People's Democratic Republic of Algeria

Ministry Of Higher Education and Scientific Research

University of 20 Août 1955-Skikda
Faculty of Sciences
Department of Computer Science

Thesis for obtaining the Master's degree in Computer Science

Specialty:

Systèmes Informatiques - SI

Theme:

Access Control System of "Cinq laboratoires" Using Face Recognition

Author:

Sabrina Bouzobra

Supervisors:

Dr. Soufiane Boulehouache

Dr. Adel Lahsasna

June 2022

Acknowledgements

"First and last Thanks to Allah"

Thank you to my supervisors Dr. ***Boulehouache Soufiane****, as well as* ***Dr. Lahsasna Adel****, for their guidance throughout this project.*

Thank you to my committee members,

Thank you to all my professors, especially Mr. Bouhouche, Mr. Kissmoune, Mr. Mazouzi, Mr.Alouche, Mr.Benoudina and Mr.Zghida for their participation in this work.

Thank you to my parents ***Yamina*** *and* ***Said*** *for your endless support, you have always stood behind me encouraging me all the time during these three last year until this big moment of reaching my goal and get my master degree. May Allah preserve you and save you.*

Thank you to my friend Mira, Wassila, Imane and my colleagues Safia, Lamia and Karima for their support and encouragement.

Thank you to my nieces ***Bissane*** *and* ***Razane*** *for making me happy when I was exhausted.*

I would like also to thank all those who helped me directly or indirectly to carry out this work.

Thanks to myself.

Sabrina

Abstract

Security in laboratories represents a big challenge and major concern for authorities and governments worldwide, and the use of traditional systems to address this issue has not solved the problem. Face recognition technology is the latest and one of the most effective authentication AI-based technologies to be adopted for keyless access control systems. In our current research work, an access control system was developed using deep learning to restrict access to the laboratory to authorized members only. The data set was collected from 09 subjects who are working in the department of computer science, university of Skikda. In order to avoid the bias in our testing results, the data set was resampled using five cross-validation method (5-cv). The results achieved by our system were presented along with discussions and analysis of the key findings and features of the system. The proposed system achieved an accuracy rate as high as 97%. Finally, some challenging cases like pose variations and face with occlusion were presented to show how our system could handle these cases efficiently. Where the results proved good efficiency.

Keywords: Access Control system, Biometrics, Face Recognition, Deep Learning, Convolutional Neural Network. Metric deep learning.

ملخـص

تمثل السلامة في المختبرات تحديًا كبيرًا واهتمامًا كبيرًا للسلطات والحكومات في جميع أنحاء العالم، واستخدام الأنظمة التقليدية لحل هذه المشكلة لم يعد كافياً. تقنية التعرف على الوجوه هي أحدث التقنيات و واحدة من أكثر تقنيات المصادقة القائمة على الذكاء الاصطناعي فاعلية والتي يتم اعتمادها لأنظمة التحكم في الوصول بدون مفتاح أو أي وسيلة مادية أخرى. في عملنا البحثي الحالي، تم تطوير نظام التحكم في الوصول باستخدام التعلم العميق لتقييد الوصول إلى المختبر للأعضاء المصرح لهم فقط. تم جمع مجموعة البيانات من 09 عناصر من الذين يعملون في قسم علوم الحاسب بجامعة سكيكدة. من أجل تجنب التحيز في نتائج الاختبار لدينا، تمت إعادة تشكيل مجموعة البيانات باستخدام طريقة التحقق الخماسي المتقاطع (5-cv). أين تمّ تقديم النتائج المتحصل عليها مع المناقشات وتحليل النتائج والخصائص الرئيسية للنظام. حيث حقق النظام المقترح معدل دقة يصل إلى 97٪، أخيرا تم عرض بعض التحديات التي تواجه النظام مثل اختلاف الوضعية، و تباين الوجه مع الانسداد لإظهار كيف يمكن لنظامنا التعامل مع هذه الحالات وكفاءة. ولإظهار كيف يمكن لنظامنا التعامل مع هذه الحالات. حيث اثبتت النتائج كفاءة جيدة.

Résumé

La sécurité dans les laboratoires représente un grand défi et une préoccupation majeure pour les autorités et gouvernements dans le monde entier, et l'utilisation des systèmes traditionnels pour résoudre ce problème n'a pas été suffisante. La technologie de reconnaissance faciale est la technologie la plus récente et l'une des plus efficaces technologies d'authentification basées sur l'IA, adopter pour les systèmes de contrôle d'accès sans clé. Dans notre travail de recherche en cours, un système de contrôle d'accès a été développé en utilisant l'apprentissage profond pour restreindre l'accès au laboratoire aux seuls membres autorisés. Le jeu de données a été collecté à partir du 09 sujets qui travaillent dans le département d'informatique, Université de Skikda. Afin d'éviter le biais dans nos résultats de test, l'ensemble de données a été ré-échantillonné à l'aide de la méthode de cinq cross- validation (5-cv). Les résultats obtenus par notre système ont été présentés avec discussions et analyse des principaux résultats et caractéristiques du système. Le système proposé atteint un taux de précision aussi élevé que 97%. Enfin, certains challenges comme la variation de pose, et visage avec occlusion ont été présentés pour montrer comment notre système pouvait gérer ces cas efficacement. Où les résultats ont prouvé une bonne efficacité.

Table of Content

List of Figures

List of Abbreviations

ACS	Access Control System
AI	Artificial Intelligence
ANN	Artificial Neural Networks
BACS	Biometric Access Control System
CMC	Cumulative Match Characteristic
CNN	Convolutional Neural Network
CPU	Central Processing Unit
DET	Detection Error Tradeoff
EER	Equal Error Rate
FAR	False Acceptance Rate
FFNN	Feed Forward Neural Network
FMR	False match rate
FNIR	False Negative Identification Rate
FNMR	False Non-Match Rate
FRR	False Rejection Rate
FTA	Failure To Acquire Rate
FTC	Failure To Capture Rate
FTX	Failure To Extract Rate
GMR	Genuine Match Rate
GPU	Graphical Processing Unit
HOG	Called Histogram Of Oriented Gradients
ICA	Independent Component Analysis
KNN	K-Nearest Neighbor
LDA	Linear Discriminant Analysis
LSTM	Long Short-Term Memory
ML	Machine Learning
MLP	Multilayer Perceptron
PCA	Principal Component Analysis
PIN	Personal Identification Number
RNN	Recurrent Neural Network
ROC	Receiver Operating Characteristic
SVM	Support Vector Machine
TAR	True Acceptance Rate
TIR	True Identification Rate
UML	Unified Modeling Language

List of Tables

General Introduction

The world has become more security conscious. That awareness extends to secure buildings, facilities, institutions such as banks, laboratories, etc. New guidelines and approaches are promulgated every year using different technologies and approaches to enhance the security in it» (Council, National Research, 2011, p. 256). Security in laboratory is a very big issue and challenge for authorities and government because of its importance. The level of the security increase depending on their specification. In a laboratory the theft of documents or equipment, which can be very expensive, or possible contamination by viruses, all of these can happen within a research laboratory, and can cause enormous losses. Hence safety in a research laboratory becomes an important topic. The case of Covid-19 is not far, where suspicions revolve around the source of the virus, which may have come out of a laboratory. Restraint access for persons that only have authorization to laboratory with an efficient system must be a solution to guarantee the security in there.

The classical systems used to access control like systems based-on object such as tokens include keys, Fobs and ID Cards, or systems based-on information such codes include password, pass-phrase and PIN number suffer from many issues which are for example magnetic cards can become corrupted or unclear, keys can be duplicated and passwords can be forgotten or guessed. So we have to use other solution instead of physical tools, or to combine them with other solutions to involve the security. The widespread of several technologies such as biometrics, artificial intelligence, image processing and computer vision open the door to several solutions to implement an efficient access control system for buildings.

Biometric access control can be the possible solution to overcome the drawbacks of these systems because it is one of the most popular types of security systems on the market, it combines security and convenience in a way that no other access control system can, However instead of using metal locks on doors, we will use some ‘’Biometric Security System’’ that is easily able to tell whether the person known or unknown can get the access and can easily identify them by scanning their face or another human characteristic such as iris, fingerprint, retina etc.

In 1960, scientists began identifying the physiological components of acoustic speech and phonic sounds. This was the precursor to modern voice recognition technology. - **The 1990s**—biometric science takes off. - **The 2000s—**rollout of biometric tech. **- 2020—**the

explosion. However, the COVID-19 outbreak had a moderate impact on the biometric technology market attributed to the rising fear of the spread of a virus through touching the surface of biometric devices such as fingerprint scanners and others. However, the demand for zero surface biometric devices such face recognition and others simultaneously, as consumers used these devices for authentication purposes without the fear of the spread of the virus.

Facial recognition is one of the biometric approaches that employs automated methods to verify or recognize the identity of a living person based on his physiological characteristics (IJCSMC, January 2013). Deep learning is a subfield of maching learning and AI, it can be used as an efficient technique for face recognition, it can solve some issues and challenges of this method and provide a good performance of the system, it is based on Deep Neural network.

Our goal is to develop a Biometric Access Control System in the laboratory "Cinq Laboratoires" located at the University of Skikda, using face recognition based on deep metric learning in order to restraint the access only for the authorizes persons "lab members" in real-time. Using identification type of recognition.

To reach our goal, this work employs different tools for each process component of the system which are enrollment (model generation) and recognition (identification): in the process of enrollment, we have used the HOG method for pre-processing step (face detection & alignment), in the feature extraction and encoding step (training) we have used CNN Architecture especially residual network (ResNet34). While in the recognition process in the classification step, we have used a simple machine learning algorithm, which, is KNN.

Our system recognizes a person who requests access, where it is the case, a sliding door will be open automatically and the access is granted, otherwise, the door will not be open and the access is denied.

In the literature, there are many approaches and methods based on different LCA, LDA, SVM, CNN...etc algorithms which give different results depending on the performance of the algorithm or also the quality and size of the database, and there are even methods combined between several solutions to increase the performance of the recognition system, for example combining methods such as PCA+LDA or PCA+SVM (Chapter 2 part1).

Our work is divided into two parts:

Part I: literature review: it consists of two chapters; the first one represents the biometric access control system with all necessary details found in the literature. The second one is about our proposed solution which is face recognition based on deep learning, with a good explanation of these concepts and how they work.

Part II: Conception and Implementation, consists of two chapters as well as, the first one is our conception of the proposed solution seen in the previous chapter with details to facilitate the implementation of our system. The second one is about all the tools used to implement the proposed system and the experimentation and evaluation of the implemented system, in addition, analysis and discussion of the results were represented to evaluate the performance obtained of this system.

Part I: Access Control System and Face Recognition: a review of literature

Chapter 1: Biometric Access Control System

1. Introduction:

This chapter aims to familiarize with three concepts, first, the access control system, its basics components such as entrance door as a part of it, then the access control system in the laboratory as a part of its security, and finally, we present the biometric access control system as the most promising option for recognizing individuals in the recent years, we will discuss with details the biometric technology, types of biometric technology, and the biometric access control system components, operation and architecture, we also discuss the biometric performance evaluation (verification and identification performance metrics), and we conclude the chapter with an overview on its application, market, current adoption, and trends of biometric access control system.

2. Access Control System:

2.1 Definition:

According to (Teh, Ling, & Cheong, 2013), « An access control system is simply defined as any technique used to control passage to or from any area or entrance, such as a residential area, an office and the like».

According to (techopedia.com, September 5, 2018), An access control system (ACS) is a type of security that manages and controls who or what is allowed entrance to a system, environment or facility. It identifies entities that have access to a controlled device or facility based on the validity of their credentials. An ACS is primarily a physical operation implemented within high security areas, such as data centers, government/military

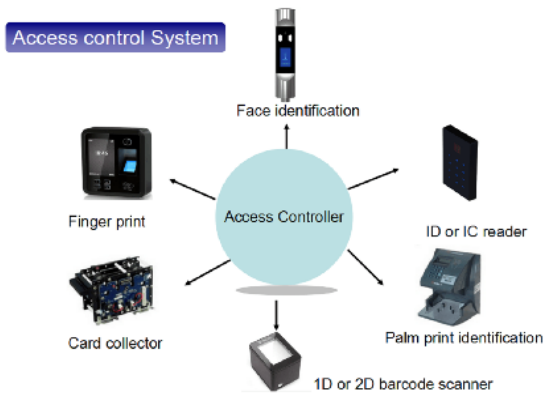


Figure 1: Access Control system Technologies

2.2 Access Control Systems Types:

1. What the user *has* (object-based):

The user must possess a specific object to be granted entry. Common types of object-based access tokens include: Keys, Fobs and ID Cards.

2. What the user *knows* (information-based):

The user must know a specific code to be granted entry. Common types of information-based access codes include: Password, Pass-phrase and PIN number.

3. Who the user *is* (biometric-based):

The user identifies themselves with their own body using biometrics. Common biometric modalities include Palm vein, Fingerprint, Iris scan, Face recognition...etc. (Karlskin, keyo, Sep 16, 2021).

2.3 Basic Components of an Access Control System:

(United Security usi, n.d., p. 16 April 2018)The most common basic components of an access control system are:

2.3.1 Access Cards:

In an access control system, access cards will take the place of keys. People will scan their access cards to gain access to the building or certain areas of the building. Each access card will have its own unique code, allowing you to control access for each.

2.3.2 Card Readers:

The card reader is the device that will read access cards in order to grant access. There are different types of card readers, some requiring card insertion, some requiring swipes, and some only needing cards to pass in proximity to the reader.

2.3.3 Keypads:

Access control keypads are another method of entry. Instead of scanning an access card, we will punch in a code on a numeric keypad. In order to gain entry, we must enter in the correct passcode. Keypads are sometimes used instead of card readers, and sometimes in conjunction with card readers.

2.3.4 Alarm Systems:

Fire alarms, burglary alarms, and intrusion detection alarms are often integrated with access control. If an unauthorized person attempts to enter the doors controlled by electric lock hardware, the access control system can signal the alarm to go off.

2.3.5 Field Panels:

Field panels are the control panels that connect all other parts of the access control system such as card readers, keypads, hardware, and more. Field panels are used to process access control activity for the whole building. The number of panels will be determined by the size of the building, the size of the system, and the extent to which the system is used.

2.3.6 Access Control Software:

The access control software is the brain of the entire system. It is the central database and file manager for the system. It records system activity and distributes information to and from the field panels in the building. This software runs on a traditional computer.

2.4 Entrance Identification Technologies:

(Kisi, nd)Access control systems can range from small, relatively simple one-door affairs to highly complex, computer-operated systems capable of handling hundreds of doors and tens of thousands of individually encoded identification credentials. A basic system usually consists of:

A. **A central processing unit (CPU):** Is the brains of the system and is programmed with data on each user.
B. **An input device at each protected door:** The most fascinating part of the system, and the different types of input devices used in today's systems are 1.Keypads, 2.Card readers and 3.Biometric readers.



Figure 2 Types of Readers

C. **An identification credential assigned to each user**: There are at least nine different card-encoding technologies available: magnetic stripe, Wiegand, proximity, barium ferrite, infrared, bar code, Hollerith, "smart" card, and optical storage.
D. **A printer:** is often included to provide a record of all activity.

2.4.1 Door Entry Systems:

(British security industry association, April 2016) There is often confusion between the terms door entry system and access control. This is in part because access control often includes door entry control as a feature and many systems sold as door entry systems include simple elements of access control.

Figure 3 Different door entry systems

3. Access Control System in laboratory:

Let us take for example an in-depth study of the safety system in a laboratory that handles hazardous chemicals. This study carried out by the National Research Council of the National Academies in Washington, approved by the Governing Board of the National Research Council and supported by the U.S. Department of Energy and others institutes and companies in 1981, and it was reviewed and updated several times till 2011 the last updated version. For us, and for our purpose, which is the access control system within laboratory, we will just focus on electronic security that include the access control system, in turn required a biometric technology.

3.1 Security in Laboratory:

3.1.1 Security Basics:

According to (Council, National Research, 2011), Security has become an important component of laboratory and it requires four integrated domains to improving it:

Physical or architectural security—doors, walls, fences, locks, barriers, controlled roof access, and cables and locks on equipment;

Electronic security—access control systems, alarm systems, password protection procedures, and video surveillance systems;

Operational security—sign-in sheets or logs, control of keys and access cards, authorization procedures, background checks, and security guards; and

Information security—passwords, backup systems, shredding of sensitive information».

3.1.2 Security Levels:

The choice and implementation depend on the level of security in laboratory, The following (tables) is one example of a management system for laboratory security, which illustrates how an institution or firm might set three security levels based on operations and materials.

- Normal (Security Level 1):

TABLE 10.1 Security Features for Security Level 1

Physical	• Lockable doors and windows
Operational	• Lock doors when not occupied • Ensure all laboratory personnel receive security awareness training • Control access to keys, use judgment in providing keys to visitors

Table 1: Security Features for Normal Level (Council, National Research, 2011)

- Elevated Security (Security Level 2):

TABLE 10.2 Security Features for Security Level 2

Physical	• Lockable doors, windows, and other passageways • Door locks with high-security cores • Separate from public areas • Hardened doors, frames, and locks • Perimeter walls extending from the floor to the ceiling (prevent access from one area to the other over a drop ceiling)
Operational	• Secure doors, windows, and passageways when not occupied • Ensure all laboratory personnel receive security awareness training • Escort visitors and contractors, consider an entry log
Electronic	• Access control system recommended • Intrusion alarm recommended where sabotage, theft, or diversion is a concern

Table 2: Security Features for Elevated Level (Council, National Research, 2011)

- High Security (Security Level 3):

TABLE 10.3 Security Features for Security Level 3

Physical	• Lockable doors, windows, and other passageways • Door locks with high-security cores • Separate from public areas • Hardened doors, frames, and locks • Perimeter walls extending from the floor to the ceiling (prevent access from one area to the other over a drop ceiling) • Double-door vestibule entry
Operational	• Secure doors, windows, and passageways when not occupied • Ensure all laboratory personnel receive security awareness training • Escort and log in visitors and contractors • Lock doors, windows, and passageways at all times • Inspect items carried into or removed from the laboratory • Have an inventory system is in place for materials of concern. • Perform background checks on individuals with direct access to the materials of concern or within the control zone.
Electronic	• Access control system that records the transaction history of all authorized individuals • Biometric personal verification technology recommended • Intrusion alarm system • Closed-circuit television cameras for entrance and exit points, materials storage, and special equipment

Table 3: Security Features for High Level (Council, National Research, 2011)

By comparing the three tables above, we can notice that the access control system is among the recommended features for security in both elevated and high levels, this last, in turn, requires also a biometric personal verification technology.

Comparing the physical security system using passwords and PINs, cards, tokens and so, we can notice that the password can be forgotten or guessed, for example magnetic cards can become corrupted or unclear, keys can be duplicated. However, an individual's biological traits cannot be misplaced, forgotten, stolen or forged.

Therefore, to ensure security in a laboratory, especially to restraint the access of individuals for authorized person only, we have to call a biometric technology in our access control system (entry door system).

4. Biometric Access Control System:

Nowadays the identification of individuals in different places and especially those whom are too sensitive and require a high level of security such as laboratory, is coming a great challenge whether for establishments, personal properties or for the security market.

4.1 Definition:

According to (Karlskin, keyo, Sep 16, 2021), « Biometric access control is one of the most popular types of security systems on the market, and for good reason: it combines security and convenience in a way that no other access control system can».

However instead of using metal locks on our doors, we will use some ‘’Biometric Security System’’ that is easily able to tell whether the person known or unknown can get the access and can easily identify them by scanning their face or another human characteristic such as iris, fingerprint, retina etc.

4.2 Biometric[1] Technology:

4.2.1 Definition:

(Heather Murray, 11 Dec 2007) Biometric technology is often described as the ultimate form of transnational surveillance, producing a readable body through a process of measurement, enrollment, and identification/verification.

4.2.2 Brief history:

The science behind scanning sensors improved to almost perfect. The use of smartphones went through the roof. Modern biometric technology began in the 1960s, evolving into high-tech scanners that read bio-markers with an accuracy touching 100%. In 2020, biology-based science is disrupting the authentication industry. (loginid.io, Feb 2021).

1960s—exploration.

In 1960, scientists began identifying the physiological components of acoustic speech and phonic sounds. This was the precursor to modern voice recognition technology.

In 1969, the Federal Bureau of Investigation (FBI) pushed for automated fingerprint identification which led to the study of minutiae points to map unique patterns and ridges.

1970s and 1980s—FBI funding.

[1] The word ‘biometrics’ is derived from the ancient Greek bios (life) and metrikos (measure). (Datta, Banerjee, & Madhura Datta, 2016)

By 1975, the first scanners to extract fingerprint points were prototyped, funded by the FBI. Digital storage costs were prohibitive, so the National Institute of Science and Technology (NIST) worked on compression and algorithms:

The work at NIST led to the development of the M40 algorithm, the first operational matching algorithm used at the FBI. Used to narrow the human search, this algorithm produced a significantly smaller set of images that were then provided to trained and specialized human technicians for evaluation. Developments continued to improve the available fingerprint technology. The NIST advanced speech, ocular, and face recognition by filing patents for iris identification and subcutaneous blood vessel patterns. Mugshots were digitized to databases.

The 1990s—biometric science takes off:

The 1990s was a boom time for biometrics. Algebraic equations revealed that less than one hundred values could differentiate normalized face images. The National Security Agency (NSA) formed the Biometric Consortium. The Department of Defense (DoD) partnered with the Defense Advanced Research Products Agency (DARPA) to fund face recognition algorithms for commercial markets

The 2000s—rollout of biometric tech:

West Virginia University (WVU) established the first bachelor's program in Biometric Systems and Computer Engineering,.. he European Biometric Forum was established to address market adoption and fragmentation barriers. Face recognition became the accepted global biometric authenticator for passports and other Machine Readable Travel Documents (MRTDs).

2020—the explosion:

Over the last 60 years, the barriers that made biometric authentication inaccessible for consumers, enterprises, and even code developers were removed. Then, something happened in 2020, momentum the moving parts came together, catapulting biometric authentication into the mainstream.

4.2.3 Types of Biometric Technology:

There are several biometric techniques, which have been analyzed and evaluated, therefore each has their strengths and limitations, and while they can be used in different applications, biometrics have been often used in security. We can mostly label biometrics into three groups:

A. Biological biometrics:

Use traits at a genetic and molecular level. These may include features like DNA or your blood, which might be assessed through a sample of your body's fluids.

B. Morphological biometrics:

Involve the structure of your body. More physical traits like your eye, fingerprint, or the shape of your face can be mapped for use with security scanners.

C. Behavioral biometrics:

Are based on patterns unique to each person. How you walk, speak, or even type on a keyboard can be an indication of your identity if these patterns are tracked.

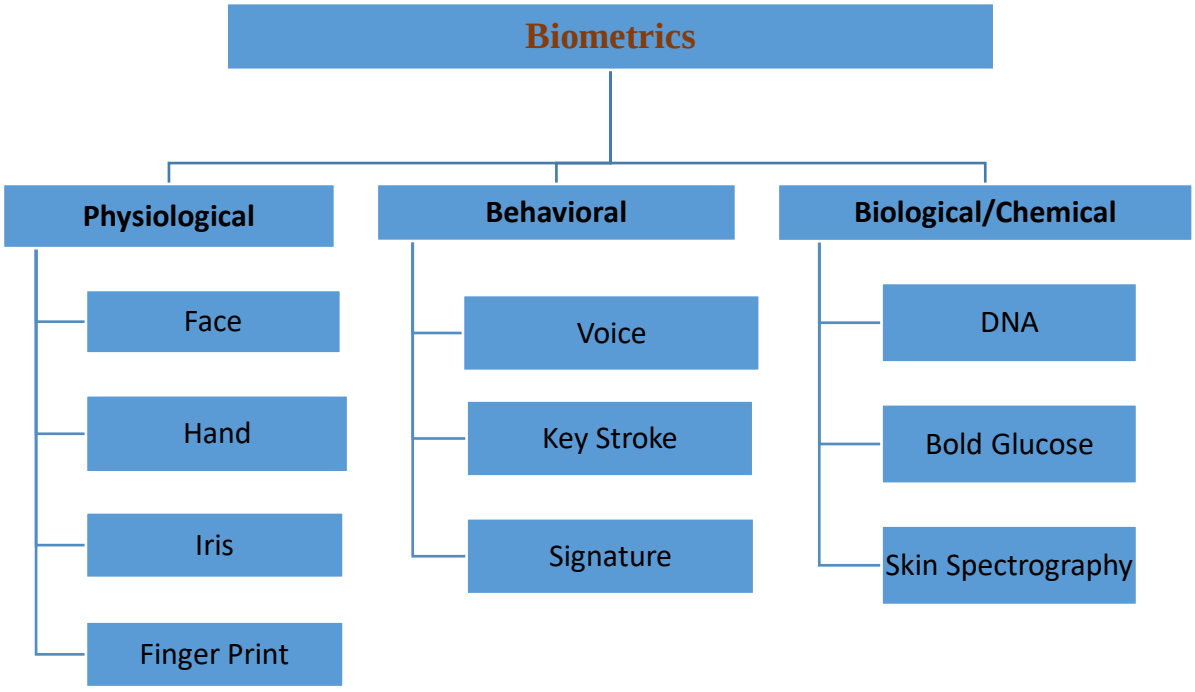


Figure 4: Types of Biometric Technology

4.2.4 Biometric Characteristics:

(Damer, April 2018), Modern biometric recognition, as seen from a computer science perspective, is define das the automated recognition of individuals based on their behavioral or biological characteristics. In order to choose a certain characteristic, it has to be evaluated in terms of the desired properties. These properties are:

- **Universality:** a biometric system has to be designed so that it is able to cover the largest possible ratio of the population. Related problems can include the absence or degraded quality of a certain biometric characteristic in a number of the population.

- Uniqueness: a biometric system has to assure to represent different individuals in highly distinct manner.

- Performance: the decision errors produced by biometric system is minimized and the computational efficiency is maximized.

- Permanence: the performance of a biometric system should be consistent over time. The ageing of certain biometric characteristics is a major challenge.

- Collectability: the biometric characteristics should be measurable and the quantitative results are reproducible.

- Acceptability: the convenience for the user is considered and the usability is maximized.

- Circumvention: collecting and replicating a fake biometric sample/template is hard.

4.3 Biometric Technology Modalities:

Biometric system depends on the biometric feature used to recognize an individual. The most common are:

4.3.1 Fingerprint Recognition:

Fingerprint scan is the most widely used biometric technology. Fingerprint (optical, silicon, ultrasound, touch less) uniqueness can be defined by analyzing the trivia of a human being. Trivia include sweat pores, distance stuck between ridges, bifurcation. It is probable that the likelihood of two individuals having the same fingerprint is less than one in billion. There are several sub-methods in fingerprinting, with changeable degrees of accuracy and correctness. Various can even detect when a live finger is present. Fingerprinting method has been developed over the years.

Figure 5: Finger print samples

Advantages: Very high accuracy, non-invasive biometric technique. Most economical biometric PC user authentication technique, it is one of the most developed biometrics, Easy to use, Small storage space required for the biometric template and also reduces the size of the database, It is standardized.

Disadvantages: For some people it is extremely intrusive, because is at rest related to criminal verification, it can be compose mistakes with the dryness or dirty of the finger's skin, as well as with the age (is not appropriate with children, because their fingerprint changes quickly),

Image captured at 500 dots per inch (dpi). Resolution: 8 bits per pixel, it demands a large memory space. (Bhatia, Biometrics and Face Recognition Techniques, May 2013)**.**

4.3.2 Voice Recognition:

Voice recognition technology does not measure the visual features of the human body. In voice recognition sound sensations of a person is measured and compared to an existing dataset. The person to be identified is usually required to speak a secret code, which facilitate the verification process.

Figure 6: Voice sample

Advantages: Non intrusive, high social capability, less verification time is about five seconds and not expensive technology**.**

Disadvantages: A person's voice can be easily recorded and used for unauthorized PC or network, Low accuracy, an illness such as a cold can change the voice of a person, which makes identification difficult or impossible. (Bhatia, Biometrics and Face Recognition Techniques, May 2013).

4.3.3 Signature recognition:

Signature recognition is the process used to recognize an individual's hand-written or signature. Dynamic signature verification technology uses the behavioral biometrics of a hand written signature to confirm the identity of a computer client. Analyzing the speed, shape, stroke, and pen pressure and timing information during the act of signing natural does this.

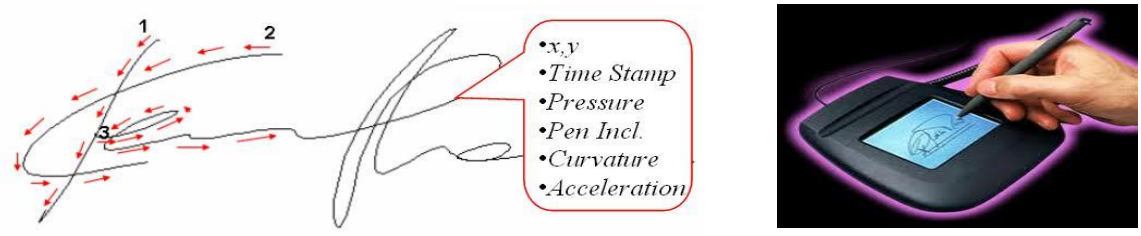


Figure 7: Signature sample

Advantages: Non-intrusive, less time of verification about 4 to 5seconds, inexpensive technology.

Disadvantages: Error rate: 1 in 50. (Bhatia, Biometrics and Face Recognition Techniques, May 2013)

4.3.4 Palm recognition:

In palm recognition, a 3-dimensional image of the hand is collected and the feature vectors are extracted and compared with the database feature vectors. These devices are bulky but identification is done in a short time.

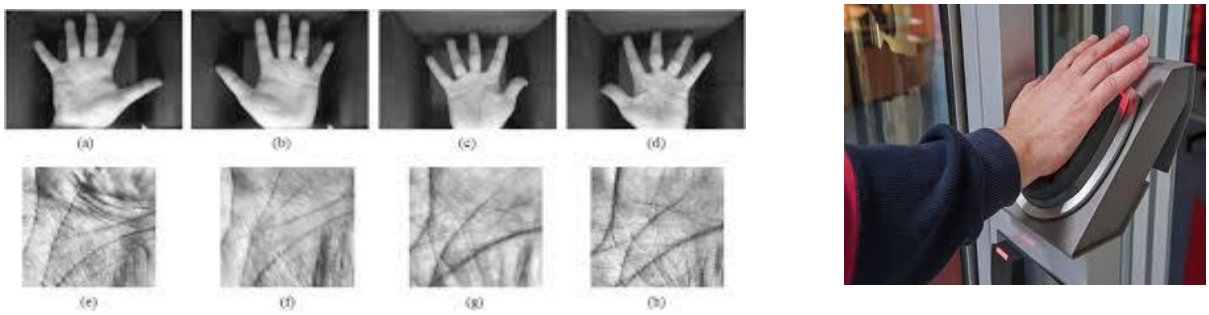


Figure 8: Palm sample

Advantages: can achieve good performance in terms of speed and accuracy and can have practical application to personal identification and verification as a new biometric technology. (Dr.hlaing-Htake-Khaung-Tin, December 2012).

4.3.5 Hand Geometry:

Hand geometry has 3-D image of top, sides of hand and fingers is collected, and the feature vectors are extract and compared with the dataset feature vectors.

It is recognition devices are bulky but identification is done in a 2-3 second. User places hand, palm-down, on an 8 x 10 metal surface with five guidance pegs. Pegs confirm that fingers are positioned correctly and also verify correct hand position.

Figure 9: Hand geometry technology

Advantages: It requires special hardware; it can be easily integrated into other devices or systems, It has no public attitude problems as it is associated most commonly with authorized access, a large amount of data are stored in database to uniquely identify a user, allow it to be used with Smartcards.

Disadvantages: Very expensive, Considerable size, it is not valid for arthritic person; they cannot put the hand on device. (Bhatia, Biometrics and Face Recognition Techniques, May 2013)

4.3.6 Iris scan:

The iris scans process start to get something on film. For this a specialized camera is required, naturally very close to the subject, not above three feet, uses an infrared imager to illuminate the eye and capture a very high-resolution photograph. Complete process takes only few seconds (approximately 1 to 2 sec) and provides the details of the iris knowingly produce, recorded and stored in dataset for future identification and verification. The quality of iris image does not get affected due to the presence of the contact lens and eyeglass.

Advantages: Very high accuracy, Verification time is generally less than 5 seconds, The

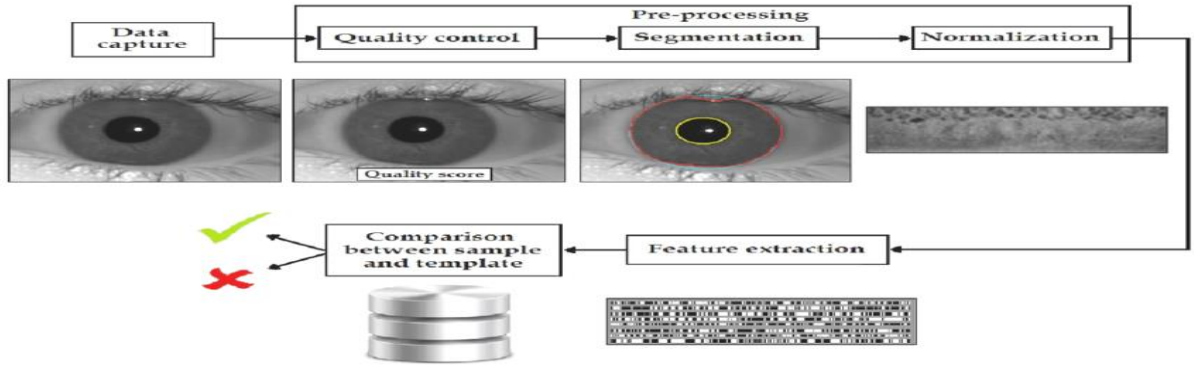


Figure 10: Iris recognition Technology

eye from a dead person would deteriorate too speedy to be valuable, so no extra protection have to been taken with retinal scans to be sure the user is a living human being.

Disadvantage: Too much movement of head or eye, wear colored contacts. (Bhatia, Biometrics and Face Recognition Techniques, May 2013)

4.3.7 Face Recognition:

Facial recognition is a form of computer vision that uses faces to attempt to identify a person or verify a person's claimed identity.

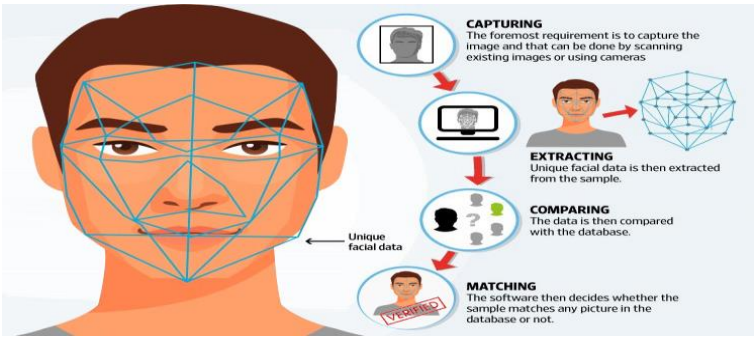


Figure 11: Face Recognition System

Facial recognition is including five steps to complete their process (Bhatia, Biometrics and Face Recognition Techniques, May 2013):

Step1: acquiring the image of an individual's face: there are two ways to acquire image: 1) Digitally scan an existing photograph, 2) Acquire a live picture of a subject.

Step2: locate image of face: software is used to locate the faces in the image that has been obtained.

Step3: analysis of facial image: software measures face according to is peaks and valleys; focuses on the inner area of the face identified as the "golden triangle", valleys are used to create a face print with their nodal points.

Step4: comparison: the face print created by the software is compared to all face prints the system has stored in its database.

Step5: match or no match: software decides whether or not any comparisons from step 4 are close enough to declare a possible match.

Advantages: the main motivation for face recognition is that, it is considered fast, a passive, non-intrusive system to verify and identify people.

Disadvantages: there are some issues of face recognition system such as illumination, pose, occlusion, ageing, interclass similarity (cannot distinguish twins or relatives)...etc.

While there are other technologies such as: retina scan, Vein pattern recognition, Gait recognition, Thermogram, Keystroke recognition, Ear geometry recognition ,Skin reflection recognition, Lip motion recognition, Body odor recognition…

Acquire Image → Locate Face → Face Detection → Face Recognition → Person

- The table below summarizes the characteristic features of biometric technology

Characteristics	Fingerprints	Hand Geometry	Retina	Iris	Face	Signature	Voice
Easy of Use	high	high	Low	Medium	Medium	High	High
Error Incidence	Dryness, dirt, age	Hand injury, age	Glasses	Lighting	Lighting, age, glasses, hair	Changing signature	Noise, colds
Accuracy	High	High	Very high	Very high	High	High	High
User Acceptance	Medium	Medium	Medium	Medium	Medium	High	high
Long Term Stability	High	Medium	high	high	Medium	Medium	Medium

Table 4: Characteristic Feature of Biometric Technology

- The table below shows the evaluation of each biometric characteristic in terms of the desired properties : (Bhatia, Biometrics and Face Recognition Techniques, May 2013)

Biometrics	Universality	Uniqueness	Permanence	Collectability	Performance	Acceptability	Circumvention
Face	H	L	M	H	L	H	L
Fingerprint	**M**	**H**	**H**	**M**	**H**	**M**	**H**
Hand Geometry	M	M	M	H	M	M	M
Keystrokes	L	L	L	M	L	M	M
Hand veins	M	M	M	M	M	M	H
Iris	**H**	**H**	**H**	**M**	**H**	**L**	**H**
Retinal scan	**H**	**H**	**M**	**L**	**H**	**L**	**H**
Signature	L	L	L	H	L	H	L
Voice	M	L	L	M	L	H	L
Face	H	H	L	H	M	H	H

Table 5: Comparison of various biometric traits

4.4 Biometric System:

4.4.1 Definition:

A Biometric system is fundamentally a pattern recognition system that recognizes a person by determining the authentication by using his different biological features i.e. Fingerprint, retina-scan, iris scan, hand geometry, and face recognition are leading physiological biometrics and behavioral characteristic are Voice recognition, keystroke-scan, and signature-scan.

4.4.2 Biometric System Components:

Biometric system consists of both hardware and software elements. Hardware generally includes electronic components and sensors to be able to read data out of specific patterns, software portion makes use of algorithms to enhance and recognize this data to generate a template unique to the individual it comes from. (bayometric.com, n.d.).

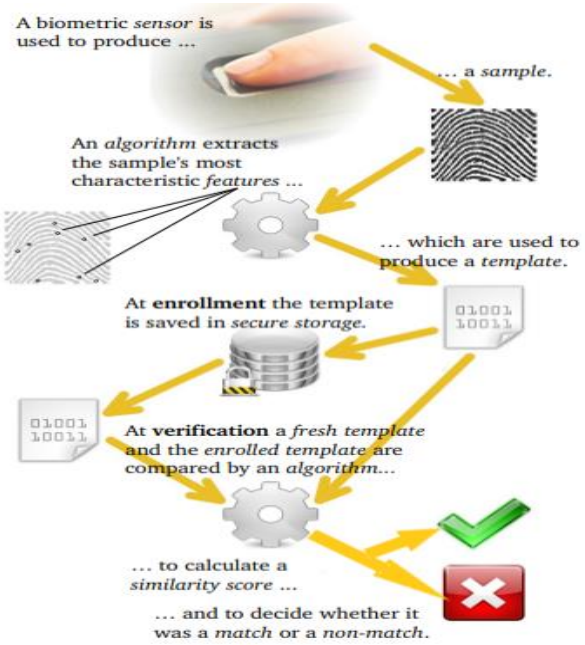


Figure 12: Biometrir system Operations

4.4.3 Biometric System Operations:

Biometric system is designed specifically to map a particular biometric trait, i.e. fingerprint recognition systems cannot process iris or retina patterns. However, all biometric systems work on the principle of mapping patterns with the help of technology. A person's physiological or behavioral patterns are mapped, stored and compared later to verify or identify the owner of the pattern.

Each biometric system comprises two distinct processes, Enrollment process and Recognition (Authentication) process. (depicted in Figure 13).

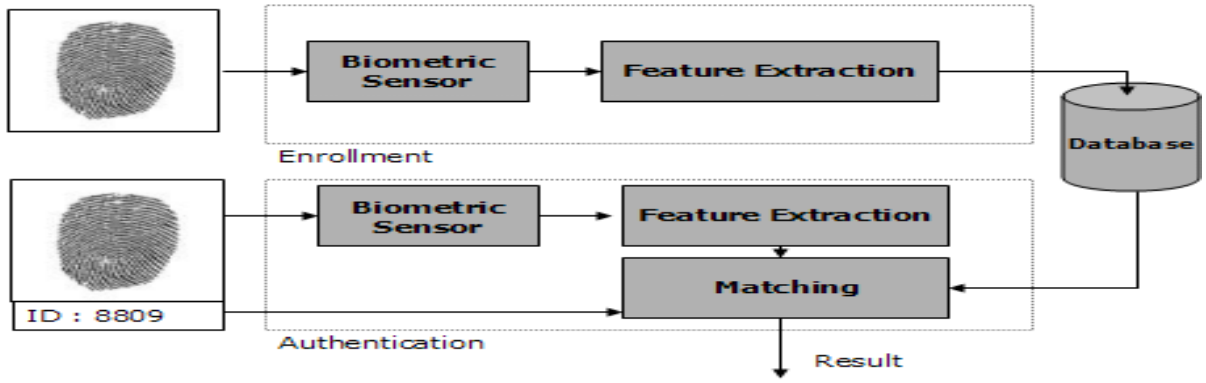


Figure 2: Biometric System Processes

A- Enrollment Process:

A biometric sensor captures the biometric characteristics. In a typical biometric work-flow, a set of features are then extracted from each captured sample. This feature set (template) is stored in a database as a reference. (Damer, April 2018)

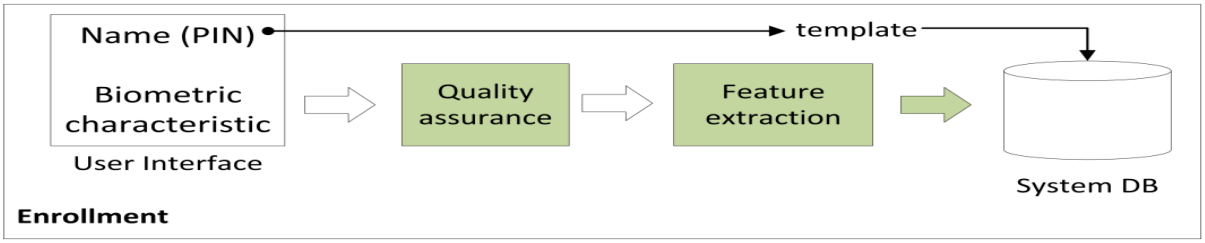


Figure 14: Enrollment operation

B- Recognition/Authentication Process:

Features are extracted from the captured sample to be recognized (fresh template) are compared to a previously acquired biometric feature set (reference template) stored in a database. A comparison results in a similarity score. Analyzing the similarity score (or a set scores) results in a biometric decision. A biometric decision, and thus a biometric system, can be a verification one or an identification one.

- **B1. Biometric verification:**

Is the use of biometrics information to verify a person claimed identity. The identity can be claimed using a smart card, a user name, or an identification number. Here, the system will verify that the claimed identity belongs to the user by comparing his/her biometric characteristics with a stored (and associated to this identity) biometric template. Therefore, the comparison (similarity measure) is only performed once for each identity claim and thus, the verification problem is usually referred to as a **1 : 1** comparison process. (Damer, April 2018)

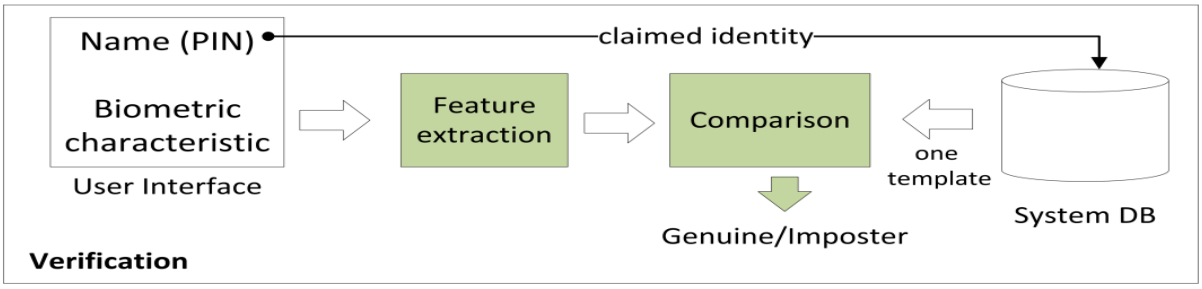


Figure 15: Verification process

- **B2. Biometric identification:**

Tries to assign an identity to an unknown person based on the captured biometric characteristics. This requires comparing the captured biometric characteristic to all the enrolled subjects. Therefore, the identification is referred to as a **1 : N** comparison process, where N is the number of the enrolled subjects. Biometric identification can be categorized into open-set and close-set identification:

• Close-set identification system:

Is confident that the captured subject is one of the enrolled subjects, and thus can report the best matching identity as the identified one.

• Open-set identification:

Implies that the system does not guarantee that the captured subject is already enrolled. Therefore, the open-set identification system has to verify one of the top matched identities to be the identity of the captured subject or to point out the fact that the subject is not enrolled.

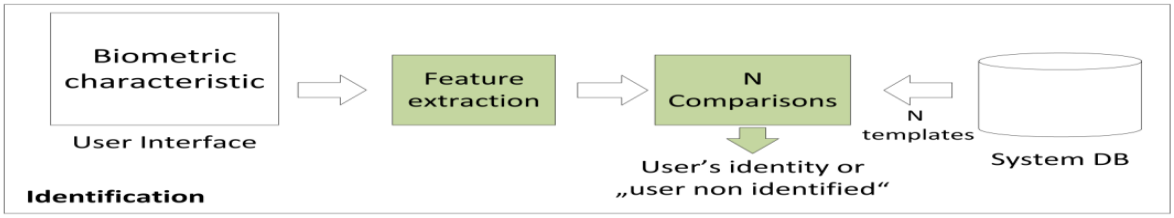


Figure 16: Identification process

To include a subject in a database, an enrollment process has to take place. Enrollment includes providing a trusted identity, capturing and quality proofing a biometric characteristic, feature extraction, and finally storing the identity information and the extracted features (template) in a database. (Damer, April 2018)

- **Verification Vs Identification:**

Verification is normally faster than identification and more secure but requires multiple reading technologies at the access point. Care must be taken with biometric readers where high security is required that there is an acceptably low FAR (False Acceptance Rate) for identification or low FMR (False Match Rate) for verification. (bsia, British security industry association, April 2016).

5. Biometric System Architecture:

It consist of the following modules (Guennouni, Mansouri , & Ahaitouf, October 19th, 2018).

5.1.1 The capture module:

That represents the entry point of the biometric system and consists in acquiring the biometric data in order to extract a digital representation. This representation is used later in the following phases of the system.

5.1.2 The module of signal processing:

Makes it possible to optimize the processing time and the digital representation acquired in the enrollment phase in order to optimize the processing time of the verification phase and the identification.

5.1.3 The storage module:

That contains the biometric templates of the system enrollees.

5.1.4 The matching module:

That compares the data extracted by the extraction module with the data of the registered models and determines the degree of similarity between the two biometric data.

5.1.5 The decision module:

That determines whether the similarity index returns through the matching module is sufficient to make a decision about the identity of an individual.

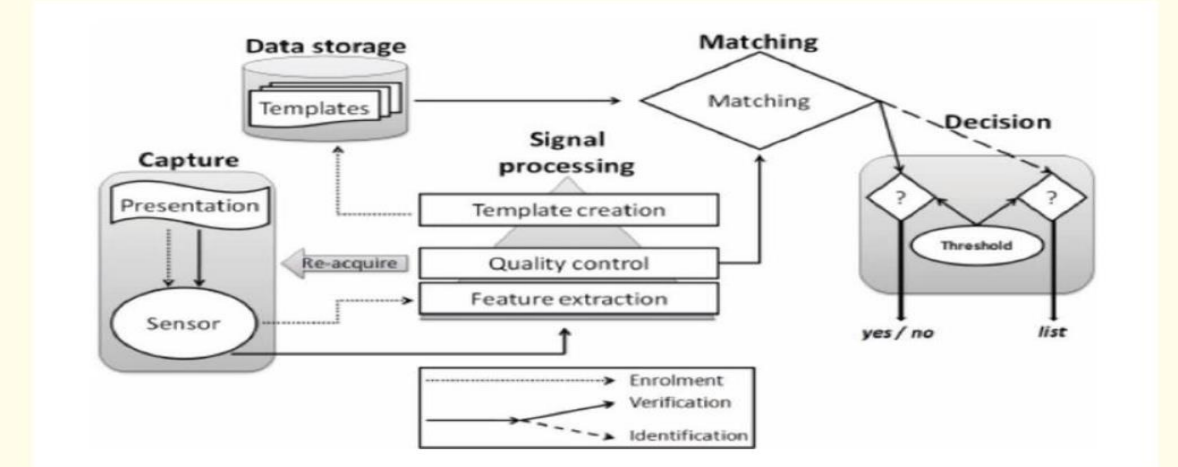


Figure 17: Biometric System Architecture

6. Biometric Performance Evaluation:

(Damer, April 2018)While most biometric characteristics are theoretically discriminant, automatic biometric systems make wrong verification or identification decisions in some cases. To build a performance comparison between different biometric systems, a set of performance metrics are defined for different operational scenarios. This section presents the relevant metrics and is largely based on the international standard ISO/IEC 2382-37:2012.
Some of the errors in biometric systems occur in an early stages of their operation, namely in the capture or feature extraction stages. These errors are:

- **Failure to capture rate (FTC):** proportion of failures of the biometric capture process to produce a captured biometric sample that is acceptable for use.
- **Failure to extract rate (FTX):** proportion of failures of the feature extraction process to generate a template from the captures sample.
- **Failure to acquire rate (FTA):** proportion of verification or identification attempts for which the system fails to capture or locate an image or signal of sufficient quality.

FTA is a result of FTC or FTX and can be given as: **FTA = FTC+FTX∗(1−FTC).**
The FTC, FTX, and FTA are related to the performance of the complete biometric system and thus are usually neglected when measuring the performance of the biometric algorithms. Performance metrics of the biometric algorithms can be categorized depending on the operational mode, verification or identification.

6.1 Verification performance metrics:

Biometric verification builds its binary decision (genuine/imposter) by thresholding the level of similarity between the probe and the reference samples (comparison score). Assigning a threshold value is a tradeoff between security and convenience. Security requires a low rate of imposter users accepted as genuine (false positives) and convenience requires a low rate of genuine users rejected as imposters(false negatives). This tradeoff can be summarized by the following biometric verification metric:

- **False match rate (FMR):** rate of zero-effort imposter attempt samples falsely verified as genuine (match) to the compared sample.
- **False non-match rate (FNMR):** rate of zero-effort genuine attempt samples falsely verified as imposter (non-match) to the compared sample.

The genuine match rate (GMR) is another metric used in the literature and points out to the same properties as the FNMR.

- The GMR is given by $GMR = 1-FNMR$.

The discussed FMR, FNMR, and GMR are error rates that describe the algorithm verification performance. Therefore, they do not consider errors introduced by the whole system, i.e. FTA. A set of similar verification metrics that theoretically consider FTA errors is also used in the literature, these metrics are the false acceptance rate (FAR) and false rejection rate (FRR) and are given by

$$\textbf{FAR} = FMR*(1-FTA) \text{ and } \textbf{FRR} = FTA+FNMR*(1-FTA).$$

The true acceptance rate (TAR) is derived from the FRR and is given as $\textbf{TAR} = 1-FRR$

In the literature FAR and FMR are often used interchangeably as well as FNMR and FRR. The only difference is that FAR and FRR consider samples failed to be acquired.

The so-far discussed verification metrics are dependent on the comparison score threshold used to make the genuine/imposter decision. This threshold builds a tradeoff between the FAR and FRR (FMR and FNMR). Therefore, the performance of biometric algorithms is only comparable by fixing one of both error rates and comparing the other.

To avoid this and provide one generalized metric, the equal error rate (EER) is used. EER is the common value of FAR or FRR at the operational point (threshold) where they are equal. This tradeoff between error rates are interesting because different applications require different levels of security and convenience from the same biometric algorithm/system.

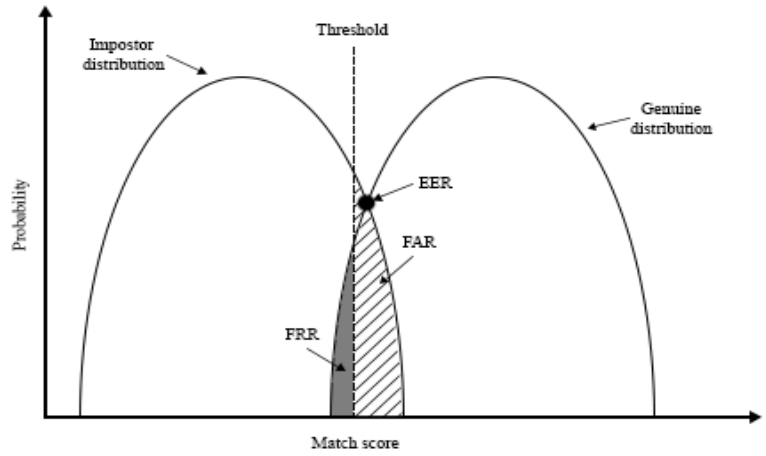


Figure 18: Verification performance metrics

To present a wide range of performance operation points, receiver operating characteristic (**ROC**) and detection error tradeoff (DET) curves plots the performance at different decision threshold points. An ROC curve plots FAR (x-axis) versus TAR (y-axis) (or FMR vs. 1-FNMR) at all possible decision thresholds. This allows the system integrator to informatively choose the threshold that best fits the required security convenience tradeoff.

A **DET** curve provides the same information as an ROC curve by plotting FAR (x-axis) versus FRR (y-axis) (or FMR vs. FNMR).

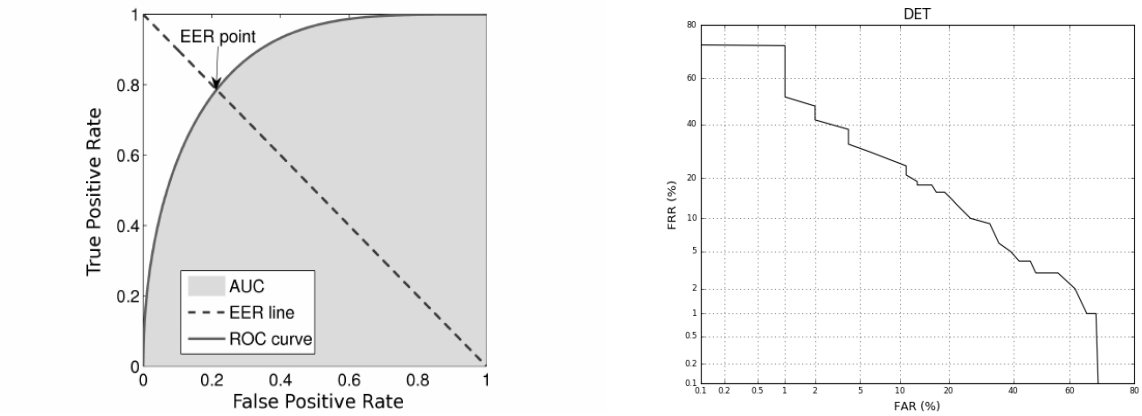


Figure 19: performance at different decision threshold points using ROC and DET

6.2 Identification performance metrics:

As we have seen before, biometric identification systems are categorized into: open-set identification and close-set identification.

The cumulative match characteristic (CMC) curve is used in the literature to present the close-set identification performance. CMC plots the number of considered top ranks (x-axis) versus the ratio of identification processes where the correct identity was found in the considered top ranks (x-axis). This ratio will be referred to as the true identification rate (TIR) at a certain rank r.

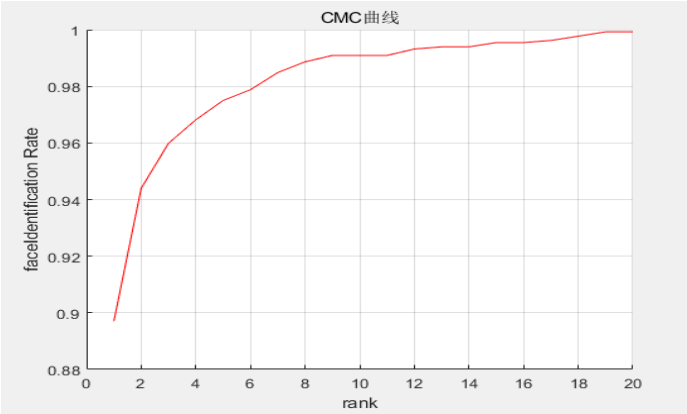


Figure 20: close-set identification performance

Open-set identification implies verifying if one of the identities in the reference database belongs to the probed subject. This leads to building a performance metric based on the verification performance of the biometric system and the size of the reference database. Two metrics with a verification-like tradeoff are usually used, the false positive identification rate (FPIR) and the false negative identification rate (FNIR). While considering the FTA, FPIR for a reference database containing N records is given by:

$$FPIR = (1-FTA)*(1-(1-FMR)N)$$

And for a zero FTA and a typically small FMR, FPIR can be approximated to:

FPIR = N∗FMR.

Similarly, **FNIR** can be approximated to:

FNIR = N∗FNMR

Therefore, the errors of an open-set identification system grow linearly with the reference database size. Along with the demand for large scale biometric systems, the motivation for achieving lower verification error rates is higher than ever. This drives the utilization of multi-biometrics to enhance performance, which is a key step in enabling large-scale biometric systems. (Damer, April 2018)

7. Applications of biometric systems:

Biometric systems can be used in a large number of applications. The following domains use biometric solutions to meet their respective needs:

- **Legal applications:**

Justice and law enforcement: Biometric technology and law enforcement have a very long history. 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.

- **Government applications:**

Border control and airport: 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.

Healthcare: 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.

- **Commercial applications:**

Security: 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.

Finance: Among the most popular applications of biometric technology, financial identification, verification, and authentication in commerce help make banking…

8. Biometric Market:

According to (Sava, Feb 25, 2022), The global biometric system market is forecast to amount to about 36.6 billion U.S. dollars in 2020. The market is expected to grow rapidly in the coming years, reaching a size of 68.6 billion U.S. dollars by 2025.

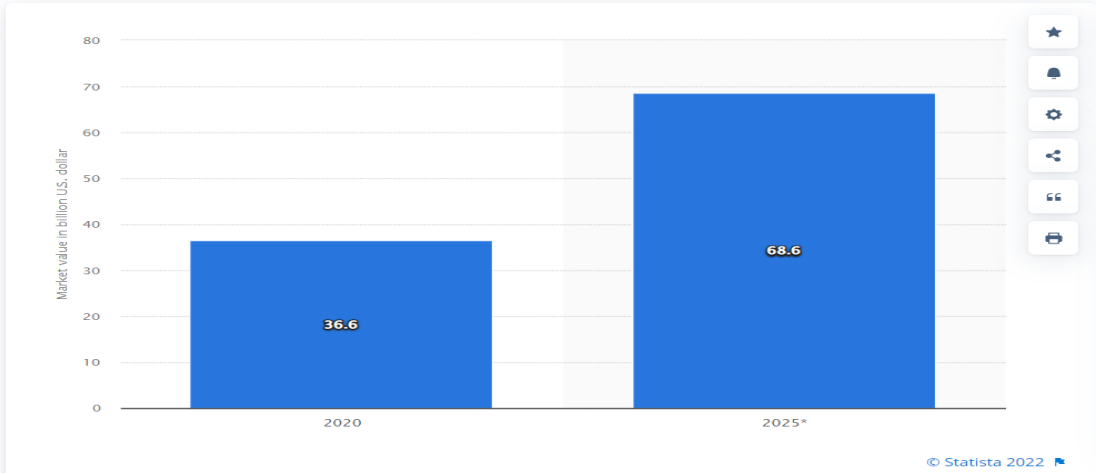


Figure 21: Biometric Market value- Forecast (2020-2025)

9. Current Biometric Adoption and Trends:

Less than a decade ago, consumers still feared biometric applications as clandestine extensions of government and law enforcement. Business initiatives relying on biometric applications once failed across market sectors, but that trend seems to be changing as younger consumer generations are now surrounded by smartphones and wearables. Consumer biometric acceptance and adoption by consumers seems to be changing as more companies use biometrics to identify and authenticate users.

Research on consumer attitudes about using biometric authentication has been mixed. The ease of using a biometric for authenticating your identity in daily transactions is a strong incentive for consumers who are inconvenienced by the need to memorize difficult, secure passwords. On the other hand, studies have found that privacy concerns were a strong influence on consumers' reluctance to use biometric authentication systems[2]. Despite some consumer concerns about the privacy of the biometric data, biometric authentication is a rapidly accelerating market and forecasts predict this growth to continue.

[2] 1 Clodfelter, R. (2010). Biometric technology in retailing: Will consumers accept fingerprint authentication? Journal of Retailing and Consumer Services, 17(3), 181-188; Morosan, C. (2012). Voluntary Steps toward Air Travel Security: An Examination of Travelers' Attitudes and Intentions to Use Biometric Systems. Journal of Travel Research, 51(4), 436-450.

According to (German & Barber, Sep 2017), « …we located and analyzed 53 different marketplace and government applications of biometric technology in order to describe various aspects of their usage. In this report, we analyze the following characteristics of biometric adoption: »

9.1 Biometric Use:

Fingerprints are still most common, used in such disparate settings, for employee cash register access, Bank, for ATM transactions, as well as for purchasing authentication at several campus retailers. Face recognition is highly prevalent as well, used for unlocking mobile apps and searching databases.

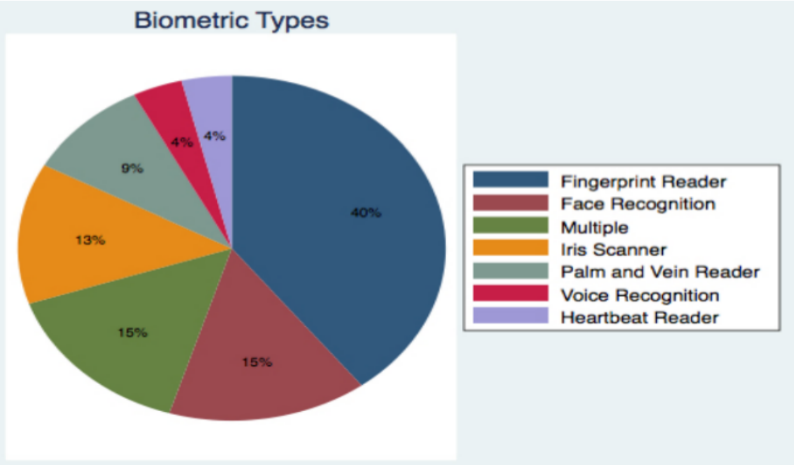


Figure 22: use of types of biometrics

- User Domain:

User Domain refers to the general domain of services and activities in which users employ biometrics to identify themselves. This differs from 'Market Sector' in that the user domain is a broad area of activity including workplace and governmental interactions.

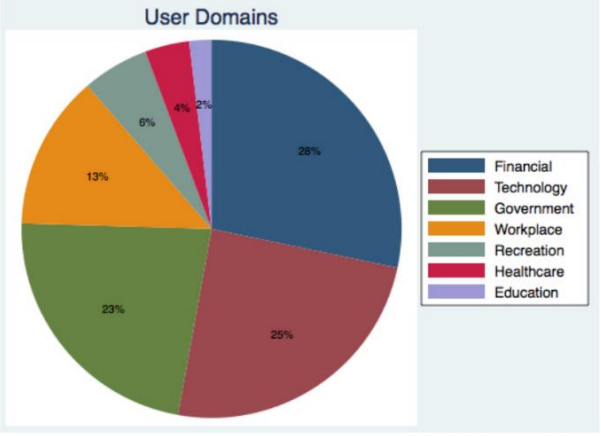


Figure 23: Classification of Biometrics by User domain

- Platforms:

The Figure bellow shows variation among the different platforms used to authenticate individuals. Due to the nature of certain services and access privileges, much of the current biometric authentication is performed on-site. From worker identification to in-store purchasing, more and more entities are utilizing the benefits of biometric technology to identify those with access to their systems. But mobile device usage is on the rise.

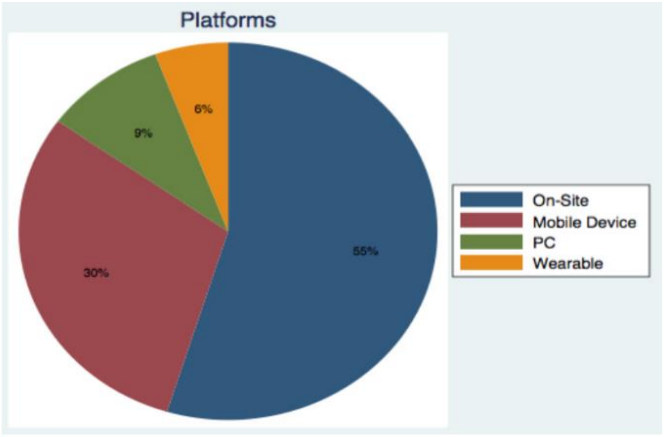


Figure 24: Classification of Biometrics by Platforms

- Targeted Activities:

Marketers and innovators are rolling out new ways to use biometrics on a regular basis. Figure below shows the various activities targeted for biometric authentication. Security for personal devices is a common use for biometrics in today's market. iProov's Verifier app, available on both Apple and Android devices, allows users to authenticate and unlock their devices via facial recognition, financial sector authentication, and government services are on the rise. In addition, healthcare providers and recreational use biometric technology to track individuals throughout their systems and properties.

Figure 25:Biometric Classification by Targeted Activities

10.Conclusion:

This chapter discussed the general properties of a biometric system. It also presented an overview on multi-biometric technologies supported by a look at the related works in the literature, including some strong and weak points of each technology. We also learned about the architecture of a biometric system and the evaluation of its performance where we have introduced the concept of verification and identification and metrics commonly used to describe its performance (verification and identification metrics), furthermore we presented different domains of application of biometric systems and an overview of the biometrics market.

Based on the literature report of (German & Barber, Sep 2017), we can conclude that the rapid expansion of biometric technologies has led to a similar explosion in biometric services and applications. Fingerprint scanners are still most common, with face recognition and iris scanners not far behind. The majority of uses of biometric authentication are performed on-site, but mobile device usage is a fast growing area as well.

However, the COVID-19 outbreak had a moderate impact on the biometric technology market attributed to the rising fear of the spread of a virus through touching the surface of biometric devices such as fingerprint scanners and others. However, the demand for zero surface biometric devices such face recognition, and others simultaneously, as consumers used these devices for authentication purposes without the fear of the spread of the virus. (AlliedMarketResearch.com, n.d.).

According to (Patel, Rathod, & Shah, November 2012), " Facial feature recognition is one of the most effective, highly authenticated and easily adaptable biometric security systems". Face recognition has the benefit of being a passive, non-intrusive system to verify personal identity in a "natural" and friendly way.

In the following chapter, we will discuss in detail the face recognition system using deep learning as method of authentication of individuals On-Site, in real time.

Chapter 2: Face Recognition and Deep Learning

1. Introduction:

This chapter consists of two parts, the first part presents face recognition system with details, such as architecture, methods and applications, then we have reviewed two states of the art of different face recognition systems using different algorithms, and then we compared the accuracy of these systems. While the second part presents deep learning as a sub-field of machine learning, in turn sub-field of artificial intelligence, that is required to implement face recognition system with other subfield of A.I like computer vision. We have focused on artificial neural networks (ANN) and its different architectures, especially CNN architecture, because it is the most performant architecture in face recognition and it is the state of the art for researchers nowadays.

2. Face Recognition:

2.1 Definition:

(IJCSMC, January 2013), Face recognition is a biometric approach that employs automated methods to verify or recognize the identity of a living person based on his physiological characteristics.

(Richa & Josan, January 2015), Face Recognition is the task of identifying the detected face as a known face or not. It is the application of Digital Image Processing and Computer Vision. Computer Vision is an advance branch of Artificial Intelligence and it is attained by Machine Learning. So, we can say that Face Recognition System is implemented with the help of Machine Learning, Computer Vision and Image processing.

2.2 Face Recognition System Architecture:

In order to develop a useful and applicable face recognition system several factors need to be take in hand.

1. The overall speed of the system from detection to recognition should be acceptable.
2. The accuracy should be high.
3. The system should be easily updated and enlarged, that is easy to increase the number of subjects that can be recognized. (Parmar & Mehta, January 2014).

The main components involved in designing a typical face recognition system are shown in the following figure:

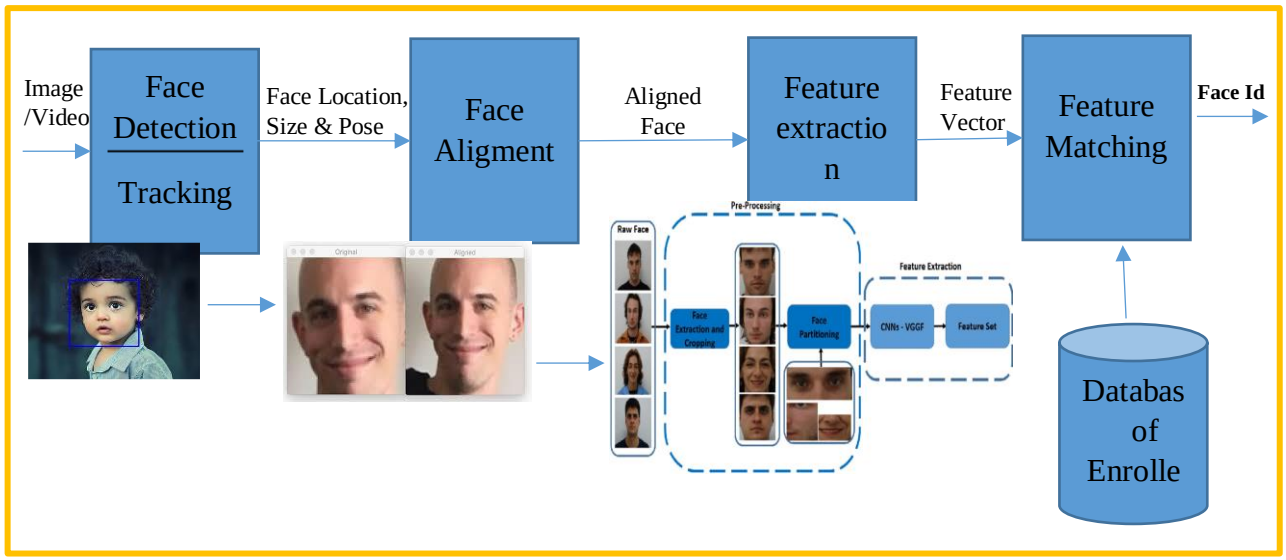


Figure 26: Face recognition processing flow

2.2.1 Face Detection:

This process separates the facial area from the rest of the background image. In the case of video streams, faces can be tracked using a face-tracking component. To detect the face from the image there are several methods:

1) Knowledge-based Method:

The rule-based method uses the knowledge of human to get the information about the typical face. Usually, the rules capture the relationships between facial features to design the location of the features in the face.

2) Template Matching Method:

In this, several standard patterns of a face are stored in the database or the system to describe the face as a whole or the facial features separately. The link between an input image and the stored patterns are evaluated for detection. These methods have been used for both face localization and detection.

3) Appearance based Method:

In contrast to template matching, the models are learned from a set of training images which should capture the representative variability of the appearance face. These learned models are then used for detection and are mainly designed for face detection.

4) Block rank patterns:

In this, a block rank pattern is generated by dividing two gradient magnitude images into nine(3×3) blocks and then a face is roughly detected by these 3×3 block rank patterns generated from the gradient magnitude images.

5) Viola Jones Face Detector:

The Viola–Jones face detector is the first object detection framework which provides competitive object detection rates in real-time. It was developed in 2001.

6) Feature Invariant Approach:

This algorithm aims to find structural features that exist even when the pose, viewpoint, or lighting conditions vary, and then we use this approach to locate face.

2.2.2 Face Alignment:

This process focus on finding the best localization and normalization of the face; where the detection step roughly estimates the position of the face, this step outlines the facial components, such as face outline, eyes, nose, ears and mouth. Afterwards normalization with respect to geometrical transforms such as size and pose, in addition to photometrical properties such as illumination and grey scale take place.

2.2.3 Feature Extraction:

After the previous two steps, feature extraction is performed resulting in effective information that is useful for distinguishing between faces of different persons and stable with respect to the geometrical and photometrical variations.

2.2.4 Face Matching:

The extracted features are compared to those stored in the database, and decisions are made according to the sufficient confidence in the match score. (Jha, nd)

2.3 Face Recognition Methods:

In the beginning of the 1970's, face recognition was treated as a 2D pattern recognition problem. The distances between important points where used to recognize known faces, e.g. measuring the distance between the eyes or other important points or measuring different angles of facial components. (Parmar & Mehta, January 2014). There are three approaches used in face recognition.

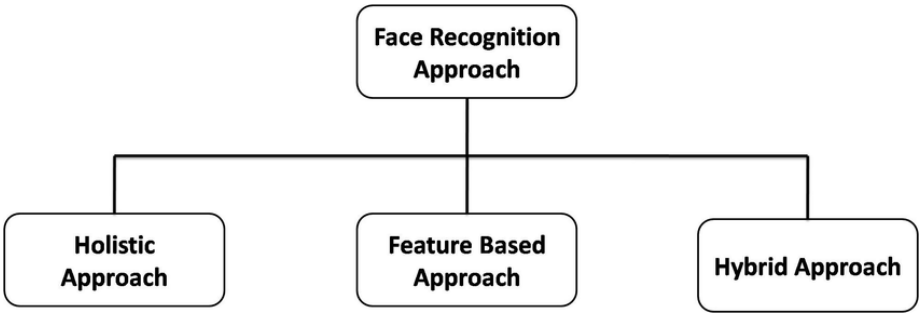


Figure 27: Face recognition Approaches

2.3.1 Holistic Matching Methods:

In holistic approach, the complete face region is taken into account as input data into face catching system. One of the best example of holistic methods are Eigenfaces, Principal Component Analysis, Linear Discriminant Analysis and independent component analysis etc.

2.3.1.1 Eigen Faces3 (also known as Karhunen- Loève expansion/eigenvector):

The first successful demonstration of machine recognition of faces was made by Turk and Pentland in 1991 using eigen faces. Their approach covers face recognition as a two-dimensional recognition problem. The different stages in an eigenface based recognition system are :

The first stage is to insert a set of images into a database, these images are names as the training set and this is because they will be used when we compare images and when we create the eigenfaces.

The second stage is to create the eigenfaces. Eigenfaces are made by extracting characteristic features from the faces. The input images are normalized to line up the eyes and mouths. They are then resized so that they have the same size. Eigenfaces can now be extracted from the image data by using a mathematical tool called Principal Component Analysis (PCA). When the eigenfaces have been created, each image will be represented as a vector of weights. The system is now ready to accept entering queries and the weight of the incoming unknown image is found and then compared to the weights of those already in the system. If the input image's weight is over a given threshold it is considered to be unidentified. The identification of the input image is done by finding the image in the database whose weights are the closest to the weights of the input image. The image in the database with the closest weight will be returned as a hit to the user of the system.

[3] The word eigenface coined by German ―Eigen wert‖ The ―Eigen‖ literally mean characteristic and ―wert‖ mean value. (Lal, et al., 2018)

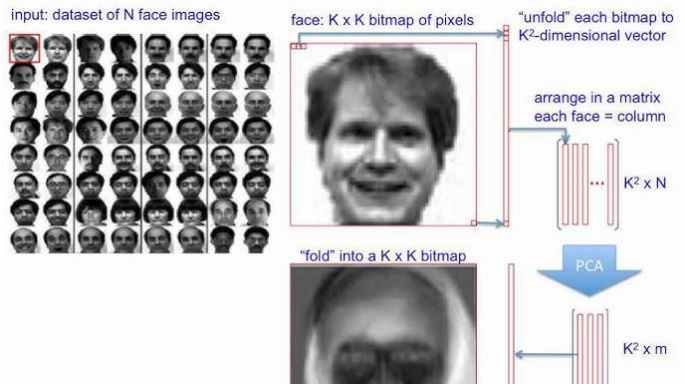


Figure 28: Chart of the eigenface-based algorithm

2.3.1.2 Linear Discriminate Analysis (LDA):

The Fisher's Linear Discriminate also called Linear Discriminate Analysis discovers few features to distinguish faces of different or same individual. Few features is found by maximizing the Fisher Discriminate Criterion (Fisher 1936), which is achieved by maximizing the grouping of individual faces whilst minimizing the grouping of different individual faces. Therefore by grouping faces of the same individual these features can be used to determine the identity of individuals. (Karamizadeh, Shahidan , & Zamani, September -2013).

2.3.1.3 Principal Component Analysis (PCA):

The principal component analysis (PCA) is a kind of algorithms in biometrics. It is a statistics technical and used orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables. PCA also is a tool to reduce multidimensional data to lower dimensions while retaining most of the information. (Karamizadeh, Shahidan , & Zamani, September -2013).

2.3.1.4 Artificial Neural Networks (ANN):

The attractiveness of using neural networks could be due to its non linearity in the network. Hence, the feature extraction step may be more efficient than the linear Karhunen-Loève methods. One of the first artificial neural networks (ANN) techniques used for face recognition is a single layer adaptive network called WISARD which contains a separate network for each stored individual (T.J. Stonham,1984). The way in constructing a neural network structure is crucial for successful recognition. It is very much dependent on the intended application.

For face detection, multilayer perceptron(K.K. Sung and T. Poggio,1995) and convolutional neural network (S. Lawrence, C.L. Giles, A.C. Tsoi, and A.D,1997) have been applied and for

face verification, (J. Weng, J.S. Huang, and N. Ahuja,1993) is a multi-resolution pyramid structure.

ANN provides an effective feature recognition technique, and it has been widely used after emergence of Artificial Intelligence. This consists of network, where neurons are arranged in the form of layers. Accuracy of face recognition has been boosted with the aid of better deep network architectures and supervisory methods. And recently few remarkable face representation learning techniques are evolved. Using these techniques, deep learning (Fig. 29) has got much closer to human performance.

One of the most viable feature of Neural Networks is it lessens the complexity. It learns from the training samples and then works fine on the images with changes in lighting conditions and increases accuracy.

Training is a precursor step to get the desired results from the system as user point of view. After feature extraction, classifiers for face recognition such as the Radial Basis Function and Feed Forward Neural Network (FFNN) are the implemented.

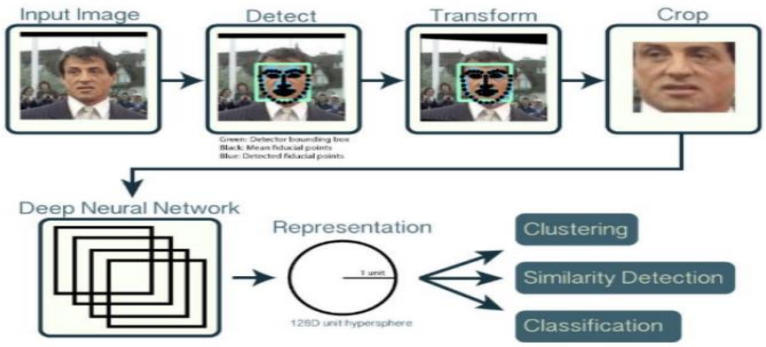


Figure 29: Face recognition using Neural Network

2.3.1.5 Support Vector Machine (SVM):

A Support Vector Machine (SVM) is formally defined by a separate hyperplane as a discriminatory classifier. When given the labeled training data, an optimal hyperplane is produced by the algorithm that categorizes new examples. A Support Vector Machine (SVM) performs classification by building an N dimensional hyperplane that divides data into two categories in an optimal way. SVM is very closely linked to neural networks. The primary objective of SVM data analysis is to find the ideal hyperplane that divides vector clusters in away that instances with one classification of the specified variable are on one side of the plane and instances with the other category are on the other side of the plane. The support vectors are the vectors near the hyperplane. The figure shows the face recognition system by using SVM. (Solomon, Meena, & Kaur, April-June 2019).

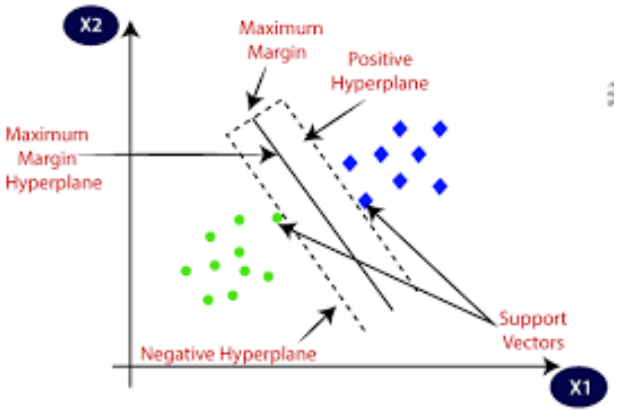


Figure 30: Support Vector Machine Algorithm

2.3.2 Feature-Based (Structural) Methods:

In this methods local features such as eyes, nose and mouth are first of all extracted and their locations and local statistics (geometric and/or appearance) are fed into a structural classifier. A big challenge for feature extraction methods is feature "restoration", this is when the system tries to retrieve features that are invisible due to large variations, e.g. head Pose when we are matching' a frontal image with a profile image. Distinguishes between three different extraction methods:

1. Generic methods based on edges, lines, and curves.
2. Feature-template-based methods.
3. Structural matching methods that take into consideration geometrical Constraints on the features.

Figure 31: Structural features for face recognition

2.3.2.1 Geometrical Feature Matching:

Geometrical feature matching techniques are based on the computation of a set of geometrical features from the picture of a face. The fact that face recognition is possible even at coarse resolution as low as 8x6 pixels when the single facial features are hardly revealed in detail, implies that the overall geometrical configuration of the face features is sufficient for recognition. The overall configuration can be described by a vector representing the position and size of the main facial features, such as eyes and eyebrows, nose, mouth, and the shape of face outline. One of the pioneering works on automated face recognition by using geometrical features was done by (T. Kanade) in 1973. Their system achieved a peak performance of 75% recognition rate on a database of 20 people using two images per person, one as the model and the other as the test image. References (A.J. Goldstein, L.D. Harmon, and A.B. Lesk,1971 & Y. Kaya and K. Kobayashi,1972) showed that a face recognition program provided with

features extracted manually could perform recognition apparently with satisfactory results. Reference (R. Bruneli and T. Poggio) in 1993 automatically extracted a set of geometrical features from the picture of a face, such as nose width and length, mouth position, and chin shape. There were 35 features extracted form a 35 dimensional vector. The recognition was then performed with a Bayes classifier. They reported a recognition rate of 90% on a database of 47 people. Reference (I.J. Cox, J. Ghosn, and P.N. Yianios,1996) introduced a mixture-distance technique which achieved 95% recognition rate on a query database of 685 individuals. Each face was represented by 30 manually extracted distances. Reference (B.S. Manjunath, R,1992) used Gabor wavelet decomposition to detect feature points for each face image which greatly reduced the storage requirement for the database. Typically, 35-45 feature points per face were generated. The matching process utilized the information presented in a topological graphic representation of the feature points. After compensating for different centroid location, two cost values, the topological cost, and similarity cost, were evaluated. The recognition accuracy in terms of the best match to the right person was 86% and 94% of the correct person's faces was in the top three candidate matches.

In summary, geometrical feature matching based on precisely measured distances between features may be most useful for finding possible matches in a large database such as a Mug shot album. However, it will be dependent on the accuracy of the feature location algorithms. Current automated face feature location algorithms do not provide a high degree of accuracy and require considerable computational time.

2.3.3 Hybrid Methods:

Hybrid face recognition systems use a combination of both holistic and feature extraction methods. Generally 3D Images are used in hybrid methods. The image of a person's face is caught in 3D, allowing the system to note the curves of the eye sockets, for example, or the shapes of the chin or forehead. Even a face in profile would serve because the system uses depth, and an axis of measurement, which gives it enough information to construct a full face. The 3D system usually proceeds thus: Detection, Position, Measurement, Representation and Matching.

In Case the 3D image is to be compared with an existing 3D image, it needs to have no alterations. Typically, however, photos that are put in 2D, and in that case, the 3D image need a few changes. This is tricky, and is one of the biggest challenges in the field today.

2.4 Face Recognition Applications:

The applications of facial recognition range from static ("mug shots") to dynamic, uncontrolled face identification in a cluttered background (subway, airport), some face recognition applications are mentioned below:

- **Face Identification:** Face recognition systems identify people by their face images. It establish the presence of an authorized person rather than just checking whether a valid identification (ID) or key is being used or whether the user knows the secret personal identification numbers (Pins) or passwords.
- **Access Control:** In many of the access control applications, such as office access or computer logon, the size of the group of people that need to be recognized is relatively small. The face pictures are also caught under natural conditions, such as frontal faces and indoor illumination. The face recognition system of this application can achieve high accuracy without much co-operation from user.
- **Information Security:** In the area of information Security, it can be used as personal device logon, border checkpoints, database security, file encryption, intranet security, medical records, banks, biometric-log-in.
- **Image database investigations**: Searching image databases of licensed drivers, benefit recipients, missing children, immigrants and police bookings.
- **General identity verification:** Electoral registration, banking, electronic commerce, identifying newborns, national IDs, passports, employee IDs.
- **Surveillance:** Like security applications in public places, surveillance by face recognition systems has a low user satisfaction level, if not lower. Free lighting conditions, face orientations and other divisors all make the deployment of face recognition systems for large scale surveillance a challenging task.

2.5 Issues & Challenges:

There are various issues and challenges due to which the recognition rate may drops:

a) Pose Variation:

Pose variance is yet another hurdle in achieving a successful face recognition system. People pose differently every time they take a picture.

There is no standardized rule for taking a pose. Therefore, it makes more difficult to distinguish and recognize the faces from images with varying poses. Pose variations degrade the

performance of the facial features. In addition, many systems work under inflexible imaging conditions and as a result it affects the quality of gallery images.

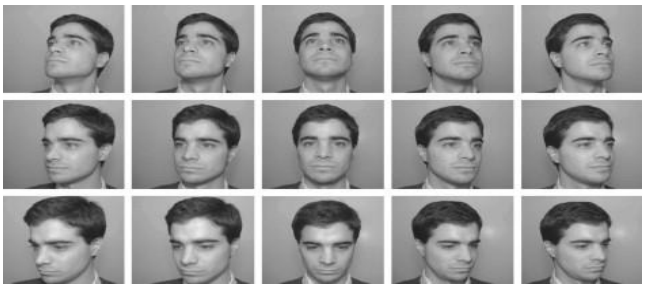


Figure 32: pose variation samples

The methods dealing with variation in pose can be divided into two kinds i.e. **multi-view face recognition** and **face recognition across pose**. Multi-view face recognition can be considered as an annexure of frontal face recognition in which gallery image of every pose is considered. On the other hand, across a pose in face recognition, yield face with a pose which has never been exposed before to a recognition system. A good face recognition approach should provide good pose tolerance and capability to recognize different poses. (Lal, et al., 2018).

b) **Illumination:**

It is the variation in the angle of light falls on the face. It changes the result so far that the difference between the two pictures of same person is large than the difference between the two different persons.

Illumination is an observable property and effect of light. It may also refer to lightning effect or the use of light sources. Global illuminations are algorithms which have been used in 3D computer graphics. Illumination variation also badly affects the face recognition system. Thus it has been turned an area of attention for many researchers. However, it becomes tedious task to recognize one or more persons from still or video images. But it can be quite easy to extract desired information from images when they are taken under a controlled environment along with uniform background.

Figure 33: Illumination samples

Also there are three methods that can be implemented to deal with illumination problem. **They are gradient**, **gray level** and **face reflection field estimation techniques**. Gray level transformation technique carries out in-depth mapping with a non-linear or linear function. Gradient extraction approaches are used to extract edges of an image in gray level. As illumination is a factor that heavily affects the performance of recognition system obtained via face images or videos. These techniques are developed to suppress the effect of illumination. (Lal, et al., 2018)

c) **Ageing:**

Aging is an inevitable natural process during the lifetime of a person as compared to other facial variations. Aging effect can be observed under main three unique characteristics:

1) The aging is uncontrollable: It cannot be advanced or even delayed and it is slow and irreversible.

2) Personalized aging Signs: Every human passes through different aging patterns. And these rely on his or her genes and many other factors, such as health, food, region, and weather conditions.

3) The aging signs depend on time: The face of person at a specific age will affect all older faces, but unaffected in younger age. (Lal, et al., 2018)

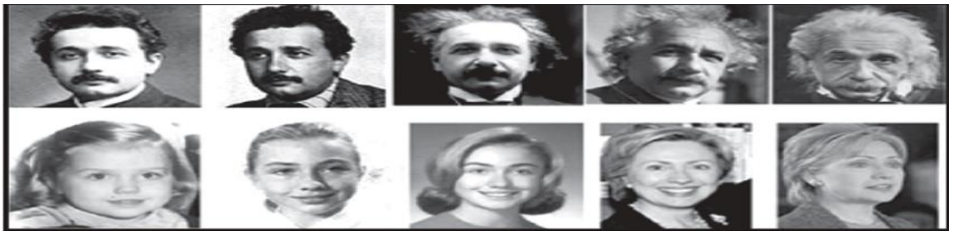


Figure 34: Aging samples

d) **Facial Expression:**

It changes due to change in facial gestures like sadness, happiness, angry, tired etc.

Figure 35: Facial expression samples

e) **Occlusion:**

It is due to the presence of glasses, beards, stroll or caps which covers the face or object.

Figure 36: Occlusion Examples

f) **Low Resolution:**

Image of very small face has low resolution, which consists of very less information which may drops the recognition rate of the system.

Figure 37: Low resolution example

g) **Imaging Condition:**

Different surroundings and different camera can change the quality of the image and affecting the face appearance.

h) **Interclass similarity:**

Different persons may have very similar appearance like: twins, relatives and strangers may look alike.

2.6 Comparative Analysis:

It is so complicated to choose a method among all those that exist, and that is not only due to the diversity of these methods (algorithm, approaches) ,which each of them has these strong points and these weak points, but, also to the images taken from the databases used in the learning and evaluation phase and on what criteria and in what environment they were taken (pose, illumination, age and expression) which is a big challenge for researchers today and in the future.

In the literature, there are several comparative studies between the existing methods, and we display two examples that are:

2.6.1 Study 1:

According to: (Solomon, Meena, & Kaur, April-June 2019), where, on one side they computed the accuracy of each: PCA, LDA , ICA and SVM), and on the other side they considered a hybrid system composed of PCA and LDA, PCA with SVM. Table below shows the individual accuracy of each of the systems and the accuracy for the hybrid system. The bar chart also shows the comparison between the algorithms. (is computed using Matlab code and ORL database).

ALGORITHMS USED	NUMBER OF SUBJECTS	CORRECT	INCORRECT	RECOGNITION ACCURACY (%)
PCA	20	286	34	89.4
LDA	20	292	28	91.25
ICA	20	283	37	88.4375
PCA+LDA	20	298	22	93.125
PCA+SVM	20	305	15	95.3125

Table 6: Result running the algorithms in ORL database

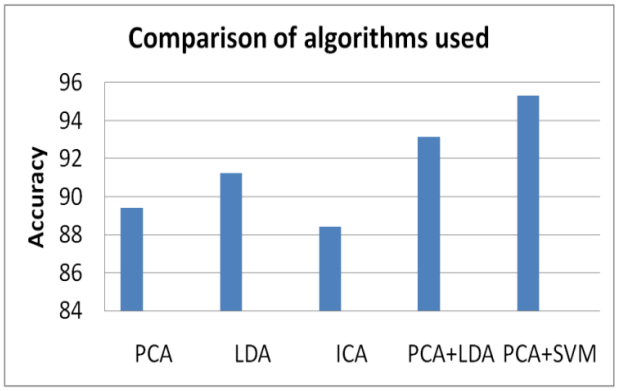


Figure 38: Algorithms accuracy comparison

2.6.1.1 Findings:

We see that the algorithms individually cannot solve all the challenges (variations), but when used together can solve the issues to an extent. Especially when combining PCA and LDA together, we find that the system can solve variation in illumination, pose and expression challenges. The age challenge has always been a very difficult one to solve as the problem with age is that, a person's face features may or may not vary in 10 years duration. ORL database used could only provide the faces with pose, illumination and expression variation.

We saw that, if the algorithms are used individually, the output produced is much less and if used in a combined form, we get much better output. The remaining issue we have to work on is the age problem, which is still an unsolved problem in face recognition system.

2.6.2 Study 2:

According to (Lal, et al., 2018) This survey studies state of the art face detection techniques and approaches (PCA, eigenfaces, ANN, SVM…) , and mentioned different testing face databases which include AT & T (ORL), AR, FERET, LFW, YTF, and Yale. The following tables show the accuracy ratio of these techniques:

Database:

Let us first look for some database used in this study to execute this methods:

ORL: The ORL Database of Faces contains 400 images from 40 distinct subjects. For some subjects, the images were taken at different times, varying the lighting, facial expressions (open / closed eyes, smiling / not smiling) and facial details (glasses / no glasses). All the images were taken against a dark homogeneous background with the subjects in an upright, frontal position (with tolerance for some side movement). The size of each image is 92x112 pixels, with 256 grey levels per pixel. (ftp://ftp.uk.research.att.com:pub/data/att_faces.tar.Z as a 4.5Mbyte compressed).

Figure 39:ORL Database sample

AR: This face database was created by Aleix Martinez and Robert Benavente in the Computer Vision Center (CVC) at the U.A.B. It contains over 4,000 color images corresponding to 126 people's faces (70 men and 56 women). Images feature frontal view faces with different facial expressions, illumination conditions, and occlusions (sun glasses and scarf). The pictures were taken at the CVC under strictly controlled conditions. No restrictions on wear (clothes, glasses, etc.), make-up, hair style, etc. were imposed to participants. Each person participated in two sessions, separated by two weeks (14 days) time. The same pictures were taken in both sessions.

Figure 40: AR Database sample

FERET: The FERET database is being used in facial recognition for system evaluation. The Face Recognition Technology (FERET) program is executed by joint collaboration between the Defense Advanced Research Projects Agency (DARPA) and the National Institute of Standards and Technology (NIST). DARPA released a high-resolution, 24-bit color version of these images in 2003. And it was tested over 2,413 still face images, representing 856 individuals. However, the main motive behind development of FERET database was to facilitate algorithm

development and evaluation. Thus initially, it requires a common database of facial images in order to develop and test for the purpose evaluation. After that, complications in image mentioned by the images should enhance.

Figure 41: FERET database

LFW: Labeled Faces in the Wild (LFW) is a database of face photographs which was mainly developed for the comprehension of unconstrained face recognition problem. The data set has over 13,000 images of faces obtained via web. And each face is labeled with the name of the person whose picture was captured. However, roughly 1680 of the pictured people contained two or more distinct photos in the data set.

Figure 42: LFW database sample

YALE: The Yale Face Database (size 6.4MB) contains 165 grayscale images in GIF format of 15 individuals. There are 11 images per subject, one per different facial expression or configuration: center-light, w/glasses, happy, left-light, w/no glasses, normal, right-light, sad, sleepy, surprised, and wink.(http://vision.ucsd.edu/datasets/yale_face_dataset_original/yalefaces.zip)

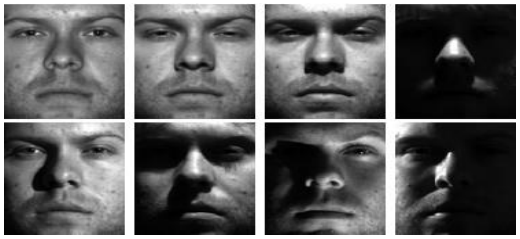


Figure 43: YALE database sample

2.6.2.1 Comparative table methods and techniques of face recognition:

S #	Year	Database	Techniqueue	Accuracy	Reference
1	2012	ORL Faces	PCA	70.0 %	Slavković et al. [20]
2	2012	Face94	PCA	100.0 %	Abdullah et al. [33]
3	2013	FRAV Face DB	Eigen Face	96.0 %	Saha, Rajib et al. [51]
4	2014	-	PCA Eigen Faces	70.0 %	Rahman, ArmanadurniAbd, et al. [21]
5	2014	Yale Database	PCA	92% to 93%	MuzammilAbdulrahman et al. [37]
6	2014	AT & T	PCA		Johannes Reschke et al. [36]
7	2016	Computer Vision Research Projects dataset	PCA	93.6 %	Md. Al-Amin Bhuiyan [34]
8	2017	EmguCV library	PCA + RMF	93.0 %	Jacky Efendi et al. [35]
9	2017	Yale Database	PCA	98.18	Riddhi A. & S.M. Shah [46]

Table 7:: Comparative Study Of Face Recognition Methods Using PCA

S #	Year	Database	Technique	Accuracy	Reference
1	2012	IIT-Dehli Database	NN Based SOM for Face recognition	88.25% to 98.3%	Raja, A. S. et al. [31]
2	2013	-	BPC and RBC Network	96.66% & 98.88%	Nandini, M. et al. [32]
3	2015	Deep ID 3		99.53%	Yi Sun, et al. [19]
4	2015	AFLW		99.00%	Haoxiang Li et al. [29]
5	2015	Multi PIE dataset	CPF	99.50%	JunhoYim et al. [52]
6	2015	AFLW		90.00%	Sachin Sudhakar Farfade [30]

Table 8: Comparative Study Of Face Recognition Methods Using ANN

S #	Year	Database	Technique	Accuracy	Reference
1	009	ORL Face Database	Least Square SVM	96%	Xie, Jianhong et al. [25]
2	2011	ORL Face Database	ICA, SVM	96%	Kong, Rui et al. [26]
3	2011	FERET Database, AT&T Database	2D-Principal Component Analysis, SVM	95.10 %	Le, Thai Hoang et al. [27]
4	2016	Yale Faces	SVM	97.78 %	Bhaskar Anand & Prashant K Shah [24]

Table 9: Comparative Study Of Face Recognition Methods Using SVM

2.6.2.2 Finding:

After analysis, it revealed that PCA is best suited technique when dimension of features is higher for original face images, whereas eigen faces image features method work well for frontal face recognition. Among face recognition methods, the most popular are Neural Networks, Support Vector Machine, Sparse Representation based Classification (SRC), Linear Regression Classification (LRC), Regularized Robust Coding (RRC) and Nearest Feature Line ((NFL). These methods provide better results when the image dimension is under 150 or more. Furthermore, it is suggested that PCA, SVM, NN and Eigen methods still need to be researched so that results that are more satisfactory could be achieved for face recognition. (Lal, et al., 2018).

Using the advanced methods such as neural networks and artificial intelligence can solve the ageing issue as well. For this purpose, In our project we have opted to use deep learning especially the technique called “metric deep learning” to improve a face recognition system in terms of performance and accuracy.

3. Deep Learning:

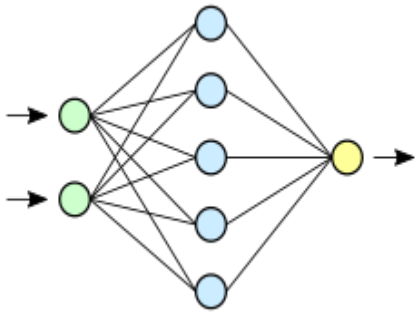
3.1 Artificial Intelligence(AI):

The central goal of Artificial Intelligence (AI) is to provide a set of algorithms and techniques that can be used to solve problems that humans perform intuitively and near automatically, but are otherwise very challenging for computers. While AI embodies a large, diverse set of work related to automatic machine reasoning. The machine learning: subfield tends to be specifically interested in pattern recognition and learning from data. Artificial Neural Networks (ANNs) are a class of machine learning algorithms that learn from data and specialize in pattern recognition, inspired by the structure and function of the brain. (Rosebrock, pyimagesearch.com/, June 18,2018)

3.2 Machine Learning (ML):

According to Arthur Samuel, Machine learning is defined as the field of study that gives computers the ability to learn without being explicitly programmed. It is used to teach machines how to handle the data more efficiently. Sometimes after viewing the data, we cannot interpret the extract information from the data. In that case, we apply machine learning. With the abundance of datasets available, the demand for machine learning is in rise. Many industries apply machine learning to extract relevant data. The purpose of machine learning is to learn from the data.

3.3 Artificial Neural Network (ANN):



Artificial Neural networks have existed since the 1940; they are interconnected networks of processing units called artificial neurons that loosely mimic the axons found in a biological brain.

In a biological neuron, dendrites receive input signals from various neighboring neurons, typically more than one thousand of them. These modified signals are then passed on to the cell body or soma of the neuron, where these signals are summed together and then passed on to the axon of the neuron. If the received input signal is more than a specified threshold, the axon will release a signal, which will be passed on to the neighboring dendrites of other neurons. (Pattanayak, 2017)

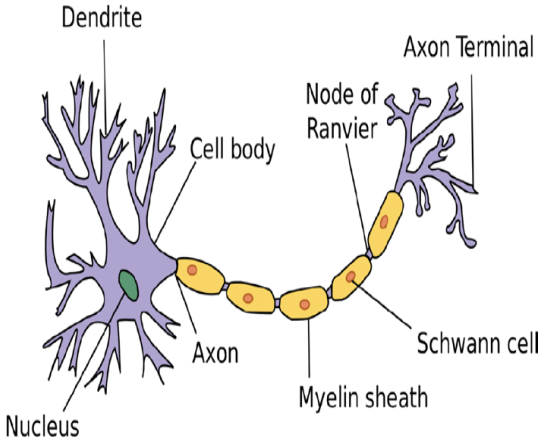


Figure 45: Structure of a biological neuron

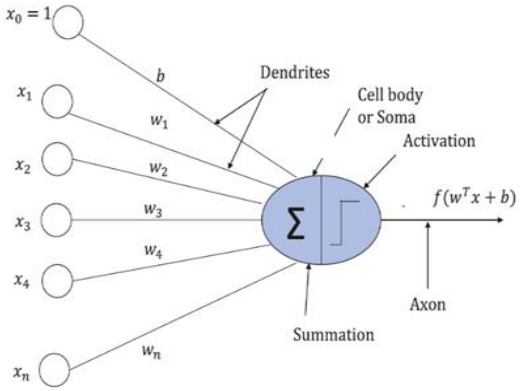


Figure 44: Structure of an artificial neuron

Artificial neuron (Perceptron[4]) units are inspired by the biological neurons, with some modifications for convenience. Much like the dendrites, the input connections to the neuron carry the attenuated or amplified input signals from other neighboring neurons. The signals are passed on to the neuron, where the input signals are summed up and then a decision is made as to what to output based on the total input received. For instance, for a binary threshold neuron, an output value of 1 is provided when the total input exceeds a pre-defined threshold; otherwise, the output stays at 0. Several other types of neurons are used in artificial neural networks, and their implementation only differs with respect to the activation function on the total input to produce the neuron output.

[4] McCulloch and Pitts in 1943 introduced the concept of perceptron as an artificial neuron that is the basic building block of the artificial neural network.

3.3.1 Evolution of artificial neural networks:

The chronology of major events in the artificial neural network community is shown in the figure below from the artificial neural network into deep learning, (Pattanayak, 2017) Evolution of artificial neural networks

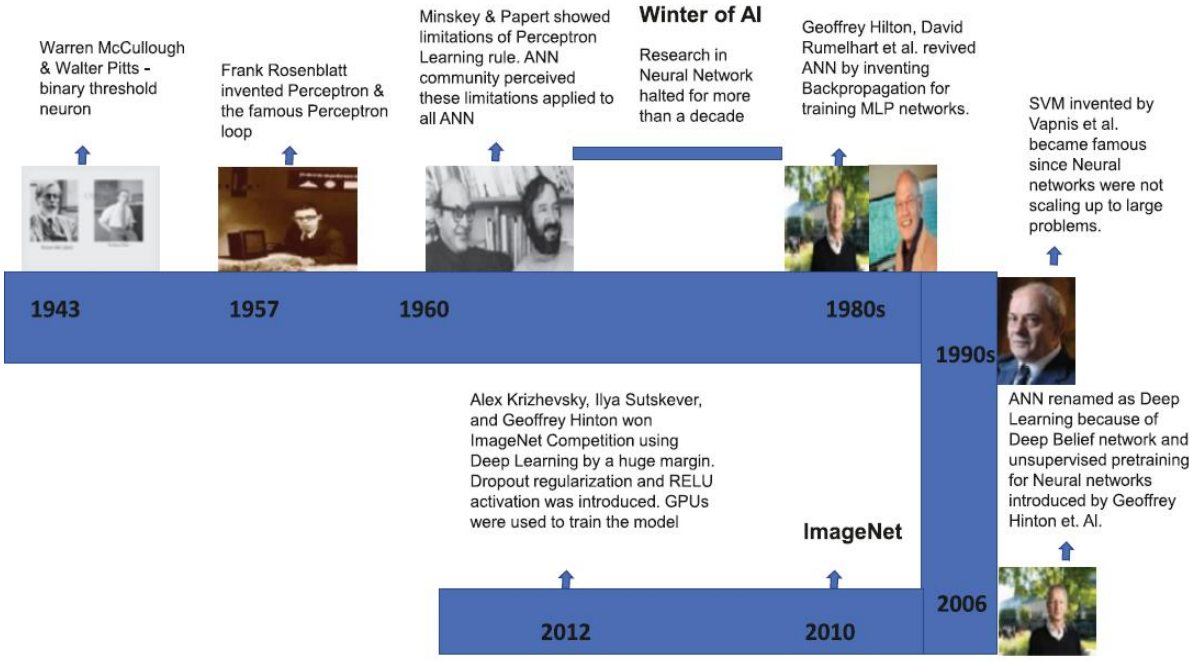


Figure 46: Evolution of artificial neural networks (1943-2012)

3.4 Deep Learning:

Deep learning is the subfield of machine learning, which is, in turn, a subfield of artificial intelligence (AI), based on artificial neural networks. Deep learning models introduce an extremely sophisticated approach to machine learning because they have been specifically modeled after the human brain. Complex, multi-layered "deep neural networks" are built to allow data to be passed between nodes (like neurons) in highly connected ways. The result is a non-linear transformation of the data that is increasingl-y abstract.

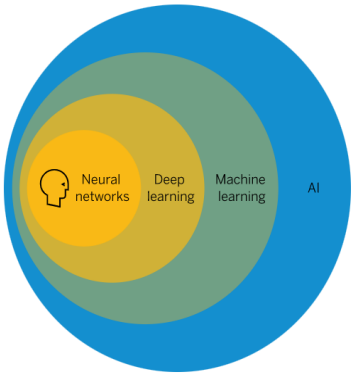


Figure 47: Deep learning subfield of AI

Deep learning methods are representation-learning methods with multiple levels of representation, obtained by composing simple but nonlinear modules that each transform the representation at one level (starting with the raw input) into a representation at a higher, slightly more abstract level. The key aspect of deep learning is that these layers are not designed by human

engineers: they are learned from data using a general-purpose learning procedure" – (Yann LeCun, Yoshua Bengio, and Geoffrey Hinton, 2015).

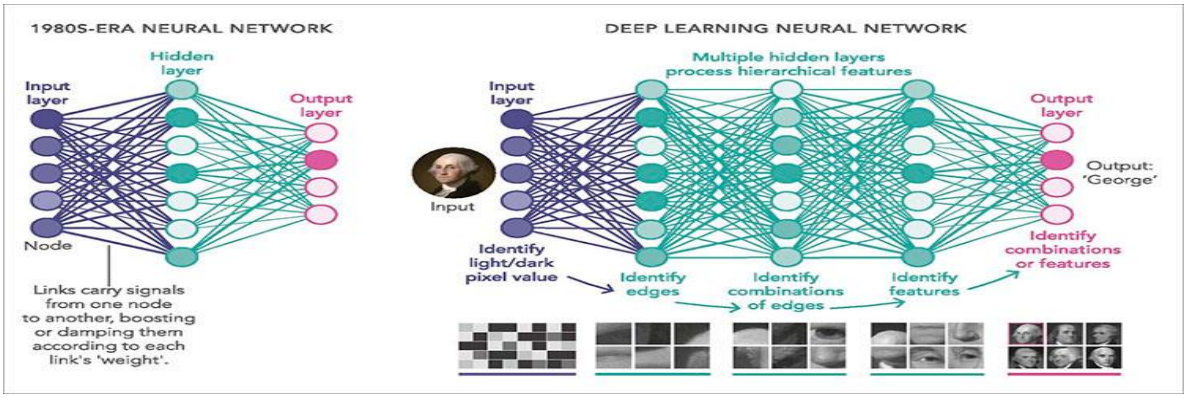


Figure 48: Deep learning Neural Network

3.4.1 Why deep learning fore face recognition:

- In response to some of the shortcomings of machine learning, and the significant advance in the theoretical and technological capabilities at our disposal today, deep learning has emerged and is rapidly expanding as one of the most exciting fields of science. It is being used in technologies such as self-driving cars, image recognition on social media platforms, and translation of text from one language to others. Deep learning is devoted to building algorithms that explain and learn a high and low level of abstractions of data that traditional machine learning algorithms often cannot. The models in deep learning are often inspired by many sources of knowledge, such as game theory and neuroscience. (Beysolow II, 2017).

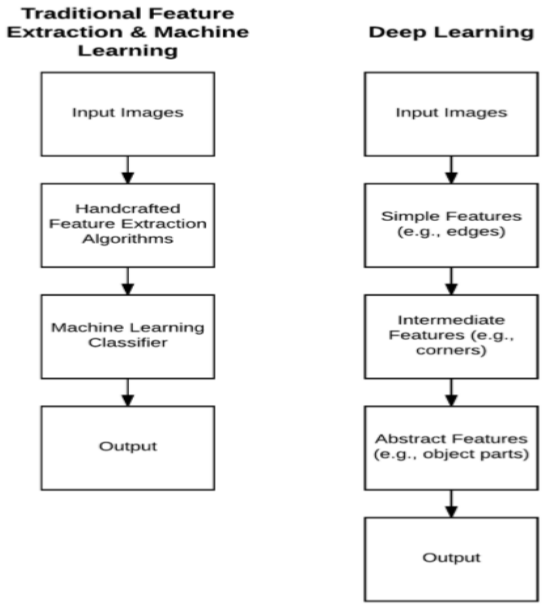


Figure 49: Traditional Machine learning vs Deep Machine learning.

Traditional process of taking an input set of images, applying hand-designed feature extraction algorithms, followed by training a machine learning classifier on the features. Deep

learning approach of stacking layers on top of each other that automatically learn more complex, abstract, and discriminating features (Rosebrock, pyimagesearch.com/, June 18,2018). One of the big differences between Deep Learning and traditional ML algorithms is that it scales well, the greater the amount of data provided, the better the performance of a Deep Learning algorithm.

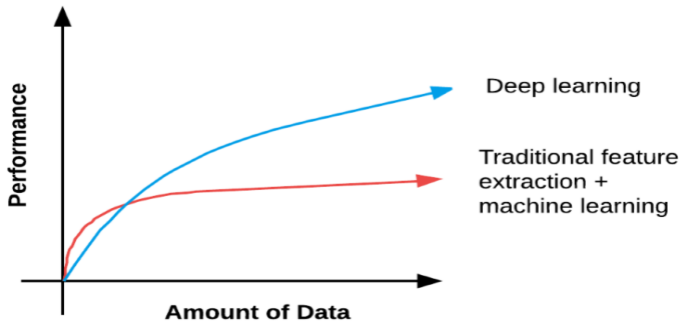


Figure 50: Performance Algorithm

- Deep learning can analyze images, videos, and unstructured data in ways machine learning cannot easily do.
- Also, GPU technology has revolutionized the deep-learning world.(GPU stands for graphical processing unit), which was initially used for gaming purposes to display more screens per second for better gaming resolution. Deep-learning networks use a lot of matrix multiplication, especially convolution, for both the forward pass and for backpropagation. GPUs are good at matrix-to-matrix multiplication; hence, several thousand cores of GPU are utilized to process data in parallel. This speeds up the deep-learning training. (Pattanayak, 2017).

3.4.2 Deep Learning Architectures:

Deep learning algorithms work with almost any kind of data and require large amounts of computing power and information to solve complicated issues. The most common deep learning algorithms are: MLP, CNNs, LSTMs, RNNs, GANs, RBFNs, MLPs, DBNs, RBMs and Autoencoders.etc

3.4.2.1 Multilayer Perceptron (Feedforward NeuralNetwork):

To address the drawback of single perceptrons, multilayer perceptrons were proposed;

It is a composition of multiple perceptrons connected in different ways and operating on distinctive activation functions to enable improved learning mechanisms. The training sample propagates forward through the network and the output error is back propagated and the error is minimized using the gradient descent method, which will calculate a loss function for all the weights in the network.

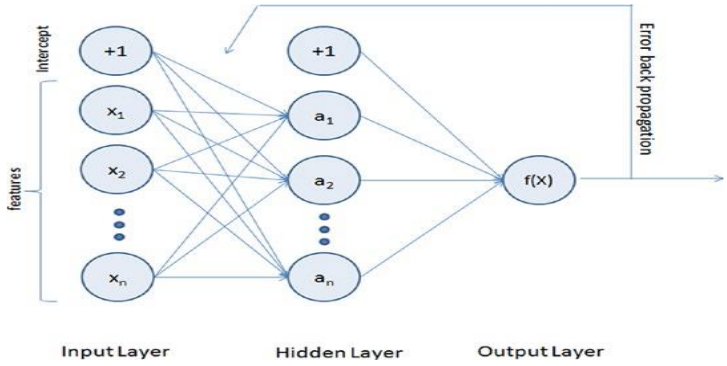


Figure 51: Multilayer perceptron representation

The activation function for a simple one-level hidden layer of a multilayer perceptron can be given by:

$$f(x) = g\left(\sum_{j=0}^{M} W_{kj}^{(2)} g\left(\sum_{i=0}^{d} W_{ji}^{(1)} x_i \right) \right)$$

Where: ***xi*** is the input and **(*Wji*)** is the input layer weights and ***Wkj*** is the weight of hidden layer.

A multilayered neural network can have many hidden layers, where the network holds its internal abstract representation of the training sample. The upper layers will be building new abstractions on top of the previous layers. So having more hidden layers for a complex dataset will help the neural network to learn better.

As it is shown in Figure 60, the MLP architecture has a minimum of three layers, that is, input, hidden, and output layers. The input layer's neuron count will be equal to the total number of features and in some libraries an additional neuron for intercept/bias.

These neurons are represented as nodes. The output layers will have a single neuron for regression models and binary classifier; otherwise it will be equal to the total number of class labels for multiclass classification models.

Note that using too few neurons for a complex dataset can result in an under-fitted model due to the fact that it might fail to learn the patterns in complex data. However,using too

many neurons can result in an over-fitted model as it has capacity to capture patterns that might be noise or specific for the given training dataset. So to build an efficient multilayered neural network, A widely accepted rule of thumb is that you can start with one hidden layer, as there is a theory that one hidden layer is sufficient for the majority of problems. Then, gradually increase the layers on a trial-and-error basis to see if there is any improvement in accuracy. The number of neurons in the hidden layer can ideally be the mean of the neurons in the input and output layers.

3.4.2.2 Autoencoders:

As the name suggests, an autoencoder aims to learn encoding as a representation of training sample data automatically without human intervention. The autoencoder is widely used for dimensionality reduction and data de-nosing.

Building an autoencoder will typically have three elements:

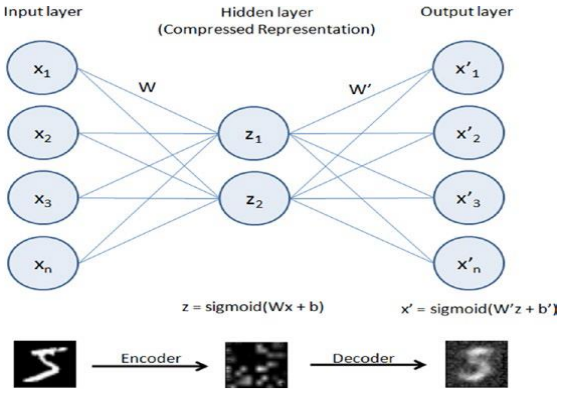


Figure 52: autoencoders Architecture

1. Encoding function to map input to a hidden representation through a nonlinear function,

z = sigmoid (Wx + b)

2. A decoding function such as

x' = sigmoid(W'y + b')

which will map back into reconstruction x' with same shape as x.

3. A loss function, which is a distance function to measure the information loss between the compressed representation of data and the decompressed representation. Reconstruction error can be measured using traditional squared error $||\mathbf{x}-\mathbf{z}||^2$.

3.4.2.3 Long Short-Term Memory (LSTM):

LSTM is an implementation of improved RNN architecture to address the issues of general RNN, and it enables long-range dependencies. It is designed to have better memory through linear memory cells surrounded by a set of gate units used to control the flow of information, when information should enter the memory, when to forget, and when to output. It uses no activation function within its recurrent components, thus the gradient term does not vanish with back propagation.

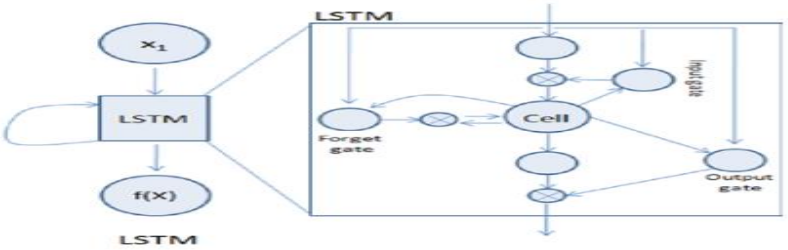


Figure 53: LSTM architecture

➔ The table below shows the LSTM components:

LSTM Component	Formula
Input gate layer: This decides which values to store in the cell state.	$i_t = \text{sigmoid}(w_i x_t + u_i h_{t-1} + b_i)$
Forget gate layer: As the name suggested this decides what information to throw away from the cell state.	$f_t = \text{sigmoid}(W_f x_t + U_f h_{t-1} + b_f)$
Output gate layer: Create a vector of values that can be added to the cell state.	$O_t = \text{sigmoid}(W_o x_t + u_i h_{t-1} + b_o)$
Memory cell state vector.	$c_t = f_t \text{ o } c_{t-1} + i_t o * \text{hyperbolic tangent}(W_c x_t + u_c h_{t-1} + b_c)$

Table 10: LSTM Components

3.4.2.4 Recurrent Neural Network (RNN)

The MLP (feedforward network) is not known to do well on sequential events models such as the probabilistic language model of predicting the next word based on the previous word at very given point. RNN architecture addresses this issue. It is similar to MLP except that they have a feedback loop, which means they feed previous time steps into the current step. This type of architecture generates sequences to simulate situation and create synthetic data, making them the ideal modeling choice to work on sequence data such as speech text mining, image captioning, time series prediction, robot control, language modeling.

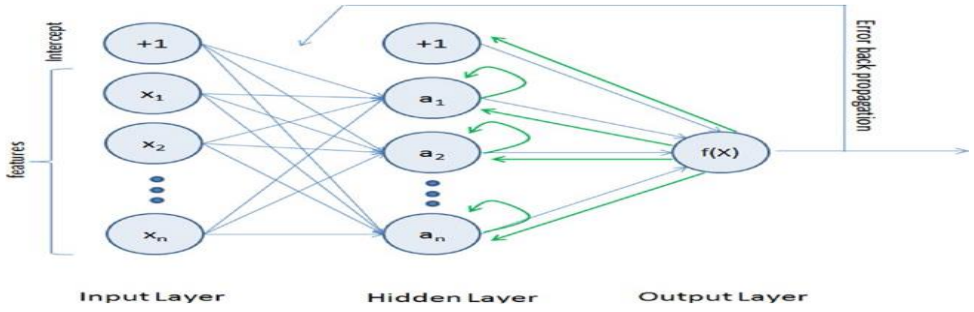


Figure 54: RNN Architecture

The previous step's hidden layer and final outputs are fed back into the network and will be used as input to the next steps' hidden layer, which means the network will remember the past and it

will repeatedly predict what will happen next. The drawback in the general RNN architecture is that it can be memory heavy, and hard to train for long term temporal dependency (i.e., context of long text should be known at any given stage). (Swamynathan, 2017)

3.4.2.5 Convolution Neural Network (CNN):

CNNs are inspired by multi-layer perceptron's. By imposing local connectivity constraints between neurons of adjacent layers, CNN exploits local spatial correlation. The core element of convolutional neural networks is the processing of data through the convolution operation. (Pattanayak, 2017)

Convolutional neural networks are distinguished from other neural networks by their superior performance with image, speech, or audio signal inputs. They have three main types of layers, which are: (IBM Cloud Education , October 2020)

a) **Convolutional Layer:**

The convolutional layer is the core building block of a CNN, and it is where the majority of computation occurs. It requires a few components, which are input data, a filter, and a feature map.

Let's assume that the input will be a color image, which is made up of a matrix of pixels in 3D. This means that the input will have three dimensions—a height, width, and depth—which correspond to RGB in an image. We also have a feature detector, also known as a kernel or a filter, which will move across the receptive fields of the image, checking if the feature is present. This process is known as a convolution. The feature detector is a two-dimensional (2-D) array of weights, which represents part of the image. While they can vary in size, the filter size is typically a 3x3 matrix; this also determines the size of the receptive field. The filter is then applied to an area of the image, and a dot product is calculated between the input pixels and the filter. This dot product is then fed into an output array. Afterwards, the filter shifts by a stride, repeating the process until the kernel has swept across the entire image. The final output from the series of dot products from the input and the filter is known as a feature map, activation map, or a convolved feature.

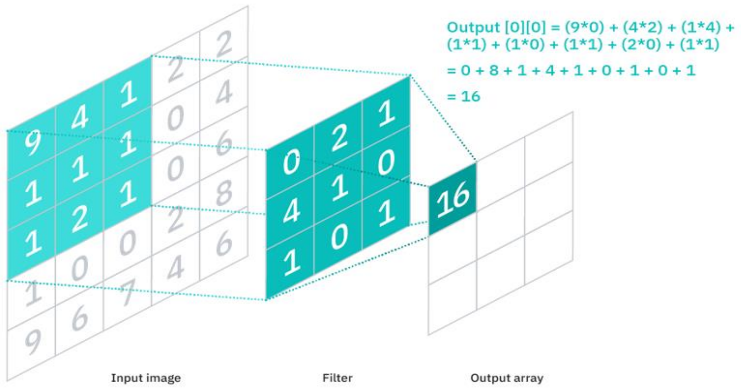


Figure 55: Convolutional Layer components

As we can see in the image above, each output value in the feature map does not have to connect to each pixel value in the input image. It only needs to connect to the receptive field, where the filter is being applied. Since the output array does not need to map directly to each input value, convolutional (and pooling) layers are commonly referred to as "partially connected" layers. However, this characteristic can also be described as local connectivity.

Note that the weights in the feature detector remain fixed as it moves across the image, which is also known as parameter sharing. Some parameters, like the weight values, adjust during training through the process of backpropagation and gradient descent. However, there are three hyper parameters which affect the volume size of the output that need to be set before the training of the neural network begins. These include:

1. The **number of filters** affects the depth of the output. For example, three distinct filters would yield three different feature maps, creating a depth of three.
2. **Stride** is the distance, or number of pixels, that the kernel moves over the input matrix. While stride values of two or greater is rare, a larger stride yields a smaller output.
3. **Zero-padding** is usually used when the filters do not fit the input image. This sets all elements that fall outside of the input matrix to zero, producing a larger or equally sized output. There are three types of padding:

Valid padding: This is also known as no padding. In this case, the last convolution is dropped if dimensions do not align.

Same padding: This padding ensures that the output layer has the same size as the input layer

Full padding: This type of padding increases the size of the output by adding zeros to the border of the input.

After each convolution operation, a CNN applies a Rectified Linear Unit (ReLU) transformation to the feature map, introducing nonlinearity to the model.

We mention also, another convolution layer can follow the initial convolution layer. When this happens, the structure of the CNN can become hierarchical as the later layers can see the pixels within the receptive fields of prior layers. Ultimately, the convolutional layer converts the image into numerical values, allowing the neural network to interpret and extract relevant patterns.

b) **Pooling Layer:**

Also known as downsampling, conducts dimensionality reduction, reducing the number of parameters in the input. Similar to the convolutional layer, the pooling operation sweeps a filter across the entire input, but the difference is that this filter does not have any weights. Instead, the kernel applies an aggregation function to the values within the receptive field, populating the output array. There are two main types of pooling:

Max pooling: As the filter moves across the input, it selects the pixel with the maximum value to send to the output array. As an aside, this approach tends to be used more often compared to average pooling.

Average pooling: As the filter moves across the input, it calculates the average value within the receptive field to send to the output array.

While a lot of information is lost in the pooling layer, it also has a number of benefits to the CNN. They help to reduce complexity, improve efficiency, and limit risk of overfitting.

c) **Fully-Connected Layer:**

The name of the full-connected layer aptly describes itself. As mentioned earlier, the pixel values of the input image are not directly connected to the output layer in partially connected layers. However, in the fully-connected layer, each node in the output layer connects directly to a node in the previous layer. This layer performs the task of classification based on the features extracted through the previous layers and their different filters. While convolutional and pooling layers tend to use ReLu functions, FC layers usually leverage a softmax activation function to classify inputs appropriately, producing a probability from 0 to 1.

3.5 Different Architectures in Convolutional Neural Networks:

There are different approaches to use CNN architecture. First is learning the model from scratch. In this case, the architecture of the pre-trained model is used and trained according to the dataset. Second is using transfer learning with features from pre-trained CNN, in cases where the dataset is large. Finally, CNN can be used via transfer learning by keeping the convolutional base in its original form and then using its outputs to feed the classifier. The pre-

trained model is used as a fixed feature extraction mechanism in cases where the dataset is small, or when the problem is similar to the one to be classified.

We will go through a few widely used convolutional neural network architectures used today. These network architectures are not only used for classification, but also, with minor modification, are used in segmentation, localization, and detection. Also, there are pre-trained versions of each of these networks that enable the community to do transfer learning or fine-tune the models. Except LeNet, almost all the CNN models have won the ImageNet competition for classification of a thousand classes.

3.5.1 AlexNet:

The AlexNet CNN architecture was developed by Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton in 2012 to win the 2012 ImageNet ILSVRC (ImageNet Large-Scale Visual Recognition Challenge). The original paper pertaining to AlexNet is titled "ImageNet Classification with Deep Convolutional Neural Networks". It was the first time that a CNN architecture beat other methods by a huge margin. Their network achieved an error rate of 15.4 percent on its top five predictions as compared to the 26.2 percent error rate for the second-best entry.

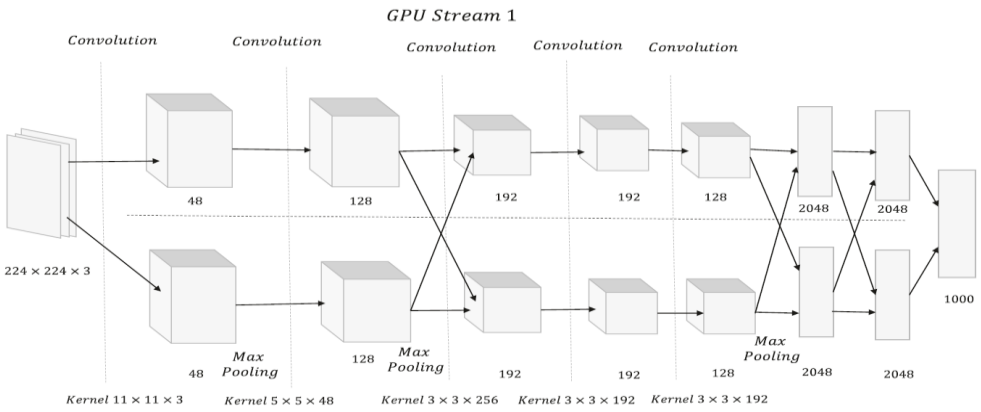


Figure 56: AlexNet CNN architecture

AlexNet consists of five convolutional layers, max pooling layers, and dropout layers, and three fully connected layers in addition to the input and output layer of a thousand class units. The inputs to the network are images of size 224´224´3. The first convolutional layer produces 96 feature maps corresponding to 96 filter kernels of size 11´11´3 with strides of four pixel units. The second convolutional layer produces 256 feature maps corresponding to filter kernels of size 5´5´48. The first two convolutional layers are followed by max pooling layers, whereas the next three convolutional layers are placed one after another without any

intermediate max pooling layers. The fifth convolutional layer is followed by a max pooling layer, two fully connected layers of 4096 units, and finally a SoftMax output layer of one thousand classes. The third convolutional layer has 384 filter kernels of size 3´3´256, whereas the fourth and fifth convolutional layers have 384 and 256 filter kernels each of size 3´3´192. A dropout of 0.5 was used in the last two fully connected layers. You will notice that the depth of the filter kernels for convolutions is half the number of feature maps in the preceding layer for all but the third convolutional layer. This is because AlexNet was at that time computationally expensive and hence the training had to be split between two separate GPUs. However, if you observe carefully, for the third convolutional activity there is cross connectivity for convolution and so the filter kernel is of dimension 3´3´256 and not 3´3´128 . The same kind of cross-connectivity applies to the fully connected layers, and hence they behave as ordinary fully connected layers with 4096 units.

The key features of AlexNet are as follows:

ReLU activation functions were used for non-linearity. They had a huge impact since RELUs are significantly easier to compute and have constant non-saturating gradients as opposed to sigmoid and tanh activation functions, whose gradients tend to zero for very high and low values of input. (Pattanayak, 2017)

3.5.2 VGG16:

The VGG group in 2014 were runners up in the ILSVRC-2014 competition with a 16-layer architecture named VGG16. It uses a deep yet simple architecture that has gained a lot of popularity since. The paper pertaining to the VGG network is titled “Very Deep Convolutional Networks for Large-Scale Image Recognition” and is authored by Karen Simonyan and Andrew Zisserman.

Instead of using a large kernel-filter size for convolution, VGG16 architecture used 3´3 filters and followed it up with ReLU activations and max pooling with a 2´2 receptive field. The inventors’ reasoning was that using two 3´3 convolution layers is equivalent to having one 5´5 convolution while retaining the advantages of a smaller kernel-filter size; i.e., realizing a reduction in the number of parameters and realizing more non-linearity because of two convolution–ReLU pairs as opposed to one. A special property of this network is that as the spatial dimensions of the input volume reduces because of convolution and max pooling, the number of feature maps increases due to the increase in the number of filters as we go deep into the network. (Pattanayak, 2017)

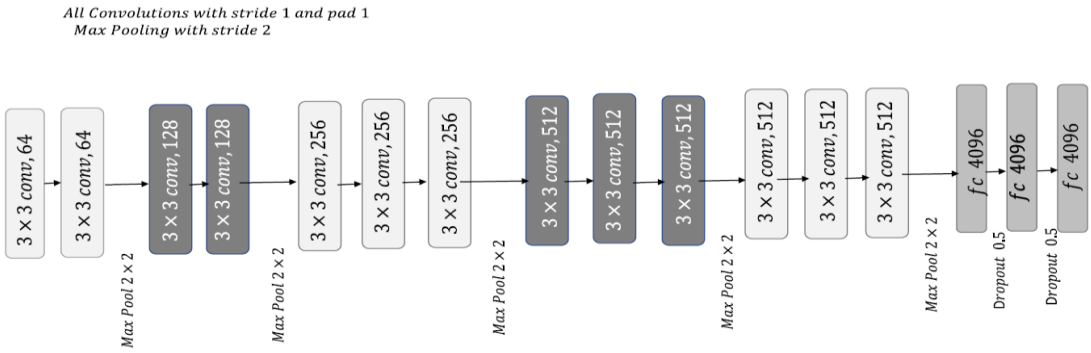


Figure 57: VGG16 architecture

Figure 57 represents the architecture of VGG16. The input to the network are images of size 224´224´3 . The first two convolutional layers produce 64 feature maps, each followed by max pooling. The filters for convolution are of spatial size 3´3 with a stride of 1 and pad of 1. Max pooling is of size 2´2 with stride of 2 for the whole network. The third and fourth convolutional layers produce 128 feature maps, each followed by a max pooling layer. The rest of the network follows in a similar fashion. At the end of the network there are three fully connected layers of 4096 units, each followed by the output SoftMax layer of a thousand classes. Dropout is set at 0.5 for the fully connected layers. All the units in the network have ReLU activations. (Pattanayak, 2017)

3.5.3 ResNet:

ResNet is a 152-layer-deep convolutional neural network from Microsoft that won the ILSVRC 2015 competition with an error rate of only 3.6 percent, which is perceived to be better than the human error rate of 5–10 percent. The paper on ResNet, authored by Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, is titled “Deep Residual Learning for Image Recognition”. Apart from being deep, ResNet implements a unique idea of residual block. After each series of convolution–ReLUs–convolution operations, the input to the operation is fed back to the output of the operation. In traditional methods while doing convolution and other transformations, we try to fit an underlying mapping to the original data to solve the classification task. However, with ResNet’s residual block concept, we try to learn a residual mapping and not a direct mapping from the input to output.

Formally, in each small block of activities we add the input to the block to the output. This is illustrated in Figure 58. This concept is based on the hypothesis that it is easier to fit a residual mapping than to fit the original mapping from input to output. (Pattanayak, 2017)

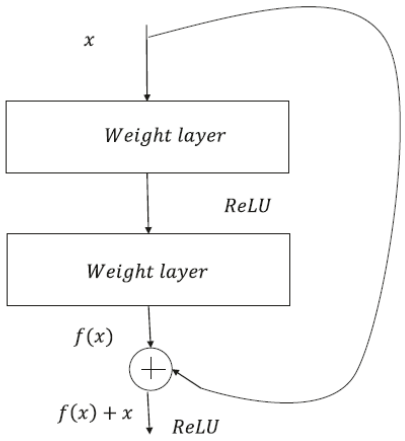


Figure 58: ResNet architecture

3.6 Deep Metric Learning:

In recent years, Metric/Distance learning using Deep learning has been shown to output highly satisfying results for many computer vision tasks such as face recognition, face verification, image classification, Anomaly detection, etc. It helps capture Non-Linear feature structure by learning a non-linear transformation of the feature space. (Agrawal, Jan 11, 2021)

3.6.1 Metric:

A Metric is a non-negative function between two points x and y{say **g(x,y)**} that describes the so-called notion of **'distance'** between these two points. A metric must satisfy several properties:

Non-negativity => $d(x,y) \geq 0$ and $d(x,y) = 0$, if $x = y$.

Triangular inequality => $d(x,y) \leq d(x,z) + d(z,y)$.

Symmetry => $g(x,y) = g(y,x)$.

3.6.2 Examples of distance metric:

A. The Euclidean Metric: In a 'd' dimensional vector space, the metric is:

$$d(x,y) = \sqrt{\cdot(xi - yi)^2}$$

B. The Discrete Metric: The metric is given by:

$$d(x,y) = \begin{cases} 1, & if\ x \neq y \\ 0, & if\ x = y \end{cases}$$

3.6.3 Metric Learning:

Metric learning is an approach based directly on a distance metric that aims to establish similarity or dissimilarity between objects, Given the data and the corresponding output labels,

the goal is to come up with a set of rules or some complex function that maps those inputs to the corresponding output labels. One of the simplest machine learning algorithms where distance information is captured is the KNN (k- Nearest Neighbours) algorithm, where the idea is to find a neighborhood of k nearest data points for a new data point and assign this data point to the class to which the majority of the k data points belong. (Mahmut KAYA Hasan & B˙ILGE, 2019).

3.6.4 Deep Metric Learning functioning:

Deep Metric Learning uses Neural Networks to automatically learn discriminative features from the images and then compute the metric. (Agrawal, Jan 11, 2021)

While metric learning aims to reduce the distance between similar objects, it also aims to increase the distance between dissimilar objects. For this reason, there are approaches, such as k-nearest neighbors, which calculate distance information, and approaches where the data is transformed into a new representation. While the metric learning approaches are moved to the transformation space with distance information, the method is basically based on a W projection matrix. Current studies are directly related to Mahalanobis distance in general. When Mahalanobis distance is transformed into the Euclidean distance, the metric learning approach is presented based on the decomposition of the covariance matrix and the use of symmetric positive definite matrices while performing these operations. (Mahmut KAYA Hasan & B˙ILGE, 2019).

3.6.4.1 Loss function used in Deep Metric Learning:

The widely used loss functions for deep metric learning are the contrastive loss and the triplet loss. Let us focus on the last.

- Triplet Network:

Triplet network is a symmetric neural network architecture but consists of three identical subnetworks that share the same sets of parameters. The learning is performed as a set of three images. (Agrawal, Jan 11, 2021):

1) The **Anchor** image (The baseline image).

2) The **positive** image (truth images of a specific person belonging to the same class as an anchor),

3) The **negative** image (images not belonging to the anchor class).

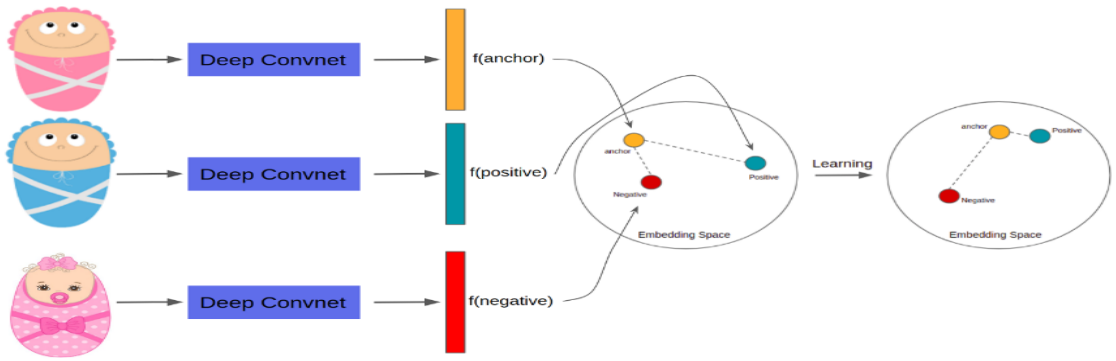


Figure 59: Triplet network

- Triplet loss:

The purpose of triplet loss is to ensure that the baseline image of a specific person is closer to all the positive sets of images (Mathematically, meaning that the Euclidean distance between an anchor and a positive image is small) and far away from the negative sets of examples.

Triplet loss function:

$$\mathrm{L} = \mathrm{L}(a, p, n) = \max(\mathrm{o}, \mathrm{m} + \|f(a) - f(p)\| - \|f(a) - f(n)\|)$$

Where:

a = anchor image, p = positive image, n = negative image, m = margin.

The embedding is represented by a 'd' dimensional vector f(.){the encoding}.

Let, d(a, p)=||f(a)-f(p)||.

When the loss will be **0**. The second term inside max(.) should be <0, i.e. d(a, n) > d(a, p) + m. If this happens, the triplet network will be quite lucky as there will be no learning. These triplets are known as Easy triplets.

When the loss will be high, i.e. when, d(a, n) << d(a, p), which has a significant contribution to the loss. This will be a hard to train on triplet. Such triplets are known as Hard triplets.

When d(a, p) < d(a, n) < d(a, p) + m, Again, the loss will be positive. (Agrawal, Jan 11, 2021)

4. Conclusion:

This chapter discussed with details two main concepts, which are face recognition and deep learning especially CNN architecture that perform well on a diverse set of image classification tasks. In addition, we have introduced also a metric deep learning such as a technique of deep learning that we will used in our project.

Part II: Design and Implementation of the System

Chapter 3: System Design

1. Introduction:

After, having carried out a bibliographic study (a state of the art) on the facial recognition methods used in access control systems, and after having seen the importance of security within laboratories in general, in this chapter we will design an access control system in a block of laboratories located at the University of Skikda based on facial recognition in order to restrict access for the authorized persons only (laboratory members). This system must be used in the main door of the building block, as well as in all the doors of the laboratories that constitute it. This chapter presented an overview of our proposed system with the necessary details. Thus, it represented all the stages of the construction process of our system, which consists of four main steps.

To design our system, we used the modeling language UML, the needs of our system were represented in the use case diagram, the operation of this system was represented in the sequence diagram, and the different classes with their relationships were represented in the class diagram.

2. Host Organization:

Les ‘’Cinqs Laboratoires’’ (i.e. the five laboratories in English): is a building located at the University of Skikda, which consists of 16 laboratories in different scientific fields, supervised by the Ministry of Higher Education and Scientific Research (MESRS). The Figure below illustrates its structure.

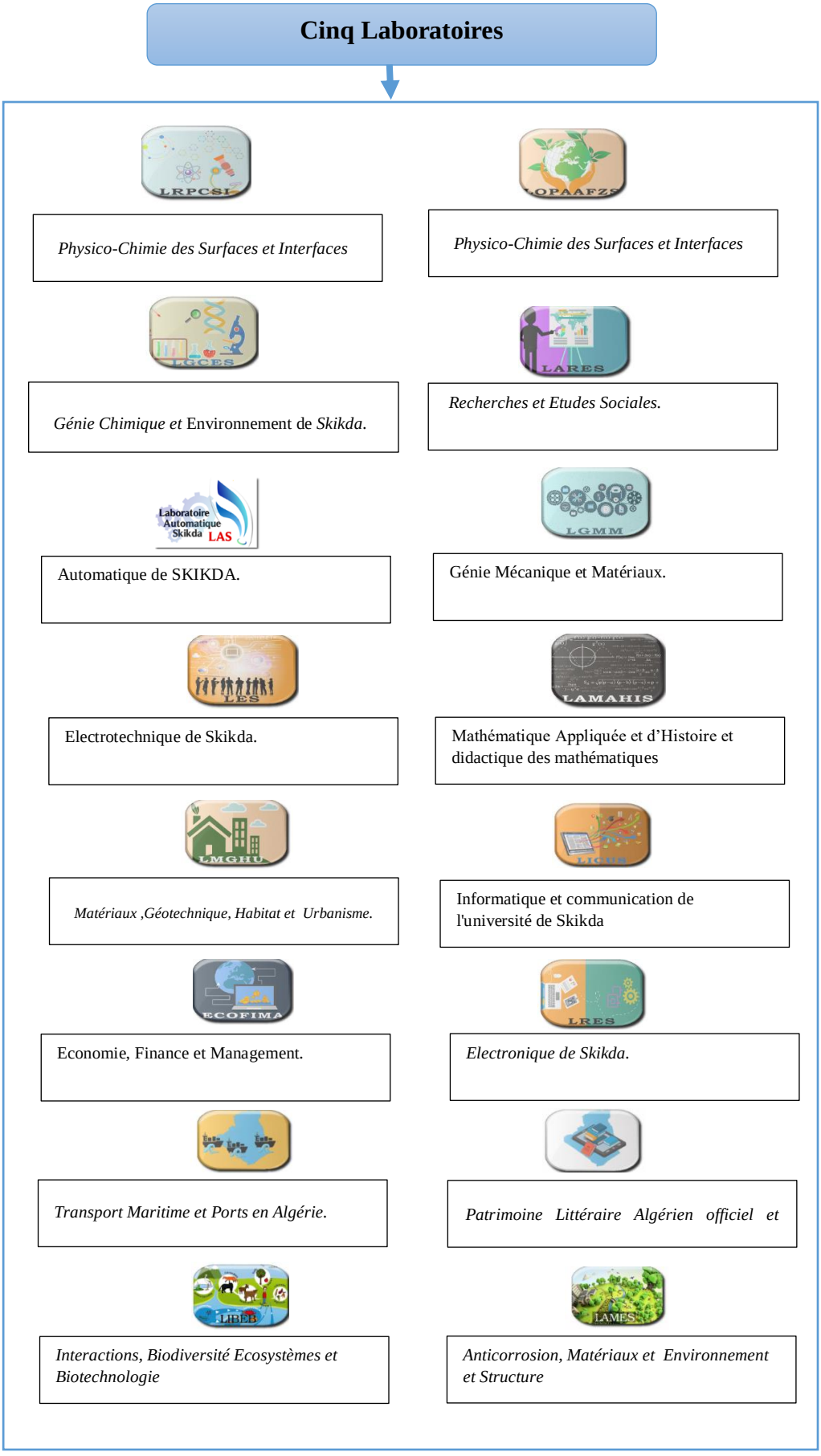


Figure 60: Cinq Laboratoires organizational chart

3. UML:

The OMG specification states (Union, 2012)‘’The Unified Modeling Language (UML) is a graphical language for visualizing, specifying, constructing, and documenting the artifacts of a software-intensive system. The UML offers a standard way to write a system’s blueprints, including conceptual things such as business processes and system functions as well as concrete things such as programming language statements, database schemas, and reusable software components.”

UML models can be directly connected to a variety of programming languages. Maps to Java, C++, Visual Basic, and so on. Tables in a RDBMS or persistent store in an OODBMS.

3.1 UML Diagrams:

It consists of 14 diagrams (UML 2.2) regrouped in 02 different sets structure diagram and behavior Diagram (Figure 69):

The use of UML diagrams aims to visualize the system from different perspectives.(Union, 2012). In our conception, we have opted for 03 diagrams, which are:

Use case Diagram: shows a set of use cases and actors and their relationships.

Sequence Diagram: shows how a use case will be implemented in terms of collaborating objects.

Class Diagram: shows UML classes and relationships.

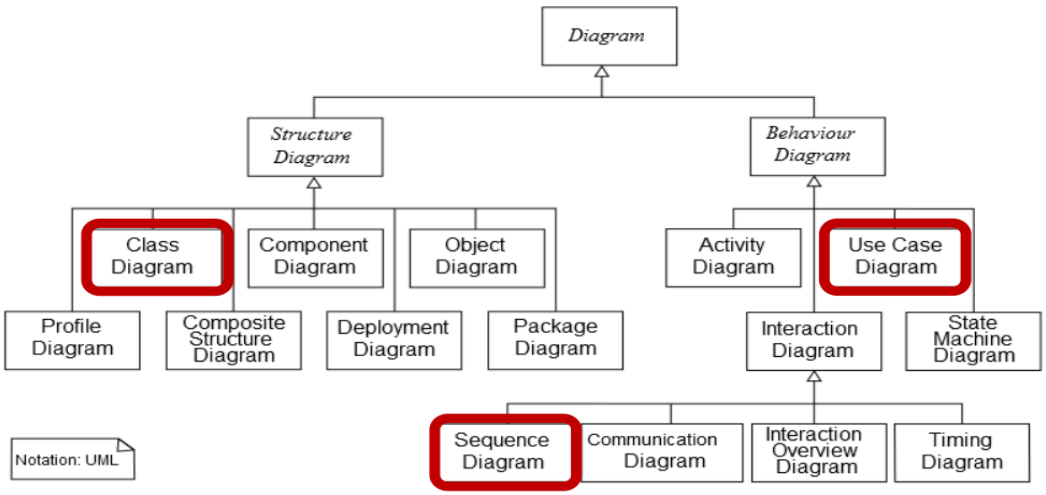


Figure 61: UML Structure

3.1.1 Use Case Diagram:

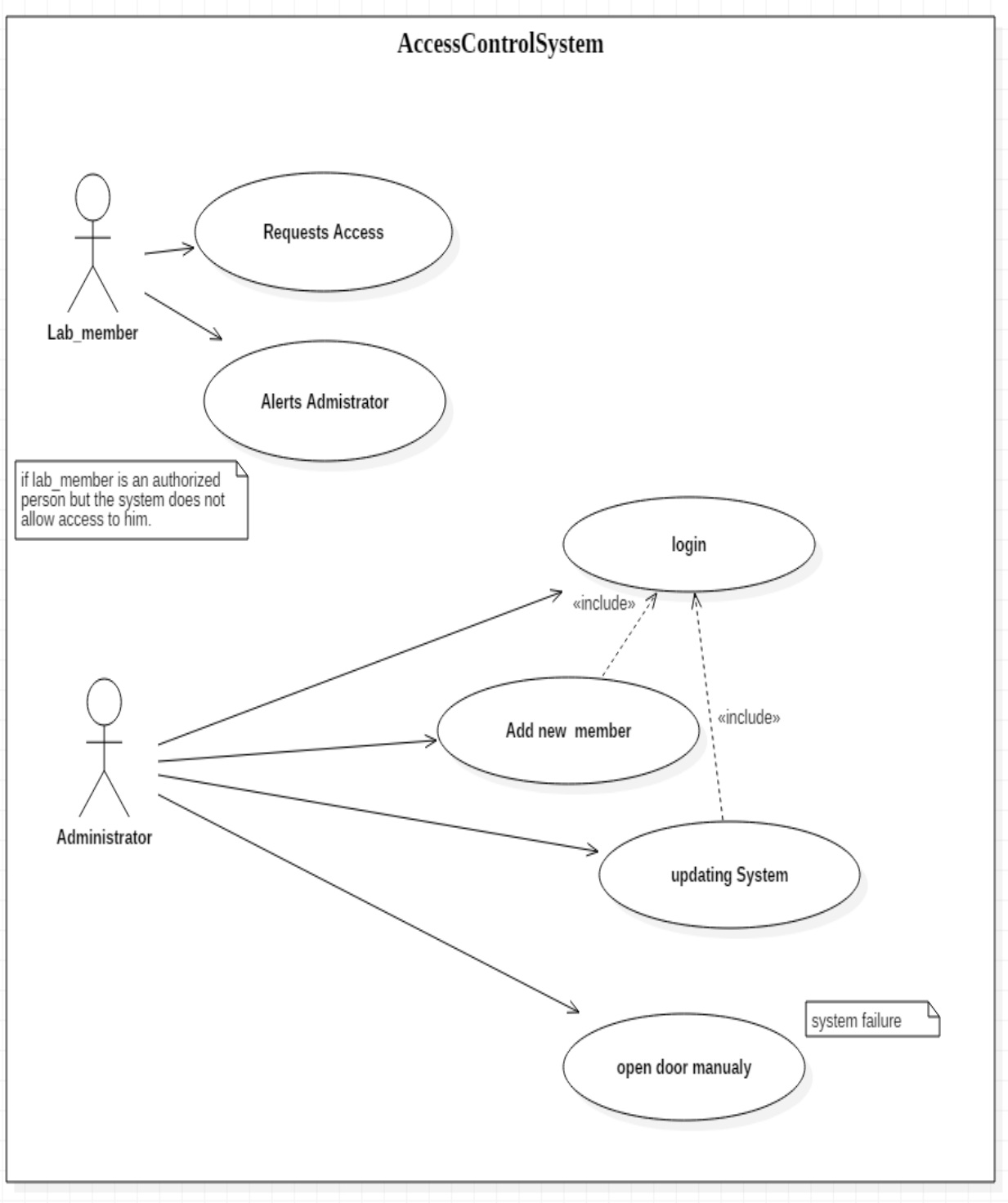


Figure 62: Use Cases Diagram

3.1.2 Class Diagram:

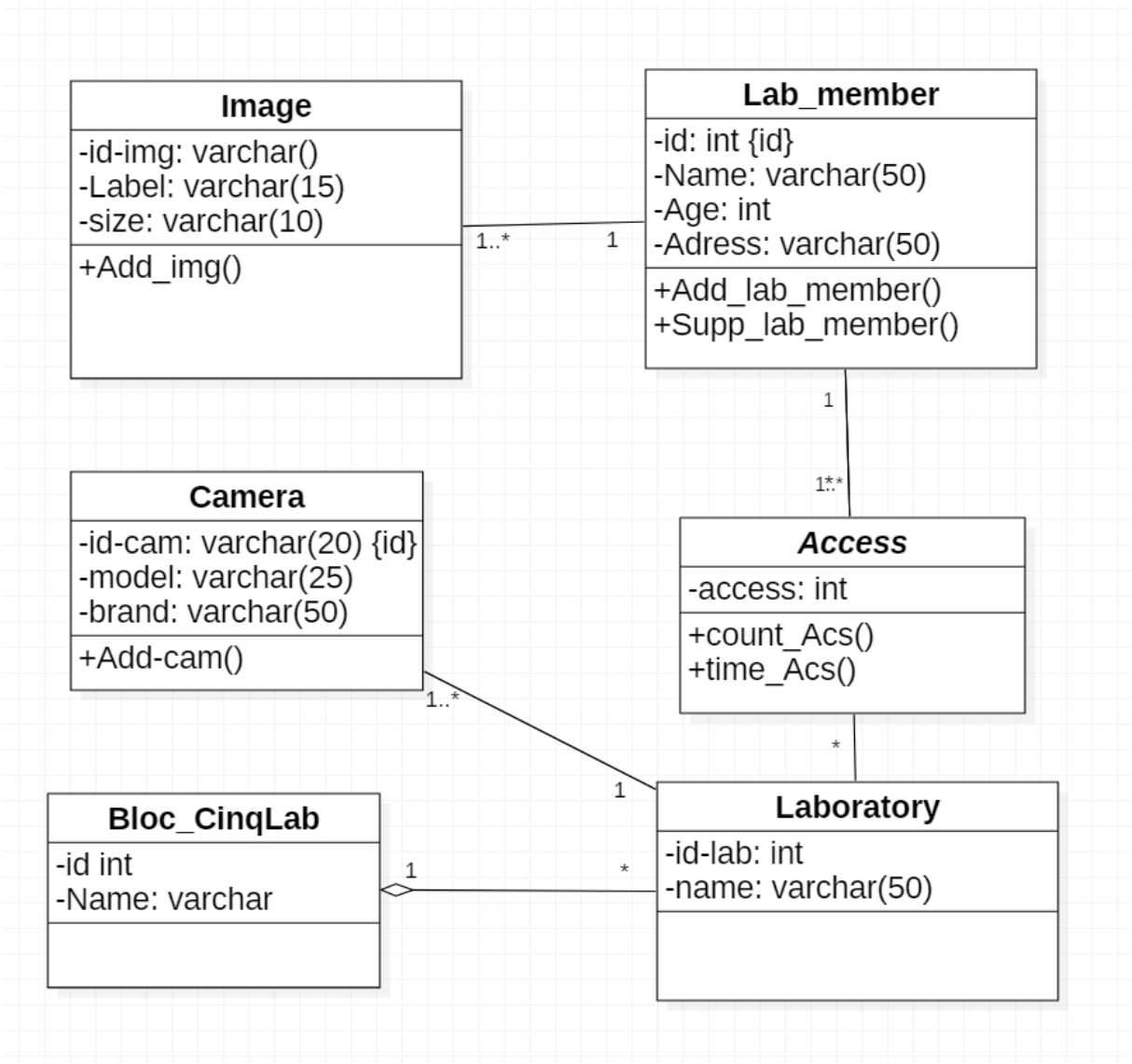


Figure 63: Class Diagram

3.1.3 Sequence Diagrams:

3.1.3.1 Sequence « Request Access »:

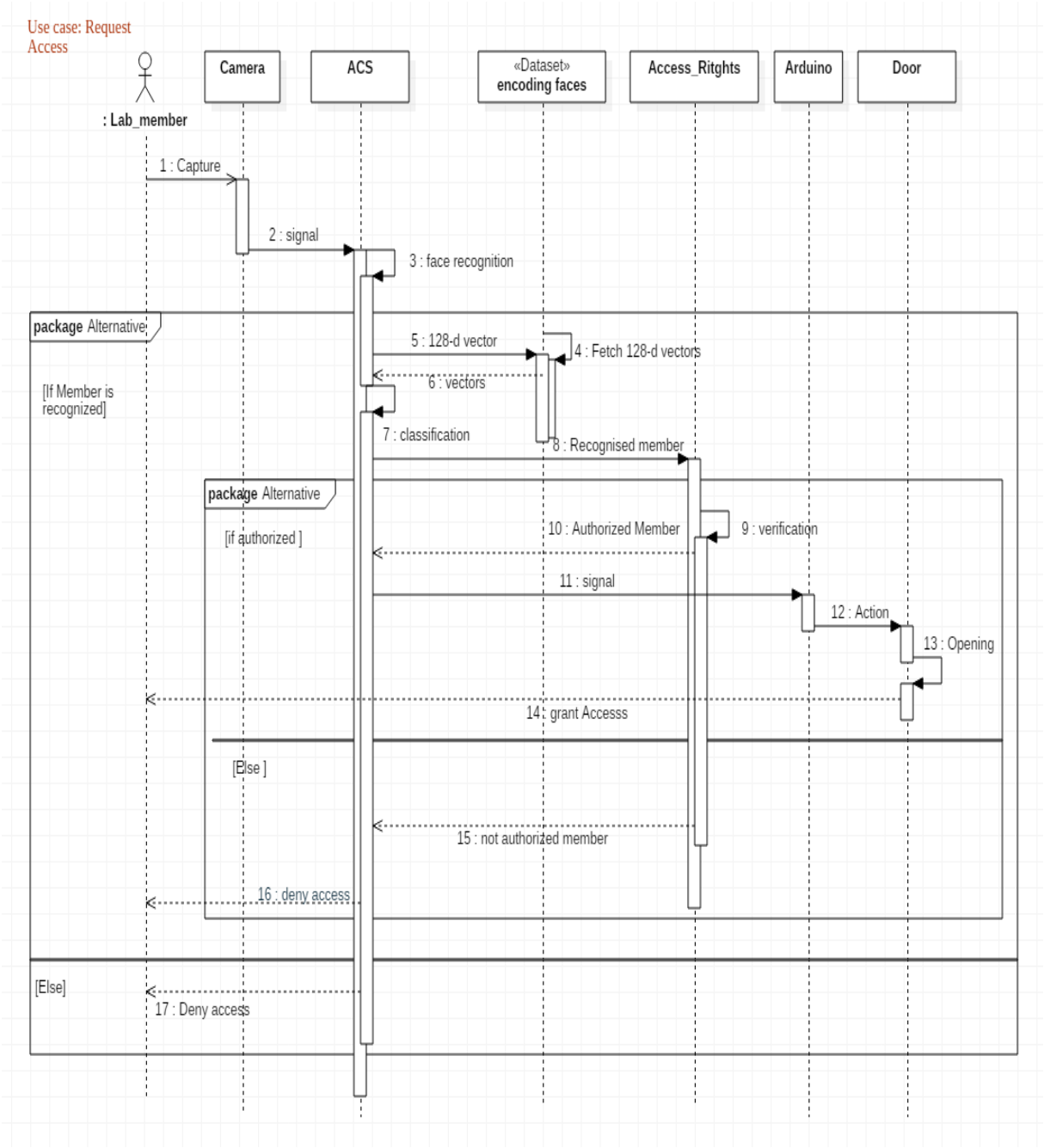


Figure 64:Request Access Sequence Diagram

3.1.3.2 Sequence « Alert Administrator »:

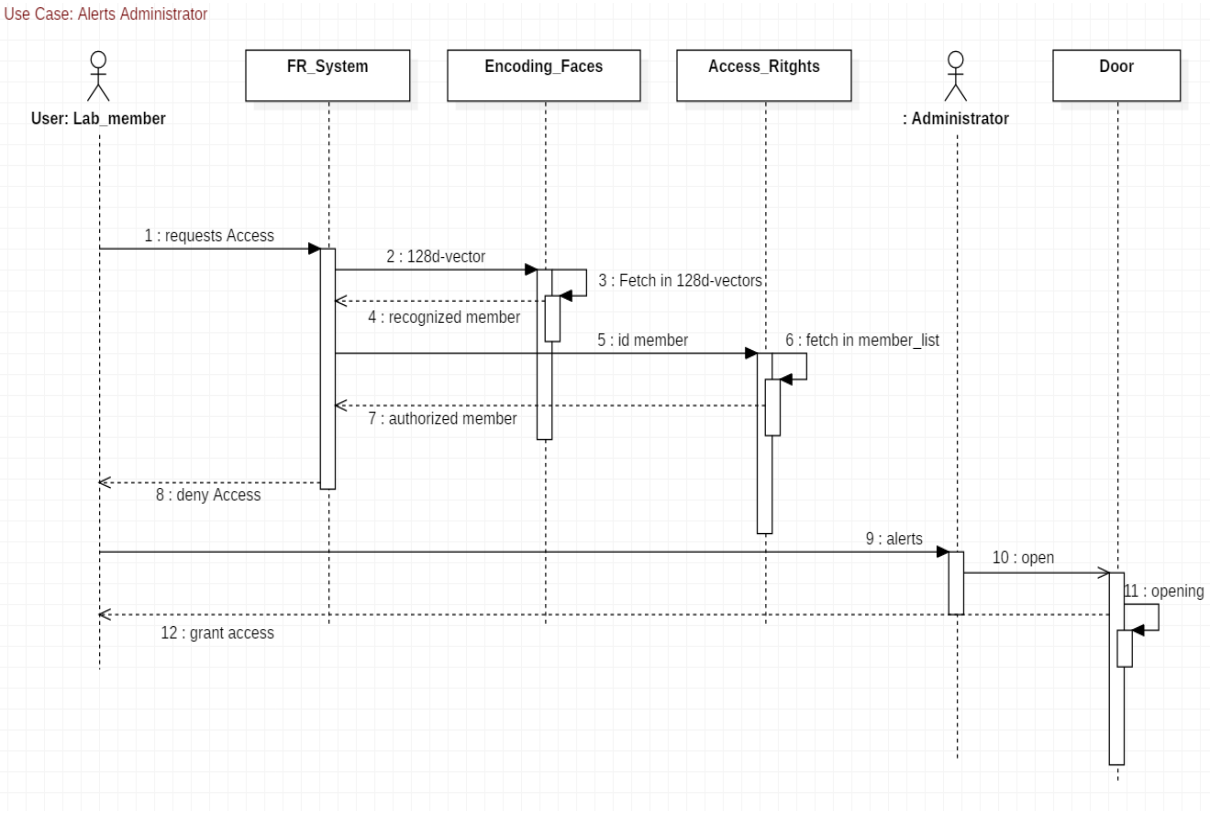


Figure 65: Alert Administrator Sequence Diagram

3.1.3.3 Sequence « Login »:

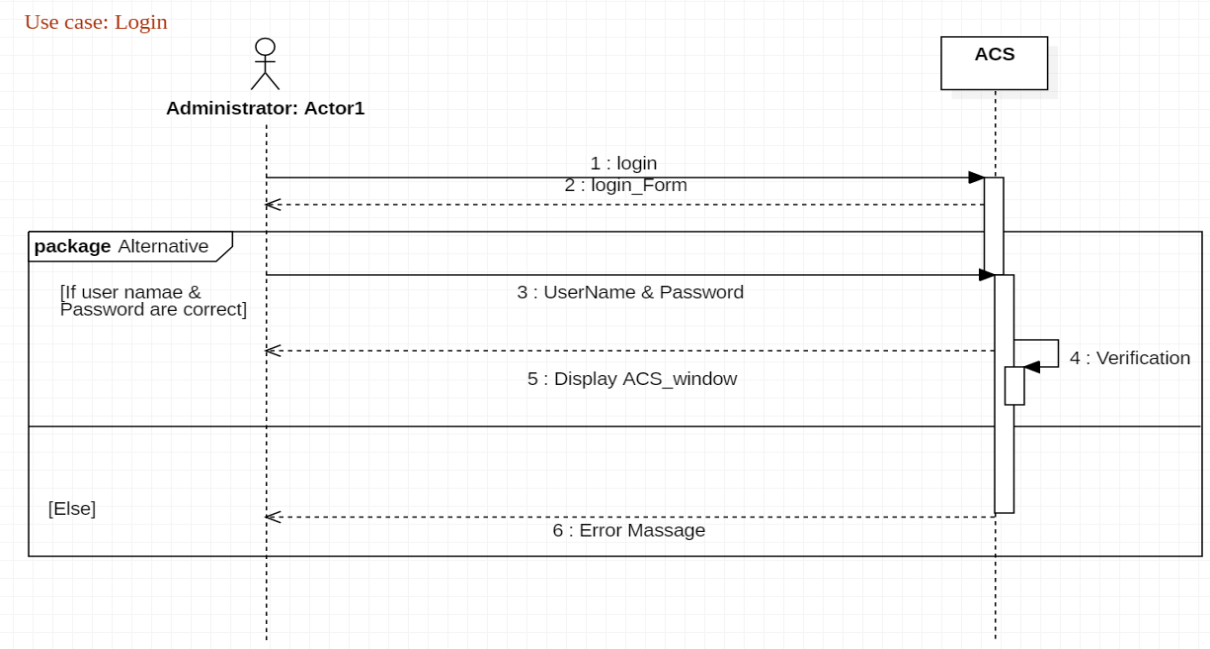


Figure 66: Login Sequence Diagram

3.1.3.4 Sequence « Add a new member »:

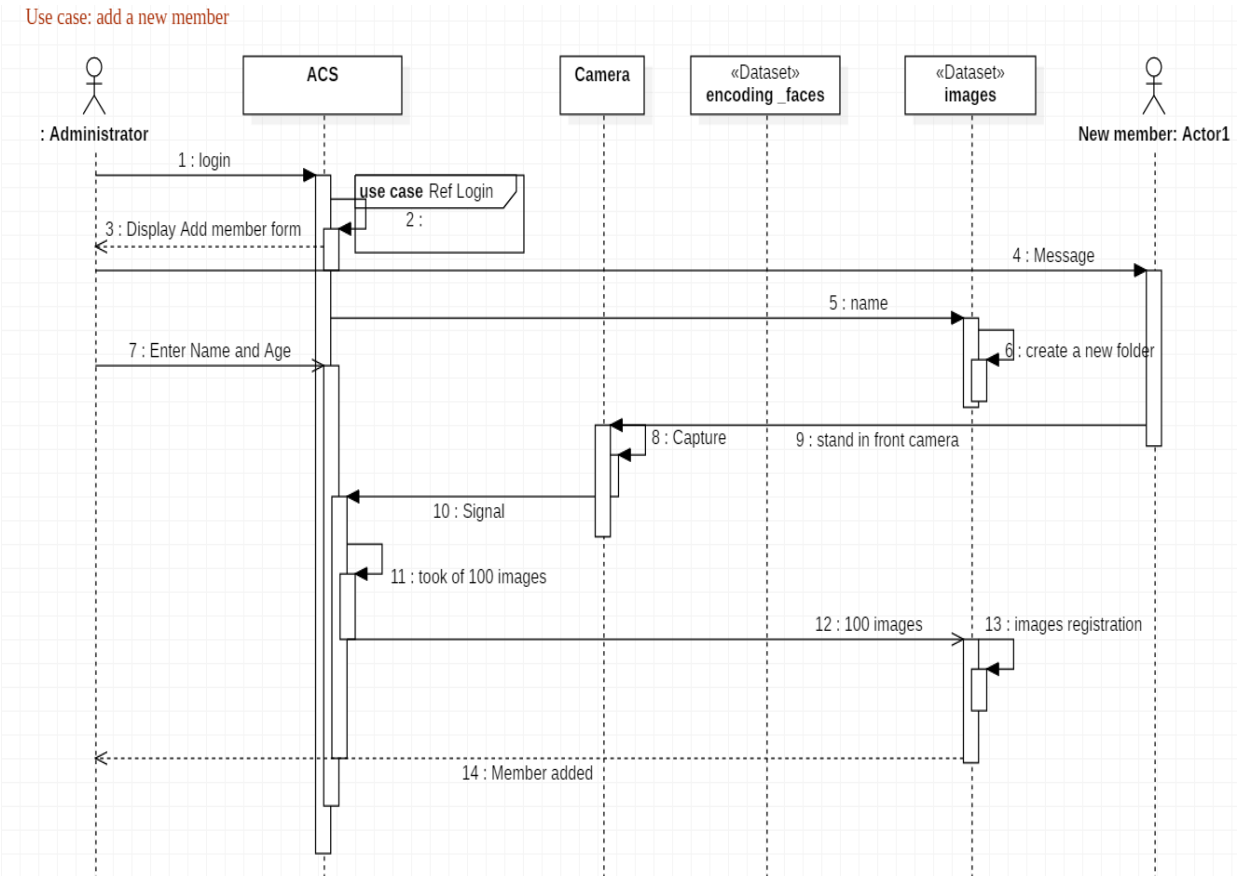


Figure 67: Add a New Member Sequence Diagram

3.1.3.5 Sequence « Updating system »:

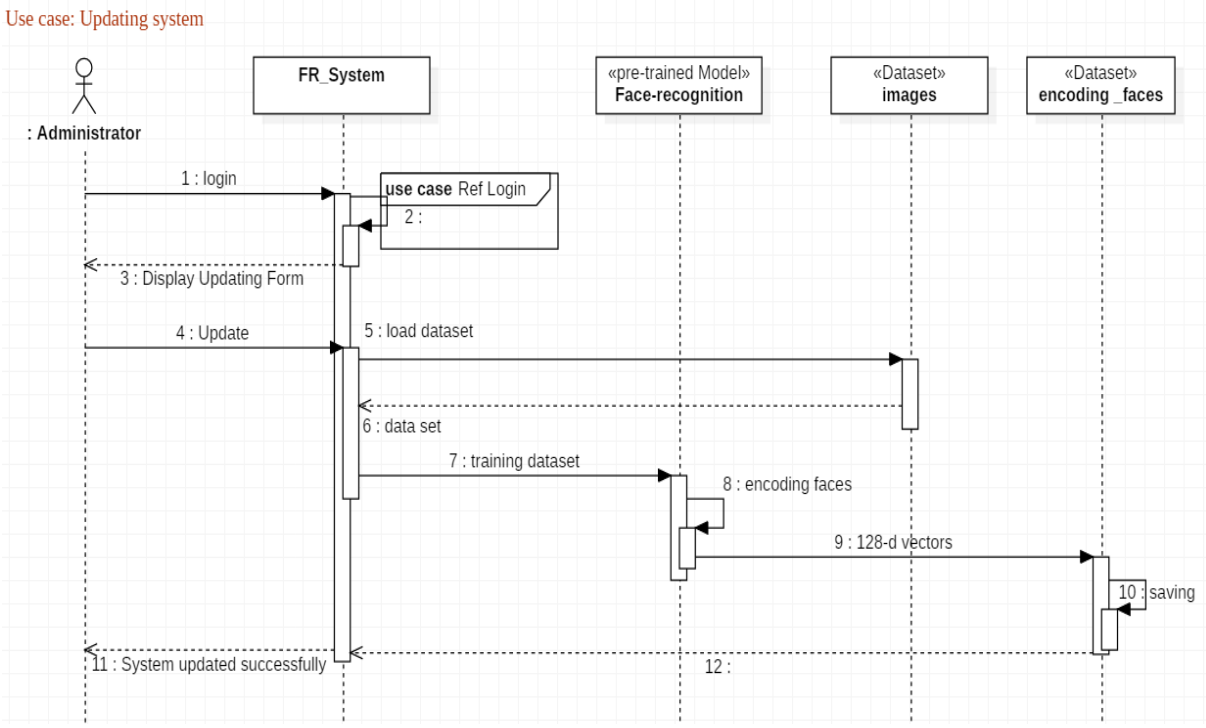


Figure 68: Updating System Sequence Diagram

4. Components of the proposed system:

In this section, we make the detailed design of our system by showing each of the four main steps, which represent the core of our proposed solution.

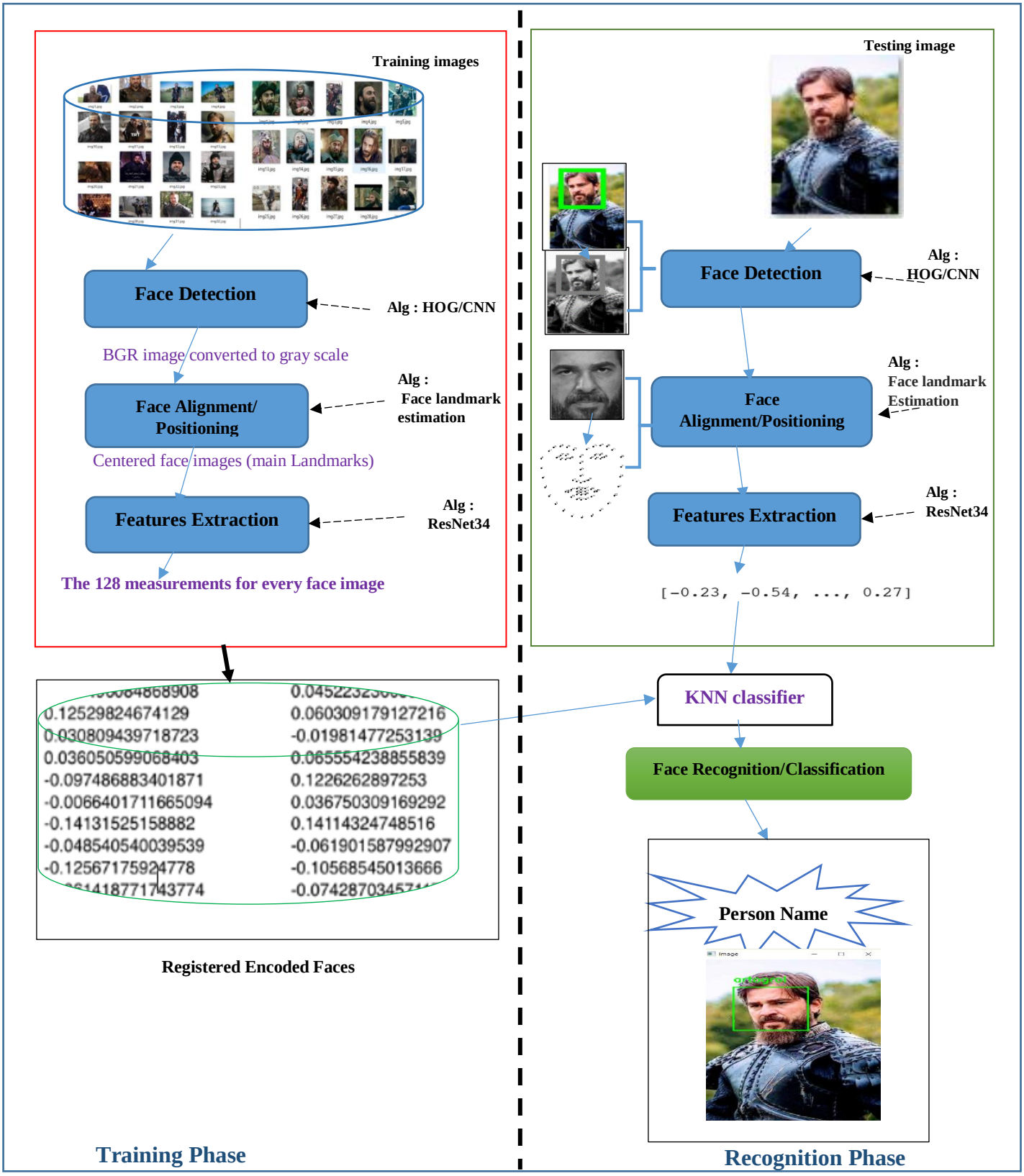


Figure 69: Proposed System Fonctioning

4.1 Step 1: Face Detection:

In this step, we are going to use a method invented in 2005 called Histogram of Oriented Gradients (HOG), it passes by:

- Find faces in an image and convert it to grayscale.
- Examines every single pixel in our image one at a time.
- For every single pixel in the image figure out how dark the current pixel is compared to the pixels, directly surrounding it in order to draw an arrow showing in which direction the image is getting darker.
- Repeat this process for every single pixel to end up with: every pixel was replaced by an arrow. These arrows are called gradients, which shows the flow from light to dark across the entire image. (Rosebrock, pyimagesearch.com/, June 18,2018). We do this process because if we analyze pixels directly, really dark images and really light images of the same person will have totally different pixel values. But by only considering the direction that brightness changes, both really dark images and really bright images will end up with the same exact representation.

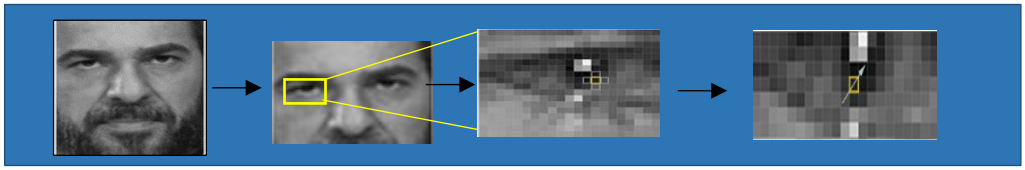


Figure 70: Pixels replaced by Gradients

Break up the image into small squares of 16x16 pixels each and in each square count up how many gradients point in each major direction (how many point up, point up-right, point right, etc…). Then it will replace that square in the image with the arrow directions that were the strongest. The last result is a very simple representation of the structure of the image called HOG version of the image (Figure 71).

Figure 71: HOG versin of image

To find faces in this HOG image, all we have to do is find the part of the image that looks the most similar to a known HOG pattern that was extracted from a bunch of other training faces. With this technique it is easily to find face in any image. (Rosebrock, pyimagesearch.com/, June 18,2018).

4.2 Step 2: Posing and Projecting Faces:

This step is crucial to solve the problem that faces turned in different directions look totally different to a computer and that by using an algorithm called face landmark estimation in order to warp each picture so that the eyes and lips are always in the sample place in the image. This approach, which the main idea is to come up with 68 specific common points (landmarks) that exist on each face. Then we will train a machine learning algorithm to be able to find these 68 specific points on any face.

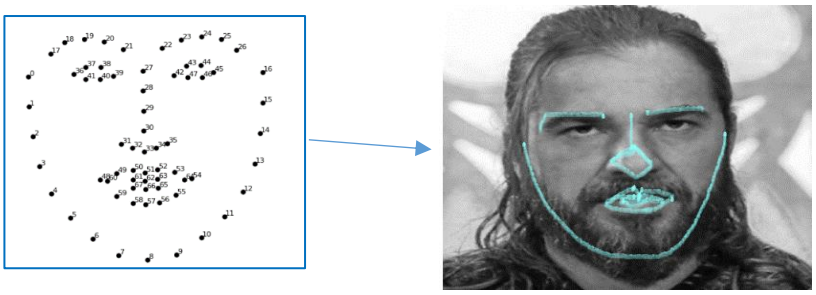


Figure 72: The 68 Landmarks

.

4.3 Step 3: Encoding Faces:

After step 2, now we have the main points of the face, which can be fed to the convolutional neural network for training, to extracting features and to classification. Our network architecture for face recognition is based on ResNet-34 from the Deep Residual Learning for Image Recognition paper by He et al[5]., but with fewer layers and the number of filters reduced by half. The network itself was trained by Davis King[6] on a dataset of ≈3 million images. On the Labeled Faces in the Wild (LFW) dataset.

4.4 Step 3.1 Feature extraction:

As we have mentioned at the beginning of this chapter, we will use deep metric learning in the training process by using a single 'Triplet' training step to generate a 128 measurement for each face in our dataset. The input, are tree images instead of one, two different images for the same person and the third image is of a totally different person,

[5] https://arxiv.org/abs/1512.03385
[6] the creator and chief maintainer of **dlib** — a toolkit for real-world machine learning, computer vision, and data analysis in C++ (with Python bindings included, when appropriate).

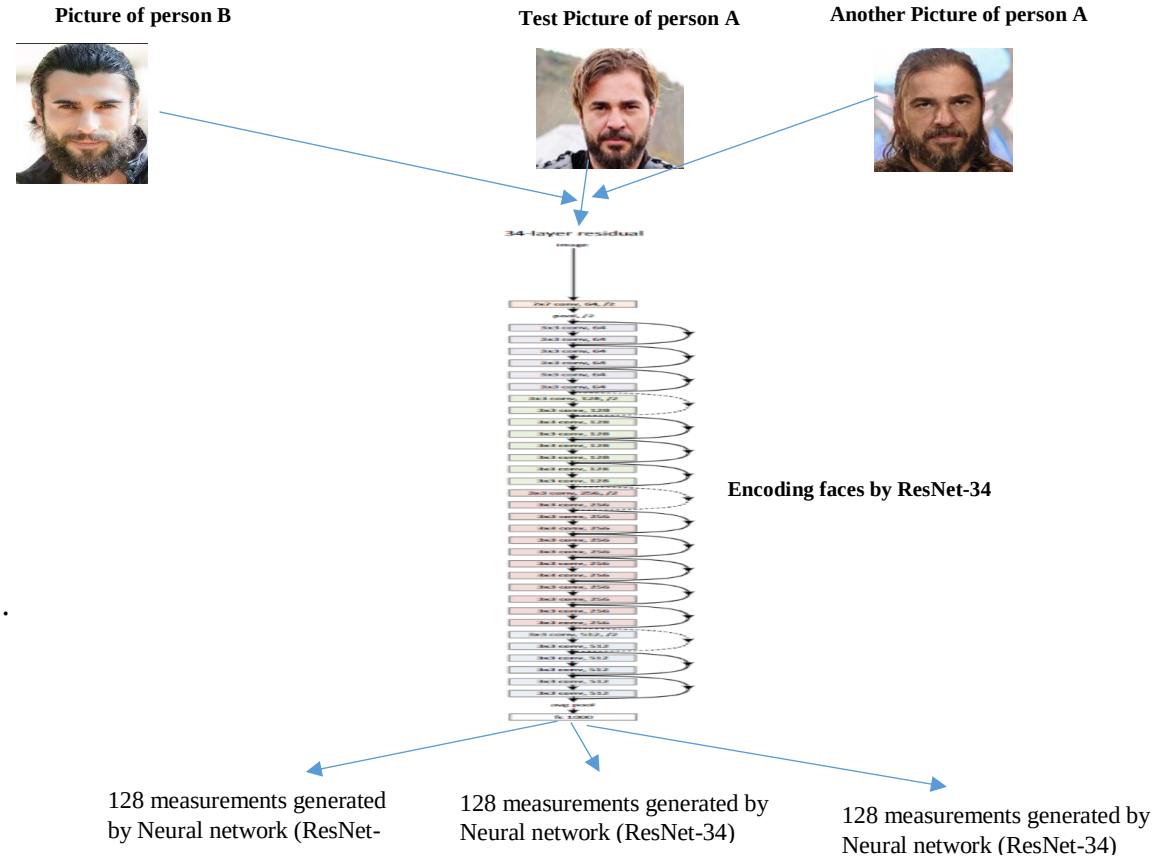


Figure 73: Triplet' training step

The algorithm compares the results of those 128 measurements of the three images then tweaks the neural network slightly so that it makes sure the measurements it generates for the first image and the second are slightly closer while making sure the measurements of the second image and of the third are slightly further apart. Machine learning people call the 128 measurements of each face an **embedding**. The exact approach for faces we are using was invented in 2015 by researchers at Google. (Rosebrock, pyimagesearch.com/, June 18,2018)

Figure 74: The 128 measurements generated from image

4.5 Step 4: Finding the person's name from the encoding (recognition):

This is the last step of the whole process of face recognition; we have used a simple machine learning algorithm, which is KNN, as classifier for the step of identification/recognition, and by finding the person in our database (encoding faces) of known person who has the closest measurements to our request image.

5. Conclusion:

We have demonstrate in this chapter our proposed system with more details from the input image until recognition to facilitate its implementation.

Chapter 4: System Implementation

1. Introduction:

The goal of this chapter is to present the implementation of our system and the main tools used for developing this system on one hand, and on the other hand, it represents the experimental results with discussion as well.

2. Overview of the Proposed System:

Our proposed system requires an image (video stream) of the person who wants access to the "Cinq laboratories" building as input, if the system recognizes the person it will grant the access for him, otherwise it will deny the access. When the system decides whether or not the captured image is close enough to declare a possible match or not (based on identification .i.e 1:N comparison process), it sends a signal to open the door [see the electronic part below] if the person is recognized, otherwise the system will not send any signal and therefore deny the access. This process is done in a real-time manner. The Figure 75 depicts an overview of the system.

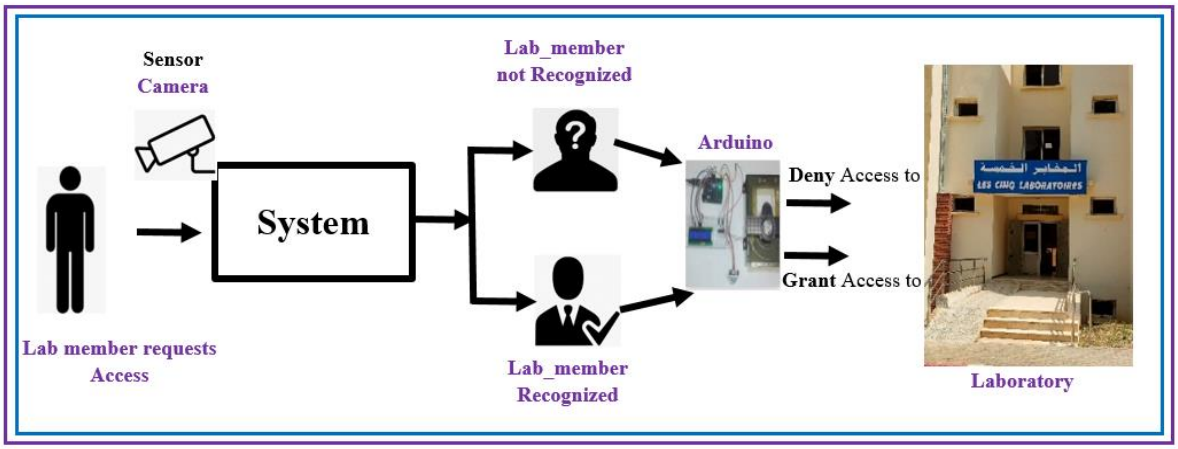


Figure 75: Overview of the proposed System

2.1 Electronic Part (Porotype):

Our realized porotype consists of the following components:

2.1.1 Arduino Mega (shield + driver A4988):

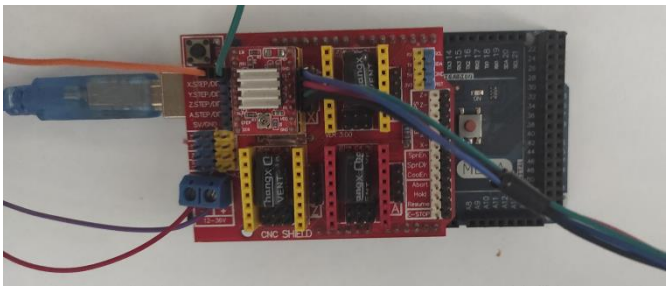


Figure 76: Arduino Mega Card

2.1.2 Sliding door Model:

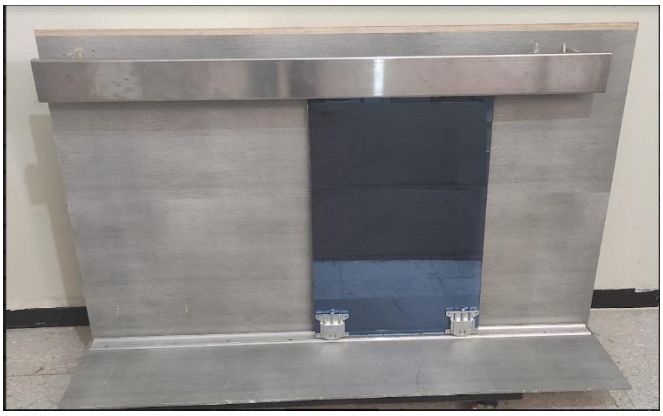


Figure 77: Sliding Door protoype

2.1.3 Power Supply:

Figure 78: Power Supply

2.2 Software and development environment:

2.2.1 Anaconda Navigator:

It is a desktop graphical user interface (GUI) included in Anaconda® distribution that allows to launch applications and easily manage conda packages, environments, and channels without using command-line commands. Navigator can search for packages on Anaconda.org or in a local Anaconda Repository. In this study, the version used is 2.1.2 (2016)[7]

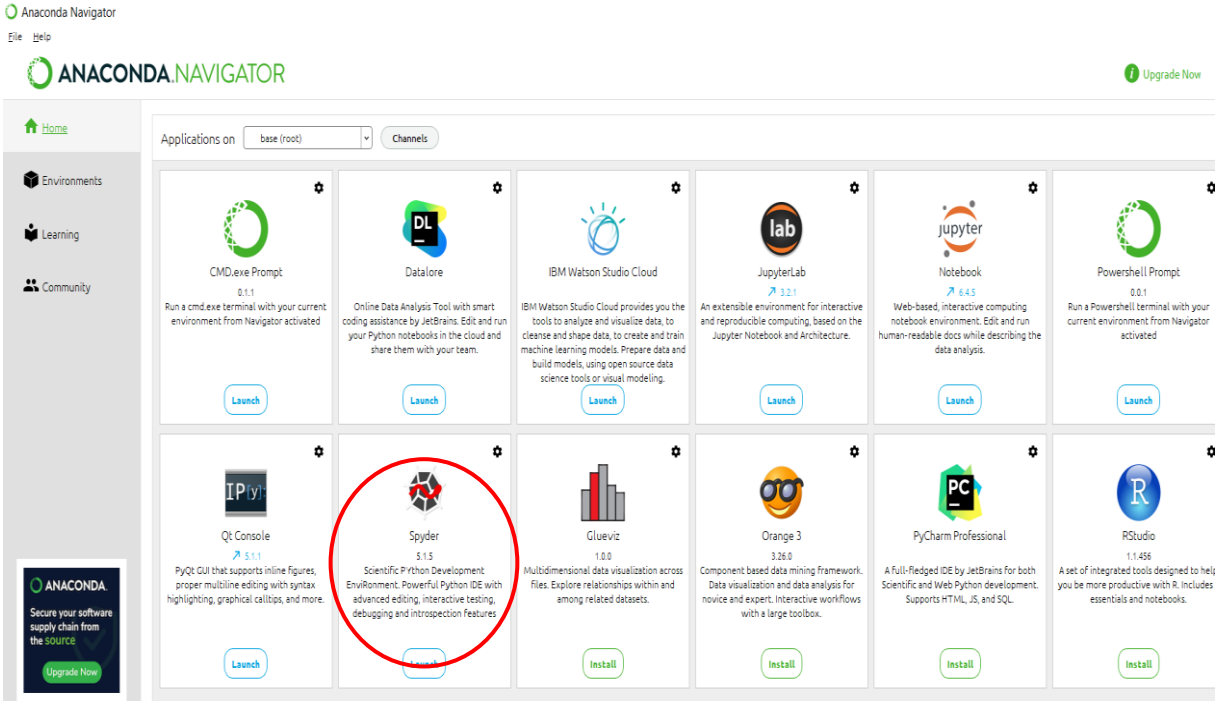


Figure 79: Anaconda Navigator

2.2.2 Spyder[8]:

It is a free and open source scientific environment written in Python, for Python, and designed by and for scientists, engineers and data analysts. It features a unique combination of the advanced editing, analysis, debugging, and profiling functionality of a comprehensive development tool with the data exploration, interactive execution, deep inspection, and beautiful visualization capabilities of a scientific package. In the present study we used version: 5.1.5

[7] https://www.anaconda.com/
[8] https://www.spyder-ide.org/

2.2.3 Python:

According to (Python, s.d.)Is a powerful programming language. It has efficient high-level data structures and a simple but effective approach to object-oriented programming. Python is an interpreted language, which can save you considerable time during program development because no compilation and linking is necessary. The interpreter can be used interactively, which makes it easy to experiment with features of the language, to write throw-away programs, or to test functions during bottom-up program development. It is also a handy desk calculator. Python enables programs to be written compactly and readably. Programs written in Python are typically much shorter than equivalent C, C++, or Java programs, for several reasons:

- the high-level data types allow you to express complex operations in a single statement;
- statement grouping is done by indentation instead of beginning and ending brackets;
- no variable or argument declarations are necessary.

Python is a programming language that supports the creation of a wide range of applications. Developers considered it as a great choice for Artificial Intelligence (AI), Machine Learning, and Deep Learning projects. It gives access to great libraries and frameworks for AI and machine learning (ML), like Numpy, Scipy, Pandas, Matplotlib; frameworks like Theano, TensorFlow, Keras and others. We have used version: 3.9.7 for 64-bit | Qt 5.9.7 | PyQt5 5.9.2 | (for Windows 10):

2.2.4 Libraries:

In our system, we used two main libraries: Dlib C++ and Face_recognition, and others libraries such as OpenCv and imulitis.

2.2.4.1 Dlib C++:

Dlib is an open source licensing modern C++ toolkit containing machine learning algorithms and tools for creating complex software in C++ to solve real world problems. Maintened by Davis King contains the implementation of “deep metric learning” which is used to construct the face embedding used for the recognition process. Dlib (pretrained model) based essentially on a version of the ResNet-34 network from the paper Deep Residual Learning for Image Recognition by He,

Zhang, Ren, and Sun[9] with 29 conv-layers (few layers removed) and the number of filters per layer reduced by half. It has an accuracy of 99.38% on the standard Labeled Faces in the Wild benchmark-LFW with the total of individuals =7485 (dataset). (Rosebrock, pyimagesearch.com/, 18 June 2018).

2.2.4.2 Face_recognition:

It is used to recognize and manipulate faces using Python. It is developed using dlib's facial recognition based on "deep learning".

2.2.4.3 OpenCV:

Is a free graphics library, originally developed by Intel, specialized in real-time image processing.

2.2.4.4 imutils:

A series of handy functions to facilitate processing functions basic image functions such as translation, rotation, resizing, skeletonization and display of Matplotlib images with OpenCV and Python 2.7 and Python 3.

2.2.5 Hardware:

The figure below displays the features of the computer used in our experimental work.

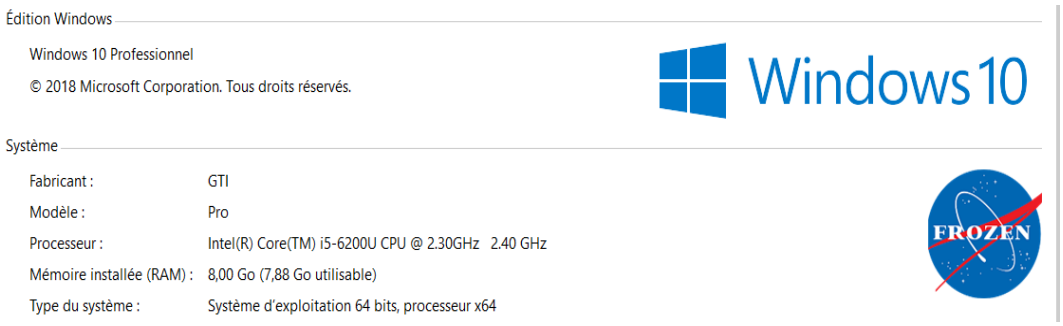


Figure 80:Computer features

[9] He, Kaiming Zhang, Xiangyu Ren, Shaoqing Sun, Jian , Deep residual learning for image recognition, 2016, DOI:10.1109/CVPR.2016.90

3. Dataset:

Our data set consists of 900 images taken from 09 subjects, with 100 images for each subject (laboratory member).

We used Logitech C270 HD WEBCAM to take the images under the following conditions.

Place: indoor place (office)

Pose: different poses

Lightning: indoor lightning

Gender: mixed

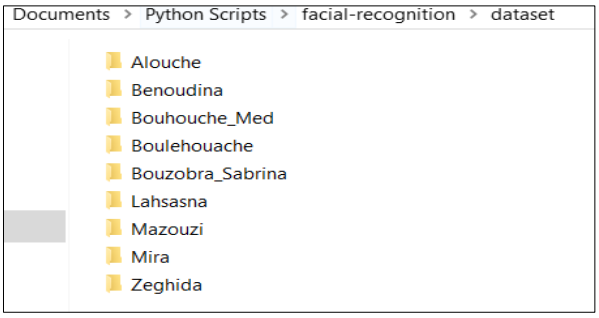


Figure 81: the 09 subjects consisting our dataset

In order to obtain reliable results, 5 cross-validation method (or 5-cv) is applied as resampling method. This method is good for avoiding bias as it allows for testing the entire dataset and determines how well the model performs. (20 images * 9 subjects* 5 times =900 images).

3.1 The 5-Cross-validation Method for testing the accuracy of the system:

Figure 82: K-kross validation method for testing

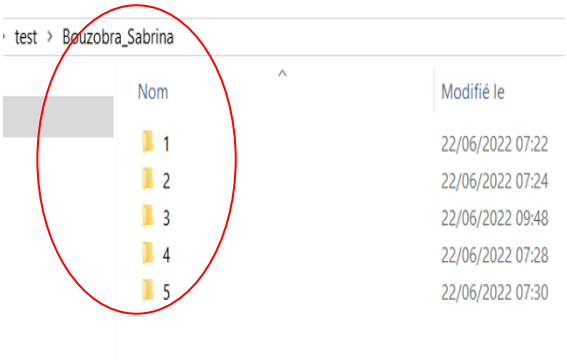


Figure 83*: Bouzobra subject- folder divided into 5 sub-folders*

3.2 Collecting Dataset:

We have used:

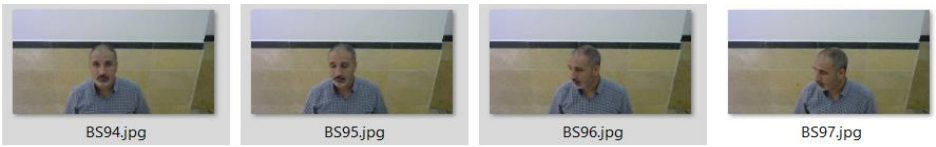
Device: Camera C270 HD WEBC
Place: indoor place (office)
Pose: different poses
Lightning: indoor lightning
Gender: mixed

3.2.1 Samples from our dataset: In the following, some smple images are provided for each subject:

Subject1: Benoudina

Subject 2: Bouhouche Med

Subject 3: Boulehouache



Subject 4: Mira

Subject 5: Lahsasna

Subject 6: Zegida

Subject 7: Mazouzi

Subject 8: Alouche

Subject 9: Bouzobra

4. Implementation:

4.1 Training Phase:

4.1.1 Processing images:

On the command prompt, we type

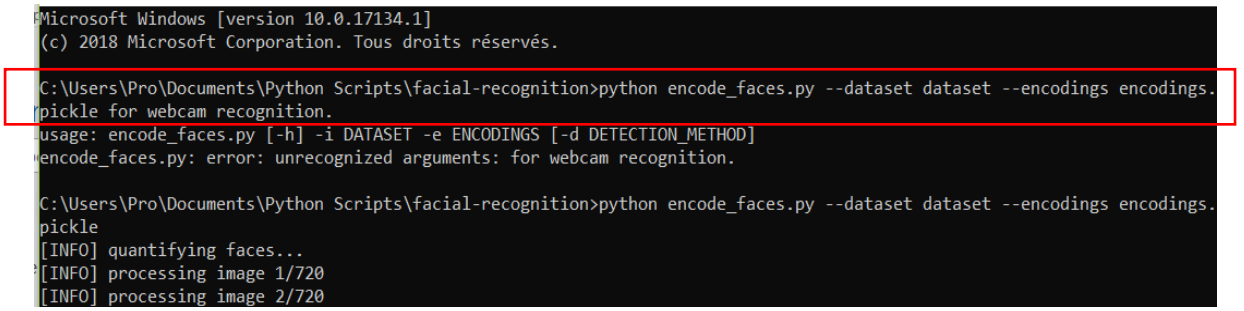


Figure 84: processing images of the training dataset

4.1.2 Encoding images:

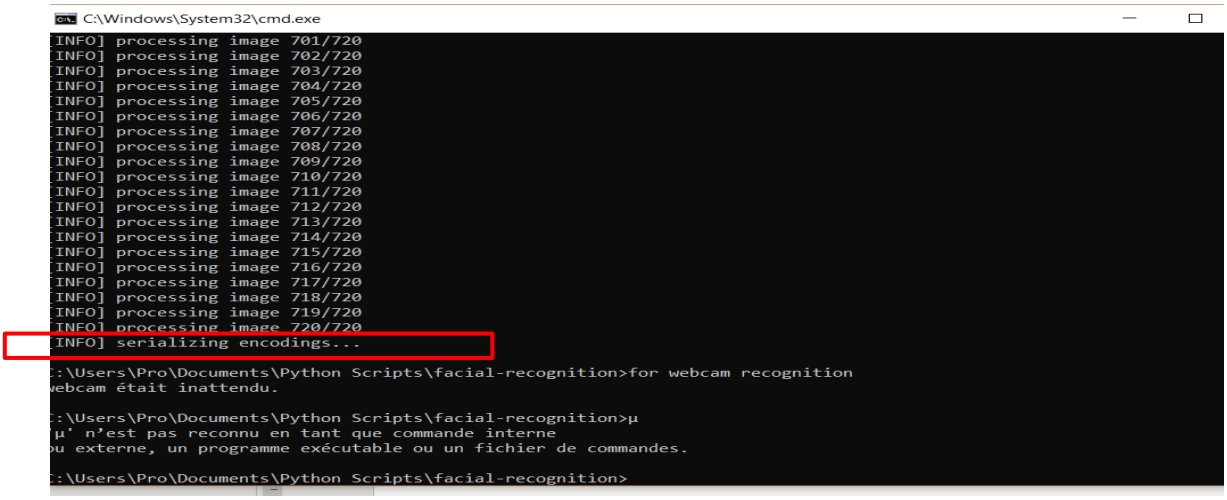


Figure 85: encoding images of the training dataset

4.2 Testing:

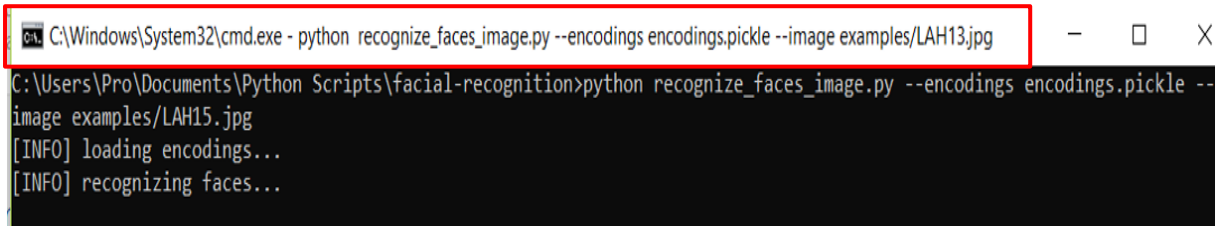


Figure 86: testing example picture: LAH13.jpg

5. Results and discussion:

After running the test command using a testing dataset in 5 iterations, we got the following results using the testing error rate as a performance metric. It can be calculated as:

Testing Accuracy (or recognition) rate = $\frac{Number\ of\ correctly\ classified\ testing\ images}{Number\ of\ testing\ images}$

5.1 First Iteration:

N°	Subject	Recognition (image)		Accuracy rate by subject (%)
		True	False	
1	Alouche	20	0	100
2	Benoudina	20	0	100
3	Bouhouche	20	0	100
4	Boulehouache	20	0	100
5	Bouzobra	20	0	100
6	Lahsasna	7	13	35
7	Mazouzi	20	0	100
8	Mira	20	0	100
9	Zeghida	20	0	100
Average rate recognition (%)				**92,78**

Table 11: Accuracy rate by subject- First iteration

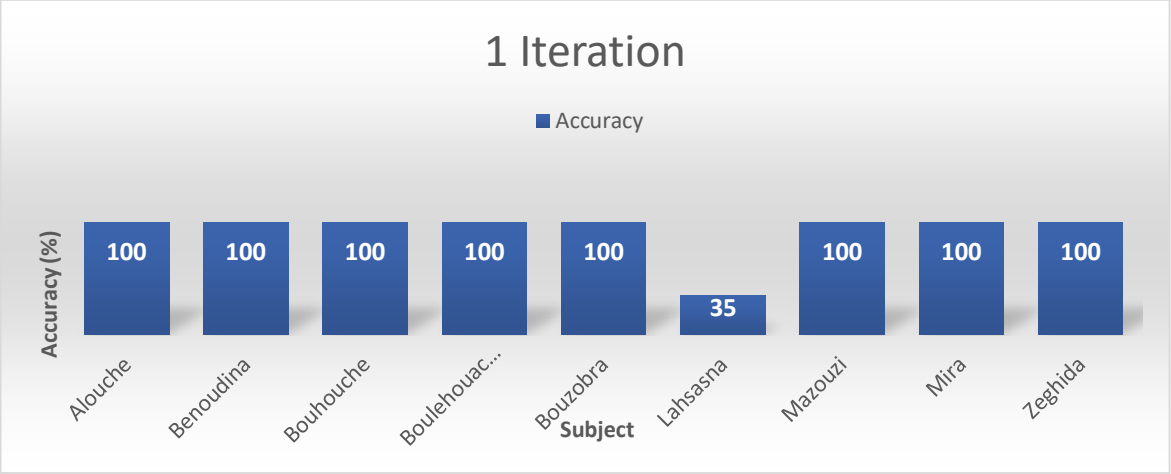


Figure 87: Accuracy rate in the first iteration

Discussion:

as 35% only. The system predicted correctly only 7 images while the other 13 images were recognized as Benoudina.


as 35% only. The system predicted correctly only 7 images while the other 13 images were recognized as Benoudina By checking the image below -in the middle- we can say that the misclassification is due to the similarity between these two subjects. (interclasses similarity).

5.2 Second Iteration:

N°	Subject	Recognition (image)		Accuracy rate by subject (%)
		True	False	
1	Alouche	20	0	100
2	Benoudina	20	0	100
3	Bouhouche	20	0	100
4	Boulehouache	20	0	100
5	Bouzobra	20	0	100
6	Lahsasna	15	5	75
7	Mazouzi	20	0	100
8	Mira	20	0	100
9	Zeghida	20	0	100
Average rate recognition (%)				**97,22**

Table 12: Accuracy rate by subject-Second iteration

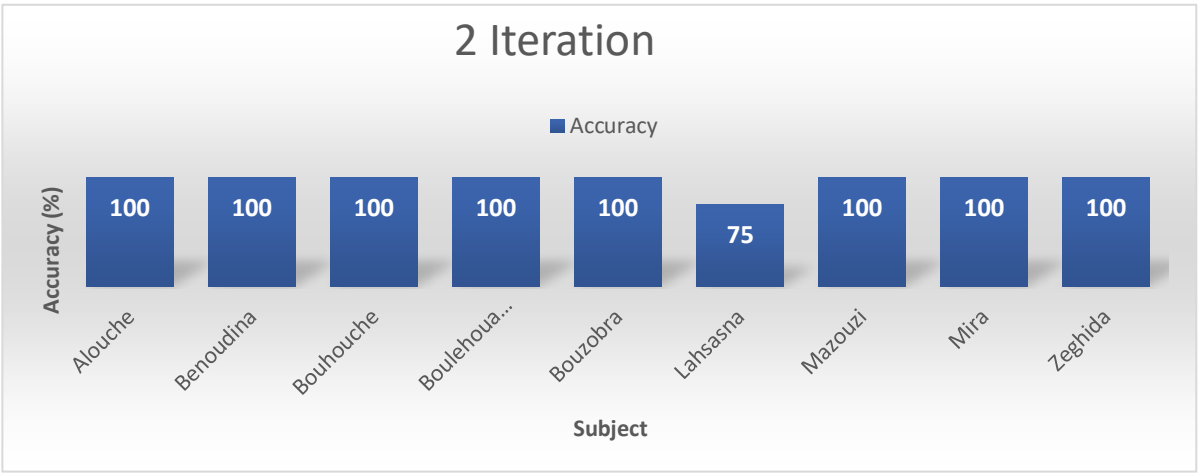


Figure 87: Accuracy rate in the second iteration

Discussion:

5.3 Third Iteration:

N°	Subject	Recognition (image)		Accuracy rate by subject (%)
		True	False	
1	Alouche	20	0	100
2	Benoudina	20	0	100
3	Bouhouche	20	0	100
4	Boulehouache	19	1	95
5	Bouzobra	20	0	100
6	Lahsasna	15	5	75
7	Mazouzi	20	0	100
8	Mira	20	0	100
9	Zeghida	20	0	100
Average rate recognition (%)				**96,67**

Table 13: Accuracy rate by subject- Third iteration

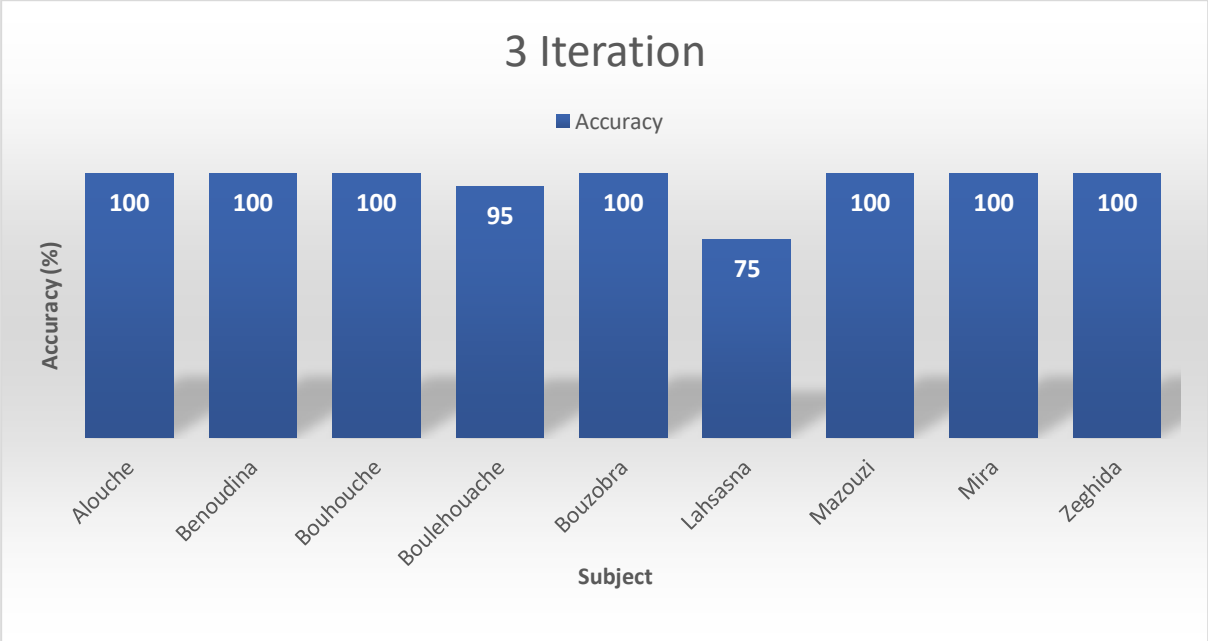


Figure 88: Accuracy rate in the third iteration

In this iteration, we can see that all the images for the 07 subjects were recognized correctly (true prediction) except for two subjects. The first one is Lahsasna whose recognition

rate was 75%, i.e. like the previous iteration and the second one is Boulehouache where the system achieved 95% accuracy. specifically, tt predicted 19 images correctly and 1 image was recognized as Bouhouche_Med.

By checking the wrongly recognized image below -in the middle- we can say that the misclassification is due to the similarity between these two subjects. (interclasses similarity).

5.4 Fourth Iteration:

N°	Subject	Recognition (image)		Accuracy rate by subject (%)
		True	False	
1	Alouche	20	0	100
2	Benoudina	20	0	100
3	Bouhouche	20	0	100
4	Boulehouache	20	0	100
5	Bouzobra	20	0	100
6	Lahsasna	17	3	85
7	Mazouzi	20	0	100
8	Mira	20	0	100
9	Zeghida	20	0	100
Average rate recognition (%)				**98,33**

Table 14: Accuracy rate by subject- Forth iteration

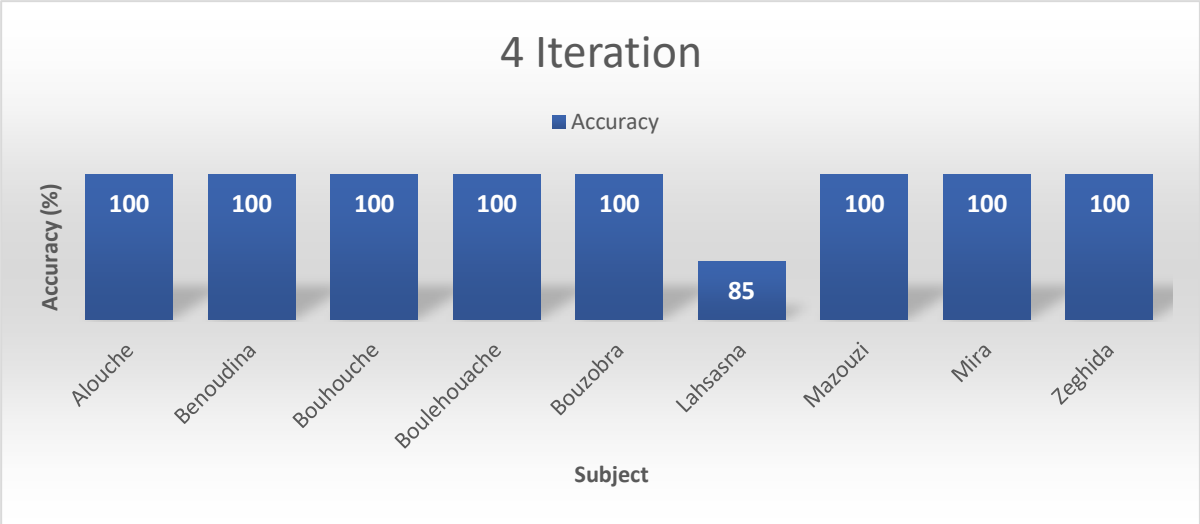


Figure 89: Accuracy rate in the fourth iteration

In this iteration and from the figure above, it can be seen that all the images of the 09 subjects were recognized correctly (true prediction) except for Lahsasna whose recognition rate was 85%. The system predicted this time 17 images correctly and 3 images were recognized as Benoudina.

5.5 Fifth Iteration:

N°	Subject	Recognition (image)		Accuracy /Subject (%)
		True	False	
1	Alouche	20	0	100
2	Benoudina	20	0	100
3	Bouhouche	20	0	100
4	Boulehouache	20	0	100
5	Bouzobra	20	0	100
6	Lahsasna	20	0	100
7	Mazouzi	20	0	100
8	Mira	20	0	100
9	Zeghida	20	0	100
Average rate recognition (%)				**100**

Table 15: Accuracy rate by subject- Fifth iteration

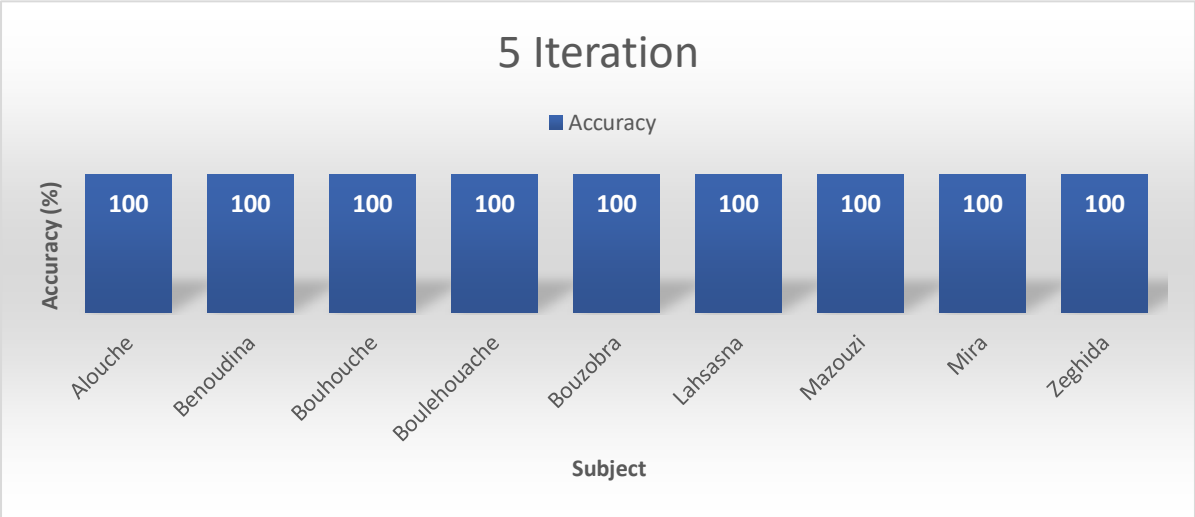


Figure 90: Accuracy rate in the fifht iteration

In this last iteration the system recognized all the images for all subjects correctly with classification accuracy of 100%.

To sum up, training the model took about 65 minutes to complete all the iterations. The testing accuracy rate achieved is relatively high (97%).

6. Analysis:

During and after training and testing phases, we can notice the following:

1. Accuracy:

The larger the amount of data, the higher the accuracy. This is one of the main features of deep learning. To test this characteristic, we run our model on a subset of our data set by reducing the number of subjects from 9 subjects (900 images) to only 5 subjects (538 images). The results are in the following table.

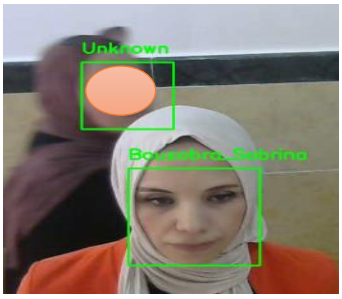
	Dataset	Images	Subjects	Accuracy

Dlib pre-trained model (based ResNet-34)	LFW	≈ 3 million	7485	99.38 %
Our model (based Dlib)	Our dataset	900	9	97 %
Our model (based Dlib)	Our dataset	538	5	94 %

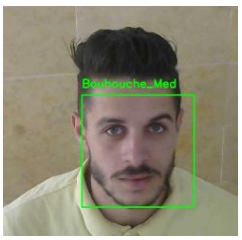
Table 16: Comparative Accuracy rate based on the amount of data

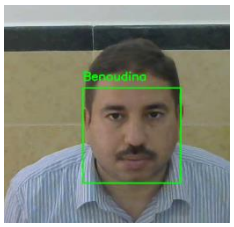
2. Recognition:

During the testing phase, using the cross-validation we tested the whole dataset, and we can state also that the system has classified all the images existing in the dataset with true or false prediction, and classify the unknown images (subject) as unknown like it is shown in the image below.



Also, as we can see in the images below when there is some degree of similarity between the faces (Interclass similarity), the model sometimes makes false predictions. This is indeed, one of the weak points of the facial recognition systems where the similarity between people can mislead the recognition system particularly in the case of twins or relatives.





3. Detection:

	Total images/Subjects	*Non detected images/subjects*	*Detected images/subjects*

Alouche	100	0	100
Benoudina	100	0	100
Bouhouche	100	4	96
Boulehouache	100	17	83
Bouzobra	100	0	100
Lahsasna	100	0	100
Mazouzi	100	38	62
Mira	100	0	100
Zeghida	100	6	94
Total	**900**	**65**	**835**
Average images Non-detected		**7%**	
Average images Detected		**93%**	

Table 17: Detection Rate

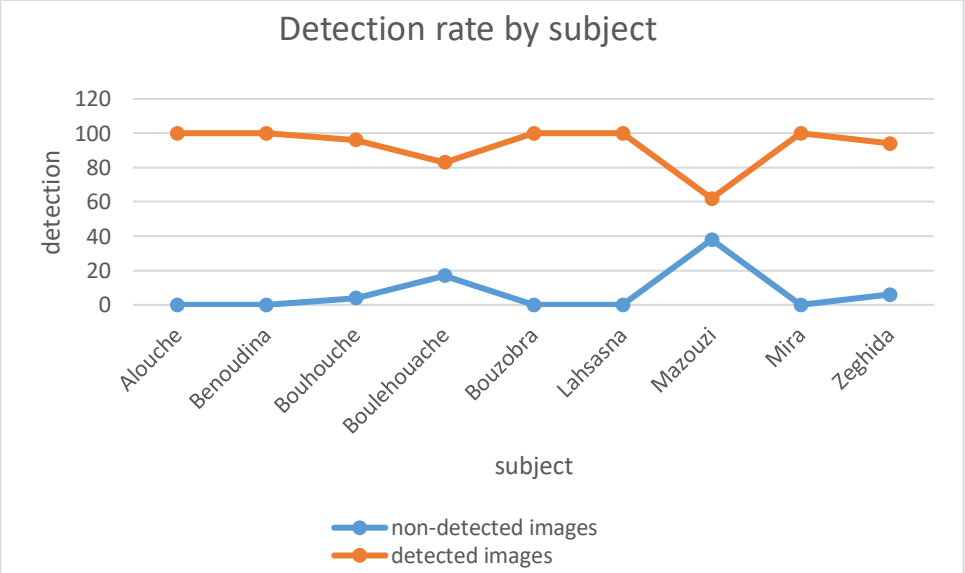


Figure 91: Detection System rate

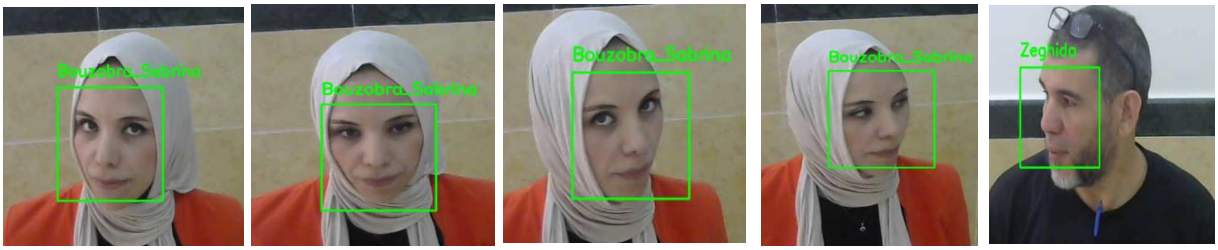
From the table we can notice that the detection rate is 93%. The following are sample images that were not detected.

During the test phase, we noticed that faces in some images were not detected, most of them belonging to subject Mazouzi (38 images), 17 images to subject Boulehouache, 6 images to subject Zeghida and 4 images to subject Bouhouche.

Analyzing these pictures, we can conclude that faces are not on the front side, or not clear, so the algorithm (HOG) in the detection step can not detect the face. Therefore, our system detection uses the frontal face detector (import from Dlib library).

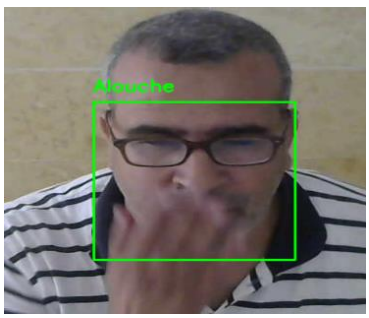
4. Pose Variation:

In our dataset images, we tried to deal with different positions of the face and check the impact of these positions on the performance of our system. we noticed that small variations in position of the subjects do not affect the performance of the recognition system. Below are sample images that were correctly recognized although the subject is not directly looking at the camera.



In addition, images of subjects with certain expressions such as smiling, blinking were successfully recognized.

Although the subject shown below puts his hand on his face. The system correctly recognizes him.



7. Conclusion:

In this chapter, we provide an overview of the implementation side of our access control system based on face recognition. In particular, a description of the data set was provided along with the development environment tools. Special attention was given to the explanation of how the 5 cross-validation method was utilized to resampling the data set in order to avoid the bias of the results. In addition, the results achieved by our system were presented along with some comments and analysis of the key findings and features of the system. The proposed system achieved an accuracy rate as high as 97%. Finally, some challenging cases like pose variation and face with occlusion were presented to show how our system could handle these cases efficiently.

General Conclusion:

In this work, we have implemented a biometric access control system for the Cinq Laboratoires, located at the University of Skikda, using face recognition based on deep learning, which is a biometric approach to solve the problem of uncontrolled access to the laboratory. This system achieved an accuracy rate as high as 97%, which is considered relatively high taking into account the small amount of our database (900 images). This accuracy can be improved, as we showed in the study, by increasing the amount of data used for training the system. In our experimental work, the accuracy rate has been increased from 94% to 97% by just adding more examples to our database. As we have seen in the previous chapter there is the issue of interclass/similarity between four subjects (Boulehouache and Bouhouche) on one side, and (Benoudina and Lahsasna) on the other side. This problem has a negative impact on the overall performance of the system. Some other challenging cases like pose variation and face with occlusion were successfully handled by our system.

Perspectives:

In the future, we intend to deepen our research in the field of facial recognition by using other algorithms and approaches or by combining several methods. For this, we can further develop our system by associating another biometric modality such as the fingerprint to get a more effective system. We plan also to propose methods to overcome the interclass/similarity between the subjects in order to make our system more reliable solution.

Bibliography

Agrawal, A. (Jan 11, 2021). Retrieved from towardsdatascience.com: https://towardsdatascience.com/the-why-and-the-how-of-deep-metric-learning-e70e16e199c0

AlliedMarketResearch.com. (n.d.). *alliedmarketresearch.com*. Retrieved from https://www.alliedmarketresearch.com/biometric-technology-market

bayometric.com. (n.d.). *https://www.bayometric.com*. Retrieved from BAYOMETRIC: https://www.bayometric.com/biometric-system-architecture

Beysolow II, T. (2017). *Introduction to Deep Learning using R.* San Francisco, California, USA. doi:10.1007/978-1-4842-2734-3

Bhatia, R. (May 2013). Biometrics and Face Recognition Techniques. *International Journal of Advanced Research in, Volume 3*(3, Issue 5), 93.

Bhatia, R. (May 2013). Biometrics and Face Recognition Techniques. *International Journal of Advanced Research in, Volume 3*(Issue 5).

BiometricUpdate.com. (Jul 12, 2013). Retrieved from BiometricUpdate.com: https://www.biometricupdate.com/201307/explainer-retinal-scan-technology

bsia, British security industry association. (April 2016). A specifier's guide to access control systems. doi:Form No,132/Issue 4

Council, National Research. (2011). *Prudent practices in the laboratory : handling and management of chemical hazards.* Washington, D.C., USA: THE NATIONAL ACADEMIES PRESS. Retrieved from www.nap.edu

Damer. (April 2018, April). Application-driven Advances in Multi-biometric Fusion.

Datta, A. K., Banerjee, P., & Madhura Datta, D. (2016). *Face Detection.* (T. &. Group, Ed.) CRC Press.

Dr.hlaing-Htake-Khaung-Tin. (December 2012). Personal Identification and Verification using Palm print Biometric. *nternational Journal of Latest Technology in Engineering Management and Applied Science (IJLTEMAS), vol. 1,*. Retrieved from ResearcheGate: https://www.researchgate.net/profile/Drhlaing-Htake-Khaung-Tin

German, R., & Barber, K. (Sep 2017). *Current Biometric Adoption and Trends.* The University of Texas, Austin. doi:UT CID Report #18-02

Guennouni, S., Mansouri , A., & Ahaitouf, A. (October 19th, 2018). *https://www.intechopen.com/.* (E. b. Catalá, Ed.) doi:10.5772/intechopen.84845

hemeinguie. (april 2016). *.(https://www.techopedia.com).* al.

IBM Cloud Education . (October 2020). Retrieved from www.ibm.com: https://www.ibm.com/cloud/learn/convolutional-neural-networks

IJCSMC. (January 2013). Face and Facial Expression Recognition -. *International Journal of Computer Science and Mobile Computing, Vol. 2, Issue. 1,*. doi:ISSN 2320–088X

Jha, A. (nd). Class Room Attendance System Using Facial. *The International Journal of Mathematics, Science, Technology and Management*.

Karamizadeh, S., Shahidan , M., & Zamani, M. (September -2013). An Overview of Holistic Face Recognition. *International Journal of Research in Computer and Communication Technology, Vol 2*(Issue 9).

Karlskin, B. (Sep 16, 2021). Retrieved from https://www.keyo.co: https://www.keyo.co/biometric-news/biometric-access-control-systems-101-everything-you-should-know#:~:text=Access%20control%20refers%20to%20the,common%20applications%20for%20access%20control.

Karlskin, B. (Sep 16, 2021). Retrieved from keyo: https://www.keyo.co/biometric-news/biometric-access-control-systems-101-everything-you-should-know#:~:text=Biometric%20access%20control%20is%20one,%2C%20turnstile%2C%20elevator%2C%20etc.

kaspersky. (n.d.). *kaspersky.com/resource-center/definitions/biometrics*. Retrieved from https://www.kaspersky.com/: https://www.kaspersky.com/resource-center/definitions/biometrics

Kisi. (n.d.). *Unlocking the future :A guide for getting started with access control.* Retrieved from https://www.getkisi.com

Lal, M., Kumar, K., Arain, R., Maitlo, A., Ali Ruk, S., & Shaikh, H. (2018). Study of Face Recognition Techniques: A Survey. *International Journal of Advanced Computer Science and Applications, 9*(6).

Mahmut KAYA Hasan , & B˙ILGE, S. (2019). Deep Metric Learning: A Survey. *MDPI*, 2-26.

Parmar, D. N., & Mehta, B. (January 2014). Face Recognition Methods & Applications. *International Journal of Computer Applications in Technology, Vol 4*.

Patel, R., Rathod, N., & Shah, A. (November 2012). Comparative Analysis of Face Recognition Approaches:. *International Journal of Computer Applications, Volume 57*(17).

Pattanayak, S. (2017). *Pro Deep Learning withTtensorFlow.* Bangalore, Karnataka, India. doi:ISBN-13 (pbk): 978-1-4842-3095-4

Python. (n.d.). Retrieved from https://www.python.org/: https://docs.python.org/3/tutorial/

Richa, & Josan, J. (January 2015). Face Recognition System – A Survey. *International Journal of Science and Research (IJSR), Volume 4 Issue1*. doi:ISSN (Online): 2319-7064

Rosebrock, A. (18 June 2018). *pyimagesearch.com/*. Retrieved from http://dlib.net/compile.html

Rosebrock, A. (June 18,2018). *pyimagesearch.com/*. Retrieved from https://pyimagesearch.com/2018/06/18/face-recognition-with-opencv-python-and-deep-learning/

Sava, J. A. (Feb 25, 2022). *https://www.statista.com*. Retrieved from https://www.statista.com/statistics/1048705/worldwide-biometrics-market-revenue/

Solomon, M. M., Meena, M. S., & Kaur, J. (April-June 2019). COMPARISON OF FACE RECOGNITION METHODS. *IJRAR- International Journal of Research and Analytical Reviews, 6*(2).

Swamynathan, M. (2017). *Mastering Machine Learning with Python in Six Steps.* Bangalore, Karnataka, India. doi:10.1007/978-1-4842-2866-1

techopedia.com. (September 5, 2018). Retrieved from techopedia: https://www.techopedia.com/definition/29707/access-control-system-acs

The National Academies Press Washington, D.C. (2011). *Prudent practices in the laboratory:handling and management of chemical hazards.*

United Security usi. (n.d.).). *(USI),(April 16, 2018). https://inbound.usisecurity.com*. Retrieved from https://inbound.usisecurity.com/blog/what-are-the-components-of-an-access-control-system