatively small. For example, if the nucleus size in a particular convolutional layer is 3 x 3, meaning there are only 9 weights, the local connection confirms that the enabled CNN produces a strong response to capture local entities (such as curves, edges, and edges in the input image)) [45].

**Convolutional Layer**

The main task of a convolutional layer is to recognize the local relationships of features in the previous layer and assign shapes to feature maps. Due to the convolution of the neural network, the image is divided into perceptrons, creating a local reception field, and finally, the perceptron is compressed into a characteristic map of m2 × m3 size. Therefore, this map stores information about the position of elements in the image and their correspondence to filters. Therefore, each filter has a spatial shape due to its position in the volume to which it is applied. [49]. The function of ConvNet is to convert the image into an easy to process format without losing the key

functionality to get the right predictions. This is important when you need to design an architecture that supports large amounts of data as well as functional training. [50]. Mathematical model of convolutional layer presented in Equation 3.1.

$$X_m^L = f(\sum_m x_m^{L-1} * W_{mn}^L) + b_n^L \tag{3.1}$$

Suppose we have an image [32x32x3] in which the first and second dimensions represent width and height and the third dimension represents the size of the color channel of the image, e.g. B. RGB (red green blue). ... If we have a filter of dimension [4x4] (width and height), every neuron of the convolutional layer has a weight tensor of dimension [4x4x3]. The last clock is 3 because the input has 3 channels and a different filter is required for each input channel. Also, if the convolutional layer consists of 32 neurons, the size of the layer becomes [32x4x4x3], each parameter can be trained [36]. As mentioned above, the filter passes through the input image and the output size depends on the length of the slide, also known as the pitch. When the step is set to 1, the filter scrolls one pixel for each dot product operation. If the distance is set to n, the filter is shifted by n pixels after each dotted product operation. Increasing the step value by 28 results in less output as more pixels pass through the filter. Another parameter of the convolution operation is called padding. This parameter is filled with zeros around the input boundary to maintain the spatial size of the input volume so that the width and height of the input and output are the same. Spacing and padding are hyperparameters for folding layers and should be optimized for the problem. The scrolling convolution filter can be applied to different sizes depending on the input. In a one-dimensional (1D) convolution, the filter flows along an axis that can be transient. The result of a one-dimensional convolution method is a one-dimensional matrix. With 2D convolution, the filter runs along two axes (x, y) and creates a 2D matrix. Convolutional 3D scrolls the filter along 3 axes (x, y, z) and the output has a 3D volume. Most levels of convolution apply 2D convolutions to 3D data because there is a filter for the third dimension. For 3D data, the 2D filter moves along the x and y axes. Hence the output is also a 2D matrix with a 3D volume. As shown in Figure 2.6, the connectivity model of convolutional layers also differs from fully connected layers. Folding layers are locally connected layers. If the size of the input image is [32x32x3] which contains 3 color channels and the neuron in the convolutional layer contains [4x4] filters, each channel has its own filter with a total size of [4x4x3]. Each filter [4x4] within

the neuron is rotated with an input image channel. Therefore, convolutional layers use the local interconnection model instead of assigning each filter to each input channel, as is the case with fully interconnected layers [47].

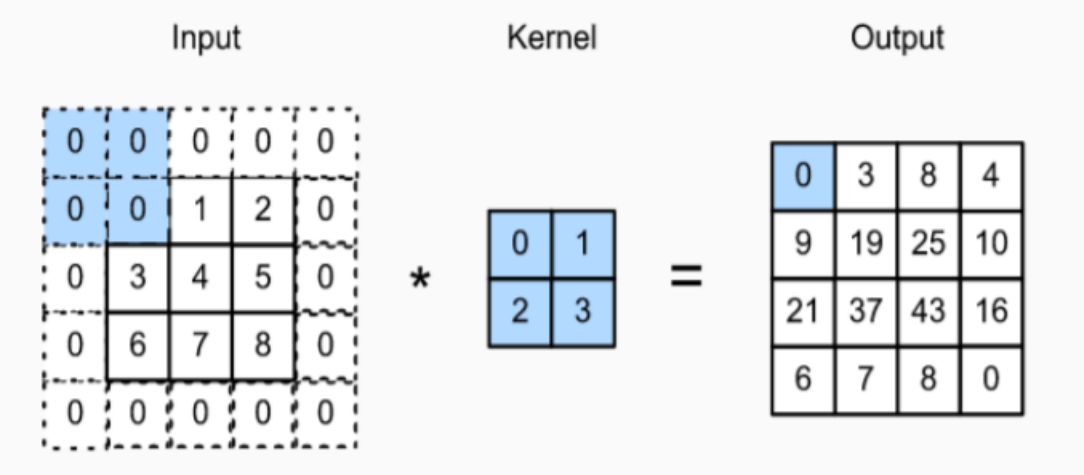


**Figure 3.2:** Activation Function [48].

### A. Local Connectivity

When it comes to multidimensional inputs like images, as we saw above, it is impractical to connect neurons to all neurons in the previous range. Instead, we connect each neuron to only one local area of the input volume. The spatial extent of this connection is a hyperparameter called the receptor field of the neuron (this corresponds to the size of the filter). The length of the gutter along the depth axis is always the same as the depth of the inlet volume. It is important to re-emphasize this asymm etry in the way we work with measurements of spatial dimensions (width and height) and depth: connections are local in space (in width and height), but are always complete throughout the entire depth of the volume entrance [49].

### B. Parameter Sharing

The parameter sharing scheme is used to control the number of parameters at the convolution level. Using the example above, you can see that the first conv-layer contains 55 * 55 * 96 = 290,400 neurons each with 11 * 11 * 3 = 363 weights and 1 slope. In total, this only adds up to 290400 * 364 = 105705600 parameters in the first level of ConvNet. It is clear that it is a lot.

It turns out that with reasonable assumptions, the number of parameters can be reduced significantly. If the function is useful for computation in one spatial location (x, y), then in another location (x2, y2). That is, setting a 2D depth portion as a depth portion (e.g., volume [55 x 55 x 96] has 96 depth portions each of [55 x 55]) to constrain neurons to each depth portion. Use the same weights and offsets. With this parameter sharing scheme, the first convection layer in this example only has 96 sets of unique weights (one for each depth layer), making a total of 96 * 11 * 11 * 3 = 34,848 weights. unique or 34,944 parameters. There are. (+96 bias). Or 55 * 55 neurons in each depth segment now use the same settings. In fact, as it propagates backwards, each neuron in the volume computes a gradient for its weight, but these gradients are added to each depth segment and only update one set of weights for each segment [50].

C.Fully Connected Layers

Lastly, after applying multiple convolutional and max pooling and at the end of convolutional and pooling layers, the high-level features in the neural network are finished by fully connected layers. Where in fully connected layer each single neuron is connected to each single neuron from the preceding layer. There are no layers after a fully connected layer [51,52, 53]. Also, not anymore, they have spatial located (it can be visualizing as one-dimensional. The main goal behind the Fully Connected layer is to utilize their features for identifying the fed date into different classes dependent on dataset for training, for the identification work, a lot of the features from convolutional and pooling layers may be best, likewise may be blends of those features stunningly better [54].

$$E(W) = \left[\sum_{i=1}^{m}\sum_{r=1}^{N} 1\{y^{(i)} = r\} \, log \frac{\exp(w_r^T x^{(i)})}{\sum_{j=1}^{N} \exp(w_r^T x^{(i)})}\right] \tag{3.2}$$

**D. Non-linear Activation Function**

Activation functions are really important to learn complex and non-linear mappings between the input and output. CNNs without non-linear activation functions would simply be a linear regression model, unable to model complicated non-linear functions [55]. The main purpose of using CNN is to make sense of something which is complex, high dimensional and non-linear such as videos, audio, images etc. This shows the importance of activation function in CNNs. There are

different kinds of non-linearities such as Sigmoid, Tanh-hyperbolic tangent and Rectied linear units (ReLU) [56]. Figure. 3.3 shows the plot of all the three functions.

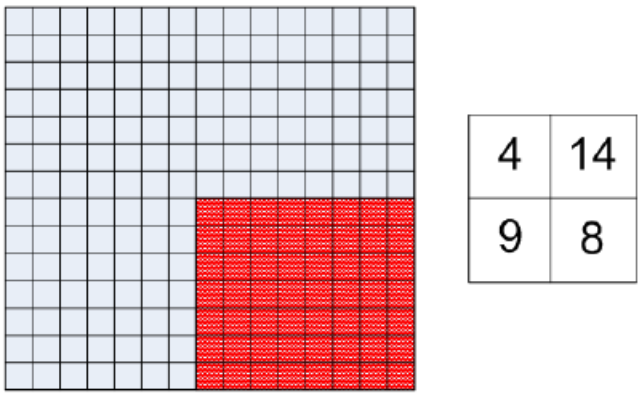


**Figure 3.3:** Activation Function [57]**.**

This function mathematically can represented as shown below:

$$f(x) = max(0, x) \qquad 3.3)$$

Why ReLU is important: The goal of ReLU is to give ConvNet non-linearity. I want ConvNet to know the real world data, so there are no negative linear values.

### E. Stride

Stride is the number of pixels in the input matrix. If the step is 1, we move the filters 1 pixel at a time. If the step is 2, we move the filters 2 pixels each, and so on. The picture below shows that a two-step step will work [58, 59].

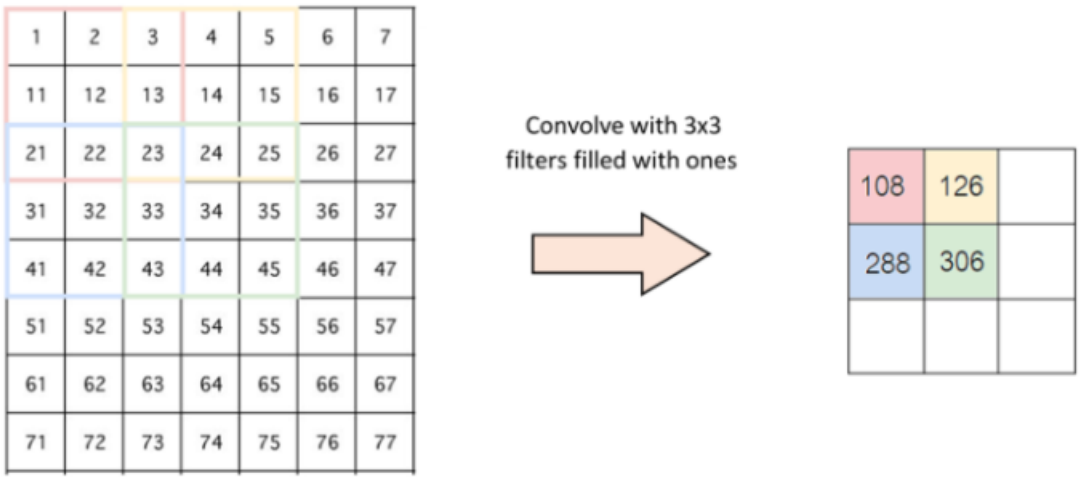


**Figure 3.4:** Stochastic Gradient Descent with Momentum algorithm [64]

## 3.3 GRADIENT DESCENT OPTIMIZATION ALGORITHMS

- **Stochastic Gradient Descent with Momentum**

SGD and its variants are probably the most used optimization algorithms for machine learning, especially for deep learning tasks [60, 61, 62]. However, It can be slow without acceleration method such as momentum. When loss curvature is high and gradients are consistent but small, momentum acceleration method is designed to accumulate the gradients using moving average of past gradients. Momentum method introduce a velocity term that helps to accelerate consistent gradient direction. Momentum coefficient $\alpha \in (0; 1]$ adjusts how previous gradient contributions affect the current momentum step. 0.5 and 0.9 are typical values of momentum coefficient. SGD with momentum update is defined in Figure 3.5.

**Algorithm 1** Stochastic Gradient Descent with Momentum

**Require:** Learning rate $\epsilon$, momentum parameter $\alpha$

**Require:** Initial parameter $\theta$, initial velocity $\vartheta$

**while** training stopping criterion not met **do**

Sample a mini-batch of m examples from training set: $x_1, x_2 \ldots x_m$ with corresponding labels $y_1, y_2 \ldots y_m$

Compute gradient estimation $g \leftarrow \frac{1}{m} \nabla_\theta \sum_i (L(f(x_i; \theta), y_i))$

Compute velocity $\vartheta \leftarrow \alpha\vartheta - \epsilon g$

Apply parameter update $\theta \leftarrow \theta + \vartheta$

**end while**

**Figure 3.5:** Stochastic Gradient Descent with Momentum algorithm [65]

**Adaptive Learning Rate Optimizers**

Learning rate is the most sensitive parameter that strongly affects performance of a model. It is also known that model cost highly relies on some parameters from model's parameter space. This brings up the idea of using separate learning rates per parameter based on their sensitivity, also adapting per parameter learning rates automatically. Earlier heuristic is increasing the learning rate if partial derivative sign of loss with respect to a model parameter remains the same, oppositely, decreasing the learning rate in case of gradient changes sign [66]. In this approach, parameters with larger partial derivatives have accelerated decrease compared to parameters with small partial derivatives. Even though this approach helps some models, keeping historical partial derivatives from beginning of training to end can introduce earlier and excessive decrease in the effective learning rate for some cases. Rmsprop modifies AdaGrad algorithm to address its excessive decrease on learning rate by weighted moving average of gradient history. Instead of AdaGrad, Rmsprop keeps moving average of gradients history with a new hyper-parameter that adjusts how historical gradients impact moving average besides current gradient. Aside from AdaGrad and Rmsprop, Adam optimizer combines Rmsprop and momentum optimizers benefits. Adam optimizer computes two historical moving average estimates which keep gradients and squared of gradients respectively [67]. There is also correction of initial bias of moving average of gradients

and square of gradients in Adam that is also an addition to Rmsprop algorithm. Although Adam is capable to work with different models with default values, some cases still requires global learning rate and other hyper-parameters tuning.

### 3.4 TRANSFER LEARNING

Transfer learning is essential machine learning tool for solving the underlying problem of poor training data. We try to transfer knowledge from the supply area to the target area, weakening the belief that training and testing data lead to significant positive effects in many areas that are difficult to improve due to poor training data [68, 69]. There are many pretrained organizations and the reasons for using pretrained models are listed below:

A. First, it needs additional computational power to train the large models on huge datasets.

B. Second, high time taking to train the network up to several of days.

There are several categories of transfer learning listed below:

C..Instances-based deep transfer learning:

Instance-based portable deep learning is a group of training courses in the target area, limited to the resource area by applying a specific approach to weight correction and applying the appropriate weighting principles to these selected examples. Possibilities to choose as an add-on for.

C. Mapping-based deep transfer learning: Mapping-Driven Deep Learning describes how to map source and destination field examples to new functional areas. Examples for both domains in this new data area are and are potentially suitable for corporate DNNs.

D. Network-based deep transfer learning: Reconditioning of a restricted network previously trained in the domain of the deep learning provider of the distribution network, including the parameters related to that network configuration, indicates that it will be broadcast as part of the DNN applied to the target domain. He says: “NN is comparable to the processing mechanisms of the human brain and an iterative and continuous step of generalization.

Researchers present several previously formed networks, and these networks are used in many classification problems [70].

### 3.4.1 AlexNet

AlexNet is a CNN channel hosted by Alex Krizhevsky. AlexNet was involved in ImageNet's core work on image recognition in 2012. The network was ranked in the top five errors with a 15.3% rate, down 10.8% from its completion rate.

AlexNet is the name of a convolutional network that has had a great impact on machine learning, especially in deep learning computer vision applications. As you know, we won the 2012 ImageNet LSVRC competition by a wide margin (26.2% (second place) with an error rate of 15.3%). Network - Yang LeCun et al. But LeNet was deeper in a convolutional layer that applies multiple filters layer by layer. Consists of 11x11, 5x5.3x3, line, max connection, interrupt, data increase, ReLU trigger, SGD with pulse. Activating Connected ReLU after fully connected layer with all stacking layers. AlexNet trained with two Nvidia Geforce GTX 580 GPUs simultaneously for six days, splitting the network into two channels.

AlexNet has eight levels. The convolutional layer was the first 5 geotags, some of which had the maximum level of clustering, the fc layer shown in the last three, and the activation function [71] was not applied. Shown improved training accuracy for Thane and Sigmoid. The AlexNet structure presented in the Figure 3.6.

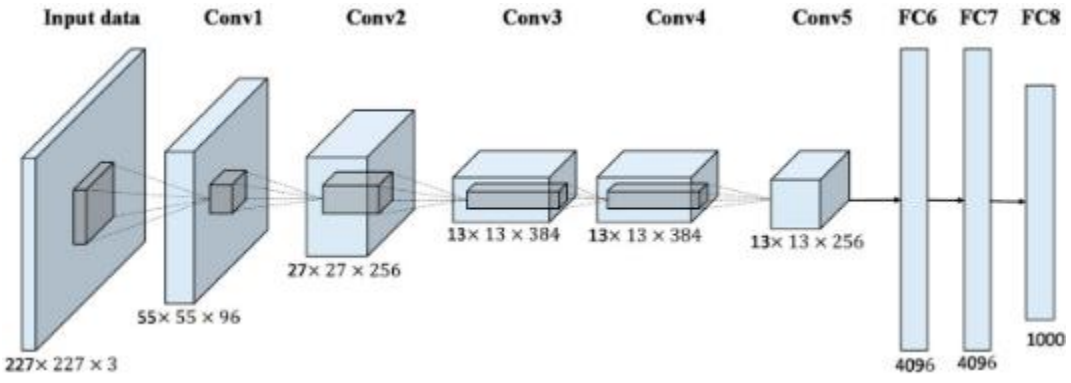


**Figure 3.6:** AlexNet Structure [72]

### 3.4.2 GoogleNet

A major improvement in GoogleNet uses a framework called Inception [73]. Inception is typically a network architectured network, and the optimal low-density LAN architecture repeats itself spatially from start to finish. Three inception architectures are presented for different environments. Inception usually calculates the layers in front of the expensive 3x3 and 5x5 turns with 1 × 1 turns [74]. The structure of GoogleNet represented in Figure 3.7.

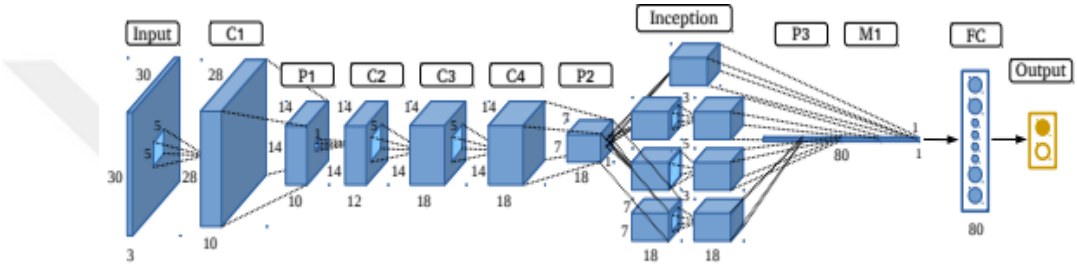


**Figure 3.7:** GoogleNet Structure [74]

The GoogleNet Architecture is 22 layers deep, with 27 pooling layers included. There are 9 inception modules stacked linearly in total. The ends of the inception modules are connected to the global average pooling layer. Below is a zoomed-out image of the full GoogleNet architecture. Furthermore, the structure of the GoogleNet presented in detail dorm in Table 3.1.

**Table 3.1:** GoogleNet Structure [75].

| type | patch size/ stride | output size | depth | #1×1 | #3×3 reduce | #3×3 | #5×5 reduce | #5×5 | pool proj | params | ops |
|---|---|---|---|---|---|---|---|---|---|---|---|
| convolution | 7×7/2 | 112×112×64 | 1 | | | | | | | 2.7K | 34M |
| max pool | 3×3/2 | 56×56×64 | 0 | | | | | | | | |
| convolution | 3×3/1 | 56×56×192 | 2 | | 64 | 192 | | | | 112K | 360M |
| max pool | 3×3/2 | 28×28×192 | 0 | | | | | | | | |
| inception (3a) | | 28×28×256 | 2 | 64 | 96 | 128 | 16 | 32 | 32 | 159K | 128M |
| inception (3b) | | 28×28×480 | 2 | 128 | 128 | 192 | 32 | 96 | 64 | 380K | 304M |
| max pool | 3×3/2 | 14×14×480 | 0 | | | | | | | | |
| inception (4a) | | 14×14×512 | 2 | 192 | 96 | 208 | 16 | 48 | 64 | 364K | 73M |
| inception (4b) | | 14×14×512 | 2 | 160 | 112 | 224 | 24 | 64 | 64 | 437K | 88M |
| inception (4c) | | 14×14×512 | 2 | 128 | 128 | 256 | 24 | 64 | 64 | 463K | 100M |
| inception (4d) | | 14×14×528 | 2 | 112 | 144 | 288 | 32 | 64 | 64 | 580K | 119M |
| inception (4e) | | 14×14×832 | 2 | 256 | 160 | 320 | 32 | 128 | 128 | 840K | 170M |
| max pool | 3×3/2 | 7×7×832 | 0 | | | | | | | | |
| inception (5a) | | 7×7×832 | 2 | 256 | 160 | 320 | 32 | 128 | 128 | 1072K | 54M |
| inception (5b) | | 7×7×1024 | 2 | 384 | 192 | 384 | 48 | 128 | 128 | 1388K | 71M |
| avg pool | 7×7/1 | 1×1×1024 | 0 | | | | | | | | |
| dropout (40%) | | 1×1×1024 | 0 | | | | | | | | |
| linear | | 1×1×1000 | 1 | | | | | | | 1000K | 1M |
| softmax | | 1×1×1000 | 0 | | | | | | | | |

### 3.4.3 ResNet101

The thread depth has increased, but the accuracy will not necessarily improve. However, there are some near-clean issues in ResNet. Greater depth, which requires a change in weight at the edges of the net, slightly improves the estimate of the stock layer [76]. Add-ons are required parameters. The structure of ResNet101 presented in Figure 3.8.

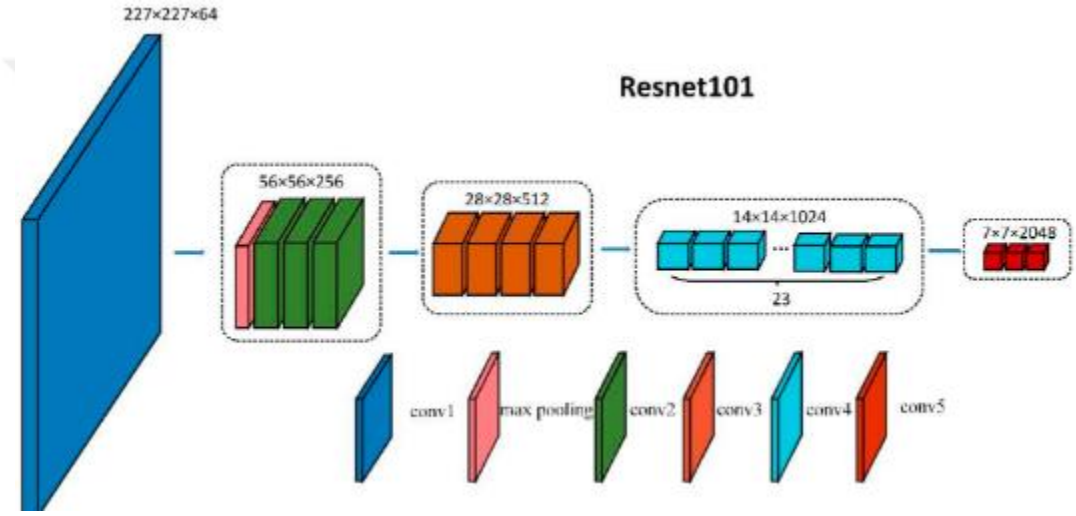


**Figure 3.8:** ResNet101 Structure [77]

### 3.4.4 DenseNet201

DenseNet was recognized for the astonishing presentation on the cutthroat object recognition standard tasks like ResNet101 and Alexnet. The initial effort to solve huge scale image recognition task by deep CNN, earned the first place and second place in localization and recognition assignment separately. DenseNets have numerous exciting benefits: they improve the vanishing-gradient problem, improve feature reproduction, promote feature recycle, and significantly decrease the number of parameters [78]. The structure of DenseNets201 shown in Figure 3.9.

| Layers | Output Size | DenseNet-121($k = 32$) | DenseNet-169($k = 32$) | DenseNet-201($k = 32$) | DenseNet-161($k = 48$) |
|---|---|---|---|---|---|
| Convolution | 112 × 112 | 7 × 7 conv, stride 2 | | | |
| Pooling | 56 × 56 | 3 × 3 max pool, stride 2 | | | |
| Dense Block (1) | 56 × 56 | [1 × 1 conv, 3 × 3 conv] × 6 | [1 × 1 conv, 3 × 3 conv] × 6 | [1 × 1 conv, 3 × 3 conv] × 6 | [1 × 1 conv, 3 × 3 conv] × 6 |
| Transition Layer (1) | 56 × 56 | 1 × 1 conv | | | |
| | 28 × 28 | 2 × 2 average pool, stride 2 | | | |
| Dense Block (2) | 28 × 28 | [1 × 1 conv, 3 × 3 conv] × 12 | [1 × 1 conv, 3 × 3 conv] × 12 | [1 × 1 conv, 3 × 3 conv] × 12 | [1 × 1 conv, 3 × 3 conv] × 12 |
| Transition Layer (2) | 28 × 28 | 1 × 1 conv | | | |
| | 14 × 14 | 2 × 2 average pool, stride 2 | | | |
| Dense Block (3) | 14 × 14 | [1 × 1 conv, 3 × 3 conv] × 24 | [1 × 1 conv, 3 × 3 conv] × 32 | [1 × 1 conv, 3 × 3 conv] × 48 | [1 × 1 conv, 3 × 3 conv] × 36 |
| Transition Layer (3) | 14 × 14 | 1 × 1 conv | | | |
| | 7 × 7 | 2 × 2 average pool, stride 2 | | | |
| Dense Block (4) | 7 × 7 | [1 × 1 conv, 3 × 3 conv] × 16 | [1 × 1 conv, 3 × 3 conv] × 32 | [1 × 1 conv, 3 × 3 conv] × 32 | [1 × 1 conv, 3 × 3 conv] × 24 |
| Classification Layer | 1 × 1 | 7 × 7 global average pool | | | |
| | | 1000D fully-connected, softmax | | | |

**Figure 3.9:** DenseNet201 Structure [79]

### 3.4.5 VGG-16

Oxford University Simonyan and Zisserman are developing sixteen floors, sixteen of which three are fully connected at the folding level. A CNN using only a path with a 3 × 3 filter and pad, and the path is up to VGG-16 with a 2 × 2 filter of 2. To reduce the number of parameters in the CNN, a small filter used 3x3 is applied to all convolutional layers with a filter with an error rate of 7.3% [80]. VGG-19 is capable of recognizing various images, classifying videos, finding objects, etc. The structure of VGG-16 shown in Figure 3.10.

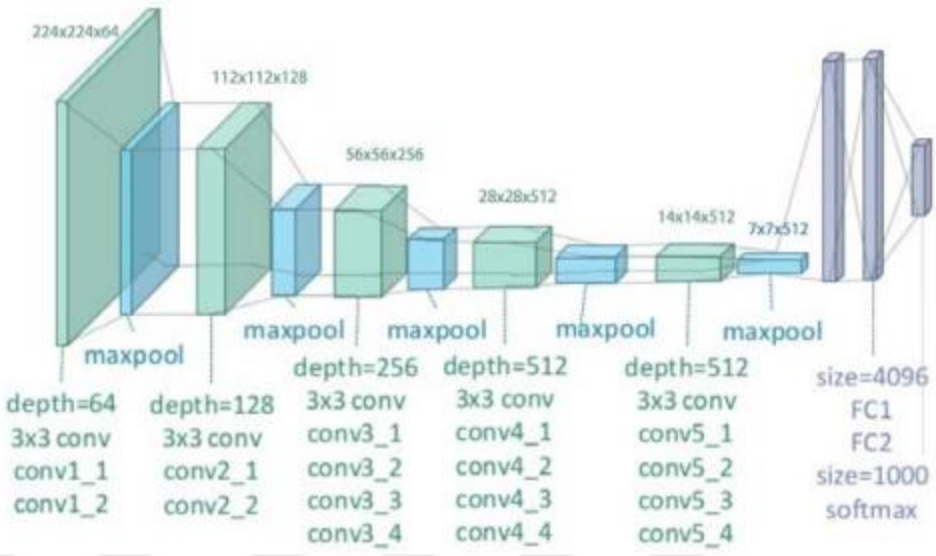


**Figure 3.10:** VGG16 Structure [81]

## 3.5 HUMAN IDENTIFICATION WITH FINGER VEIN IMAGE USING CNN BASED KNN AND DT METHOD

In this section, an efficient method based CNN and KNN are presented for human action recognition with action images using deep learning. In the first stage, we prepare the vein images which each vein images of each person saved in one folder. The aim of this stage is seperae the data of each person which can analyse and learned with CNN model. In the second stage, the framework separed the images into train and test groups automaticly. The aim of this stage is to train the model with selected data and test it with other data to ensure that the model learn well or what is the learning rate of the model. Then, the network that we applied consist from two parts feature extraction part and classifier. The feature extraction part which is done by using CNN the aim of this part is extracted high level features such as arc and Circle. The pretrained networks used in this section to avoid overfitting and to obtain best features for classification rate. In the second part of the network we applied KNN, DT and Extreme learning machines (ELM) which are not applied in previous studies for finger vein image classification. Then, the proposed method structure presented in the Figure 3.11.

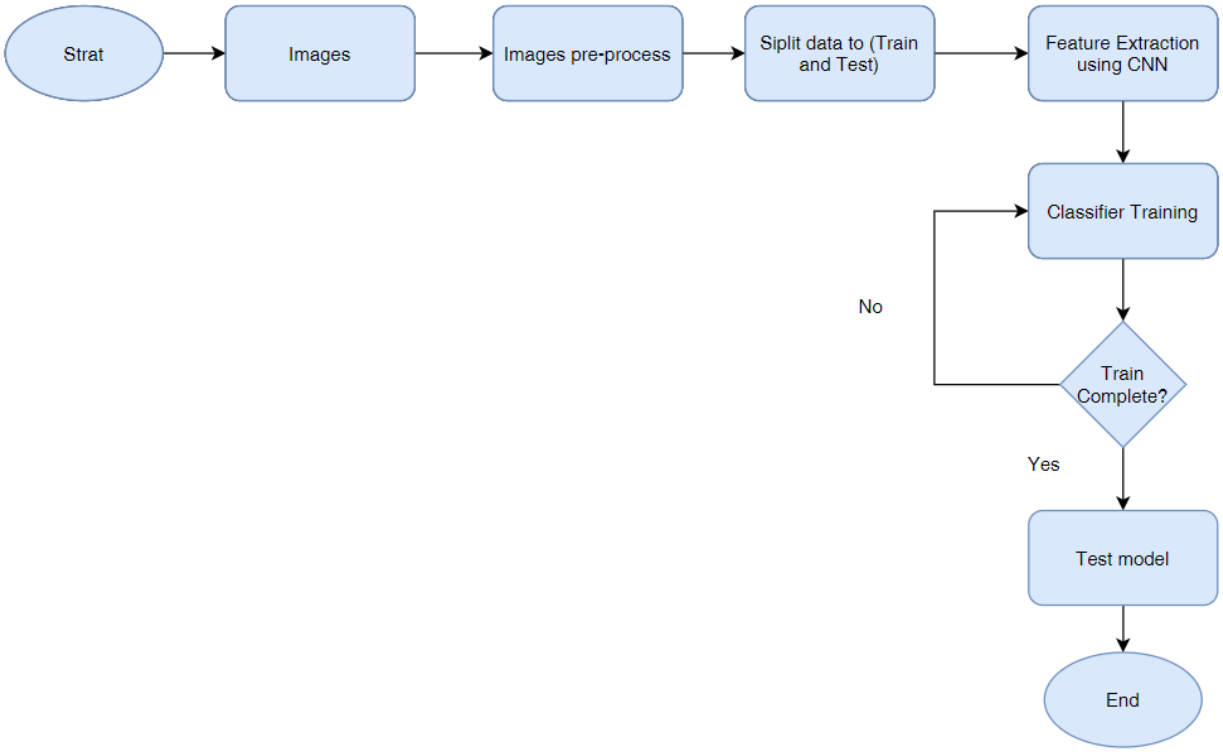


**Figure 3.11:** Proposed Method.

# 4. EXPERIMENTS AND DISSCUSION

Human activity classification systems aim to indicate the action that is carried out by defining the environment in which we are living with the data obtained from humans. This way about the person Information can be obtained when requested. Human movements in general; It is divided into three main groups as short events, basic activities and complex activities. If we give an example to the movements called short activities; While it can be called the transition from sitting position to another position, walking is a basic activity. Finally, movements performed in association with different objects and individuals are complex movements. Playing an instrument and various sports movements are examples of complex activity groups. In this study, the movements within the basic activities group were classified. Within this activity group, 6 basic movements are specified. These; It includes three static activities, including standing, reaching, and sitting, and three dynamic activities, including walking, climbing and climbing stairs. The basic description of the constructed structure is shown in Figure 4.1.

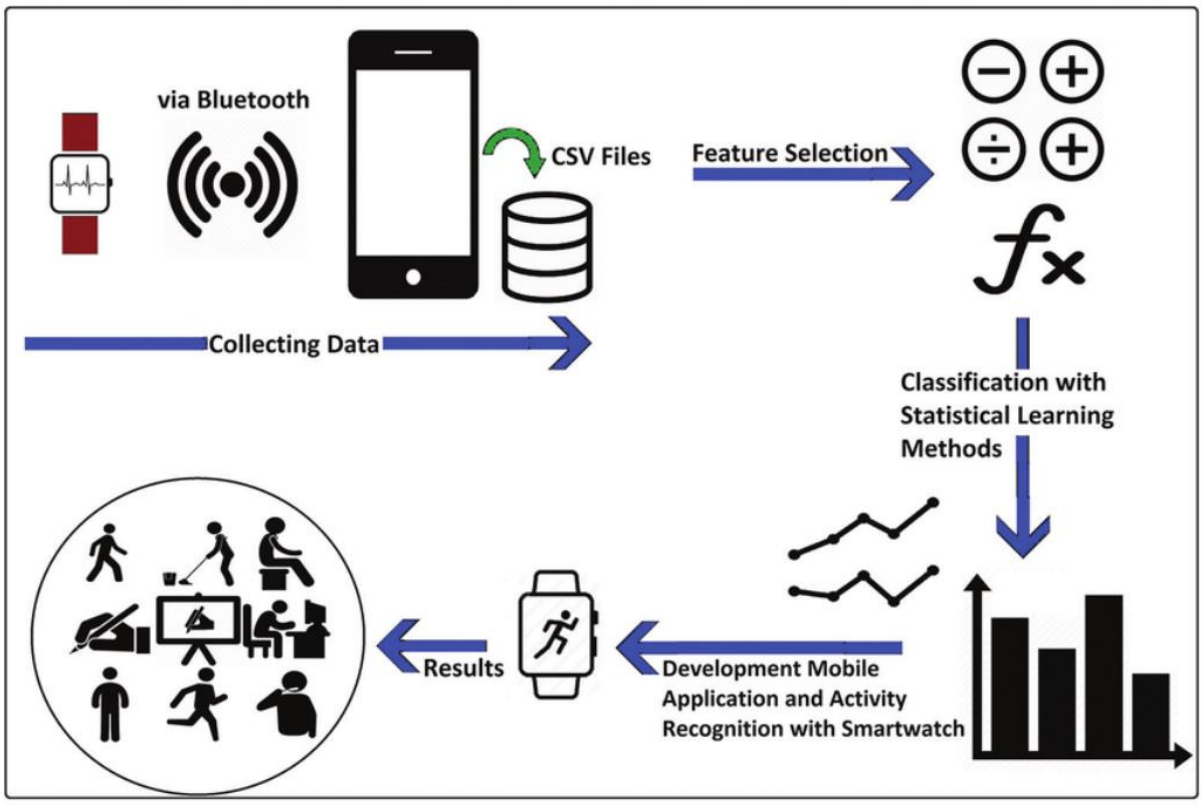


**Figure 4.1:** Random Subsampling [82].

Several experiments are executed in various scenarios, the experiment and the result were assessed using various measurements, the act of various experiments was compared and the results were highlighted. The experimental implemented by using MATLAB2020 as tool. The dataset features that explained in the chapter 3 used as input to the proposed system that written by using MATLAB2020.

## 4.1 EVALUATION PARAMETERS

In this section four parameters are calculated for face recognition problem to evaluate our proposed method the calculated statistical parameters are: Accuracy is how close a measured value is to the actual (true) value. The confusion matrix is used to evaluate the results of our method by calculating several parameters such as: True positives (TP), True negatives (TN), false positives (FP), and false negatives (FN).

## 4.2 RANDOM SUBSAMPLING

In this method, random sub-sampling validation is track for a secure amount of K iterations. Throughout apiece iteration it used random sampling minus spare, in instruction to choice a secure amount of S examples that brand up the examination set and are excepted from the training procedure of the perfect. The experiential test statistics are then be around over all iterations. The Figure 4.1 represented the process of the validation using Random Subsampling.

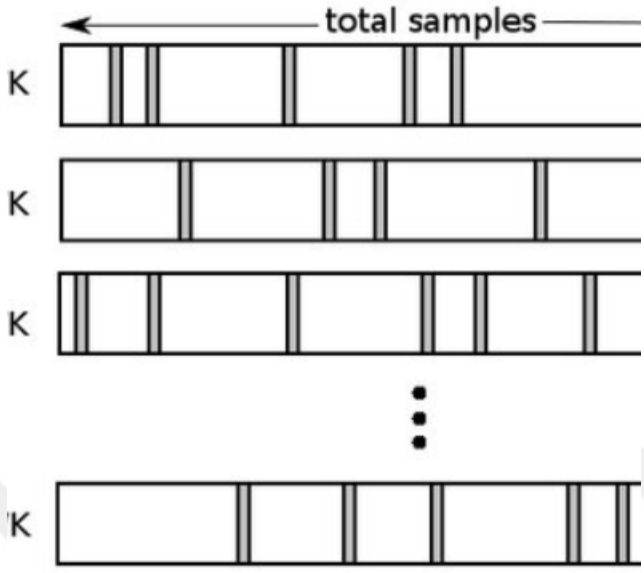


**Figure 4.2:** Random Subsampling [83].

## 4.3 IMPLEMENTATION

Four experiments were conducted to evaluate the performance and classification of a CNN-based system for HAR using two sensors on a Samsung Galaxy S2 smartphone. The first and second experiments are repeated with filter size and number of filters, respectively. The third experiment explores the possibility of changing the number of layers in the network for the CNN algorithm to see if it can improve network performance. The last experiment is to increase the number of training periods for the data used. The level of precision was recorded for each scenario. The learning and testing phase mainly works for:

HAR containing 10,298 samples for six daily human activities. While 70% of the data set is used in the training phase, the remaining 30% is used in the testing phase. Each sample was taken from a smartphone battery and gyro sensors.

In this section, the results of our method presented and discussed in both the accuracy and the execution time for two methods that's applied for human identification with finger vein image. The main evaluation parameter for evaluating our method was calculated as shown in Eq (4.1)

$$Acc = \frac{TP + TN}{Total} \tag{4.1}$$

Where ACC represented the accuracy of the system, TP presented the true positive, TN represented the true negative, and Total represented the total estimation of the system and the results presented in Figure 4.2 and 4.3.

For post-test, the number of periods is increased from 300 to 1000 to determine the trend in system accuracy. The X-axis represents the number of periods that we have increased from 300 to 1000, and the Y-axis represents the level of precision that each epoch has achieved in our experience. The number of periods is the number of iterations throughout the training until the completion of the CNNet training algorithm. In this experiment, each increase in age lasts longer than the previous one, and the average yield for this test is 98.70%. During our experience, we have found

You need to take care of determining the correct amount of workouts because multiple workouts will not show significant changes or trends in system performance. To speed up our solution, just use a GPU instead of a CPU.

Three experiments applied to evaluate the presented methods CNN-based DT, KNN, and ELM. The results of these studies presented below and compared which show that the ELM based CNN presented best results.

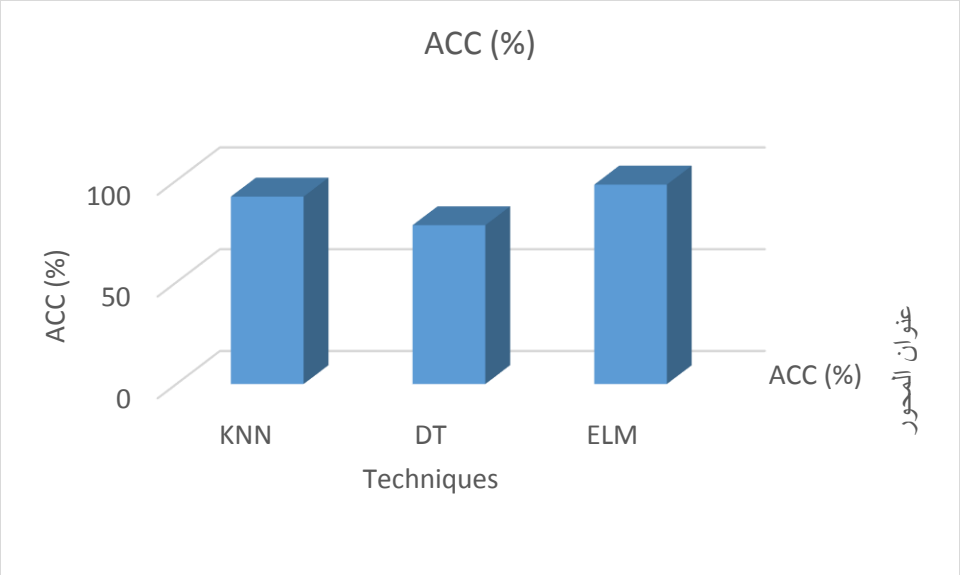


**Figure 4.3:** Comparisons of several techniques executed in this study

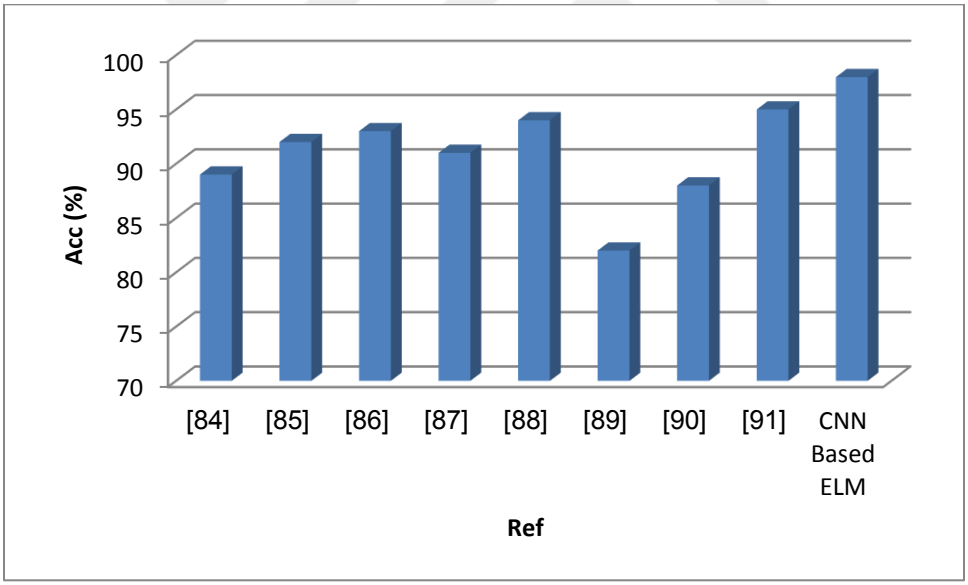


**Figure 4.3:** Comparisons with previous studies

Our method presented best results than all common studies that are published in well known journals. Our method combines the power of the CNN in the feature extraction stage with the power of the CNN in the classification stage. These resaons lead to presented best results than previous studies in both accuracy and execution time.

# 5. CONCLUSION

In this study, training and tests were carried out to classify 6 human activities of UKSB cells created with a dataset created by collecting accelerometer and gyroscope data within smart phones. In these tests, by giving different input parameters to UKSB cells, the highest value was obtained as 97.98% success rate. Compared to other artificial neural networks (eg multilayer perceptron), CNN stands out because it can automatically process attributes from the data set and automatically analyze the temporal relationship. For these reasons, CNN was preferred in our study. In this review, the current status of the articles that do the activity recognition process is shown. First, the introduction of activity recognition, the sensors used, simple and complex activities, the general method used and then the steps in the activity recognition process were discussed in detail. The current situation in this area has been revealed and the studies conducted have been compared in many respects. It has been determined that the success of new generation deep learning techniques compared to classical methods, battery consumption, model update alternatives, data set sizes have not yet been analyzed in detail and comparatively. Another obvious point is to recognize mixed activities. Few activity recognition studies have studied mixed activities and only one of the results is positive. Along with mixed activity, another potential field of study is overlapping activities. For example, driving while talking on the phone, listening to music while walking, watching movies while traveling. Another issue similar to mixed activities is the recognition of social, collective activities rather than one-person activity. One of the problems encountered in the articles is the device orientation and location problem. Few articles have provided both device orientation and activity recognition independent of location. More attention should be paid to this issue in order to make an application that will be used by people in daily life. The potential study issue regarding the use of sensors is to choose a dynamic sensor. Instead of using fixed and the same sensors to recognize all activities, using the sensor dynamically according to the activity can give positive results. In the studies conducted so far, the problems of recognizing daily activities and determining the type of transportation have been discussed separately. These two problems should be taken together, especially in order to keep track of a person's movement integrity throughout the day.

# REFERENCES

[1] Statistica, https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/ (Ziyaret Tarihi: 31 Aralık 2018).

[2] Cisco, VNI Global Fixed and Mobile Internet Traffic Forecasts, https://www.cisco.com/c/en/us/solutions/service-provider/visual-networking-index-vni/index.html#~mobilevni (Ziyaret Tarihi: 10 Mayıs 2018)

[3] Ranao C., Cho S., Human Activity Recognition Using Smartphone Sensors With Two-Stage Continious Hidden Markov Models, International Conference on Natural Computation (ICNC), Xiamen, China, Aug. 19-21, 2014.

[4] Lester J., Choudhury T., Borriello G., A Practical Approach to Recognizing Physical Activities, International Conference on Pervasive Computing, Ireland, May 07-10, 2006, 1-16.

[5] Bayat A., Pomplun M., Tran D. A., A study on human activity recognition using accelerometer data from smartphones. Procedia Computer Science, 34, 450-457.

[6] Hsieh C.Y., Huang C.N., Liu K.C., Chu W.C., Chan C.T., A machine learning approach to fall detection algorithm using wearable sensor, Advanced Materials for Science and Engineering (ICAMSE), Tainan, Taiwan, Nov. 12-13, 2016, 707-710.

[7] Catalbas B., Yucesoy B., et al, A comparative study of classification methods for fall detection, Signal Processing and Communications Applications Conference (SIU), Trabzon, Turkey, Apr. 23-25, 2014, 1315-1318.

[8] Erdogan S.Z., Bilgin T.T., A data mining approach for fall detection by using k-nearest neighbor algorithm on wireless sensor network data, IET Communucations, 2012, 6(18), 3281–3287.

[9] Lee T., Mihailidis A., An intelligent emergency response system: preliminary development and testing of automated fall detection, Journal of telemedicine and telecare, 2005, 11(4), 194-198.

[10] Rougier C., Meunier J., St-Arnaud A., Rousseau J., Robust video surveillance for fall detection based on human shape deformation, IEEE Transactions on circuits and systems for video Technology, 2011, 21(5), 611-622.

[11] Kurban O.C., Yildirim T., Neural network based daily activity recognition without feature extraction, Signal Processing and Communications Applications Conference (SIU), Apr. 23-25, 2014, 567-570.

[12] Wei L., Shah S.K., Human Activity Recognition using Deep Neural Network with Contextual Information, International Conference on Computer Vision Theory and Applications (VISAPP), 2017, 34-43.

[13] Bao L., Intille S., Activity recognition from user-annotated acceleration data, International Conference on Pervasive Computing, Berlin, Germany, 2004, 1-17.

[14] Banos O., Damas M., Pomares H., Prieto A., and Rojas I., Daily living activity recognition based on statistical feature quality group selection, Expert Systems with Applications, 2012, 39(9), 8013–8021.

[15] Ravi N., Dandekar N., Mysore P., Littman M.L., Activity recognition from accelerometer data, Innovative applications of artificial intelligence, Pittsburtg, ABD, July 09, 2005, 1541-1546.

[16] Kwapisz J.R., Weiss G.M., Moore S.A., Cell phone-based biometric identification, Theory Applications and Systems (BTAS), Washington, DC, USA, Sept. 27-29, 2010, 1-7.

[17] Kwapsiz J. R., Weiss G. M., Moore S.A., Activity Recognition Using Cell Phone Accelerometers, SIGKDD Explor. Newsl., New York, NY, USA, December, 2011, 12(2), 74-82.

[18] Chernbumroong S., Atkins A. S., and Yu H., Activity classification using a single wrist-worn accelerometer. In Software, Knowledge Information, Industrial Management and Applications (SKIMA), 2011, 1-6.

[19] Wu W., Dasgupta S., Ramirez E.E., Peterson C., Norman G.J., Classification accuracies of physical activities using smartphone motion sensors, Journal of Medical Internet Research, 2012, 14(5), 5317-5337.

[20] Anjum A., Ilyas M. U., Activity Recognition using smartphone sensors, Consumer Communications and Networking Conference (CCNC), Las Vegas, NV, USA, Jan. 11-14, 2013, 914-919.

[21] Lee Y., Cho S., Activity Recognition Using Hierarchical Hidden Markov Models on a Smartphone with 3D Accelerometer, International Conference on Hybrid Artificial Intelligence Systems (HAIS), 2011, 460-467.

[22] Kim Y., Kang B., Kim D., Hidden Markov Model Ensemble for Activity Recognition using Tri-axis Accelerometer, IEEE International Conference on Systems, Man and Cybernetics (SMC2015), Kowloon, China, Oct. 9-12, 2015, 3036-3041.

[23] Ranao C., Cho S., Recognizing Human Activities from Smartphone Sensors Using Hierarchical Continuous Hidden Markov Models, International Journal of Distributed Sensor Networks, 2017, 13, 1-16.

[24] Han C.W., Kang S.J., Kim N.S., Implementation of hmm-based human activity recognition using single triaxial accelerometer. IEICE transactions on fundamentals of electronics, communications and computer sciences, 2010, 93(7), 1379-1383.

[25] Mannini A., Sabatini A.M., Machine learning methods for classifying human physical activity from on-body accelerometers, Sensors, Feb 1, 2010, 10(2), 1154-1175.

[26] Bao L., Intille S., Activity recognition from user-annotated acceleration data, International Conference on Pervasive Computing, Berlin, Germany, 2004, 1-17.

[27] Janine George, Geoffroy Hautier, Chemist versus Machine: Traditional Knowledge versus Machine Learning Techniques, Trends in Chemistry, 2020, ISSN 2589-5974, https://doi.org/10.1016/j.trechm.2020.10.007

[28] Christopher R. Cox, Emma H. Moscardini, Alex S. Cohen, Raymond P. Tucker, Machine learning for suicidology: A practical review of exploratory and hypothesis-driven approaches, Clinical Psychology Review, Volume 82, 2020, 101940, ISSN 0272-7358, https://doi.org/10.1016/j.cpr.2020.101940

[29] Kilinç, D., Borandağ, E., Yücalar, F., Tunali, V., Şimşek, M., & Özçift, A., "KNN algoritması ve R dili ile metin madenciliği kullanılarak bilimsel makale tasnifi", 2016.

[30] HongFang Zhou, JiaWei Zhang, YueQing Zhou, XiaoJie Guo, YiMing Ma, A feature selection algorithm of decision tree based on feature weight, Expert Systems with Applications, Volume 164, 2021, 113842, ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2020.113842.

[31] Qinghua Tao, Zhen Li, Jun Xu, Na Xie, Shuning Wang, Johan A.K. Suykens, Learning with continuous piecewise linear decision trees, Expert Systems with Applications, 2020, 114214, ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2020.114214.

[32] Guoquan Li, Linxi Yang, Zhiyou Wu, Changzhi Wu, D.C. programming for sparse proximal support vector machines, Information Sciences, Volume 547, 2021, Pages 187-201, ISSN 0020-0255, https://doi.org/10.1016/j.ins.2020.08.038.

[33] Jiale Huang, Tao Jin, Menglin Liang, Houlei Chen, Prediction of heat exchanger performance in cryogenic oscillating flow conditions by support vector machine, Applied Thermal Engineering, Volume 182, 2021, 116053, ISSN 1359-4311, https://doi.org/10.1016/j.applthermaleng.2020.116053.

[34] Ingo Gühring, Mones Raslan, Approximation rates for neural networks with encodable weights in smoothness spaces, Neural Networks, 2020, ISSN 0893-6080, https://doi.org/10.1016/j.neunet.2020.11.010.

[35] Sara Kaviani, Insoo Sohn, Study of scale-free structures in feed-forward neural networks against backdoor attacks, ICT Express, 2020, ISSN 2405-9595, https://doi.org/10.1016/j.icte.2020.11.004.

[36] Zhuo Long, Xiaofei Zhang, Li Zhang, Guojun Qin, Shoudao Huang, Dianyi Song, Haidong Shao, Gongping Wu, Motor Fault Diagnosis Using Attention Mechanism and Improved AdaBoost Driven by Multi-sensor Information, Measurement, 2020, 108718, ISSN 0263-2241, https://doi.org/10.1016/j.measurement.2020.108718.

[37] Amin Shahraki, Mahmoud Abbasi, Øystein Haugen, Boosting algorithms for network intrusion detection: A comparative evaluation of Real AdaBoost, Gentle AdaBoost and Modest

AdaBoost, Engineering Applications of Artificial Intelligence, Volume 94, 2020, 103770, ISSN 0952-1976, https://doi.org/10.1016/j.engappai.2020.103770.

[38] Xiwu Zhang, Lei Wang, Yan Su, Visual Place Recognition: A Survey from Deep Learning Perspective, Pattern Recognition, 2020, 107760, ISSN 0031-3203, https://doi.org/10.1016/j.patcog.2020.107760.

[39] Zhuo Zhang, Yan Lei, Xiaoguang Mao, Meng Yan, Ling Xu, Xiaohong Zhang, A study of effectiveness of deep learning in locating real faults, Information and Software Technology, Volume 131, 2021, 106486, ISSN 0950-5849, https://doi.org/10.1016/j.infsof.2020.106486.

[40] Chen Y, Shen C. "Performance analysis of smartphone-sensor behavior for human activity recognition". IEEE Access 5, 3095–3110, 2017.

[41] Campbell A, Choudhury T. "From smart to cognitive phones". IEEE Pervasive Comput. 11 (3), 7–11, 2012.

[42] Clarkson B.P. Life patterns: Structure from wearable sensors (Ph.D. thesis), Massachusetts Institute of Technology, 2002.

[43] Zhang Z, Ji L,Huang Z, Wu J. "Multi-Model Adaptation for Thigh Movement Estimation Using Accelerometers". IET Signal Procesing, 5:709-716, 2011.

[44] Muscillo R, Schmid M, Conforto S, D'Alessio T. "An Adaptive Kalman-Based Bayes Estimation Technique to Classify Locomotor Activities in Young and Elderly Adults Through Accelerometers". Medical Engineering & Physics, 32:849-859, 2010.

[45] Yang M, Zheng H, Wang H, McClean S, Newell D. "iGAIT: An interactive accelerometer based gait analysis system". Computer Methods and Programs in Biomedicine, 108:715-723, 2012.

[46] Chung P.Y.M, Ng G.Y.F. "Comparison Between an Accelerometer and a Three-Dimensional Motion Analysis System for The Detection of Movement". Physiotherapy, 98:256-259, 2012.

[47] Mathie M.J, Celler B.G, Lovell N.H, Coster A.C.F. “Classification of basic Daily movements using a triaxial accelerometer”. Medical & Biological Engineering & Computing, 42:679-687, 2004.

[48] Avci A, Bosch S, Marin-Perianu M, Marin-Perianu R, Havinga P. “Activity recognition using inertial sensing for healthcare, well being and sports applications: A survey”. 23rd International Conference on Architecture of Computing Systems (ARCS), pp. 1-10, VDE 2010.

[49] Yang J.-Y, Wang J.-S, Chen Y.-P. “Using acceleration measurements for activity recognition: An effective learning algorithm for constructing neural classifiers”. Patternrecognitionletters, 29(16), 2213-2220, 2008.

[50] Ravi N, Dandekar N, Mysore P, Littman M.L.“Activity recognition from accelerometer data”. AAAI, 5, 1541-1546, 2005.

[51] De la Vega L.G.M, Raghuraman S, Balasubramanian A, Prabhakaran B. “Exploring unconstrained mobile sensor based human activity recognition”. 3rd International Workshop on Mobile Sensing, 8-11, Philadelphia USA, 2013.

[52] Sağbaş E.A, Ballı S. “Akıllı Telefon Algılayıcıları ve Makine Öğrenmesi Kullanılarak Ulaşım Türü Tespiti”. Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi, Cilt 22, Sayı 5, Sayfalar 376- 383, Ocak 2016.

[53] Wang J, Chen R, Sun X. “Generative models for automatic recognition of human daily activities from a single triaxial accelerometer”. WCCI 2012 IEEE World Congress on Computational Intelligence, 10- 15, June 2012.

[54] Bao L, Intille S.S. “Activity recognition from userannotated acceleration data”. Pervasive Computing, 1-17, 2004.

[55] Aggarwal K.K, Ryoo M.S. “Human activity analysis: A review”. ACM Computing Surveys, 43 (3), 1–16, 2011.

[56] Reyes-Ortiz J.L, Oneto L, Sama A, Parra X, Anguita D. “Transition-Aware Human Activity Recognition Using Smartphones”, Neurocomputing, Volume171 Issue C, Pages 754-767, January 2016.

[57] Tural K, Akdoğan E. "Akıllı Telefonların Algılayıcılarının Verilerini Kullanarak Yapay Sinir Ağları ile İnsan Hareketlerinin Sınıflandırılması". Otomatik Kontrol Ulusal Toplantısı TOK, cilt.1, no.1, ss.479-483, İstanbul, Türkiye, 21-23 Eylül 2017.

[58] Reyes-Ortiz J.L, Oneto L, Ghio A, Sama A, Anguita D, Parra X. "Human Activity Recognition on Smartphones with Awareness of Basic Activities and Postural Transitions". Artificial Neural Networks and Machine Learning, ICANN, Lecture Notes in Computer Science. 8681, 177–184, Springer 2014.

[59] Su X, Tong H, Ji P. "Activity Recognition with Smartphone Sensors". Tsinghua Science and Technology, International Journal of Information Science, vol. 19, no. 3, pp. 235-249, 2014.

[60] Sağbaş E.A, Ballı S. (2015) Akıllı Telefon Sensörlerinin Kullanımı ve Ham Sensör Verilerine Erişim. Akademik Bilişim, 158-164, Eskişehir, Türkiye, 4-6 Şubat 2015.

[61] Gunduz H, Cataltepe Z. "Borsa Istanbul (BIST) daily prediction using financial news and balanced feature selection". Expert Systems with Applications, 42(22), 9001–9011, 2015.

[62] Murata S, Namikawa J, Arie H, Sugano S, Tani J. "Learning to Reproduce Fluctuating Time Series by Inferring Their Time-Dependent Stochastic Properties: Application in Robot Learning via Tutoring". IEEE Transactions on Autonomous Mental Development (JIF: 1.348), Vol. 5, Issue 4, pp. 298–310, 2013.

[63] Özden C, Acı Ç. Analysis of injury traffic accidents with machine learning methods: Adana case. Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi, 24(2), 266-275, 2018.

[64] Mikolov T, Karafiat M, Burget L, Cernocky J, Khudanpur S. "Recurrent neural network based language model". In INTERSPEECH, 1045-1048, 2010.

[65] Feng W, Guan N, Li Y, Zhang X, Luo Z. "Audio visual speech recognition with multimodal recurrent neural networks". pp. 681-688, 2017.

[66] Hochreiter S, Schmidhuber J. "Long-short term memory, Neural computation". 9(8), 1735–1780, 1997.

[67] Greff K, Srivastava R.K, Koutnik J, Steunebrink B.R, Schmidhuber J. “LSTM: A Search Space Odyssey”. IEEE Transactions on Neural Networks and Learning Systems, Vol 28, Issue 10, Pp 2222- 2232, 2017.

[68] UCI veri kümesi websitesi. “Human Activity Recognition Using Smartphones Data Set”. https://archive.ics.uci.edu/ml/datasets/human+ac tivity+recognition+using+smartphones. (03.09.2019)

[69] Python programlama dili websitesi. https://www.python.org/. (03.09.2019).

[70] Anaconda Veri Bilimi Platformu websitesi. https://www.anaconda.com/. (03.09.2019). [71] PyCharm IDE websitesi. https://www.jetbrains.com/pycharm/. (03.09.201 9).

[72] Google Tensorflow API websitesi. https://www.tensorflow.org/. (03.09.2019).

[73] Keras API websitesi. https://keras.io/. (03.09.2019)

[74] Anguita D, Ghio A, Oneto L, Parra X, Reyes-Ortiz J.L. "A Public Domain Dataset for Human Activity Recognition Using Smartphones". ESANN proceedings, European Symposium on Artificial Neural Networks, Computation a Intelligence and Machine Learning. Bruges (Belgium), 24-26 April 2013, i6doc.com public, ISBN 978-2-87419-081-0. http://www.i6doc.com/en/livre/?GCOI=28001100 131010.

[75] Ignatov A. "Real-time human activity recognition from accelerometer data using Convolutional Neural Networks". Applied Soft Computing, Volume 62, 2018, Pages 915-922, ISSN 1568-4946, 2018. https://doi.org/10.1016/j.asoc.2017.09.027.

[76] Zhao Y, Yang R, Chevalier G, Xu X, Zhang Z. “Deep Residual Bidir-LSTM for Human Activity Recognition Using Wearable Sensors”. Mathematical Problems in Engineering, Article ID 7316954, 13 pages, 2018. https://doi.org/10.1155/2018/7316954.

[77] Kolosnjaji B, Eckert C. "Neural network-based userindependent physical activity recognition for mobile devices". In Proceedings of the IDEAL: 16th International Conference, pp. 378-386, 2015.

[78] Ballı A, Sağbaş EA. “Akıllı saat algılayıcıları ile insan hareketlerinin sınıflandırılması”. Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 21(3), 980-990, 2017.

[79] Ayu MA, Ismail SA, Matin AFA, Mantoro T. “A comparison study of classifier algorithms for mobile-phone’s accelerometer based activity recognition”. Procedia Engineering, 41, 224-229, 2012.

[90] Zhao Z, Chen Y, Liu J, Shen Z, Liu M. “Cross-people mobilephone based activity recognition”. IJCAI'11 Proceedings of the Twenty-Second İnternational Joint Conference on Artificial Intelligence, Barcelona, Spain, 16-22 July 2011.

[81] Borazio M, Laerhoven KV. “Using time use with mobile sensor data: a road to practical mobile activity recognition?”. MUM 2013, Luleå, Sweden, 02-05 December 2013.

[82] Sousa PAC, Gaber MM, Krishnaswamy S, Gomes JB, Menasalvas E. “MARS: a personalised mobile activity recognition system”. 14th International Conference on Mobile Data Management, Bengaluru, India, 23-26 July 2012.

[83] Vo QV, Hoang MT, Choi D. “Personalization in mobile activity recognition system using k-medoids clustering algorithm”. International Journal of Distributed Sensor Networks, 9(7), 315841, 2013.

[84] Yan Z, Subbaraju V, Chakraborty D, Misra A, Aberer K. “Energy-efficient continuous activity recognition on mobile phones: an activity-adaptive approach”. 16th International Symposium on Wearable Computers, Newcastle, UK, 18-22 June 2012.

[85] Lee J, Kim J. “Energy-efficient real-time human activity recognition on smart mobile devices”. Mobile Information Systems, 2016, 1-12, 2016.

[86] Acharjee D, Mukherjee A, Mandal JK, Mukherjee N. “Activity recognition system using inbuilt sensors of smart mobile phone and minimizing feature vectors”. Microsystem Technologies, 22(11), 2715–2722, 2016.

[87] Abdallah ZA, Gaber MM, Srinivasan B, Krishnaswamy S. “Adaptive mobile activity recognition system with evolving data streams”. Neurocomputing, 150(PA), 304-317, 2015.

[88] Anguita D, Ghio A, Oneto L, Parra X, Reyes-Ortiz JL. “Human activity recognition on smartphones using a multiclass hardware-friendly support vector machine”. International Workshop on Ambient Assisted Living, Vitoria-Gasteiz, Spain, 3-5 December 2012.

[89] Sztyler T, Stuckenschmidt H, Petrich W. “Position-aware activity recognition with wearable devices”. Pervasive and Mobile Computing, 38(2), 281-295, 2017.

[90] Kemilly Dearo Garcia, Cláudio Rebelo de Sá, Mannes Poel, Tiago Carvalho, João Mendes-Moreira, João M.P. Cardoso, André C.P.L.F. de Carvalho, Joost N. Kok, An ensemble of autonomous auto-encoders for human activity recognition, Neurocomputing, Volume 439, 2021, Pages 271-280, ISSN 0925-2312, https://doi.org/10.1016/j.neucom.2020.01.125.

[91] Aswin K Ramasubramanian, Syed M. Aiman, Nikolaos Papakostas, On using human activity recognition sensors to improve the performance of collaborative mobile manipulators: Review and outlook, Procedia CIRP, Volume 97, 2021, Pages 211-216, ISSN 2212-8271, https://doi.org/10.1016/j.procir.2020.05.227.