KASDI MERBAH OUARGLA UNIVERSITY

Faculty of New Information and Communication Technologies

Department of Computing and Information Technology

Professional Master Thesis

Field: Mathematics and Computer Science

Branch: Computer Science

Specialty: Network Administration and Security

Presented by: Rechache Saher

Theme

Face Attributes Classification Based on Deep Neural Network

In front of the Jury

Supervisor:	BENKADDOUR Mohammed Kamel	UKM OUAGLA
President:	AMIRAT.H	UKM OUARGLA
Examiner:	BELHADJ.M	UKM OUARGLA

2021/2022

Acknowledgment

First, I would like to thank "ALLAH" our creator for giving me strength to accomplish this work.

I would like to express my special thanks of gratitude to my teacher (MR BENKADDOUR Mohammed Kamel) for accepting to supervise me, and for his incontestable instructions, patience, skills, and remarks that I have benefited and I was able to discover the world of scientific research in the field of artificial intelligence and biometric techniques. I would also like to thank the members of the jury for their precious time devoted to study my work.

My thanks and gratitude to our teachers who throughout the years of study have passed on their knowledge to us.

A special feeling of gratitude to my loving parents, my family my friends for their support and assistance through the provision of the comfort conditions.

Table of Contents

Table of Figures

Tables

LIST OF ACRONYMS

2D: 2 Dimensions

3D: three Dimensions

AAM: Active Appearance Models

AI: Artificial Intelligence

ANN: Artificial Neural Network

CNN: Convolutional Neural Network

DL: Deep Learning

DNN: Deap Neural Networkes

FAR: False Acceptance Rate

FMR: False Match Rate

FNMR: False Non-Match Rate

FRVT: Face Recognition Vendor Test

GPU: Graphical Processor Unit

KPCA: Kernel Principle Component Analysis)

MDR: medical image recognition,

ML: Machin Learning

PCA: principal component analysis

QDA: Quadratic Discriminant Analysis

SVM: Support Vector Machine

GENERAL INTRODUCTION

GENERAL INTRODUCTION

Biometrics is a system for recognizing people based on one or several physiological or behavioral traits. This kind of data is used to identify the user and assign an access ID. Biometric access control systems are convenient for users because information carriers are always with them and cannot be lost or fabricated. They are considered to be reliable systems because even in case of a data breach, biometric credentials like fingerprints or a face scan are almost impossible to use for unsanctioned access, unlike username and password combinations.

Face recognition has been one of the active research problems studied in computer vision for decades due to its many practical applications, for instance, in the automotive industry, security, retail, beautification, and social networks there have been many developments in this field. Due to the COVID-19 pandemic, which has caused many people to wear face masks to prevent infection, it has become urgent to meet the challenge of detecting faces wearing masks. the study assisting not only in detecting faces wearing masks but also in predicting gender and age estimation, In particular, gender and age are facial characteristics that can be very useful for a multitude of applications, for example an automatic gender and age prediction system is used to profile customers who are interested in a product or for target advertising.

Mask detection, Gender prediction and age estimation are tasks that humans perform naturally and effortlessly in their daily lives. Endowing a computer with this reasoning ability is still a great challenge.

In order to give computers the feature of face recognition and the automatic classification of objects, several methods have been proposed in recent years. Many CNN models have been created in order to give highly resolution and perfect accuracy. In our work, we opted for our model convolutional neural networks (CNN) based on deep learning to make model for age estimation and gender prediction and face mask detection withe the same architecture and deferent image training part

The objective of this work is to study and analyze the architecture of CNN to implement an application capable of predicting gender and estimating age and detect face Mask in real time from facial images. To do this, we have structured our thesis in four chapters:

We started by first chapter that presents Biometrics and its types with how it helps the world in the present time, it also contains facial recognition that we depend on it in computer vision, under that we mentioned a bread lines of age estimation, gender prediction and face mask detection.

In chapter2 which is the state of art, it was about approach of pattern recognition and how we recognize the patterns from data specifically from images

Continuing with chapter 3 Techniques used in our work that explains the most important part that we basing on it in the project, it defines the Artificial intelligent and Machin learning and deep learning and the relationship between them and after that we presented neural networks and specified convolutional neural networks

We sealed with the chapter4 we will explain our experimental results obtained by age estimation and gender prediction and face mask detection classification systems followed by a discussion with interpretation of the results.

CHAPTER1 BIOMETRICES

1.1. Introduction

Faced with document fraud and identity theft, new threats such as terrorism or cybercrime and faced with the logical evolution of international regulations, new technological solutions are gradually being implemented. Among these technologies, biometrics quickly stood out as the most relevant for identifying and authenticating people, reliably and quickly, based on unique biological characteristics.

We will introduce in this chapter some basic definitions related to biometrics. Then we will give the motivations and the objectives of this theme which mainly concerns the facial modality to estimate the age and the gender of a person.

1.2. Biometrics

1.2.1. What Are biometrics?

Biometrics is the measurement and statistical analysis of people's unique physical and behavioral characteristics. The technology is mainly used for identification and access control or for identifying individuals who are under surveillance. The basic premise of biometric authentication is that every person can be accurately identified by intrinsic physical or behavioral traits. The term *biometrics* is derived from the Greek word’s *bio*, meaning *life*, and *metric*, meaning *to measure* [1] .

1.2.2. Why use biometrics?

Biometric authentication and its uses in modern-day tech and digital applications has a number of advantages [2]:

- High security and assurance: Biometric identification provides the answers to “something a person has and is” and helps verify identity.
- User experience: convenient and fast.
- Non transferrable: Everyone has access to a unique set of biometrics.
- Spoof proof: Biometrics are hard to fake or steal.

High security and assurance

Biometrics provide increased levels of assurance to providers that a person is real by verifying a tangible, real-world trait as both something the user has and something the user is. Most user’s passwords and PINs and personal identifying information have likely been compromised with a data breach, meaning, billions of accounts can be accessed by fraudsters who retain the answers to traditional authentication methods. Introducing biometric authentication into the process adds in a road-block for fraudsters that only a real, authorized user can circumnavigate - though a fraudster may know a person uses their dog's name and some lucky numbers for most of their online accounts, they can't use their fingerprint to unlock an account [2] if they can't provide it on the spot. Additionally, biometrics can only be provided by living, breathing people - at this point in time, a robot would have a hard-time passing an iris scan.

User experience is convenient and fast

While the internal processes for biometric authentication is technical, from a user's point of view it's incredibly easy and quick. Placing a finger on a scanner and unlocking an account in seconds is faster than typing out a long password that has multiple special characters. In addition, forgetting a password is a common mistake of most users. The chances of you forgetting your own biometrics? [2] Never!

Non transferable

Biometric authentication requires its input is present upon authorization. [2] You can't transfer or share a physical biometric digitally – the only way to utilize most biometric authentication systems is with a physical application.

Near spoof proof

Biometrics like face patterns, fingerprints, iris scanning, and others are near-impossible to replicate with current technology. There's a one in 64 billion chance that your fingerprint will match up exactly with someone else's. Said a different way, [2] you have a better chance winning the lottery than having the same fingerprint as a hacker trying to get into your account that's secured by biometrics.

1.2.3. Types of biometrics

The two main types of biometric identifiers are either physiological characteristics or behavioral characteristics.

Biometric recognition is the individual's presentation of his unique biometric parameter and the process of comparing it with the entire database of available data. Biometric readers are used to retrieving this kind of personal data.Facial recognition [3]

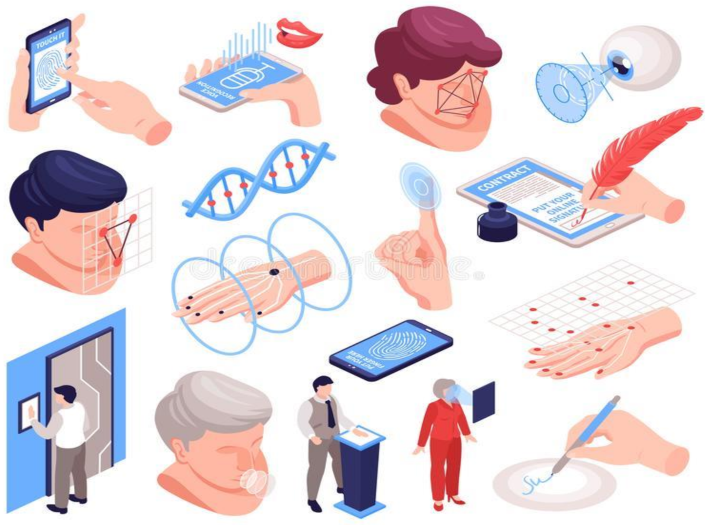

Figure 1_ 1 : Types of biometrics

1. **Physical identification methods:** are based on the analysis of the invariable physiological characteristics of a person.

These characteristics include:

- **Face shape and geometry** (technologies for recognizing two-dimensional images of faces drawn from photographs and video sequences work with these identifier). Thanks to the growth of multimedia technologies, can see more and more video cameras installed on city streets and squares, [4] airports, train stations, and other crowded places, determining this direction's development.
- **Fingerprints** (the most widespread, convenient, and effective biometric technology is built on the use of these identifiers). The advantages of fingerprint access are ease of use, convenience, and reliability. [4] Although the false identification rate is about 3%, the unauthorized access probability is less than 0.00001% (1 in 1,000,000).
- **The shape and structure of the skull** (for greater euphony, companies operating in this area prefer to talk about human recognition technologies based on a three-dimensional face model).
- **Retina** (rarely used as an identifier). When identifying by the retina, the angular distribution of blood vessels on the retina's surface relative to the blind spot of the eye and other features are measured. The retina's capillary pattern differs even in twins and can be used with excellent personality identification success [4].
- **The iris of the eye** (patent restrictions constrain the spread of the technology in which this identifier is used). The advantage of iris scanners is that they do not

require the user to focus on the target because the iris pattern is on the eye's surface. The eye's video image can be scanned at a distance of less than 1m.

- **Palm, hand, or finger geometry** (used in several narrow market segments).
- **Facial thermography, hand thermography** (technologies based on the use of these identifiers have not become widespread).
- **Drawing of veins on the palm or finger** (the corresponding technology is becoming popular, but due to the high cost of scanners, it is not yet widely used) [4].
- **DNA** (mainly in the field of specialized expertise).

2. **Behavioral identification methods:** are based on the analysis of a person's behavioral characteristics. The characteristics inherent in each person in behavioral methods of user identification are divided by:

- **Signature recognition** (for identification, simply the degree of coincidence of the two pictures is used). By painting and dynamic characteristics of writing (for identification, a convolution is built, which includes information on painting, time characteristics of painting application, and statistical characteristics of the dynamics of pressure on the surface) [3].
- **Keystroke dynamics.** The method is generally similar to that described above. Instead of a signature, a certain codeword is used (when the user's password is used for this, such authentication is called two-factor). The dynamics of the code word's set is the main characteristic used to construct the convolution for identification.
- **Speaker recognition.** It is one of the oldest biometric technologies. Its development has intensified, a great future, and widespread use in constructing «intelligent buildings» are predicted. There are many ways to construct a voice identification code; as a rule, these are various combinations of the voice's frequency and statistical characteristics [3].
- **Gait recognition.** This should be categorized as exotic. It seems that this direction is a dead-end due to the poor repeatability of the feature and its weak identification.

Figure 1_ 2: Recognition types

In theory, we say that a physiological or behavioral criterion can be used to recognize individuals if it satisfies the following conditions:

- **Universality**: Every individual accessing the application should possess the characteristics. As an example, we can't use the iris characteristics to identify blind persons, [3] as we can't use signature in an environment where most of the population don't write.
- **Uniqueness**: The underlying characteristics should be sufficiently different across individuals to be able to distinguish between two persons.
- **Permanence**: The biometric characteristics should be resistant to changing in time at least with respect to the operating recognition system period. A trait that changes significantly over time is not a useful biometric.
- **Measurability**: The biometric characteristics must be quantitatively measurable to be further processed by a machine. Suitable devices connected to the machine can be used to acquire and digitize the biometric trait to be transferred later to the recognition system [5].

in practice,there are other important conditions:

- **Performance**: The application that uses the biometric characteristics must ensure an acceptable degree of performance. This includes the matching accuracy/time as well as the resources devoted to build the overall recognition system.
- **Acceptability**: this indicates how much people that are intended to be identified using these characteristics are willing to cooperate with the system by presenting their biometric [3].
- **Circumvention**: It measures the robustness of the system; i.e., how much is easy to fool the system to make it taking wrong decision or to compromise information about the user's biometric data [5].

1.2.4. Advantages and disadvantages of biometrics

The use of biometrics has plenty of advantages and disadvantages regarding its use, security and other related functions.

Advantages

- hard to fake or steal, unlike passwords
- easy and convenaient to use
- generally, the same over the course of a user's life
- non-transferable
- efficient because templates take up less storage.

Disadvantages

- It is costly to get a biometric system up and running.
- If the system fails to capture all of the biometric data, it can lead to failure in identifying a user.
- Databases holding biometric data can still be hacked.
- Errors such as false rejects and false accepts can still happen.
- If a user gets injured, then a biometric authentication system may not work -- for example, if a user burns their hand, then a fingerprint scanner may not be able to identify them.

1.2.5. How do biometrics Works

A biometric system is basically a pattern recognition system that uses an individual's biometric data. There are three modules in a biometric system, the verification, identification and verification module.

While biometric systems can combine authentication, verification, and identification, there are some key differences between those three facets. Namely, identification asks, “who are you?” while verification asks, “Is there data associated with you?” Authentication asks, “Are you who you say you are?

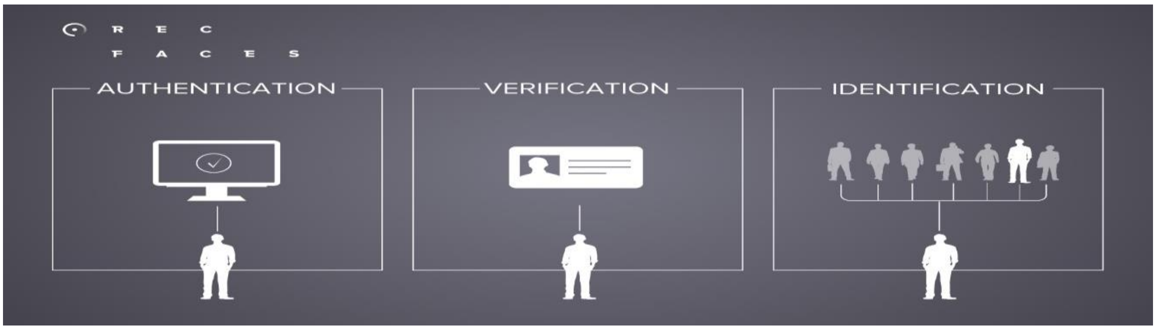

Figure 1_ 3 : Modules of a biometric system

- **Biometric authentication**

Biometric authentication’s aim is to verify that you are who you are supposed to be. With such systems, a computer will scan a person for inherent attributes – for instance, a face recognition template, and will then compare the individual’s characteristics to a template stored within a database. If the scanned attributes match the template, the person is allowed into the system [4].

The process of biometric authentication looks like this:

1/ Enrollment. A reference sample is collected from an individual – perhaps a photo, a writing sample, a retina scan, or a fingerprint. The biometrics sample created by specialized algorithms from these data is called a template, and it is either stored in a database or on a card, or it is managed by an authenticating authority.

2/ Live Sample. Now that the template is in place, the user provides a live sample as part of the authentication process. For instance, they may insert a card containing face recognition data in a machine and then make a photo of the user [4].

3/ Comparison. In order to complete the authentication process, Step 2's live sample is compared to the template. If there is a verified match, the user is authenticated and may access the system.

- **Biometric identification**

Biometric identification can be applied to digital and physical scenarios, and it's a solution that is used in defense, law enforcement, and border control. With identification, there is a database that contains physical characteristics of a vast number of people for instance, the FBI's repository stores the height, hair color, weight, eye color, scars, and tattoos of over 70,000,000 criminal records. With authentication, a person's features are compared to one specific template or With identification, however, the person's features are matched against the entire database [4].

- **Biometric verification**

Biometric verification is often confused with authorization – however, there is a subtle difference between the two processes. While authentication indicates that a person has the same biometric features as somebody who is already in the system, verification can conclusively prove that their online identity is linked with their real-life identity [4].

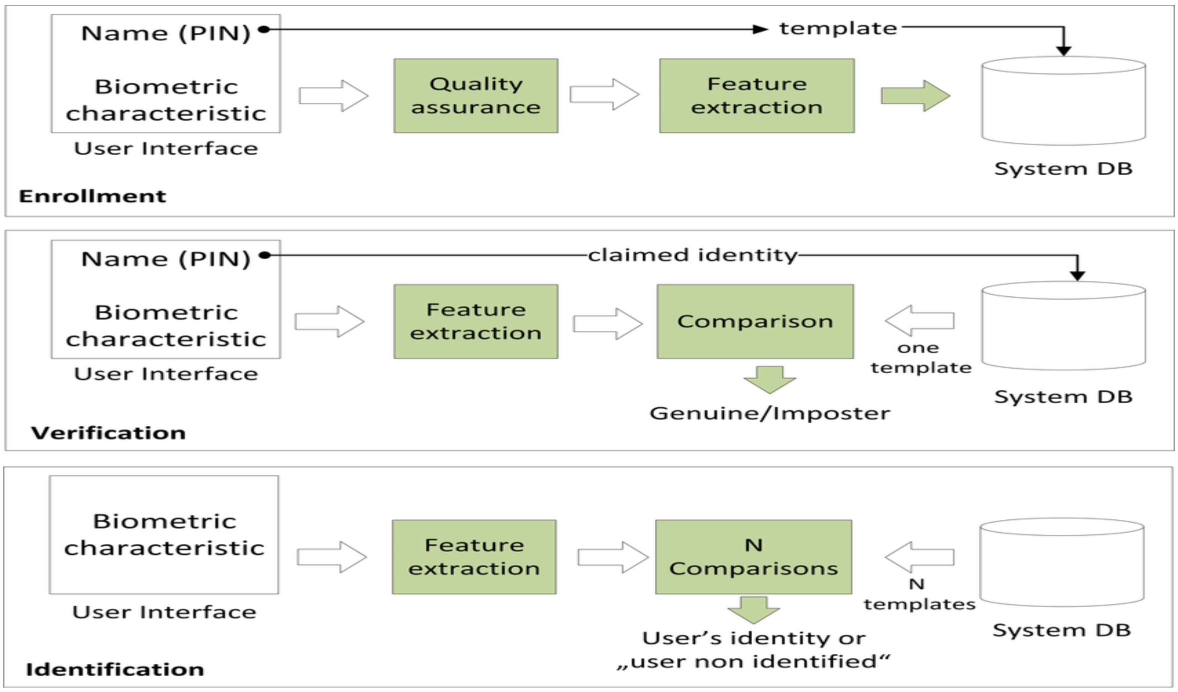

Figure 1_4: the main operations of a biometric system

1.2.6. biometric system architecture

Biometric system architecture has the following main components:

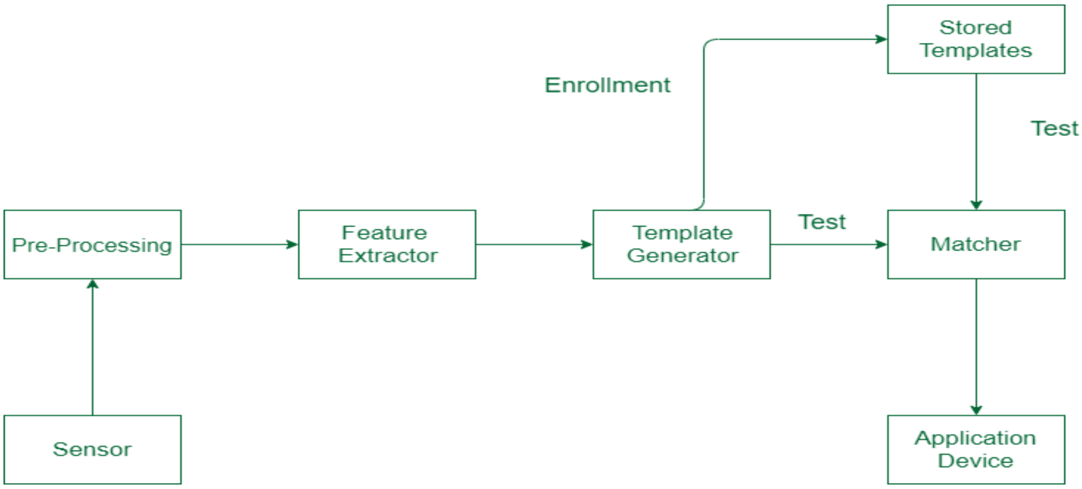

Figure 1_ 5 : Biometric Architecture

1. **sensor:** The sensor is the first block of the biometric system which collects all the important data for biometrics. It is the interface between the system and the real world. Typically, it is an image acquisition system, but it depends on the features or characteristics required that it has to be replaced or not [6].
2. **Pre-processing:** It is the second block that executes all the pre-processing. Its function is to enhance the input and to eliminate artifacts from the sensor, background noise, etc. It performs some kind of normalization.
3. **Feature extractor:** This is the third and the most important step in the biometric system. Extraction of features is to be done to identify them at a later stage. The goal of a feature extractor is to characterize an object to be recognized by measurements.
4. **Template generator:** The template generator generates the templates that are used for authentication with the help of the extracted features. A template is a vector of numbers or an image with distinct tracts. Characteristics obtained from the source groups come together to form a template. Templates are being stored in the database for comparison and serve as input for the match [6].
5. **Matcher:** The matching phase is performed by the use of a match. In this part, the procured template is given to a matcher that compares it with the stored templates using various algorithms such as Hamming distance, etc. After matching the inputs, the results will be generated.

6. **Application device:** It is a device that uses the results of a biometric system. The Iris recognition system and facial recognition system are some common examples of application devices [6].

1.2.7. Applications of biometric systems

Biometric systems can be used in a large number of applications. For security reasons, biometrics can help make transactions, and everyday life is both safer and more practical. The following domains use biometric solutions to meet their respective needs:

- **Legal applications:** Biometric technology and law enforcement have a very long history, and many very important innovations in identity management have emerged from this beneficial relationship. Today, the biometrics applied by the police force is truly multimodal. Fingerprint, face, and voice recognitions play a unique role in improving public safety and keeping track of the people we are looking for [3].
- **Government applications:** A key area of application for biometric technology is at the border. Biometric technology helps to automate the process of border crossing. Reliable and automated passenger screening initiatives and automated SAS help to facilitate international passenger travel experience while improving the efficiency of government agencies and keeping borders safer than ever before. In the field of healthcare, biometrics introduces an enhanced model. Medical records are among the most valuable personal documents; doctors need to be able to access them quickly, and they need to be accurate. [3] A lack of security and good accounting can make the difference between timely and accurate diagnosis and health fraud.
- **Commercial applications:** As connectivity continues to spread around the world, it is clear that old security methods are simply not strong enough to protect what is most important. Fortunately, biometric technology is more accessible than ever, ready to provide added security and convenience for everything that needs to be protected, from a car door to the phone's PIN [3].

1.2.8. Biometric System Performance

Biometric system manufacturers claim high system performance which is practically difficult to achieve in actual operating environments. The possible reasons are, tests conducted in controlled environment setups, limitations on hardware, etc.

For example, a voice recognition system can work efficiently only in quiet environment, a facial recognition system can work fine if lighting conditions are controlled, and candidates can be trained to clean and place their fingers properly on the fingerprint scanners [2].

However, in practice, such ideal conditions may not be available in the target operating environment.

Performance Measurements

The performance measurements of a biometric system are closely tied to False Reject Rate (FRR) and False Accept Rate (FAR) [2].

- **FRR** is also known as **Type-I error** or False Non-Match Rate (FNMR) which states the likelihood of a legitimate user being rejected by the system.
- **FAR** is referred to as **Type-II error** or False Match Rate (FMR) which states the likelihood of a false identity claim being accepted by the system.
- An ideal biometric system is expected to produce zero value for both FAR and FRR. Means it should accept all genuine users and reject all fake identity claims, which is practically not achievable [2].
- **FAR** and **FRR** are inversely proportional to each other. If FAR is improved, then the FRR declines. A biometric system providing **high FRR ensures high security**. If the FRR is too high, then the system requires to enter the live sample a number of times, which makes it less efficient.
- The performance of current biometrics technologies is far from the ideal. Hence the system developers need to keep a good balance between these two factors depending on the security requirements [2].

1.3. Face Recognition

Facial recognition has turned into a popular tool to authenticate the identity of an individual. In modern times, this technology has been used in various sectors and industries to prevent ID fraud and identity theft. If you want to know more about facial recognition and how you can implement it into your business, read it [7].

While the concept of facial recognition is not new, technological improvement has led to a massive proliferation of this technology. Today, you can see its usage everywhere – in passport offices to detect fraud in passports and travel visas, in airports to screen passengers; in police stations to scan the criminal data; in ATMs/banks to ensure the security of users; in organizations to keep track on the attendance of the employees, and more.

In addition to this, some web applications, social networking websites, and smartphones also use this technology to prevent data misuse. Probably, your Smartphone also has a face recognition feature to unlock it [8].

Regardless of where this technology has been used, its role is to verify a human identity and prevent fraud.

1.3.1. What is facial recognition

To put it simply is a biometric technology that identifies and verifies an individual from a digital image and video. It captures, examines, and then compares a person's facial details.

1.3.2. Challenges in face recognition

As the range of application is expanding day by day, the complexity of the system is increasing as well. This in fact affects the efficiency of the system. In this section of the paper, we shall discuss the different challenges of face recognition systems that are present today. These challenges are related to the face image which is given as the input to the system. The algorithms used or this process varies from application to application. There are many reasons that are responsible for variation in faces. These sources of variation are classified into two main factors [8]. They are:

1. Intrinsic factors: It is due to the physical nature of the face and not dependent on the observer. Intrinsic factors are further divided into intrapersonal and interpersonal. Intrapersonal is caused due to variation in face appearance of an individual, for example ageing, facial expression and facial paraphernalia (facial hair, cosmetics, glasses etc.)

2. Extrinsic factors: This is caused due to the variation in face appearance due to the interaction of light with the face and the observer. This will include illumination, pose, scale and imaging parameters (resolution, focus, imaging, noise etc.) [9].

Following are the common challenges seen in face recognition system can have while detecting a face.

- **Pose variation**

  Variation in pose causes significant problems in detecting a face. Pose variation can be due to change in observing angle of the observer and also due to rotation in the head position. These variations can cause a serious problem in identifying the input image. Many of the systems can tolerate small variations such as small rotations in angles. But it will be difficult when it comes to large rotational angles. The database usually consists of face images of frontal view of the faces. Since the existing FRSs are very sensitive to pose variation, pose correction is essential and could be achieved by means of efficient techniques aiming to rotate the face and/or to align it to the image's axis [9].

Figure 1_ 6: Variation in poses

- **Variation in illumination**

Variations of illuminations could reduce the efficiency of FRS. For moderate levels of lighting of the background, face detection and recognition are much difficult to perform. Variation in illumination can vary the total magnitude of light intensity being reflected back from an object. On the other hand, higher light levels could lead to over-exposure of the face and (partially) undetectable facial patterns. There have been many algorithms such as equalization techniques that are available now to get rid of this problem to an extent. Sometimes even multiple algorithms can be used in a face recognition system to tolerate the issue of illumination. But in case of extents, it is not desirable to depend on these techniques [9].

Figure 1_ 7 : Variation in Illumination

- **Variation in expression**

Some variation in the face images can be caused due to difference in expression influenced by the individual’s state of emotion. Therefore, it is important to recognize different facial expressions for evaluating the emotional state. Human expressions consist of macro-expressions such as, disgust, anger, happiness, fear, sadness or surprise, and other involuntary, rapid facial patterns. These facial changes can be computed with the help of dense optical flow. Cosmetics and hair styles can also be included in this challenge as changing hair style and putting make-up can also cause variation in facial expression [8].

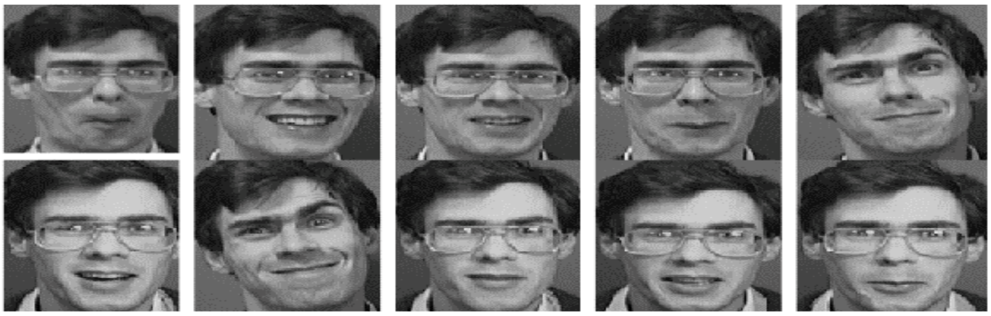

Figure 1_ 8: Variation in Expression

- **Aging**

Another reason for the changes in the appearance of the face could be the aging of the human face and could affect the entire process of face recognition; if the time between each image capture is large, there will be significant changes in the person. As per various study conducted by scientists, in every 10 years there will be significant changes in an individual’s face appearance. The fig. 5 shows the change in an individual’s face at different ages. It is not just the shape and lines of a face that gets modified over time; there will be changes in hairstyles as well.

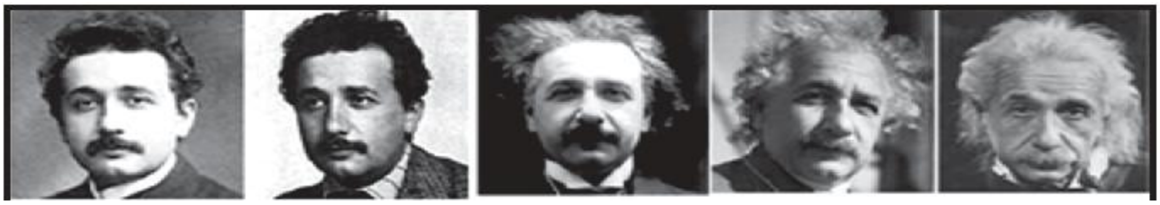

Figure 1_ 9 : Aging

- **Occlusions**

Variation in facial appearance can also be caused due to presence of objects that such as occlusion that partially cover the face. This makes it a difficult task for the system to classify the image. Although the face is found, it may be difficult to recognize it due to some hidden facial parts, making it difficult to recognize features. This challenge can be seen in real world application were acquiring persons talking on the phone, wearing glasses, scarf, hats etc. or having their faces covered with hands [9].

Figure 1_ 10: Occlusions

- **Similar Faces**

This is usually a not so common challenge. But we have seen that even humans find it difficult to identify people with similar faces. Hence, we can imagine the difficult situation for computer to identify similar face individuals. Especially identical twins with similar facial features, shape etc. this becomes a difficult task for the face recognition system to identify the individual. This will cause an increase in false recognition rate (FRR) as well [8].

Figure 1_ 11 : Similar Faces

- **Image Resolution**

Another important issue with face recognition system is the varying quality and resolution of the images given as input. Many factors can affect the resolution of an image. The environment, the performance quality of the acquiring system and many other reasons can be mentioned as factors that are responsible for varying resolution of the image. If the resolution is good, then the recognition process will be much easier and efficient. So, we can say that resolution is directly proportional to the efficiency of the face recognition system [8].

Figure 1_ 12: Variation in resolution

1.3.3. Advantages and disadvantages of facial recognition

Facial recognition technology is a controversial topic, and it is likely to be used in many more ways than we can imagine. Here are some pros and cons of facial recognition [10].

Advantages

Facial recognition is a technology that can benefit society, including increasing safety and security, preventing crimes, and reducing human interaction. [10] Here are some pros of facial recognition:

- Helps find missing people
- Protects businesses against theft
- Improves medical treatment
- Strengthens security measures
- Makes shopping more efficient
- Reduces the number of touchpoints
- Improves photo organization

disadvantages

Facial recognition is an innovative technology that has the power to change our future. However, like any innovation, some consequences and risks are involved when implementing this new system in society [10].

- Threatens privacy
- Imposes on personal freedom
- Violates personal rights
- Data vulnerabilities
- Misuse causing fraud and other crimes
- Technology is still new
- Errors can implicate innocent people
- Technology can be manipulated

1.3.4. How Facial Recognition Works

Any facial recognition algorithm uses biometrics to map out facial features captured in a video still or a photograph. That information is then compared to a database of faces. There are four general steps in the process, which we'll explain further [10].

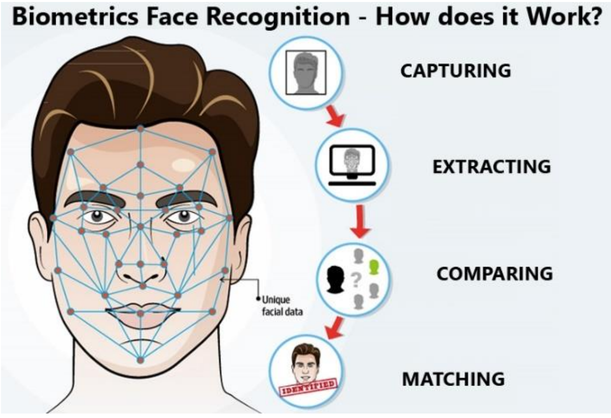

Figure 1_ 13 : Biometrics Face Recognition – How does it Work?

Step 1: Face detection

First, a camera will detect and recognize a human's face – one that can either be in a crowd or alone. It is most easily detected when the person is looking straight at the camera. However, modern technological advances allow face recognition software to still work if the person's. is angled slightly [8].

Step 2: Face analysis

After detection and recognition, a photo will capture the face and will then be analyzed. The majority of face recognition technology use 2D images instead of 3D. This is because 2D photos are more readily correlated with public photos or pictures in a database (these are typically 2D as well). During analysis, the face will be separated into distinguishable landmarks – we can call these nodal points. A human face has eight nodal points. Face recognition technology will analyze each of these points – for example, the distance between your eyebrows [7].

Step 3: Converting an image into data

After analysis, each nodal point becomes a number in the application database. The entire numerical code is referred to as a faceprint. Just like how everybody has a unique thumbprint, everyone also has a unique faceprint.

Step 4: Matching

The final step of the process is finding a match. Your faceprint is compared to a database of other facial codes. The number of faces that are compared depends on the database and how many databases the software has access to. For instance, the FBA has access to 21 state databases, with 641 million photos across them. The facial recognition technology then identifies a match for your exact facial features – it returns the user with the found match and other relevant information, such as an address and a name [7].

1.4. Gender Recognition and Age Estimation

Interest in automatically predicting people's gender and age from face image has grown steadily in recent years. Automatic gender and age classification are also required whenever it is necessary to identify the gender and age of an audience, for example to analyze the effectiveness of advertising. Such prediction can also be useful for creating human-computer interfaces where system behavior is tailored to a specific user based on a number of factors, including gender and age. In this part, we provide an overview of existing techniques and works on gender and age prediction from face images.

1.4.1. Age Estimation

the Adience dataset this method achieved 50.7% exact accuracy and 84.7% one-off accuracy. This method also addresses the problem of gender classification where they achieved accuracy of 86.8%. Many real world applications such as embedded systems require a more efficient neural network architectures. Authors Yang et al. introduce a network architecture called Soft Stagewise Regression Network (SSR-Net). The method classifies age in stages where each stage refines the classification made by previous stage. The best result this method achieved was mean absolute error of 3.16 using model trained on IMDB-WIKI dataset and evaluated on MORPH 2 dataset [11].

1.4.2. Gender Prediction

The problem of gender prediction and classification is closely related to face detection. One of the first method for gender classification was SEXNET introduced by Golomb et al. It classified images sampled at 30 x 30 pixels using a fully connected neural network with 3 layers of 900, 40 and 900 neurons. Another method introduced by Moghaddam and Yang uses Support Vector Machines to classify gender. The method used subsampled images of 21 x 21 pixels and classifed them using SV M with radial basis function kernel. More recently the gender classification problem has been treated as a subproblem of age prediction. Current approaches use the same neural network architecture for age as well as gender prediction. Gender can be easily modelled as a binary classification problem. If the gender is modelled using regression we can view the output number as confidence [11].

1.5. Face Mask Detection

Paravision’s biometric facial recognition has been found to have the second-best accuracy for identifying people with masks among algorithms tested by the U.S. National Institute of Standards and Technology (NIST) in a Face Recognition Vendor Test (FRVT) with error rates below three percent when the subject’s nose and mouth are covered.

All the tested algorithms were developed before the COVID-19 pandemic made mask use ubiquitous around the world [12].

Deep Glint finished first across all nine facial-occlusion categories, but Paravision took second in each, and was within 0.5 percent of the lead. The leaderboard is dominated by Chinese and Russian companies, and Paravision ranked first among those from the U.S., UK, Europe and Japan by what it calls “a substantial margin.”

False non-match rates (FNMR) for coverings worn slightly below the nose with Paravision’s algorithm were 1.24 percent to 1.35 percent, at a false match rate (FMR) of 0.0001 percent. With the subject’s nose covered, FNMR ranged from 1.81 percent to 3.27 percent at the same FMR. Light blue full-width face masks resulted in an FNMR of 2.81 percent [12].

NIST tested 89 algorithms for the effectiveness at identifying people wearing masks, and found that the best ones had error rates of between 5 and 50 percent.

Since the pandemic outbreak, Paravision has also developed algorithms for mask and social distancing detection.

“So many aspects of NIST FRVT provide value and insight to Paravision and the rest of the industry,” states Paravision Chief Product Officer Joey Pritikin. “While we’re always thrilled to perform well, evaluations that assess performance in the toughest situations like heavy angles or with face masks—are the most exciting to us, as they support what we consistently find in the field: We shine in challenging, real-world conditions [13].”

Paravision says it has been ranked number three in 1:N identification in previous NIST FRVT results, and is the top provider from the U.S., UK and Europe for both 1:1 verification and 1:N matches [13].

1.5.1. Face Mask detection application

Airports Hospitals Offices

Figure 1_ 14: .Face Mask detection application

1.6.conclusion

In this chapter, we have presented the different technologies used in biometric systems for the identification of people, then we have described the mode of operation of a biometric system, as well as the biometric technologies with their advantages and disadvantages. This study allowed us to see that face recognition is attracting more and more interest from the scientific community, because it presents several challenges and technological locks. Finally, we have highlighted the various difficulties inherent in the automatic recognition of faces, which allowed us to clearly define the issues addressed in this thesis, in particular the invariance to illumination.

CHAPTER2 STATE OF ART

APPROACHES USED IN PATTER RECOGNITIO

2.1. Introduction

In this chapter, we briefly described some of the most important or commonly used techniques in the field of pattern recognition, starting with the definition of pattern recognition, followed by it's features, and architecture, next we presented the four basic models followed in PR, after that we mentioned the algorithms that are consisted all in pattern recognition witch depends on supervised and unsupervised learning then we presented several types of classification algorithms, at the end we mentioned clustering algorithms that helps in gaining overall distribution of patterns.

2.2. What is Pattern Recognition?

The practice of distinguishing the patterns using Artificial Intelligence and Machine Learning tools with algorithms. The classification and identification methods of data supported by the prior statistical knowledge about the retrieved patterns as inputs. Applications such as facial expression recognition, speech recognition, MDR, medical image recognition, etc., are a part of PR systems. Two types of techniques are very prominent in the AI and ML dictionary. These are Classification and clustering without which the definition is incomplete [14].

The raw data which acts as input is fed to be processed, and then it is converted into machine-understandable codes. Then these codes undergo the training process, which includes these steps of Classification and Clustering. Although the terms are synonymous, yet they are different in terms of their practical application. The supervised learning and unsupervised learning govern these techniques [14].

2.2.1. Features of Pattern Recognition

- They recognize the patterns accurately with greater precision.
- They are able to recognize and classify even the unknown and unfamiliar entities.
- The accuracy can be relied upon, and the shapes and objects can be explored from different angles.
- The identification is eased by various unknown patterns in the world of science.
- They are able to detect the minute differences and recover the original pattern in case of missing data.
- They help to study and research the unknown domains of science in medical fields [15], data sciences, etc.

2.2.2. How Patter Recognition Works?

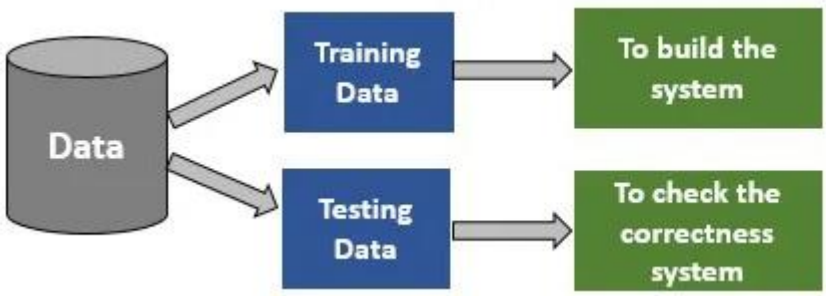

Figure 2_ 1 : How pattern recognition works?

The working of pattern recognition depends on the various notions of supervised and unsupervised learning approaches. This whole phase cycle reveals the working of the Pattern recognition approach [15].

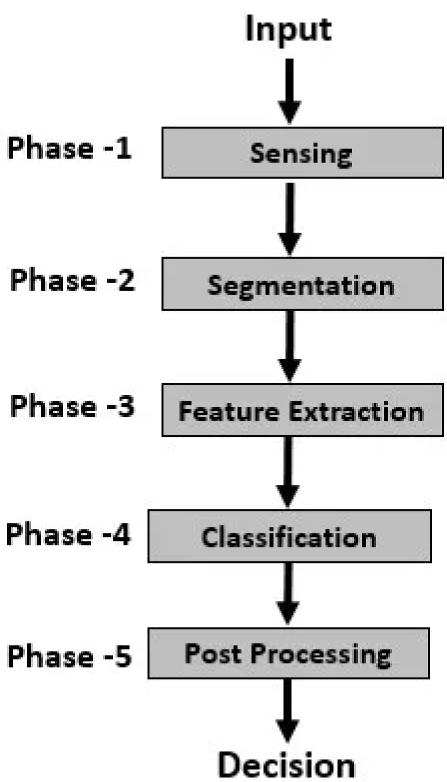

Figure 2_ 2 : Pattern recognition phases

The different phases are as follows:

Phase 1: This phase translates inputs to the analogous signal data.

Phase 2: This phase helps to isolate the sensed input data and eliminate the noise.

Phase 3: This measures the entities and the objects on its properties and sends the, for further procedure of classification [16].

Phase 4: It assigns the sensed object to category.

Phase 5: It takes other consideration to decide for appropriate action.

2.3. Pattern Recognition Models

There are four basic models followed in pattern recognition

- Statistical Techniques
- Structural Techniques
- Template Matching
- Neural Network Approach

2.3.1. Statistical Algorithm Model

In this model, the pattern is termed in the form of features. These Features are selected in a way that different patterns take space without overlapping. It is able to predict and recognize the probabilistic nature. It works so nicely that the selected features are helping the formation of clusters. It analyses the probability distribution, decision boundaries, etc., for the patterns. The machine learns and adapts accordingly. Then these patterns are projected to further processing, training. Then we apply testing patterns for recognition of patterns. This leads to further classification methods. The various schemes used in it are Baye's Decision Rule, PCA, etc [17].

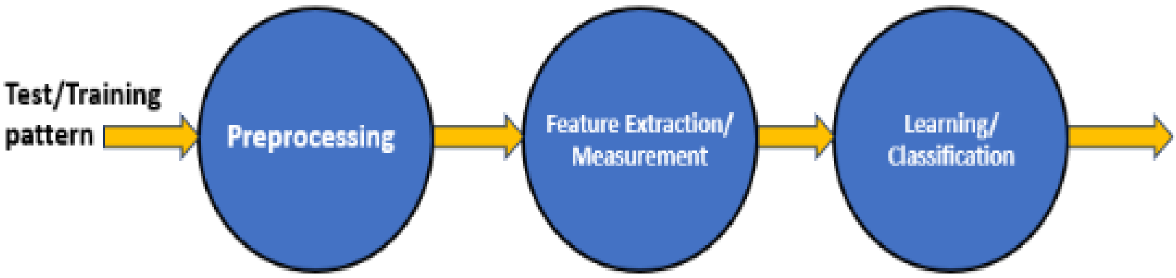

Figure 2_ 3 : Statistical Algorithm Model

2.3.2. Structural Algorithm Model

These models are also named as structural models for pattern recognition and are based on the relation between features. Here the patterns are represented by structures which can take into account more complex relations between features unlike the numerical feature sets used in statistical pattern recognition models [18]. Also the patterns used in this model forms a hierarchical structure composed of sub-patterns. In this model, the patterns to be recognized are called primitives and the complex patterns are represented by the inter-relationship formed between these primitives and the grammatical rules associated with this relationship In syntactic pattern recognition, a similarity is associated between the structure of patterns and the syntax of a language. The patterns are the sentences belonging to a language, primitives are the alphabet of the language, and using these primitives, the sentences are generated according to the grammar. Thus, the very complex patterns can be described by a small number of primitives and grammatical rules [19] . This approach is considered to be an appealing model in pattern recognition because, in addition to classification, it also provides a description of how from the primitives the given pattern is constructed due to its hierarchical structure. This paradigm has been used in situations where the patterns have a definite structure which can be captured in terms of a set of rules . The implementation of a syntactic model approach, however, leads to many difficulties because of the segmentation of noisy patterns (to detect the primitives) and the inference of the grammar from training data. This may yield a combinatorial explosion of possibilities to be investigated, demanding a very large training sets and huge amount of computational efforts [20] .

2.3.3. Template Matching Algorithm Model

The model of Template matching is simplest. It is most primitive of all the models. The model is used to determine similarity among two images. The pattern matched is being stored in templates, and the templates are given flexibility for scalar and rotational changes. The competence of this model relies upon the already stored templates in the database. We take correlation function to be the function of recognition in this case, and later it is optimized according to the availability of the training set. [21] The only problem with this model is that this approach is not as efficient while working in distorted patterns.

2.3.4. Neural Network Based Algorithm Model

Neural networks can be viewed as a parallel computing systems consisting of an extremely large number of simple processors with many interconnections between them. Typically, a neural network or to be more specific [16], an artificial neural network (ANN) is a self-adaptive trainable process that is able to learn and resolve complex problems based on available knowledge. An ANN-based system behaves in the same manner as how the biological brain works; it is composed of interconnected processing elements that simulate neurons. Using this interconnection, each neuron can pass information to another. Artificial Neural network models attempt to use some organizational principles such as learning, generalization, adaptivity, fault tolerance and distributed representation, and computation in the network of weighted directed graphs in which the artificial neurons forms the nodes of the model and the directed edges (with weights) are connections between neuron outputs and neuron inputs. [19] The weights applied to the connections results from the learning process and indicate the importance of the contribution of the preceding neuron in the information being passed to the following neuron . [20] The main characteristics of all the neural networks are that they possess the ability to learn complex nonlinear input-output relationships, use sequential training procedures, and adapt themselves to the data. The following diagram is a two layer neural network with one input layer constituting of three neurons and one output layer with two neurons and corresponding weights are assigned in between them [17].

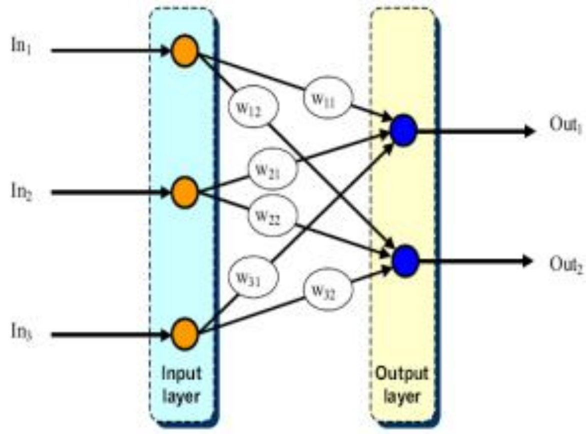

Figure 2_ 4 : cnn weights

2.4. Pattern Recognition Algorithms

The field of pattern recognition has been explored widely by a number of researchers who as a result have developed various algorithms. The design pattern of all these algorithms consists of three basic elements, i.e., [20] data perception, feature extraction and classification. There are various different techniques to implement these three basic elements. So which technique is chosen for each element in design cycle defines the algorithm characteristic of the pattern recognition algorithm.

This is the design cycle of a basic pattern recognition model. Algorithms for pattern recognition depend on the type of label output, on whether learning is supervised or unsupervised.

2.4.1. Supervised learning

Supervised learning is a process of allotting a function to some desired category as learnt from supervised training data. Here the training data consist of a set of training examples where each set consist of a pair consisting of an input object and a desired output value.

A supervised learning algorithm learns from this training pair relationship and produces an inferred function [19].

In simple terms, in supervised learning, there is a teacher who provides a category label or cost for each pattern in the training set which is used as a classifier. So basically a supervised learning method is used for classification purpose.

In the given figure, the input image consist of a mixture of two alphabets, i.e., A and B. Then the classification algorithm classifies the input to two different categories.

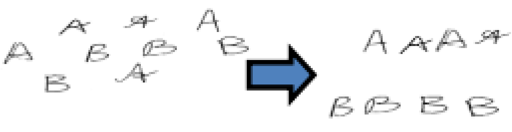

Figure 2_ 5 : Supervised learning

Here a set of combined input is classified using supervised learning approach.

2.4.2. Unsupervised learning

Un-supervised learning can be defined as the problem of trying to find out the hidden structure in an unlabeled data set. Since the examples given to the learner are unlabeled, each algorithm itself classifies the test set . [16] In simple terms, here no labeled training sets are provided and the system applies a specified clustering or grouping to the unlabeled datasets based on some similarity criteria. So an unsupervised learning method is used for clustering. Here the input consists of some unlabeled values whose distinguishing feature is initially not known. [9] The following input consists of such a combination with all values technically same but still its clusters are formed using some metric which is different for each algorithm.

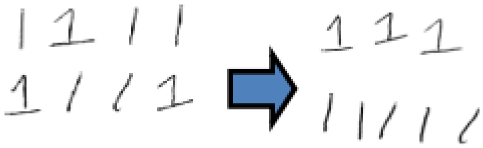

Figure 2_ 6 : Un-supervised learning

Here clusters are formed in the output

2.5. Classification algorithms

2.5.1. Linear discriminant analysis (LDA)

It is used to find a linear combination of features which characterizes or separates two or more classes of objects or events. LDA is a parametric approach in supervised learning technique. It was initially used for dimensionality reduction and feature extraction, and later moved for classification purpose also. [8]LDA easily handles the cases where the within-class frequencies are unequal and their performances had been examined on randomly generated test data. Thus it maximizes the ratio of between-class variance to the within-class variance in any particular data set thereby guaranteeing maximal separability. Linear discriminant analysis has a close relation with Principal Component Analysis (PCA). Both methods are used for dimensionality reduction. LDA have been proven better algorithm when compared with PCA. [11] The prime difference between LDA and PCA is that PCA does more of feature classification and LDA does data classification. In working with PCA, the location of the original data set changes when transformed to a totally different space whereas LDA doesn't change the location but only tries to provide more class separability and draw a decision region between the given classes [10].

2.5.2. Quadratic Discriminant Analysis (QDA)

It is used in machine learning and statistical classification to separate measurements of two or more classes of objects or events by a quadric surface [8]. It is a more general version of a linear classifier. QDA is a parametric approach in supervised learning which models the likelihood of each class as a Gaussian distribution, then uses the posterior distributions to estimate the class for a given test point. The Gaussian parameters for every class can be estimated from training points with maximum likelihood (ML) estimation. The simple Gaussian model assumption is best suited to cases when one does not have much information to characterize a class, i.e., if there are too few training samples to infer much about the class distributions. Also, when the number of training samples is small compared to the number of dimensions of each training sample [14], the ML covariance estimation can be ill-posed. There exist some solutions to resolve this ill-posed estimation; one is to regularize the covariance estimation and another is to use Bayesian estimation.

2.5.3. Maximum entropy classifier (multinomial logistic regression)

Maximum entropy classifier (multinomial logistic regression) In statistics, a maximum entropy classifier model is a regression model which generalizes logistic regression by allowing more than two discrete outcomes. This forms a model that is used to predict the probabilities of the different possible outcomes of a categorically distributed dependent variable, given a set of independent variables (which may be real-valued, categorical-valued etc.) [21]. The actual goal of the multinomial logistic regression model is to predict the categorical data. Maximum entropy classifiers are commonly used as an alternative to Naive Bayes classifiers because they do not require statistical independence of the independent variables (commonly known as features) that serve as the predictors. This algorithm may not be appropriate to learn large number of classes since it is slower than for a Naive Bayes classifier. Multinomial logistic regression is a particular solution to the classification problem that assumes that a linear combination of the observed features and some problem-specific parameters can be used to determine the probability of each particular outcome and the best values of the such parameters for a given problem are usually determined from some training data [11].

2.5.4. Decision trees

It is considered to be a decision support tool that uses a tree-like structure or model of decisions and all its possible consequences. It is one way to display an algorithm.

These trees are basically used in operations research, mostly in decision analysis, to help identify a strategy most likely to reach a goal [13].

In this process, a decision tree and the closely related influence diagram is used as a visual and analytical decision support tool where the expected values of competing alternatives are calculated. Decision trees are a simple, but very powerful form of multiple variable analysis.

The trees provide unique capabilities which act to be supplement, complement, and substitute for

- Traditional statistical forms of analysis (such as multiple linear regressions)

• A lot of data mining tools and techniques (such as neural networks)

• The recently developed multidimensional forms of reporting and analysis found in the field of business intelligence [15] The decision trees are produced by algorithms which identify various ways of splitting the data set into branch-like segments. These segments form an inverted decision tree which starts with a root node at the top of the tree. Each node starting from root contains the name of field which is also called object of analysis. The decision rule is discovered based on a method that extracts the relationship between the object of analysis (that serves as the target field in the data) and one or more fields that serve as input fields to create the branches or segments [9]. The values of the input field are used to estimate the likely value of the target field which can also be termed as an outcome, response, or dependent field or variable. Once the relationship is found, then one or more decision rules can be derived which describe the relationships between inputs and targets [20]. Then these decision rules can be used to predict the values of new or unseen observations which contain values for the inputs, but might not contain values for the targets.

2.5.5. Kernel Estimation & K-Nearest Neighbors

In the field of pattern recognition, the k-nearest neighbor algorithm (k-NN) is a method for classifying objects based on closest training examples in the feature space. K-NN is a type of example-based learning, or lazy learning where the function is only approximated locally and all the computation is deferred until classification. This algorithm is one of the simplest machine learning algorithms in which an object is classified using a majority vote of its neighbors and the object is then assigned to the class which is most common amongst its k-nearest neighbors [15]. Here the neighbors are taken from a set of objects for which the correct classification is known. These neighbors can be assumed as a training set for this algorithm, though no explicit training step is required [19]. The learning in this model is based on storing all the training instances which corresponds to points in an n-dimensional Euclidean space along with their class labels and classification is delayed till a new instance is arrived. As the new unlabeled query instance or vector arrives, the classification is performed by assigning the label which is most frequent among the k-training samples nearest to that query point [8].

There are some variations that can be performed on this algorithm. These variations start with 1-NN where k=1 and the object are simply assigned to the class of its nearest neighbor. Then we have k-NN approach where the value of k is chosen randomly. Here we find k closest training points to the test instance according to some metric (mostly used metric is the Euclidean distance) and then perform classification operation. The best choice of k generally depends on the data itself However larger value of k reduces the effect of noise on classification but makes the boundaries between classes less distinct. So, a good choice of k is required which can be achieved by some heuristic technique called cross-validation [20].

2.5.6. Naive Bayes classifier

Naive Bayes classifier is a simple, probabilistic and statistical classifier which is based on Bayes theorem (from Bayesian statistics) with strong (naive) independence assumptions and maximum posteriori hypothesis. As Bayesian classifiers are statistical in nature [22], they can predict the probability of a given sample belonging to a particular class. The underlying probability model to this classifier can be termed more appropriately as an "independent feature model" because a naive Bayes classifier assumes that the effect of an attribute value on a given class is independent of the values of the other attributes. Such an assumption is called class conditional independence. It is made to simplify the computation involved and, in this sense, is considered "naive" [14]. We can explain this classifier with a small example. A fruit is considered to be an apple if it is red in color, round in shape, and around 5" in diameter. Although these features depend on each other or upon the existence of the other features, a naive Bayes classifier takes all of these properties to independently contribute to the probability that this fruit is an apple. The naive Bayes classifier is trained using a supervised learning approach that just requires consideration of each attribute in each class separately. So, the training in naive Bayes classifier is considered to be very easy and fast. To estimate the parameters in naive Bayes model it uses the principle of maximum likelihood method in many practical applications [16]. Testing in this algorithm is also very straightforward and simple; just look the tables and calculate conditional probabilities with normal distributions. The advantage of Naive Bayes model is that it only requires a small amount of training data to estimate the parameters, i.e., means and variances of the variables which are necessary for classification.

2.5.7. Artificial Neural Networks

It is an interconnected network of a group of artificial neurons. An artificial neuron can be considered as a computational model which is inspired by the natural neurons present in human brain. Unlike natural neurons [19], the complexity is highly abstracted when modeling artificial neurons. These neurons basically consist of inputs (like synapses), which are further multiplied by a parameter known as weights (strength of each signal), and then computed by a mathematical function which determines the activation of the neuron. After this there is another function that computes the output of the artificial neuron (sometimes in dependence of a certain threshold) [16]. Thus, the artificial networks are formed by combining these artificial neurons to process information.

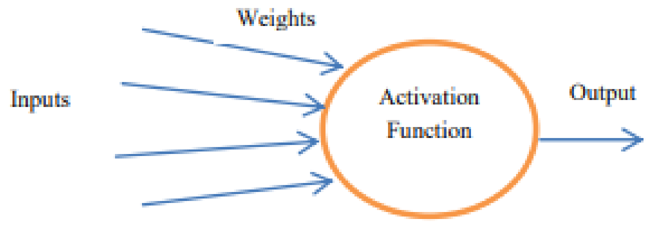

Figure 2_ 7 : artificial neurons

We can train ANN for best matched solution; ANN can perform fuzzy matching and provides the optimal solution. It also acts as a classifier in pattern recognition. It falls under the category of supervised learning where the model initially learns from the training data set and then classifies the test image using the learnt knowledge [15].

2.5.8. Support Vector Machine

A Support Vector Machine (SVM) performs classification by constructing an N-dimensional hyperplane that optimally separates the data into two categories. A support vector machine (SVM) is used in computer science for a set of related supervised learning methods that analyze input data and learns from it and then use it for performing classification and regression analysis [10]. The standard SVM is a two-class SVM which takes a set of input data and predicts the possible class, for each input, among the two possible classes the input is a member of, which makes it a non-probabilistic binary linear classifier. Given the set of training examples where each one of them is marked as belonging to one of the two categories, the SVM training algorithm builds a model that assigns new examples into one category or the other. SVM is an efficient method of finding an optimal hyperplane for separating non-linear data also [9]. Presently, the traditional two-class SVM is also used in multiclass classification where the data to be classified may belong to any one class among a number of classes.

2.6. Clustering algorithms

2.6.1. Hierarchical Clustering

It is a process used in data mining concept where it can be defined as a method of cluster analysis which works to build a hierarchy of clusters [16]. It is a widely used data analysis tool. The idea behind hierarchical clustering is to build a binary tree of the data that successively merges similar groups of points and visualizing this tree provides a useful summary of the data. Hierarchical clustering strategies generally fall into two types:

- Agglomerative: This is a "bottom up" approach of hierarchical clustering where each observation starts with one single cluster, and then pairs of clusters are merged as one move up the hierarchy [23]. In Agglomerative clustering, each level of the resulting tree is a segmentation of the data. Hence the algorithm results in a sequence of grouping and then it is up to the user to choose a natural clustering from this sequence.
- Agglomerative: This is a "bottom up" approach of hierarchical clustering where each observation starts with one single cluster, and then pairs of clusters are merged as one move up the hierarchy [23]. In Agglomerative clustering, each level of the resulting tree is a segmentation of the data. Hence the algorithm results in a sequence of grouping and then it is up to the user to choose a natural clustering from this sequence.

2.6.2. K-means Clustering

As a process employed in data mining, k-means clustering is defined as a method of cluster analysis which aims to partition n different observations into k different clusters in which each observation belongs to the cluster with the nearest mean. Although this problem is computationally very difficult and has been put under the NP hard problem set, there are efficient heuristic algorithms that are commonly employed and converge quickly to a local optimum. Such algorithms are similar to the expectation-maximization algorithm for mixtures of Gaussian distributions via an iterative refinement approach employed by both algorithms [8].

2.6.3. KPCA (Kernel Principle Component Analysis)

It is an extension of principal component analysis (PCA) or may also be termed as a nonlinear form of PCA. [19] Using this form of PCA one can efficiently compute principal components in a very high dimensional feature spaces related to input space by some non-linear mapping using techniques of kernel methods and functions. Particularly for clustering [21], KPCA can be used to construct a hyperplane that divides the „n" points into arbitrary clusters by making them almost always linearly separable in d >=n dimensions. Also, this nonlinear kernel PCA can be used for simple pattern recognition with a linear classifier with much better recognition rates in comparison to simple PCA.

Along with this, the computational complexity of KPCA does not grow with the dimensionality of the feature space it is working on.

2.7. Conclusion

It has always been difficult to decide which algorithm is best to classify patterns with least computational effort, least time and maximum and best results. In this chapter, pattern recognition has been defined and followed by basic models, various categories of algorithms are discussed. Pattern recognition field has a wide range of applications in the field of classification, clustering, regression, sequence labeling and parsing among which this chapter reviews the algorithms of the most applied field on pattern recognition, i.e., classification and clustering.

CHAPTER3 DEEP NEURAL NETWORKES AND CNN

3.1. Introduction

Deep learning is a subfield of machine learning, and machine learning is the subfield of artificial intelligence. The backbone of deep learning algorithm is composed of neural networks. The structure is illustrated in Fig. 1.

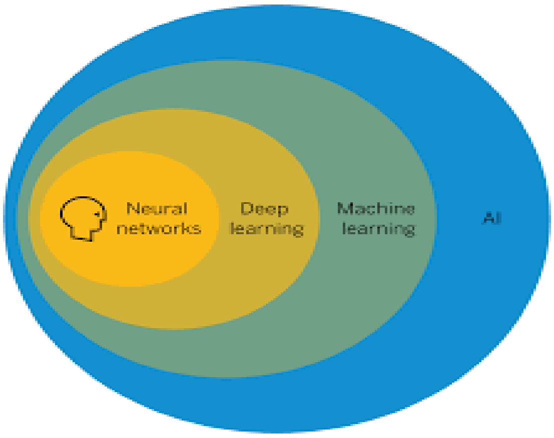

Figure 3_ 1 : Artificial intelligence subfield

Fig. 1. The illustration of the relationship between neural network, deep learning, machine learning and artificial intelligence.

In this chapter, we present a detailed study on CNN neural networks and its different components. This chapter aims to provide an introduction to the concept of convolutional neural networks, to do this it is necessary to understand the concept of the artificial neural network, where part of the chapter is dedicated to it.

3.2. Artificial Intelligence (AI)

3.2.1. What is Artificial Intelligence ?

Artificial intelligence is the simulation of human intelligence processes by machines, especially computer systems. Specific applications of AI include expert systems, natural language processing, speech recognition and machine vision [24].

3.2.2. How Does AI Work?

In general, AI systems work by ingesting large amounts of labeled training data, analyzing the data for correlations and patterns, and using these patterns to make predictions about future states. In this way, a chatbot that is fed examples of text chats can learn to produce lifelike exchanges with people, or an image recognition tool can learn to identify and describe objects in images by reviewing millions of examples [25].

AI programming focuses on three cognitive skills: learning and self-correction.

- **Learning processes.** This aspect of AI programming focuses on acquiring data and creating rules for how to turn the data into actionable information. The rules, which are called algorithms, provide computing devices with step-by-step instructions for how to complete a specific task [26].
- **Reasoning processes.** This aspect of AI programming focuses on choosing the right algorithm to reach a desired outcome.
- **Self-correction processes.** This aspect of AI programming is designed to continually fine-tune algorithms and ensure they provide the most accurate results possible.

3.2.3. Advantages and Disadvantages of Artificial Intelligence

Artificial neural networks and deep learning artificial intelligence technologies are quickly evolving, primarily because AI processes large amounts of data much faster and makes predictions more accurately than humanly possible.

While the huge volume of data being created on a daily basis would bury a human researcher, AI applications that use machine learning can take that data and quickly turn it into actionable information. As of this writing, the primary disadvantage of using AI is that it is expensive to process the large amounts of data that AI programming requires [27].

Advantages

- Good at detail-oriented jobs
- Reduced time for data-heavy tasks
- Delivers consistent results
- AI-powered virtual agents are always available

Disadvantages

- Expensive
- Requires deep technical expertise
- Limited supply of qualified workers to build AI tools
- Only knows what it's been shown
- Lack of ability to generalize from one task to another

3.2.4. What are the 4 types of Artificial Intelligence?

Arend Hintze, an assistant professor of integrative biology and computer science and engineering at Michigan State University, explained in a 2016 article that AI can be categorized into four types, beginning with the task-specific intelligent systems in wide use today and progressing to sentient systems [26], which do not yet exist. The categories are as follows:

Type 1: Reactive machines. These AI systems have no memory and are task specific. An example is Deep Blue, the IBM chess program that beat Garry Kasparov in the 1990s. Deep Blue can identify pieces on the chessboard and make predictions, but because it has no memory, it cannot use past experiences to inform future ones.

Type 2: Limited memory. These AI systems have memory, so they can use past experiences to inform future decisions. Some of the decision-making functions in self-driving cars are designed this way [14].

Type 3: Theory of mind. Theory of mind is a psychology term. When applied to AI, it means that the system would have the social intelligence to understand emotions. This type of AI will be able to infer human intentions and predict behavior, a necessary skill for AI systems to become integral members of human teams.

Type 4: Self-awareness. In this category, AI systems have a sense of self, which gives them consciousness. Machines with self-awareness understand their own current state. This type of AI does not yet exist.

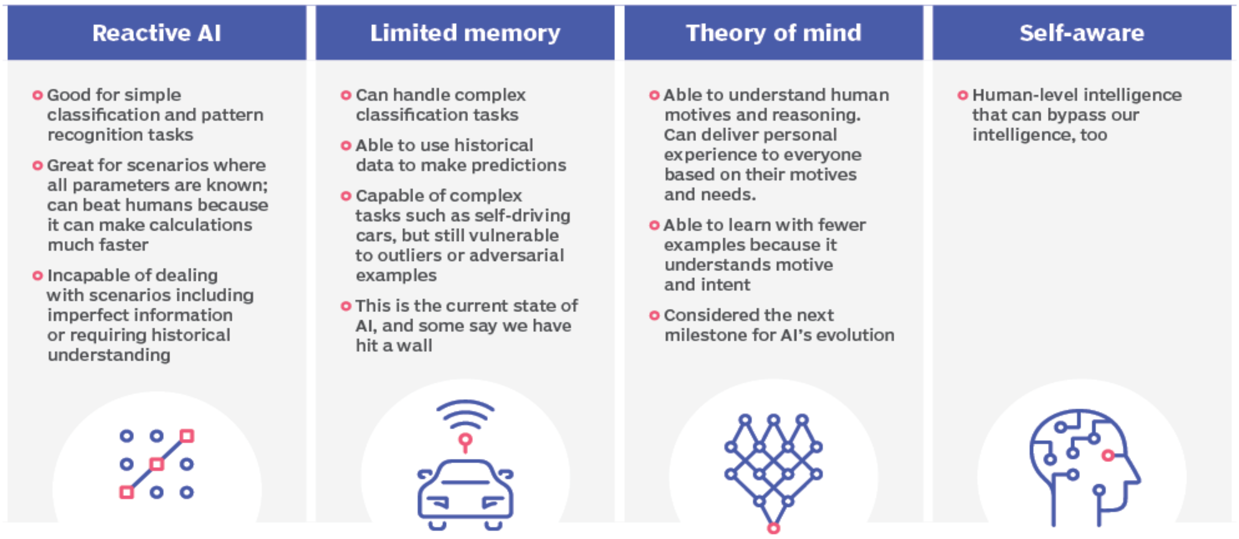

Figure 3_ 2 4 : types of Artificial Intelligence

3.2.5. What are Examples of AI?

AI is incorporated into a variety of different types of technology.

- **Automation.** When paired with AI technologies, automation tools can expand the volume and types of tasks performed. An example is robotic process automation (RPA), a type of software that automates repetitive, rules-based data processing tasks traditionally done by humans. When combined with machine learning and emerging AI tools, RPA can automate bigger portions of enterprise jobs, enabling RPA's tactical bots to pass along intelligence from AI and respond to process changes [25].
- **Machine learning.** This is the science of getting a computer to act without programming. Deep learning is a subset of machine learning that, in very simple terms, can be thought of as the automation of predictive analytics. There are three types of machine learning algorithms :
    - **Supervised learning.** Data sets are labeled so that patterns can be detected and used to label new data sets.
    - **Unsupervised learning.** Data sets aren't labeled and are sorted according to similarities or differences.
    - **Reinforcement learning.** Data sets aren't labeled but, after performing an action or several actions, the AI system is given feedback.
- **Machine vision.** This technology gives a machine the ability to see. Machine vision captures and analyzes visual information using a camera, analog-to-digital conversion and digital signal processing. It is often compared to human eyesight, but machine vision isn't bound by biology and can be programmed to see through walls, for example. It is used in a range of applications from signature identification to medical image analysis. Computer vision [27], which is focused on machine-based image processing, is often conflated with machine vision.
- **Natural language processing (NLP).** This is the processing of human language by a computer program. One of the older and best-known examples of NLP is spam detection, which looks at the subject line and text of an email and decides if it's junk. Current approaches to NLP are based on machine learning. NLP tasks include text translation, sentiment analysis and speech recognition.
- **Robotics.** This field of engineering focuses on the design and manufacturing of robots. Robots are often used to perform tasks that are difficult for humans to perform or perform consistently. For example, robots are used in assembly lines for car

production or by NASA to move large objects in space. Researchers are also using machine learning to build robots that can interact in social settings [26].

- **Self-driving cars.** Autonomous vehicles use a combination of computer vision, image recognition and deep learning to build automated skill at piloting a vehicle while staying in a given lane and avoiding unexpected obstructions, such as pedestrians.

3.3. Machine Learning

machine learning is a core sub-area of Artificial Intelligence (AI). ML applications learn from experience (or to be accurate, data) like humans do without direct programming. When exposed to new data, these applications learn, grow, change, and develop by themselves. In other words, machine learning involves computers finding insightful information without being told where to look. Instead, they do this by leveraging algorithms that learn from data in an iterative process.

The concept of machine learning has been around for a long time (think of the World War II Enigma Machine, for example). However, the idea of automating the application of complex mathematical calculations to big data has only been around for several years, though it's now gaining more momentum.

At a high level, machine learning is the ability to adapt to new data independently and through iterations. Applications learn from previous computations and transactions and use "pattern recognition" to produce reliable and informed results.

3.3.1. How Does Machine Learning Work?

Machine Learning is, undoubtedly, one of the most exciting subsets of Artificial Intelligence. It completes the task of learning from data with specific inputs to the machine. It's important to understand what makes Machine Learning work and, thus, how it can be used in the future [28].

The Machine Learning process starts with inputting training data into the selected algorithm. Training data being known or unknown data to develop the final Machine Learning algorithm. The type of training data input does impact the algorithm, and that concept will be covered further momentarily [29].

New input data is fed into the machine learning algorithm to test whether the algorithm works correctly. The prediction and results are then checked against each other.

If the prediction and results don't match, the algorithm is re-trained multiple times until the data scientist gets the desired outcome. This enables the machine learning algorithm to continually learn on its own and produce the optimal answer, gradually increasing in accuracy over time.

The process flow depicted here represents how Machine Learning works:

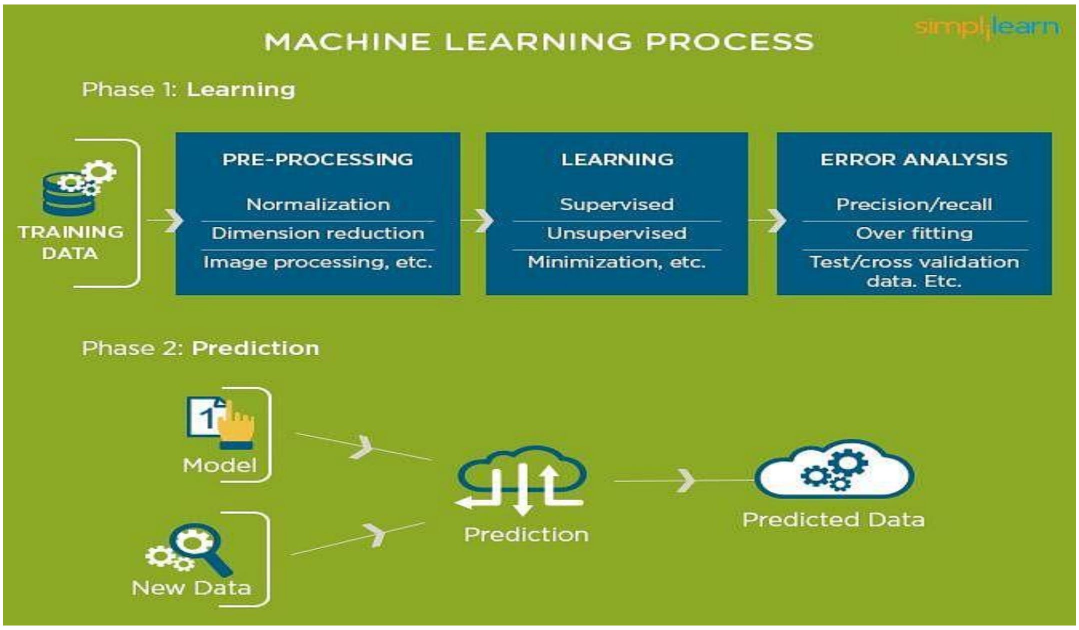

Figure 3_ 3 : how Machine Learning works

3.3.2. Different Types of Machine Learning

Machine Learning is complex, which is why it has been divided into two primary areas, supervised learning and unsupervised learning. Each one has a specific purpose and action, yielding results and utilizing various forms of data. Approximately 70 percent of machine learning is supervised learning, while unsupervised learning accounts for anywhere from 10 to 20 percent. The remainder is taken up by reinforcement learning [30].

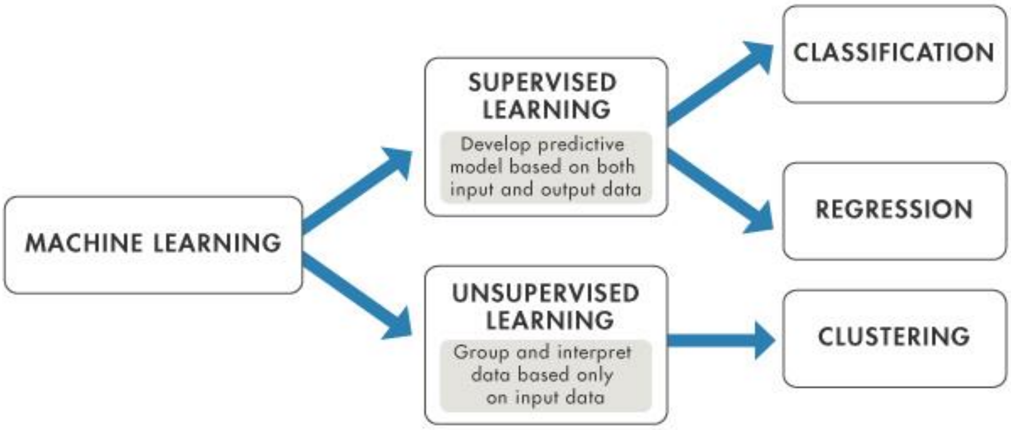

Figure 3_ 4 : Different Types of Machine Learning

3.3.2.1. Supervised learning

In supervised learning, we use known or labeled data for the training data. Since the data is known, the learning is, therefore, supervised, i.e., directed into successful execution. The input data goes through the Machine Learning algorithm and is used to train the model. Once the model is trained ba [3]sed on the known data, you can use unknown data into the model and get a new response.

- **Model:** Most machine learning involves transforming the data in some sense. We might want to build a system that ingests photos and predicts smiley-ness. Alternatively, we might want to ingest a set of sensor readings and predict how normal vs. anomalous the readings are. By model, we denote the computational machinery for ingesting data of one type, and spitting out predictions of a possibly different type. In particular, we are interested in statistical models that can be estimated from data. While simple models are perfectly capable of addressing appropriately simple problems, the problems that we focus on in this book stretch the limits of classical methods [25]. Deep learning is differentiated from classical approaches principally by the set of powerful models that it focuses on. These models consist of many successive transformations of the data that are chained together top to bottom, thus the name deep learning. On our way to discussing deep models, we will also discuss some more traditional methods.

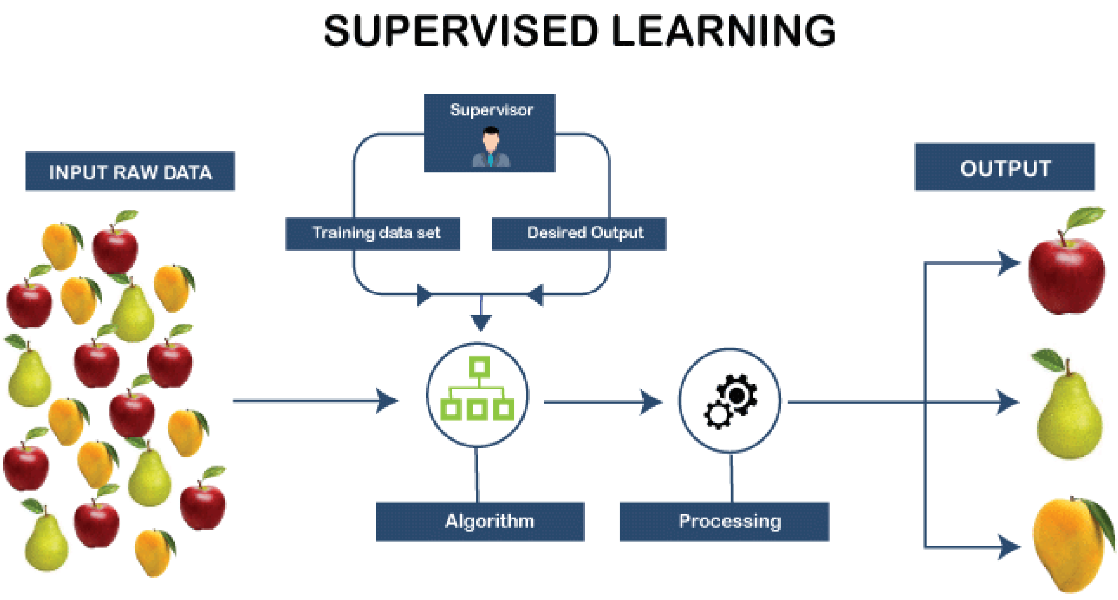

Figure 3_ 5 : Supervised learning

3.3.2.2. Unsupervised Learning

In unsupervised learning, the training data is unknown and unlabeled - meaning that no one has looked at the data before. Without the aspect of known data, the input cannot be guided to the algorithm, which is where the unsupervised term originates from. This data is fed to the Machine Learning algorithm and is used to train the model. The trained model tries to search for a pattern and give the desired response [29]. In this case, it is often like the algorithm is trying to break code like the Enigma machine but without the human mind directly involved but rather a machine.

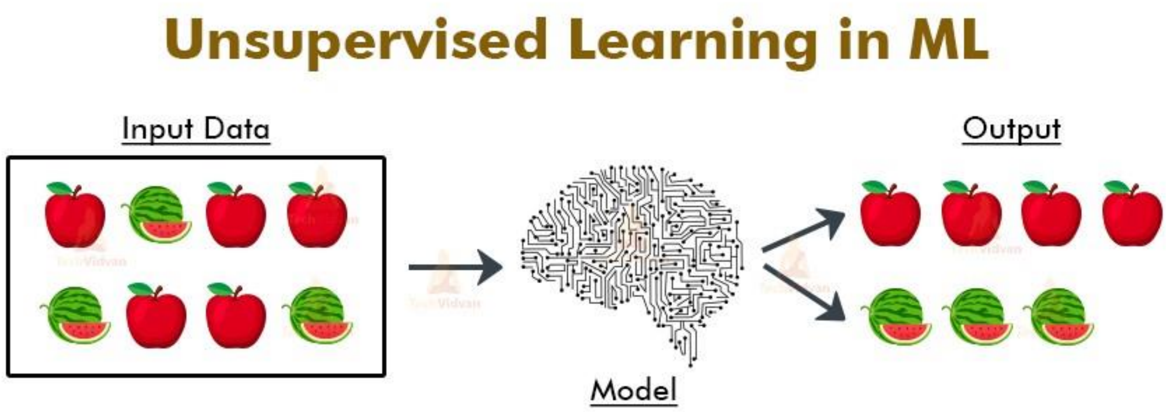

Figure 3_ 6 : unsupervised learning

3.4. Deep Learning

Deep learning is a subset of machine learning that deals with algorithms inspired by the structure and function of the human brain. Deep learning algorithms can work with an enormous amount of both structured and unstructured data. Deep learning's core concept lies in artificial neural networks, [25] which enable machines to make decisions.

The major difference between deep learning vs machine learning is the way data is presented to the machine. Machine learning algorithms usually require structured data, whereas deep learning networks work on multiple layers of artificial neural networks.

3.4.1. How Does Deep Learning Work?

1. Calculate the weighted sums.
2. The calculated sum of weights is passed as input to the activation function.
3. The activation function takes the "weighted sum of input" as the input to the function, adds a bias, and decides whether the neuron should be fired or not.
4. The output layer gives the predicted output.
5. The model output is compared with the actual output. After training the neural network, the model uses the backpropagation method to improve the performance of the network. The cost function helps to reduce the error rate.

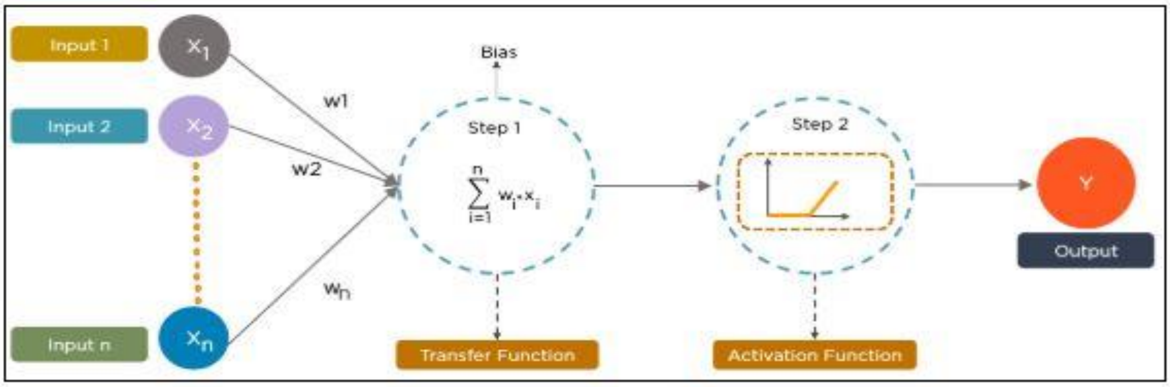

Figure 3_ 7 : How Does Deep Learning Work?

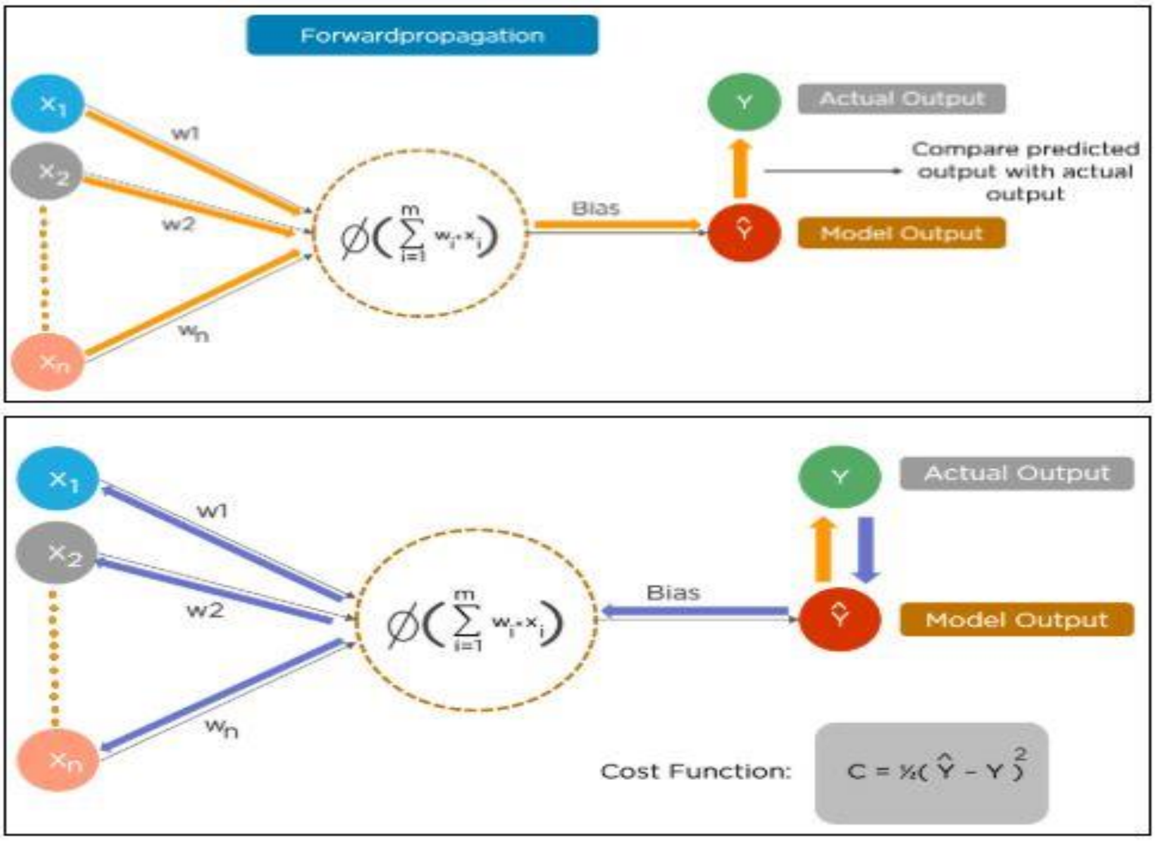

Figure 3_ 8 : How Does Deep Learning Work?

In the following example, deep learning and neural networks are used to identify the number on a license plate. This technique is used by many countries to identify rules violators and speeding vehicles [31].

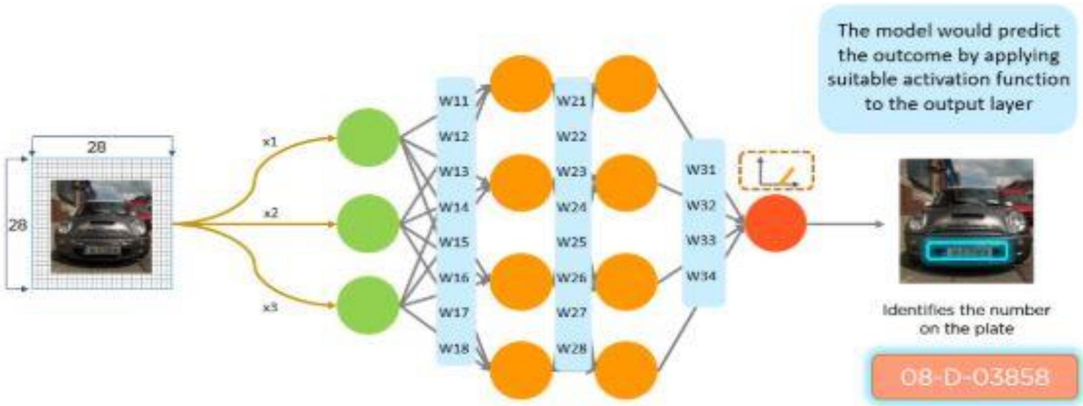

Figure 3_ 9 : How Does Deep Learning Work?

3.4.2. Importance of deep learning

Deep learning is strongly connected to artificial intelligence and machine learning. AI deep learning makes machines artificially intelligent to think or act like humans. Machine learning is just making a machine learn from its experience and improve its performance. Then, deep learning pops up – a cause of all of the miraculous discoveries the world has come across in recent years. The structure of deep learning is a bit similar to a human brain – there are many neurons and nodes like neurons [32].

3.4.3. Deep Learning Applications

Figure 3_ 10 : Deep Learning Applications

- Cancer tumor detection
- Captionbot for captioning an image
- Music generation
- Image coloring
- Object detection

3.5. Deep Neural Network

Deep learning is a Neural Network consisting of a hierarchy of layers, whereby each layer transforms the input data into more abstract representations. These series of layers, between input and output, identify the input features and create a series of new features based on the data, [32] just as our brain. In deep learning the more layers a network has, the higher the level of features it will learn. The output layer combines all these features and makes a prediction. This is different from a classical Artificial Neural Network [33].

Artificial neural network is only good at learning the weights of a network with one hidden layer but does not contain multiple hidden layers and hence it cannot learn complex features. Deep learning can be expensive and require massive datasets to train itself on. Since, in deep learning, more the neurons (cells in hidden layers) are, the more features it creates, and correspondingly it needs more data to train on. **Figure 4** present a simple deep neural network with three hidden layers.

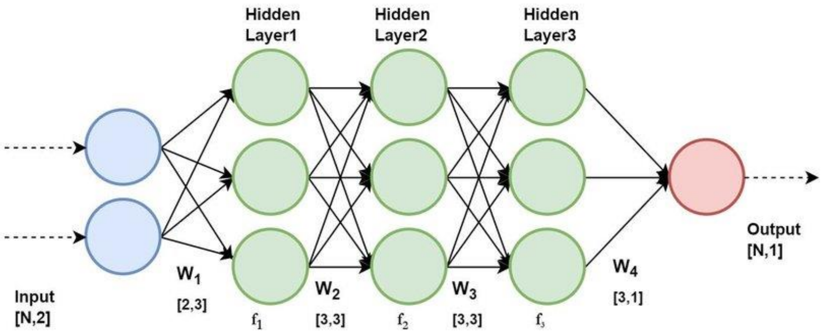

Figure 3_ 11 : Deep neural network with three hidden layers

Any Deep neural network will consist of three types of layers:

- The Input Layer
- The Hidden Layer
- The Output Layer

1.The input layer

It receives all the inputs and the last layer is the output layer which provides the desired output.

2. Hidden Layers

All the layers in between these layers are called hidden layers. There can be n number of hidden layers. The hidden layers and perceptrons in each layer will depend on the use-case you are trying to solve [34].

3. Output Layers

It provides the desired output.

3.6. Artificial intelligence versus machine learning, deep learning, and neural networks

	Artificial intelligence	Machine learning	Deep learning	Neural networks
What is it	Intelligence demonstrated by machines	A subset of artificial intelligence	A subset of machine learning	An approach to artificial intelligence
What does it use	It studies ways to build programs so that machines can solve problems	It provides systems the ability to automatically learn and improve from experience	It imitates the workings of the human brain in processing data so the system can create patterns	It analyses factors with a structure that is similar to the human neural system
Where is it used	Siri, Tesla, Alexa, Netflix, Face detection and recognition, Recommendation algorithms, Google maps	Virtual assistants, Traffic predictions, Social media with people you may know suggestions, Medical diagnosis	Self-driving cars, Visual recognition, Virtual assistants, Financial fraud detection	Vehicle control, trajectory prediction, social network filtering, Fault detectors and simulations, Product design analysis, Dynamic modelling

Figure 3_ 12 : Artificial intelligence versus machine learning, deep learning, and neural networks

3.7. Neural networks

An artificial neural network is an application, non linear with respect to its parameters θ that associates to an entry x an output y = f(x, θ). For the sake of simplicity, we assume that y is unidimensional, but it could also be multidimensional. This application f has a particular form that we will precise. The neural networks can be use for regression or classification. As usual in statistical learning, the parameters θ are estimated from a learning sample. The function to minimize is not convex, leading to local minimizers. The success of the method came from a universal approximation theorem due to Cybenko (1989) and Hornik (1991). Moreover, Le Cun (1986) proposed an efficient way to compute the gradient of a neural network, called backpropagation of the gradient, that allows to obtain a local minimizer of the quadratic criterion easily [33].

3.7.1. Artificial Neuron

An artificial neuron is a function fj of the input x = (x1, . . . , xd) weighted by a vector of connection weights wj = (wj,1, . . . , wj,d), completed by a neuron bias bj , and associated to an activation function φ, namely

$$y_j = f_j(x) = \phi(\langle w_j, x\rangle + b_j).$$

Several activation functions can be considered.

• The identity functions

$\Phi(x) = x$

• The sigmoid function (or logistic)

$$\phi(x) = \frac{1}{1 + \exp(-x)}.$$

• The hyperbolic tangent function ("tanh")

$$\phi(x) = \frac{\exp(x) - \exp(-x)}{\exp(x) + \exp(-x)} = \frac{\exp(2x) - 1}{\exp(2x) + 1}.$$

• The hard threshold functions

$\Phi\beta(x) = 1x\geq\beta.$

• The Rectified Linear Unit (ReLU) activation function

$\Phi(x) = \max(0, x).$

Here is a schematic representation of an artificial neuron where $\sum = (\omega_j , x_i) + b_j$.

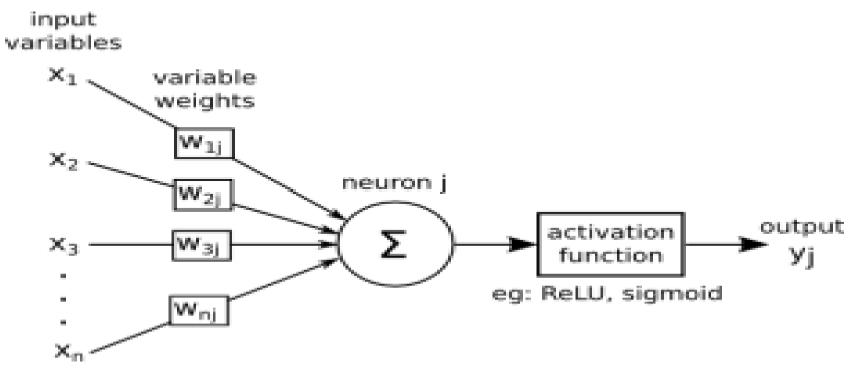

Figure 3_ 13 : artificial neuron where ∑ = (ωj , xi)+ bj

The Figure 3-32 represents the activation function described above.

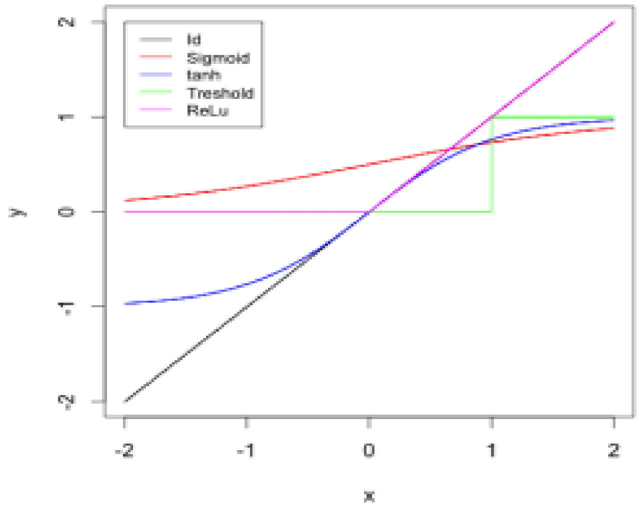

Figure 3_ 14 : Activation functions

Historically, the sigmoid was the mostly used activation function since it is differentiable and allows to keep values in the interval [0, 1]. Nevertheless, it is problematic since its gradient is very close to 0 when |x| is not close to 0. The Figure 3 represents the Sigmoid function and its derivative [25].

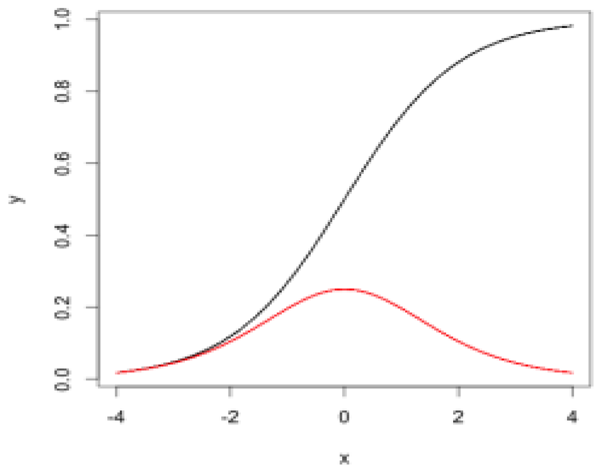

Figure 3_ 15 : Sigmoid function (in black) and its derivatives (in red)

With neural networks with a high number of layers (which is the case for deep learning), this causes troubles for the backpropagation algorithm to estimate the parameter (backpropagation is explained in the following) [22]. This is why the sigmoid function was supplanted by the rectified linear function. This function is not differentiable in 0 but in practice this is not really a problem since the probability to have an entry equal to 0 is generally null. The ReLU function also has a sparsification effect. The ReLU function and its derivative are equal to 0 for negative values, and no information can be obtain in this case for such a unit, this is why it is advised to add a small positive bias to ensure that each unit is active. Several variations of the ReLU function are considered to make sure that all units have a non vanishing gradient and that for x < 0 the derivative is not equal to 0. Namely

$$\Phi(x) = \max(x, 0) + \alpha \min(x, 0)$$

where α is either a fixed parameter set to a small positive value, or a parameter to estimate.

3.7.2. Convolutional neural networks

For some types of data, especially for images, multilayer perceptron's are not well adapted. Indeed, they are defined for vectors as input data, hence, to apply them to images, we should transform the images into vectors, losing by the way the spatial information's contained in the images, such as forms. Before the development of deep learning for computer vision, learning was based on the extraction of variables of interest, called features, but these methods need a lot of experience for image processing. The convolutional neural networks (CNN) introduced by LeCun have revolutionized image processing, and removed the manual extraction of features. CNN act directly on matrices [33], or even on tensors for images with three RGB color channels. CNN are now widely used for image classification, image segmentation, object recognition, face recognition ...

Figure 3_ 16 : Image annotation

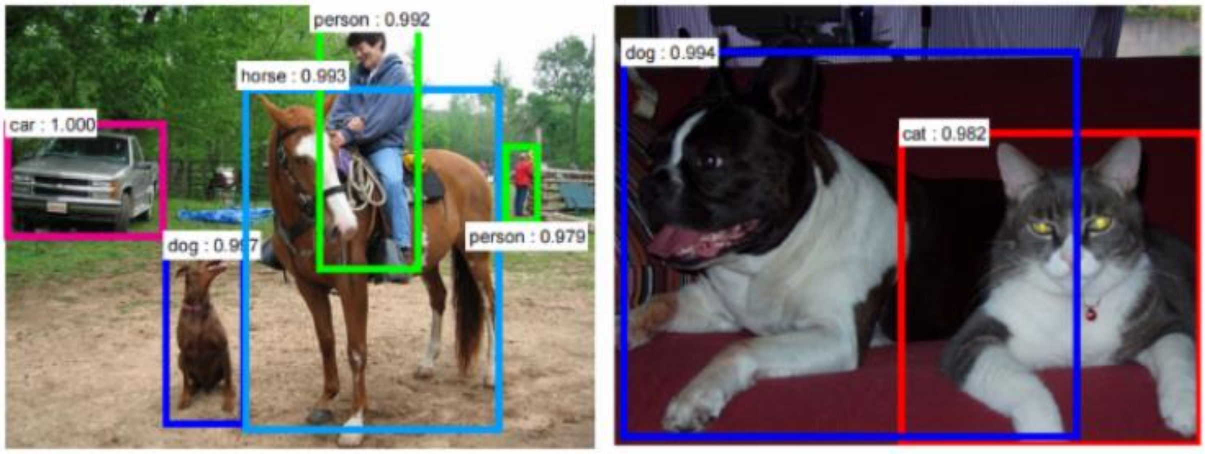

Figure 3_ 17 : Image Segmentation

3.7.2.1. Layers in CNN

A Convolutional Neural Network is composed by several kinds of layers, that are described in this section [25] : convolutional layers, pooling layers and fully connected layers.

3.7.2.2. Convolution layer

The discrete convolution between two functions f and g is defined as

$$(f * g)(x) = \sum_{t} f(t)g(x+t).$$

For 2-dimensional signals such as images, we consider the 2D-convolutions

$$(K * I)(i,j) = \sum_{m,n} K(m,n)I(i+n, j+m).$$

K is a convolution kernel applied to a 2D signal (or image) I.
As shown in Figure 8, the principle of 2D convolution is to drag a convolution kernel on the image. At each position, we get the convolution between the kernel and the part of the image that is currently treated. Then, [26] the kernel moves by a number s of pixels, s is called the stride. When the stride is small, we get redondant information. Sometimes, we also add a zero padding, which is a margin of size p containing zero values around the image in order to control the size of the output. Assume that we apply C0 kernels (also called filters), each of size k × k on an image. If the size of the input image is Wi × Hi × Ci

(Wi denotes the width, Hi the height, and Ci the number of channels, typically Ci = 3), the volume of the output is W0 × H0 × C0, where C0 corresponds to the number of kernels that we consider, and

$$W_0 = \frac{W_i - k + 2p}{s} + 1$$

$$H_0 = \frac{H_i - k + 2p}{s} + 1.$$

If the image has 3 channels and if Kl (l = 1,…, C0) denote 5 × 5 × 3 kernels (where 3 corresponds to the number of channels of the input image), the

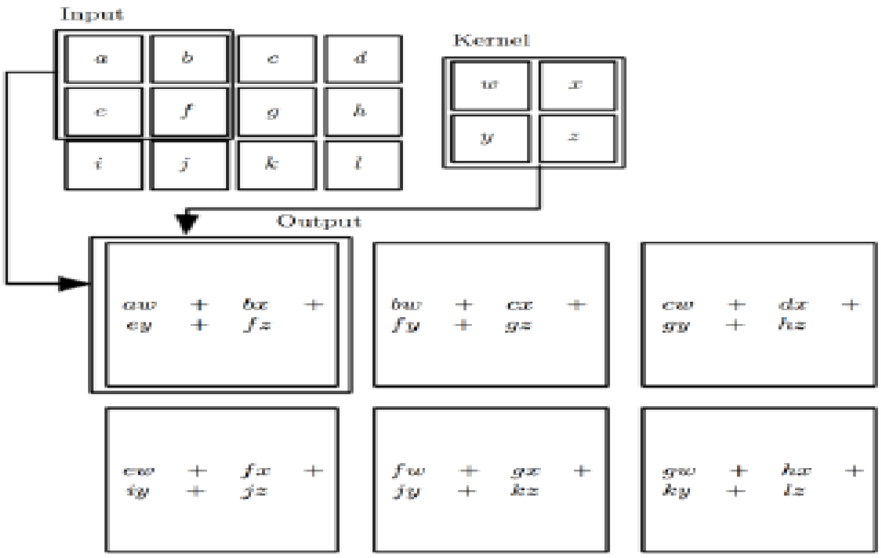

Figure 3_ 18 : 2D convolution

convolution with the image I with the kernel Kl corresponds to the formula:

$$K_l * I(i,j) = \sum_{c=0}^{2} \sum_{n=0}^{4} \sum_{m=0}^{4} K_l(n,m,c) I(i+n-2, i+m-2, c).$$

More generally, for images with Ci channels, the shape of the kernel is (k, k, C^i ,C^0) where C0 is the number of output channels (number of kernels) that we consider. This is (5, 5, 3, 2) in Figure 10. The number of parameter associated with a kernel of shape (k, k, C^i ,C^0) is (k × k × C^i + 1) × C^0.

The convolution operations are combined with an activation function _ (generally the Reluactivation function) [15]: if we consider a kernel K of size k×k, if x is a k × k patch of the image, the activation is obtained by sliding the k × k window and computing

z(x) = Φ(K * x + b), where b is a bias.

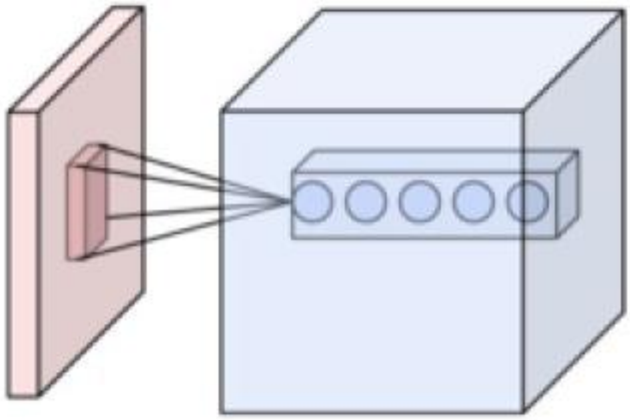

Figure 3_ 19 : 2D convolution - Units corresponding to the same

This is in the convolution layer that we find the strength of the CNN, indeed, the CNN will learn the filters (or kernels) that are the most useful for the task that we have to do (such as classification). Another advantage is that several convolution layers can be considered : the output of a convolution becomes the input of the next one.

3.7.2.3. Pooling layer

CNN also have pooling layers, which allow to reduce the dimension, also referred as subsampling, by taking the mean or the maximum on patches of the image (mean-pooling or max-pooling). Like the convolutional layers, pooling layers acts on small patches of the image, we also have a stride. If we consider 2×2 patches, over which we take the maximum value to define the output layer, and a stride $s = 2$, we divide by 2 the width and height of the image. Of course, it is also possible to reduce the dimension with the convolutional layer, [25]by taking a stride larger than 1, and without zero padding

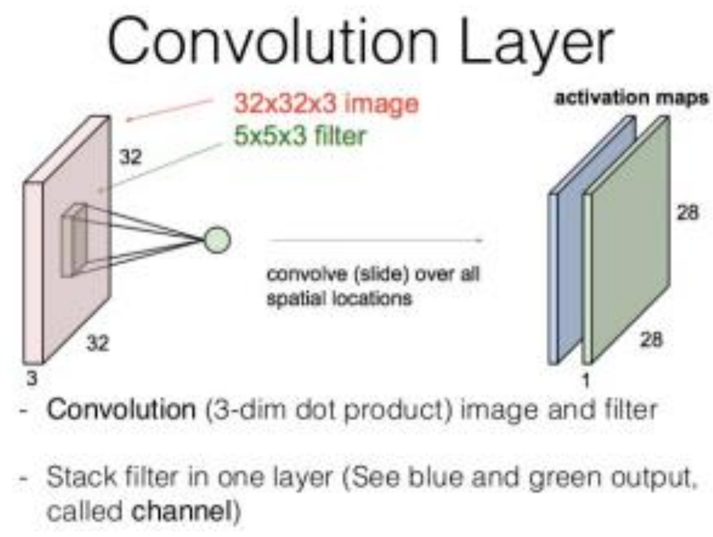

Figure 3_ 20 : Convolutional layer

but another advantage of the pooling is that it makes the network less sensitive to small translations of the input images.

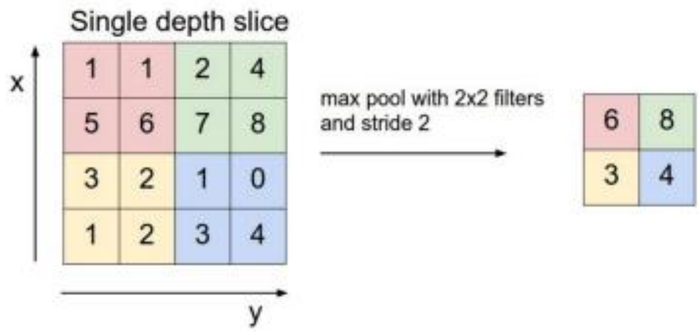

Figure 3_ 21 : Maxpooling and effect on the dimension

3.7.2.4. Fully connected layers

After several convolution and pooling layers, the CNN generally ends with several fully connected layers. The tensor that we have at the output of these layers is transformed into a vector and then we add several perceptron layers.

3.7.3. Architectures

We have described the different types of layers composing a CNN. We now present how these layers are combined to form the architecture of the network. Choosing an architecture is very complex and this is more engineering that an exact science. It is therefore important to study the architectures that have proved to be effective and to draw inspiration from these famous examples. In the most classical CNN, we chain several times a convolution layer followed by a pooling layer and we add at the end fully connected layers. The LeNet network, proposed by the inventor of the CNN, [15] Yann LeCun is of this type, as shown in Figure 12. This network was devoted to digit recognition. It is composed only on few layers and few filters, due to the computer limitations at that time.

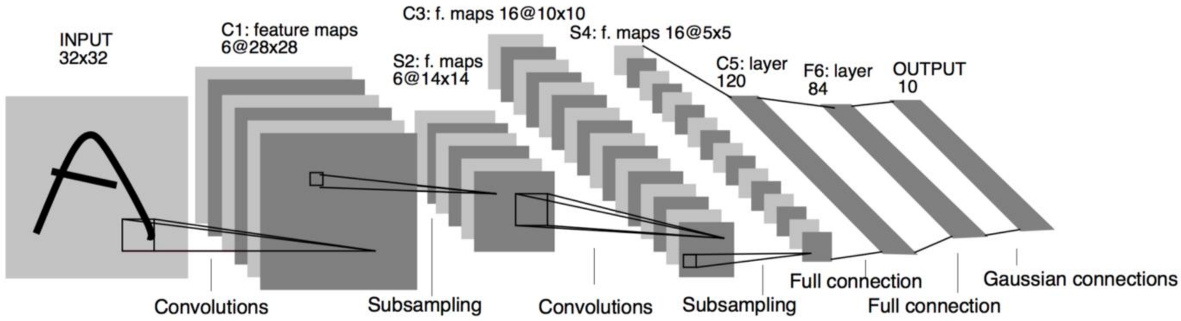

Figure 3_ 22 : Convolutional neural network.

A few years later, with the appearance of GPU (Graphical Processor Unit) cards, much more complex architectures for CNN have been proposed, like the network AlexNet that won the ImageNet competition and for which a simplified version is presented in Figure 13. This competition was devoted to the classification of one million of color images onto 1000 classes. The resolution of images was 224 × 224. AlexNet is composed of 5 convolution layers, 3 max-pooling 2 × 2 layers and fully connected layers. As showed if Figure 13, the kernel shape of the first convolution layer is (11, 11, 3, 96) with a stride of s = 4, and the first output shape is (55, 55, 96).

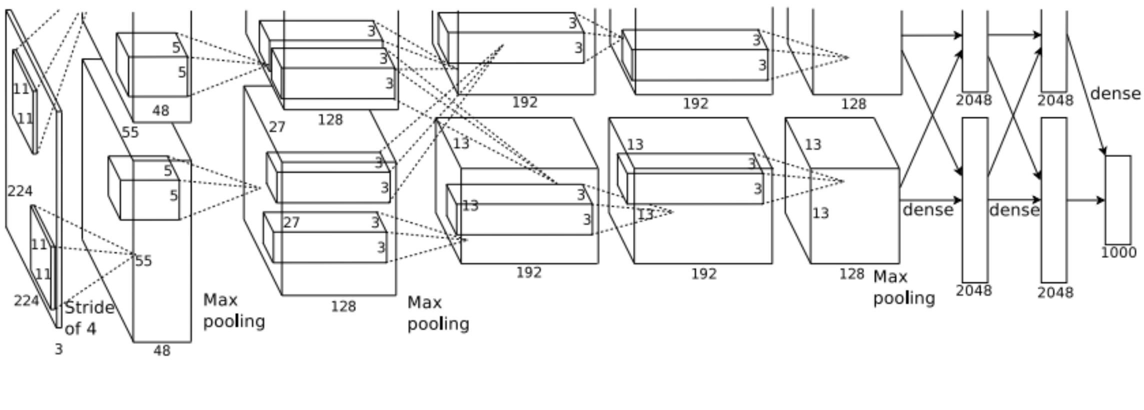

Figure 3_ 23 : CNNArchitectures

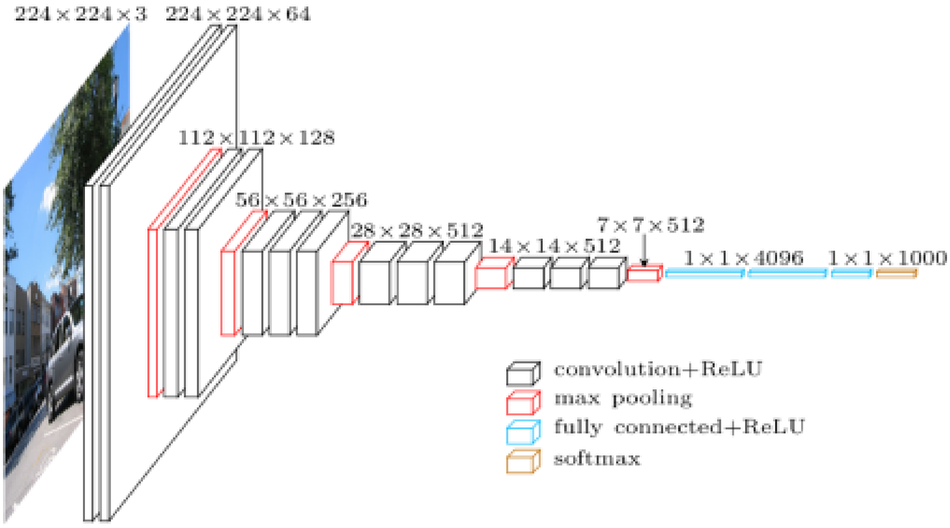

Figure 3_ 24 : architectures for CNN

The network that won the competition in 2014 is the network GoogLeNet, which is a new kind of CNN, not only composed on successive convolution and pooling layers, but also on new modules called Inception, which are some kind of network in the network.The most recent innovations concern the ResNet networks .The originality of the ResNets is to add a connection linking the input of a layer (or a set of layers) with its output. In order to reduce the number of parameters, [9]the ResNets do not have fully connected layers. GoogleNet and ResNet are much deeper than the previous CNN, but contain much less parameters. They are nevertheless much costly in memory than more classical CNN such as VGG or AlexNet.

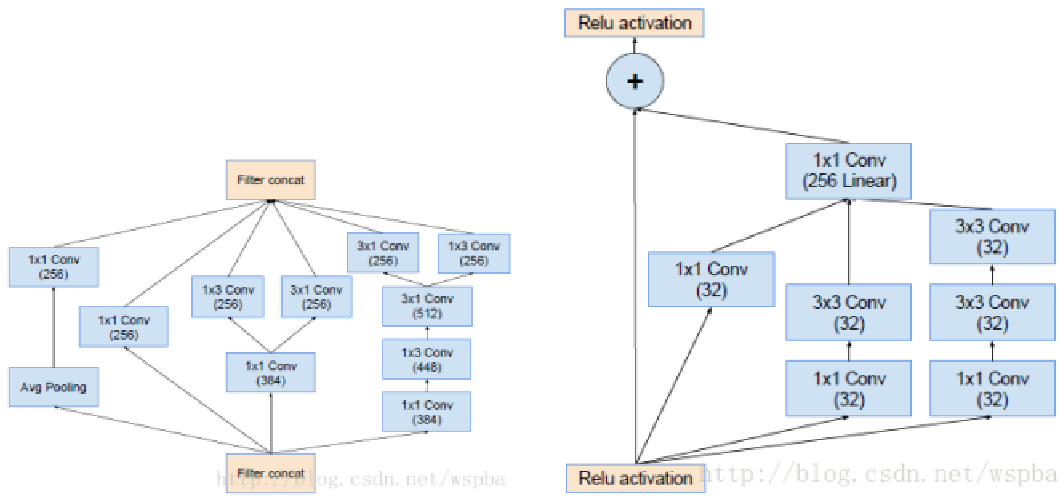

Figure 3_ 25 : inception modules,

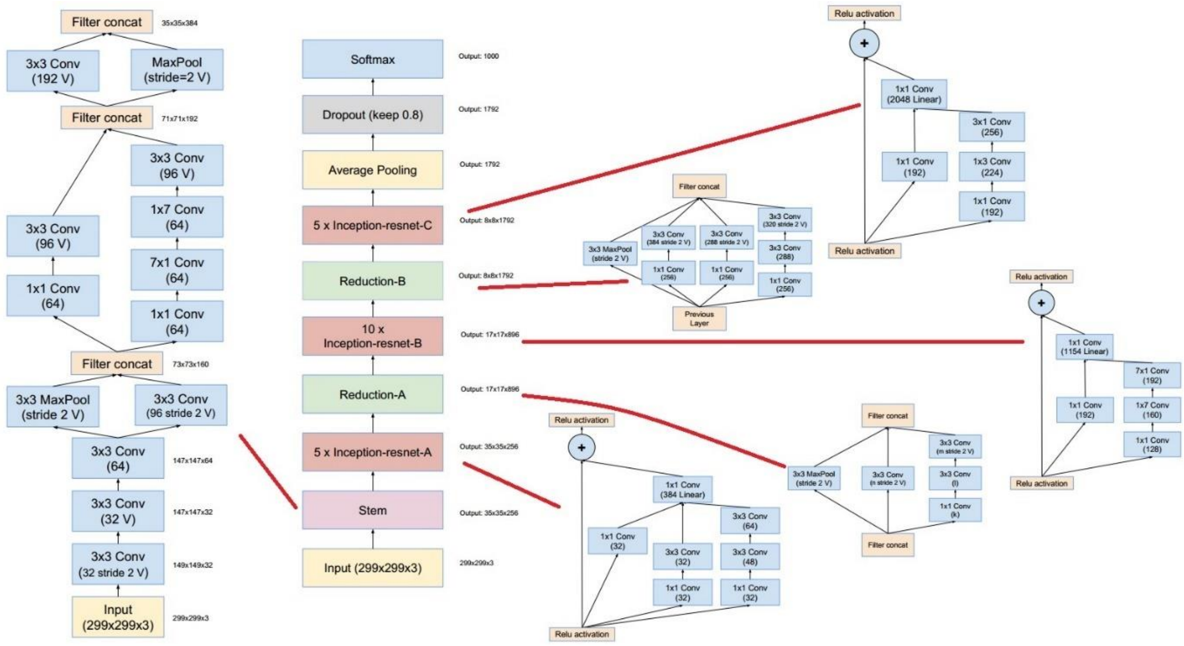

Figure 3_ 26 : Inception-v4,

3.8. Conclusion

as all the four terms discussed are inextricably connected, it is worth remembering the differences between them.

The main idea behind them is to make our lives much easier. In today's reality with the continuous growth of advanced technology, there is no doubt, we will be surrounded by more and more various devices and applications where Artificial Intelligence principles are used. That creates as well a huge opportunity for all, who'd like to become an expert in this exciting field and work for global technology leaders.

In this chapter we presented the concepts related to artificial intelligence, machine learning, deep learning and neural network and the relationship they have with, after we presented the neural networks and their different types then we focused our attention on CNN convolutional neural networks, their structure, and its different layers, we finally presented some examples of architectures. This detailed study of the CNN model allows us to better understand its process and to propose our CNN model which will be the subject of Chapter 4.

CHAPTER4 EXPERIMENTS, DISCUSSION AND APPLICATION

4.1. Introduction

After addressing the theoretical aspect in the previous chapters, we move on to the design and implementation of our system developed during work for gender prediction, age estimation and face mask detection of individuals based on the convolutional neural networks.

4.2. Face DataSets

4.2.1. UTK Face Dataset

UTKFace dataset contains 23,708 RGB images of faces in JPG format of size 200x200 pixels each. The labels of images are embedded in the file names, formatted like [age]_[gender]_[race]_[date&time].jpg. The distribution and range of ages within the dataset are shown in the bar charts below. The distribution and range of ages within the dataset are shown in the bar charts below.

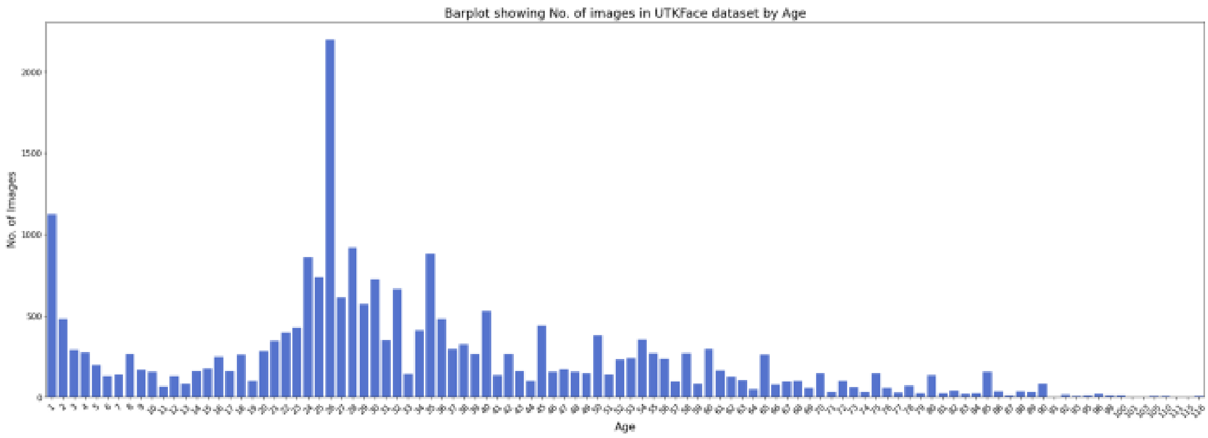

Figure 4_ 1 : Bar chart sowing no of img in UKTface dataset by age

Figure 4_ 2 : Examples of images in UTK Face Dataset

4.2.2. Facial-Age DataSet

Facial age dataset contains 10000 RGB images of faces in png format of size 200x200 Pixel each. The images are separated into folders and the folder names correspond to the age labels of images inside those folders. The distribution and rang of ages within the dataset are shown in the bar charts below.

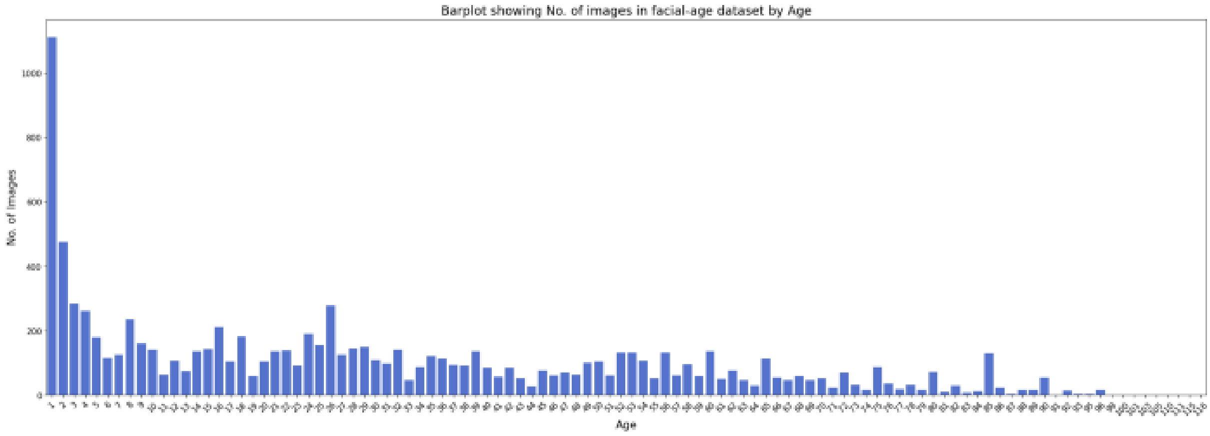

Figure 4_ 3 : Bar chart of facial age DS BY AGE

4.2.3. Age Dataset

Since both Facial age and UTKF are datasets already provided .so they decided to merge the them together and convert all 33,486 images to a standard JPG format. The distribution of images after combining the two datasets is shown in the bar charts below.

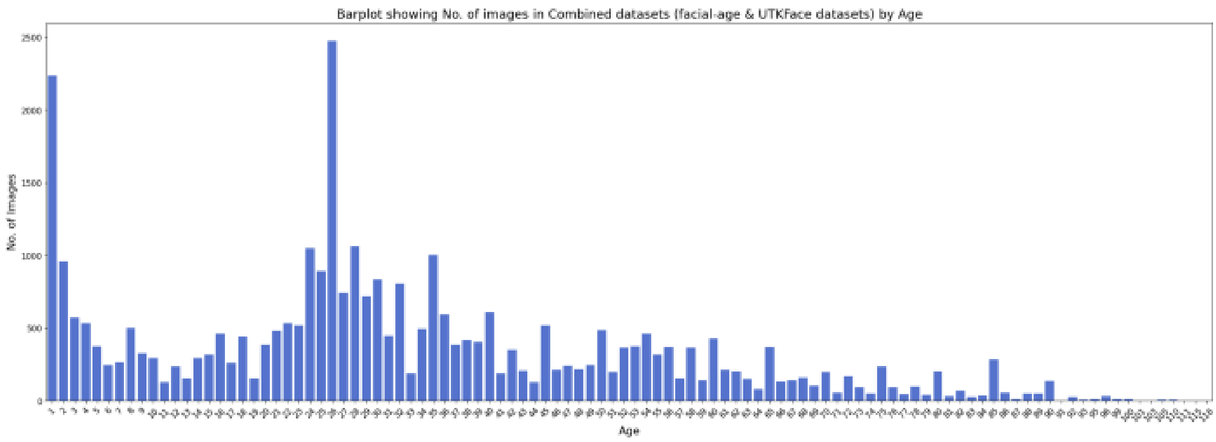

Figure 4_ 4 : Bar chart of combining facial age DS with UKTface DS by age

After combining the two datasets, the next step was to divide the different age labels into classes of age-ranges.

Class label	Age-ranges (classes)	No. of images	Class balance (%)
0	1 - 2	3192	9.53
1	3 - 9	2816	8.41
2	10 - 20	3136	9.37
3	21 - 25	3474	10.37
4	26 - 27	3217	9.61
5	28 - 31	3063	9.15
6	32 - 36	3086	9.22
7	37 - 45	3207	9.58
8	46 - 54	2802	8.37
9	55 - 65	2796	8.35
10	66 - 116	2697	8.05

Figure 4_ 5 : AGE RANGES

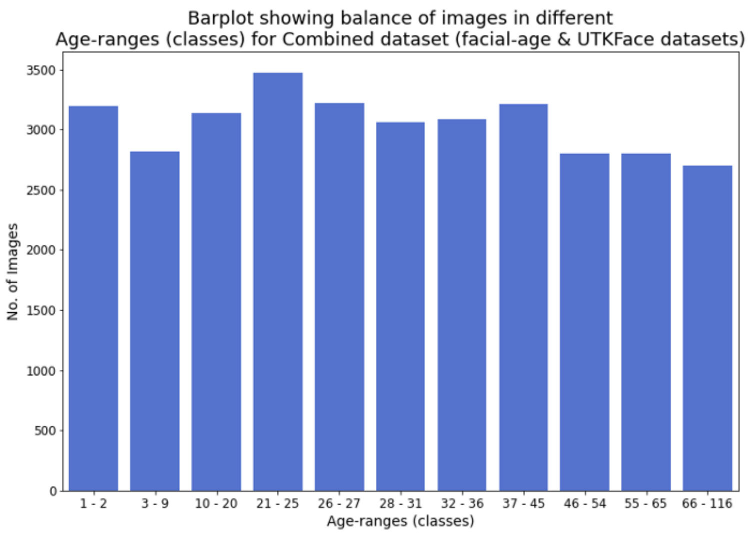

Figure 4_ 6 : Distribution of images into classes of age-ranges

- **Image Data Augmentation**

Data augmentation will do slight modifications on all our training set images we have, by rotating the original image 20 degrees and 40 degrees clockwise and anticlockwise. Plus, it mirrors our original image and then again do the rotations likewise.

Figure 4_ 7 : Image augmentation

So, with this, we have 10 total images from each base image, labelled the same as the original image. Or in other words our training dataset becomes 10x of what we originally had.

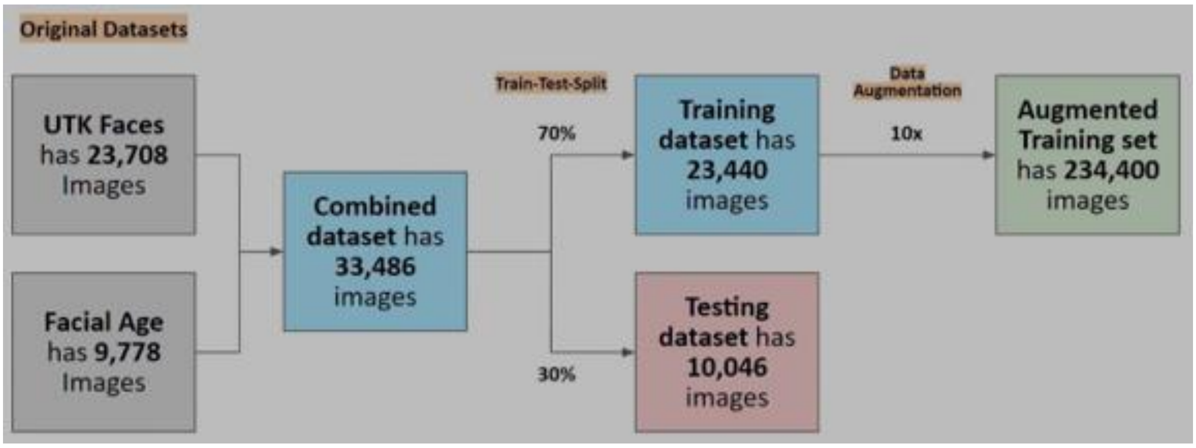

Figure 4_ 8 : Training set augmentation

4.2.4. Facemask Dataset

This dataset consists of 4095 images belonging to two classes with mask: 2165 images without mask: 1930 images the images used were real images of faces wearing masks

Figure 4_ 9Example of images in Facemask Dataset

4.2.5. Separation of Dataset

To implement our application, it is necessary to divide the databases into two parts, the first for training in order to train the CNN models and the second for the test to value the classification rates of the system. I used the sklearn.model_selection.train_test_split method to split the utkface dataset and age dataset into 70% training data (23,440 images) and 30% testing data (10,046 images). And divide also mask dataset into training and testing splits using 75% of the data for training and the remaining 25% for test set.

System	DataSet	TrainigSet	TestingSet
Age Estimation	Age DataSet	234400	10046
Gender Prediction	UTKface DataSet	17781	5927
FaceMask Detection	FaceMask DataSet	3072	1023

Table 1 : DATA SEPARATION

4.3. The work environment

In order to carry out this project, I used material with main characteristics are as follows:

- A laptop Lenovo i3 CPU 1.70GHz
- Graphic carde Intel(R) HD Grafics 4400
- RAM 4.00GB
- Windows 10 Pro

4.3.1. programming Tools

- Python: Python is an interpreted, object-oriented, high-level programming language with dynamic semantics. Its high-level built-in data structures, combined with dynamic typing and dynamic binding, make it very attractive for Rapid Application Development, as well as for use as a scripting or glue language to connect existing components together. Python's simple, easy to learn syntax emphasizes readability and therefore reduces the cost of program maintenance. Python supports modules and packages, which encourages program modularity and code reuse. The Python interpreter and the extensive standard library are available in source or binary form without charge for all major platforms, and can be freely distributed.

4.3.2. Libraires used

- **TensorFlow:** is an open-source end-to-end platform for creating Machine Learning applications. It is a symbolic math library that uses dataflow and differentiable programming to perform various tasks focused on training and inference of deep neural networks. It allows developers to create machine learning applications using various tools, libraries, and community resources.
- **OpenCV:** (Open-Source Computer Vision Library) is an open-source computer vision and machine learning software library. OpenCV was built to provide a common infrastructure for computer vision applications and to accelerate the use of machine perception in the commercial products.
- **Keras:** is a high-level neural network library, written in Python and able to run on TensorFlow or Theano. It was developed with the aim of allowing rapid experimentation, to be able to go from idea to result as quickly as possible. We used keras because it allows easy and fast prototyping (thanks to full modularity), it supports convolutional neural networks, and runs seamlessly on CPU and GPU
- **Face Recognition:** Recognize and manipulate faces from Python or from the command line with the world's simplest face recognition library.
- Built using dlib's state-of-the-art face recognition built with deep learning. The model has an accuracy of 99.38% on the Labeled Faces in the Wild benchmark.
- This also provides a simple face recognition command line tool that lets you do face recognition on a folder of images from the command line!
- **Numpy:** The Numpy library allows you to perform numerical calculations with Python. It introduces easier management of N-dimensional arrays
- **Streamlit:** Streamlit is an open source app framework in Python language. It helps us create web apps for data science and machine learning in a short time. It is compatible with major Python libraries such as scikit-learn, Keras, PyTorch, SymPy(latex), NumPy, pandas, Matplotlib etc.

4.4. Age Estimation System

Estimation of age consists in classifying a person according to classes of age-ranges

Wich is 11 classes. the classification process by CNN networks and Keras Library

4.4.1. The effect of architecture

During our experiments, we created three configuration models of CNN convolutional networks with different architectures trained on the basis of age dataset images.

➢ **Architecture of model N°1**

Oue age estimation CNN model shall be defined and trained by:

- Importing training set and test set from age dataset
- Training dataset is already augmented and has 234400 images
- Gray scaling images instead of using RGB color images
- Defining our intuitively distributed classes of age ranges
- Using 30 epochs on our optimized CNN architecture, comprising of:
    - Defining the architecture of the sequential neural network
    - An input convolutional layer (with 32 filters) paired with an averagpooling layer
    - One convolutional layer (with 64 filters) and averagpooling layer
    - A Globalaveragepooling layer

    A Globalaveragepooling layer be for going into Dense layers below, it gives number of outputs equal to number of filters in last convolutional layer above (64)

    - One fully connected layer with 39 nodes

    One danse layer with 39 nodes so as to taper down the number of nodes from number of outputs of Globalaveragepooling layer above towards number of nodes in output layer below (7)

    - An output fully connected layer with 7 nodes with SoftMax activation function

    Evry filter is size of (3x3) and each layer followed by RELU activation function that forces the neurons to return positive values.

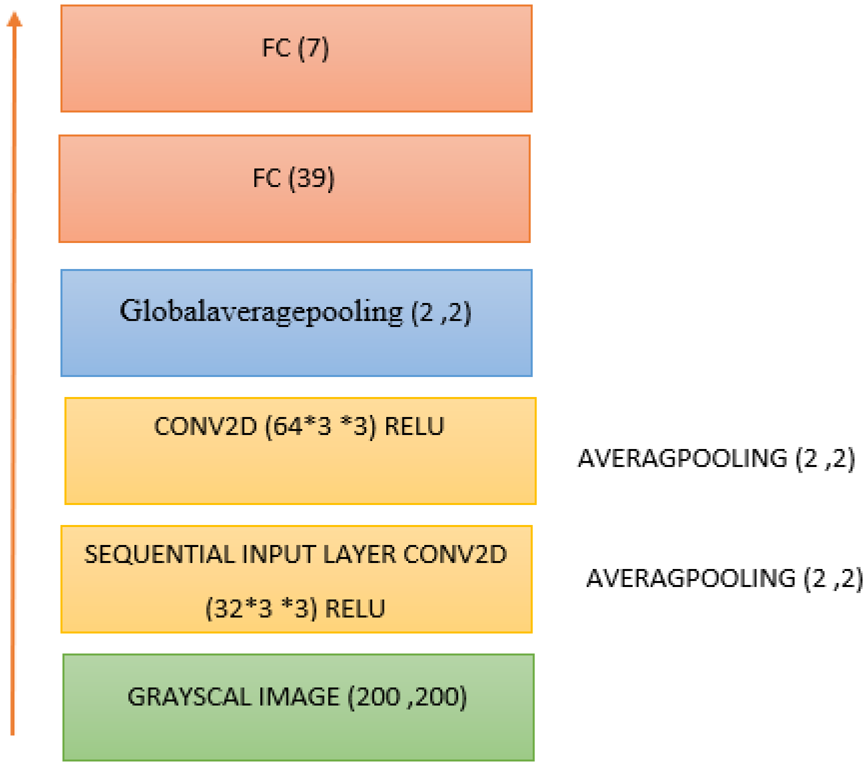

Figure 4_ 10 : CNN model1

➢ **Architecture of model N°2**

Oue age estimation CNN model 2 is similar to the model 1 but we added one mor convolutional layer and increase the number of filters to double (64) and changing the number of nodes in first danse layer to 72.

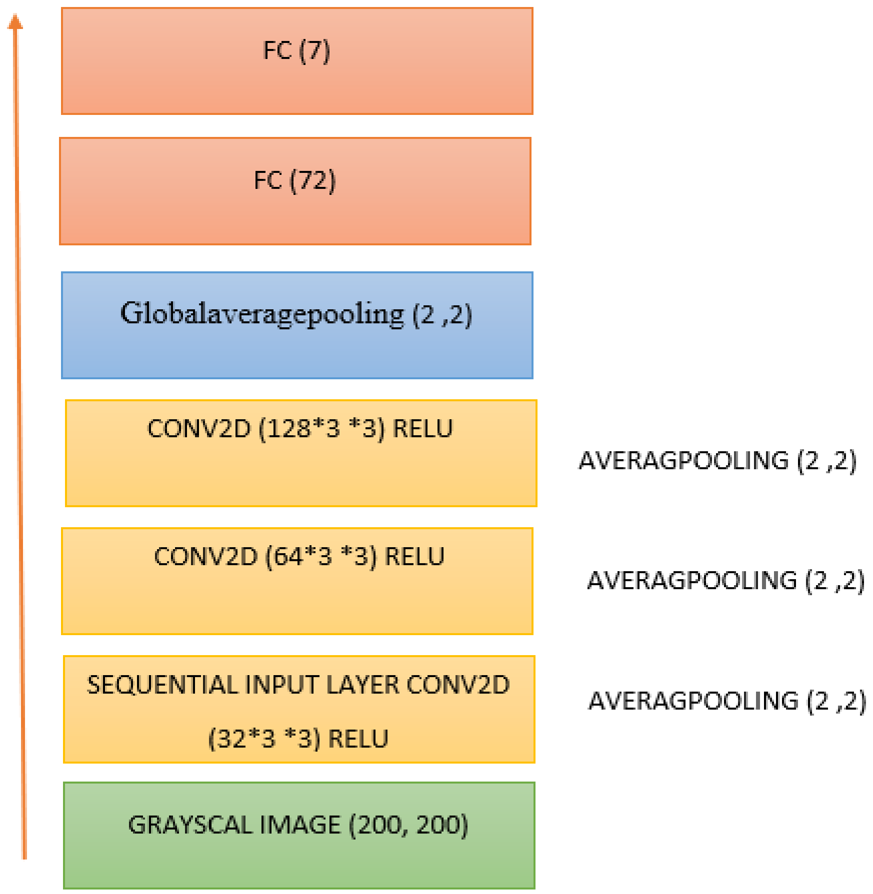

Figure 4_ 11 : CNN model 2

- **Architecture of model N°3**

Oue age estimation CNN model 3 similar to the model 2 but we added also one mor convolutional layer with increasing the number of filters to double (256) and changing the number of nodes in first danse layer to 132

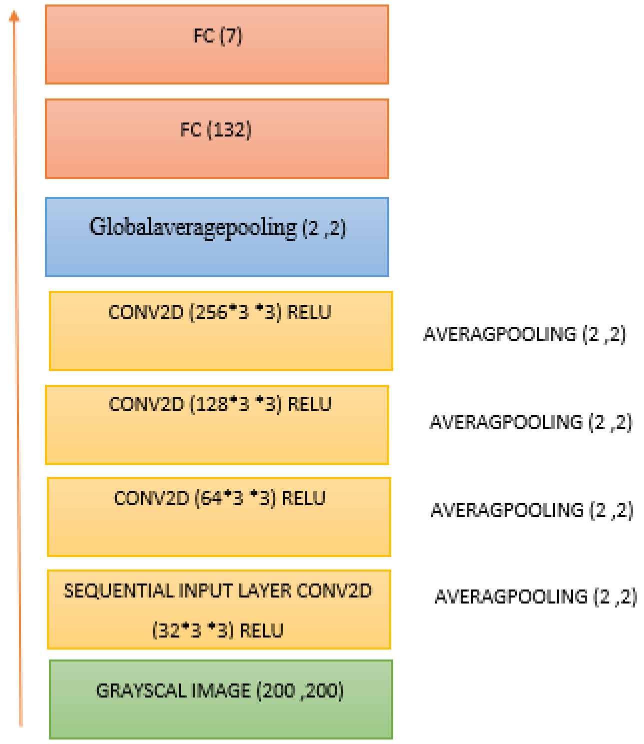

Figure 4_ 12 : CNN model 3

4.4.2. Resultes

The results obtained by the three CNN architectures proposed on the age database in 30 epochs are presented in the following table

Model	Trainset	Testset	Accuracy obtained on the trainset	Accuracy obtained on testset	Train loss value	Test loss value
N°1	234400	10046	**60.34%**	**58%**	**0.094**	**0.964**
N°2	234400	10046	**74.5%**	**73.25%**	**0.056**	**0.849**
N°3	234400	10046	**98%**	**97.3%**	**0.004**	**0.248**

Table 2 : The results obtained by the three CNN architectures

Table above illustrates the classification rates obtained on the three models used in 30 epochs, the number of training data from the age database is 33,486 samples distributed between test and training, the results obtained are expressed in terms of accuracy rate in the learning and testing as well as loss value.

Model N°3 has the best results, this is due to the depth of the network and the number of convolution layers used in this model compared to (model N°1 and N°2and N°3). Thus, the model N°2 presents acceptable results compared to the model N°1, difference is based on the number of filters used.

The resultes clearly shown in the curves below

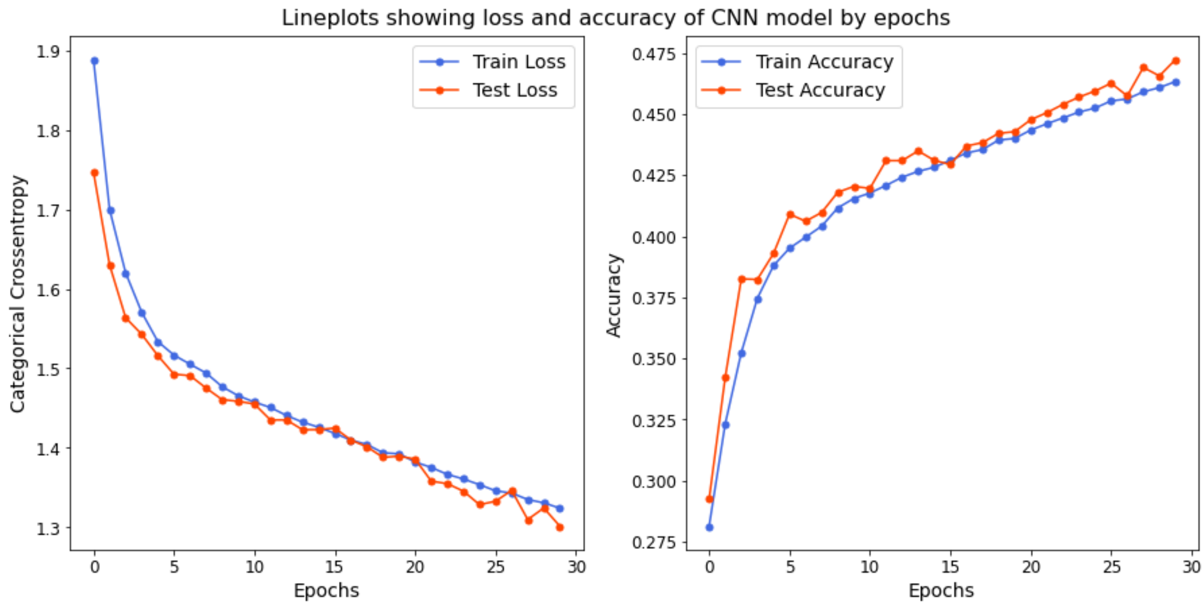

Figure 4_ 13 : result curves of model

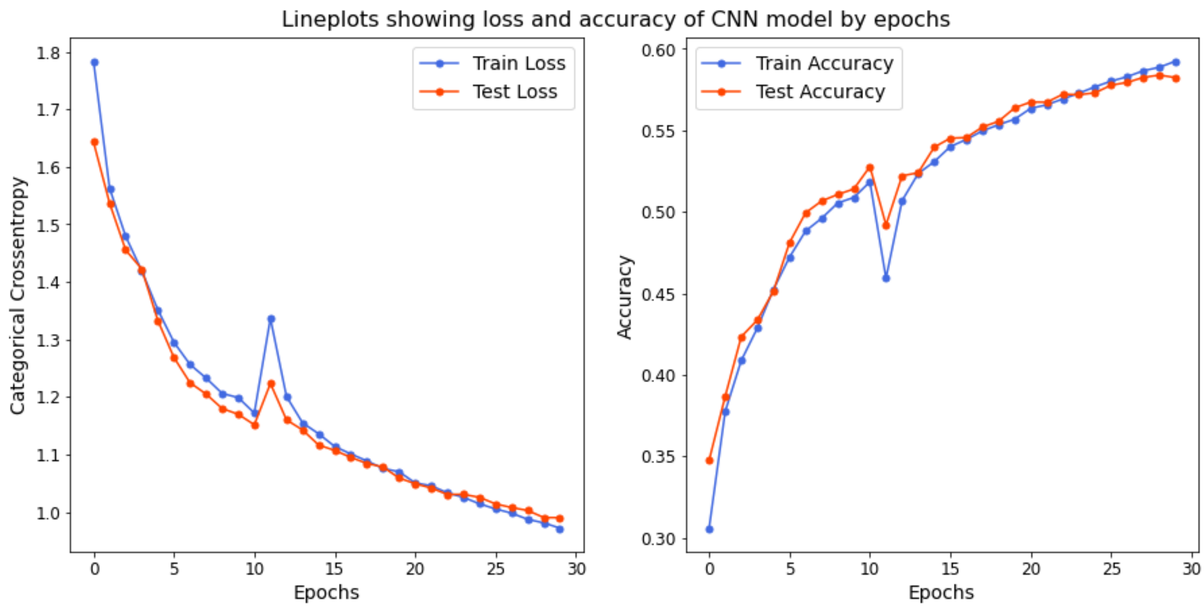

Figure 4_ 14 : result curves of mode2

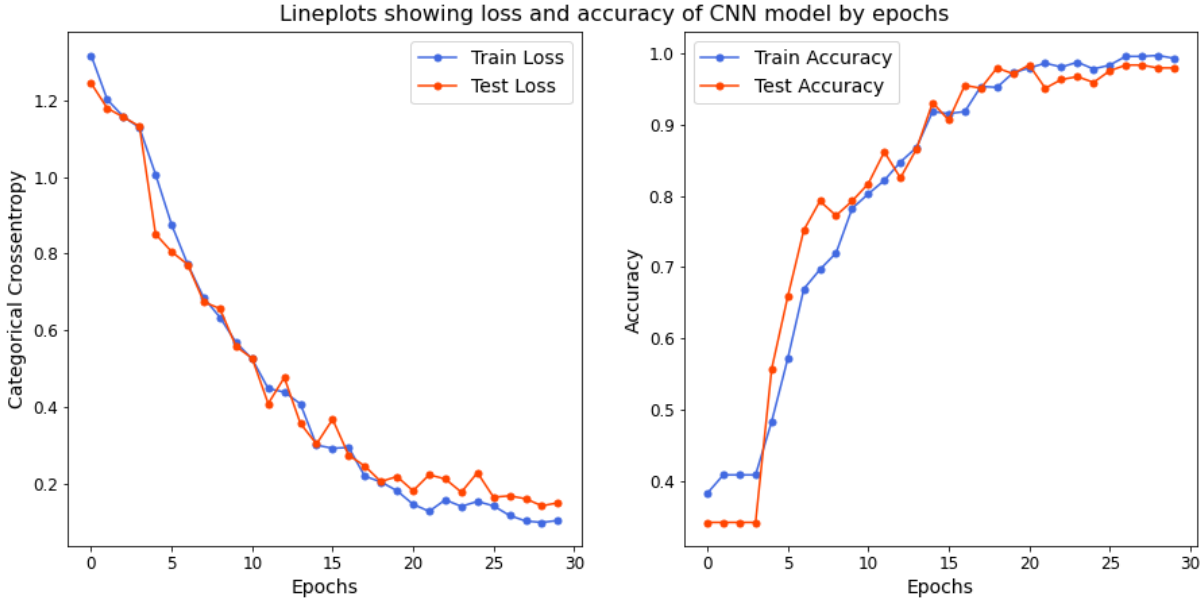

Figure 4_ 15 : result curves of model 3

After analyzing the results obtained, we see that:

• The learning and test accuracy increases with the number of epochs, this reflects that at each epoch the model learns more information.

• Model N°3 is more efficient with an accuracy rate in the test phase . followed by model N°2 and model N°1 the least efficient .

4.4.3. Effect of the number of training images on the recognition rate

In this experiment, we changed the number of training and test images in order to see the effect of the amount of training image on the classification rate. In this part we take up the architecture of model n°3 which will be trained on the age database with different galleries including a variation in the number of training and test images, table below summarizes our results

Gallery	Trainset	Testset	Test accuracy
Gallery1	3000	900	72.4%
Gallery2	9000	2700	85%
Gallery3	15000	4500	92.7%
Gallery4	234400	10046	98%

Table 3 : Effect of the number of training images on the recognition rate

From the results obtained, we find that the recognition accuracy is positively proportional to the number of training and test images.

4.5. Gender Prediction System

The classification task for gender prediction is a binary classification, where a session is labeled as male and female category

In this experiment we applied the architecture of the age model described for gender prediction with the use of a sigmoid function in the output layer, after the training of the CNN network on the UKTFACE dataset we have noticed that the performances obtained are relatively acceptable like it shown in the table below

Dataset	Trainset	Testset	Accurancy
UKTface dataset	17781	5927	93.15%

Table 4.: Gender Prediction System performence

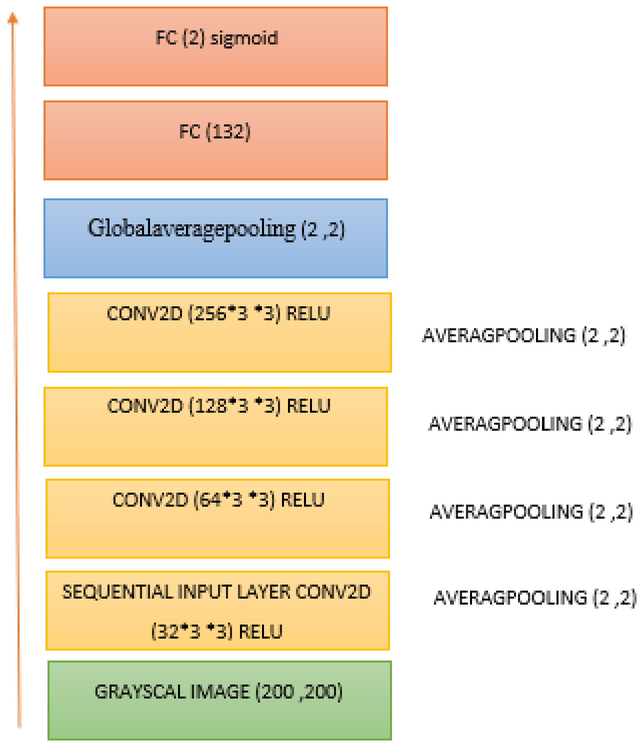

Figure 4_ 16 : Gender Prediction model

4.6. Face Mask Detection

Like gender prediction face mask detection, the classification task is a binary classification, where a session is labeled as with mask and without mask category

In this experiment we applied the architecture of the gender model described for face mask detection with the use of a sigmoid function in the output layer, after the training of the

CNN network on the Facemask dataset we have noticed that the performances obtained are relatively acceptable like it shown in the table below

Dataset	Trainset	Testset	Accurancy
FaceMask dataset	3072	1023	91.1%

Table 5 : Face Mask Detection performance

4.7. Application

For tasting our work in real life, we developed a web application that predict age and gender and detect face mask. We used for building the web application Streamlit platform

The web application contains

- **Age and gendre prediction**
    - Real time prediction
    - Image prediction
- **Face Mask detection**
    - Real time detection
    - Image detection

Interface: There is tow selectboxs for detection which is age and gender prediction and face mask detection.

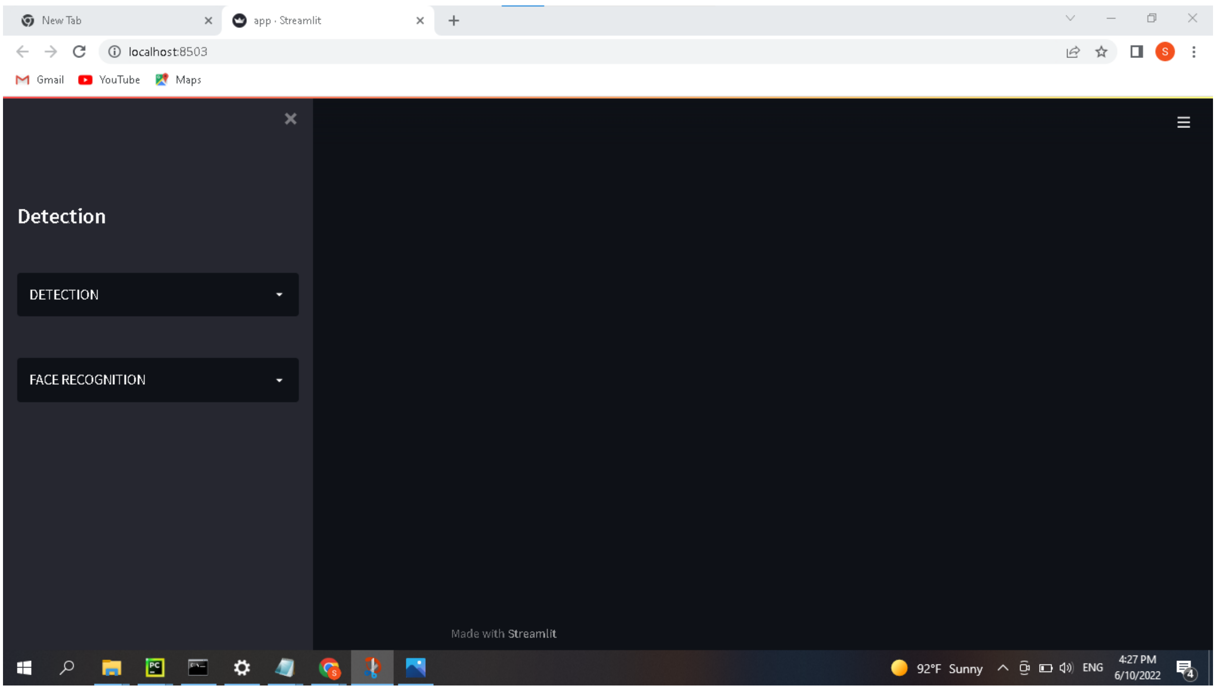

Figure 4_ 17 : The application interfaces

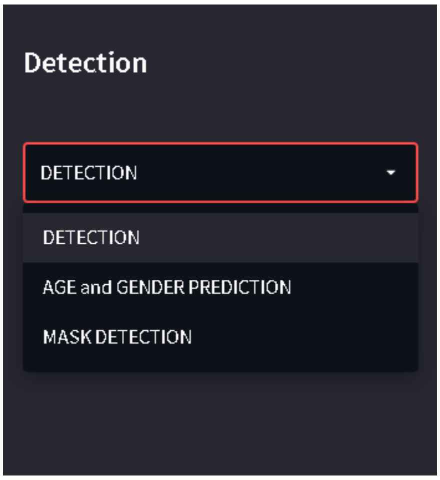

Figure 4_ 18 : Menu

Age and gender prediction: When selecting age and gender prediction it will appear another select box for real time prediction and image prediction

Run on image : In image predication there is browse file button to search for image to select

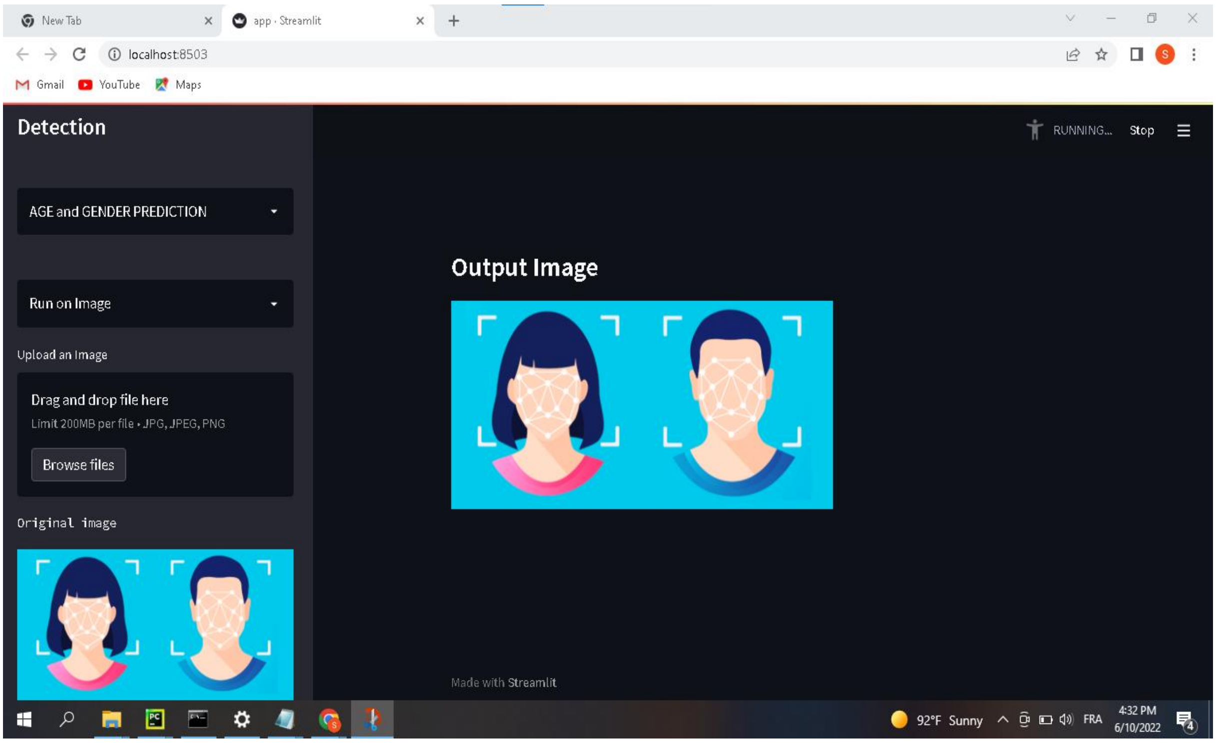

Figure 4_ 19 : Age and gender image prediction interface

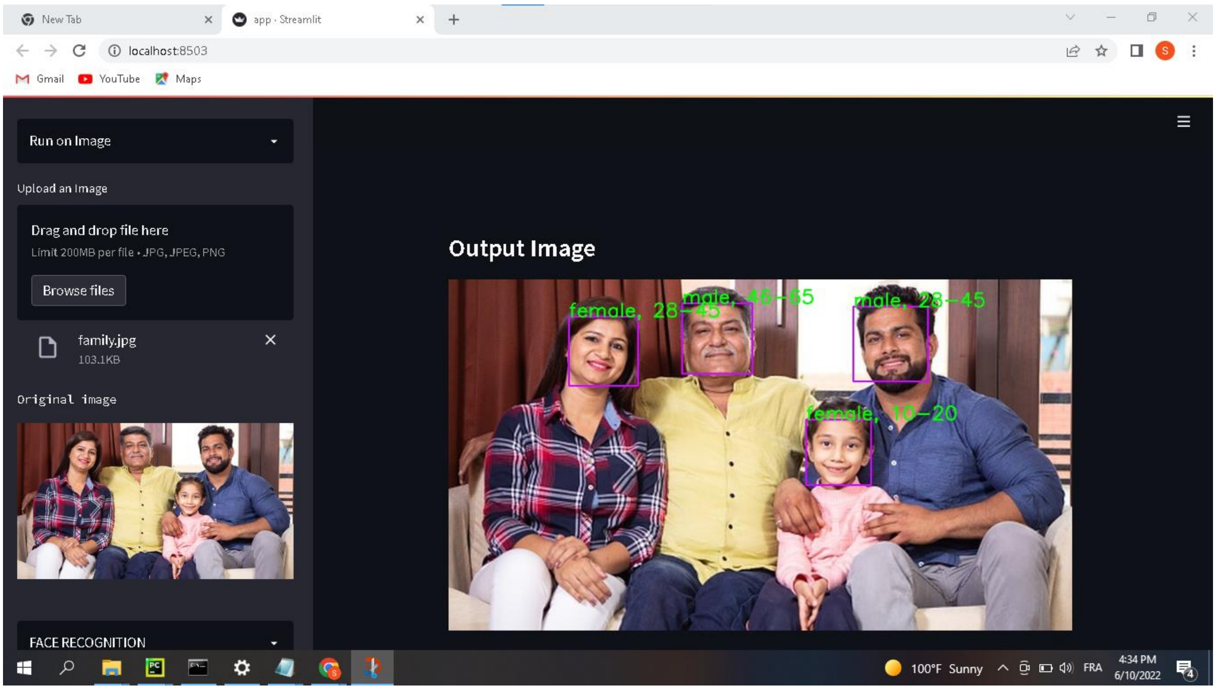

Figure 4_ 19 : Age and gender image prediction

- Run on video: it will appear a Run check box to start a video and detect the face

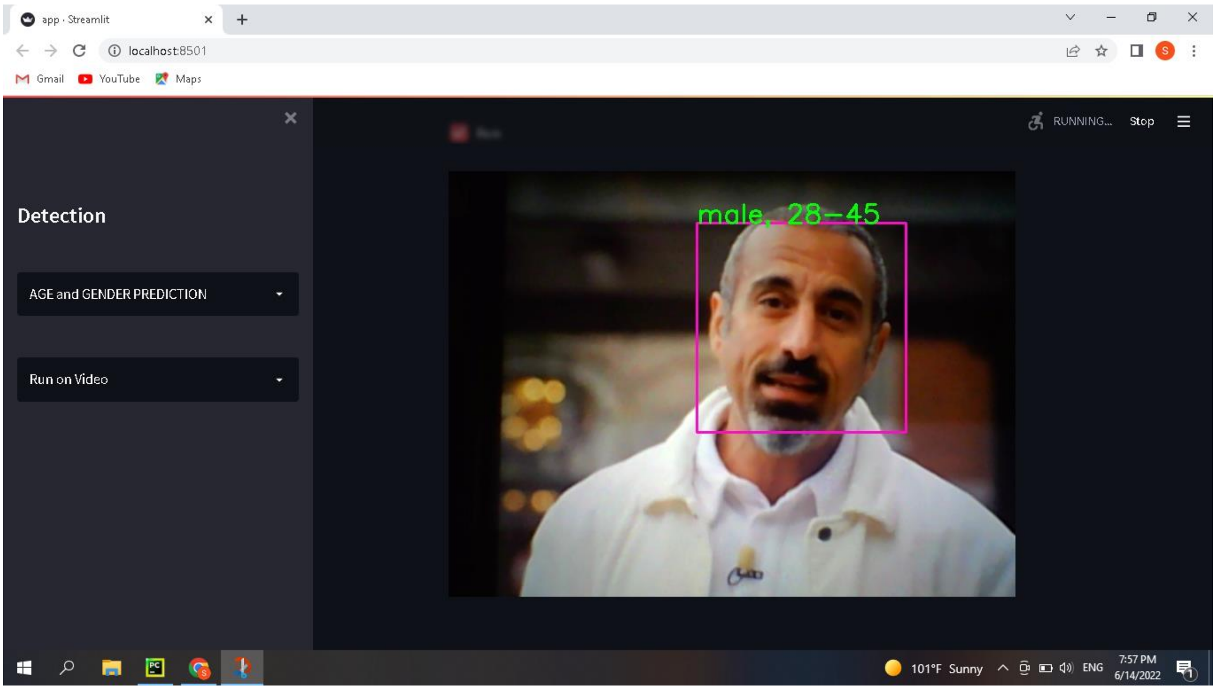

Figure 4_ 20 : Run on video

Face mask detection: In Face mask detection there is also a select box for video detection and image detection

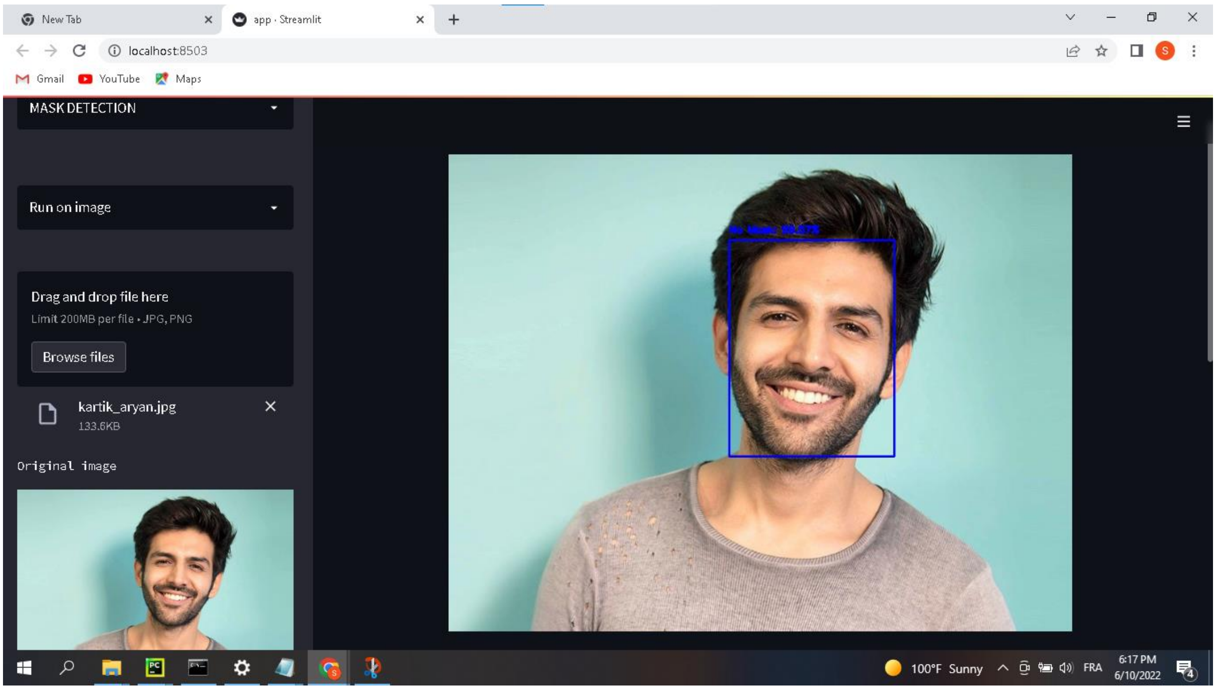

Figure 4_ 21 : Face Mask image detection

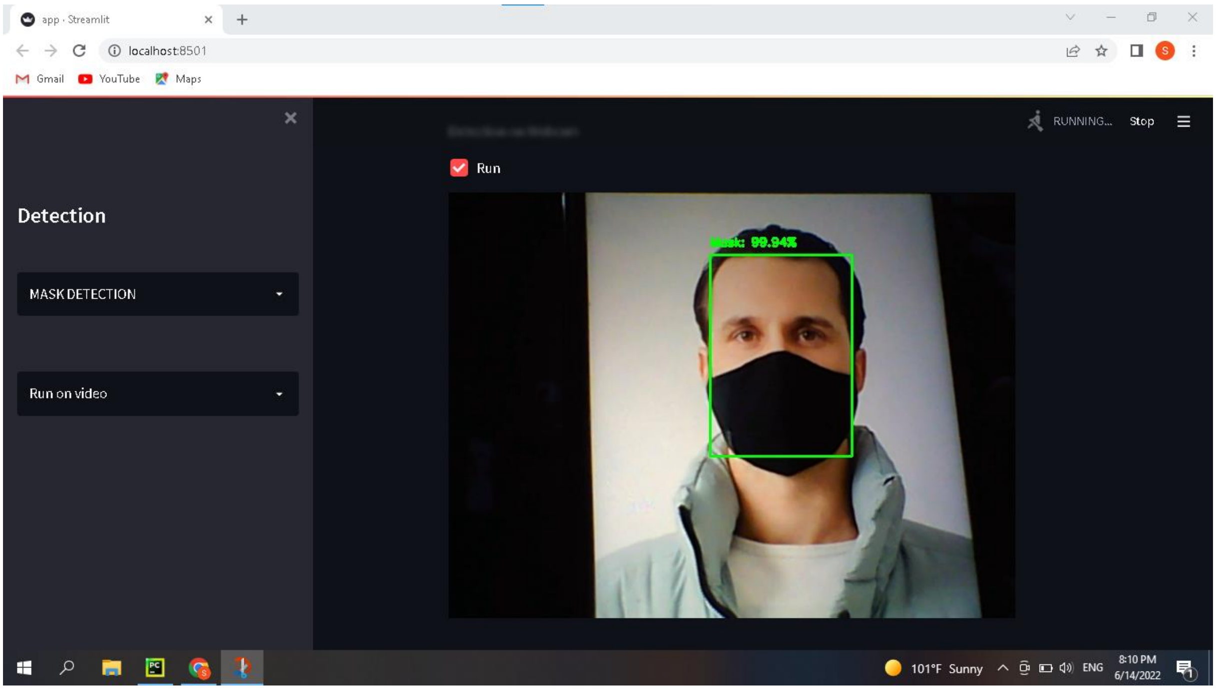

Figure 4_ 22: Run on video

4.8. Conclusion

In this chapter, we presented a gender and age prediction and face mask detection systems based on Deep neural networks. Our system is applied and tested on age estimation the we created three deferent models after training and testing them we came up with best one.

Finally, we developed a web application based on the model that we trained

GENERAL CONCLUSION

GENERAL CONCLUSION

biometrics quickly emerged as the most relevant to identify and authenticate people reliably and quickly, Facial recognition as a basic biometric technology has attracted particular interest, as it provides an unobtrusive and non-intrusive means of detection, identification and verification without the need for subject knowledge or consent.

The objective of this work is to design and implement a web application, able in real time to predict the gender and age of a person and detect face mask that can be used in social networks and billboards by responding and hospitals and schools beside the crowded places to speed and robustness requirements.

This study was divided into three parts: the first part consists in studying convolutional neural networks in order to propose an age estimation model, the second part is devoted to applying the most efficient model on gender prediction. and face mask detection, from a facial image, and the third part aims to integrate a real-time gender and age prediction system and face mask detection, we have validated the systems developed in this work on several datasets of UKT, AGE and FACE MASK datasets.

References

[1] A. S.Gillis, "Biometrics".

[2] "Performance evaluation of behavioral biometric system," 2009.

[3] Shilpa_Shrivastava, "Biometrics types and its applications," 2013.

[4] "recfaces About biometrics," 4 11 2014. [Online]. Available: https://recfaces.com/articles/types-of-biometrics.

[5] Foudil_Belhaj, "Biometric system for identification and authentication (Doctoral disyertation,)" 2017.

[6] Ashusharma, "Biometric system architecture," 3 12 2021.

[7] Bai_yuan_ding, "Desing of facial recognirion system based on visual communication effect, " culture and art managment of hunan university, korea, 2021.

[8] m. figen_ozen, "a face recognation system based on eigenfaces method, " halic university electonical and elicronics engineering depart,ent, sisli turky, 2011.

[9] f. f. ozen, "a new license plate recognition system based on probabilistic nueral networks, " halic university electonical and elictronic engineering department , sisli turky, 2020.

[10] w. wei_leng_deng, "deepface recognition," 2020.

[11] L. S. T. Maroua, "Reconnaissance de forme et classification automatique, " Kasdi merbah ouargla, 2019-2020.

[12] M. L. H. N. E. M.Khalifad, "A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic, " 1 January 2021.

[13] S. D. M. Menaka, "Real time data analysis of face mask detection and social distance measurement, " pariss, 20 February 2021.

[14] E.R.Dvies, Computer Vision, 5th Edition 14_11_2017.

[15] Bc.matus_Namesny, "Gander and Age Classification in Camera Data, " masaryk university, brno, 2019.

[16] M. Gurubasava, "Age and Gender Detection ising Python,"*International journal of Interdisciplinary Innovative Research and Devlopment(IJIIRD),* Vols. VOL-5, no. 01, 2020.

[17] P. Manavjeet, "Classification in pattern Recognition," *International Journal of Advanced Research in Computer Science and Software Engineering,* vol. VOLUM3, no. issue4, 4_4_2013.

[18] P. Pedamkar, "EDUCBA," [Online]. Available: educba.com/pattern-recognition.

[19] Jiri_Kepka, "The Current Approaches in Pattern Recognition," *KYBERNETIKA,* vol. 30 , p. 159_176.

[20] B. B. Sina_Massoud_Ansari, "Costal Image classification and pattern recogniation, " Gemine ViVone, Tairua_Beach New Zealand, 5 11 2021.

[21] P. Dieu_tien_Bui, HandBook of Probabilistic Models, 1st Edition 5_10_2019.

[22] a. a. m. m. e. azimi, "age and gender from face images using attentional convolutional networkes, " new york university , 7_12_2020.

[23] L. q. guanjing_huo, "Image Classification using Biomi,etic Patter Recognation zith Convolutional neural netwarkes, " *Co,putational Intelligence and Neuroscience,* vol. 2017, 2017.

[24] H. BAUM, "AN INTRODUCTION TO ARTIFICIAL INTELLIGENCE".

[25] "Artificial Intelligence versus Machine Learning, Deep Learning, and Neural Networks," Computer & Data Articles, September 2, 2021.

[26] s. kumar, "age anf gender detection, " galgotias university, 2020.

[27] Z. Saleh, Artificial Intelligence Definition, Ethics and Standards, The British University in Egypt.

[28] j. h. d. kirsch, machin learning for dummies, john wiley and , 2018.

[29] s. shwlev, understanding machin learning from theory to algorithms, cambridge university press, 2014.

[30] a. smola, "introduction to machin learning," college of engineering and computer science australian national university, australia, 2010.

[31] "introduction to deep learning," texas instru,ents, 2018.

[32] i. g. y. b. a. courville, "deep learning," errata.

[33] l. y, "neural networkes and introduction to deep learning," 2015.

[34] a. z. m. l. a. jsmola, "dive into deep learning," 25_3_2022.

ABSTRACT

In recent years, face recognition technology has become a hot topic in the field of pattern recognition. The human face is one of the most important human biometric characteristics, which contains a lot of important information, such as identity, gender, age and so on.

The goal of this work is to develop a system of age estimation, gender prediction and face mask detection by using convolutional neural networks based on a face image or real-time video. In this project, we created three CNN models with different architectures (the number of filters, the number of convolution layers...), the results obtained showed that the CNN significantly improved the performance of system as well the accuracy of the recognition.

Keywords: intelligent systems, biometrics, facial recognition, convolutional neural networks, gender prediction, age estimation, face mask detection.

ABSTRAIT

Ces dernières années, la technologie de reconnaissance faciale est devenue un sujet brûlant dans le domaine de la reconnaissance des formes. Le visage humain est l'une des caractéristiques biométriques humaines les plus importantes, qui contient de nombreuses informations importantes, telles que l'identité, le sexe, l'âge, etc.

L'objectif de ce travail est de développer un système d'estimation de l'âge, de prédiction du sexe et de détection de masque facial en utilisant des réseaux de neurones convolutifs basés sur une image de visage ou une vidéo en temps réel. Dans ce projet, nous avons créé trois modèles CNN avec des architectures différentes (le nombre de filtres, le nombre de couches de convolution...), les résultats obtenus ont montré que le CNN améliorait significativement les performances du système ainsi que la précision de la reconnaissance.

Mots-clés : systèmes intelligents, biométrie, reconnaissance faciale, réseaux de neurones convolutifs, prédiction du sexe, estimation de l'âge, détection du masque facial

ملخص

في السنوات الأخيرة، أصبحت تقنية التعرف على الوجوه موضوعًا ساخنًا في مجال التعرف على الأنماط. يعد الوجه البشري من أهم خصائص القياسات الحيوية للإنسان، حيث يحتوي على الكثير من المعلومات المهمة، مثل الهوية والجنس والعمر وما إلى ذلك.

الهدف من هذا العمل هو تطوير نظام لتقدير العمر والتنبؤ بالجنس واكتشاف قناع الوجه باستخدام الشبكات العصبية الاتفافية بناءً على صورة الوجه أو الفيديو في الوقت الفعلي. في هذا المشروع، أنشأنا ثلاثة نماذج CNN ببنيات مختلفة (عدد المرشحات، عدد طبقات الالتفاف ...)، وأظهرت النتائج التي تم الحصول عليها أن CNN حسنت أداء النظام بشكل كبير بالإضافة إلى دقة التعرف.

الكلمات المفتاحية: الأنظمة الذكية، القياسات الحيوية، التعرف على الوجه، الشبكات العصبية الاتفافية، التنبؤ بالجنس، تقدير العمر، كشف قناع الوجه.