Algeria of Republic Democratic People's Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université Kasdi Merbah - Ouargla

Faculté des Nouvelles Technologies de l'Information et la Communication

D ´ département de l'informatique et Technologies de l'information

Master's dissertation

Field *: Mathematics and Computer Science*

Branch *: Computer Science*

Specialty : *: Network Administration and Security*

Thème

Convolutional Autoencoder for distorted facial images recognition

Supervisor

• Dr.Azzaoui Hanane

Presented By

• Siagh Aya

• Bennouna Safa

2023/2024

Acknowledgements

Thank Allah, the Most Merciful, for granting us the health, strength, and perseverance to complete this work. May He guide us on the right path and bless us with His grace.

We extend our heartfelt gratitude to our family for their unwavering support and encouragement throughout this journey. Their love and patience have been our constant source of strength.

Special thanks to Madame Azzaoui Hannan for her encouragement and kindness.

We want to thank the jury members for their time and effort in examining our work, as well as all of the Computer Science department's instructors for their devotion and patience.

Finally, we want to thank everyone who has helped, advised, or encouraged us on our journey. Your contributions have proven essential. May Allah bless our mother with health and longevity for her unwavering patience and support.

ABSTRACT

Facial recognition technology has advanced considerably in recent years, but it encounters difficulties in accurately identifying distorted facial images. In this master's thesis, we propose a system aimed at improving facial recognition using Convolutional Autoencoders (CAE) to process distorted images. This unsupervised deep learning approach allows us to effectively denoise and reconstruct distorted facial images, thereby improving recognition accuracy.

We chose Convolutional Autoencoders for their performance in detecting deformation and reconstructing images. We trained the CAE on a custom dataset comprising both distorted and undistorted facial images, generated using the Photopea application and systematically organized. A classifier based on VGG16 was then used to differentiate clear faces from distorted faces, leveraging the enhanced images produced by the CAE.

The experiments conducted on the distorted face dataset yielded favorable results, demonstrating significant improvements in image quality and classification accuracy. The system application was tested and validated, showing its potential to enhance the robustness and reliability of facial recognition systems in security applications.

Keywords : Distorted faces, Facial recognition, Autoencoder, Convolutional Autoencoder, VGG16, Security applications.

RESUME

La technologie de reconnaissance faciale a considérablement avancé ces dernières années,mais elle rencontre des difficultés à identifier avec précision les images faciales déformées. Dans ce mémoire de master, nous proposons un système visant à améliorer la reconnaissance faciale en utilisant des Autoencodeurs Convolutionnels (CAE) pour traiter les images déformées. Cette approche d'apprentissage profond non supervisée nous permet de débruiter et de reconstruire efficacement les images faciales déformées, améliorant ainsi la précision de la reconnaissance.

Nous avons choisi les Autoencodeurs Convolutionnels pour leurs performances en détection la déformation et en reconstruction d'images. Nous avons entraîné le CAE sur un ensemble de données personnalisé comprenant des images faciales déformées et non déformées, générées à l'aide de l'application Photopea et organisées systématiquement. Un classificateur basé sur VGG16 a ensuite été utilisé pour différencier les visages nets des visages déformés, en tirant parti des images améliorées produites par le CAE.

Les expérimentations menées sur l'ensemble de données de visages déformés ont donné des résultats favorables, démontrant des améliorations significatives de la qualité des images et de la précision de la classification. L'application du système a été testé et validé, montrant son potentiel à renforcer la robustesse et la fiabilité des systèmes de reconnaissance faciale dans les applica- tions de sécurité.

Mots clés : Visages déformés, Reconnaissance faciale, Autoencoder, Auto-Encodeur Convolutionnel, VGG16, Applications de sécurité.

الملخص

وقد أحرزت تكنولوجيا التعرف على الوجه تقدما كبيرا في السنوات الأخيرة، ولكنها تواجه صعوبات في التعرف بدقة على الصور المشوهة للوجه. في أطروحة الماجستير هذه، نقترح نظاماً يهدف إلى تحسين التعرف على الوجه باستخدام رموز تلقائية متحولة **Convolutional Autoencoder(CAE)** لمعالجة الصور المشوهة.

ويتيح لنا هذا النهج التعليمي العميق غير الخاضع للإشراف أن ننحرف ونعيد بناء الصور المشوهة للوجه على نحو فعال، مما يحسن من دقة التعرف.

لقد اخترنا مبرمجين آليين متحولين لأداءهم في كشف التشويه وإعادة بناء الصور. ودرّبنا المركز على مجموعة بيانات مخصصة تضم صوراً مشوّهة وغير مشوّهة على حد سواء للوجه، ونظّمت بشكل منهجي. ثم استُخدم مصنّف يستند إلى **VGG16** لتمييز الوجوه الواضحة عن الوجوه المشوّهة، والاستفادة من الصور المحسّنة التي أنتجها المركز.

وقد أسفرت التجارب التي أجريت على مجموعة بيانات الوجه المشوهة عن نتائج مواتية، مما يدل على حدوث تحسن كبير في نوعية الصورة ودقة التصنيف كما تم تطبيق النظام والتحقق منه، مما يبين قدرته على تعزيز قوة وموثوقية نظم التعرف على الوجه في التطبيقات الأمنية.

الكلمات الرئيسية : الوجوه المشوهة، التعرف على الوجه، **Autoencoder** ، **Convolutional Autoencoder** ، **VGG16** ، التطبيقات الأمنية.

Table des matières

Liste des figures

Liste des tableaux

0.1 GENERAL INTRODUCTION

Facial recognition systems, in the digital technological framework today, have turned out to be the most important security solution to improve personal identification in various aspects.

However, these systems get rendered to less effective states following the introduction of a lot of distortions in facial images, thereby posing serious problems in identification. It is against this background that this thesis addresses this critical problem by proposing the development and implementation of a robust system using Convolutional Autoencoders.

The principal aim of this research is to enhance the accuracy and reliability of facial recognition systems when analyzing distorted images. This goal is pursued through the innovative application of CAEs, which are adept at denoising and reconstructing facial images, thereby ensuring higher precision in recognition tasks.

The methodology encompasses the training of the CAE on a custom dataset comprising both distorted and undistorted images, followed by the deployment of a VGG16-based classifier to assess the quality and clarity of the reconstructed images. By integrating theoretical knowledge with empirical analysis, this study demonstrates substantial improvements in image quality and recognition accuracy, thereby underscoring the potential of Convolutional Autoencoders in overcoming the limitations of traditional facial recognition technologies, particularly in security-sensitive environments.

This thesis not only contributes to the academic discourse on advanced machine learning techniques in image processing but also proposes practical solutions that could be integrated into existing security systems, thereby offering a dual advantage of theoretical enrichment and practical utility.

This thesis is organized into several chapters, each dedicated to different aspects of utilizing Convolutional Autoencoders for the recognition of distorted facial images. The structure is designed to guide the reader through the stages of research, from theoretical foundations to practical implementations and analysis. The following is an overview of each chapter :

Chapter One - Overview of Biometrics : This chapter begins by providing an overview of biometrics, reviewing its history, significance, and primary role in defining digital identity. Afterward, we discuss various biometric measures and how they are used to secure data and privacy.

Chapter Two - Convolutional Neural Networks and Autoencoders : The second chapter focuses on the theoretical and practical foundations of Convolutional Neural Networks (CNNs) and Autoencoders, explaining how they work and their applications in the field of computer vision.

Chapter Three - Facial Recognition : This chapter addresses facial recognition techniques, concentrating on the challenges these techniques face when dealing with distorted images and how to overcome them using self-convolutional neural networks.

Chapter Four - Implementation and Results : The research concludes with a detailed presentation of the practical implementation of the developed system, analyzing the obtained results and discussing their effectiveness in improving facial recognition accuracy.

Biometrics

1

Sommaire

1.1 Introduction :

The "Biometrics" chapter delves into a sophisticated realm of technology and security, utilizing distinct physical characteristics and behavioral patterns to uniquely identify people.

Understanding the foundations of biometrics and identifying the many uses that improve security, safeguard privacy, and improve user experiences are the main topics of this chapter.

We will go over the main points that need to be covered, such as the many kinds of biometrics, their benefits and drawbacks, the architecture of biometric systems, and their functioning. This chapter's goal is to provide the reader a thorough grasp of biometric technology and how important it is to boosting user experience and digital security.

1.2 History of biometrics

Biometrics has been a concern for centuries. Proving one's identity reliably was done using several techniques. From prehistory man knew the uniqueness of fingerprints, which meant that signatures by fingerprints were sufficient to prove the identity of an individual. Indeed, two centuries before Christ, the Emperor Ts-In-She authenticated certain sealed with the fingerprint.

At the beginning of the nineteenth century, in France, Alphonse Bertillon launched the first steps of the scientific police. He proposed the first method of biometrics that can be described as a scientific approach : bertillonage allowed the identification of criminals through several physiological measures.

At the beginning of the twentieth century, biometry was rediscovered by William James Herschel, an English officer who had the idea of having his subcon-tractors sign their fingerprints to find them easily in case of unhonored contracts.

As a result, police departments have begun using fingerprints as a unique and reliable feature to identify an individual. Biometrics is constantly growing especially in the field of secure identity docu-ments such as the national identity card, passport, or driving license. This technology is running on new platforms, including chip cards based on the microprocessor.[16]

1.3 What Are biometrics ?

Biometrics refer to the study of physical or behavioral characteristics used for the identification of a person. It involves measuring unique biological traits for identification or verification purposes.

Physiological biometrics focus on physical characteristics like fingerprints, hand geometry, iris, face, and DNA, while behavioral biometrics analyze actions like keystrokes, signatures, and voice patterns.

Biometric systems use these characteristics to provide authentication for computer-based security systems, allowing individuals to be identified based on who they are rather than what they have or know. [34]

1.4 Why use biometrics ?

Biometrics is used for authentication and verification purposes in numerous systems. Here are a few motives why biometrics is used : [12]

1.4.1 Enhanced security :

Biometrics provides a better stage of protection compared to conventional authentication methods like passwords or PINs. It is hard to forge or mirror biometric traits, making it extra secure against identity robbery or fraud. [30]

1.4.2 Privacy protection :

Biometric systems aim to protect sensitive information by not storing raw biometric data. Instead, they store extracted templates that contain discriminative information for recognition purposes.[13]

1.4.3 Operability in different environments :

Biometric structures can use multiple sensors to gather the same biometric modality, letting them operate in extraordinary environments.
For example, a face reputation gadget may also use each a visible spectrum digicam and a close to infrared digital camera to facilitate biometric recognition in a middle of the night environment. [35]

1.4.4 Scalability :

Biometric structures may be without difficulty scaled to accommodate a huge variety of customers. They are extensively used in various programs, inclusive of private, commercial, and governmental identity management. [30]

1.4.5 Convenience :

Biometrics gives a convenient way for customers to authenticate themselves. Instead of remembering and entering passwords or sporting bodily tokens, customers can clearly use their biometric tendencies, which includes a fingerprint or face, to verify their identity. [12]

1.4.6 Non-transferable :

Biometric traits cannot be easily transferred or shared, unlike passwords or tokens. This makes biometrics more secure against identity theft or unauthorized access. [12]

1.5 Types of biometrics :

Biometrics can be classified into most important sorts : physiological biometrics and behavioral biometrics :

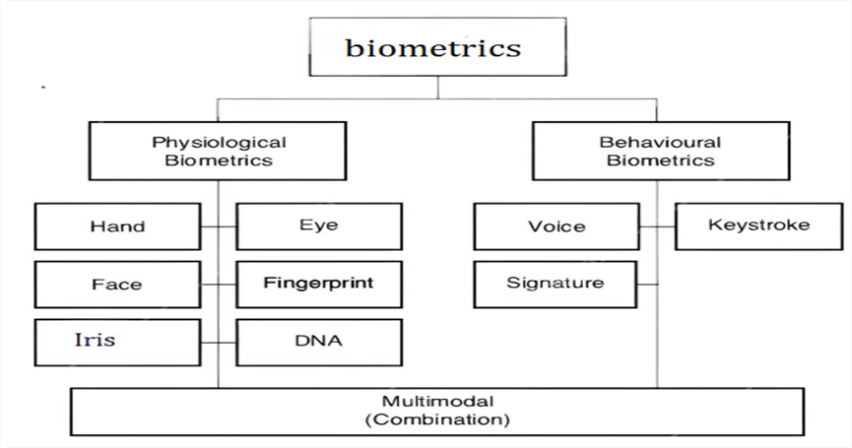

FIGURE 1.1 – *Types of biometric*

1.5.1 Physical identification methods :

Physiological biometrics involve physical characteristics of a person for identification or verification purposes. Some examples include :

Fingerprints :

FIGURE 1.2 – *Fingerprints*

A fingerprint is a pattern of ridges and furrows located on the tip of each finger. Fingerprints were used for personal identification for many centuries, and the matching accuracy was very high. Patterns have been extracted by creating a colored impression of the fingertip on paper. Today, compact sensors provide digital images of these patterns. Fingerprint recognition for identification acquires the initial image through live scanning of the finger by direct contact with a reader device. [34]

Hand Geometry :

FIGURE 1.3 – *Hand Geometry*

Hand geometry systems produce estimates of certain measurements of the hand such as the length and the width of fingers.
Various methods are used to measure the hand. These methods are most commonly based either on mechanical or optical principle.
The latter ones are much more common today. The hand geometry is used for identification and recognition of a person . [34]

Iris :

FIGURE 1.4 – *Lris*

Iris popularity is a shape of biometric identification that identifies people based at the one of a kind patterns of their irises. The colored part of the eye that surrounds the scholar is known as the iris, and it's miles composed of extraordinary patterns of ridges, furrows, and different traits that may be used for identity purposes .
The iris can be distinguished from other people's irises by comparing its ridges and furrows.
Iris recognition technology works by taking a picture of the subject's iris and analyzing its individual characteristics before comparing those characteristics to a library of pre-existing iris templates. [1]

Face :

FIGURE 1.5 – *Face*

Facial recognition is the most natural means of biometric identification. The approaches to face recognition are based on shape of facial attributes, such as eyes, eyebrows, nose, lips, chin and the relationships of these attributes. As this technique involves many facial elements ; these systems have difficulty in matching face images .[34]

DNA :

FIGURE 1.6 – *DNA*

(Deoxyribonucleic Acid) sampling is rather intrusive at present and requires a form of tissue, blood or other bodily sample. This method of capture still has to be refined. So far the DNA analysis has not been sufficiently automatic to rank the DNA analysis as a biometric technology. The analysis of human DNA is now possible within 10 minutes. As soon as the technology advances so that DNA can be matched automatically in real time, it may become more significant. At present DNA is very entrenched in crime detection and so will remain in the law enforcement area for the time being.[34]

1.5.2 Behavioral identification methods :

Behavioral biometrics focus on the actions of a person, considering inner variants like mood and health condition. Examples of behavioral biometrics include :

Keystroke :

FIGURE 1.7 – *Keystroke*

Keyboard is the part that helps us to communicate with computer.
People use keyboard in different ways. Some people type fast, some slow. The speed of the typing also depends on the mood of a person and a time of a day. Biometric keystroke recognition is a technology of recognizing people from the way they are typing.
It is rather important to understand that this technology does not deal with "what" is written but "how" it is written.[34]

Signature :

FIGURE 1.8 – *Signature*

The way a person signs his or her name is known to be characteristic of that individual.
Signature is a simple, concrete expression of the unique variations in human hand geometry. Collecting samples for this biometric includes subject cooperation and requires the writing instrument.

Signatures are a behavioral biometric that change over a period of time and are influenced by physical and emotional conditions of a subject.
In addition to the general shape of the signed name, a signature recognition system can also measure pressure and velocity of the point of the stylus across the sensor pad. [34]

Voice :

FIGURE 1.9 – *Voice*

The features of an individual's voice are based on physical characteristics such as vocal tracts, mouth, nasal cavities and lips that are used in creating a sound. These characteristics of human speech are invariant for an individual, but the behavioral part changes over time due to age, medical conditions and emotional state. Voice recognition techniques are generally categorized according to two approaches :
1) Automatic Speaker Verification (ASV) and 2) Automatic Speaker Identification (ASI).
Speaker verification uses voice as the authenticating attribute in a two-factor scenario. [34]

1.6 Advantages and disadvantages of biometrics

1.6.1 Advantages of Biometrics :

Improved security :

Biometric systems, mainly the ones primarily based on deep getting to know techniques, can offer a higher stage of safety in comparison to traditional handcrafted processes. They can perform at decrease False Match Rates (FMRs), reducing the probabilities of fake matches and unauthorized access. [30]

Enhanced privacy protection :

Biometric systems utilizing deep learning-based techniques can significantly increase the level of privacy protection. They can conceal soft-biometric information and minimize the risk of privacy breaches. [30]

Improved biometric performance

Deep learning-based methods have been shown to improve biometric performance for systems based on different biometric characteristics. This means that biometric systems utilizing deep learning can potentially operate with higher accuracy and reliability. [30]

Compatibility with various biometric characteristics :

Deep mastering-primarily based biometric structures have been correctly applied to numerous biometric characteristics, consisting of face recognition and gait analysis. This makes them versatile and adaptable to distinctive use cases . [30]

1.6.2 Disadvantages of Biometrics :

Vulnerability to attacks :

Deep learning-based biometric systems can be vulnerable to certain attacks due to their enhanced generalization capabilities. These attacks can exploit the artificial biometric samples created by Generative Adversarial Networks (GANs) and launch dictionary attacks on biometric verification systems.[30]

Detectability of privacy enhancement methods :

Some soft-biometric privacy enhancement methods can be detectable and vulnerable to sophisticated attacks. This raises concerns about the effectiveness and reliability of these methods in providing privacy protection.[30]

Lack of comprehensive empirical comparison :

While deep studying-based totally strategies have proven promising results, there's nevertheless a need for a comprehensive empirical evaluation among earlier handcrafted techniques and modern-day deep learning-based totally methods to assess their performance and effectiveness. [30]

Potential for false matches :

Despite improvements in accuracy, there's still a opportunity of false matches in biometric structures. This can cause unauthorized access or denial of get admission to to legitimate users. Continuous improvements and improvements are required to decrease the occurrence of false suits .[30]

1.7 Biometric system architecture :

The generic architecture of a biometric system consists of five main modules as depicted in Figure 10 :[10]

1.7.1 Capture module :

It consists of capturing the biometric raw data in order to extract a numerical representation. This representation is then used for enrollment, verification or identification.

1.7.2 Signal processing module :

It allows the reduction of the extracted numerical representation in order to optimize the quantity of data to store during the enrollment phase, or to facilitate the processing time during the verification and identification phases. This module can have a quality test to control the captured biometric data.

1.7.3 Storage module :

It is used to store biometric individuals' templates.

1.7.4 Matching module :

It is used to compare the extracted biometric raw data to one or more previously stored biometric templates. The module therefore determines the degree of similarity (or of divergence) between two biometric vectors.

1.7.5 Decision module :

It is used to determine if the returned index of similarity is sufficient to determine the identity of an individual.

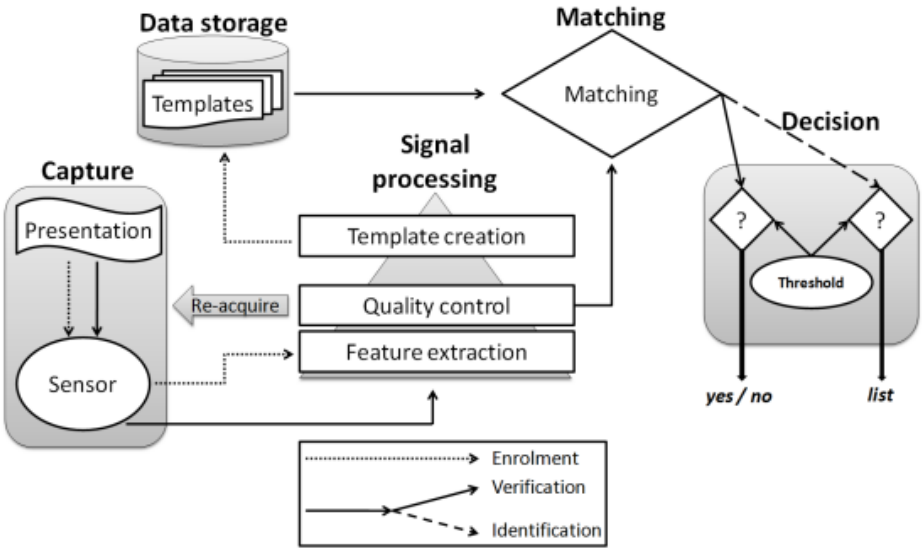

FIGURE 1.10 – *Generic architecture of a biometric system*

1.8 How do biometrics Works :

A biometric system is basically a pattern recognition system that uses an individual's biometric data. There are three modules in a biometric system, the verification, identification and verification module. While biometric systems can combine authentication, verification, and identification, there are some key differences between those three facets. Namely, identification asks, "who are you?" while verification asks, "Is there data associated with you?" Authentication asks, "Are you who you say you are? [31]

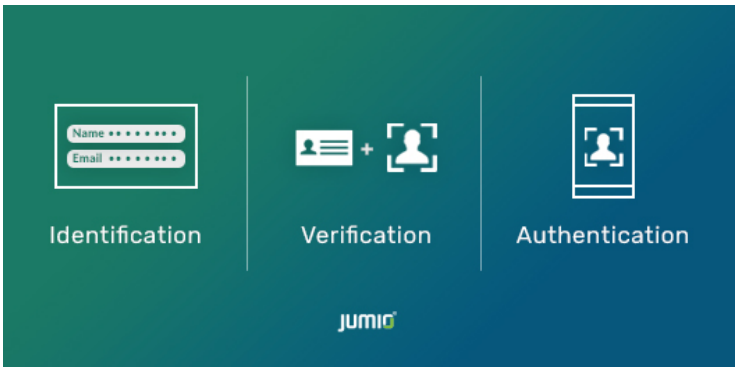

FIGURE 1.11 – *Modules of a biometric system*

1.8.1 Biometric authentication

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

The process of biometric authentication looks like this : [31]

1- Enrollment :

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

2- Live Sample :

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

3- Comparison :

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

1.8.2 Biometric identification

Biometric identification can be applied to digital and physical scenarios, and it's a solution that is used in defense, law enforcement, and border control. With identification, there is a database that contains physical characteristics of a vast number of people for instance, the FBI's repository stores the height, hair color, weight, eye color, scars, and tattoos of over 70,000,000 criminal records. With authentication, a person's features are compared to one specific template or With identification, however, the person's features are matched against the entire database . [31]

1.8.3 Biometric verification

Biometric verification is often confused with authorization – however, there is a subtle difference between the two processes. While authentication indicates that a person has the same biometric features as somebody who is already in the system, verification can conclusively prove that their online identity is linked with their real-life identity . [31]

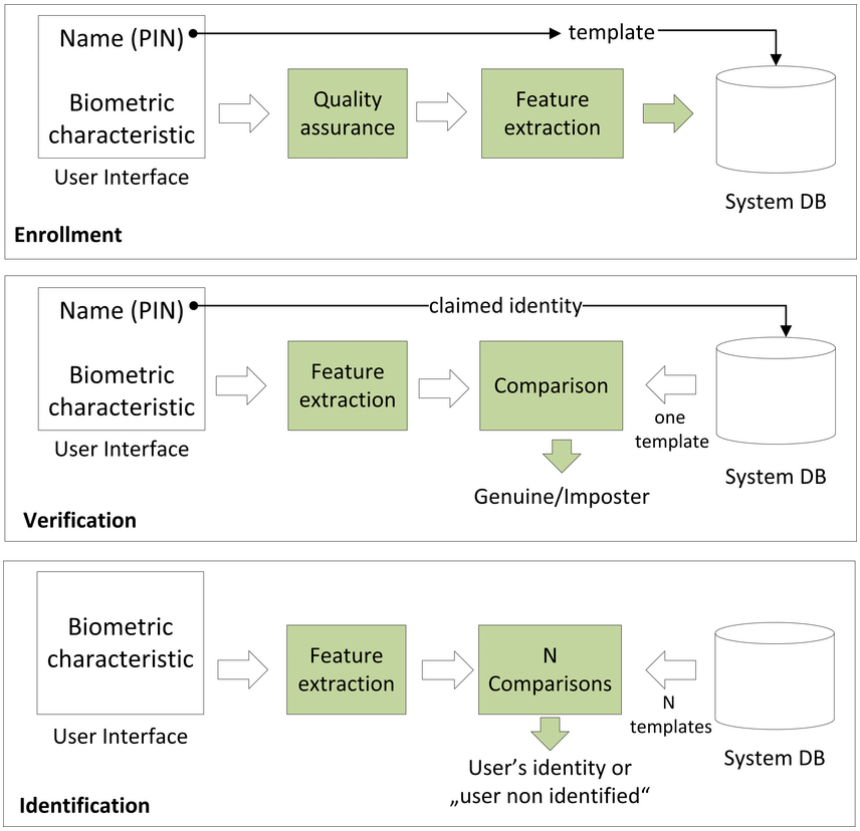

FIGURE 1.12 – *the main operations of a biometric system*

1.9 Applications of biometric systems :

Biometric systems have a wide range of applications in various fields. Some common applications include : [34]

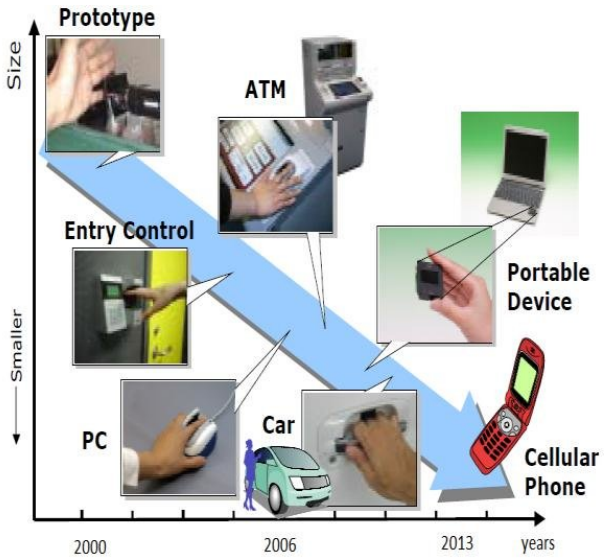

FIGURE 1.13 – *Applications of biometric systems*

1.9.1 Fingerprint Recognition :

- Large-scale Automated Fingerprint Imaging Systems (AFIS) used by law enforcement.
- Fraud prevention in entitlement programs.
- Physical access control for doors.
- Logical access to computer systems.
- Administering drugs and controlled substances to patients.

1.9.2 Hand Geometry :

- Access control systems.
- Time and attendance applications.

1.9.3 Iris Recognition :

- Airline passenger screening.
- Inmate identification in correctional facilities.
- Border security.
- Facility access control.
- Computer login.
- ATMs and grocery stores.

1.9.4 Facial Recognition :

- Controlled environments like casinos for identifying card counters.
- Screening individuals to see if they are known to the system.
- Fraud prevention during visa or driver's license applications.

1.9.5 Voice Recognition :

- Access control.
- Banking.
- Government offices.
- Entertainment applications.
- Smart cards, PIN, and other security purposes.

1.10 Performance of biometric systems

The performance of biometric systems can be evaluated using the Mean Square Error (MSE) and the Structural Similarity Index (SSIM). These metrics provide insights into the quality and similarity of perturbed images compared to the original image.[11]

— **MSE** measures the average squared difference between the perturbed image and the original image. A lower MSE value indicates a higher similarity between the two images.

— **SSIM** measures the structural similarity between the perturbed image and the original image. It ranges from 0 to 1, where 1 represents a perfect match between the two images.

figure 14 provides the rate of perturbations applied to images in each experimental scenario, along with the corresponding MSE and SSIM values. These metrics can be used to compare the proportion of perturbations and the quality of the perturbed images in different scenarios. [11]

Metric	Scenarios					
	(ii) Gaussian noise	(iii) Laplacian noise	(iv) Spread transformation	(v) Imploding transformation	(vi-a) OTB	(vi-b) OTB
MSE	712.48	361.78	493.52	464.07	1785	2830
SSIM	0.21	0.30	0.23	0.58	0.20	0.13

FIGURE 1.14 – *The rate of perturbations applied to images in each experimented scenario.*

Conclusion :

The chapter on "Biometrics" provides a thorough analysis of the integration of technology and security, focusing on the use of unique behavioral and physical characteristics for personal identification. The chapter outlines the various uses and intricate workings of biometric systems during its investigation, highlighting both the benefits—like increased convenience and security—and drawbacks—like vulnerability to intrusions and the requirement for ongoing development. The chapter highlights the vital role biometrics plays in bolstering digital security and improving user experiences by synthesizing these findings, opening the door for a more sophisticated knowledge of this crucial field.

Convolution Neural Network 2

Sommaire

2.1 Introduction

In this second chapter of this research, we focus on the core of artificial intelligence and deep learning technologies. This chapter holds a strategic position in our work as it is dedicated to the Convolutional Autoencoders and their function in data analysis and obtaining accurate visual characteristics.

The primary purpose of this chapter is to enhance the knowledge of the process of converting raw data into meaningful information, which is crucial for creating better and more effective recognition systems. Thus, after outlining the history and the different approaches of Machine Learning and offering an in-depth view of Deep Learning, the reader is equipped with the theoretical knowledge to comprehend the latest advancements in this field.

A considerable portion of this chapter is devoted to discussing the uses of autoencoders with emphasis on the use in handling distorted images which is one of the most critical areas of enhancing the performance of face recognition and identity authentication systems. The possibility to restore the necessary information from the distorted images is the important step towards the creation of the more reliable and efficient electronic protection.

2.1.1 History

AI, ML, DL technologies have progressed tremendously in a short time and, consequently, have transformed almost every technical aspect of modern society, hence, the lives of people.

Convoulution Neural Network is a deeply learned network that has originally derived from the natural visual observation. Hubel and Wiesel described simple cells in the animal visual cortex responsible for detecting light. According to Fukushima, the message appeared in 1980, which could be viewed as a leading step towards the development of CNN.

CNN's basic structure was established by LeCun and others in 1990 and the authors designed LeNet-5 for handwriting digit recognition. LeNet-5 also has layers like the other neural networks and can too be trained using backpropagation.

As such, it can very well reproduce the original image to be used in visions pattern recognition, with little or no preprocessing. After capturing the images, Zhang et al. used the SIANN to recognize the characters in the picture. However, due to availability of limited data and limited power of computers, their networks had issues with complex computations.

The evolutions that have been developed since 2006 are the ones that deal with the challenges in the training of deep CNNs. Krizhevsky et al. proposed Alex net as a typical CNN architecture for image classification and also as a model with simplified designs with many enhancements. The design of Alex Net is also akin to LeNet-5 ; however, it has additional layers.

Based on the performance of Alex Net several other works have been proposed such as ZFNet, VGGNet, Google Net, ResNet. The new trend in the evolution of architecture is deeper networks such as ResNet that have triumphed the ILSVRC 2015 contest.[9]

2.2 Machine Learning

2.2.1 Introduction of Machine Learning

Machine learning as defined by Michie and 3 others (D. Michie 1994) is broadly defined as involving Automatic computing procedures based on logic or binary operations that learn a task from a series of examples.

Here we are interested only in the classification and it is still ambiguous what belongs to the Machine Learning field. More emphasis has been given on steps in a decision tree where classification takes place step by step. These are capable of representing the most complex problem given sufficient data.

Other approaches, which are currently being worked out and seem to be more versatile, at least in principle, are the genetic algorithms and inductive logic procedures (ILP) where learning may take place in more complicated levels where the types of attributes may be different and more intricate and it is possible that learning takes place on different levels : there may be hierarchies of attributes and classes and so on. Machine learning or data mining tries to produce such classifying expressions that any human being can easily understand easily.

They must be similar enough to human rationality to give information on the decision-making process. As with statistical techniques, background knowledge may be used in development, but operation is realised without explicit human control.
To learn is :

to acquire understanding, skills in, or proficiency with, something through practice or acquire knowledge about or learn or become proficient in some skill in some art or craft or trade to gain by experience, example, or practice of an ability or a skill in to put into ones memory ; cause to remember, learn.

In its essence, Machine Learning describes a methodology for developing computer systems which are capable of improving their performance as a result of learning from experience.

Automated extraction of useful information from a body of data by building good probabilistic models. Ideally suited for areas with lots of data in the absence of a general theory. [2]

2.3 ML Approaches

The main groups of learning algorithms can be identified : Supervised, unsupervised, Semi-supervised, and Reinforcement learning.

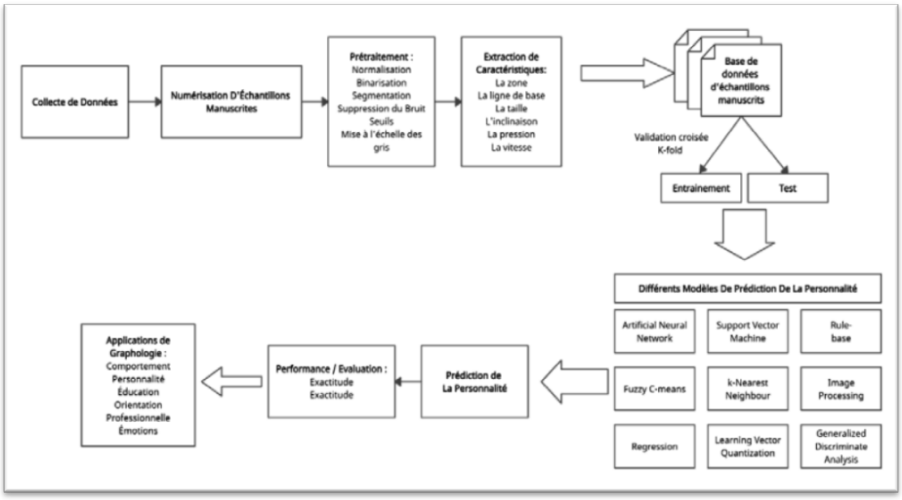

FIGURE 2.1 – *ML Technique*

There are various classifications of ML, they are classified as follows ; supervised learning, unsupervised learning, Reinforcement learning (RL).

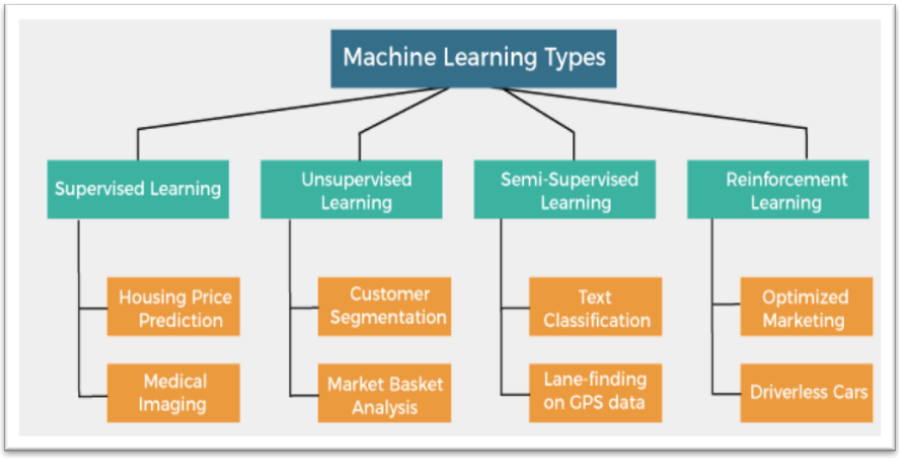

FIGURE 2.2 – *The four different Machine Learning algorithms*

2.3.1 Supervised learning

All the necessary information for training in supervised learning is that every observation in the dataset has a label.

This label could be binary and categorical for example, results from cancer tests where the patient either has a benign tumor or a malignant tumor or even quantitative, for instance an increased gene expression fold change. The rationale of calling these methods as supervised can be attributed to the fact that, they require supervision in form of labeled data.

Supervised learning specifically focuses on the utilization of several input features to estimate the value of one or several desirable outcomes.

While supervised learning is the most common domain of the ML applications and has been fulfilled most of the successful applications of the concept, it requires the large amount of the labeled data, which has been one of the major hurdles in developing such models in the scena-

rios where creation of the large-scale labeled dataset takes a significant amount of time. This, in addition to the fact that manual labeling is tiresome and time-consuming, means that it can only be done by specialists that, in some fields like health care, are scarce and costly. [24]

2.3.2 Unsupervised learning

It means that only the input samples are provided to the learning system, e. undefined (such as clustering and estimation of probability density function).

One can use unsupervised methods when these challenges occur although building large-scale datasets may be unfeasible thereby limiting the application of the full potential of supervised learning.

The unsupervised learning is used to identify relationships within the input data points. is also used to find a dense representation for high-dimensional data points. These approaches do not need labeled data as input which would have been time consuming and required a lot of manpower for labeling.

Some of the applications of unsupervised learning include clustering, association mining, and dimensionality reduction[24]

2.3.3 Semi-supervised Learning

Machine learning is a form of supervised learning in combination with unsupervised learning whereby some portions of the datasets are partially prelabeled.

The labeled portion is used to estimate the unlabeled segment ; e. g. , when labeled segment = 100% and unlabeled segment = 0%, the subsequently calculated optimal number of clusters would be based on the labeled portion only without considering the unlabeled portion. g. , text/image retrieval systems).

Semi supervised learning is also a form of learning where only small portion of the training data used to create an ML model is labeled while other samples are unlabeled.

Semi supervised learning is aimed at enhancing the model performance for those cases in which only a few labeled samples are available whereas there are many unlabeled samples available. The unlabeled samples can be used for the steps for adjustment of all models to enhance their performance.[24]

2.3.4 Reinforcement learning

Reinforcement Learning (RL) is a subfield of Machine Learning (ML) that aims at creating an entity known as an agent with the probability of getting the highest amount of total reward through sample actions in an environment.

An RL system includes an agent that gets feedback from it's actions with respect to the series of action-state-action, now often referred to as action sequence up to the terminal point like wining a game of chess or go.

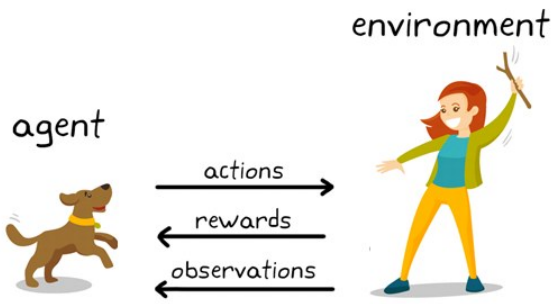

FIGURE 2.3 – *Reinforcement learning*

Some examples of environments include the games itself or a simulated environment such as in the case of autonomous cars. an important factor in RL is that it has been shown to outcompete human players in games because it is much easier and faster to run many simulated games with trial-and-error, especially since it optimizes the agent's performance.

Unlike supervised, unsupervised or semi-supervised learning techniques RL does not work on training data-sets but gains experience by trial and error. This approach obviously does not require data labeling but at the same time makes the use of RL approach only possible in the situations where simulating trial-and-error procedures is easy.[24]

2.4 Deep learning

2.4.1 Introduction to Deep Learning

Deep Learning is a branch of the machine learning that mostly use artificial neural networks that are particularly structured with different layers and nodes. These networks mimic the characteristics and functionality of human brain networking.

Deep learning algorithms can be described as similar to the nervous system since the neuron connects to others and sends information. The current generations of these models shall function in layers, although a basic deep learning model has at least three of this layers. Each layer takes processed information from the layer below it and gives output to the layer above it.

For most operations, the difference between machine learning and deep learning tend to be centered in feature extraction. In the case of machine learning, it entails feature extraction wherein the human inputs the features of the data while, in deep learning, the model identifies the features on its own. Also, deep learning models are capable of increasing their performance with different data as compared to traditional machine learning where the models attain saturation at a particular point.

Hence, deep learning can be considered as a subdiscipline of machine learning, which itself is a subdiscipline of AI.

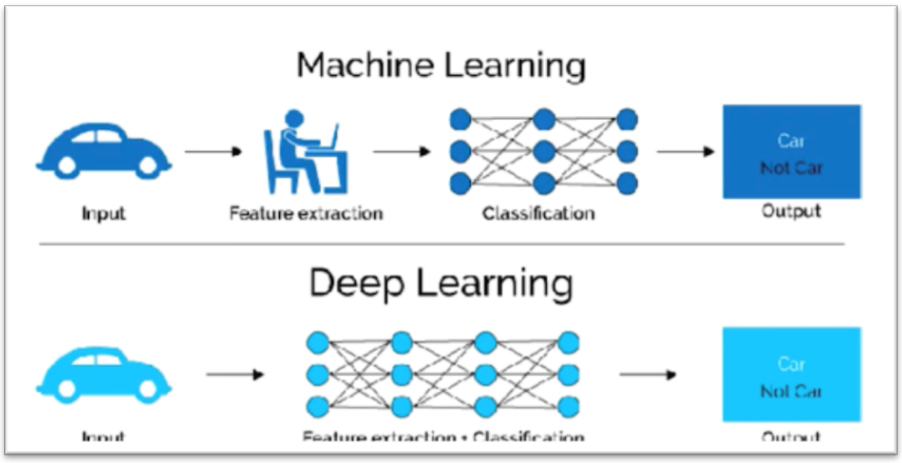

FIGURE 2.4 – *Difference between Deep Learning and Machine Learning*

2.4.2 From perceptron to deep learning

The organization of deep learning has occurred through three significant waves.

The first tide initiated with artificial neural networks with perceptron initiating in 1958 but

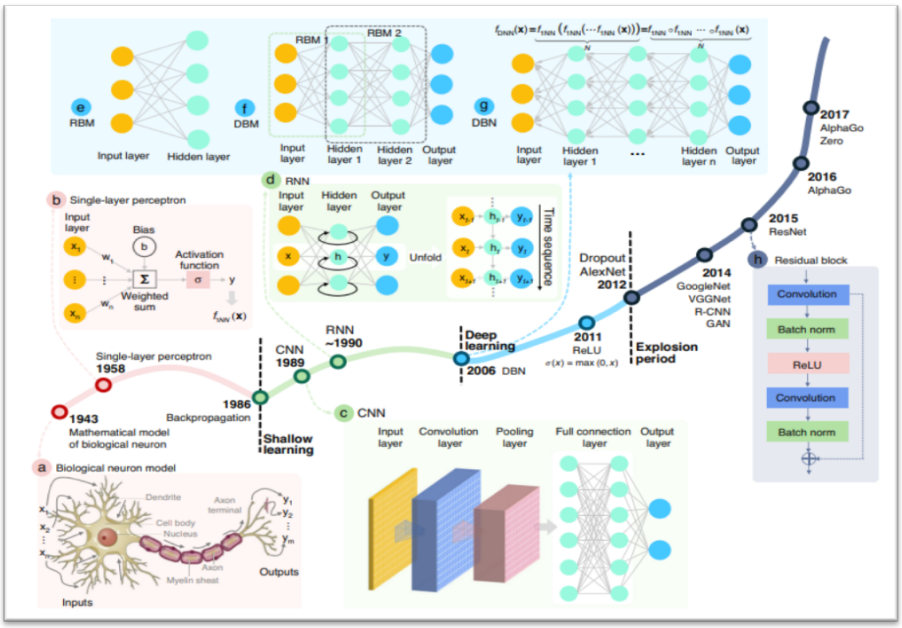

FIGURE 2.5 – *From perceptron to deep learning*

later being abandoned due to the presence of errors.

The second regime began in 1986 with the ANN backpropagation algorithm or, as it was called, "shallow learning." In 1989, LeCun et al. 's CNN fundamentally changed computer vision.

This phase also saw the development of LSTMs, RNNs, and other important models and it was also during this phase that issues surrounding model quality, limitations of deep MoEDALs, and other problems were raised.

The third wave had new features like GPU acceleration in the computer hardware and large labeled datasets in the deep learning architectures.[37]

2.5 Deep Learning techniques

This type of connection architectures has been around for over 70 years or so, but with the new emerging architectures as well as GPUs , they seem to play a significant role in AI.

Deep learning is a broad family of techniques that can be applied to many of the data problems, and now studying a fast-growing rate due to such resources as deeply layered structures, graphic cards, and big data.

This section takes a look at deep learning architectures over the past 20 years in which LSTM networks and CNNs are categorized as the earliest but most popular deep learning models. [7]

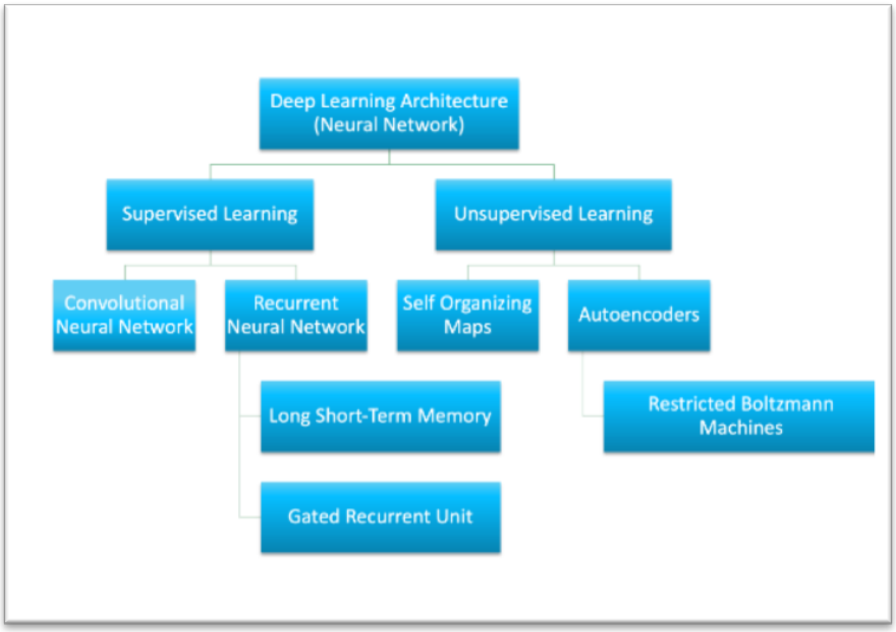

FIGURE 2.6 – *Deep Learning Architectures*

2.5.1 Supervised deep learning

Supervised learning is the problem space in which the target to be predicted is clearly labeled within the data used for training.

This section looks at some Supervised deep-learning Architectures :

1- Convolution Neural Network (CNN)

A Convolutional Neural Network (CNN) is a specific type of feedforward neural network designed to process data with a specific structure, most commonly two-dimensional data such as images. CNNs are particularly effective at feature extraction from raw data using convolutional layers. Unlike traditional feature extraction methods, CNNs can automatically learn to

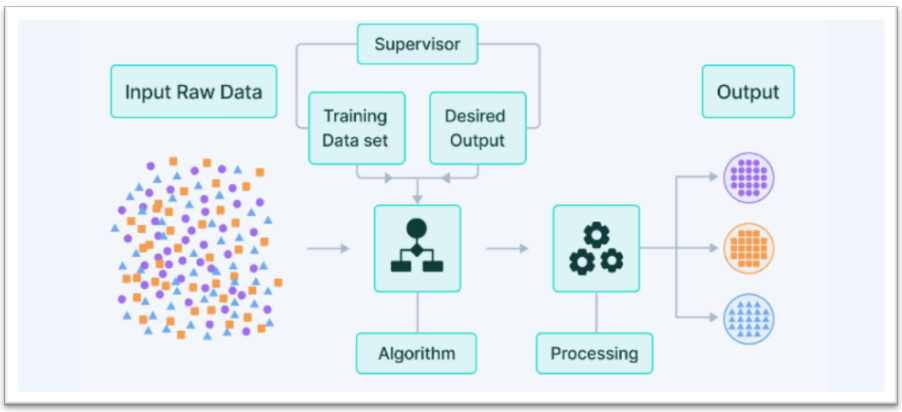

FIGURE 2.7 – *Supervised DL*

detect features from the input data.

The Architecture of CNNs is inspired by the biological vision system, processing images into abstract layers that are passed through convolutional layers. In biological terms, a neuron corresponds to an artificial neuron in a CNN, and it processes an image in a representative manner.[23]

- CNN Architecteur

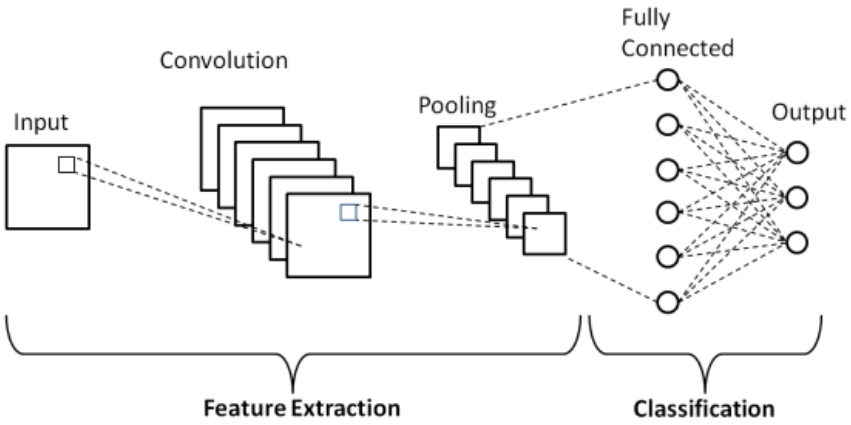

FIGURE 2.8 – *CNN Architecteur*

CNN is one of the distinguished categories of deep learning algorithms and plays a very crucial role particularly in handling the graphic data like images and videos. Like the neurobiology of visual cortex, popular CNNs employ several layers of convolutional layers using filters followed by fully connected layers in which subsampling also can be employed in-between.[9]

- Key Components of CNNs :

1. Convolutional Layer :

- The simplest module of a CNN, capturing pixel values and basic features of an input image.
- Adds the kernels to the height and weight of the input data and then computes the dot product to make a 2D activation map.
- Acquires how filters work in a way that certain characteristics elicit distinct spatial patterns at specific locations.[9]

2. Non-linearity Layer :

- Adds and subtracts input and applies nonlinear functions like ReLU, sigmoid, Tanh to the input signal for the next layer.[9]

3. Pooling Layer :

- Scales down the input to shrink the spatial dimension of the representations lowering the quantity and density of parameters and computations.
- Usually employs max pooling to avoid such a problem as overfitting and to carry out calculations more efficiently.[9]

4. Fully Connected Layer :

- Standard deep neural network layer that computes the final predictions for a decision or to a continuous variable depending on the type of problem one is solving.
- Every neuron in the output layer is linked to all the activations from the previous layer in that layer.[9]

5. Loss/Classification Layer :

- Assists in the training phases and measures the difference between the actual and predicted labels of an instance.
- Go through several loss functions (for instance, softmax, cross-entropy) depending on the problem at hand.

The CNNs are highly suited to process databases where a large number of nodes and parameters are required to be trained and are frequently applied in image processing.[9]

2.5.2 Recurrent Neural Network (RNN)

RNN is a kind of artificial neural network specialized in sequence learning. Similar to other neural networks, RNNs are made up of multiple layers with multiple weights and multiple biases.

In an RNN, the nodes are connected in a direction cyclic way, meaning the network is able to take a stream of information as illustrated in this architecture allowing RNNs to accommodate

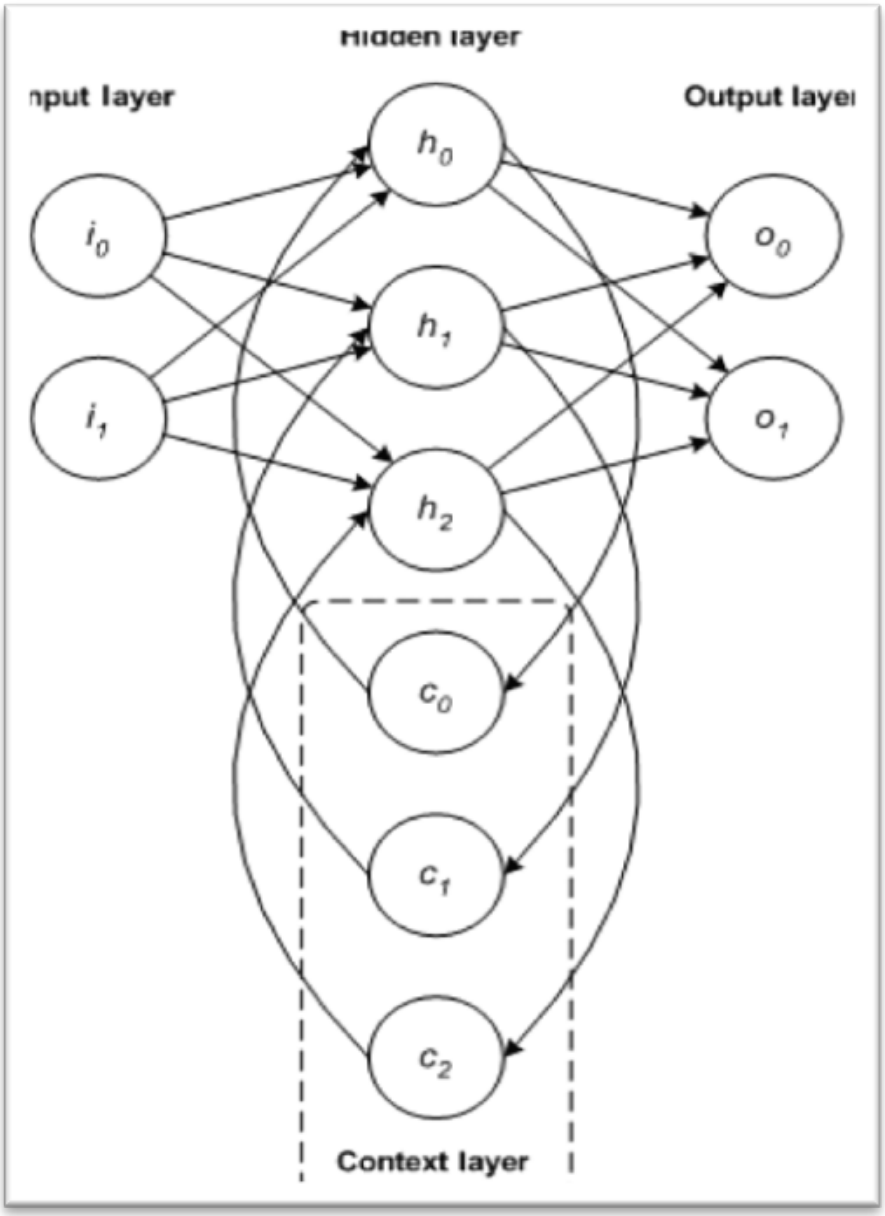

FIGURE 2.9 – *RNN Architecture*

the temporal dynamic pattern, thus applying to NL and speech recognition tasks.**Figure**2.9.

1. One to One : A single input connected to a single output, like Image classification.
2. One t many : A single input linked to output sequences, like image captioning, includes several words from a single image.
3. Many to One : Series of inputs generating single output, like Sentiment Analysis.
4. Many to many : series of inputs yielding series of outputs, like video classification
5. It is also widely used in language translation, and conversation modeling.

Example applications : speech recognition and handwriting recognition.[7]

2.5.3 LSTM networks

Proposed by Hochreiter and Schmidhuber in 1997, Long Short Term Memory (LSTM) is a technique which has recently emerged as Recurrent Neural Network (RNN) design in speech recognition used in smartphones and IBM Watson® products.

Neural networks introduces memory cells in LSTM where the memory holds an information for certain time period depending on the input given to it and this is different from the other traditional neural networks look in **figure** 2.10.
These cells have three gates :

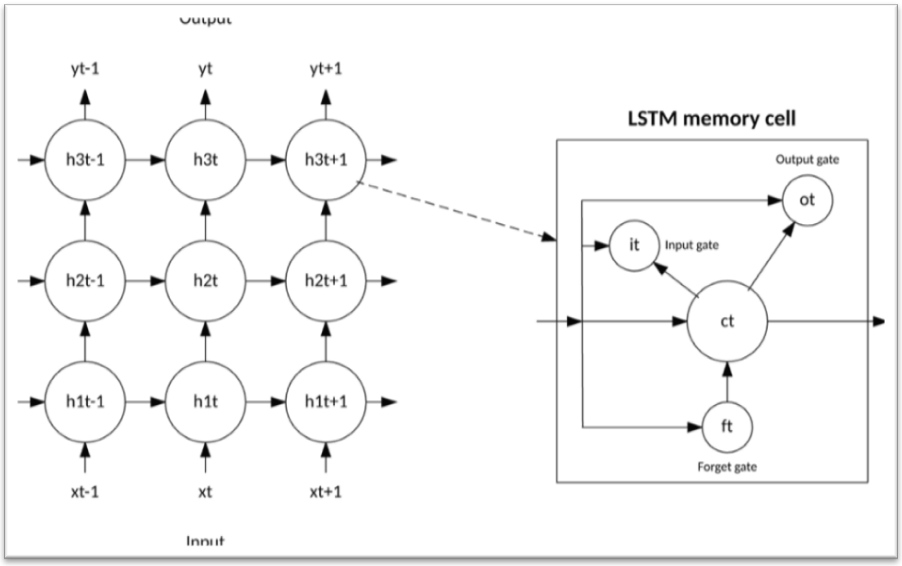

FIGURE 2.10 – *LSTM Network architecture.*

- This includes the input gate or the ability of the network to incorporate fresh information.

- the forget gate which determines the time to remove relegated information.

- the output gate which determines a time to emulsify an output.[7]

2.5.4 GRU Network

It should also be noted that one of the simplest LSTM analogs, called the gated recurrent unit, was proposed in 2014.

The full form of this model is Gated Recurrent Unit or simply GRU and this has two gates and does not have the output gate that is present in the LSTM model.

GRU is less complex than the LSTM, and it only takes a shorter amount of time to train it ; it may also be faster than the LSTM in its operations. However, with more data when deployed the LSTM is more flexible than other models and can help yield better performance ; see the **figure** 2.11

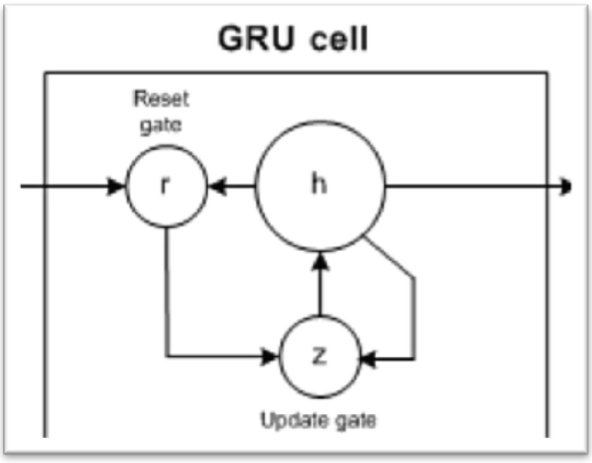

FIGURE 2.11 – *GRU Cell*

Example applications : Text and voice; recognition of handwriting and gestures; image description.[7]

2.5.5 Unsupervised deep learning

It can be described as the problem space of the learning techniques where input data carries no target label.

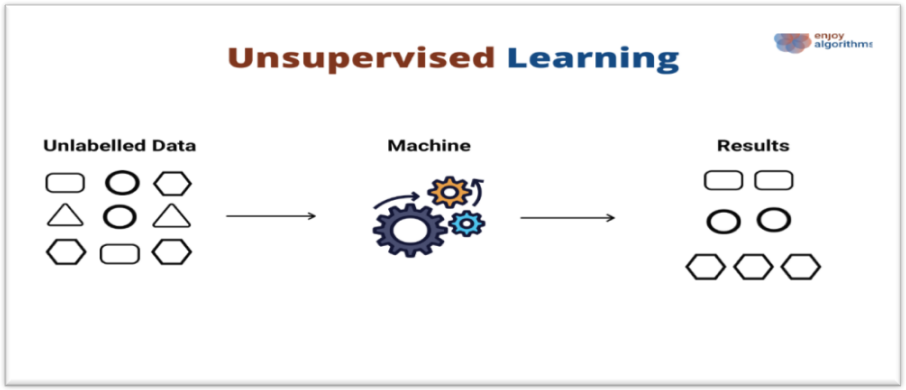

FIGURE 2.12 – *Unsupervised DL*

This section looks at some unsupervised deep learning architectures :

1. Self-organized maps

SOMs are applied to reduce the dimensionality, and categorize high-dimensional input data into a two-dimensional output, grading, and visualization.

Example applications : Dimensionality reduction, clustering high-dimensional inputs to 2-dimensional output, radiant grade result, and cluster visualization.[7]

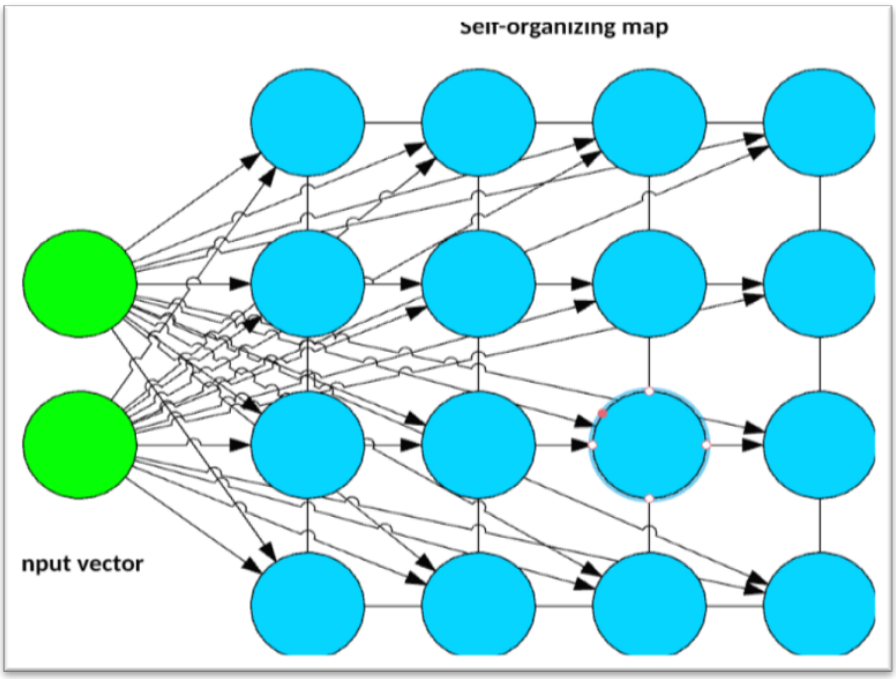

FIGURE 2.13 – *Self-Organized Technique*

2. Restricted Boltzmann Machines

Structured RBMs were introduced by Paul Smolensky in 1986; initially, he named them as Harmonium. An RBM is a neural network that has two layers; the input layer as well as the

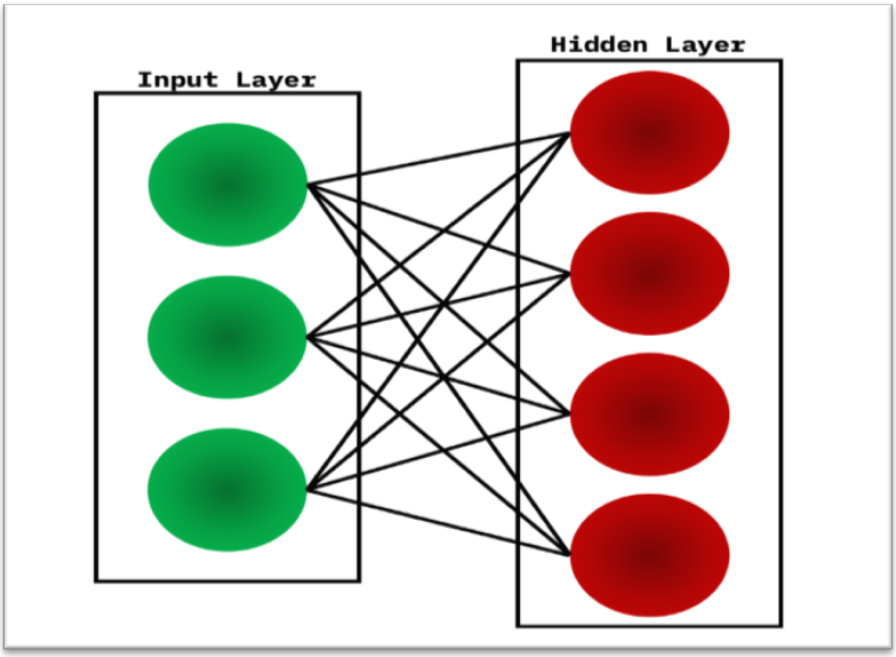

FIGURE 2.14 – *Restricted Boltzmann Machines Technique*

hidden layers and every neuron in the hidden layers connects to every neuron in the visible layers. Unlike typical Boltzmann Machines, the nodes in RBMs of one layer do not connect with each other which makes the calculation easier.

Some of the application of RBMs include ; Dimensionality reduction and Collaborative filtering.[7]

3. Deep belief networks

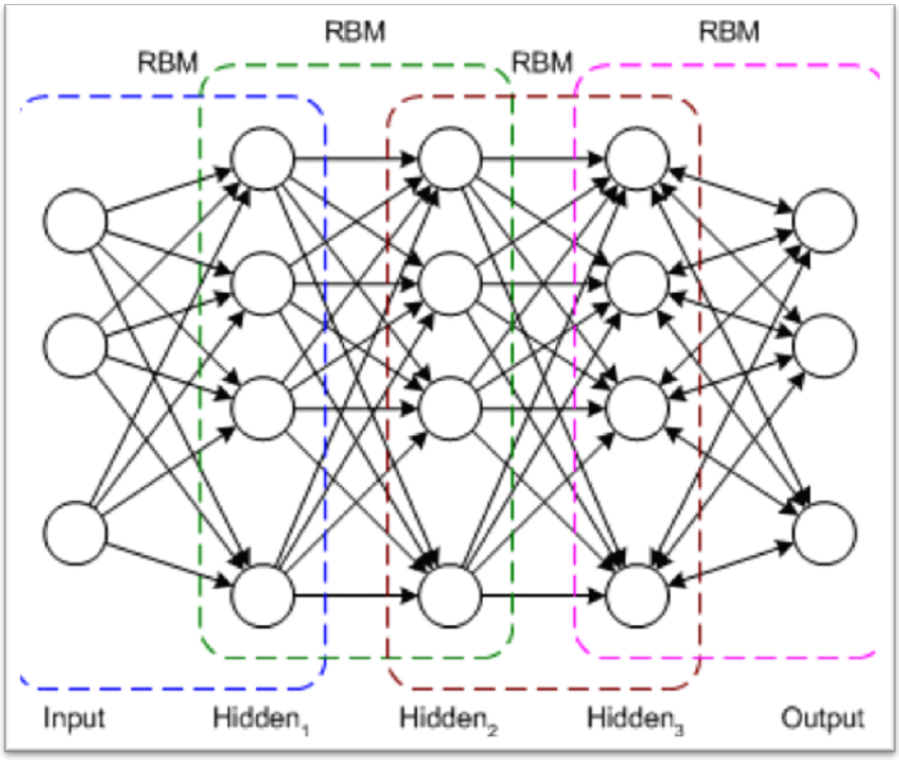

FIGURE 2.15 – *Deep belief networks Tech*

A Deep Belief Network is a kind of multilayered neural network with a distinctive feature focusing on training. In fact, what is implanted below is actually the stack of Restricted Boltzmann Machines (RBMs).

In a DBN, the first layer takes raw sensory data into the network, while the intermediate layers have to learn the hidden structures ; the last layer is of classifiers.

DBNs has been applied in image recognition information retrieval and natural language processing and the failure prediction among others.

4. Autoencoders

Originally, Autoencoders were invented by Geoffrey Hinton in 1986, as exemplar architectures of deep learning employed for unsupervised learning. They are optimal in data compression and in being able to learn useful data abstractions.

Autoencoders are flexible to process generic or image data and are normally utilized in the format of Convolutional Neural Networks (CNNs) in the process of dimensionality decrease.

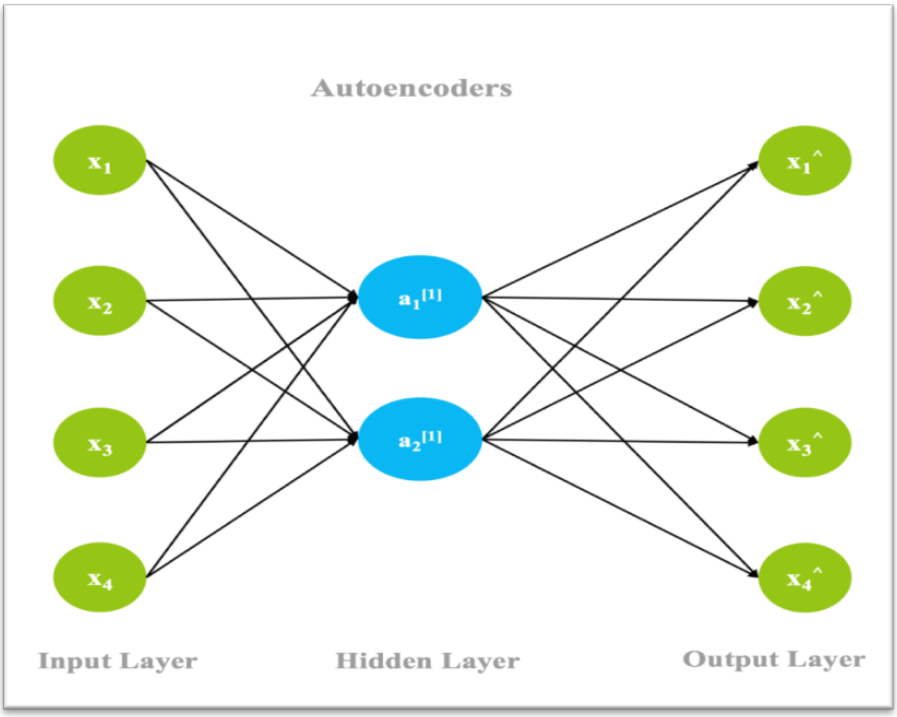

FIGURE 2.16 – *Autoencoders Technique*

In implementation, the input layer is transformed to a hidden layer by applying an encoding function and the number of nodes is smaller than that of the input layer. In this position, it holds a compressed form of the input image, which is original to this layer. The final layer of this neural network tries to predict the input layer utilizing a decoder mechanism.
[32]

During the hiding phase, the actual output from the layer is calculated using an error function in case of back-propagation, or the actual difference between the input and output layer is calculated and the weights are adjusted to minimize the error.

Autoencoders on the other hand differ from most unsupervised learning approaches since there is no set data against which the output could be compared, but rather the auto-encoders learn independently through the use of backward propagation.Because of this, AE falls under the Self-Supervised Learning Algorithms category.[32]

Autoencoder Architecture

Autoencoder in Simple Terms :

1. What's Inside ?

- An autoencoder is like a two-part machine : Since, there are two blocks in this scheme, they are referred to as the encoder and the decoder.
- The two parts are utilized when we train the autoencoder.

2. Slimming Down :

- If we wish to reduce the data i. e. compressing a large picture into a smaller one, then we only need the encoder.
- Thus, in the case of shrinking data, the only the encoder part should be employed.

3. How it Works :

- Think of the structure of autoencoder as a composition of numerous floors.
- At the bottom we find the input floor, that is where the data enters the system.
- When moving upwards through each floor, there are lesser rooms (neurons) and finally, we are led to the top floor, or the bottleneck.
- This top floor, though is not as large as the main one, contains the most valuable items.

4. Expanding Back :

- Once we get to the top floor we begin descending the building.
- Each floor (decoder layers) has more rooms (neurons) than the one directly above it until the bottom floor.
- The output floor at the bottom thus has the same number of rooms (neurons) as the input artificial floor.

Thus, the autoencoder is somewhat like a building with an elevator : data get shrunk as they are ascended through the architecture and expanded back upon descent.

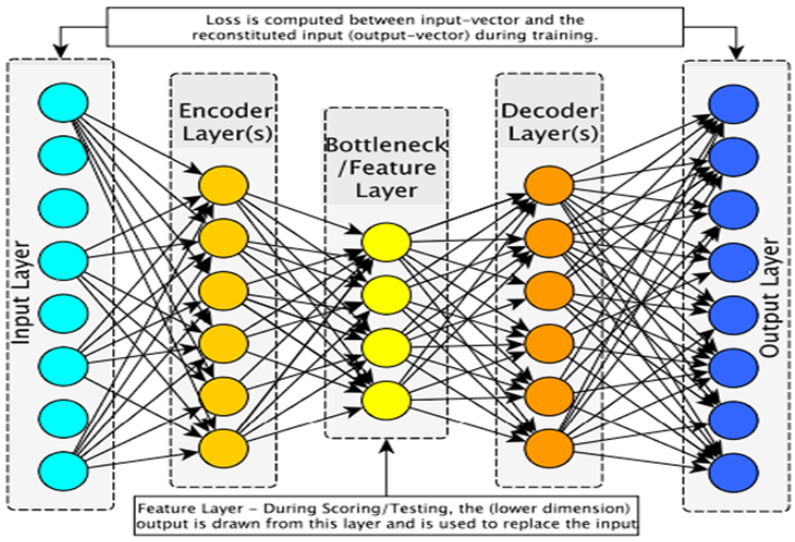

FIGURE 2.17 – *A schematic of an AE showing an encoder connected to the input layer*

Their primary use is to learn in an unsupervised manner an "informative" representation of the data for eventualities like clustering. The problem, about which I want to write is to study the functions

$$A : \mathbb{R}^n \to \mathbb{R}^p \quad \text{(encoder)} \quad \text{and} \quad B : \mathbb{R}^n \to \mathbb{R}^p \quad \text{(decoder)} \quad \text{that satisfy :}$$

$$\arg\min_{A,B} E\left[\Delta(x, B \circ A(x))\right]$$

Autoencoders may be trained end-to-end or gradually layer by layer. In the latter case, they are "stacked" together, which leads to a deeper encoder. This is done with convolutional autoencoders .[32]

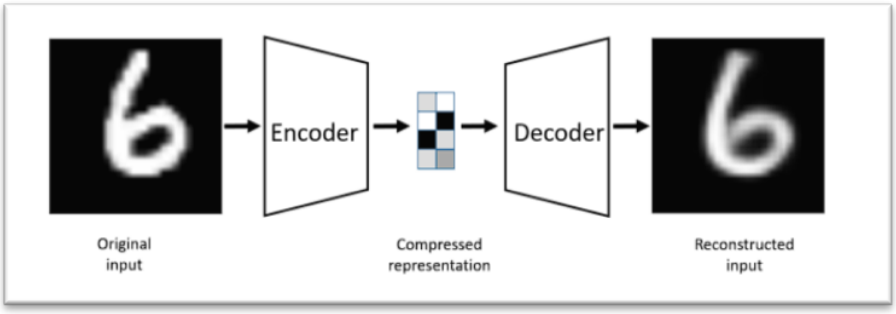

FIGURE 2.18 – *An autoencoder example. The input image is encoded to a compressed*

Loss Function

When training the Autoencoder, the main goal is to reduce the reconstitution error between the output (the last decoder layer) and the input (the first encoder layer). This ensures that even after compressing the data through the encoders, the bottleneck layers capture most of the relevant patterns in the data, allowing for effective reconstruction after decoding. For continuous input, squared error loss is commonly used as one of the loss functions for Autoencoders, given as :

$$L(x,\, x') = \|\, x \,-\, x'\|^2 = \|(x - W'((Wx + b) \,+\, b'\|^2$$

where,
x is the input,
W, b is the weight vector and bias of the encoder, and
W', b are the weight vector and bias of the decoder.

When considered in batches of n samples, the mean square error for the batch could be computed by the formulae.

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^{n} L(xi, x'i)$$

We train a mono-layer neural network, and the hidden value h is predicted by the output z of the input x. This is a task called reconstruction .
- Therefore, the criterion is to minimize the reconstruction error L(x ;z).
- The hidden layer must contain information relating to the reconstruction.
For example, if the hidden layer contains fewer units than the input, so to properly reconstruct, it must learn to summarize the entry, as all of the necessary information is contained in the hidden layer. As a result, characteristics relevant to the reconstruction must be extracted.[36]

During training, the Autoencoder aims to minimize the reconstitution error between the output and input layers. This ensures that the bottleneck layers effectively capture the relevant patterns in the data for reconstruction. The squared error loss function is commonly used, and

the mean square error for a batch can be computed accordingly.

In training a single-layer neural network, the hidden value is predicted by the output of the input. This is known as reconstruction, and the objective is to minimize the reconstruction error. Therefore, the hidden layer must contain information necessary for accurate reconstruction, even if it has fewer units than the input.

Before training an autoencoder, four hyperparameters need to be set.Four hyper-parameters should be set before training an auto-encoder :

Training Autoencoder

1) Code size :

The code size or the size of the bottleneck, it is the most important hyper-parameter used to tune the auto-encoder. The bottleneck size decides how much the data has to be compressed. This can also act as a regularisation term.

2) Number of layers :

Like all neural networks, an important hyper-parameter to tune auto-encoders is the depth of the encoder and the decoder. While a higher depth increases model complexity, a lower depth is faster to process.

3) Number of nodes per layer :

The number of nodes per layer defines the weights we use per layer. Typically, the nodes number decreases with each subsequent layer in the auto-encoder as the input to each layer becomes smaller across the layers.

4) Reconstruction Loss :

The loss function we used to train the auto-encoder depends on the type of input and output we want the auto-encoder to adapt.

The most popular loss functions for the reconstruction of image data are mean squared error (MSE) Loss and L1 Loss. If the inputs and outputs are within the range [0,1], as in MNIST [Den12] (DATABASE of handwritten digits), Binary Cross Entropy can also be used as the reconstruction loss.

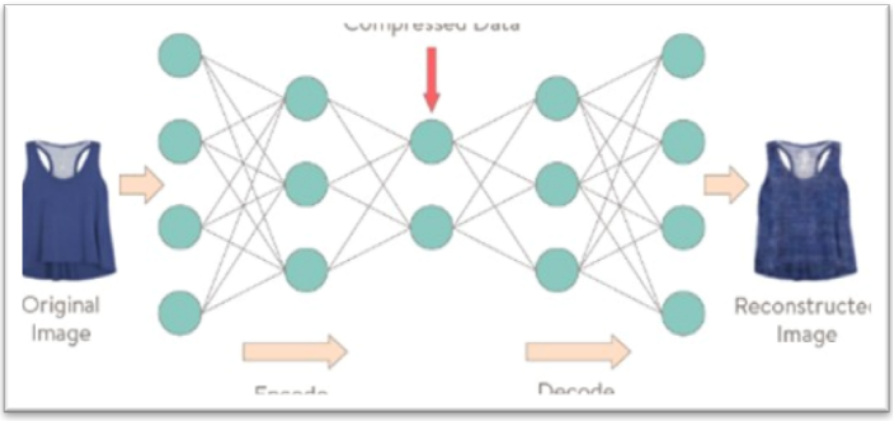

Figure 2.19 – *Principle of Autoencoder on image.*

Type of Autoencoders

There are many types of Autoencoders, ranging from simple to denoising autoencoders used for various purposes. Here are the popular autoencoders :

1. Undercomplete autoencoders.
2. Sparse autoencoders.
3. Contractive autoencoders.
4. Denoising autoencoders.
5. Deep convolutional autoencoder.

a- Undercomplete Autoencoders

Undercomplete Autoencoder is one of the simplest types of Autoencoders. It is an Autoencoder whose code dimension is less than the input dimension.

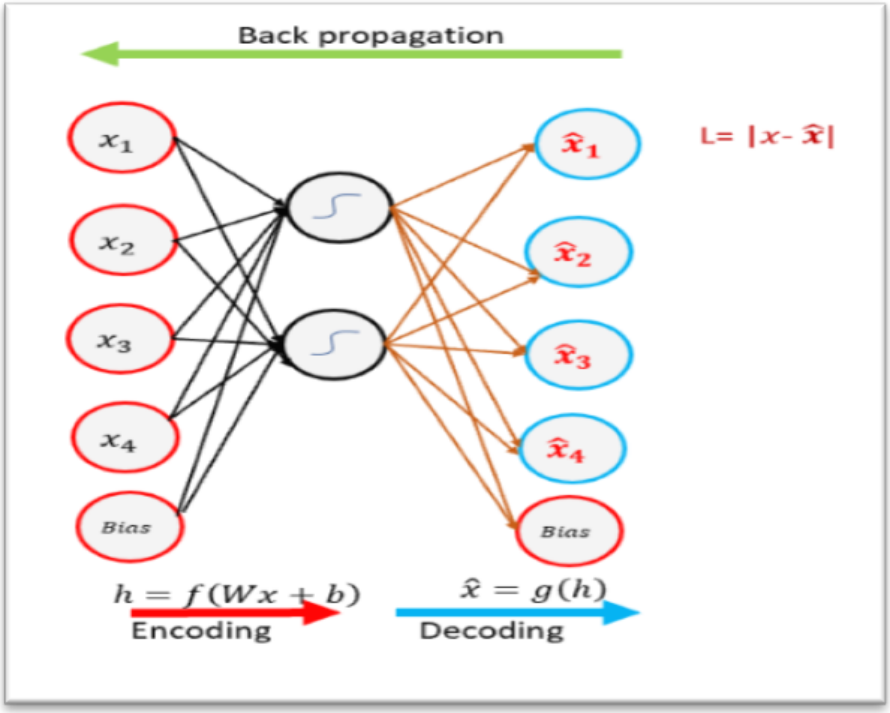

FIGURE 2.20 – *Undercomplete Autoencoder- Hidden layer has smaller dimension than input layer*

As the latent code has far fewer neurons than the input, we force the encoder to compress the data. If the input features were all independent from one another, this dimensionality reduction and subsequent reconstruction would be a tough task.

However, if there is some underlying structure in the data , this can be learned and consequently leveraged when forcing the input through the latent code.

In the undercomplete autoencoder, the encoder essentially projects the data into a lower-dimensional space (performing the task of dimensionality reduction). Indeed, if an autoencoder uses the linear activation function in combination with the mean squared error loss function, the resulting model tends to perform equally to the PCA algorithm.

Learning an undercomplete representation forces the autoencoder to capture the most salient features of the training data.

The way it works is very straightforward, Undercomplete autoencoder takes in an image and tries to predict the same image as output, thus reconstructing the image from the compressed bottleneck (code) region.[32]

Objective is to minimize the loss function by penalizing the g(f(x)) for being different from the input x.

$$\mathbf{L = |x + x\text{^}|}$$
$$\mathbf{L = |x - g(f(x))|}$$

b- Sparse Autoencoder : 2.21

Sparse Autoencoders are similar to the undercomplete autoencoders in that they use the same image as input and ground truth. However, the means by which the encoding of information is regulated is significantly different. The sparse autoencoder is regulated by changing the number of nodes at each hidden layer[32].

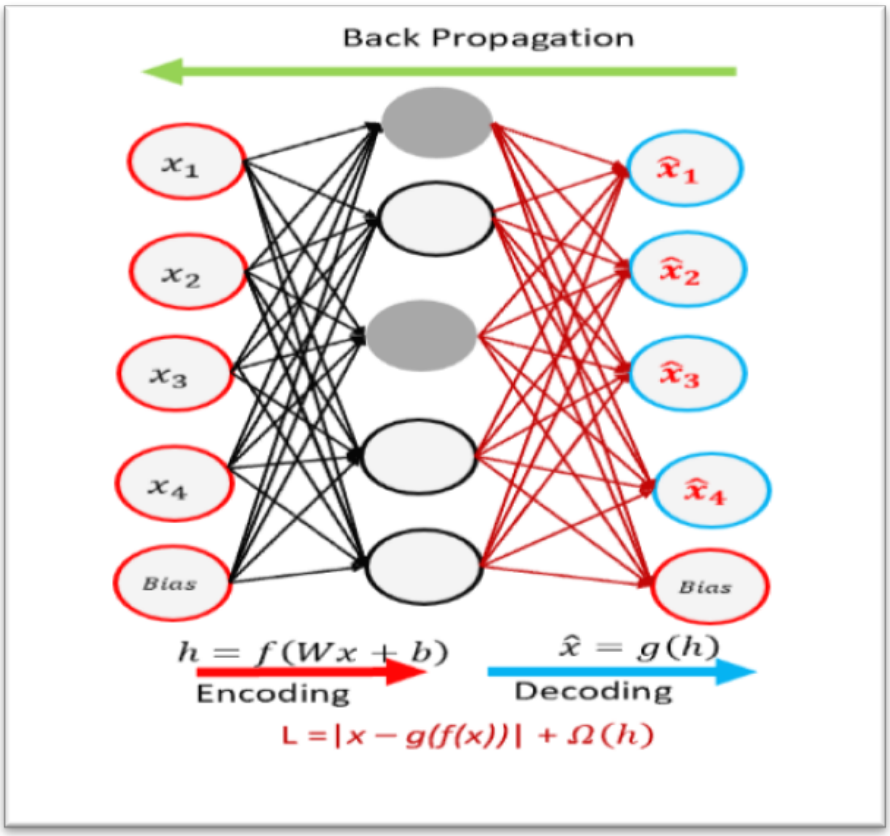

FIGURE 2.21 – *Sparse Autoencoder*

Sparse autoencoders have a sparsity penalty, (h), a value close to zero but not zero. Sparsity penalty is applied on the hidden layer in addition to the reconstruction er- ror. This prevents overfitting. [14]

$$L = |x - g(f(x))| + \Omega(h)$$

Sparse autoencoders take the highest activation values in the hidden layer and zero out the rest of the hidden nodes. This prevents autoencoders to use all of the hidden nodes at a time and forcing only a reduced number of hidden nodes to be used. As we activate and inactivate hidden nodes for each row in the dataset. Each hidden node extracts a feature from the data.[14]

c- Contractive Autoencoder :2.22

Similar to other autoencoders, contractive autoencoders perform task of learning a representation of the image while passing it through a bottleneck and reconstructing it in the decoder.

The contractive autoencoder also has a regularization term to prevent the network from learning the identity function and mapping input into the output.

Contractive autoencoders work on the basis that similar inputs should have similar encoding and a similar latent space representation. It means that the latent space should not vary by a considerable amount for minor variations in the input [36].

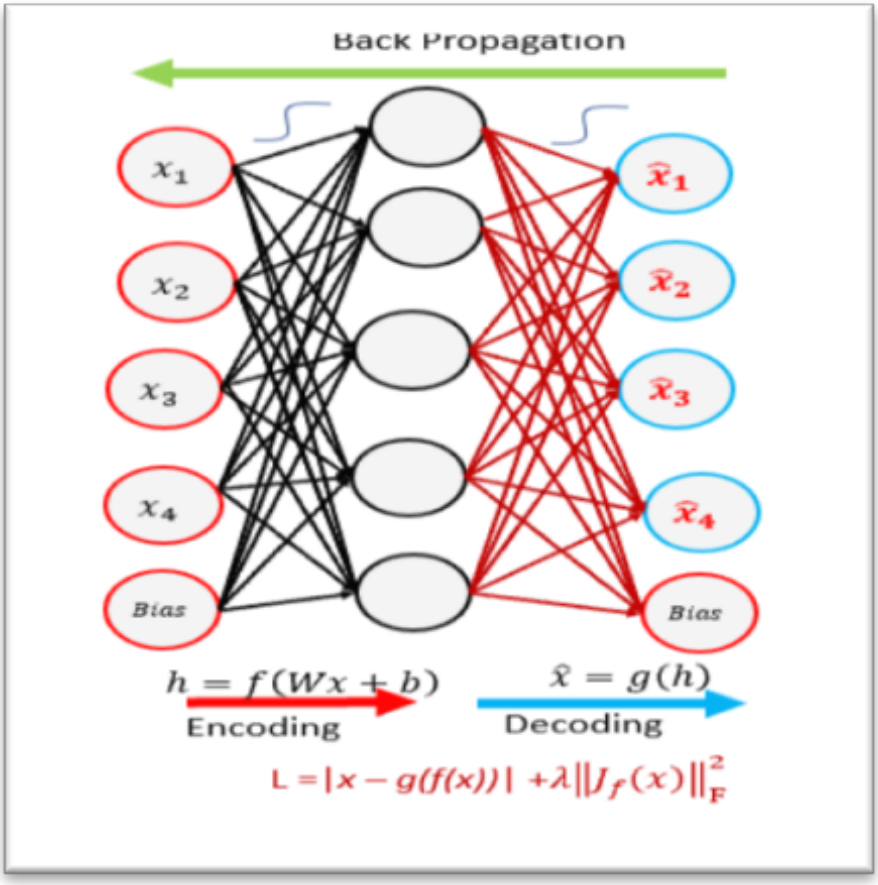

FIGURE 2.22 – *Contractive Autoencoders*

Robustness of the representation for the data is done by applying a penalty term to the loss function.

Contractive Autoencoders surpasses results obtained by regularizing autoencoder using weight decay or by denoising. CAE is a better choice than denoising autoencoder to learn useful feature extraction.

Penalty term generates mapping which are strongly contracting the data and hence the name contractive autoencoder.[14]

d- Denoising Autoencoder :

The denoising autoencoder (DAE) is an autoencoder that receives a corrupted data point as input and is trained to predict the original, uncorrupted data point as its output. Autoencoders that remove noise from an image. [4]

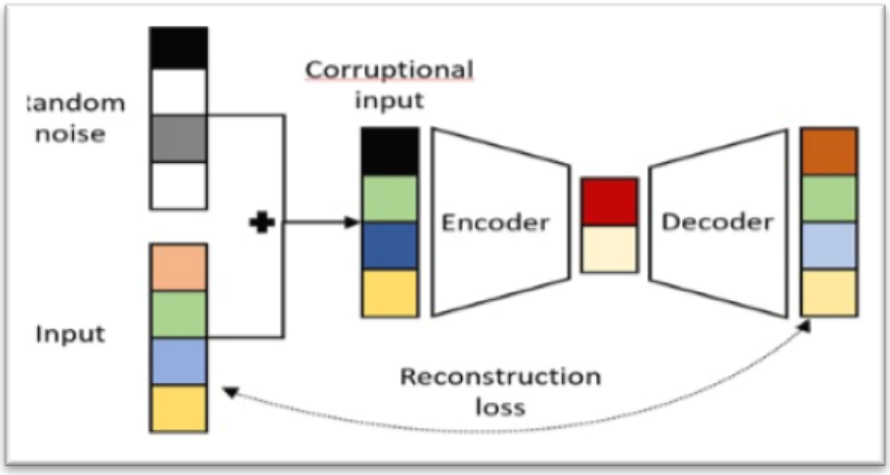

FIGURE 2.23 – *Denoising Autoencoders Principle*

In denoising autoencoders, they feed a noisy version of the image, where noise has been added via digital alterations. The noisy image is fed to the encoder-decoder architecture, and the output is compared with the ground truth image.
[14]

The DE removes noise by learning a representation of the input where the noise can be filtered out. 2.24

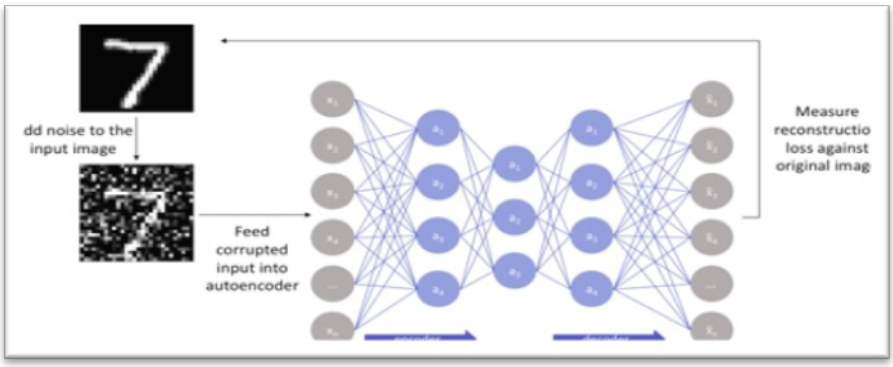

FIGURE 2.24 – *Denoising Autoencoder example.*

e- Deep convolutional Autoencoder :

Convolutional Autoencoder is a variant of Convolutional Neural Networks that are used as the tools for unsupervised learning of convolution filters. [21]

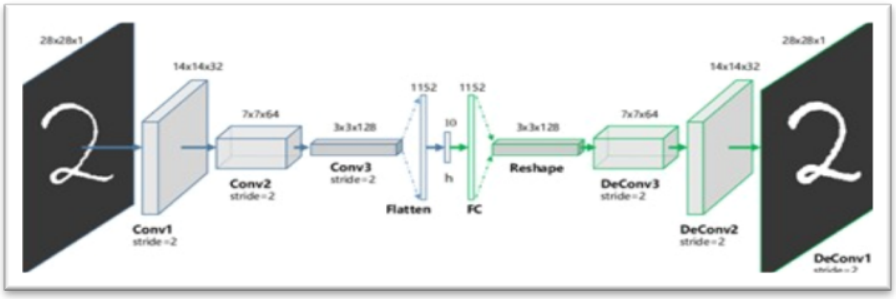

FIGURE 2.25 – *The structure of Convolutional Autoencoder.*

They are generally applied in the task of image reconstruction to minimize reconstruction errors by learning the optimal filters. Once they are trained in this task, they can be applied to any input in order to extract features.[21]

2.6 Application of Autoencoders

a- Dimensionality reduction :

Undercomplete autoencoders are those that are used for dimensionality reduction. These can be used as a preprocessing step for dimensionality reduction as they can perform fast and accurate dimensionality reductions without losing much information. Furthermore, while dimensionality reduction procedures like PCA can only perform linear dimensionality reductions, undercomplete autoencoders can perform large-scale non-linear dimensionality reductions. [32]

b- Image denoising :

Autoencoders like the denoising autoencoder can be used for performing efficient and highly accurate image denoising. Unlike traditional methods of denoising, autoencoders do not search for noise, they extract the image from the noisy data that has been fed to them via learning a representation of it. The representation is then decompressed to form a noise-free image.

Denoising autoencoders thus can denoise complex images that cannot be denoised via traditional methods. [32]

c- Generation of image and time-series data :

Variational Autoencoders can be used to generate both image and time series data. The parameterized distribution at the bottleneck of the autoencoder can be randomly sampled to generate discrete values for latent attributes, which can then be forwarded to the decoder, leading to of image data generation.VAEs can also be used to model time series data like music.[32]

d- Anomaly Detection :

Anomaly detection is another unsupervised task, where the objective is to learn a normal profile given only the normal data examples and then identify the samples not conforming to the normal profile as anomalies.This can be applied in different applications such as fraud detection, system monitoring,...etc.[32]

e- Image distorted :

Deep Autoencoders, especially the RDONet as an example of hierarchical compressive Autoencoder, are rather effective in dealing with distorted images and offer high functionality and accuracy.

These Autoencoders do not try to fix the distortions as conventional Approaches, but learn the image in a compressed form from the distorted data. This representation is then expanded to get an improved and more representative image than the compressed one. While encoding, RDONet brings the best balance between rate and distortion on the same frame while maintaining the vital details despite distortions. This is useful since, with this approach, Autoencoders can be able to handle the distortions well while preserving the quality and integrity of the original image.[6]

Conclusion

In conclusion, The importance of Autoencoders is discussed as a key element of deep learning. There, we have discussed the history of this technology and how it has managed to make so much progress across several sectors especially where the extraction and analysis of features from the visual data is of importance.

Convolutional Autoencoders may be used to improve low-quality pictures, which opens up new possibilities for the implementation of face and other biometric identification systems. The result not only improves images but, especially, it works to recover critical information from images that have changed, opening up new areas for research and creativity in the field of electronic security.

Face Recognition

3

Sommaire

3.1 Introduction

This chapter examines the possibilities and difficulties of facial recognition technology. In the first section of the chapter, face recognition technology is defined as a biometric technique that uses distinctive facial traits to identify individuals. It then discusses the difficulties, such as different poses, different lighting, different facial emotions, aging, occlusions, similar faces, and picture accuracy. It also emphasizes the benefits like effectiveness, biometric security, and a wide range of applications as well as the drawbacks like the requirement for big databases and challenges with recognition in certain scenarios.

In conclusion, it provides an explanation of the steps involved in using facial recognition technology, from face detection and analysis to digital data conversion.

3.2 What is facial recognition ?

Facial recognition is a biometric identification technique that uses facial features for identity verification. It has gained importance in various fields such as security, biometric identification, and smart applications. Some key points about facial recognition are :

1. Facial recognition utilizes facial landmarks, texture, and shape for identity verification.
2. It is widely used in security systems for access control and surveillance.
3. Facial recognition has applications in biometric identification, where it can be used to verify the identity of individuals in databases.
4. It is also used in smart applications like facial recognition-based authentication on smartphones and personalized marketing.
5. However, facial recognition can be affected by variations caused by facial expressions and low-light conditions.

Overall, facial recognition plays a crucial role in enhancing security measures, improving biometric identification systems, and enabling smart applications.[26]

3.3 The main applications of face recognition technologies :

The main applications of face recognition technologies include :

1. Identity verification : Face recognition can be used to verify a person's identity, such as for access control to secure areas or for authentication purposes.

2. Surveillance and security : Face recognition can be used in surveillance systems to identify individuals in real-time or to search for specific individuals in large databases.

3. Personalized experiences : Face recognition can be used to personalize experiences, such as in retail settings where it can be used to offer personalized recommendations or targeted advertisements based on a person's identity.

4. Law enforcement : Face recognition can be used by law enforcement agencies to identify suspects or to match faces in surveillance footage with known individuals.

5. Attendance and access management : Face recognition can be used for attendance tracking in schools or workplaces, as well as for access management to restricted areas. Overall,

face recognition technologies have a wide range of applications in various industries and sectors, offering convenience, security, and efficiency.[8]

3.4 Challenges in recognizing distorted faces :

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

Intrinsic factors :

It is due to the physical nature of the face and not depende nt on the observer. Intrinsic factors are fu rther divided into intrapersonal and interpersonal.
Intrapersonal is caused due to variation in face appearance of an indi vidual, for example ageing, facial expression and facial paraphernalia (facial hair, cosmetics, glasses etc.) .

Extrinsic factors :

This is caused due to the variation in face appearance due to the interaction of light with the face and the observer. This will include I illumination, pose, scale and imaging parameters (resolution, focus, imaging, noise etc.).

Following are the common challenges seen in face recognition system can have while detecting a face :[25]

3.4.1 Pose variation :

Variation in pose causes significant problems in detecting a face. Pose variation can be due to change in observing angle of the observer and also due to rotation in the head position. These variations can cause a serious problem in identifying the input image. Many of the systems can tolerate small variations such as small rotations in angles.
But it will be difficult when it co mes to large rotational angles. The database usually consists of face images of frontal view of the faces. Since the existing FRSs are very sens itive to pose variation, pose correction is essential and could be achieved by means of efficient techniques aim ing to rotate the face and/or to align it to the image's axis. [25]

FIGURE 3.1 – *Pose variation*

3.4.2 Variation in illumination :

Variations of illuminations could reduce the efficiency of FRS. For moderate levels of lighting of the background, face detection and recognition are much difficult to perform. Variation in illumination can vary the total magnitude of light intensity being reflected back from an object. On the other hand, higher light levels could lead to over-exposure of the face and (partially) und detectable facial patterns.
There have been many algorithms such as equalization techniques that are available now to get rid of this problem to an extent. Sometimes even multiple algorithms can be used in aface recognition system to tolerate the issue of illume nation. But in case of extents, it is not desirable to depend on these techniques. [25]

FIGURE 3.2 – *Variation in illumination*

3.4.3 Variation in expression :

Some variation in the face images can be caused due to difference in expression influenced by the individual's state of emotion.
Therefore, it is important to recognize different facial expressions for evaluating the emotional

state. Human expressions consist of macro-expressions such as, disgust, anger, happiness, fear, sadness or surprise, and other involuntary, rapid facial patterns.
These facial changes can be computed with the help of dense optical flow. Cosmetics and hair styles canalso be included in this challenge as changing hair style and putting make-up canalso cause variation in facial expression. [25]

FIGURE 3.3 – *Variation in expression*

3.4.4 Ageing :

Another reason for the changes in the appearance of the face could be the aging of the human face and could affect the entire process of face recognition ; if the time between each image capture is large, there will be significant changes in the person.
As per various study conducted by scientists, in every 10 years there will be significant changes in an individual's face appearance. The fig. 5 shows the change in an individual's face at different ages. It is not just the shape and lines of a face that gets modified overtime ; there will be changes in hairstyles as well. [25]

FIGURE 3.4 – *Enter Caption*

3.4.5 Occlusions :

Variation in facial appearance can also be caused due to presence of objects that such as occlusion that partially cover the face. This makes it a difficult task for the system to classify the image. Although the face is found, it may be difficult to recognize it due to some hidden facial parts, making it difficult to recognize features.
This challenge can b e seen in real world application where acquiring persons talking on the phone, wearing glasses, scarf, hats etc or having their faces covered with hands. [25]

FIGURE 3.5 – *Occlusions*

3.4.6 Similar Faces :

This is usually a not so common challenge. But we have seen that even humans find it difficult to identify people with similar faces. Hence we can imagine the difficult situation for

computer to identify similar face individuals.
Especially identical twins with similar facial features, shape etc. this becomes a difficu lt task for the face recognition system to identify the individual. This will cause an increase in false re cognition rate (FRR) as well. [25]

FIGURE 3.6 – *Similar Faces*

3.4.7 Image Resolution :

Another important issue with face recognition system is the varying quality and resolution of the images given as input. Many factors can affect the resolution of an image. The environment, the performance quality of the acquiring system and many other reasons can be mentioned as factors that are responsible for varying resolution of the image. If the resolution is good, then the recognition process will be much easier and efficient.
So we can say that resolution is directly proportional to the efficiency of the face recognition system. [25]

FIGURE 3.7 – *Image Resolution*

3.5 Advantages and disadvantages of facial recognition

3.5.1 Advantages of Facial Recognition :

Facial recognition technology offers several advantages, including :[28]

Efficient Identification :

Facial recognition allows for quick and accurate identification of individuals based on their facial features, eliminating the need for complex passwords or fingerprints.

Biometric Security :

It serves as a form of biometric security that is intelligent, rapid, and reliable, enhancing security measures in various applications.

Versatile Applications :

Facial recognition technology can be applied in a wide range of contexts, including security systems, access control, surveillance, and even in healthcare and educational sectors.

Advanced Techniques :

The use of techniques like multi-task cascaded convolutional neural networks (MTCN) and FaceNet for feature extraction and verification contributes to improving the accuracy and effectiveness of facial recognition systems.

3.5.2 Disadvantages of face recognition :

Facial recognition technology, while highly advanced and widely used, faces several disadvantages. Some of the key drawbacks include :[28]

• Difficulty in identifying faces behind masks, glasses, or other obstructions :

Facial recognition models often struggle to accurately identify individuals when their faces are partially obscured. This limitation can impact the overall accuracy and effectiveness of facial recognition systems.

• Requirement for a large number of datasets and substantial storage space :

Implementing facial recognition technology effectively requires a significant amount of data for training and storage. This can lead to challenges in managing and processing the vast datasets needed for accurate facial recognition.

• Low accuracy levels due to challenges like illumination, face posture, and partial occlusion :

Factors such as varying lighting conditions, facial angles, and partial obstructions can significantly reduce the accuracy of facial recognition systems. Overcoming these challenges is crucial for improving the overall performance of facial recognition technology.

• Impact of external factors like the COVID-19 pandemic :

With the increasing use of masks for infection prevention, automatic facial recognition systems may face additional difficulties in accurately identifying individuals. This external factor can further hinder the effectiveness of facial recognition technology.

3.6 How Facial Recognition Works :

Facial recognition is a process that involves identifying and verifying a person's identity based on their facial features. The process typically involves the following steps : [22]

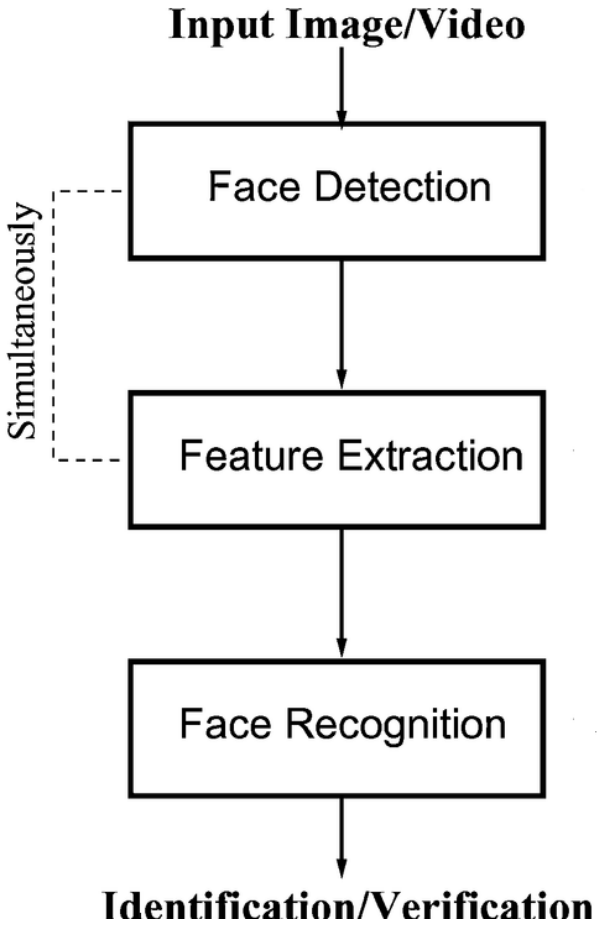

Figure 3.8 – *How Facial Recognition Works*

3.6.1 Face detection :

In this step, algorithms are used to detect and locate faces in an image or video. This is done by analyzing patterns and features such as the shape of the face, the position of the eyes, nose, and mouth.

3.6.2 Feature extraction :

Once the face is detected, algorithms extract specific facial features that are unique to each individual. These features can include the distance between the eyes, the shape of the nose, and

the contours of the face. Deep neural networks are often used in this step to learn and extract these features automatically.

3.6.3 Face matching :

In this step, the extracted facial features are compared to a database of known faces to find a match. Deep neural networks can be used to compare the extracted features with the features stored in the database and determine the similarity between them.

3.6.4 Verification or identification :

Depending on the application, facial recognition can be used for verification or identification purposes. In verification, the system compares the extracted features with a single reference image to determine if they match. In identification, the system compares the extracted features with multiple reference images to find a match.

In addition to deep neural networks, other algorithms and techniques such as biometric fingerprinting can also be used in facial recognition. Biometric fingerprinting involves capturing and analyzing unique patterns in the skin ridges and valleys of a person's face. This technique can enhance the accuracy and reliability of facial recognition systems.

3.7 The use of deep neural networks in facial recognition :

Deep neural networks, such as Convolutional Neural Networks (CNNs), can be used to improve the accuracy of facial recognition and distinguish between people in the following ways : [19]

3.7.1 Feature Extraction :

CNNs are capable of extracting high-level features from facial images, such as facial landmarks, textures, and patterns. These features can be used to create a unique representation of each individual's face, enabling accurate recognition and distinguishing between different people.

3.7.2 Discrimination Enhancement :

By training CNNs on a large dataset of labeled face images, the network can learn to differentiate between different individuals based on their unique facial features. The network can learn to identify subtle differences in facial attributes, such as gender, race, and age, which can further enhance the accuracy of facial recognition.

3.7.3 Shared Learning Parameters :

CNNs can be trained to simultaneously predict facial attributes and recognize face images using shared learning parameters. This shared learning allows the network to leverage the information from both tasks to improve overall prediction performance. By jointly training on

facial attributes and face recognition, the network can learn to better distinguish between individuals.

3.7.4 Auxiliary Information :

Facial attributes, such as gender and race, can be used as auxiliary information to assist CNN features extracted from face images. By concatenating facial attribute features with face features, the discrimination of the CNN for face recognition can be increased. This process helps the CNN classifier better recognize face images and improves the accuracy of facial recognition.

Overall, deep neural networks, particularly CNNs, offer a powerful approach to improving the accuracy of facial recognition and distinguishing between people. By leveraging the capabilities of CNNs for feature extraction, discrimination enhancement, shared learning parameters, and auxiliary information, facial recognition systems can achieve higher accuracy and reliability.

Conclusion :

This chapter determined that although facial reputation technology has several complex issues that impact its overall effectiveness, it is a strong tool for enhancing protection and identity verification.

This chapter covered the concept of facial recognition and the challenging scenarios that affect the machine's accuracy, including position variations, lighting, facial emotions, aging, occlusions, and similar faces.

We also looked at the advantages, such speed, effectiveness, and biometric security, as well as the drawbacks, like the requirement for a large number of data and issues identifying partially covered faces or in dimly lit environments. By elucidating the mechanisms underlying this generation's functioning, it will become evident that continual improvements to the algorithms and techniques used to overcome those challenges and ensure better accuracy and effectiveness.

Proposed Method And Experimental Results

4

Sommaire

4.1 Introduction

This general Approach applied for the distorted faces images analysis is described in this chapter within the context of Convolutional Autoencoder and VGG16 classifier.

In this section, we first introduce the used data source and provide some basic information about the dataset ; Then, the main preprocessing steps which has been conducted in this study will be afforded.

Therefore, the expected outcome of the current research is to establish that there is a possibility of improving the quality of the available images, as well as the classification accuracy, by adopting the proposed method.

4.2 Dataset

The dataset used in this study consists of facial images from two categories : clean and distorted . These images were collected and preprocessed for the sake of simulating the real-life situations of the facial images degradation, for instance because of noise or occlusion, or because of the lossy compression.

— **Properties** : The images are grayscale and resized to a fixed dimension of 256x256 pixels.
— **Number of Images** : The dataset is divided into training and testing sets :

a- Training Set : Contains (train-clean) clean images and (train-noisy) distorted images.
b- Testing Set : Contains (test-clean) clean images and (test-noisy) distorted images.

FIGURE 4.1 – *Original faces images*

4.3 Database Division

For the “Convolutional Autoencoder Model” The dataset is divided into Two parts :

• The training images (training set), will be used to train the AE. This represents 280 images

FIGURE 4.2 – *Distorted faces images*

of real Personnes.

• The testing images (testing set), will be used to test .This represents 140 images of real Personnes

	Train	Test
Original faces	280	140
Distorted faces	280	140

TABLE 4.1 – *Model Dataset*

4.4 Dataset Preprocessing

Preprocessing methods were used to normalize the dataset and get it ready for model training. These steps include :

• **Normalization** :

All images were normalized by scaling pixel values to the range [0, 1].

• **Resizing** :

Images were resized to 256x256 pixels to ensure consistency in input dimensions.

• **Data Augmentation** :

ImageDataGenerator was applied to add to the training data, including rescaling, in order improve its variety.

4.5 Environment

For the implementation of the proposed methodology in my thesis, We used Google Colab that is a cloud computing platform.

4.5.1 Google Colab Environment :

- Google Colab can be considered as a cloud-based environment for Python development that offers the utilization of computing resources, such as GPUs and TPUs, which are useful for training deep learning algorithms.[27]

4.5.2 Python Libraries :

- **NumPy :** Used for performing numeric calculations, including operations with arrays of many dimensions.[17]

- **Matplotlib :** Used for making graphics, plotting graphs, as well as displaying images.[20]

- **TensorFlow :** It is utilised as the basis for constructing and training the deep learning models for the definition of and training of the Autoencoder.[29]

- **Keras :** Used as an advanced level of the deep learning library operating on TensorFlow with the simple interface for constructing the neural networks.[18]

4.5.3 Other Libraries and Modules :

- **Tensorflow keras preprocessing image :** Used to import images and convert them into a loading and preprocessing format.[33]

- **Tensorflow keras layers :** Employed in creating architectural layers of neural networks.[33]

- **Tensorflow keras models :** Used in defining and compiling the neural network models as imported.[33]

- **Tensorflow keras callbacks :** It is used in the training of models for callbacks including ModelCheckpoint and other callbacks for displaying images.[33]

- **Tensorflow keras applications VGG16 :** Imported for the purposes of utilizing the pre-trained VGG16 model as the base for the classifier architecture.[33]

- **Tensorflow keras optimizers :** Essential when you want to set optimization algorithms while compiling your model.[33]

- **Tensorflow keras preprocessing image ImageDataGenerator :** Used for data enhancement and in producing batches of image data for training and validation purposes.[33]

- **PIL Image** : Used in working with images for instance, imported for the purpose of displaying image files.[15]

4.5.4 Google Colab-Specific Modules :

- **Google colab :** Used for traversing files and resources within the context of the Google Colab environment.[5]

- **Google colab files :** Used for copying files from local system to Colab setting or vice versa.[5]

4.5.5 Operating System Interaction :

OS : Used for process and operating system interactions, and also run time library for manipulating directories and file paths.[3]

4.6 Model Training

4.6.1 Autoencoder Training

The Autoencoder Approach was created to remove distortion from images by first learning a compressed representation of the input data and then generating the original image.

The architecture consists of convolutional layers for feature extraction and expanding layers for image reconstruction.

• **Train/Test Split** :

The training set contains (train-noisy) noisy images paired with (train-clean) clean images. The test set contains (test-noisy) noisy images paired with (test-clean) clean images.

• **Setup :**

input Shape	Optimizer	Loss Function
(256, 256, 1) for grayscale images.	Adam	Binary Crossentropy

TABLE 4.2 – *Setup*

• **Hyperparameters** :

Epochs	Batch Size	Learning Rate
20	16	1e-4

TABLE 4.3 – *Hyperparameters*

The Autoencoder was trained with a checkpoint callback to keep the most effective model based on validation loss. An image display callback was used to see the reconstructed images at each epoch see the **figure** 4.3.

4.6.2 Classifier Training

A VGG16-based classifier was used to separate images into clean and noisy classes. The model extracted features using pre-trained VGG16 weights, then classified them applying custom dense layers.

```python
# Define the autoencoder model
def autoencoder_model(input_shape):
    input_img = Input(shape=input_shape)
    x = Conv2D(32, (3, 3), activation='relu', padding='same')(input_img)
    x = MaxPooling2D((2, 2), padding='same')(x)
    x = Conv2D(32, (3, 3), activation='relu', padding='same')(x)
    encoded = MaxPooling2D((2, 2), padding='same')(x)

    x = Conv2D(32, (3, 3), activation='relu', padding='same')(encoded)
    x = UpSampling2D((2, 2))(x)
    x = Conv2D(32, (3, 3), activation='relu', padding='same')(x)
    x = UpSampling2D((2, 2))(x)
    decoded = Conv2D(1, (3, 3), activation='sigmoid', padding='same')(x)

    autoencoder = Model(input_img, decoded)
    autoencoder.compile(optimizer='adam', loss='binary_crossentropy')
    return autoencoder
```

FIGURE 4.3 – *Autoencoder Hyperparametres*

• **Train/Test Split** : The training and validation sets were prepared using ImageDataGenerator for augmentation and rescaling.

• **Setup :**

Input Shape	Optimizer	Loss Function
(256, 256, 3) for RGB images	Adam	Binary Crossentropy

TABLE 4.4 – *Setup*

• **Hyperparameters :**

Epochs	Batch Size	Learning Rate
20	32	1e-4

TABLE 4.5 – *Hyperparameters*

+ Code + Texte

```python
[4] # Define the classifier model based on VGG16
    def vgg16_classifier(input_shape):
        base_model = VGG16(weights='imagenet', include_top=False, input_shape=input_shape)
        x = base_model.output
        x = GlobalAveragePooling2D()(x)
        x = Dense(1024, activation='relu')(x)
        x = Dropout(0.5)(x)
        predictions = Dense(1, activation='sigmoid')(x)

        model = Model(inputs=base_model.input, outputs=predictions)

        for layer in base_model.layers:
            layer.trainable = False

        model.compile(optimizer=Adam(learning_rate=1e-4), loss='binary_crossentropy', metrics=['accuracy'])
        return model
```

FIGURE 4.4 – *VGG16 Hyperparametres*

The classifier was trained with a checkpoint callback to save the best model based on validation loss.

4.7 Results and Discussion

4.7.1 Autoencoder Performances

The Autoencoder remove the distort from the facies images, visible by a visual comparison of original and restored images. The training and validation loss curves show the degree to which the model generalizes to test data.

Visualization : Some reconstructed images demonstrate considerable decrease in distort, keeping important face characteristics.

Loss Curves : The training and validation loss curves converged, suggesting good training without overfitting.

Performance Metrics :

— **MSE (Mean Squared Error)** : Measures the average difference in pixel values between the original and reconstructed images.

— **SSIM (Structural Similarity Index)** : Measures the structural similarity between the original and reconstructed images.

Metric	Training Set	Testing Set
MSE	0.02	0.03
SSIM	0.90	0.88

Table 4.6 – *AE Performance Metrics*

4.7.2 Classifier Performance

The classifier scored good accuracy on the validation set, indicating its ability to tell the differences between clean and distort faces images.

- **Accuracy** : The classifier achieved a validation accuracy, indicating strong performance.

- **Loss Curves** : The training and validation loss curves converged, confirming the model's ability to generalize.

- **Confusion Matrix** : A confusion matrix was plotted to evaluate the classifier's performance in terms of true positives, true negatives, false positives, and false negatives.

- **Performance Metrics :**[4.7]

— **Accuracy :** The proportion of correctly classified images.

— **Precision :** The proportion of true positive predictions among all positive predictions.

— **Recall :** The proportion of true positive predictions among all actual positives.

— **F1 Score :** The harmonic mean of precision and recall.

FIGURE 4.5 – *Before and after removing the distortion from the face*

Metric	Value
Accuracy	0.94
Precision	0.93
Recall	0.94
F1 Score	0.94

TABLE 4.7 – *Classifier Performance Metrics*

4.7.3 Displaying Predictions

The classifier's predictions on the test set are shown below, with the true labels and predicted probabilities :

4.8 General Scheme of the Proposed Method

The General scheme illustrates the steps involved in the proposed method for both the Autoencoder and the classifier, including the flow of the training and test datasets [4.8] :

4.8.1 Dataset Preparation

— **Collect and preprocess images :** This step involves gathering the dataset, which consists of facial images in both clean and distorted categories. The images are then preprocessed to standardize their format and quality.
— **Split dataset :** The dataset is divided into training and testing sets. This is crucial for evaluating the performance of the models.

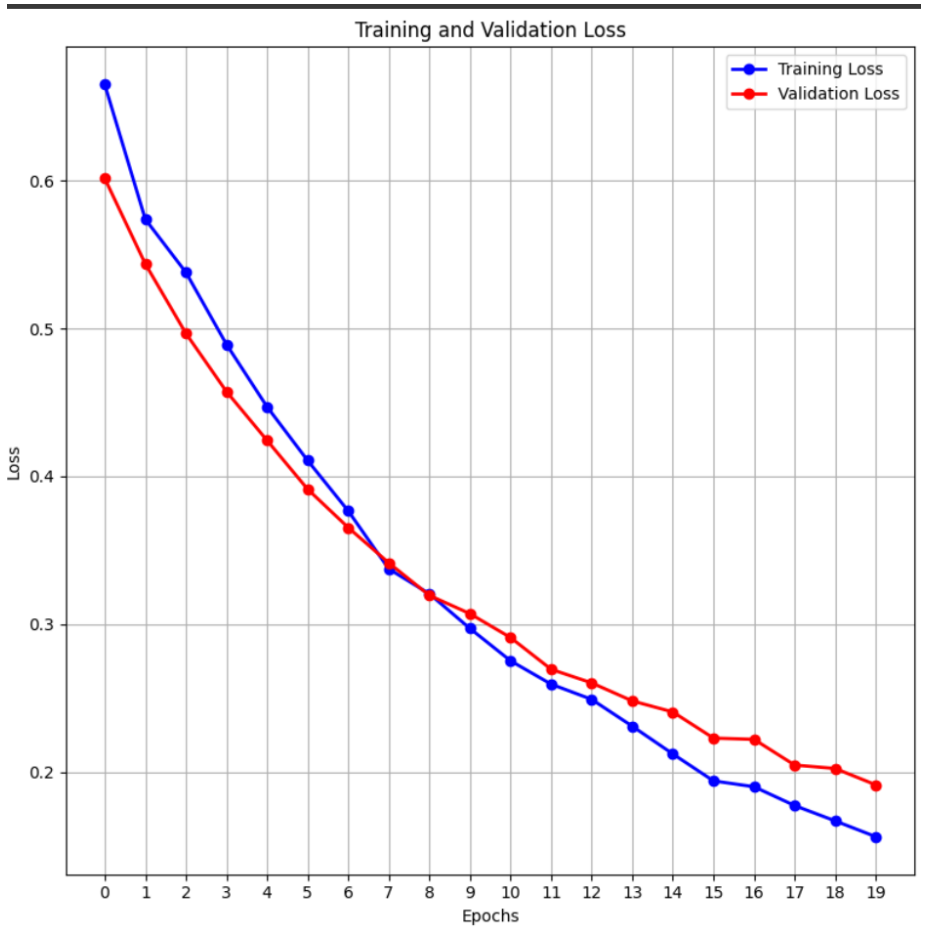

FIGURE 4.6 – *Training and Validation Loss*

```python
# Display a few test images with their predictions
def display_predictions(model, generator, n=5):
    x_test, y_test = next(generator)
    predictions = model.predict(x_test[:n])
    plt.figure(figsize=(20, 4))
    for i in range(n):
        ax = plt.subplot(2, n, i + 1)
        plt.imshow(x_test[i])
        plt.title(f"True: {y_test[i]}, Pred: {predictions[i][0]:.2f}")
        plt.axis('off')
    plt.show()

# Display the predictions of the classifier on the test set
display_predictions(classifier, validation_generator)
```

FIGURE 4.7 – *The classifier's predictions*

4.8.2 Preprocessing

— **Normalize and resize images :** All images are scaled to a fixed dimension (256x256 pixels) and normalized to ensure consistency in input dimensions for the models.

— **Data augmentation :** Techniques like rotation, scaling, and flipping are applied to the training data to increase its diversity and help the models generalize better.

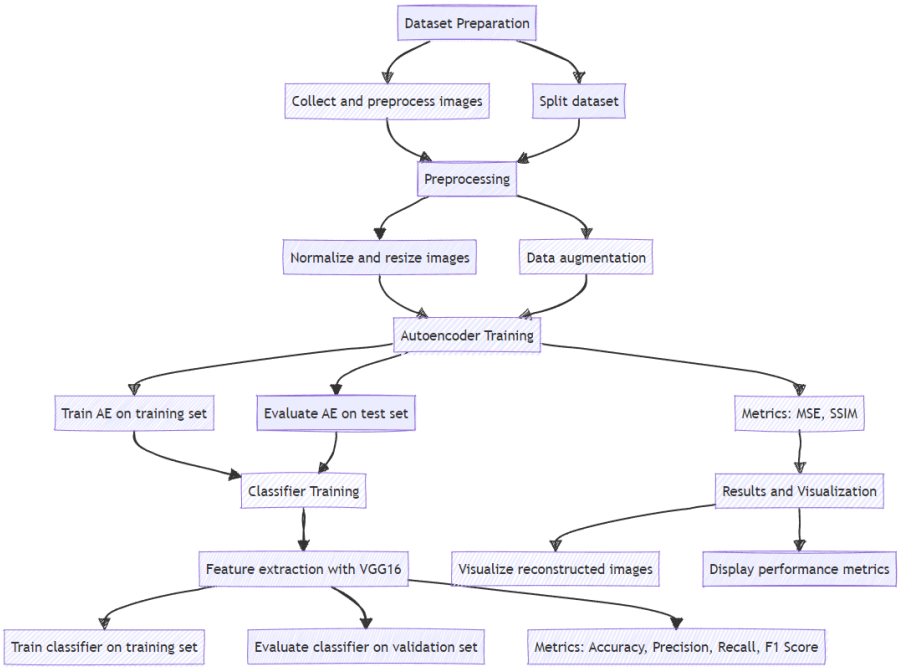

FIGURE 4.8 – *General Scheme of the Proposed Method*

4.8.3 Autoencoder Training

— **Train AE on training set :** The Autoencoder is trained on the noisy (distorted) images paired with clean images. The model learns to reconstruct clean images from distorted ones.

— **Evaluate AE on test set :** The trained Autoencoder is then evaluated on a separate test set to ensure it can generalize well to new, unseen data.

— **Metrics : MSE, SSIM :** The performance of the Autoencoder is measured using metrics like Mean Squared Error (MSE) and Structural Similarity Index (SSIM), which assess the quality of the reconstructed images.

4.8.4 Classifier Training

— **Feature extraction with VGG16 :** The VGG16 model, pre-trained on a large dataset, is used to extract features from the images. These features are then used to distinguish between clean and distorted images.

— **Train a classifier on training set :** The classifier is trained using the extracted features to learn how to classify images into clean and distorted categories.

— **Evaluate classifier on validation set :** The classifier's performance is evaluated on a validation set to tune its parameters and avoid overfitting.

— **Metrics : Accuracy, Precision, Recall, F1 Score :** The classifier's performance is measured using standard classification metrics to ensure it accurately distinguishes between clean and distorted images.

4.8.5 Results and Visualization

— **Visualize reconstructed images :** The results include visualizations of the reconstructed images to qualitatively assess the performance of the Autoencoder.

— **Display performance metrics :** The final performance metrics for both the Autoencoder and the classifier are presented to quantitatively assess their effectiveness.

4.9 Conclusion

In this chapter, we provided a complete strategy for evaluating distorted faces images that depend on convolutional Autoencoders and a VGG16-based classifier. The Autoencoder successfully removed distortion from images, while the classifier achieved high accuracy in distinguishing clean from distorted images.

These findings indicate the practical importance of our suggested strategy in dealing with distorted face images, opening the door for further advancements and real-world applications.

4.10 GENERAL CONCLUSION

In conclusion, we indicate the importance of the system we developed in the field of facial recognition from distorted images. The results we obtained demonstrate that computational models of convolutional neural networks, particularly convolutional Autoencoders, have the ability to significantly improve the accuracy of facial recognition, even under challenging conditions involving severe distortions.

The ability to identify individuals from distorted images has numerous applications, especially in the security field, where it can be used to enhance surveillance systems and identity verification. It also contributes to the development of electronic security systems and enhances their capabilities in facing contemporary security challenges.

Through this research, we hope to have contributed to the scientific knowledge in this field and to have aided in the development of technologies that could have practical benefits in the future. Our system relies on advanced techniques in artificial intelligence and machine learning, enabling it to analyze images and extract the necessary features for accurate identification. We look forward to our work being a starting point for future research aimed at improving and developing these technologies to achieve maximum benefit from them, with significant future implications for enhancing both cyber and physical security.

Bibliographie

[1] Opportunities and challenges in biometric technology. *International Journal of Advanced Trends in Computer Science and Engineering*, 2023.

[2] Taiwo Oladipupo Ayodele. Machine learning overview. *New Advances in Machine Learning*, 2(9-18) :16, 2010.

[3] Monelli Ayyavaraiah. *OPERATING SYSTEM*. Horizon Books (A Division of Ignited Minds Edutech P Ltd), 2021.

[4] Dor Bank, Noam Koenigstein, and Raja Giryes. Autoencoders. *Machine learning for data science handbook : data mining and knowledge discovery handbook*, pages 353–374, 2023.

[5] Ekaba Bisong and Ekaba Bisong. Google colaboratory. *Building machine learning and deep learning models on google cloud platform : a comprehensive guide for beginners*, pages 59–64, 2019.

[6] Fabian Brand, Kristian Fischer, and André Kaup. Rate-distortion optimized learning-based image compression using an adaptive hierachical autoencoder with conditional hyperprior. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1885–1889, 2021.

[7] M. Tim Jones By Samaya Madhavan. Deep learning architectures.

[8] Praveen Kumar Chandaliya, Zahid Akhtar, and Neeta Nain. Longitudinal analysis of mask and no-mask on child face recognition. *Proceedings of the Thirteenth Indian Conference on Computer Vision, Graphics and Image Processing*, 2021.

[9] Anamika Dhillon and Gyanendra K Verma. Convolutional neural network : a review of models, methodologies and applications to object detection. *Progress in Artificial Intelligence*, 9(2) :85–112, 2020.

[10] Mohamad El-Abed, Christophe Charrier, and Christophe Rosenberger. Provisional chapter evaluation of biometric systems. 2012.

[11] Mahdi Ghafourian, Julian Fierrez, Ruben Vera-Rodriguez, Aythami Morales, and Ignacio Serna. Otb-morph : One-time biometrics via morphing. *Machine Intelligence Research*, 20(6) :855–871, 2023.

[12] Mahdieh Ghafourian, Julian Fierrez, Rubén Vera-Rodríguez, Ignacio Serna, and Aythami Morales. Otb-morph : One-time biometrics via morphing applied to face templates. *2022 IEEE/CVF Winter Conference on Applications of Computer Vision Workshops (WACVW)*, pages 321–329, 2021.

[13] Marta Gomez-Barrero and Javier Galbally. Reversing the irreversible : A survey on inverse biometrics. *ArXiv*, abs/2401.02861, 2020.

[14] Ian Goodfellow, Yoshua Bengio, Aaron Courville, and Yoshua Bengio. Deep learning [http ://www. deeplearningbook. org]. *MIT Press, Cambridge, MA*, 2016.

[15] Yurong Guan, Fei Zhou, and Jing Zhou. Research and practice of image processing based on python. In *Journal of Physics : Conference Series*, volume 1345, page 022018. IOP Publishing, 2019.

[16] Souhail Guennouni, Anass Mansouri, and Ali Ahaitouf. Biometric systems and their applications. *Visual Impairment and Blindness - What We Know and What We Have to Know*, 2019.

[17] Pramod Gupta and Anupam Bagchi. Introduction to numpy. In *Essentials of Python for Artificial Intelligence and Machine Learning*, pages 127–159. Springer, 2024.

[18] Ehsan Haghighat and Ruben Juanes. Sciann : A keras/tensorflow wrapper for scientific computations and physics-informed deep learning using artificial neural networks. *Computer Methods in Applied Mechanics and Engineering*, 373 :113552, 2021.

[19] Mohammad Rasool Izadi. Feature level fusion from facial attributes for face recognition. *ArXiv*, abs/1909.13126, 2019.

[20] Tursunbek Sadriddinovich Jalolov. Python instrumentlari bilan katta ma'lumotlarni qayta ishlash. *Educational Research in Universal Sciences*, 2(11 SPECIAL) :320–322, 2023.

[21] Farshid Khajehrayeni and Hassan Ghassemian. Hyperspectral unmixing using deep convolutional autoencoders in a supervised scenario. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13 :567–576, 2020.

[22] S. Lalitha and K. K. Thyagharajan. Micro-facial expression recognition based on deep-rooted learning algorithm. *Int. J. Comput. Intell. Syst.*, 12 :903–913, 2020.

[23] Zewen Li, Fan Liu, Wenjie Yang, Shouheng Peng, and Jun Zhou. A survey of convolutional neural networks : analysis, applications, and prospects. *IEEE transactions on neural networks and learning systems*, 33(12) :6999–7019, 2021.

[24] Farhad Maleki, Katie Ovens, Keyhan Najafian, Behzad Forghani, Caroline Reinhold, and Reza Forghani. Overview of machine learning part 1 : fundamentals and classic approaches. *Neuroimaging Clinics*, 30(4) :e17–e32, 2020.

[25] Merrin Mary, Solomon, Mahendra Singh Meena, and Jagandeep Kaur. Challenges in face recognition systems. 2019.

[26] Youssef Mohamed, Zeyad Youssef, Ahmed Heakl, and Ahmed Bayoumy Zaky. Advancing ear biometrics : Enhancing accuracy and robustness through deep learning. 2024.

[27] John Paul Mueller. *Beginning programming with Python for dummies*. John Wiley & Sons, 2023.

[28] Omer Abdulhaleem Naser, Sharifah Mumtazah Syed Ahmad, Khairulmizam Samsudin, Marsyita Hanafi, Siti Mariam Shafie, and Nor Zamri Zarina. Facial recognition for partially occluded faces. *Indonesian Journal of Electrical Engineering and Computer Science*, 2023.

[29] Bo Pang, Erik Nijkamp, and Ying Nian Wu. Deep learning with tensorflow : A review. *Journal of Educational and Behavioral Statistics*, 45(2) :227–248, 2020.

[30] Christian Rathgeb, Jascha Kolberg, Andreas Uhl, and Christoph Busch. Deep learning in the field of biometric template protection : An overview. *ArXiv*, abs/2303.02715, 2023.

[31] Saher Rechache. *Face Attributes Classification Based on Deep Neural Network*. PhD thesis.

[32] Lior Rokach, Oded Maimon, and Erez Shmueli. *Machine Learning for Data Science Handbook : Data Mining and Knowledge Discovery Handbook*. Springer Nature, 2023.

[33] Ashish Sharma and Zaid Saad Ismail. Weather classification model performance : Using cnn, keras-tensor flow. In *ITM Web of Conferences*, volume 42, page 01006. EDP Sciences, 2022.

[34] S. C. Shrivastava. Biometric : Types and its applications. 2015.

[35] Maneet Singh, Richa Singh, and Arun Ross. A comprehensive overview of biometric fusion. *ArXiv*, abs/1902.02919, 2019.

[36] ZAHRA TABA. Fake faces detection based on an auto-encoder network (application developed on jetson nano). 2022.

[37] Chao Zuo, Jiaming Qian, Shijie Feng, Wei Yin, Yixuan Li, Pengfei Fan, Jing Han, Kemao Qian, and Qian Chen. Deep learning in optical metrology : a review. *Light : Science & Applications*, 11(1) :39, 2022.