PEOPLE’S DEMOCRATIC REPUBLIC OF ALGERIA

MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH

M’HAMED BOUGARA UNIVERSITY –BOUMERDES

FACULTY OF TECHNOLOGY

Electrical Systems Engineering Department

Master's thesis

Presented by:

Mr SADAT Abdelmalek

Field of Study: Electronic

Option: Electronic of embedded systems

TITLE

Face recognition based on AI

Defended on 11 /07 /2024 in front of the jury composed of :

Noureddine	MESSAOUDI	Pr	University of Boumerdes	Chair
Samia	BELKACEM	MCA	University of Boumerdes	Supervisor
Samah	RIACHE	MCA	University of Boumerdes	Examiner

2023/2024

DEDICATION

My Sincere Gratitude to the Great Man, Mr. Sadat Abdelmalek may the passing days etch your name inlines of shimmering gold.

To My Dear Father, a Man of HonorIn the quiet corners of our hearts, the memory of my father resides like a cherished treasure. May Allah’s mercy envelop his soul, and may he rest in eternal peace.

To Zahoo, Nesrine, Ilham (Riri), Souhila, Chahira, Ratiba, and Fateh Thank you all for being in my life it’s like a beautiful melody that fills every corner of my heart. Your unwavering support and boundless love have been my guiding light. You are the most precious part of my journey.

And to S. Baki Ilham, my trusted ally on the path and my journey mate.

THANKS

First and foremost, I express my heartfelt gratitude to Almighty God, who has bestowed upon us health and determination to complete our work and excel in it. Praise be to God.

I extend my profound appreciation to Mrs. Belkacem Samia for her unwavering support, invaluable advice, and steadfast presence by my side.

Additionally, I want to thank my family whole heartedly for their tremendous support. May God protect each one of you on my behalf.

ملخص

توضح هذه الأطروحة تفاصي لإنشاء نظا مقو يللتعرف على الوجه يستفيد من الذكاءالاصطناعي (AI) وبرمجة بايثون. كان هدفنا الأساسي هو تحقيق التعرف على الوجوه بدقة عالية عبر ظروف متنوعة. أسفرت الاختبارات المكثفة على مجموعة بيانات تضم أكثرمن 1500 فرد عن نتائج مبهرة ، حيث وصلت معدلات النجاح إلى ذروتها بنسبة 93.1%. وهذا يؤكد قدرة النظام على الأداء بفعالية في ظل ظروف مختلفة ،بما في ذلك الزواي االمختلفة وبيئات الإضاءة والفئات العمرية وتسريحات الشعروالإكسسوارات وأنماط الماكياج.

كلمات مفاتيح

ذكاءالاصطناعي،معالجة الصور،التعرف على الوجوه،الشبكات العصبية التلافيفية،التعلم العميق, مخطط التدرجات الموجهة.

ABSTRACT

This thesis details the creation of a robust facial recognition system leveraging artificial intelligence (AI) and Python programming. Our primary aim was to achieve highly accurate face recognition across diverse conditions. Extensive testing on a dataset comprising over 1500 individuals yielded impressive results, withsuccess rates reaching a peak of 93.1%. This underscores the system's capability to perform effectively under varying conditions, including different angles, lighting environments, age groups, hairstyles, accessories, and makeup styles.

Key Words

AI, image processing , face recognition, convolutional neural networks, deep learning, Histogram ofOriented Gradients.

Résumé

Cette thèse détaille la création d'un système de reconnaissance faciale robuste tirant parti de l'intelligence artificielle (IA) et de la programmation Python. Notre objectif principal était d’obtenir une reconnaissance faciale très précise dans diverses conditions. Des tests approfondis sur un ensemble de données comprenant plus de 1 500 individus ont donné des résultats impressionnants, avec des taux de réussite atteignant un sommet de 93,1 %. Cela souligne la capacité du système à fonctionner efficacement dans diverses conditions, notamment sous différents angles, environnements d'éclairage, tranches d'âge, coiffures, accessoires et styles de maquillage.

Mots clés

IA, traitement d'images, reconnaissance faciale, réseaux de neurones convolutifs, apprentissageprofond,Histogramme des gradients orientés.

CONTENTS

List of abbreviation

AC: alternating Current.

ACS: Access control system.

ADTV: advanced definition televisions.

AGI: Artificial general intelligence.

AI : Artificial Intelligence.

APIs: Application Programming Interfaces

ASI: achieving super intelligence.

ATM. AWG: American wire gauge.

ATM: the automated teller machine

Bit: Binary digit (basic unit of information).

BLE: Bluetooth Low Energy.

BMP: Bitmap.

BMP: Bitmap.

BYLC: Berkeley Vision and Learning Center

CLI: Command line interface.

CNN: Convolutional neural network.

CNNs: Convolutional Neural Networks.

CNTK: Microsoft cognitive toolkit.

ConvNets: Convolutional neural networks.

Cpu: Central processing unit.

CSI: Camera Serial Interface.

CV: Computer Vision

DC power: Direct Current power.

DSI: Display Serial Interface.

DVD: digital versatile disc.

DVI: digital visual Interface.

DVRs: digital video recorders.

EM: Expectation-Maximization.

ENFSI: the European Network of Forensic Science Institutes.

FISWG: the Facial Identification Scientific Working Group.

FN: False Negative.

FP: False Positive.

FPS: Frames Per Second

FRT: face recognition technique.

FRVT: Face Recognition Vendor Test.

FTP: stands for "File Transfer Protocol.

GB: Gigabyte.

GHz: Gigahertz.

GIF: Graphics Interchange Format.

GMM: Gaussian Mixture Models.

GPIO: General Purpose Input/Output.

GPU: Graphics Processing Unit.

GUI: Graphical user interface .

HD: High definition.

HDMI: High-Definition Multimedia Interface.

HDT: High-definition televisions

HOG: Histogram of Oriented Gradients.

ID: identity document.

IDEs: integrated development environments.

IDLE: Integrated Development and Learning Environment.

IEEE: Institute of Electrical and Electronics Engineers.

iOS: iPhone Operating System.

IOT: Internet of Things.

JPEG: Joint Photographic Experts Group.

LAN: Wireless Local Area Network.

LBP: Local Binary Patterns.

LCD: Liquid crystal display.

LFW: The Labelled Face in the Wild.

LPDDR: Low Power Double Data Rate.

LUT: Look-up Table.

MAP: Maximum *A Posteriori*.

Mbps: stands for "Megabits per second.

MHz: Megahertz.

MIPI: Mobile Industry Processor Interface.

ML: Machine learning.

MMAL: Multi-Media Abstraction Layer.

MMC: Mobile memory card.

MP: Mega pixel. Ms: Milliseconds.

NIST: National Institute of Standards and Technology.

OpenCV: Open source computer vision library.

PC: Personal Computer.

PCA: Principal Component Analysis.

PCIe: peripheral component interconnect express.

PGM: Portable Gray Map.

PhD: Philosophiæ doctor.

PIL: Python Imaging Library.

PNG: Portable Network Graphics. PoE: Power-over-Ethernet.

POIs: Persons of interest. PPM: Portable Pixma

RAM: Random Access Memory.

RBF: Radial basis function.

RTC: stands for "Real-Time Clock.

SATA: Serial advanced technology attachement.

SD: Secure Digital.

SDHC: Secure Digital high capacity.

Sdxc: Secure Digital extended capacity.

SIFT: scale invariant and feature transform.

SoC: System-on-chip.

SOM: Self-Organizing Map.

SPI: Serial Peripheral Interface.

SSH: Secure SHell.

SSH: Secure Shell.

SVM: Support Vector Machines.

TIFF: Tagged Image File Forma.

TN: True Negative.

TP: True Positive.

USB: Universal Serial Bus.

VNC: Virtual Network Computing.

Web: World Wide Web.

Wi-Fi: Wireless Fidelity.

YTF: Youtube Face.

List of figures

List of tables

GENERAL INTRODUCTION

GENERAL INTRODUCTION

Artificial intelligence (AI) plays a crucial role in the 21st century. This rapidly growing field benefits efrom tchnological advancements that pave the way for innovative and promising applications. Among these advancements, convolutional neural networks (CNNs) have revolutionized image processing, particularly in the field of facial recognition. In this paper, we will explore the fundamental concepts of AI and facial recognition, emphasizing their interaction and implications.

Facial recognition is a biometric technology that identifies or verifies individuals by analyzing and comparing patterns based on the person's facial features. It uses unique characteristics such as the distance between the eyes, the shape of the cheekbones, and the contour of the lips, ears, and chin to create a digital map of the face. [1]

Facial recognition begins with image capture, where a photo or video of a face is taken. Next, the system performs face detection, locating and isolating the face within the image. Following this, feature extraction takes place, where key facial features are identified and extracted. This process often involves converting the image into a set of numerical values or a facial template. Once the features are extracted, they are compared against a database of known faces. Finally, in the recognition/verification step, a match is found (identification) or the person's claimed identity is verified. [2]

Artificial Intelligence (AI) involves the simulation of human cognitive processes in machines, enabling them to perform tasks that typically require human intelligence. This includes the ability to reason, learn from experience, and make decisions. AI systems are built using various advanced algorithms and technologies, which allow them to process information and perform complex functions. [3]

AI encompasses various technologies that enable machines to simulate human cognitive functions. Hereare four fundamental components of AI:

Machine Learning (ML) is a core component of AI, consisting of algorithms that allow computers to learn from and adapt to new data without being explicitly programmed. ML models improve their performance over time as they are exposed to more data, making accurate predictions or decisions based on patterns and trends. [4]

Deep Learning (DL) is a specialized subset of machine learning that uses neural networks with many layers, known as deep neural networks. These networks are capable of learning high-level abstractions in data, enabling tasks such as image and speech recognition with a high degree of accuracy. Deep learning models are particularly effective in handling large and complex datasets. [5]

Natural Language Processing (NLP) enables machines to understand, interpret, and respond to human language in a meaningful way. NLP combines computational linguistics with machine learning to process and analyze large amounts of natural language data. This allows AI systems to perform tasks such as language translation, sentiment analysis, and conversational interactions. [6]

Computer Vision is a field within AI that focuses on enabling machines to interpret and understand visual information from the world. By using algorithms to process and analyze images and videos, computer vision systems can perform tasks such as object detection, facial recognition, and scene understanding. This technology is integral to applications like autonomous vehicles, surveillance systems, and medical imaging. [7]

Together, these components form the backbone of AI, driving advancements across various industries and transforming the way we interact with technology.

AI algorithms handle extensive data processing to discern unique facial features from large image datasets. This initial phase involves preprocessing to standardize inputs, followed by advanced feature extraction techniques like Scale-Invariant Feature Transform (SIFT) or Histogram of Oriented Gradients (HOG). [8]

Deep learning models, such as convolutional neural networks (CNNs), excel in pattern recognition within facial features. These models employ layers of convolution, pooling, and fully connected networks to identify intricate facial patterns and expressions with high precision. Non-linear activation functions like ReLU enrich learning capabilities, crucial for capturing diverse facial nuances effectively. [9]

AI-driven facial recognition systems continually advance through supervised and unsupervised learning methods. Supervised learning fine-tunes models on labeled datasets, employing backpropagation and gradient descent to refine accuracy and mitigate errors. Unsupervised learning aids in clustering similar facial attributes, fostering adaptive learning and ongoing model enhancement. [10]

This collaborative approach between AI, machine learning, and deep learning empowers facial recognition systems to evolve dynamically. They adeptly adapt to varying environmental conditions such as lighting changes, diverse facial angles, and expressions, marking significant strides in intelligent technology applications across security, authentication, and personalized services domains.

We have structured our thesis into three chapters. In the first chapter, we begin by introducing the general concept of facial recognition, covering its history, importance, mechanisms, and the challenges faced by facial recognition systems.

In Chapter 2, we delve into the types and definitions of artificial intelligence, exploring various AI techniques employed in facial recognition. Specifically, we focus on Convolutional Neural Networks (CNNs), Deep Learning, and Facial Landmark Detection. Additionally, we discuss the theoretical foundations of AI within the context of facial recognition, comparing tradition.

In Chapter 3, we provided a detailed presentation of our face detection code using the face_recognition library, a Python package designed specifically for facial recognition and manipulation tasks. This library relies on the dlib library, which integrates advanced face recognition methods based on deep learning principles. Additionally, comprehensive experimental results were presented, addressing various challenges that facial recognition systems may encounter.

CHAPTER 1:

FACIAL RECOGNITION

Chapter 1: Facial Recognition

1. Introduction

During daily life, each of us identifies various faces throughout the day. So when we encounter a person, our brain searches our memory and checks whether this person is registered or not, which is an easy task for humans. Is it the same for a machine?

In this first chapter, we will precisely outline the main aspects of our work. We will explain general concepts about facial recognition, how a facial recognition system operates, discuss some techniques used, provide examples, and explore the application domains.

2. General presentation of facial recognition

Facial recognition systems are a sub-field of AI technology that can identify individuals from images and video based on an analysis of their facial features. Today, facial recognition systems are powered by deep learning, a form of AI that operates by passing inputs through multiple stacked layers of simulated neurons in order to process information. These neural networks are trained on thousands or even millions of examples of the types of problems the system is likely to encounter, allowing the model to “learn” how to correctly identify patterns from the data. Facial recognition systems use this method to isolate certain features of a face that has been detected in an image like the distance between certain features, the texture of an individual’s skin, or even the thermal profile of a face and compare the resulting facial profile to other known faces to identify the person.

Broadly, facial recognition systems can be used to accomplish one of two different tasks: verification or identification. Verification (also known as 1:1 matching) is used to confirm that a person is who they say they are. An example of verification is when a person uses their face to unlock their smartphone, sign into a banking app, or verify their identity when boarding a plane. A sample image is taken of a person’s face during login, which is then compared to a known image of the person they claim to be. Facial recognition algorithms tend to have good accuracy on verification tasks, because the subject usually knows they are being scanned and can position themselves to give their cameras a clear view of their face.

Identification (also known as 1:N or 1:many matching) is when software takes an unknown face and compares it to a large database of known faces to determine the unknown person's identity. Identification can be used on "cooperative" subjects who know they are being scanned, or "uncooperative" ones who do not. The latter has drawn the most attention, due to the fear that law enforcement or private businesses will use the technology to remotely gather data about individuals without their knowledge. However, remote identification can also be used to identify suspects from surveillance footage, track down missing persons or the victims of kidnapping, and improve private sector services. Remote identification systems tend to have lower accuracies compared to verification systems, because it is harder for fixed cameras to take consistent, high-quality images of individuals moving freely through public spaces [11].

3. History of face recognition

Facial recognition has a fascinating history, marked by major technological advances and growing concerns about privacy.

Nowadays, facial recognition may seem ubiquitous, but it hasn't always been the case. The history of facial recognition dates back to the 1950s and 1960s, but research on automatic facial recognition is considered to have started in the 1970s when Goldstein et al. published "Identification of human faces," which was a rudimentary attempt at facial identification. This method proposed studying 21 subjective facial characteristics, such as hair color and lip thickness, to identify a face in a photograph, but the main drawback was that these characteristics were not only subjective but also manually calculated.

In 1987, Sirovich and Kirby published their seminal work, "A Low-Dimensional Procedure for the Characterization of Human Faces," followed by Turk and Pentland in 1991 with "Face Recognition Using Eigenfaces." Research studies on facial recognition have multiplied since the early 1990s, driven by hardware advancements and the increasing importance of security-related applications.

Progress in image-based facial recognition techniques can be broadly divided into four main phases of conceptual development, as outlined by Wang and Deng, which reflect the historical development of major methods (OUBAHA, 2021):

1. Holistic or appearance-based approaches utilize the entire facial region and employ linear or non-linear methods to represent the face in a lower-dimensional subspace. One of the earliest successful methods is Eigenfaces, developed by Turk and Pentland. Other approaches include those using linear subspaces, manifold learning, and sparse representations.

2. Local feature-based facial recognition methods became popular after the 2000s and use manually crafted features to describe the face, such as Gabor features, local binary patterns (LBP), and their variants.

3. Methods utilizing learning-based local descriptors emerged after the 2010s, learning discriminative image filters using shallow techniques.

4. Deep learning-based methods gained popularity after the significant success of AlexNet in the ImageNet competition in 2012 and brought a new perspective to the facial recognition problem. Unprecedented stability has been achieved for facial recognition, with performances similar to those of humans on large-scale datasets collected in unconstrained contexts.

Deep learning is now responsible for unprecedented accuracy in facial recognition. Specialized architectures have been trained and have achieved accuracies very close to or even surpassing human performance. This accuracy has reached 97.35% and exceeded 99.80% in just three years.

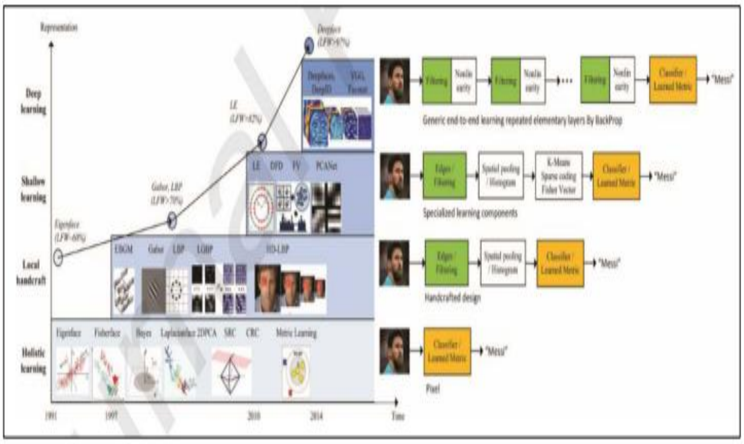


Figure 1. 18 : History of face recognition

4. Importance of facial recognition in various fields

Due to its diverse applications and advantages, facial recognition is crucial in various fields. [12],[13].

1. Security and Surveillance: used for personal identification and tracking to improve security

public places, businesses and sensitive areas, thus helping to prevent crime.

2. Identity management: effectively implement physical access control and digital security, and provide reliable solutions

and ways to quickly verify a person's identity.

3. Mobile technology: integrated into mobile devices to unlock phones and optimize access

Enhance personal data security.

4. Commerce and Marketing: Used to personalize the customer experience and enable retailers to analyze Understand consumer preferences and improve marketing strategies through facial reactions.

5. Healthcare: Used for patient identification to facilitate access to medical records and improve

health management, thereby contributing to more efficient and safer healthcare.

6. Border Control: Used to enhance national security by identifying and identifying travelers

Suspicious activity at the border.

7. Education: Used in the education industry to monitor examination rooms, control campus access, and improve student and staff safety.

5. Contextualizing facial recognition in embedded systems

Contextualizing facial recognition in embedded systems requires the integration of this technology into autonomous computers. This technical application is crucial in many fields due to its many advantages. [2]

1. Embedded systems for security: In the context of security, facial recognition is integrated into embedded systems to identify and track individuals. These systems improve security in environments such as government offices, critical facilities and sensitive areas.

2. Physical Access Control: Regarding personal recognition, facial recognition has been implemented in built-in systems for physical access control, providing an efficient and contactless method to provide access to restricted areas.

3. Integration with embedded mobile devices: Embedded mobile devices such as smartphones and tablets include facial recognition for features such as secure device unlocking and personal data protection.

4. Resource Optimization: Contextualization of face recognition in embedded systems requires resource optimization, especially in terms of computing power, storage and energy consumption, to ensure efficient performance of embedded devices with limited resources.

5. Internet of Things (IoT) Applications: Facial recognition can be integrated into IoT-connected embedded systems, enabling applications such as smart tracking, secure access control, and personalizing human interactions with IoT devices.

6. Real-time processing: Contextualizing facial recognition in embedded systems often requires real-time processing capabilities to ensure fast and accurate results, especially in applications such as security and surveillance.

7. Integration with smart embedded devices: Facial recognition can be used in smart embedded devices like smart security cameras to enable advanced features like emotion recognition and context recognition

6. Disadvantages and limitations of facial recognition system

Aforementioned advantages and application, facial recognition system has drawbacks and limitations revolving around concerns over its effectiveness and controversial applications. Take note of the following disadvantages: [3]

1. Issues About Reliability and Efficiency: A notable disadvantage of facial recognition system is that it is less reliable and efficient than other biometric systems such as fingerprints. Factors such as illumination, expression, image or video quality, and software and hardware capabilities, can affect its reliability or accuracy and overall system performance.
2. Further Reports About Its Reliability: Several reports have pointed out the ineffectiveness of some systems. An advocacy organization noted that the systems used by law enforcement in the United Kingdom had an accuracy rate of only 2 percent. Implementations in London and Florida did not result in better law enforcement according to another report.
3. Concerns About Possible Racial Bias: A study by the American Civil Liberties Union revealed that the Rekognition technology from Amazon failed nearly 40 percent false matches in tests that involved people of color. The system has been criticized for

perpetuating racial bias due to false matches. This is another disadvantage of facial recognition technology.

4. Varied Performance Across Different Systems: The capabilities of a particular system are also dependent on the underlying technologies. Some Android smartphones have an inferior or a less reliable system because they depend on the front-facing cameras. iPhones and iPads have a better implementation because they use infrared for three-dimensional mapping.
5. Possible Concerns with Privacy Laws: Possible conflict with privacy rights is also another key disadvantage of facial recognition. Illinois has a Biometric Information Privacy Act that requires affirmative consent for companies to collect biometric data. The fact that the system is used for mass identification has also been criticized for mass surveillance and profiling.

7. Facial recognition

7.1. Definition of facial recognition

Figure 1.19: Definition of Facial Recognition

Facial recognition is a technology that involves identifying or verifying a person based on their facial features. It uses biometric data such as a person's unique facial patterns to identify them. The process involves capturing and analyzing facial features, often using facial landmarks to create a digital representation that can be compared to stored data for identification or verification.

In the context of facial recognition, two modes of operation are generally distinguished: real-time mode (online) and post-event mode (offline). In post-event mode, the recognition system collects information about each detected face, storing this data in an easily accessible database. In case of need (online mode), a user can access this data and select a specific face for authentication or identification.

Regardless of the mode of operation, the facial recognition system performs several essential operations: acquiring the facial image, preprocessing the data, detecting facial features, extracting these features (features), classification, and decision-making (authentication or recognition) (AKCHA, 2020).

7.2. The mechanism of facial recognition

The mechanism of facial recognition is listed below:

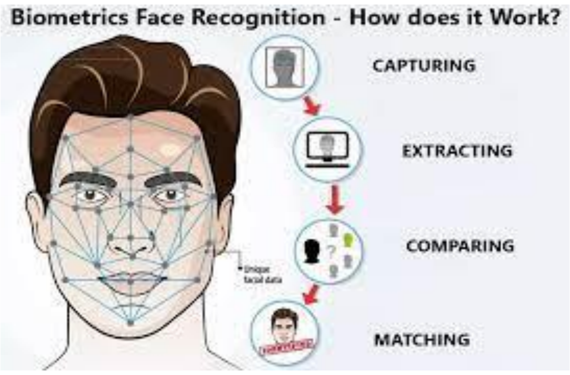


Figure 1. 20 : Biometrics Face Recognition – How does it Work ?

1. Image Acquisition : Image acquisition is the first step in a recognition system, accomplished by capturing the image of the individual from the external world using devices such as a camera or a digital camera.

2. Preprocessing : The preprocessing phase follows the acquisition phase. It involves preparing the face image in such a way that it becomes usable. The objective is to eliminate

noise in the input image caused by the quality of the devices used during acquisition, retaining only the relevant information and thus preparing the image for the next stage.

3. Face Detection : This step generally involves determining the possible presence of a face in the image and locating it in such a way as to obtain a region of interest (ROI) face where feature vector extraction can be performed.

4. Feature Extraction : This involves extracting the intrinsic characteristics of the face, often referred to as landmarks in English. These pieces of information are unique to each individual and can accurately represent them. These features must be relevant and unique for each individual, forming the core of the facial recognition system.

5. Classification : It involves determining one of the models based on the characteristics of a set of extracted faces in the previous step to assess the similarities they share. Each model is a unique representation of distinct features. Several individuals may possess common characteristics, and individuals with similar traits are grouped into the same class.

6. Decision : During this step, the system evaluates the identity inferred from an individual's characteristics by comparing it with the identities of stored models, assessing the degree of similarity between the characteristics of the queried individual and those of the stored models) (AKCHA , 2020).

8. Future applications of facial recognition

Facial recognition technology is rapidly evolving and has a wide range of potential future application.hereare some areas where it could be increasingly used:

1.Lie Detection And Age Verification : Several countries have come up with lie-detection programs that enable face recognition to discern what’s true from an individual’s expression. The technology allows investigators to determine if a suspect is speaking the truth or not. An initiative to develop a virtual lie detector test that individuals who are traveling to the EU might take from their homes utilizing a live webcam was financed by the European Union.

The technology analyses an individual's facial expressions to identify the extent to which they are being truthful. Identification of age will be helpful in online programs, much as a lie detection. When enrolling for exams, institutions, or government programs, it correctly verifies an individual's age using online applications.

2. Payment And Crypto Currency : Using facial recognition to make payments is now no longer an unrealistic goal. The whole payment procedure is completely contactless and takes only a few minutes. You also don't need to carry a mobile device or a debit or credit card, or input a PIN; simply scan your face. In contrast to passwords, which can be easily generated and broken, your face is the sole password for accessing financial accounts or conducting transactions.

Cryptocurrencies are a game-changing technology that has exploded in popularity in the past few years. Many businesses, however, have used facial recognition to certify the real identities of stakeholders as well as investors in order to reduce the risks they face.

3. Personal Assistant Becomes More Intelligent With New Insights : Artificial intelligence-powered virtual assistants such as Siri, Alexa, and Cortana have improved their ability to execute simple activities such as texting, arranging scheduled meetings, and streaming music. Nevertheless, when combined with facial recognition technology, these personal assistants grow more receptive and personalized, making them a feasible and viable replacement to current robotic incarnations. These technologies might be trained to detect feelings along with other reactionary social indications, allowing them to communicate more effectively.

4. AI Enhanced Facial Capabilities : Another development that we're keeping an eye on is the rising usage of artificial intelligence (AI) in face recognition systems. By enabling the application to gain insight and adjust to diverse scenarios, AI might be able to make face recognition substantially more precise and effective.

Systems for access control are going to get faster, more flexible, and smarter in the near future when AI is used in face verification. Face identification technology, for instance, may include video authorization, multi-factor authorization, and additional features to improve security and privacy.

5. Facial Recognition Prevents Hacking : Typically, the majority of vital transactions, including website logins, online payments, and other protected areas, are facilitated by facial recognition technology. Face recognition is an excellent option in the digital age where PII (personally identifiable information) is easily accessible on the dark web. It safeguards people’s privacy while they move about in both real and virtual settings.

Face authentication is a potential way to strengthen people’s safety and confidentiality because passwords and other credentials are frequently stolen or exploited and can’t be utilized to determine the real identity of a person. Although these advances must, of course, be backed by safety regulations, the general benefits related to user protection make this laborious operation worthwhile to study and enhance.

6. Driver Monitoring : Another prominent trend is the adoption of facial recognition systems in automobiles. Tiredness is one of the most common causes of vehicle accidents worldwide. Facial recognition is now starting to change that. Facial recognition techniques are now utilized to monitor the focus of drivers across long distances. If the driver looks to be dozing off, the technology can instantly slow the car down and provide the driver with a sound alarm.

Unlocking the car with the face is a novel and dependable method of reducing theft. Other family members might also be restricted or granted permission by car owners. In this manner, individuals can keep their little kids from driving, and if an unauthorized individual enters the vehicle, the security system will stop it from starting. Face recognition can additionally be used to customize car specifications for an individual in particular. It can, for instance, stream their favorite podcast or modify their seat position and temperature.

7. Facial Recognition With Masks : Face identification has become more difficult as a result of the rise in the use of facial masks brought on by the COVID-19 epidemic, however, facial recognition technology has replied by becoming more innovative. According to certain research, masked-face recognition can be 99% accurate. For instance, NEC Corp. of Japan has

introduced a facial recognition system that it claims has an accuracy percentage of over 99.9% and needs just one second to identify faces.

To verify an individual's identity, the device first assesses whether they are wearing a mask before concentrating on the exposed areas of their face. Several setups in the US and Europe presently make use of this recognition software. It is used in restaurants and hotels to check if their employees are wearing masks. The system is currently being implemented at airports as well.

8. Bottom Line : Facial recognition systems are probably going to start showing up in our daily lives as they continue to advance and develop. Because of its contactless aspect and ease of use, face recognition technology is now chosen above other forms of biometric identification like fingerprint scanning, speech recognition, structure recognition, and skin texture recognition. At present, surveillance and attendance are the main applications of face recognition. Although it's hard to predict if this technology will be adopted by a significant portion of the population, the benefits that these solutions may bring about are extremely substantial. How it will strengthen individuals and businesses will only get clear with time)(Krithika, 2023).

9. Challenges of face detection

With the diversity of image types and sources, the human skin color can vary significantly, making accurate skin detection challenging. The challenges associated with skin detection can be attributed to the following factors:

1. Lighting Variations: The variation in lighting during face capture makes the recognition task more challenging. A face of the same individual with two different lighting levels may be recognized as two different individuals.

Figure 1. 21 : Lighting Variations

2. Ethnic Diversity: The appearance of skin color varies from person to person due to physical differences among human racial groups.

Figure 1. 22 : Human racial groups

3. Imaging Conditions: When an image is captured, factors such as camera characteristics (sensor response, lenses) affect the appearance of the skin. In general, different color cameras do not necessarily produce the same color appearances for the same scene under the same imaging conditions.

Figure 1. 23 : Lighting effect on the image.

4. Image Editing and Reproduction: Some images have already been captured using color filters, making the processing of color information even more challenging.

Figure 1. 24 : The difference between filters and reality.

5. Makeup: Affects the appearance of skin color.

Figure 1. 25 : The difference between Makeup and reality.

6. Aging: Human skin varies from fresh and elastic to dry and rough with wrinkles.

Figure 1. 26 : Face recognition struggles to recognise us after five years of ageing

7. Complex Background: Another challenge arises from the fact that many objects in the real world have colors similar to that of the skin. The diversity of backgrounds is practically limitless, leading to false detections by the skin detector.

8.Pose variation : Pose variation is considered a major challenge for facial recognition systems. When the face is in profile within the image plane (orientation < 30°), it can be normalized by detecting at least two facial features (passing through the eyes). However, when the rotation exceeds 30°, geometric normalization is no longer possible.

Figure 1. 27 : Change of angle (poses).

9. Presence or absence of structural components : Specific aspects such as beard, mustache, glasses, hair style, and color cause significant changes in the structural components of the face, including shape, color, size, etc.

Figure 1. 28 : Changing specific aspects

10. Partial occlusions : The face may be partially obscured by objects in the scene or by the wearing of accessories such as glasses, scarves, etc. In the context of biometrics, proposed systems must be non-intrusive, meaning they should not rely on active cooperation from the subject.

11. Facial Expressions Alteration: The face is a non-typical element. The emotion conveyed by the face, in addition to the transformation resulting from speech, results in a perceptible significant change, and the number of possible structures cannot realistically be expanded(Bouzit ,2019) ,(AKCHA ,2020) , (Boukerrouche , 2018).

Figure 1. 29 : Facial Expression Change

10.Methods of facial recognition evaluation

Facial recognition systems are typically evaluated using a variety ofmethods. Here are some commonlyused methods:

10.1. Introduction to evaluation metrics

Human face is one of the most informative means of communication in our societal life. Unlike, face recognition by humans to comprehend their peers possess a natural phenomenon, but recognizing facial geometry through machine is still a challenging problem. Face recognition is the task of recognizing an individual using digital facial image. The progress of face recognition technology over the past two decades has been substantial, as benchmarked by the National Institute of Standards and Technology (NIST)1 . The release of the Multiple Biometric Evaluation 2010 report of NIST shows that the most accurate face recognition technique (FRT), the chance of identifying the unknown subject is about 92% when searching a gallery of 1.6 million faces. For other population sizes, this accuracy rate decreases linearly with the logarithm of the population size.

Face recognition techniques (FRT) are more accurate on the images that are collected with careful cooperation of the subjects, active compliance by the photographer to the image collection under controlled environment, and a proper review by an official. The compressed facial images such as passport, visa and ID card are subject to losses, while less compressed images such as mugshot are generally of higher resolution, exhibit considerable pose, illumination and expression variations. Identifying person in an uncontrolled environment is still a challenge for facial recognition reliability. The FRT performance deteriorates significantly when variations are found in illumination, facial pose and expression3,4. Other factors such as image resolution, orientation and blurring, time delay or facial aging, and occlusion such as partial covering of face by clothing, shadows and obstructions also contribute to face recognition errors(Agrawala, 2014).

10.2. Images and contrast methods

Facial recognition is employed in three primary tasks, contingent upon the available hardware and the intended objective:

10.2.1. Images

For each of these tasks, the two main categories of images compared against each other are traces and references. Traces, referring to physical and digital remnants of contentious activities (Ribaux, 2014), are images of unknown origin recorded at the time of the events and subsequently collected. They allow for the visualization of one or more individuals and points of interest. For instance, in an investigation related to the fraudulent use of a credit card following a pickpocketing incident, surveillance camera recordings at the time of the theft as well as those from the automated teller machine (ATM) are provided to the police upon request. Such images offer substantial information to investigators. It is thus possible to trace the movements of persons of interest (POIs) before or after the events, to distinguish their morphology, clothing, and faces, and then, if the quality permits, to compare the face of the POIs to reference images, such as a suspect's photograph (verification) or a database to search for potential suspects (identification).

References, on the other hand, consist of images of known origin and good quality. These are typically identification documents, driver's licenses, or mugshots. When an investigation identifies one or more persons of interest (POIs) through trace images and legal proceedings require the work of an expert to present conclusions to the court, it is necessary to acquire reference images to compare them to the traces.

Similar to fields like digital forensics, the comparison between trace images and reference images can be done either manually or using automated systems. In the following sections, we briefly outline these approaches, emphasizing their practical advantages and limitations.

10.2.2. Manual methods

Manual comparison is used for tasks requiring the analysis of few images simultaneously, such as identity checks (comparing a person with an identification document) and criminal investigations (comparing a surveillance video image with the mugshot of a suspect). Specialized literature describes fouranalysis methods for manual face comparison by an expert: holistic, morphological, photo-anthropometric, and overlaying of the two compared images (Ali, Veldhuis, & Spreeuwers, 2012).

Holistic analysis involves a global description of facial features, without any measurements. This method is commonly used by officers during customs checks, for example. It allows for quick comparison but is highly dependent on image quality and the difficulty of the comparison (strong resemblance between different individuals). In a judicial context, it is recommended to use this method only initially, combining it with the following approaches due to its weak and variable performance. Morphological analysis involves describing the shapes and proportions of facial elements (forehead, eyes, nose, mouth, eyebrows, ears, cheeks, chin) as precisely as possible. This also includes describing facial marks (wrinkles, scars, freckles, moles, tattoos, etc.), as well as hair color and length, hair implantation, and hair density. In forensic examinations, this practice is considered the more efficient, applicable to images of varying quality, and within the reach of non-specialized individuals. However, it suffers from a lack of transparency and standardization, as well as excessive reliance on the expert's innate ability to recognize faces and the image quality (Noyes, Phillips, & O'Toole, 2017; Peng, 2019; Phillips et al., 2018). The innate "talent" of a person in recognizing faces is increasingly prevalent in studies such as Towler et al.'s (2019), where the authors highlight the ineffectiveness of facial recognition training programs dedicated to experts. The third approach, photo- anthropometric analysis, relies on measuring distances between certain points on the face such as the base of the chin, the center of the mouth and eyes, the corners of the lips, the edges of the nostrils, etc. (Moreton & Morley, 2011). Lastly, the technique of image overlay aims to visualize more directly the morphological and anthropometric similarities and discrepancies of faces in the two images.

Regarding the applicability of these two methods, the Facial Identification Scientific Working Group (FISWG) and the European Network of Forensic Science Institutes (ENFSI) recommend not using anthropometric and image overlay approaches alone for comparison purposes, and instead advocate for combining multiple approaches (ENFSI, 2018; FISWG, 2012). This is due to:

– The need to use images taken under optimal conditions, which is rarely the case;

- The highly time-consuming nature of these processes;
- The advanced training required by the operator, assuming it is effective;
- THe lack of reliability of the results. In a judicial context, experts base their comparisons on Approaches validated and recommended by competent organizations. These organizations regularly conduct performance tests to evaluate the results of individual protocols and adjust their methodological recommendations accordingly.

10.2.3. Automated systems

The increasing number of image comparisons required in certain domains has naturally led to the automation of the process. Thus, an investigator can compare a trace image to all reference images in a database of known offenders.

To better understand the issues related to these systems, we first offer a brief overview of their operation.

Automatic systems are developed based on two main algorithmic models: rule-based engines and machine learning (Buchanan, 2005). Rule-based algorithms extract targeted information from an image, as defined by the developer, and compare it to evaluate the degree of similarity between the information in image A and that in image B. The entire process is transparent to the algorithm developer and can also be made transparent to the user if the entire code is made available to the public (open-source algorithms). Until recently, all algorithms were based on this functionality. However, in recent years, there has been an emergence and rapid development of algorithms based on machine learning processes, particularly deep learning (LeCun, Bengio, & Hinton, 2015). Learning-based systems are trained to detect and recognize faces from large databases provided by the developer (or by the operator if the algorithm is open source). During this phase, the system collects a wealth of information and generates its own rules to detect, analyze, and compare two faces to assess their similarity. Therefore, the performance of these systems is closely linked to the data used for training. For example, a system trained only on high-quality frontal ID photographs will perform poorly on low-quality surveillance camera images with varying angles (Peng, 2019).

When used by public or private institutions or in criminal investigations, these systems represent valuable time-saving aids, and any potential errors can be controlled post hoc by an operator or investigator. Nevertheless, in the criminal context, such errors are prohibited due to their potentially significant impact (wrongful incarceration or release of a suspect). This is why automatic facial recognition systems are still underutilized, and numerous research efforts are currently focused on developing methods to utilize their results as evidence in court (Ali et al., 2012; Jacquet&Champod, 2020; Rodriguez, Geradts, &Worring, 2018).

The performance of the most commonly used or recently developed facial recognition systems is evaluated by the National Institute of Science and Technology (NIST) in its Face Recognition Vendor Test (FRVT) reports (Maëlig,2021).

10.3. The face detection approaches

Several approaches have been developed for face detection, which are divided into four categories, some use shape, or are based on color (skin color), and others are based on facial appearance, or the last one is a combination of all the previous approaches. The four subdivisions of face detection approaches are:

- Knowledge-based approach.

- Template-matching.

- Appearance-based approach.

- Approach based on invariant features.

10.3.1. Appearance-based approach

The methods of this approach generally rely on machine learning techniques, learning models that are later used for face detection using a set of images representing facial space variation.

The problem of face detection for this approach is considered as a classification problem between two classes: Face Class and Non-face Class.

Appearance-based approach methods rely on statistical analysis techniques and machine learning to find appropriate features of face images and non-face images.

Several techniques have been used for this approach such as Eigenface, Neural Networks, and Support Vector Machine (SVM).

1.Eigenface :Turk and Pentland were the first to develop the method in 1991, Eigenface, which would later become one of the most well-known methods for face detection. The principle of this method is to project an image into a space, then calculate the Euclidean distance between the original image and its projection. Encoding an image in a space serves to degrade the information contained in the image. After evaluating the distance, which is compared to a pre-set threshold, if the information loss is greater, it implies that the image is not well represented in the space and does not contain a facial area: a non-face class.

Figure 1. 30 : Linear Combination of EigenFaces

The advantage of this method is that it yields very promising results, but the computation is time-consuming.

Neural networks : The principle of face detection through classification based on neural networks involves using two sets of images, one for facial images and the other for non-facial images, to train the neural network. An interchangeable window scans the entire input image. This window introduced to thenetwork will be classified into two classes: face class and non-face class.

The face detection technique based on neural networks, as described by Rowley et al, is divided into twosteps:

Face localization using a neural network.

Verification of the obtained results

Support Vector Machine (SVM): One of the earliest statistical methods based on information theory for face detection, SVM is considered a novel classifier model for polynomial function learning, neural networks, or radial basis function (RBF). Most of the aforementioned learning classifiers are based on minimizing the "empirical error," SVM operates with another principle called "structural risk minimization," which aims to minimize potential overfitting on generalized errors. During training, for each pair of pixels in the training set, a histogram is used to create probability functions for face and non-face classes because pixel values depend on the values of their neighbors. For training, Colmenarez and Huang used a large set of 11*11 pixel images of both faces and non-faces. The learning results form a Look-up Table (LUT) with probability ratios, aimed at enhancing performance.

10.3.2. Knowledge-based approach

This approach relies on the various components of the face and the relationships between them. Thus, the relative positions of different key elements such as the mouth, nose, and eyes are measured to then serve for classification into face or non-face classes. The challenge with this type of method is that it's difficult to uniquely define a face. If the definition is too detailed, some faces will be missed, whereas if it's too general, the false positive rate will increase.

The primary goal of this approach is face localization. In 1997, Kotropoulous and Pitas utilized a technique based on well-defined rules to locate facial features using a projection method, which was developed by Kanade in 1973 for detecting facial contours.

Yang and Huang developed a hierarchical method for face detection based on the observation that "When the resolution of a facial image is low, the face area becomes uniform." The process begins with a low-resolution image and a set of rules, from which a set of candidate faces is deduced after applying the entire set of rules.

The low-resolution image, along with the face candidates, allows for verifying the presence of facial features through the calculation of local minima. Unfortunately, the number of false detections using this method is significant.

10.3.3. Template matching

Template matching enables the detection of a face or part of a face by learning from a standard example of the face. The idea behind this approach is to calculate the correlation between the candidate faces (each part of the input image) and the template. Templates can be defined manually or parametrized using functions. Although all faces have the same structure, variations in face distance, position, and size pose robustness issues for this approach, particularly concerning variations in light and scale.

The face detection procedure for this method involves two steps:

- First step: Detection of candidate face regions.
- Second step: Examination of details to determine the necessary facial features.

According to Sinha, the use of a set of invariants to describe a face model for determining invariants to changes in brightness allows for characterizing different components of the face such as the eyes, forehead, etc. The algorithm used calculates the luminance ratio between different regions of the face and respects the direction of these ratios.

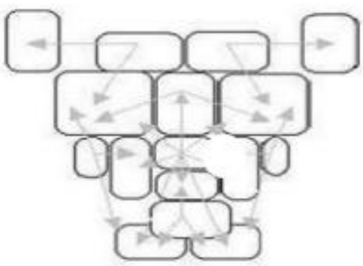


Figure 1. 31 : Face Model Composed of 16 Regions and 23 Directions

This predefined model is decomposed into 16 regions and 23 relations, which are divided into two groups. One group consists of 11 relations representing essential relationships, while the other group consists of 12 relations representing confirmation relationships. Each arrow in the model represents a relationship between two regions. A relationship between two regions is verified if and only if the degree of correspondence exceeds a defined threshold. The face cannot be determined or located unless the number of essential and confirmation relationships also exceeds a defined threshold.

In contrast to Sinha's technique,Yuille et al. used a deformable template to model facial features. They created an elastic template model adaptive to facial features such as eyes, mouth, etc. The parameterized template of this technique describes facial features. From these two experiments, we can deduce two face detection techniques belonging to the template matching-based detection approach, which are:

- Predefinedtemplate
- Deformable or elastictemplate

10.3.4. Approaches based on invariant features

The main objective of this approach is face localization. The algorithms of this approach aim to find and search for structural features (of the face) under various conditions such as lighting changes, easy position changes, expression changes, etc. Afterward, they study invariant features to locate the human face. The algorithm developed by D. Silva et al. is a typical example of feature-based methods.

The algorithms of this approach can be divided into two families:

- Based on facial features
- Based on skin color analysis

1. Based on facial features: Algorithms in this family utilize an initial hypothesis about the position of the top of the face. Subsequently, the search algorithm traverses the face from top to bottom to find the axis of the eyes characterized by a sudden increase in contour density (measured by the black/white ratio along horizontal planes). The distance between the top of the face and the eye plane is then used as a reference length to construct a template.

This template covering features such as the eyes and mouth is initialized from the input image. The initial shape of the template is obtained using anthropometric length while respecting the reference length.

The flexible template is then adjusted relative to the final feature positions using a fine-tuning algorithm employing a contour-based cost function. Although these algorithms succeed in detecting features of different ethnicities since they do not rely on grayscale and color information, they fail to detect these features correctly if the facial image contains glasses or if hair covers the forehead. [16]

The most well-known techniques for this family are: (By Texture, Facial Features, Multi-Features)

- Texture : Human texture is distinctive and can be used to separate faces from other objects. Augusteijin and Skufca developed a face detection method on an image based solely on texture. Texture calculation is done using second-order features on 16 × 16 pixel sub-images. In this method, three types of features are considered: skin, hair, and the rest of the facial components.
- Facial Features : This technique uses edge plans and heuristics to remove all edge groups except those representing facial contours. An ellipse is deduced as a boundary between the background and the face. This is described as being formed from discontinuity points in the luminance (intensity) function of the image. The basic principle is to recognize objects in an image based on known contour models. To accomplish this task, two methods will be presented: the Hough transform (which extracts and locates groups of points with certain characteristics, equation of a well-determined shape) and the Hausdorff distance (which aims to measure the distance between two separate sets of points).
- Multi-features : Various methods combine multiple facial features to locate or detect faces. Most utilize global properties such as skin color, face size, and shape to identify candidates. They then verify these candidates using local features such as eyebrows, nose, and lips.

2. Methods Based on Skin Color Analysis : Detection methods based on skin color analysis are efficient and fast. They reduce the search space for the facial region in the image. Additionally, skin color is robust information against rotations, scale changes, and partial occlusions. Several color spaces can be used to detect human skin; in the image, pixels with skin color are identified. The effectiveness of detection primarily depends on the chosen color space.

Figure 1. 32 : Result of skin color detection

10.4. Comparison between different approaches to face detection

Face detection has evolved significantly over the years , with various approaches being developed ,here's acomparison of some of the different approaches :(Bouzit Dhikra, 2019).

Table 1. 2: Represents the advantages and disadvantages of each method

Approach	Advantages	disadvantages
Approaches based on appearance	- Eigenface yields good results. - Neural networks are robust to noise.	- Significant computation time required. - Difficult to construct. - Challenging learning phase to conduct.
Knowledge-based approaches"	-Reduce the necessary computation time by using subsampled images.	- Causes numerous false detections. - Difficult to translate human knowledge into well-defined rules.
Template-matching	- Simple in terms of detection process. - Yields fairly encouraging results.	- Update with each change in orientation.
Approches basées sur des caractéristiques invariantes	- Resistant to minor changes in lighting and facial position	-The use of skin color-based method requires additional processes to

	- Skin color reduces the search area.	complete face detection.

10.5. True positive, false positive, false negative, true negative

In facial recognition, as in any binary classification problem, the terms True Positive (TP), False Positive (FP), False Negative (FN), and True Negative (TN) are used to evaluate the performance of the system. Here's what each term means:

1. True Positive (TP): This occurs when the facial recognition system correctly identifies a face as belonging to a specific individual when it does indeed belong to that individual. In other words, it correctly detects the presence of a face.

2. False Positive (FP): This happens when the facial recognition system incorrectly identifies a face as belonging to a specific individual when it does not actually belong to that individual. In other words, it incorrectly detects a face.

3. False Negative (FN): This occurs when the facial recognition system fails to identify a face as belonging to a specific individual when it does indeed belong to that individual. In other words, it fails to detect a face that is present.

4. True Negative (TN): This happens when the facial recognition system correctly does not identify a face as belonging to a specific individual when it does not actually belong to that individual. In other words, it correctly identifies the absence of a face.

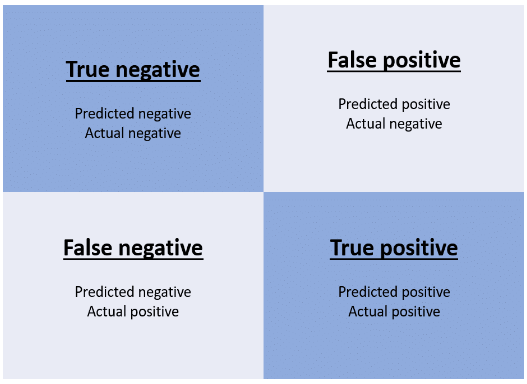


Figure 1. 33: True Positive, False Positive, False Negative, True Negative

These terms are often used in the context of confusion matrices or performance evaluation metrics such as accuracy, precision, recall, and F1 score.

precision measures the ratio of true positives to the sum of true positives and false positives, recall measures the ratio of true positives to the sum of true positives and false negatives, and accuracy measures the ratio of correct predictions (true positives and true negatives) to the total number of predictions.

- U.S. gov't accountability office, face recognition technology: fbi should better ensure privacy and accuracy, gao-16-267 5 (2016) [hereinafter 2016 GAO REPORT]. For further discussion on face matching and specific matching (a true match, a true mismatch, a false positive, and a false negative), see BUOLAMWINI ET AL., supra note 24, at 12–14

10.6. Importance of accuracy and error rates in evaluation

Accuracy and error rates are of paramount importance in the evaluation of facial recognition systems. Under ideal conditions, these systems can achieve near-perfect accuracy, as evidenced by precision scores reaching up to 99.97% in standard tests such as the Face Recognition Vendor Test (FRVT) conducted by the National Institute of Standards and Technology (NIST).

However, this optimal precision is only achievable in perfect conditions where variables such as lighting and positioning remain constant, and where subjects facial features are well-defined and unobscured. In practice, precision rates are often much lower. For instance, the FRVT observed a significant increase in the error rate of a leading algorithm, rising from 0.1% when matching with high-quality mugshots to 9.3% when matching with photos taken "in the wild," where various external factors can affect the quality of the capture.

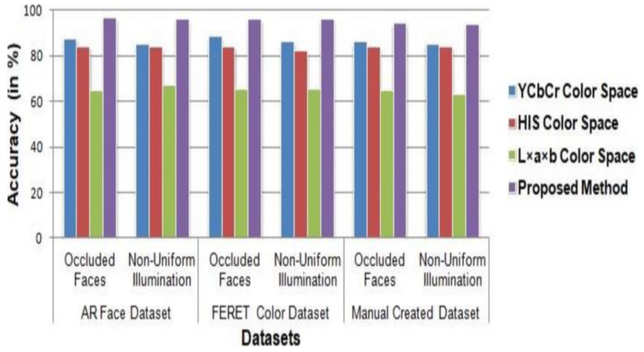


Figure 1. 34: Comparison of various face detection algorithms (with Accuracy rate)

It is important to consider the impact on precision when algorithms are adjusted to minimize false positives. Confidence thresholds are commonly used for this purpose, but they may result in reduced match rates by disregarding correct but less reliable matches. In contexts where human review is anticipated, these thresholds can be adjusted to maximize the system's sensitivity. However, strict confidence thresholds remain necessary in situations where errors could have serious consequences, such as in cases of misidentification by law enforcement.

Furthermore, sensitivity to external factors becomes evident when considering the performance of facial recognition algorithms on faces recorded in surveillance video sequences. The results reveal significant variation among vendors, underscoring the need for stricter regulations and adequate oversight to ensure proper configuration and use of systems.

In conclusion, while facial recognition systems can achieve high levels of precision under ideal conditions, challenges persist in real-world environments, particularly regarding error rates and sensitivity to external factors. A balanced approach, considering both accuracy and associated risks, is essential for the responsible deployment of this technology [1].

11.Conclusion

In this chapter, we outlined the main aspects of our work, such as basic concepts and techniques used (state of the art). We highlighted critical points to address and identified areas to be covered in the subsequent stages of our project implementation. In the next chapter, we will delve into artificial intelligence to specify the work steps and elucidate their main functionalities.

CHAPTER 2 :

Artificial Intelligence (AI)

Chapter2:Artificial Intelligence (AI)

1. Introduction

Artificial Intelligence is a method of making a computer, a computer-controlled robot, or a software think intelligently like the human mind. AI is accomplished by studying the patterns of the human brain and by analyzing the cognitive process. The outcome of these studies develops intelligent software and systems. Inthis first chapter, we will about fartificial intelligence, how it werk, there types, techniques ,and it use in facial recognition..

2. Definition of artificial intelligence

Artificial intelligence (AI) is a branch of computer science. It involves developing computer programs to complete tasks which would otherwise require human intelligence. AI algorithms can tackle learning, perception, problem-solving, language-understanding and/or logical reasoning. AI is used in many ways within the modern world, from personal assistants to self-driving car. Artificial intelligence (AI) is evolving rapidly. While science fiction every so often portraits AI as robots closely as possible to humans. [15]

3. Traits of artificial intelligence

- Artificial intelligence is a broad field with many caractéristiques, here are some key traits :(Ziyad , 2019)
- AI Capable of predicting and adapting: AI uses algorithms that discover patterns from huge amounts ofinformation.
- Makes decisions on its own: AI is capable to augment human intelligence, deliver insights and improveproductivity.
- Continuous learning: AI uses algorithms to construct analytical models. From those algorithms, AI technology will find out how to perform tasks through innumerable rounds of trial and error.
- AI is forward-looking: AI is a tool that allows people to reconsider how we analyze data and integrateinformation, and then use these insights to make better decisions.
- AI is capable of motion and perception.

4. Types of artificial intelligence

4.1. Based on capabilities :

artificial intelligence (AI) can be broadly classified into three types :[14].

Artificial narrow AI :Artificialnarrow Intelligence, also known as weak AI, what we refer to as narrow AI is the only type of AI that exists today. Any other form of AI is theoretical. It can be trained to perform a single or narrow task, often far faster and better than a human mind can. However, it can't perform outside of its defined task. Instead, it targets a single subset of cognitive abilities and advances in that spectrum. Siri, Amazon's Alexa and IBM Watson are examples of Narrow AI. Even OpenAI's ChatGPT is considered a form of Narrow AI because it's limited to the single task of text-based chat.

General AI :Artificial General Intelligence (AGI), also known as Strong AI, is today nothing more than a theoretical concept. AGI can use previous learnings and skills to accomplish new tasks in a different context without the need for human beings to train the underlying models. This ability allows AGI to learn and perform any intellectual task that a human being can.

Super AI :Super AI is commonly referred to as artificial superintelligence and, like AGI, is strictly theoretical. If ever realized, Super AI would think, reason, learn, make judgements and possess cognitive abilities that surpass those of human beings. The applications possessing Super AI capabilities will have evolved beyond the point of understanding human sentiments and experiences to feel emotions, have needs and possess beliefs and desires of their own.

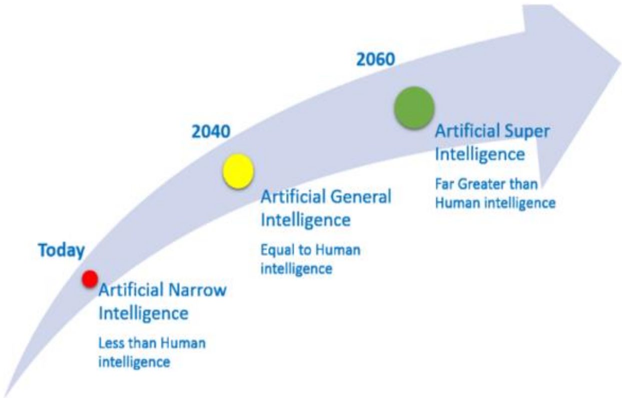


Figure 2. 2: Types of AI

4.2. AI based on functionalities

The four types of AI based on functionalities :

Underneath Narrow AI, one of the three types based on capabilities, there are two functional AI categories:

Reactive machine AI :Reactive machines are AI systems with no memory and are designed to perform a very specific task. Since they can't recollect previous outcomes or decisions, they only work with presently available data. Reactive AI stems from statistical math and can analyze vast amounts of data to produce a seemingly intelligence output.

Examples of Reactive Machine AI

- IBM Deep Blue: IBM's chess-playing supercomputer AI beat chess grandmaster Garry Kasparov in the late 1990s by analyzing the pieces on the board and predicting the probable outcomes of each move.
- The Netflix Recommendation Engine: Netflix's viewing recommendations are powered by models that process data sets collected from viewing history to provide customers with content they're most likely to enjoy.
- Limited memory AI :Unlike reactive machine AI, this form of AI can recall past events and outcomes and monitor specific objects or situations over time. Limited memory AI can use past- and present-moment data to decide on a course of action most likely to help achieve a desired outcome. However, while Limited Memory AI can use past data for a specific amount of time, it can't retain that data in a library of past experiences to use over a long-term period. As it's trained on more data over time, Limited Memory AI can improve in performance.

Examples of Limited Memory AI :

- Generative AI: Generative AI tools such as ChatGPT, Bard and DeepAI rely on limited memory AI capabilities to predict the next word, phrase or visual element within the content it's generating.
- Virtual assistants and chatbots: Siri, Alexa, Google Assistant, Cortana and IBM Watson Assistant combine natural language processing (NLP) and Limited Memory AI to understand questions and requests, take appropriate actions and compose responses.
- Self-driving cars: Autonomous vehicles use Limited Memory AI to understand the world around them in real-time and make informed decisions on when to apply speed, brake, make a turn, etc.
- Theory of Mind AI :Theory of Mind AI is a functional class of AI that falls underneath the General AI. Though an unrealized form of AI today, AI with Theory of Mind functionality would understand the thoughts and emotions of other entities. This understanding can affect how the AI interacts with those around them. In theory, this would allow the AI to simulate human-like relationships. Because Theory of Mind AI could infer human motives and reasoning, it would personalize its interactions with individuals based on their

unique emotional needs and intentions. Theory of Mind AI would also be able to understand and contextualize artwork and essays, which today's generative AI tools are unable to do.Emotion AI is a theory of mind AI currently in development. AI researchers hope it will have the ability to analyze voices, images and other kinds of data to recognize, simulate, monitor and respond appropriately to humans on an emotional level. To date, Emotion AI is unable to understand and respond to human feelings.

– Self-Aware AI: Self-Aware AI is a kind of functional AI class for applications that would possess super AIcapabilities. Like theory of mind AI, Self-Aware AI is strictly theoretical. If ever achieved, it would have the ability to understand its own internal conditions and traits along with human emotions and thoughts. It would also have its own set of emotions, needs and beliefs.Emotion AI is a theory of mind AI currently in development. Researchers hope it will have the ability to analyze voices, images and other kinds of data to recognize, simulate, monitor and respond appropriately to humans on an emotional level. To date, emotion AI is unable to understand and respond to human feelings.

5. AI and facial recognition

1.facial recognition: Facial recognition is a way of identifying or confirming an individual's identity using their face. Facial recognition systems can be used to identify people in photos, videos, or in real-time.[17]

Figure 2.2 : facial recognition

Facial recognition systems utilize various AI techniques to identify and verify individuals based on their facial features.

1. Convolutional Neural Networks (CNNs): are Artificial Intelligence algorithms based on multi-layer neural networks that learns relevant features from images, being capable of performing several tasks like object classification, detection, and segmentation (Eduardo, 2019).

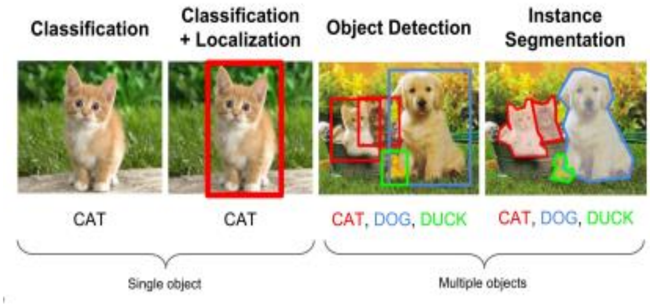


Figure 2. 3: Convolutional Neural Networks (CNNs)

2. Deep Learning: Deep learning, or deep neural networks, refers to a subfield of machine learning, which itself is a subset of artificial intelligence (AI), as illustrated in Figure 2.3. Artificial intelligence encompasses the concept of machines performing tasks that are considered complex by humans, such as reasoning, task planning, or structured knowledge representation.

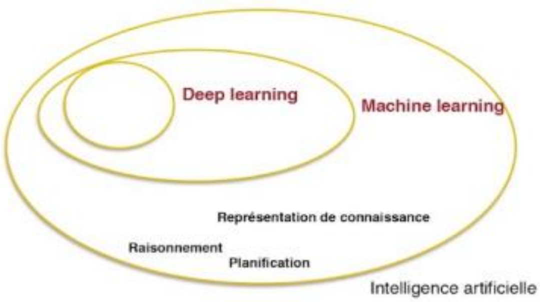


Figure 2. 4: Artificial intelligence and its subdomains

Deep learning, or deep neural networks, currently encompasses the most effective and high-performing methods applied within the machine learning community (Clément, 2021).

3.Facial Landmark Detection: Facial landmark detection aims to detect the location of predefined facial landmarks, such as the corners of the eyes, eyebrows, the tip of the nose. It has drawn much attention recently as it is a prerequisite in many computer vision applications. For example, facial landmark detection can be applied to a large variety of tasks, including face recognition , head pose estimation , facial reenactment and 3D face reconstruction.

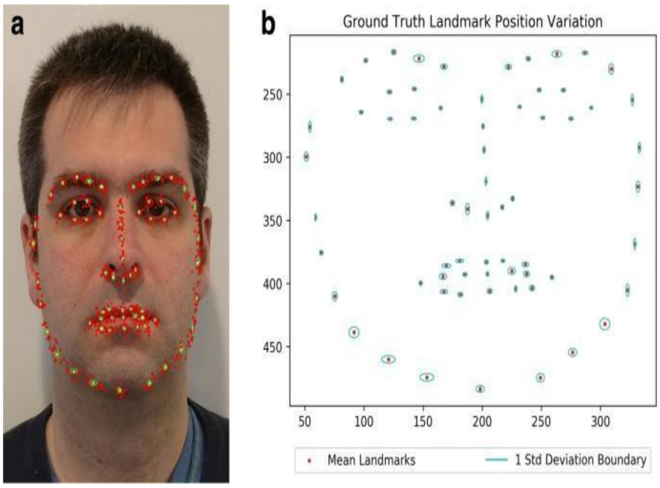


Figure 2.5: A review of image-based automatic facial landmark identification techniques.

Recent advances in facial landmark detection mainly focus on learning discriminative features from abundant deformation of face shapes and poses, different expressions, partial occlusions, and others.

A very typical framework is to construct features to depict the facial appearance and shape information by the convolutional neural networks (ConvNets) or hand-crafted features, and then learn a model, that is., a regressor, to map the features to the landmark locations . Most of them apply a cascade strategy to concatenate prediction modules and update the predicted locations of landmarks progressively (Xuanyi).

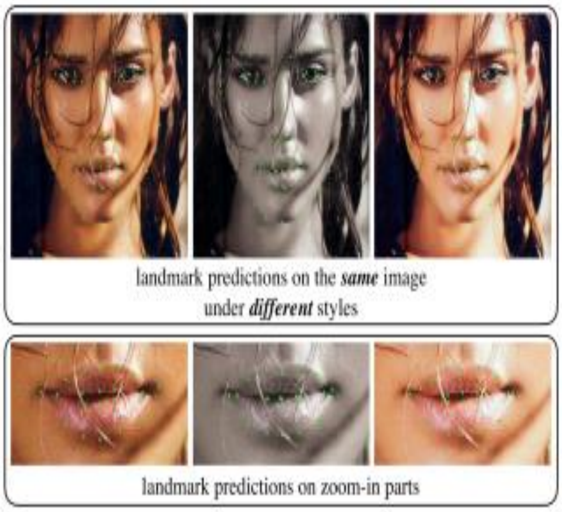


Figure 2. 6: A face image in three different styles and the locations of the facial landmarks predicted by a facial landmark detector on them.

4. 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. It covers standard deviation, covariance, and eigenvectors. This background knowledge is meant to make the PCA section very straightforward, but can be skipped if the concepts are already familiar.

The statistical information published in the area of facial recognition technology utilizing the PCA method reveals the significance of using this method for identifying and verifying facial features. Figure 1 below reveals the amount of publications that have used the words ‘face recognition’ and ‘PCA’ in their headings (Sasan, 2013)

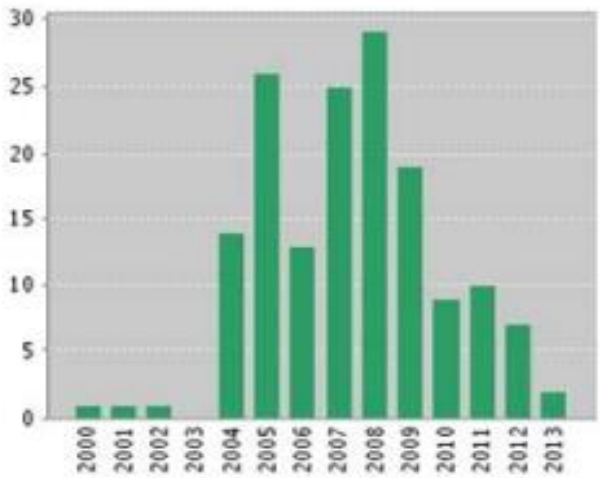


Figure 2. 7: Number of publication utilizing

6. Local Binary Patterns (LBP)

Local Binary Pattern was introduced by Timoojala (put the refrence). The standard version of the LBP of a pixel is formed by thresholding the 3X3 neighborhood of each pixel value with the center pixel's value. Let gc be the center pixel gray level and gi (i=0,1,..7) be the gray level of each surrounding pixel. Figure 2. 6 illustrate the basic LBP operation. If gi is smaller than gc, the binary result of the pixel is set to 0 otherwise set to 1. All the results are combined to get 8 bit value. The decimal value of the binary is the LBP feature.

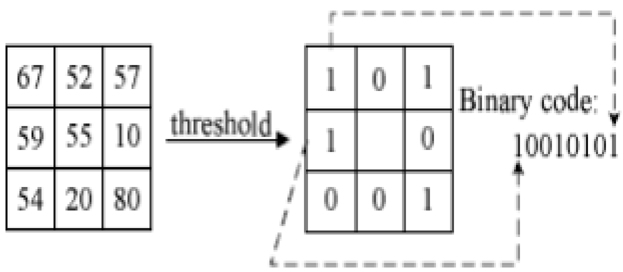


Figure 2. 8 : Illustration of Basic LBP operator

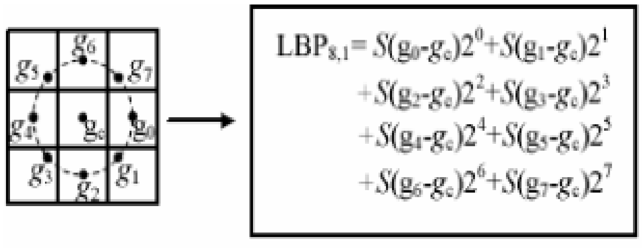


Figure 2. 9: The LBP operator of a pixel's circular neighborhoods with r=1 ,p=8

Bilinear interpolation method is used for a sampling point does not fall in the center of the pixel. Let LBPp,r denote the LBP feature of a pixel 's circularly neighborhoods, where r is the radius and p is the number of neighborhood points on the circle .From Figure 2. 7 we can write,

$$LBP_{p,r} = \sum_{p=0}^{p-1} s(g_i - g_c)\, 2^i \text{ , } S(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

The concept of uniform patterns is introduced to reduce the number of possible bins. Any LBP pattern is called as uniform if the binary pattern consists of at most two bitwise transitions from 0 to 1 or vice versa. For example if the bit pattern 11111111(no transition) or 00110000 (two transitions) are uniform where as 10101011 (six transition) are not uniform. The uniform pattern constraint reduces the number of LBP pattern from 256 to 58 and it is very useful for face detection (K.Meena, 2011).

7. Support Vector Machines (SVM)

SVM is a binary classification method that finds the optimal linear decision surface based on the concept of structural risk minimization.Support Vector Machine (SVM) is a supervised machine learning algorithm used for both classification and regression. Though we say regression problems as well it's best suited for classification. The main objective of the SVM algorithm is to find the optimal hyperplane in an N- dimensional space that can separate the data points in different classes in the feature space. The hyperplane tries that the margin between the closest points of different classes should be as maximum as possible. The dimension of the hyperplane depends upon the number of features. If the number of input features is two, then the hyperplane is just a line. If the number of input features is three, then the hyperplane becomes a 2-D plane. It becomes difficult to imagine when the number of features exceeds three.

- **Support Vector Machine Terminology :**

- **Hyperplane:** Hyperplane is the decision boundary that is used to separate the data points of different classes in a feature space. In the case of linear classifications, it will be a linear equation i.e. wx+b = 0.
- **Support Vectors:** Support vectors are the closest data points to the hyperplane, which makes a critical role in deciding the hyperplane and
- margin.
- **Margin**: Margin is the distance between the support vector and hyperplane. The main objective of the support vector machine algorithm is to maximize the margin. The wider margin indicates better classification performance.
- **Kernel**: Kernel is the mathematical function, which is used in SVM to map the original input data points into high-dimensional feature spaces, so, that the hyperplane can be easily found out even if the data points are not linearly separable in the original input space. Some of the common kernel functions are linear, polynomial, radial basis function(RBF), and sigmoid.
- **Hard Margin:** The maximum-margin hyperplane or the hard margin hyperplane is a hyperplane that properly separates the data points of different categories without any misclassifications.
- **Soft Margin:** When the data is not perfectly separable or contains outliers, SVM permits a soft margin technique. Each data point has a slack variable introduced by the soft-margin SVM formulation, which softens the strict margin requirement and permits certain misclassifications or violations. It discovers a compromise between increasing the margin and reducing violations.
- **C:** Margin maximisation and misclassification fines are balanced by the regularisation parameter C in SVM. The penalty for going over the margin or misclassifying data items

is decided by it. A stricter penalty is imposed with a greater value of C, which results in a smaller margin and perhaps fewer misclassifications.

- **Hinge Loss:** A typical loss function in SVMs is hinge loss. It punishes incorrect classifications or margin violations. The objective function in SVM is frequently formed by combining it with the regularisation term.
- **Dual Problem:** A dual Problem of the optimisation problem that requires locating the Lagrange multipliers related to the support vectors can be used to solve SVM. The dual formulation enables the use of kernel tricks and more effective computing.

- **Mathematical intuition of Support Vector Machine :**

Consider a binary classification problem with two classes, labeled as +1 and -1. We have a training datasetconsisting of input feature vectors X and their corresponding class labels Y.

The equation for the linear hyperplane can be written as:

$w^T x + b = 0$

The vector W represents the normal vector to the hyperplane. i.e the direction perpendicular to the hyperplane. The parameter **b** in the equation represents the offset or distance of the hyperplane from the origin along the normal vector **w**. The distance between a data point x_i and the decision boundary can be calculated as:

$$d_{i=}\frac{W^T x_i + \mathrm{b}}{\| w \|}$$

Where ||w| represents the Euclidean norm of the weight vector w. Euclidean norm of the normal vector WFor Linear SVM classifier:

$$\hat{y}\begin{cases} 1: w^T x + b \geq 0 \\ 0: w^T x + b < 0 \end{cases}$$

Optimization:

$$\frac{minimize}{w,b}\frac{1}{2}w^t w = \frac{minimize}{W,b}\frac{1}{2}\|W\|^2$$

Subject to $y_i(w^T x_i + b) \geq 1$ for I = 1,2,3, ...,m

The target variable or label for the i[th] training instance is denoted by the symbol ti in this statement. And ti=-1 for negative occurrences (when yi= 0) and ti=1positive instances (when yi = 1) respectively. Because we require the decision boundary that satisfy the constraint: $(wTxi + b) \geq 1$

For Soft margin linear SVM classifier:

Subject to $(w^T x_i + b) \geq 1 - C_i$ and $C_i \geq 0$ for I = 1,2,3, ,m

Dual Problem: A dual Problem of the optimisation problem that requires locating the Lagrange multipliers related to the support vectors can be used to solve SVM. The optimal Lagrange multipliers α(i) that maximize the following dual objective function

$$\text{Maximize}_{\alpha} : \frac{1}{2}\sum_{i\to m}\sum_{i\to m}(\alpha_i \alpha_j t_i t_j \, k(x_i, x_j)) - \sum_{i\to m}\alpha_i$$

Where

αi : is the Lagrange multiplier associated with the ith training sample.

K(xi, xj) is the kernel function that computes the similarity between two samples xi and xj. It allows SVM to handle nonlinear classification problems by implicitly mapping the samples into a higher-dimensional feature space. The term ∑αi represents the sum of all Lagrange multipliers. The SVM decision boundary can be described in terms of these optimal Lagrange multipliers and the support vectors once the dual issue has been solved and the optimal Lagrange multipliers have been discovered. The training samples that have i > 0 are the support vectors, while the decision boundary is supplied by:

W = $\sum_{i\to m} \alpha_i t_i \mathrm{k}(x_i, x_j)$ + b

$(w^T x_i - b) = 1 \leftrightarrow \mathrm{b} = w^T x_i - t_i$

Types of Support Vector Machine :

Based on the nature of the decision boundary, Support Vector Machines (SVM) can be divided into twomain parts:

- **Linear SVM:** Linear SVMs use a linear decision boundary to separate the data points of different classes. When the data can be precisely linearly separated, linear SVMs are very suitable. This means that a single straight line (in 2D) or a hyperplane (in higher dimensions) can entirely divide the data points into their respective classes. A hyperplane that maximizes the margin between the classes is the decision boundary.
- **Non-Linear SVM:** Non-Linear SVM can be used to classify data when it cannot be separated into two classes by a straight line (in the case of 2D). By using kernel functions, nonlinear SVMs can handle nonlinearly separable data. The original input data is transformed by these kernel functions into a higher- dimensional feature space, where the data points can be linearly separated. A linear SVM is used to locate a nonlinear decision boundary in this modified space.
- **Gaussian Mixture Models (GMM):** A Gaussian Mixture Model (GMM) is a parametric probability density function represented as a weighted sum of Gaussian component densities. GMMs are commonly used as a parametric model of the probability distribution of continuous measurements or features in a biometric system, such as vocal-tract related spectral features in a speaker recognition system. GMM parameters are estimated from training data using the iterative Expectation-Maximization (EM) algorithm or Maximum *A Posteriori* (MAP) estimation from a well-trained prior model.[5]
- here are some techniques that extend the foundational principles of Gaussian Mixture Models(GMMs):(Ralph)
- **PCA Based Face Recognition:** Among the numerous face recognition algorithms introduced in recent years, the eigenface approach proposed by Turk and Pentland is one of the most influential ones. A face image, if interpreted as a vector, defines a point in a high dimensional space. Different face images share a number of similarities with each other, so the points representing these images are not randomly distributed in the image space. The key idea of the recognition process is to map the face images into an

appropriately chosen lower dimensional subspace and perform classification by distance computation. If we restrict ourselves to a linear dimensionality reduction, the optimal solution is provided by principal component analysis . The basis functions of the lower dimensional "face space" are formed by the eigenvectors of the covariance matrix of the set of training images corresponding to the largest eigenvalues. In the context of face recognition these eigenvectors are called "eigenfaces".

- **Pose Invariant Face Recognition :** The eigenface approach works reasonably well only in mugshot settings where the input space is restricted to frontal face images. An extension by Pentland et al deals with the problem of multiple head poses by building separate eigenspaces for nine different views. In the recognition stage they first determine the subspace which is most representative for the test image and then find the closest match between this image and a model in the chosen subspace. A different approach was proposed by Graham and Allinson . They built a common eigenspace from faces of all views and observed that a face which continuously changes pose between the two profile views forms a convex curve in the subspace. Using a radial basis function network they were able to exploit this fact and recognize faces in previously unseen views.

  Cootes et. al proposed Active Appearance Models, which combine shape and gray-level appearance. During localization a generic face model is deformed to fit the input face.

- **Growing Gaussian Mixture Models :** Given a face image $\varkappa$ and classes C_k we are interested in the probability $p\ (C_k|\varkappa)$, that $\varkappa$ belongs to class C_k , where each class represents a different person. Using Bayes' rule we link the posterior probability $p\ (C_k|\varkappa)$ to the class conditional probability $(C_k|\varkappa)$. Assuming equal class priors we can determine the most likely class for an image $\varkappa$ with a maximum likelihood estimation

$$C^{n} = \arg \max_{k} p(\varkappa|C_k) \approx \arg\max_{k} p\ (\varkappa|C_k)$$

We model the class conditional probability $p\ (\varkappa|C_k)$ with a Gaussian mixture model:

$$p\ (\varkappa|C_k) = \sum_{J=1}^{M} P\ (\varkappa\ |\ j)p(j)$$

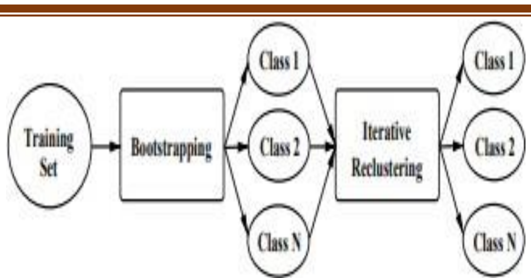
where each mixture component $P\ (\varkappa \mid j)$is a Gaussian distribution with mean μ_j and covariance matrix Σ_j

$$p(\varkappa \mid j) = \frac{1}{(2\pi)^{\frac{d}{2}}\left|\Sigma_j\right|^{\frac{1}{2}}} e^{-\frac{1}{2}(z-\mu_j)^T \Sigma_j^{-1} (z-\mu_j)}$$

Given a predefined number of mixtures K the models can be trained using the EM algorithm(the EM algorithm is an iterative method that starts with some initial estimate of parameters ,and then proceeds to iteratively update these parameters until convergenceis detected , thealgorithme consists of two steps the expectation ‘E’ step and the maximization ‘M’ step .

In the context of GMM the EM algoritm is used to estimate theparameters of the GMM (the means,variances and weights of the gaussian components)that maximize the like lihood of the observed data [19]). The problem lies in the determination of M. The input to the recognizer is comprised of a stream of face images captured in the meeting room. As people move freely about the room any head pose can occur. The number of different model categories and therefore the optimal choice of M is unknown. To address this problem a growing scheme for the GMM was implemented. The process is initialized with a single mixture component for each model computed from the sample mean and covariance matrix of the training set. In order to reduce the number of parameters that have to be estimated in each step, diagonal covariance matrices $\Sigma_J = \sigma^2$ are used. As we perform principal component analysis prior to recognition we can assume that the different feature dimensions are uncorrelated. The algorithm then proceeds in two steps. During bootstrapping the training samples are evaluated using the model for the respective class only. From a pool of samples with low probability we randomly draw a vector and use it as a seed point for a Gaussian mixture. The training set is clustered according to the previous mean and the new seed point using neural gas clustering, which is an extension to kmeans clustering.

The model parameters are then re-estimated using the EM algorithm. After two bootstrapping iterations the algorithm trains the models discrimatively. It evaluates the training samples using all models and records those which are misclassified. From the



misclassified samples the example with the highest probability is used as the starting point for a new Gaussian mixture. The same procedure of re-clustering and parameter re-estimation is then iteratively applied until a local minimum in the number of misclassified samples for each model is found. Figure 2.8 shows an overview over the training procedure.

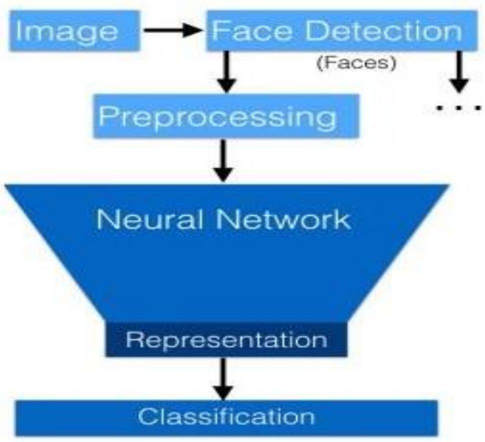
Figure 2.10: Overview of the training procedure

8. Programs and therories related to artificial intelligence

8.1 AI programs used in facial recognition

Some of major AI programs used in facial recognition.

– OpenFace: OpenFace is in mobile scenarios where a user's real-time face recognition system adapts depending on context. Our key design consideration is a system that gives



high accuracy with low training and predictiontimes.

Figure 2.11: Logic flow for face recognition with a neural network.

OpenFace provides the logic flow presented in Figure 2.9 to obtain low-dimensional face representations for the faces in an image.

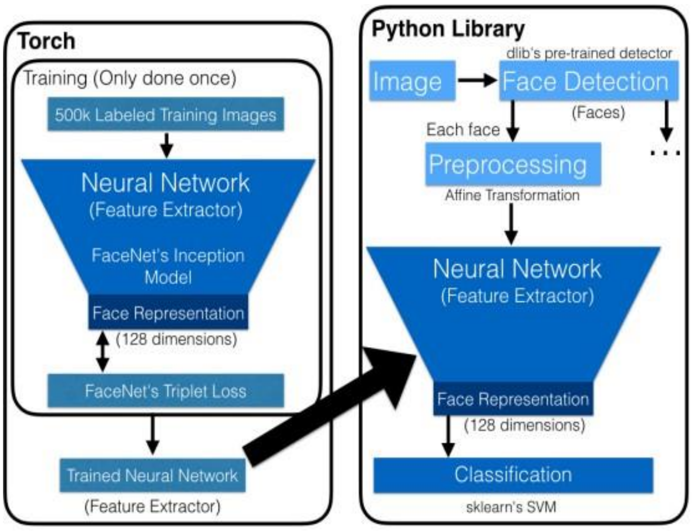


Figure 2.12: OpenFace's project structur.

Figure 2. 10, highlights OpenFace's implementation. The neural network training and inference portions use Torch [CKF11], Lua [IDFCF96] and luajit [Pal08]. Our Python [VRDJ95] library uses numpy [Oli06] for arrays and linear algebra operations, OpenCV [B+00] for computer vision primitives, and scikit-learn [PVG+11] for classification. it also provide plotting scripts that use matplotlib [H+07]. The project structure is agnostic to the neural network architecture and currently use FaceNet's architecture [SKP15]. and use dlib's [Kin09] pre-trained face detector for higher accuracy than OpenCV's detector. and compile native C code with gcc [Sta89] and CUDA code with NVIDIA's LLVM-based [LA04] nvcc(Brandon , 2016).

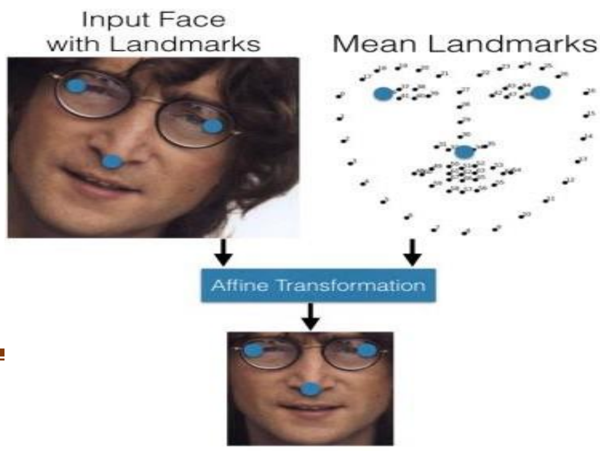


Figure 2.13: OpenFace's affine transformation. The transformation is based on the large blue landmarksand the final image is cropped to the boundaries and resized to 96 × 96 pixels.

– Dlib: Dlib is an open source suite of applications and libraries written in C++ under a permissive Boost license. Dlib offers a wide range of functionality across a number of machine learning sectors, including classification and regression, numerical algorithms such as quadratic program solvers, an array of image processing tools, and diverse networking functionality, among many other facets.Dlib also features robust tools for object pose estimation, object tracking, face detection (classifying a perceived object as a face) and face recognition (identifying a perceived face).[14]

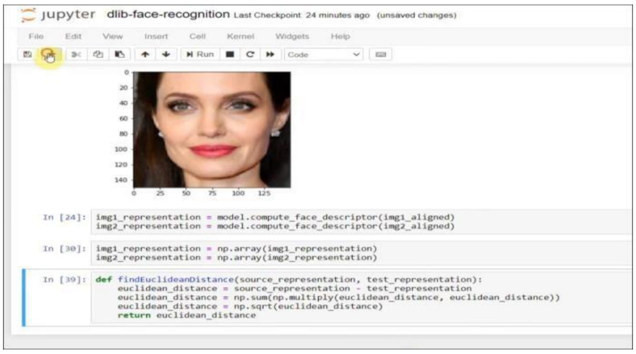
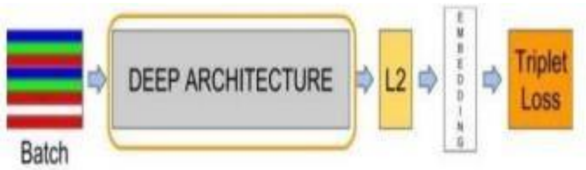


Figure 2.14: Dlib being used in a Jupyter notebook as part of a novel facialrecognition framework.

– FaceNet: FaceNet is a method that uses deep convolutional networks to optimize its embedding, compared to using intermediate bottleneck layers as a test of previous deep learning approaches. This method is called one- shot learning. In more detail, this method can use a small sample of face images to produce the initial model, and when there are new models, the initial model can be used without retraining. FaceNet directly trains the face using the Euclidean space where the distance consists of similarities between facial models. When the results of similarities between face models are obtained, it will be easy to carry out face recognition and classification using FaceNet attached become feature vectors.

FaceNet applies triplets by matching face to face with the online novel triplet mining method. Of course, this triplet consists of a collection of anchor images, where each image consists of positive and negative images. FIGURE 2.13 shows the structural model used in FaceNet. FaceNet consists of batch layers as input and deep architecture which is deep CNN followed by L2 normalization, that become the result of face embedding .FaceNet also pursued by the triplet loss when the training process, see FIGURE 2.14 . Triplet loss training methods have three main elements namely anchor, positive and negative. This triplet loss works by minimizing the distance between anchors positively and maximizing the distance between anchors negatively. Where this positive has the same identity as the anchor and



negative has a different identity from the anchor

Figure 2.15: The model structure of the FaceNet.

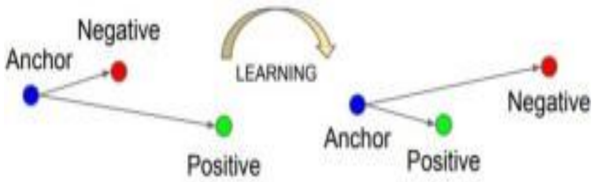


Figure 2.16: The triplet loss training.

FaceNet trains its output directly into concise 128- dimensional embedding by apply triplet based loss method depend on LMNN. It formed by two thumbnails of compared faces and thumbnails that do notmatch and the loss aim to distinguish between positive and negative pairs using a range of limit.

Thumbnails were cut tightly on the face field, it didn't need 2D or 3D adjustment, apart from the ratio andtranslation implemented (Ivan).

– DeepFace: DeepFace is a machine learning-based face recognition framework that helps bridge the gap between machine learning and software engineering. In the DeepFace framework, there are models of face recognition, such as Facenet .ArcFace .SFace, VGG-Face, OpenFace, Dlib, Human-Beings, and DeepID. DeepFace uses the Labelled Face in the Wild (LFW) and Youtube Face (YTF) datasets for training existing models. Table 2.1

contains the accuracy of each model in DeepFace. All available models can have an accuracy above 90% in testing with LFW and YTF datasets. Only OpenFace has an accuracy below 95%, at 93.80%. Facenet512 is the model that has the best accuracy of 99.65% using the LFW dataset (Andrian, 2023).

Tableau 2.1: Comparison between Deep Face model

Model	LFW Score	YTF Score
Facenet512	99.65%	-
Facenet	99.20%	-
SFace	99.60%	-
ArcFace	99.41%	-
Dlib	99.38%	-
VGG-Face	98.78%	**97.40%**
Human-Beings	97.53%	-
OpenFace	93.80%	-
DeepID	-	**97.05%**

8.2 Theories underlying ai and its use in facial recognition

The use of AI in facial recognition is grounded in various theories and methodologies. Here are some key theories underlying AI and its application in facial recognition, along with relevant references:

- **Machine Learning :** Machine learning (ML) is the scientific study of algorithms and statistical models that computer systems use to perform a specific task without being explicitly programmed. Learning algorithms in many applications that's we make use of daily. Every time a web search engine like Google is used to search the internet, one of the reasons that work so well is because a learning algorithm that has learned how to rank web pages.These algorithms are used for various purposes like data mining, image processing, predictive analytics, etc. to name a few. The main advantage of using machine learning is that, once an algorithm learns what to do with data, it can do its work automatically. In this paper, a brief review and future prospect of the vast applications of machine learning algorithms has been made(Batta ,2020).

- **Deep Learning :** Deep learning is a computer-based modeling approach, which is made up of many processing layers that are used to understand the representation of data with several levels of abstraction. This review paper presents the state of the art in deep learning to highlight the major challenges and contributions in computer vision. This work mainly gives an overview of the current understanding of deep learning and their approaches in solving traditional artificial intelligence problems. These computational models enhanced its application in object detection, visual object recognition, speech recognition, face recognition, vision for driverless cars, virtual assistants, and many other fields such as genomics and drug discovery.Finally, this paper also showcases the current developments and challenges in training deep neuralnetwork(Ranjan, 2021).
- **Biometric Theories :** The term biometrics is derived from the Greek words bio meaning "life" and metrics meaning " tomeasure". Biometrics refers to the identification or verification of a person based on his/her physiological and/or behavioral characteristics. Several verification/identification based biometrics have evolved based on various unique aspects of human body, ease of acquiring the biometric, public acceptance and the degree of security required. This paper presents an overview of various biometrics in use/proposed and their applicability to different activities (Jammi, 2010).
- **Triplet Loss :** The loss function is used to measure how well your model fits into the dataset. Triplet loss is also a loss function for machine learning algorithms where a reference input (called anchor) is compared to a matching input (called positive) and a non-matching input (called negative). The distance from the anchor to the positive is minimized, and the distance from the anchor to the negative input is maximized (V.L.B. DeMel , 2021).

9. Conclusion

The use of Artificial Intelligence (AI) in facial recognition represents a significant technological advancement, providing opportunities in both the security sector and the streamlining of daily processes. This revolutionary technology has the potential to enhance the accuracy of security systems, facilitate access to services, and strengthen identity management.

However, its deployment raises ethical and privacy concerns, necessitating strict regulation and oversight. The massive collection of biometric data and potential risks of misuse underscore the need for a robust legal framework to govern the use of facial recognition.

Despite the challenges, there is no denying that AI offers substantial benefits in the realms of security and convenience. The continuous development of this technology should be guided by strong ethical principles, ensuring responsible use and respect for individual rights. It is imperative to strike a balance between the practical advantages of facial recognition and the protection of individual liberties, in order to fully harness this innovation while minimizing potential risks.

CHAPTER 3:

Facial recognition code

Chapter3: Facial recognition code

In this chapter, we will delve into the intricacies of our facial recognition implementation. We will elucidate the workings of our algorithm, which relies on face detection using convolutional neural network (CNN) methods. Furthermore, we will scrutinize the pivotal role of artificial intelligence in this process, including its training from training data and its utilization for face identification and classification.

1. facial recognition système

A facial recognition système is a technology potentially capable of matching a human face from a digital image or a video frame against a database of faces. Such a system is typically employed

to authenticate users through ID verification services, and works by pinpointing and measuring facialfeatures from a given image.(Thorat, 2010).

2. facial recognition système and image processing

Image processing involves techniques and algorithms to manipulate and analyze digital images. Face recognition specifically focuses on identifying and verifying human faces within images or videos. It utilizes various image processing techniques such as pattern recognition, feature extraction, and machine learning algorithms to achieve its goal of detecting and recognizing faces accurately.[52]

2.1 Digital Image

A digital image is a two-dimensional representation of visual information, typically in the form of a grid of pixels.It is created by sampling the continuous intensity variations of an analog scene at discrete points.Digital images can be captured using digital cameras, scanners, or generated using computer graphics software.They consist of pixel values that represent color and brightness information at each point in the image.Digital images can be in various formats, such as JPEG, PNG, BMP, etc., depending on the compression and storage requirements.[22]

2.2 Digital Image Processing

Digital image processing refers to the manipulation and analysis of digital images using algorithms and techniques to enhance, modify, or extract information from the images.It involves a series of operations like filtering, segmentation, edge detection, noise reduction, and feature extraction.The goal of digital image processing is to improve the visual quality of images, extract meaningful information, or make images suitable for specific applications like computer vision, medical imaging, and more.It can be used for tasks like image restoration, image compression, object recognition, and image analysis.Digital image processing plays a crucial role in various fields, including medicine, remote sensing, entertainment, and scientific research.[22]

2.3 Programming languages that they are used in image processing

Image processing can be done using various programming languages, but some of the commonly used languages for image processing are:[22]

- Python: Python is one of the most popular languages for image processing, thanks to libraries like OpenCV, Pillow, scikit-image, and many others.
- Matlab: Matlab provides a comprehensive environment for image processing with built-in functions and toolboxes.
- C/C++: These languages are used for performance-critical image processing applications, often in combination with libraries like OpenCV.
- Java: Java offers image processing capabilities through libraries like Java Advanced Imaging (JAI) and ImageJ.
- R: R has packages like "imager" and "EBImage" for image analysis and processing. Now, let's discuss the difference between a digital image and digital image processing:

3. The programming language used :

3.1 Python:

Python is an interpreted, object-oriented, high-level programming language with dynamic semantics. Its high-level built in data structures, combined with dynamic typing and dynamic binding, make it very attractive for Rapid Application Development, as well as for use as a scripting or glue language to connect existing components together. Python's simple, easy to learn syntax emphasizes readability and therefore reduces the cost of program maintenance. Python supports modules and packages, which encourages program modularity and code reuse. The Python interpreter and the extensive standard library are available in source or binary form without charge for all major platforms, and can be freely distributed.[23]

Figure 3.1: Python.

3.2 Why did we choose python

In the context of this project , we chose to use the Python programming language due to its simplicity and readability ,which allows us to reduce the complexity and length of the code moreover,Python is particulartly suited to emerging research fields such as artificial intelligence thanks to its vast ecosystem of specialized libraries.Finally, Python's ability to run on ARM processors, like those used in many embedded systems ,was a key factor in our decision . This has allowed us to optimize the efficiency and portability of our solution.

4. Software environment

4.1 Anaconda Navigator

Figure 3.2: Anaconda Navigator.

Anaconda Navigator is a desktop graphical user interface (GUI) included in Anaconda® Distribution that allows to launch applications and manage conda packages, environments, and channels without using command line interface (CLI) commands. Navigator can search for packages on Anaconda.org or in a local Anaconda Repository. It is available for Windows, macOS, and Linux. [52]

4.2 Spyder

Figure 3.3: spyder.

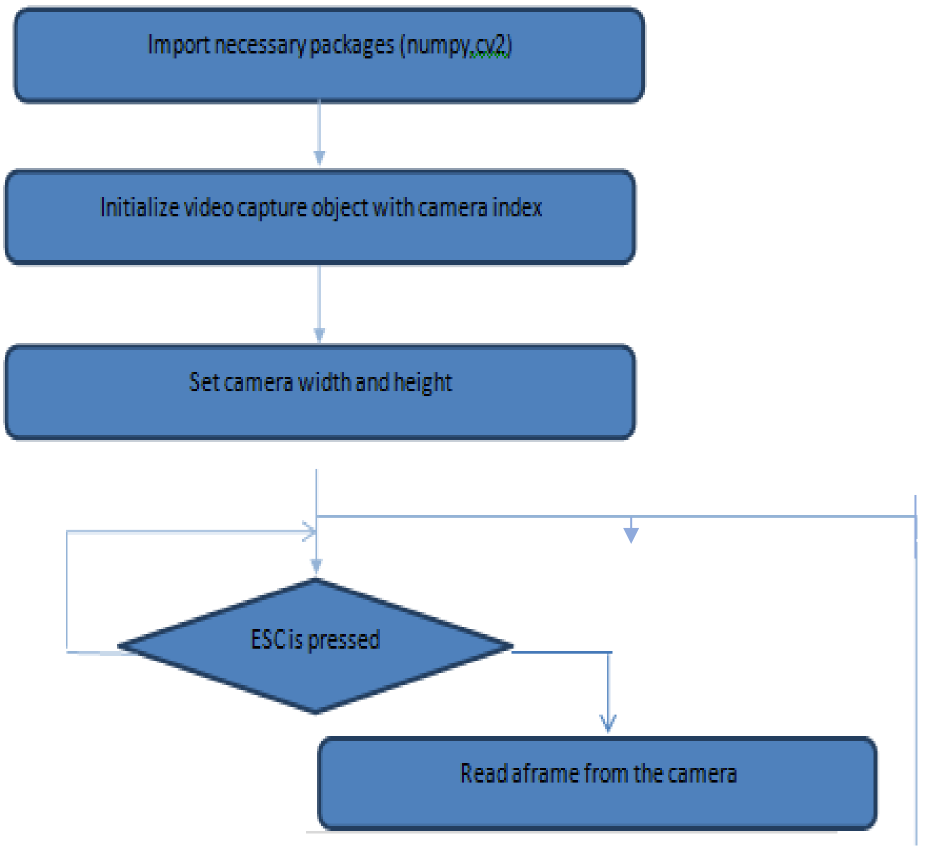
Spyder is a free and open source scientific environment written in Python, for Python, and designed by and for scientists, engineers and data analysts. It features a unique combination of the advanced editing, analysis, debugging, and profiling functionality of a comprehensive development tool with the data exploration, interactive execution, deep inspection, and beautiful visualization capabilities of a scientific package.[53]

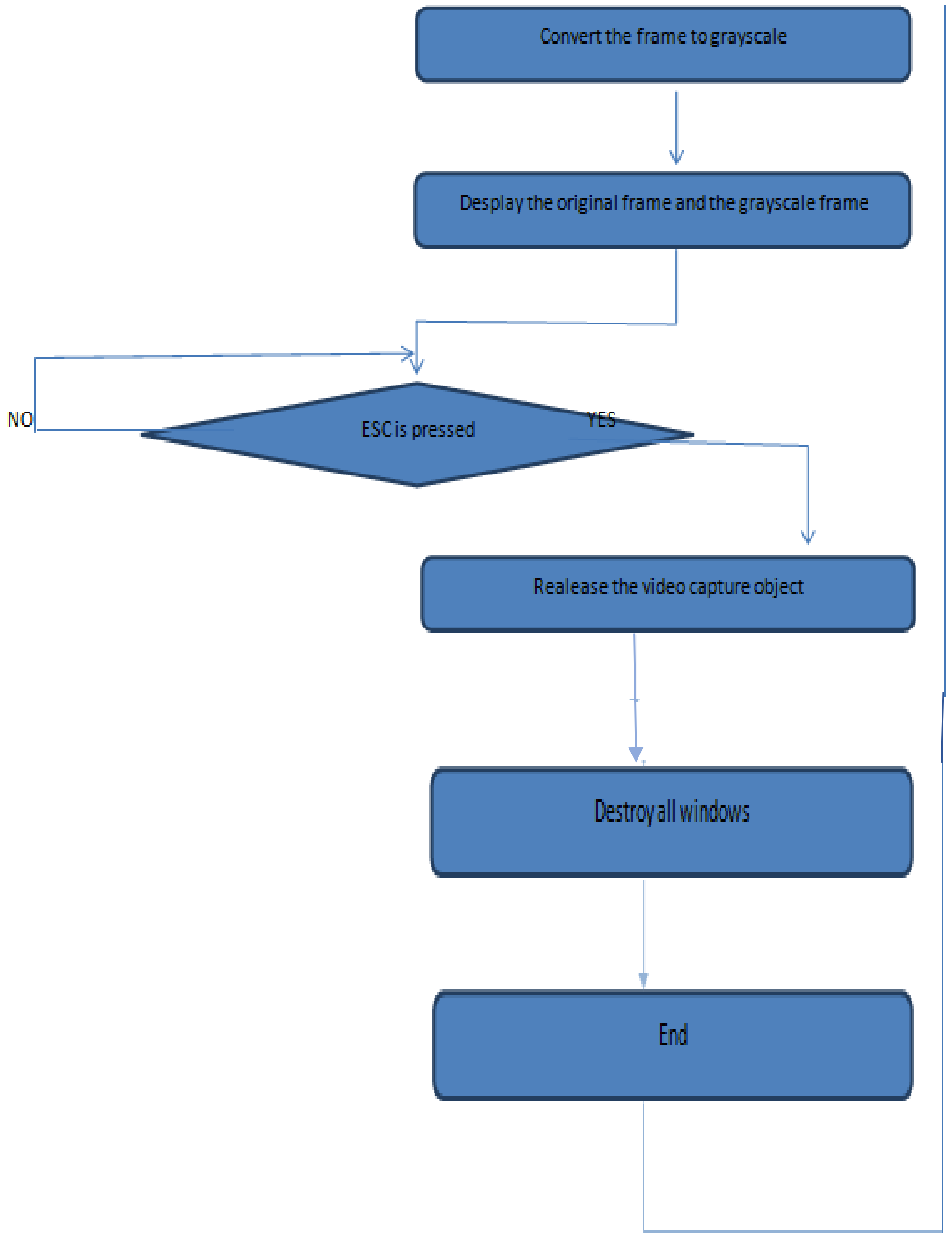
5. *Code*

5.1 Camera testing

First we will test our camera. We enter the following Python code in IDLE:

flowchart :





5.2 The code explication

The code uses OpenCV library to capture video from the default camera, then convert the video to grayscale , and display both the original video and the grayscale video in real-time.

5.3 Face Detection

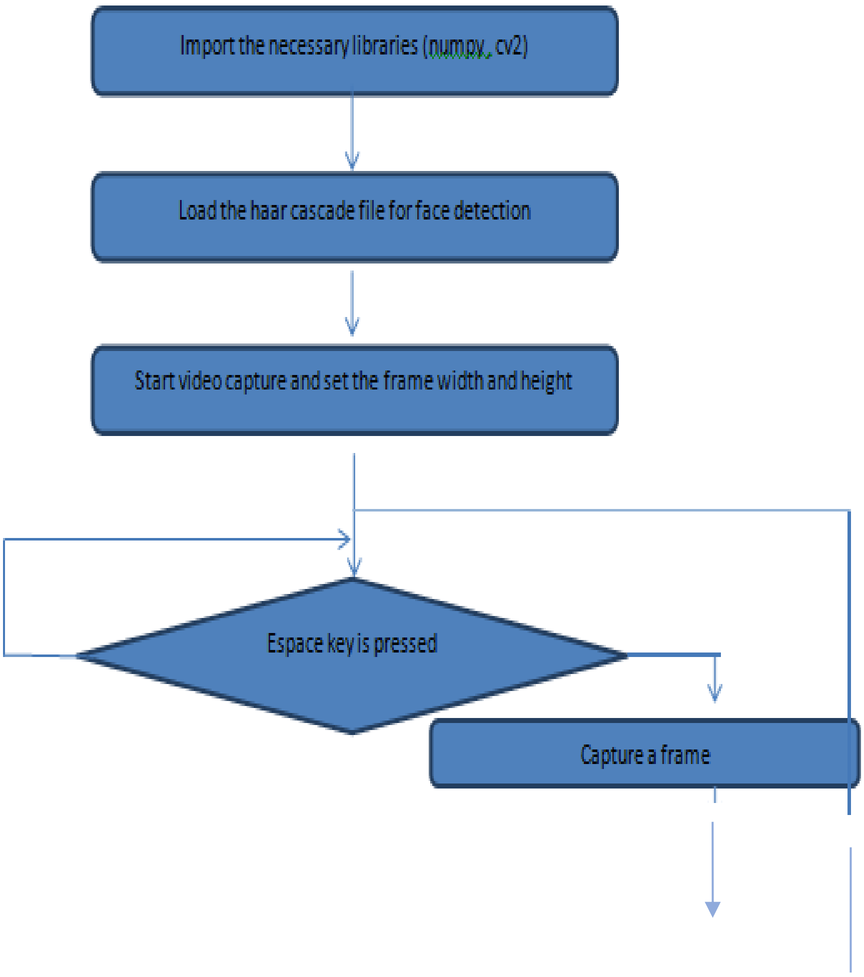
The initial phase of facial recognition involves face detection .the primary requirement is to capture anexisting face for recognition, rather than a newly captured face in the future.

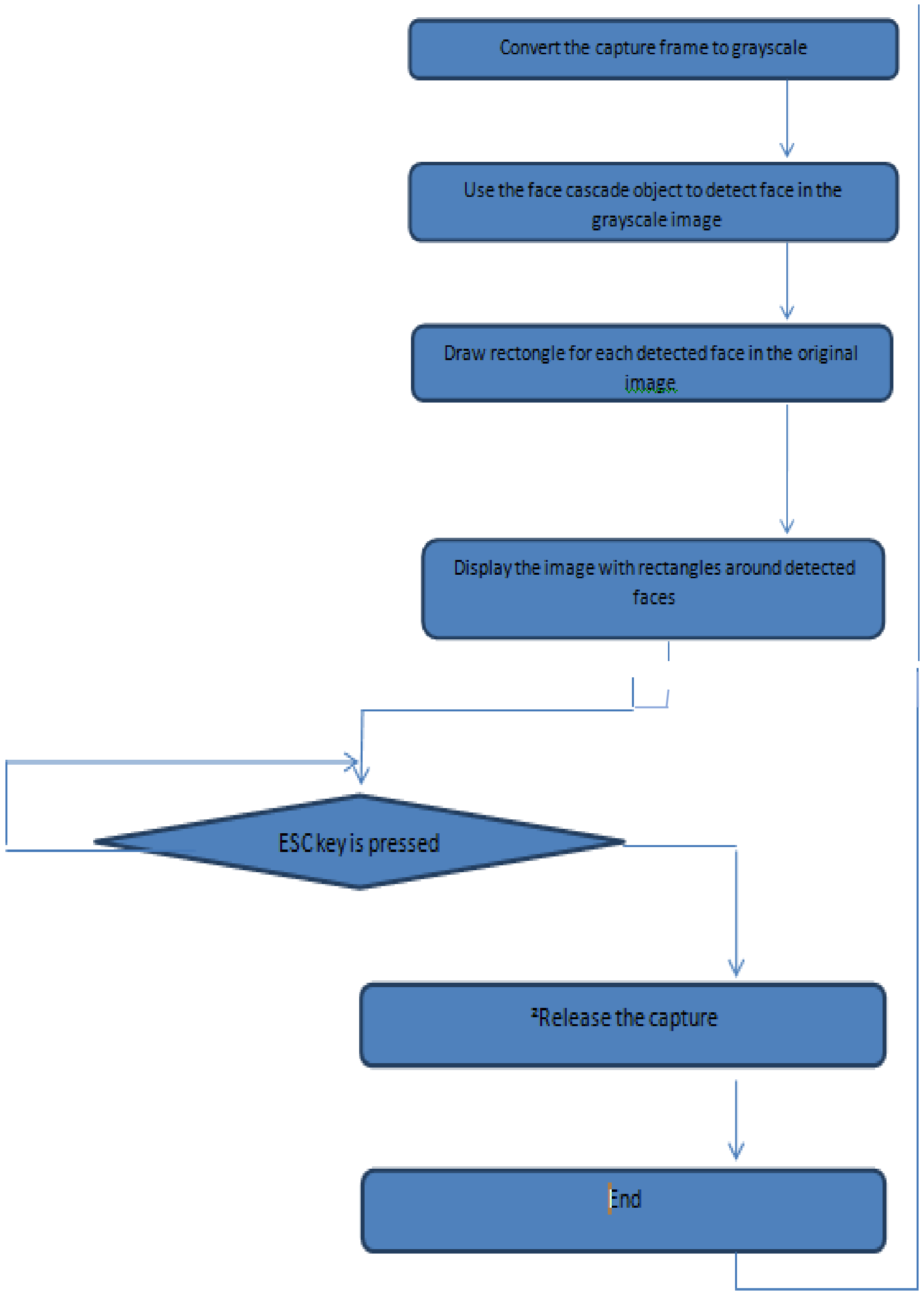
The haar cascade classifier is commonly employed for face detection (or any other object detection).thismethod is based on machine learning, where a cascade function is trained using a substantial number of positive and negative images,the trained function is then utilized to detect object in other images ,in thiscontext, we are focusing on face detection.

The algorithme initially necessities a significant number of positiveimages(images of face)andnegative images(images without faces)to train the classifier, subsequently,feature extraction is performed.

OpenCV already includes numerous Predefined face classifiers we need download the face detection modelfrom GitHub and place it in the directory.

Flowchart :


Import the necessary libraries (numpy, cv2)
Load the haar cascade file for face detection
Start video capture and set the frame width and height
Espace key is pressed
Capture a frame


Convert the capture frame to grayscale
Use the face cascade object to detect face in the grayscale image
Draw rectongle for each detected face in the original image
Display the image with rectangles around detected faces
ESC key is pressed
Release the capture
End

Technical and mathematical explication of the code:

This Python code,designed for real-time face detectin using a webcam,operates as follows: Firt, it imports the necessary libraries,numpy for numerical operations and cv2 for computer vision tasks, it then loads the haar cascade classifier, amachine learning objectdetection method,froma specified path,the script initializesa video capture obect to capture video trom the webcam and sets the frame width and height to 640 and 480 pixels,respectively.

The code enters a loop whereif Espace key is pressed it captures frames from the video,the captured frame is flipped vertically and converted to grayscale to simplify the image and reduce computational complexity the haar cascade classifier in then used to detect faces inthe grayscale image, the detectMultiScale function used with a scale factor of 1.2 and a minimum neighbor parameter of 5 and the minimum size of the detected face is set to 20x20 pixels.

For each detected face, a rectangle is drawn around the face in the original image with a color of blue (255,0,0) and a thickness of 2 pixels , the original image now with rectangles around detected faces isdisplayed in a new window.

The script checks if the ESC key has been pressed if it has, it breaks the loop and stops capturing video after the loop is exited, the video capture is releasedand all created windows are destroyed this marksthe end ofthe script.

5.4 Database creation

We will establish a database to store a set of grayscale images for each identified individual. These imageswill contain the facial features used to detect and recognize faces.

Before we begin, we need to prepare the development environment. For this, we will create a folder dedicated to this project. In this folder, we will find three Python scripts. These scripts have been specially designed for this project and are based on the face detection model we have chosen to use.

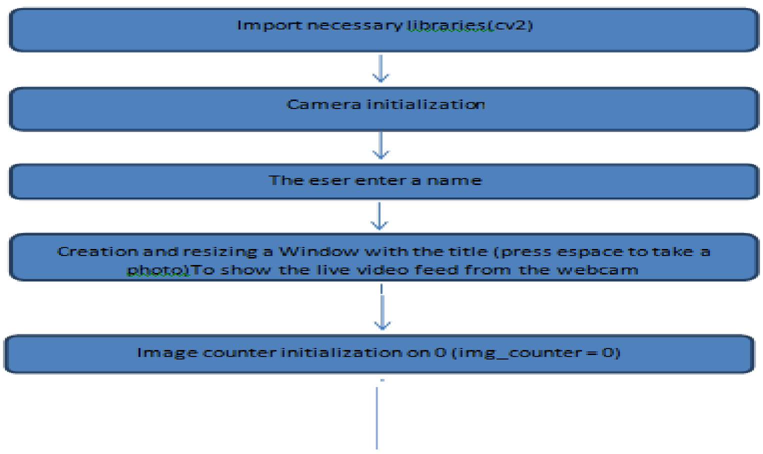
In addition to these scripts, we will also have two additional folders. The first folder will be used to store the previously mentioned database. This database will contain the grayscale images of each individual.

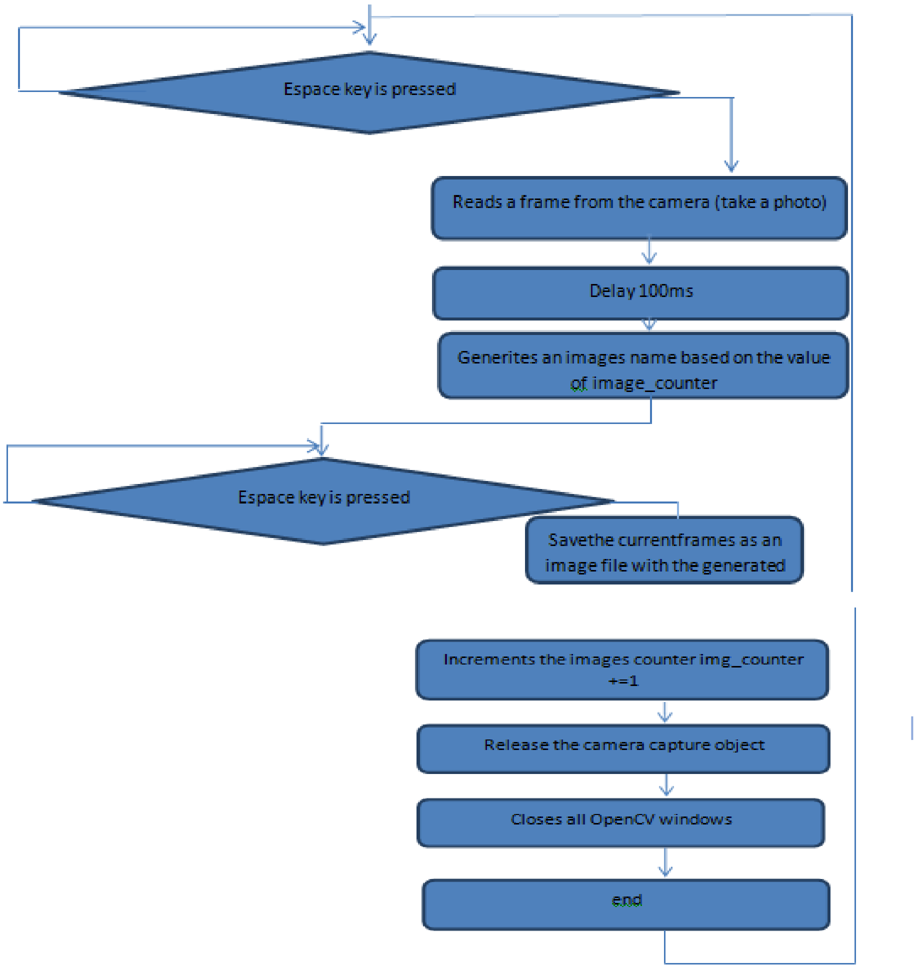
The second folder will be used for the training of the facial recognition model. It will contain the necessary data to train our model to correctly recognize faces. Images are converted to grayscale in facial recognition systems for severl reasons :[24]

- Data Simplification: Color images contain much more information than grayscale images. By converting an image to grayscale, we reduce the amount of data to process, which can make the facial recognition algorithm faster and more efficient.
- Uniformity: Color images can vary depending on lighting, shadows, etc. By using grayscale images, we can achieve a more uniform representation of faces, which can improve the accuracy of facial recognition.
- Focus on Important Features: The important features for facial recognition, such as the contours of the face, eyes, nose, mouth, etc., are often more clearly visible in grayscale than in color.

That's why, in our code, the captured image is converted to grayscale before being processed by the face detection model. This improves the accuracy and efficiency of facial recognition.

Flowchart:





the code explication:

The code capture frames from the webcam,display them in a window,and allows the user to take photos bypressing the ’space’ key , the captured images are saved in a directory named dataset with filenames like “image_0.jpg”, “image_1.jpg”, .

The code starts by importing the cv2 library stands for OpenCV(open source computer vision library) Then , the user is prompted to enter their name and the name is stored in the variable name.

Next a window with the title ‘press space to take a photo’ is created and resized to dimensions 500x300pixels .

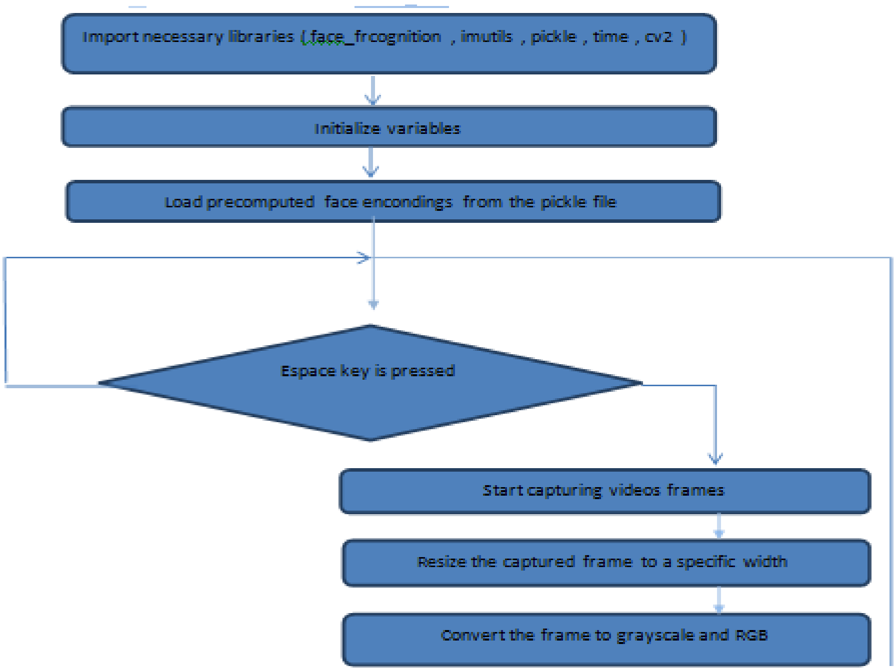
Inside the while true loop,whenEspace key is pressed the code captures frames from the camera.and if the frams is successfully read it is displayed in the named window ,when the user presses the ESCkey,the current frame is saved as an image file ,the image file is constructed as "dataset/" + name + "/image_{img_counter}.jpg", where img_counter keeps track of the image number.

Finally the named window is closed.

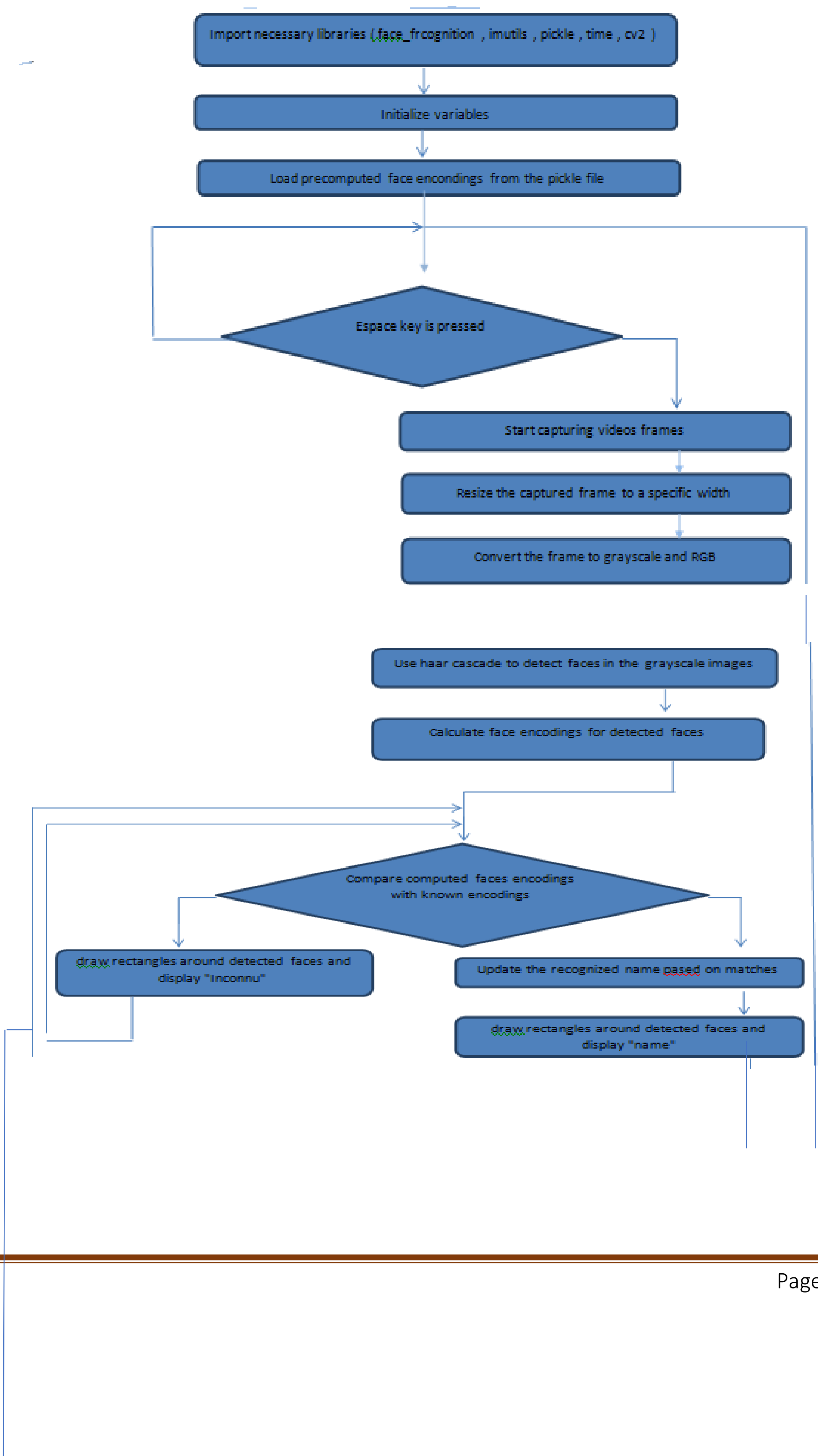
Basic facial recognition code

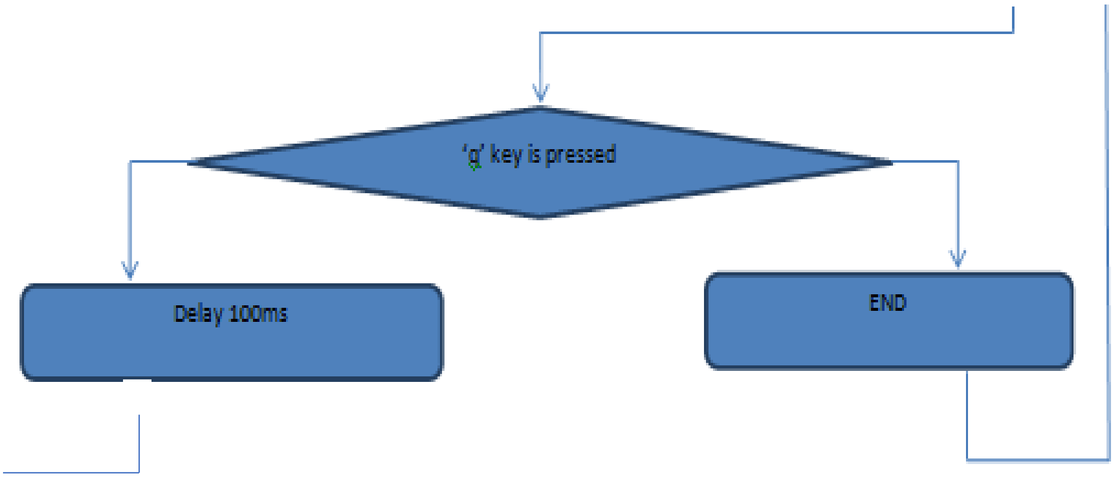
In this code, the program will draw a rectangle around every face it detects. If the face is recognized as “Mr.malek”, it will label the face as “Mr. Malek”. If the face is recognized as someone else or not recognized at all, it will label the face as “Unknown”.[57]:

flowchart :



*


Import necessary libraries (face_frcognition , imutils , pickle , time , cv2)
Initialize variables
Load precomputed face encondings from the pickle file
Espace key is pressed
Start capturing videos frames
Resize the captured frame to a specific width
Convert the frame to grayscale and RGB
Use haar cascade to detect faces in the grayscale images
Calculate face encodings for detected faces
Compare computed faces encodings with known encodings
draw rectangles around detected faces and display "Inconnu"
Update the recognized name pased on matches
draw rectangles around detected faces and display "name"



The code explication :

The code is a real time facial recognition system,it start with importing the necessary libraries and modules, imutils is used for basic image processing functions such as translation, rotation, resizing, skeletonization, and displaying Matplotlib images. Face recognition is used for face detection and recognition. cv2 is used for handling video frames.

Then the face encodings and Haar cascade files are loaded. The face encodings file contains a list of 128-d embeddings of known faces. The Haar cascade file is a trained classifier for detecting faces in an image, and A video stream is started with the default camera (camera 0). The FPS counter is also started.In the main loop, each frame from the video stream is read and resized to a width of 500 pixels. The frame is then converted to grayscale for face detection and to RGB for face recognition.The Haar cascade detector is used to detect faces in the grayscale image. This returns a list of bounding boxes around the detected faces , and For each detected face, the 128-d face encoding is computed. This encoding is then compared with the

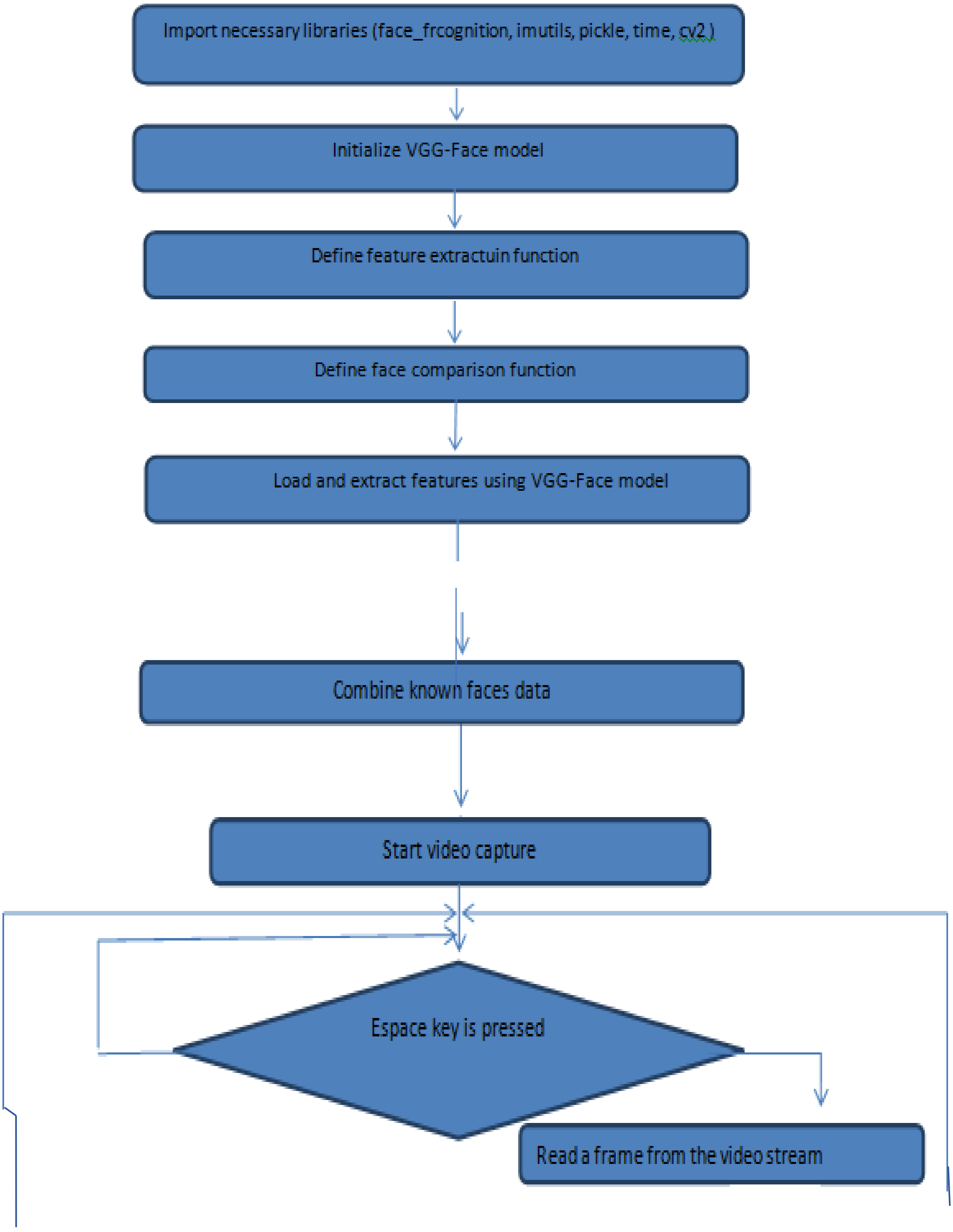
known face encodings. If there's a match, the name of the person is added to the list of names, or draw rectangles around detected faces and display "Unknown".For each detected face, a bounding box is drawn and the name of the person (if recognized) is displayed on the image and The processed frame is displayed on the screen. If the 'q' key is pressed, the loop breaks and the program ends , The FPS counter is updated after each frame and stopped after the loop ends in the end the OpenCV windows are destroyed and the video stream is stopped or the loop repeted .The mathematical aspect of this code lies in the face recognition part. The face_recognition.face_encodings function computes a 128-d vector that quantifies the face. This is done using a deep learning model which outputs these encodings (also known as embeddings). The face_recognition.compare_faces function then compares the computed face encoding with the known face encodings by calculating the Euclidean distance. If the distance is below a certain threshold, a match is found. The threshold is determined empirically and can be adjusted based on the requirements of the system. The name of the person with the most matches is then determined and displayed on the screen.

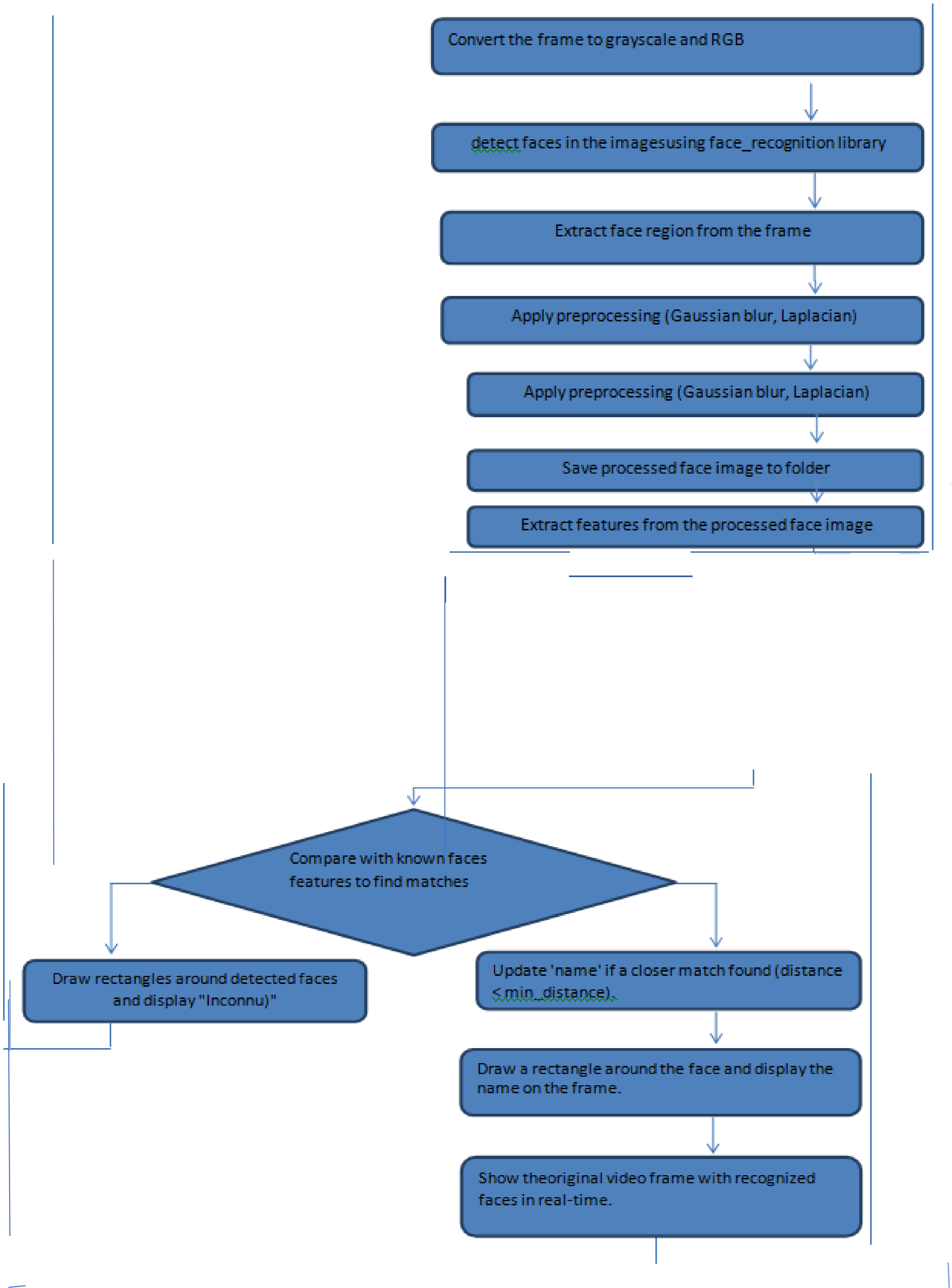
5.5 Facial recognition code withe ai

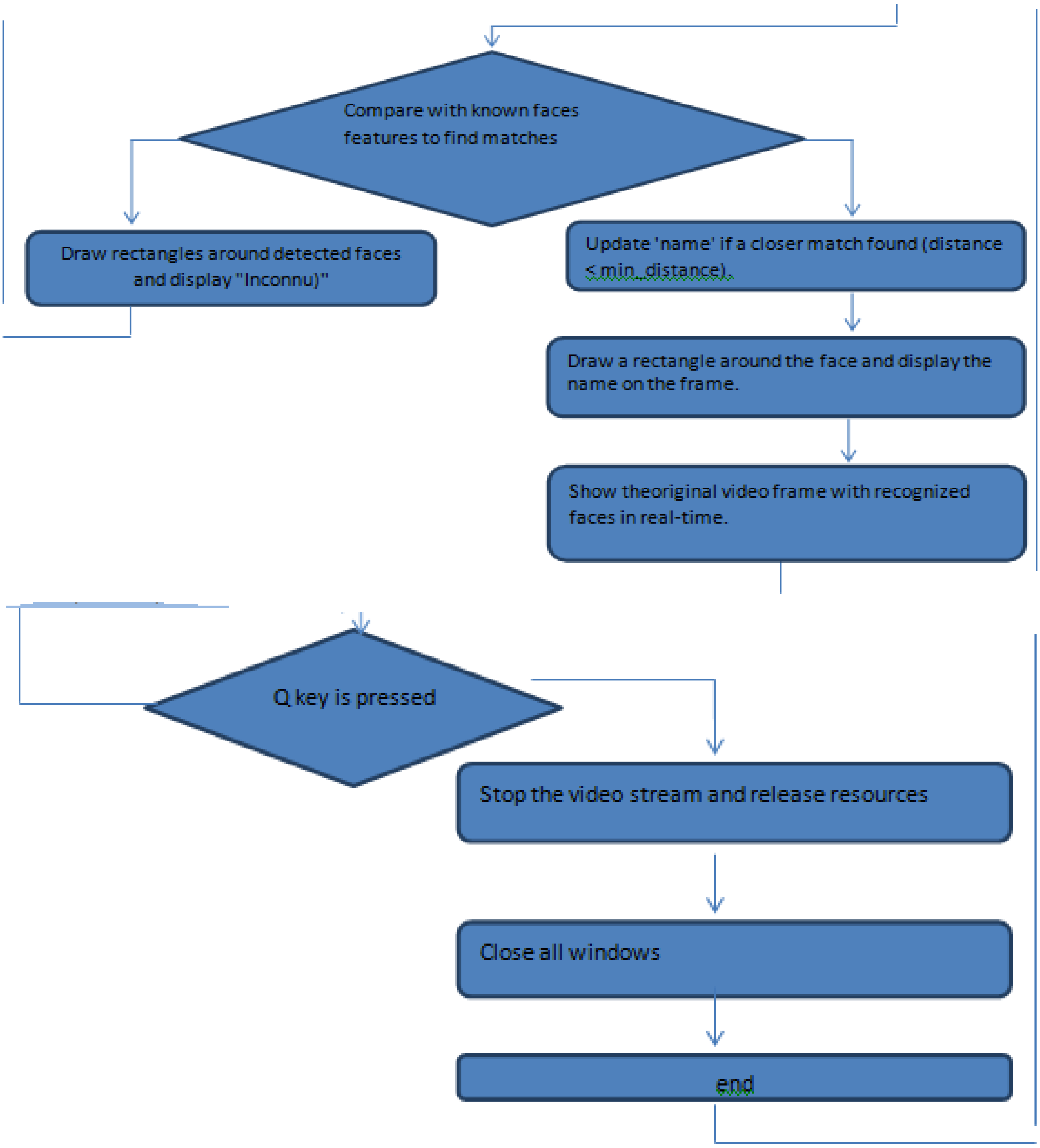
To enhance facial recognition under variable conditions, such as changes in expression, lighting, or the addition of accessories, we've utilized a more robust and advanced facial recognition model, such as Convolutional Neural Networks (CNNs) or Deep Learning architectures. These models are generally more adept at handling variations in facial images.This is an improved version of our code that uses the

`face_recognition` library, which is based on a deep learning model for better accuracy. We have created an encoding file (`encodings.pickle`) with known faces. Moreover, for enhanced performance under variable conditions, we collected facial images under different lighting conditions, with various accessories, and with diverse facial expressions when creating the encoding database.

Flowchart:


Import necessary libraries (face_frcognition, imutils, pickle, time, cv2)
Initialize VGG-Face model
Define feature extractuin function
Define face comparison function
Load and extract features using VGG-Face model
Combine known faces data
Start video capture
Espace key is pressed
Read a frame from the video stream


Convert the frame to grayscale and RGB
detect faces in the imagesusing face_recognition library
Extract face region from the frame
Apply preprocessing (Gaussian blur, Laplacian)
Apply preprocessing (Gaussian blur, Laplacian)
Save processed face image to folder
Extract features from the processed face image
Compare with known faces features to find matches
Draw rectangles around detected faces and display "Inconnu)"
Update 'name' if a closer match found (distance < min_distance).
Draw a rectangle around the face and display the name on the frame.
Show theoriginal video frame with recognized faces in real-time.



The code explication:

The code provided is a facial recognition script that uses mathematical relationships and algorithms to identify and label faces in video frames.

Librairie explications:

cv2 (OpenCV): A library for computer vision that provides tools for image and video analysis. It's used in ourcode for image manipulation and video processing tasks. [25]

Numpy: A package for scientific computing in Python. It offers support for large, multi-dimensional arrays and matrices, along with a collection of mathematical functions to operate on these arrays. [26]

Os: This module provides a portable way of using operating system-dependent functionality like reading orwriting to the file system. [27]

keras_vggface: An extension of Keras library specifically for face recognition tasks. The VGGFace model is a deep learning model pre-trained on a large dataset of faces to perform face recognition. [28]cipy.spatial.distance.cosine: Part of the SciPy library, this function computes the cosine distance between two non-zero vectors, which can be used as a similarity measure in face recognition. [29]

The face_recognition Python library : is a simple and effective tool for recognizing and manipulating faces from Python or from the command line. Built using dlib's state-of-the-art face recognition built with deep learning, the library uses a deep learning model to locate and recognize faces in images and videos.

Here's a brief overview of its capabilities:

Face Detection: Find faces in a photograph or folder full for photographs.

Face Recognition: Recognize who is in each photo.

Face Manipulation: Adjust facial attributes like smiling, opening/closing eyes, etc.

This library is widely used due to its ease of use and high accuracy in facial recognition tasks. [30]

VGGFace : The VGGFace refers to a series of models developed for face recognition and demonstrated on benchmark computer vision datasets by members of the Visual Geometry Group (VGG) at the University ofOxford. [31]

Face detection methods :

Our code uses the face_recognition library for face detection, which primarily employs methods that fall under the appearance-based category. Specifically, it uses:

Appearance-Based : The appearance-based technique, in order to discover face models, is dependent on a collection of delegate training face photos. The appearance-based approach outperforms all other methods of performance evaluation. When searching for relevant qualities in face photos, appearance-based methods depend on techniques from statistical analysis and machine learning to uncover important characteristics of face photos. This approach is also used in the extraction of facial features for the purpose of face recognition.

Following that, the appearance-based model is further subdivided into sub-methods for the purpose of facedetection, which are as follows:

- Eigenface-Based.
- Distribution-Based.
- Neural-Networks.
- Support Vector Machine (SVM).
- Sparse Network of Winnows.
- Naive Bayes Classifiers.
- Hidden Markov Model.
- Application of Information Theoretical Principles.
- Inductive Learning.[32]

HOG (Histogram of Oriented Gradients): Histogram of oriented gradients (HOG) is a feature descriptor like the Canny edge detector and scale invariant and feature transform (SIFT). It’s used in computer vision and image processing for the purpose of object detection.The technique counts occurrences of gradient orientation in the localized portion of an image.

His method is used by default in the face_recognition.face_locations function to detects faces based on the learned appearances from training images. [33]

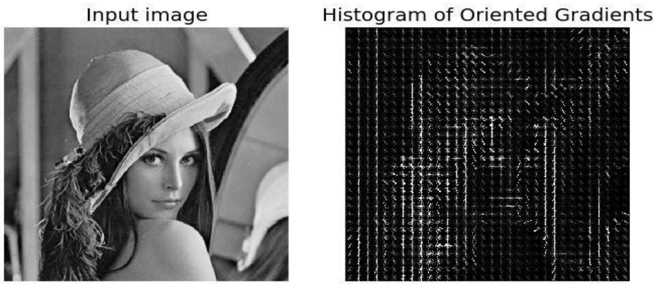


Figure 3.10 :HOG (Histogram of Oriented Gradients)

CNN (Convolutional Neural Network): A convolutional neural network (CNN or ConvNet) is a network architecture for deep learning that learns directly from data.CNNs are particularly useful for finding patterns in images to recognize objects, classes, and categories. They can also be quite effective for classifying audio, time-series, and signal data.[34]

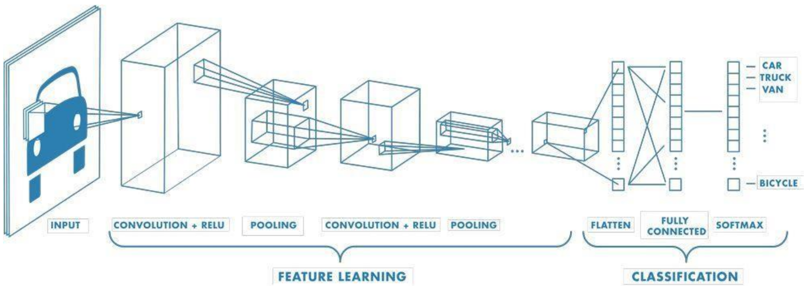


Figure 3.11: Example of a network with many convolutional layers. Filters are applied to each training image at different resolutions, and the output of each convolved image is used as the input to the nextlayer.

The face_recognition library specifically uses appearance-based methods learned from training data to detect faces in images. These methods are robust against variations in lighting, facial expressions, and orientations, making them effective for real-time applications.

Mathématiques explications:

Here's an explanation of the mathematical relationships present in the code:

Feature Extraction:

The image resizing step:

When we have a raw image (I), we need to adjust its size so that it matches the input size expected by the VGGFace model. The standard input size for many convolutional neural networks, including VGGFace, is (224 × 224) pixels.

The resizing process can be represented by: [I′=resize(I,224,224)]

(I) is theoriginal raw image.

Esize (I, 224, 224) is the function that resizes the image III to the dimensions 224 pixels by 224pixels.

(I’) is the new image after resizing.

So, this formula simply states that we are transforming our original image (I) into a new image (I’) that has a width and height of 224 pixels each. This step ensures that the image is in the correct format for theVGGFace model to process. [35]

The preprocessing step:

The preprocessing step is crucial for preparing the image to be in the exact format that the model expects.This often involves several sub-steps such as scaling pixel values to a certain range, normalizing the image data, and sometimes applying data augmentation techniques.

For example, if we’re using a VGGFace model, this function might scale pixel values from [0, 255] to a range that the network can work with more effectively, like [0, 1] or [-1, 1], and apply other necessary transformations. [16]

The formula you’ve mentioned:

- [I″=(I′,version=2)(I′=resize(I,(224,224)))]
- (I’) is the resized image from the previous step.

- (I′, version=2) specifies which version of preprocessing to apply, which can vary based on themodel's requirements.
- (I'') is the final preprocessed image ready to be fed into the model.

The VGGFace model : The VGGFace model is a type of convolutional neural network (CNN) that is specifically designed for face recognition tasks. It works by taking an input image and passing it through a series of layers that extract increasingly complex features from the image. These features are then used to represent the image in a way that can be used for recognition tasks[37]

- The formula :v = f(I'')
- (I'') is the preprocessed image ready to be fed into the model.
- (f) represents the VGGFace model function.
- (v) is the output feature vector that represents the high-level features extracted from the image bythe model

Convolutional Neural Network (CNN) :

Convolutional Neural Network (CNN) the convolutional layers are responsible for detecting patternsand features in the input image. They do this by sliding

filters, also known as kernels, across the image, which results in feature maps. These feature mapsrepresent different aspects of the image, such as edges, textures, or colors.

Pooling layers follow the convolutional layers and serve to reduce the spatial size of these feature maps. This reduction is done to decrease the number of parameters and computation in the network, which helps to control overfitting and also reduces the computational cost. Pooling layers summarize the features extracted by convolutional layers and retain only the most essential information. [38]

The function [f: \mathbb{R}^{224 \times 224 \times 3} \rightarrow \mathbb{R}^d]

represents how a CNN transforms an input image (of size 224x224 pixels with 3 color channels) into a high- dimensional feature vector ((\mathbb{R}^d)), where (d) is the dimensionality of the output feature vector. This vector captures the essential information needed for tasks such as image classification or recognition.

Cosine Similarity:Cosine similarity is a measure used to calculate the similarity between two non-zerovectors by measuring the cosine of the angle between them.

This metric effectively captures the orientation (or direction) of the vectors and not their magnitude, which makes it particularly useful in many applications such as text analysis,

$$(A, B) = \frac{A \cdot B}{|A||B|} = \frac{\sum_{i=1}^{d} A_i B_i}{\sqrt{\sum_{i=1}^{d} A_i^2}\sqrt{\sum_{i=1}^{d} B_i^2}}$$

where the length of the document (or vector magnitude) is less important than the direction in which the vector points (whichindicates the document's content). [39]The formula :

This formula calculates the cosine of the angle between two vectors (A) and (B) in a multi-dimensional space. The dot product (A, B) represents the sum of the products of their corresponding components, and (|A|) and (|B|) are the magnitudes (or lengths) of vectors (A) and (B), respectively.

- If the cosine similarity is close to 1, it means that the angle between (A) and (B) is small, and theyare considered similar.
- If it is close to 0, it means that they are orthogonal (or very dissimilar).
- If it is negative, it means that they point in largely opposite directions.

Cosine Distance:

Cosine distance is a metric derived from cosine similarity, used to quantify the dissimilarity between twovectors. [40]

It is defined as:

cosine distance (A,B) = 1−cosine similarity (A,B).

This formula essentially subtracts the cosine similarity from 1 to provide a measure of dissimilarity. Here'swhat the values indicate:

A cosine distance close to 0 means that the vectors are very similar (since their cosine similarity isclose to 1).

A cosine distance close to 1 means that the vectors are very dissimilar (since their cosine similarityis close to 0).

Thresholding for Recognition:A threshold value (t) is set to determine if a face is recognized: if (

\text{cosine distance}(A, B) < t), then face (A) is recognized as face (B).

In the context face recognition, a threshold value is used to determine whether two faces are consideredthe same based on their cosine distance.

Here's how it works:

- A feature vector is extracted from each face, representing it in a multi-dimensional space.
- The cosine distance between the feature vectors of two faces (A and B) is calculated.
- A threshold value (t) is set, which serves as a cut-off point for deciding if the faces are recognized asthe same.
- If the cosine distance between face A and face B is less than the threshold value (t), then face A is recognized as face B. This means that the smaller the cosine distance, the more similar the faces are, and ifit's below the threshold, they are considered a match.[40]
- These mathematical formulations are at the core of our facial recognition system They enable the transformation of image data into feature vectors and provide a quantitative measure for comparing thesevectors to identify faces.

The role of artificial intelligence (AI) in this code:

The role of artificial intelligence (AI) in this code is crucial for achieving accurate and efficient facialrecognition. AI techniques are employed at various stages, from detecting faces in video frames to recognizing and identifying these faces based on their unique features. Here's a detailed breakdown of howAI is utilized:

- Facial Detection:

The code uses a face detection library to locate faces within the captured video frames. This library leverages AI models, such as convolutional neural networks (CNNs), which are trained on large datasets to detect faces with high accuracy.

The face detection algorithm scans each video frame to identify regions that contain faces. It returns the coordinates (bounding boxes) of these regions, which are then used to isolate and process the faces.

- Feature Extraction with Deep Learning:

The code employs the VGG-Face model, a deep learning model based on the ResNet-50 architecture, to extract features from the detected faces. This model is pre-trained on a large dataset of faces, enabling it to generate high-dimensional feature vectors that represent unique facial characteristics. The detected face images are resized and preprocessed to fit the input requirements of the VGG-Face model. The model processes these images through multiple convolutional layers to produce feature vectors that encapsulate the distinctive features of each face.

- Face Recognition:

The code compares the extracted features of the detected faces with those of known faces using a similarity metric, specifically cosine distance. This AI technique helps in recognizing and identifying faces based on the degree of similarity between feature vectors.

The cosine distance between the feature vector of the detected face and the feature vectors of known faces is calculated. The face is identified based on the closest match. If no close match is found, the face is labeled as "Unknown."

By integrating these AI techniques, the code can accurately detect faces in real-time, extract meaningful features, and recognize faces even in varying conditions, This advanced use of AI makes the system robust and reliable for applications in security, authentication, and personalized services.

Code testing

6.1 Test N°1

In the initial test, we will compare our face recognition code results with those obtained from Shamla

Mantri and Kalpana Bapat's code, using the same database. This analysis aims to assess the performance ofour face recognition system.We use AT&T face dataset ,tha contains a set of face images taken between April 1992 and April 1994 at the lab. The database was used in the context of a face recognition project carried out in collaboration with the Speech, Vision and Robotics Group of the Cambridge University Engineering Department..[56]GTDLBench (git-disl.github.io).The original files are in PGM format, and can conveniently be viewed on UNIX (TM) systems using the ‘xv’ program. The size of each image is 92x112 pixels, with 256 grey levels per pixel. The images are organised in 40 directories (one for each subject), which have names of the form sX, where X indicates the subject number (between 1 and 40). In each of these directories, there are ten different images of that subject, which have names of the form Y.pgm, where Y is the image number for that subject (between 1 and 10).

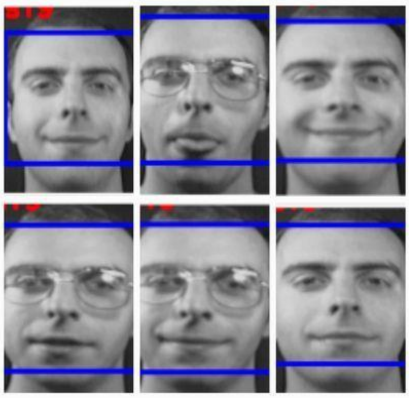
There are 10 different images of each of 40 distinct subjects. For some subjects, the images were taken at different times, varying the lighting, facial expressions (open / closed eyes, smiling / not smiling) and facial details (glasses / no glasses). All the images were taken against a dark homogeneous background with the subjects in an upright, frontal position (with tolerance for some side movement). A preview image of the Database of Faces is available.

From each database, we use 100% images for training and 100% for testing.

Table 3.1: Face recognition results for AT&T Face dataset

N°	Personne	Recognitio nrate (%)	Error rate (%)	N°	Personne	Recognitio nrate (%)	Error rate (%)
1	S1	100	0	21	S21	100	0
2	S2	90	10	22	S22	100	0
3	S3	90	10	23	S23	100	0

4	S4	100	0	24	S24	90	10
5	S5	80	20	25	S25	100	0
6	S6	100	0	26	S26	100	0
7	S7	100	0	27	S27	100	0
8	S8	100	0	28	S28	100	0
9	S9	90	10	29	S29	100	0
10	S10	90	10	30	S30	90	10
11	S11	100	0	31	S31	100	0
12	S12	100	0	32	S32	100	0
13	S13	80	20	33	S33	100	0
14	S14	90	10	34	S34	70	30
15	S15	100	0	35	S35	100	0
16	S16	100	0	36	S36	100	0
17	S17	100	0	37	S37	80	20
18	S18	100	0	38	S38	100	0
19	S19	90	10	39	S39	100	0
20	S20	100	0	40	S40	100	0



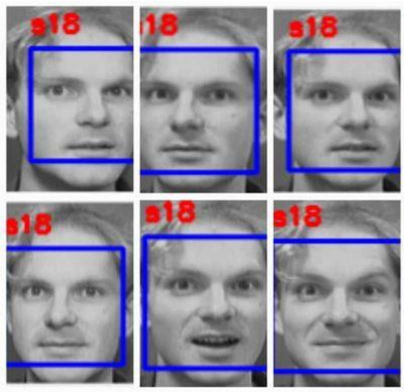
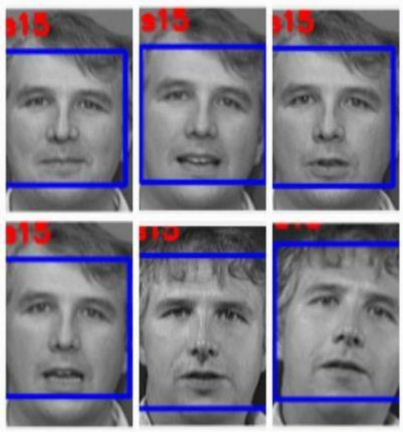


Figure 3.15: s19 facerecognitionFigure 3.16: s18 facerecognitionFigure 3.17: s17 facerecognition



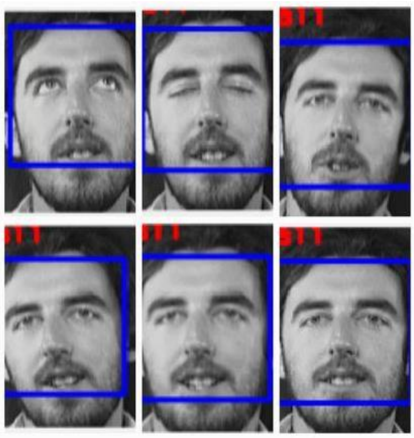


Figure 3.18: s14 facerecognitionFigure 3.19: s15 facerecognitionFigure 3.20:s11 facerecognition

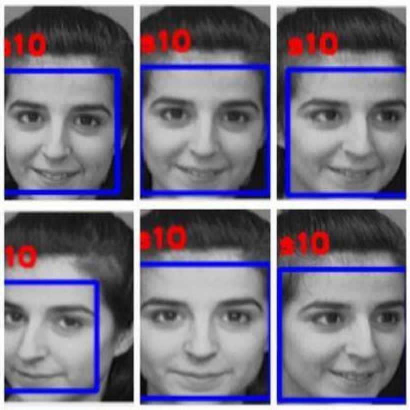


Figure 3.21: s10 facerecognition

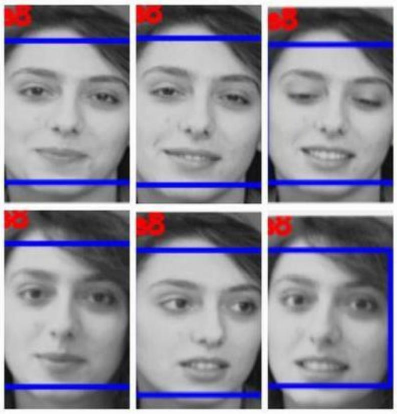


Figure 3.22: s8 facerecognition

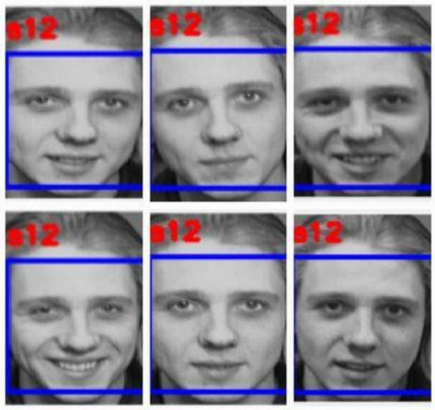


Figure 3.23: s12 facerecognition

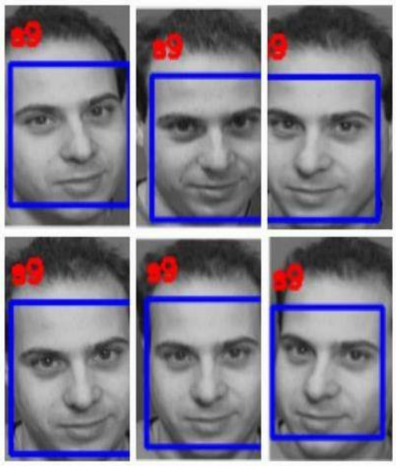


Figure 3.24: s12 facerecognition

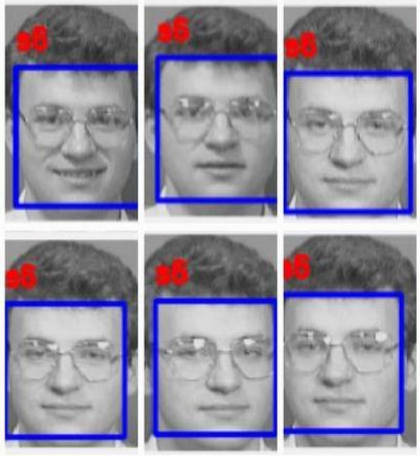


Figure 3.25: s6 facerecognition

Figure 3.26: s7 facerecognition

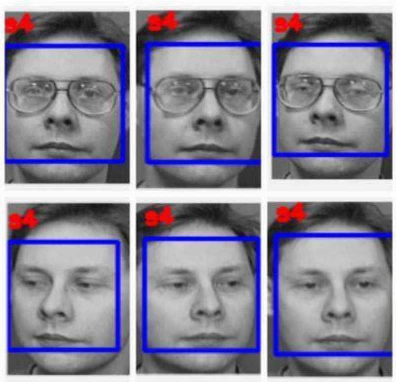


Figure 3.27: s4 facerecognition

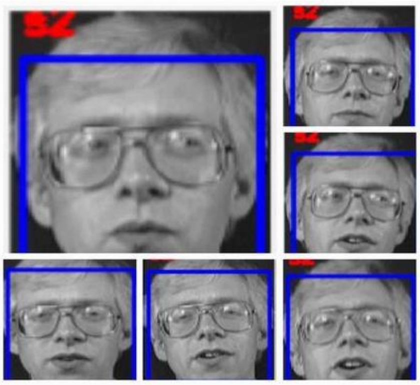


Figure 3.28: s2facerecognition

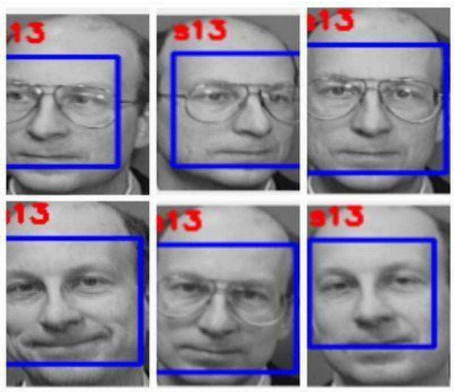


Figure 3.29: s13 facerecognition

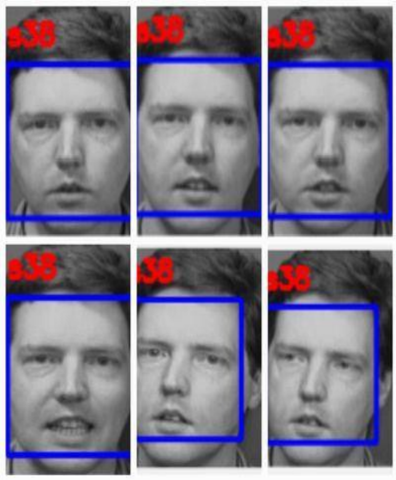


Figure 3.30: s38 facerecognition

Figure 3.31: s37 facerecognition

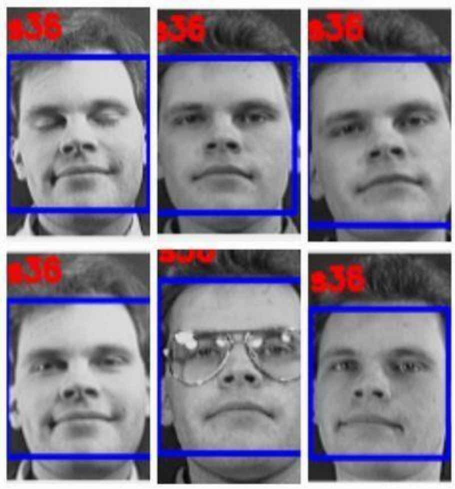


Figure 3.32: s36 facerecognition

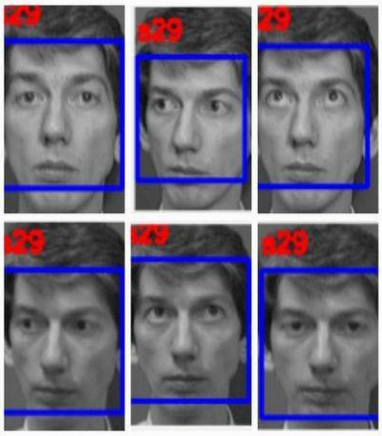


Figure 3.33: s29 facerecognition

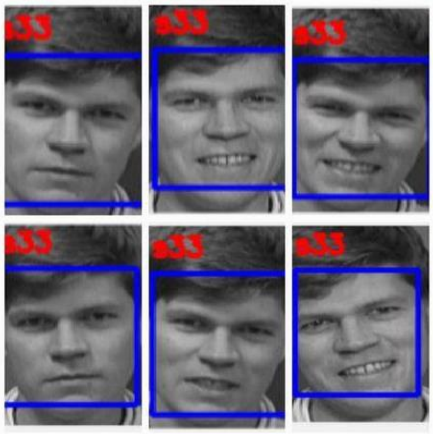


Figure 3.34: s33 facerecognition

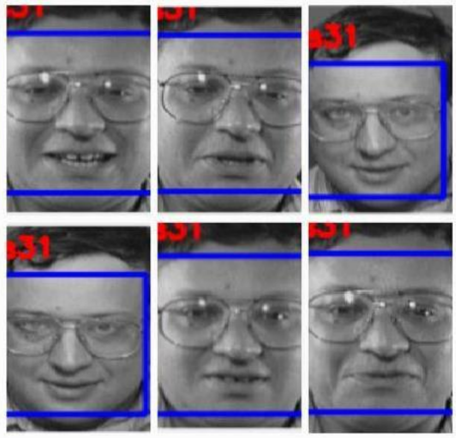


Figure 3.35: s31 facerecognition

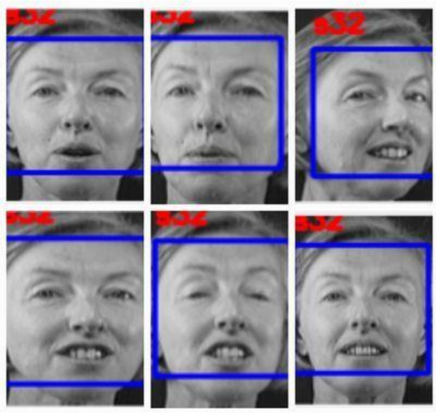


Figure 3.36: s32 facerecognition

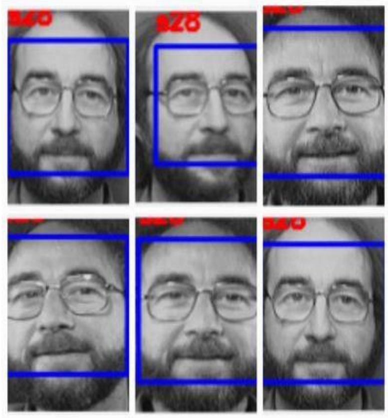


Figure 3.37: s28 facerecognition

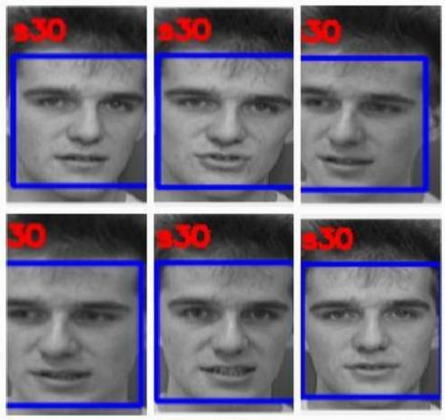


Figure 3.38: s30 facerecognition

Figure 3.39: s27facerecognition

Figure 3.40: s40 facerecognition

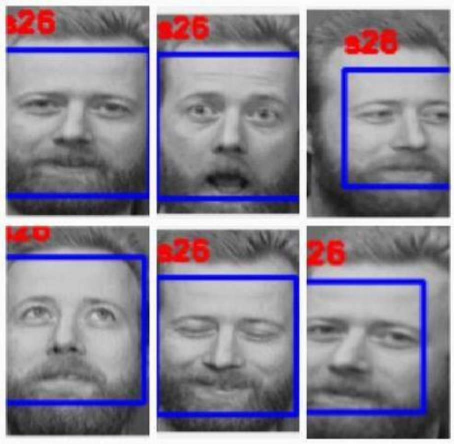


Figure 3.41: s28 facerecognition

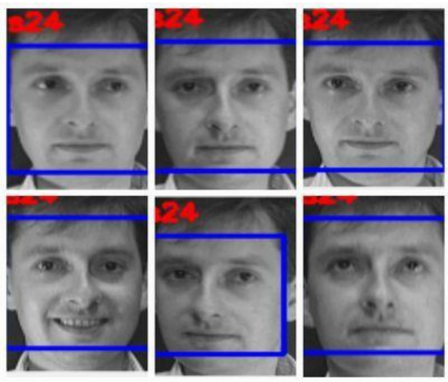


Figure 3.42: s24 facerecognition

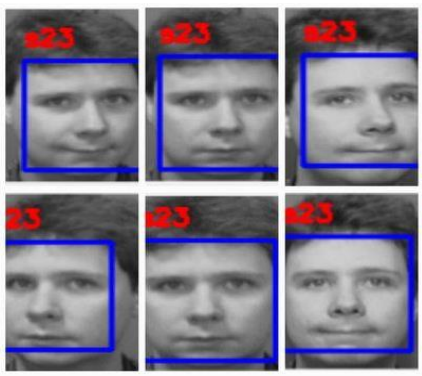
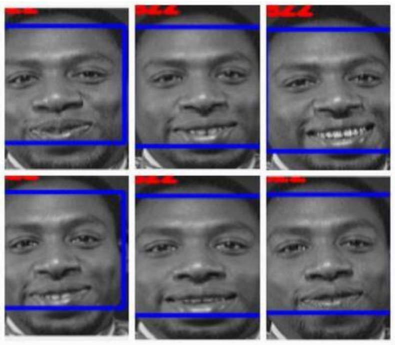
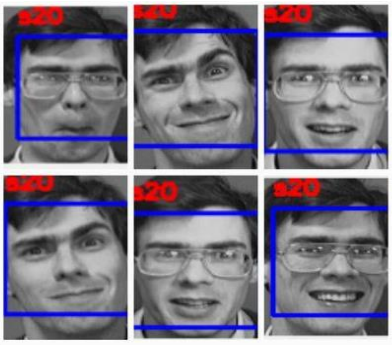


Figure 3.43: s23 facerecognition

Figure 3.44: s25 facerecognition





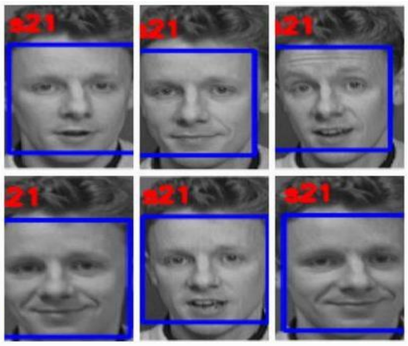
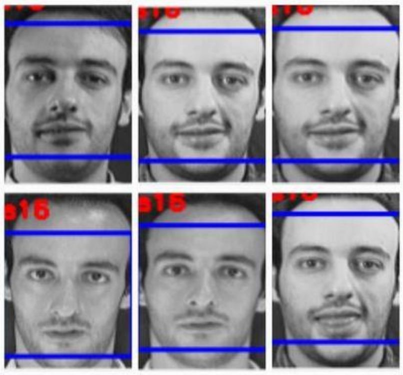
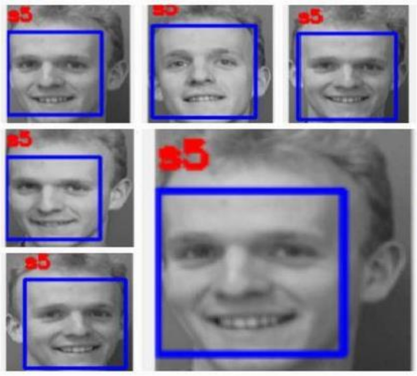


Figure 3.45: s22facerecognition **Figure 3.46: s20facerecognition** **Figure 3.47: s21 facerecognition**





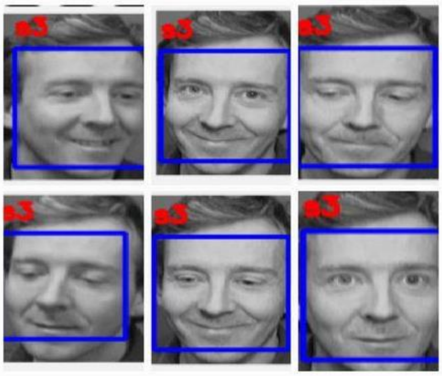
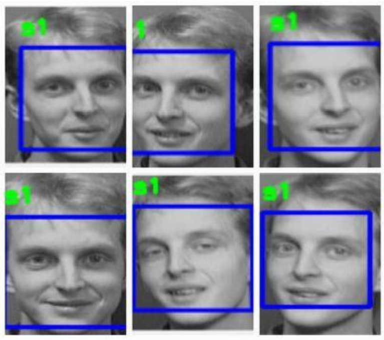


Figure 3.48: s16 facerecognition **Figure 3.49: s5 facerecognition** **Figure 3.50: s3 facerecognition**



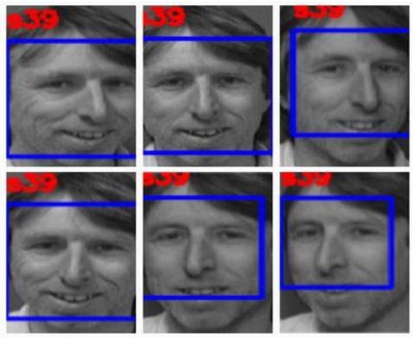


Figure 3.51: s1 facerecognition **Figure 3.52: s39 facerecognition**

Tableau 3.2: Test N°1 résults

Test	Recognition rate (%)	Error rate (%)
Available inthe database	96.5	3.5

Using my code with VGG-Face model andAT&T face dataset I could make 96.5 (%)Recognition rate.And with using the sameatabase mr finde that The highest average recognition nate achieved usung single algorithem som is of 92.40%.[55], obtained for 40 persons 400 images of AT&T database by using SOMface recognition system.

The Self-Organizing Map (SOM), also known as the Kohonen map, is a well-known type of artificial neural network. It operates in unsupervised learning mode, creating a low-dimensional representation (usually two-dimensional) of a higher-dimensional data set while preserving its topological structure. Neurons in the SOM compete to be activated, and only one neuron (the "winner") is activated, thus representing a specific group.

The Kohonen ruleallows the neuron weights to learn from input vectors, making them useful for recognition tasks. In a system like the one described, a SOM is used to classify vectors based on Sobel, DWT, and DCT features to determine whether the subject in an input image is "present" or "not present" inthe image database.[41]

The VGG-Face model, built upon the VGG architecture, is a deep convolutional neural network designed specifically for face recognition tasks. It operates by processing input face images through a series of convolutional and pooling layers. These layers progressively extract hierarchical features from the input, starting with basic features like edges and textures, and advancing to more complex facial structures andexpressions.

The network's architecture includes multiple convolutional layers followed by fully connected layers, which combine the extracted features to make final predictions related to facial identity. (Parkhi, 2015)

The VGG-Face model generally achieves higher recognition accuracy compared to SOMface, primarily because of its deep learning architecture. This capability is crucial for achieving high accuracy, which is particularly important in large-scale applications such as surveillance and authentication.

However, the VGG-Face is more complex and resource-intensive to train due to its deeper architecture andthe larger number of parameters involved.

On the other hand, SOMface may be more suitable for simpler scenarios, such as smaller datasets or resource-constrained environments. Its simplicity and lower computational requirements make it more feasible in such settings, where efficiency and simplicity outweigh the need for very high accuracy.

6.2 Test N°2

In the second test, we will explore variations in appearance, including makeup, accessories, hairstyles, facial expressions, and the effects of aging and scars. We download a database of face photographs designed for studying the problem of unconstrained face recognition. The dataset contains more than 13,000 images of faces collected from the web. It includes folders of images for each individual in the dataset, with 100 images for each available sample. For example, George W. Bush: 100 images. [42]From each database, we use 70% images for training and 30% for testing. Additionally, 50 faces photographs not available in the database are used for testing unknown individuals.

Figure 3.53: database of face photographs.

Tableau 3.3: Results obtained from the databases

N°	Personn	Recognition rate (%)	Error rate (%)
1	Elizabeth Olsen	93.3	6.7
2	robert downey jr	90	10
3	natalie portmane	86.7	13.3
4	claire holt	90	10
5	margrot robbie	80	20
6	lisa kudrow	96.7	3.7
7	jessica alba	86.7	13.3
8	hugh jackman	93.3	6.7

9	courtney cox	80	20
10	charlize theron	96.7	3.7
11	zac efron	96.7	3.7
12	tom cruise	93.3	6.7
13	brad pitt	73.3	26.7
14	billie eilish	100	0
15	andy samberg	96.7	3.7
16	amiska sharma	70	30
17	amatch bachchan	90	10
18	alia bhatt	83.3	26.7
19	akshay kumar	93.3	6.7
20	alexander daddario	96.7	3.7
	Results obtained from the 20 databases	89.3	10.7

Figure 3.54:Elizabeth Olsenfacial recognition,Figure 3.55: amiska sharmafacial recognition,Figure 3.56: tom cruisefacial recognition

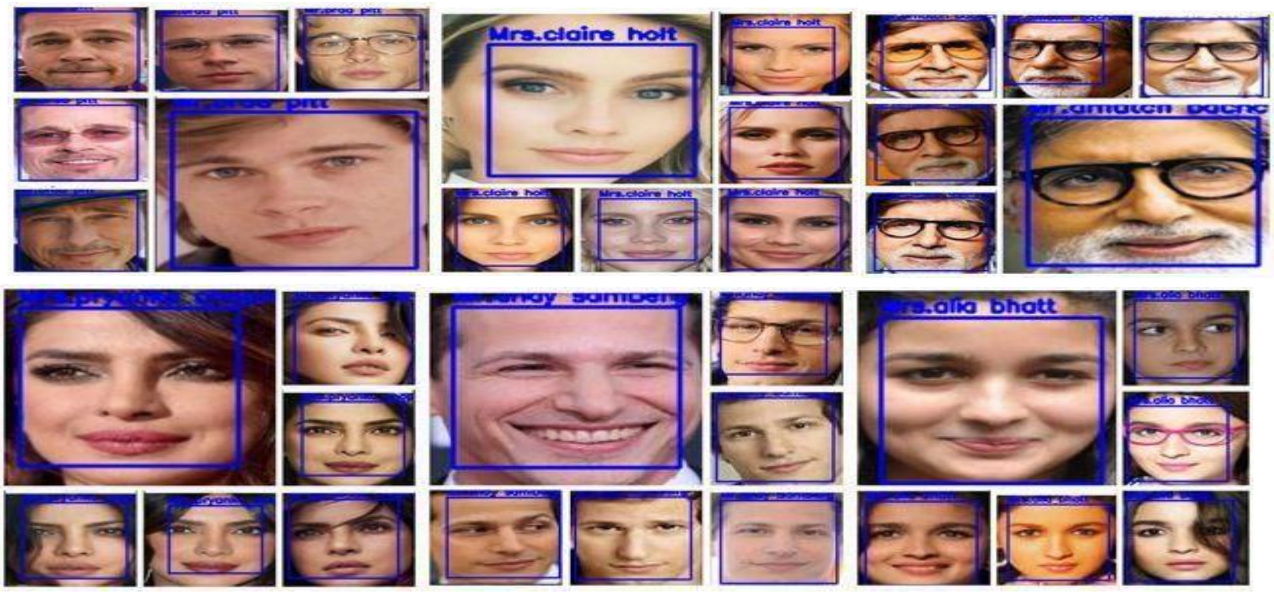


Figure 3.57: brad pittfacial recognition, Figure 3.58: claire holtfacial recognition, Figure 3.59: amatchbachchanfacial recognition

Figure 3.60: pryanka choprafacial recognition,Figure 3.61: andy sambergfacial recognition,Figure 3.62: aliabhattfacial recognition.



Figure 3.63: courtney coxfacial recognition,Figure 3.64: hugh jackmanfacial recognition, Figure 3.65: zacefronfacial recognition.


Figure 3.66: Lisa Kudrowfacial recognition, Figure 3.67: robert downey jrfacial recognition,Figure 3.68: alexander


daddariofacial recognition.

Figure 3.69: charlize theronfacial recognition,Figure 3.70: margrot robbiefacial recognition , Figure 3.71: akshay kumarfacial recognition


Figure 3.72: jessica albafacial recognition, Figure 23.73: natalie portmanefacial recognition*[2]*

Tableau 3.4: Available and not available in the database testing.

Test	Recognition rate (%)	Error rate (%)
Available in the database	89.3	10.7
Not Available in the database	84	16

Figure 3.74: personnes Available in the database facial recognition.

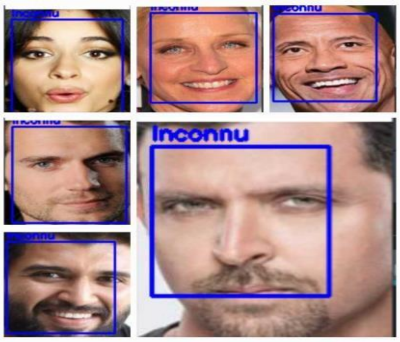


Figure 3.75: unknownfacial recognition.

- Confusion Matrix: This is a table used to describe the performance of a classification model. It contains four elements: True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN).[43]

In facial recognition systems, here's how we can classify the photos:

- True Positive (TP): The system correctly identifies a face from the database.
- True Negative (TN): The system correctly identifies that a face is not in the database.
- False Positive (FP): The system incorrectly identifies an unknown face as someone in the database.
- False Negative (FN): The system fails to identify a face that is in the database.

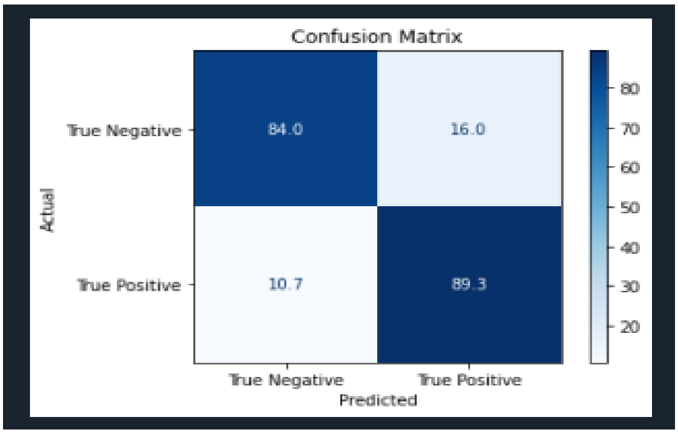


Figure 3.76: test N°2 confusion matrix

Tableau 3.5: Confusion Matrix

Photosclassify	Recognition rate (%)
True positive	89.3
false positive	16
true negative	84
false negative	10.7

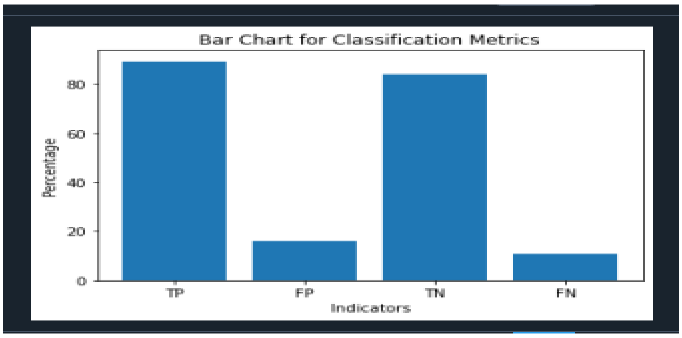


Figure 3.77: TEST N°2 bar chart for classification metrics.

Figure 3.78: True positive.

Figure 3.79: false positive.

Figure 3.80: False Negative.

Figure 3.81: True Negative.

- Precision: This measures the accuracy of the positive predictions. It's calculated as:[43]

Precision =

T P TP+ FP

- Recall (also known as Sensitivity or True Positive Rate): This measures the ability of the model to find all the relevant cases (all actual positives). It's calculated as:

$$Recall = \frac{TP}{TP+FN}$$

- ... core ... s, j the harmonic mean of precision and recall and gives a balance between them. It's calculated as:

$$F1\backslash Score = 2\times \frac{Precision \times Recall}{Precision + Recall}$$

Tableau 3.6: Metrics Summary

metrics	Results (%)
Recall	0.89
F1\ Score	0.87
Precision	0.85

The second test reveals thart our models capable of handling variation in appearance with reasonable accuracy, especially with known individuals. However, performance drops when encountering unfamiliar faces, suggesting the need for further training and optimization to enhance its robustness and reduce error rates. The true positive rate of 89.3% and true negative rate of 84% suggest that the model is generally effective in correctly identifying both known and unknown individuals. The model's high recall and balanced F1 score indicate strong potential for reliable face recognition, but continuous refinement is essential to improve precision and overall effectivenes.

6.3 Test N°3

In the third test, we will focus on addressing inadequate lighting, variations in angle, and the distance between the camera and the individuals to achieve better face recognition results. To date, we have downloaded 50 images from the web for each person, covering different sizes and face angle variations. From each database, we allocate 70% of the images for training and reserve 30% for testing. Additionally, we include 50 face photographs not available in the database to test recognition for unknown individuals.

Figure 3.82: Test N°3 data basse.

Tableau 3.7: Results obtained from the databases

N°	Personn	Recognition rate (%)	Error rate (%)
1	mr robert de niro	86.7	13.3
2	mrs Lisa Kudrow	93.3	6.7
3	mr brad pitt	100	0
4	mrs emilia clark	100	0
5	mr Jake Gyllenhaal	93.3	6.7

6	mr will smith	100	0
7	mrs angelina jolie	93.3	6.7
8	mr jackie chan	73.3	26.7
9	mr tom hardy	93.3	6.7
10	mr Chris Evans	100	0
	Results obtained from the10 databases	93.3	6.7

Figure 3.83: mr Jake Gyllenhaal

Figure 3.84: mr robert de niro

Figure 3.85: mr brad pitt

Figure 3.86: mr Chris Evans

Figure 3.87: mr tom hardy

Figure 3.88: mrs Lisa Kudrow

Figure 3.89: mr will smith

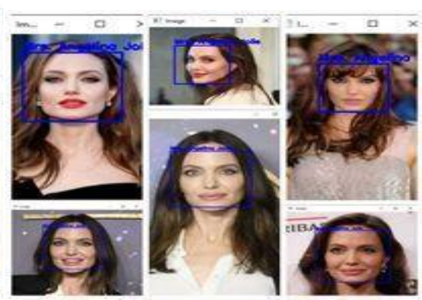


Figure 3.90: mrs angelina jolie

Figure 3.91: mr jackie chan

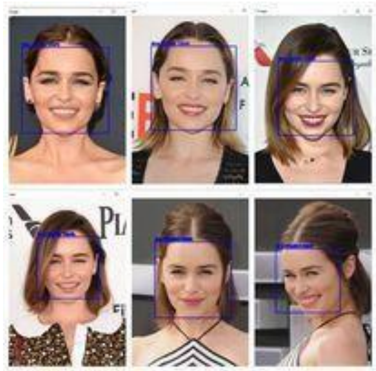


Figure 3.92: Mrs emilia clark.

Tableau 3.8: Test N°3 available and not available in the database testing

Test	Recognition rate (%)	Error rate (%)
Available in the database	93.3	6.7
Not Available in the database	94	6

- Confusion Matrix:

Tableau 3.9: Test N°3 Confusion Matrix

Photosclassify	Recognition rate (%)
True positive	93.3
false positive	6
true negative	94
false negative	6.7

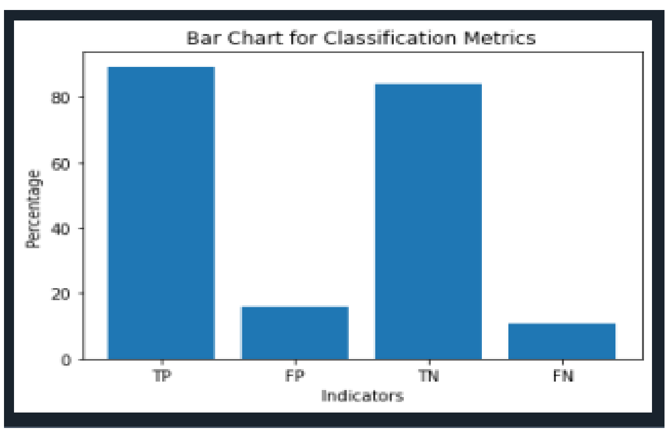


Figure 3.93: test N°3 confusion matrix.

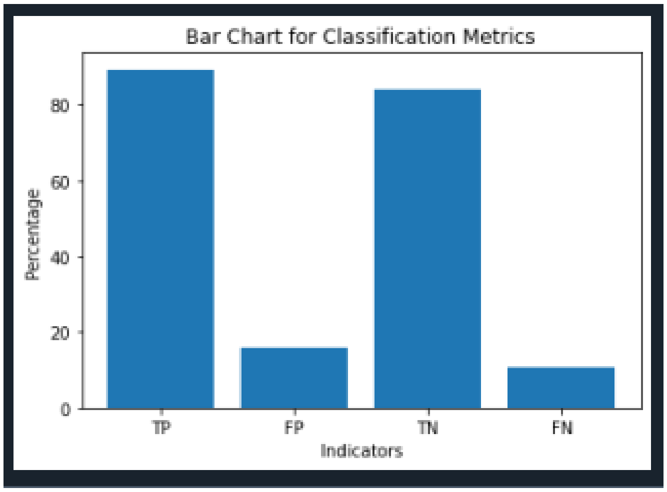


Figure 3.94: test N°3 bar chart classification metrics

Tableau 3.10: **Test n°3 metrics summary**

Metrics	Results
Precision	**0.94**
Recall	**0.93**
F1\ Score	**0.94**

Our model's metrics indicate both high accuracy and reliability. Here's a detailed evaluation of our resultsand a comparison with standard face recognition benchmarks:

- Precision: Our model achieved a precision of 0.94, meaning 94% of the faces identified as known individuals are correct. This demonstrates a low rate of false positives, which is crucial for applications where minimizing false identifications is vital. For comparison, leading face recognition systems, such as those used in security and identification, typically achieve precision rates between 0.90 and 0.95 .[49]
- Recall : With a recall of 0.93, our model correctly identifies 93% of the faces it should recognize. This lowrate of false negatives ensures most individuals are accurately identified. Industry standards for recall in face recognition models generally range from 0.85 to 0.95 , indicating that our model performs competitively.[50]

- F1 Score: Our model's F1 score of 0.94 reflects an excellent balance between precision and recall. The F1 score is a critical measure in face recognition tasks as it combines both precision and recall into a single metric. Leading face recognition models usually report F1 scores between 0.88 and 0.95 , placing our model at the higher end of this spectrum.[51]

our model's performance metrics precision, recall, and F1 score are all within or above the typical range for stateof the art face recognition systems. This suggests our model is highly effective and reliable for accurate identification.

7. How can we improve the program

To improve face id algorithm In addition to the libraries thatwe used there are several other deeplearning libraries that are popular and powerful for various applications, including facial recognition.Here are somethat you might consider:

- **Keras :** is the high-level API of the TensorFlow platform, written in Python, and capable of running on TensorFlow, CNTK, or Theano. It provides an approachable and highly productive interface for solving machine learning problems, with a particular focus on modern deep learning. Keras covers every step of the machine learning workflow, from data processing to hyperparameter tuning to deployment. This high-level neural networks API was developed with a focus on enabling fast experimentation, both with deep neural networks and other machine learning models. Its design allows for rapid prototyping and iteration, making it an excellent choice for both research and production environments. [44]
- **Caffe:** Caffe is a deep learning framework developed with cleanliness, readability, and speed in mind. It was created by Yangqing Jia during his PhD at UC Berkeley, and is in active development by the Berkeley Vision and Learning Center (BVLC) and by community contributors. [45]
- **Theano:** Theano is a Python library that allows us to evaluate mathematical operations including multi- dimensional arrays efficiently. It is mostly used in building Deep Learning

Projects. Theano works wayfaster on the Graphics Processing Unit (GPU) rather than on the CPU. [46]

Chainer :

supports various network architectures , including feed-forward nets, convnets, recurrent nets and recursive nets. It also supports per-batch architectures .its computation can include any control flow statements of Python without lacking the ability for backpropagation. This makes the code intuitive and easy to debug , allowing for the efficient evaluation of mathematical operations , including multi- dimensional arrays . It is mostly used in building deep Learning Projects. [47]

MXNet: MXNet is a multi-language machine learning (ML) library to ease the development of ML algorithms, especially for deep neural networks. Embedded in the host language, it blends declarative symbolic expression with imperative tensor computation. It offers auto differentiation to derive gradients. MXNet is computation and memory efficient and runs on various heterogeneous systems, ranging from mobile devices to distributed GPU clusters. [48]

8. Conclusion

In this project, we developed a powerful face recognition system using the Python programming language. Our goal was to achieve high accuracy in recognizing human faces. We tested it on a large dataset of over 1500 individuals, with positive results reaching up to 93.1%.

The system performed well despite variations in shooting angles, lighting conditions, different age groups, haircuts, accessories, and the presence or absence of makeup. This system heavily relies on artificial intelligence.

The role of artificial intelligence (AI) in this code is to use machine learning models (like VGGFace) to analyze and understand content within images, specifically for the purpose of recognizing human faces. The AI model has been trained on a vast amount of data to identify patterns that correspond to facial features, which it then uses to compare and recognize individual faces.

For comparison, the National Institute of Standards and Technology (NIST) reports that top facial recognition algorithms can achieve accuracy levels as high as 99.97% under ideal conditions[54].

However, real-world factors such as lighting, angles, and obstructions can significantly impact performance. Despite these challenges, our model's performance is comparable to these leading systems, demonstrating robust accuracy and reliability in diverse conditions.

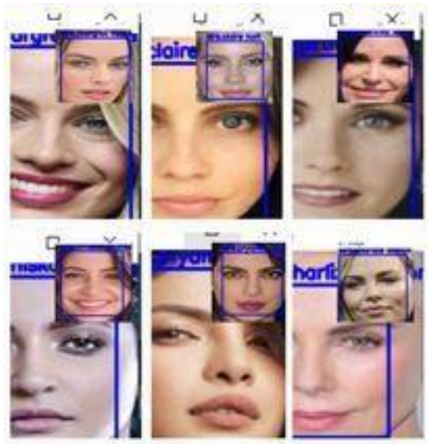


Figure 3.95: face recognition With and Without Makeup.

Figure 3.96: face detection and recognition From Different Angles.

Figure 3.97: face recognition Across Different Age Stages.

Figure 3.98: face recognition withe Partial occlusions.

Figure 3.99: face recognition withe Various Facial Expressions.

Figure 3.100: face recognition for different human racial groups.

GENERAL CONCLUSION

In this thesis, we developed a robust facial recognition system using artificial intelligence and Python programming. Our primary goal was to achieve highly accurate face recognition under diverse conditions. Through extensive testing on a dataset of over 1500 individuals, our system achieved impressive success rates, peaking at 93.1%. This underscores its effectiveness across varied angles, lighting, ages, hairstyles, accessories, and makeup styles.

Artificial intelligence, particularly machine learning models like VGGFace, played a crucial role. These models analyzed facial features, enabling precise face comparison and recognition. Trained on extensive datasets, our AI model learned intricate facial patterns, enhancing its accuracy in identifying individuals.

In conclusion, this thesis contributes significantly to facial recognition technology and highlights artificial intelligence's transformative role in advancing biometric systems. Our findings pave the way for further innovations in security, personalized user experiences, and other critical applications where precise facialrecognition is indispensable.

Bibliography

(Ziyad, 2019)Ziyad, M, (2019) , Artificial Intelligence Definition, Ethics and Standards , the british university in egypt .(PDF) Artificial Intelligence Definition, Ethics and Standards (researchgate.net)

(Clément, 2021) Clément Douarre, Rousseau David, INTRODUCTION À L’APPRENTISSAGE PROFOND (DEEPLEARNING) DE L'INTELLIGENCE ARTIFICIELLE , cultur sciences physiques , 11/10/2021 .Introduction à l’apprentissage profond (deep learning) de l'intelligence artificielle — (ens-lyon.fr)

(Xuanyi)Xuanyi Dong1, Yan Yan1, Wanli Ouyang2, Yi Yang1∗ 1University of Technology Sydney, 2 The University of Sydney , Style Aggregated Network for Facial Landmark Detection .Style Aggregated Networkfor Facial Landmark Detection (thecvf.com)

(Sasan, 2013) Sasan Karamizadeh1, Shahidan M. Abdullah1, Azizah A. Manaf1, Mazdak Zamani1, AlirezaHooman2 , An Overview of Principal Component Analysis , Journal of Signal and Information Processing, 2013 .(PDF) An Overview of Principal Component Analysis (researchgate.net)

(Ralph)Ralph Gross, Jie Yang, Alex Waibel , Interactive Systems Laboratories Carnegie Mellon University , Growing Gaussian Mixture Models for Pose Invariant Face Recognition . Growing gaussian mixture models for pose invariant face recognition - Pattern Recognition, 2000.Proceedings. 15th International Conferenceon (cmu.edu)

(Batta ,2020) Batta Mahesh , International Journal of Science and Research (IJSR) , Machine Learning Algorithms - A Review , 1, January 2020.Machine Learning Algorithms - A Review (ijsr.net)

(Ranjan , 2021)Ranjan Kumar Mishra , G. Y. Sandesh Reddy , Himanshu Pathak , TheUnderstandingofDeepLearning:A ComprehensiveReview,Article , India , 2021.The Understanding ofDeep Learning: A Comprehensive Review (hindawi.com)

(JAMMI , 2010)JAMMI ASHOK , VAKA SHIVASHANKAR , P.V.G.S.MUDIRAJ , AN OVERVIEW OF BIOMETRICS ,

jammi Ashok et. al. / (IJCSE) International Journal on Computer Science and Engineering , 2010.(PDF) AnOverview of Biometrics (researchgate.net)

(Andrian ,2023) Andrian Firmansyah , Tien Fabrianti Kusumasari , Ekky Novriza Alam , School of Industrial and System Engineering Telkom University Bandung, Indonesia , International Conference on Computer Science, Information Technology and Engineering (ICCoSITE) , Comparison of Face Recognition Accuracy of ArcFace, Facenet and Facenet512 Models on Deepface Framework , 2023 .Comparison of Face Recognition Accuracy of ArcFace Facenet and Facenet512 Models On Deepface Framework | PDF | Sensitivity And Specificity | Accuracy And Precision (scribd.com)

(V.L.B. De Mel, 2021)V.L.B. De Mel , University of Moratuwa , Triplet Loss, December 2021.(PDF) TripletLoss (researchgate.net)

(Ivan)Ivan William, De Rosal Ignatius Moses Setiadi, Eko Hari Rachmawanto , Heru Agus Santoso , Christy Atika Sari , Department of Informatics Engineering Dian Nuswantoro University Semarang, Indonesia , Face Recognition using FaceNet (Survey, Performance Test, and Comparison) .(PDF) Face Recognition using FaceNet (Survey, Performance Test, and Comparison) (researchgate.net)

(P. Jonathon) P. Jonathon Phillips, National Institute of Standards and Technology, Support Vector Machines Applied to Face Recognition. Support Vector Machines Applied to Face Recognition (neurips.cc)

(K.Meena, 2011)K.Meena, Dr.A.Suruliandi, local binary patterns and its variants for face recognition, mit, anna university, chennai. june 3-5, 2011. (pdf) local binary patterns and its variants for face recognition (researchgate.net)

(Oubaha, 2021) OUBAHA, A.&Sabri, Z,(2011), Système de reconnaissance faciale , Master’s Thesis, université sidi mohamed ben abdellah , maroc.Systeme de reconnaissance facia - Oubaha Amina - SabriZineb - Département d’Informatique Projet de - Studocu

(Akcha, 2020)Akcha, I.& Ammari, A,(2020), Développement d'un système de reconnaissance faciale,Master’s Thesis,Centre de développement des technologies avancées, CDTA.https://di.univ-blida.dz/jspui/handle/123456789/7208

(Bouzit, 2019)Bouzit, D (2019), Reconnaissance de visage basée sur une approche triangulaire . Master’s Thesis, Université de 8 Mai 1945 – Guelma.Université 8 mai 1945 - GUELMA: Reconnaissance de visage basée sur une approche triangulaire (univ-guelma.dz).

(Agrawala, 2014)Agrawala ,A ,K,&Singh, Y, N ,(2014), International Conference on Intelligent Computing, Communication & Convergence (ICCC-2014).Chance Constraint Based Multi-objective Vendor Selection Using NSGAII (sciencedirectassets.com)

(Maëlig, 2021)Maëlig,J,& Grossrieder,L(2021) , Enjeux et perspectives de la reconnaissance faciale en sciences criminelles , resses de l’Université de Montréal. (PDF) Enjeux et perspectives de la reconnaissance faciale en sciences criminelles (researchgate.net)

(Boukerrouche, 2018) Boukerrouche, Y,&Zerriouh ,A,(2018)Mise au point d’une application de détectionetreconnaissance faciale , Master’s Thesis, université abou bekr belkaid de tlemcen , faculté de technologiedépartement d’informatique.dspace.univ-tlemcen.dz/bitstream/112/13471/1/Mise-au-point- dune-application-de-detection-et-reconnaissance-faciale. Pdf

(Eduardo, 2019) Eduardo, T& Bruno, A,(November 30, 2019), KVRI Group - Vision Robotic and Images Federal University of Parana´.(PDF) Primer on artificial intelligence and robotics (researchgate.net)

Webography

[1] What is Facial Recognition | Glossary | CyberGhost VPN

[2] Face Detection vs. Face Recognition vs. Face Verification (idscan.net)

[3]Introductory Chapter: Artificial Intelligence - Challenges and Applications | IntechOpen

[4] Pattern Recognition and Machine Learning - Microsoft Research

[5] Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions | SN Computer Science (springer.com)

[6] Natural Language Processing (NLP) Solutions | IBM

[7] What is Computer Vision? | IBM

[8](PDF) Feature Extraction Techniques in Facial Expression Recognition (researchgate.net)

[9] Convolutional neural networks: an overview and application in radiology | Insights into Imaging | Full Text (springeropen.com)

[10] (PDF) Hybrid Learning Systems: Integrating Traditional Machine Learning with Deep learning Techniques. (researchgate.net)

[11] Middle East Program | CSIS

[12] 7 New Facial Recognition Technology Trends to Boom in the Future (lystloc.com)

[13] Pros and cons of facial recognition (techtarget.com)

[14] Facial Detection and Recognition With Dlib | Width.ai

[15] What is Artificial Intelligence | Pure Storage

[16] Types of Artificial Intelligence | IBM

[17] What is Facial Recognition & How does it work? (kaspersky.com)

[18] Gaussian Mixture Models | SpringerLink

[19] Support Vector Machine (SVM) Algorithm - GeeksforGeeks

[20] How is image processing in Matlab different from Python? (matlabsolutions.com)

[21] Past, Present, and Future of Face Recognition: A Review (mdpi.com)

[23] What is Python? Executive Summary | Python.org

[24] why we should use gray scale for image processing - Stack Overflow

[25] OpenCV: OpenCV modules

[26] NumPy Documentation

[27] os — Miscellaneous operating system interfaces — Python 3.12.6 documentation

[28] GitHub - rcmalli/keras-vggface: VGGFace implementation with Keras Framework

[29] Distance computations (scipy.spatial.distance) — SciPy v1.14.1 Manual

[30] Face Recognition in Python: A Comprehensive Guide | by Basil CM | Medium

[31] GitHub - aaronpp65/face-recognition-vggface2: The VGGFace refers to a series of models developed for face recognition and demonstrated on benchmark computer vision datasets by members of the Visual Geometry Group (VGG) at the University of Oxford.

[32] 4 Popular Face Detection Methods (plugger.ai)

[33]Histogram of Oriented Gradients: An Overview | Built In

[35] What Is a Convolutional Neural Network? | 3 things you need to know - MATLAB & Simulink (mathworks.com)

[36] How to deal with image resizing in Deep Learning | by Adriano Dennanni | Neuronio | Medium

[37]Best Practices for Preparing and Augmenting Image Data for CNNs - MachineLearningMastery.com

[38]How to Perform Face Recognition With VGGFace2 in Keras - MachineLearningMastery.com

[39] Convolutional Neural Network: A Complete Guide (learnopencv.com)

[40] Cosine Similarity (tutorialspoint.com)

[41] Face recognition - Sface match result evaluation - Python - OpenCV

[42] (PDF) Face Recognition using SOM Neural Network with Different Facial Feature Extraction Techniques (researchgate.net)

[43] What is the Cosine Similarity or Cosine Distance? (Measure of Angle) (datacadamia.com)

[44] Python Features - GeeksforGeeks

[45] 5 Great Raspberry Pi IDEs For Programmers and Students (makeuseof.com)

[46] Remote access - Raspberry Pi Documentation

[47] How to use VNC Connect Remote Desktop software effectively and easily (realvnc.com)

[48] What is PuTTY? A Comprehensive Guide - Rushax

[49] (PDF) Image processing and face detection analysis on face verification based on the age stages (researchgate.net)

[50] How accurate are facial recognition systems – and why does it matter? - iMEdD Lab

[51] How Accurate is Facial Recognition in 2021? | Face Recognition Accuracy Evaluation | RecFaces

[52] CiteSeerX (psu.edu)

[53] CiteSeerX (psu.edu)

[54] National Institute of Standards and Technology (nist.gov)

[55] PFE Smart Door | PDF | Biométrie | Analyse en composantes principales (scribd.com)

[56] Rapport L2 (Version Final) | PDF | Environnement de développement | Arduino (scribd.com)

[57] (PDF) Robust Real-Time Object Detection (researchgate.net)

Université 8 mai 1945 - GUELMA: Régulation intelligente du trafic routier par des feux de carrefours (univ-guelma.dz)

[58] Introducing Thonny, a Python IDE for learning programming | Request PDF (researchgate.net)

[59] Proceedings of the 15th Koli Calling Conference on Computing Education Research, Koli, Finland, November 19-22, 2015 - researchr publication