Republic of Iraq
Ministry of Higher Education and Scientific Research
University of Al-Qadisiyah
College of Computer Science and Information Technology
Department of Computer Science

Face Recognition System under Disguise and Makeup Based on Deep Learning Technique

A Thesis Submitted to
the Council of the College of Computer Science and Information Technology at the University of Al-Qadisiyah in Partial Fulfilment of the Requirements for the Degree of Master in Computer Science

by

Farah Jawad Abdulkadhim

Supervised by:

Asst. Prof. Dr. Ali Mohsin Mohammed

2021 A.D. **1443 A.H.**

بسم الله الرحمن الرحيم

{ قَالُوا سُبْحَانَكَ لَا عِلْمَ لَنَا إِلَّا مَا عَلَّمْتَنَا ۖ إِنَّكَ أَنْتَ الْعَلِيمُ الْحَكِيمُ }

صدق الله العلي العظيم

[البقرة:32]

Supervisor Certificate

*I certify that thesis entitled "**Face Recognition System under Disguise and Makeup Based on Deep Learning Technique**" is prepared and written under my supervision at the department of Computer Science / College of Computer Science and Information Technology / University of Al-Qadisiyah as a partial fulfilment of the requirements of the degree of Master in Computer Science.*

Signature:

Assistant Professor. Dr. Ali Mohsin Mohammed

Date: 30/ 11 / 2021

Head of Department Certificate

*In view of the available recommendations, I forward the thesis entitled "**Face Recognition System under Disguise and Makeup Based on Deep Learning Technique**" for debate by the examination committee.*

Signature:

Assistant. Dr. Qusay Omran Mosa

Head of the Department of Computer Science

Date: 30/ 11 / 2021

Certificate of the Examination Committee

We, the undersigned, certify that ***(Farah Jawad Abdulkadhim)*** *candidate for the degree of Master in Computer Science, has presented this thesis entitled* ***(Face Recognition System under Disguise and Makeup Based on Deep Learning Technique)*** *for debate examination. The examination committee confirms that this thesis is accepted in form and content and displays a satisfactory knowledge in the field of study based on the candidate demonstration during the debate examination held on:* ***18-November-2021.***

Signature: M. J. Jawad
Name: Dr. Majid Jabbar Jawad
Title: Professor
Date: 2 / 12 / 2021
(Chairman)

Signature:
Name: Dr. Mustafa Jawad Radif
Title: Assistant Professor
Date: / / 2021
(Member)

Signature:
Name: Mohammed Hamzah Abed
Title: Assistant Professor
Date: / /
(Member)

Signature:
Name: Dr. Ali Mohsin Mohammed
Title: Assistant Professor
Date: / /
(Supervisor)

Signature: Dhiah
Name: Dr. Dhiah Eadan Jabor Al-Shammary
Title: Assistant Professor
Date: 2 / 12 / 2021
(Dean of College of Computer Science and Information Technology)

Dedication

This work is dedicated to…

My father and My mother

I will always be your proud daughter, I ask God to protect you from all harm and I will not disappoint you

My beloved husband and children

Whom I can't force myself to stop loving.

My beloved brothers and sisters

Who stands by me when things look very difficult.

Acknowledgement

First of all, I thank "The Greatest Allah" for granting me the strength, willing and patience to accomplish this work.

My grateful deep appreciation and thanks to my supervisor Asst. Prof. Dr. Ali mohsin Al-juboori for his continuous guidance and support, for his patience, encouragement, meaningful and valuable instruction throughout the whole year during working on this thesis.

My thanks to the Academic and Administrative staff at the College of Computer Science and Information Technology.

Finally, I wish to thank my dear family for support and love from the beginning of this journey till this day.

Farah Jawad

2021

ABSTRACT

Facial recognition has been broadly used in advanced intelligent systems (i.e: smart video surveillance, intelligent access control system, and online payment). The performance of existing algorithms for automatic facial recognition is hampered by various covariates like pose variations, face aging, Masks, disguises, and makeup.

Disguises and makeup are especially used to intentional or unintentional change facial appearance to either hide one’s personal identity or impersonate someone's different identity. It is also one of the common traditional variables that people make in their daily lives. While new algorithms continue to improve performance, most face recognition systems are liable to failure when disguised or makeup altered, which is one of the most challenging factors to overcome.

With enormous capability and promising results, deep learning technology becomes attracted the greatest attention to the research in a diversity of computer vision tasks. In order to overcome this problem, the database of disguised and makeup faces (DMFD) used in this thesis. Which contains images under these two variables mentioned. The facial images are obtained from the (DMFD) dataset from (Hong Kong Polytechnic University) comes with cropped images and passes through preprocessing steps before the classification procedure.

The preprocessing phase includes performing histogram equalization to enhance the contrast value of the facial image. Afterward, the image is resized to prepare it for the feature extraction phase. The features of facial images are extracted in this thesis using Linear Discriminant Analysis (LDA) method that gives the minimum feature length and minimum extraction time.

Facial recognition system is done by using the proposed hybrid-deep learning Classifier for more precise feature learning. Also, we compared the proposed method with pre-trained model (AlexNet) and four machine learning algorithms (Naïve Bayes (NB), K- Nearest Neighbor (KNN), Random Forest (RF), and support vector machines (SVM)).

The experimental results show a high-quality performance and a perfect precision value equals to 99% when using the LDA feature extraction with the proposed hybrid-deep learning classifier on the facial images. The proposed system also was fast in the concept of speed with time equals to parts of a second in the prediction of the person.

Table of Contents

Continued

Continued

List of Figures

Continued

Continued

List of Tables

List of Algorithms

Continued

List of Abbreviation

Abbreviation	Expression
PIN	Personal Identification Number
SSN	Social Security Numbers
PDA	Personal Digital Assistant
2D	Two Dimension
MRTD	Machine Readable Travel Document
LDA	Linear Discriminant Analysis
HIS	Hue, Saturation, Intensity
HSV	Hue, Saturation, value
ATM	Automated teller machine
MSAC	M-estimator Sample Consensus
SVM	Support Vector Machine
ML	Machine learning
RBF	Radial Basis Function
CNN	Convolutional Neural Network

Continued

KNN	K-nearest Neighbor
PCA	Principal component Analysis
HOG	Histogram of oriented gradients
HCI	Human-Computer Interaction
1D	One Dimension
NB	Naïve Bayes
RF	Random Forest
NN	Neural Network
ANN	Artificial Neural Network
TP	True Positives
FP	False Positives
FN	False Negative
TN	True Negative
DMFD	Disguise and Makeup Face Dataset
RELU	Rectified Linear activation function
RGB	Red Green Blue

List of Equation

Continued

List of Publication

No.	Paper Name
1	**“Detection of Human Faces Covered with disguise and Makeup”, Journal of Al-Qadisiyah for Computer Science and Mathematics .**
2	**“Face Recognition with Disguise and makeup Variations Using Image Processing and Machine Learning”, 5th International Conference on Advances in Computing and Data Sciences (ICACDS)-2021, Springer in Communications in Computer and Information Science.**
3	**“Face Identification under Disguise and makeup based on hybrid deep learning”, 7th International Conference on Contemporary Information Technology and Mathematics, 2021 – IEEE in Computer Vision and Image Processing.**

Chapter One

Introduction and Literature Survey

Chapter One

Introduction and Literature Survey

1.1 Introduction

This chapter includes a brief overview of the image processing concept and its implementations, as well as a biometric definition, biometric recognition system concept, forms, specifications, and applications of the biometric system, literature Survey, the problem statement, thesis objectives, and thesis outlines.

1.2 Biometric Systems

In the field of information technology, ensuring that information is accessible and safe is a major concern. Since the methods of authentication and validation used in previous decades, such as PIN and passwords, are inefficient due to the risk of fraud, a solution to this problem is provided by using biometric identities. Biometrics refers to a method for distinguishing individuals based on at least one intrinsic physical or behavioural characteristic.

Biometrics refers to technologies that use to calculate and analyse human body characteristics such as hand measurement, voice pattern, fingerprints, facial patterns, and retinas for the purpose of authentication. The term 'Biometric' is derived from the Greek terms 'Bios', which means life, and 'Matron,' which means measure [1]. Compared to conventional approaches, the biometric system has many benefits. Biometric traits, unlike passwords and tokens, can not be lost or forgotten. Biometric traits are difficult to copy, exchange, distribute, or steal. Biometrics can be classified into two categories as show in figure (1.1):

1) Physiological: It is linked to body shapes such as hand geometry, DNA, fingerprints, palm prints, face recognition, iris recognition (which has mostly replaced the retina), and smell.
2) Behavioural: It is linked to an individual's behaviour, such as their speech, gait, and typing rhythm. This form of biometrics has been referred to as (behaviometrics).

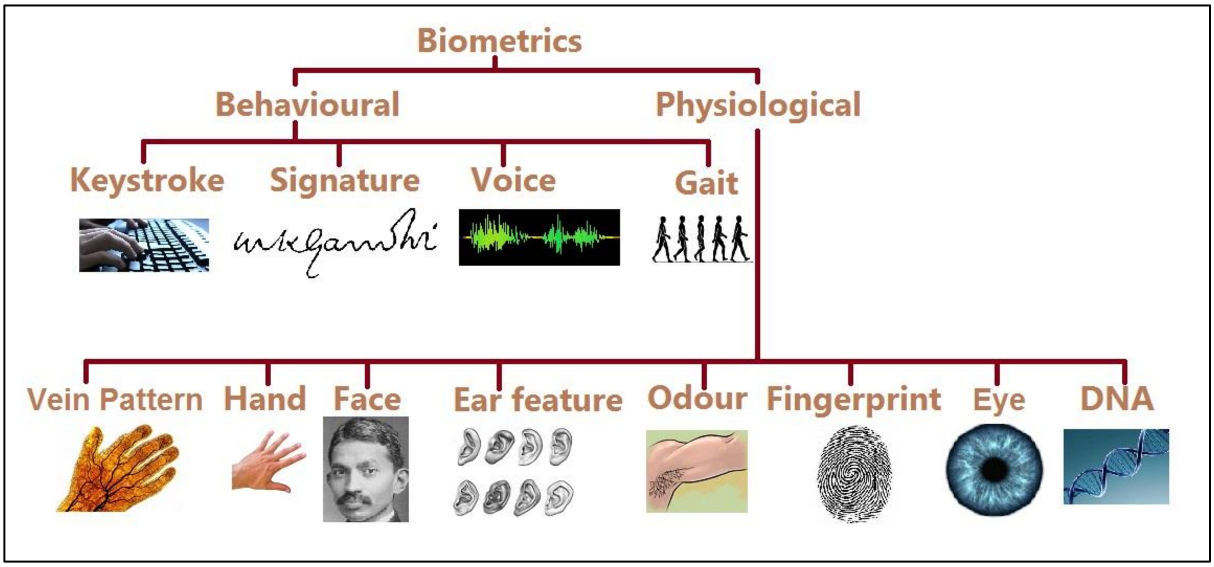


Figure (1.1): Types of Biometrics [2]

A biometric system is a pattern recognition system that acts by gathering biometric data from a person, extracting features from that data, and comparing those features to a database template(s) set. A biometric device can work in either authentication or identification, depending on the application context. The system verifies a person's identity in the verification mode via comparing the captured biometric data to the own biometric template (s) saved in system database. In identification mode, the device recognizes a user by looking for a match among all of the users' models in the database [3].

1.2.1 Biometric System Characteristics

Since biometric systems are commonly used for authentication and protection, they should have a number of characteristics [4].

1) **Collectability:** refers to the ease with which traits can be obtained and evaluated.
2) **Acceptability:** refers to the users' acceptance of the system; they should feel at ease using it.
3) **Performance:** A precise result is always expected, regardless of circumstances in which the system operates.
4) **Uniqueness:** In order for the system to perform tasks correctly and accurately, the users' characteristics must differ to some extent.
5) **Universality:** The characteristics that were used had to be present in all of the enrolled users.
6) **Permanence:** Over time, the biometric traits token for approved people should not be not alter.

1.2.2 Biometric System Applications

Biometric systems are used for authentication, security, and access control in a variety of applications. The following is a list of some of them [5]:

1) Biometric devices are commonly used in commercial applications like Internet access, E-commerce, computer network login, ATMs (or credit) cards, cell phones, physical access control, medical records management, and PDAs.

2) Biometric devices are used in government applications as national (ID) cards, driver's licenses, social security numbers (SSN),passports, welfare spending codes, border crossing systems, and surveillance systems to the identity verification.
3) Biometric devices are used in forensic applications such as corpse detection, parenthood determination, and the criminal investigation.

1.3 Biometric system based on facial recognition

Facial recognition technology has been one of the most well-known research subjects in pattern recognition and computer vision in recent decades. This interest illustrates the need for a reliable face recognition system. There are still some obstacles to address when it comes to facial recognition device research and performance enhancement, Such as pose variance, facial expression, occlusion, lighting, and aging [6]. Face detection, feature extraction, and classification are the three main stages that make up a facial recognition system. Each stage is important, but feature extraction has a big impact on the recognizing system's accuracy [7].

A face recognition system recognizes the person by their input facial images to a system, where the personal identity is a specific procedure of linking a determined person with an identity using two traditional techniques widely used: token-based methodologies and knowledge-based methodologies. Token-based methodologies depict "something owned" in order to obtain an individual identity, such as a driver's license, travel document, credit card, or ID card. Knowledge-based methodologies obtain "something you know" to conduct an individual identification, such as the person's identification number (PIN), or secret words. Both methods have drawbacks, such as tokens being stolen, squandered, lost, or forgotten, and a PIN being ignored by a legitimate user or speculated by a juggler. They are unfit to distinguish the authorized person from the juggler who obtains the authorized individual's token or awareness by fraud.

A biometric technique has been used to develop individual personality and is a more reliable method [8]. Biometry is a method of determining an individual's personality based on their physical or behavioural characteristics. There are some benefits of using a facial recognition system that is not joined with biometrics strategies. Other biometric systems are difficult to use because they require physical contact with the device and a brief delay for detection and recognition. Among the six well-known biometric systems used in Machine Readable Travel Documents (MRTD), such as registration, protection, machine specifications, renewal, surveillance, and public perception, facial features had the highest compatibility. Face recognition has the highest weighted percentage in the biometric system as compared to other biometric traits dependent on MRTD, as shown in Figure (1.2) [9].

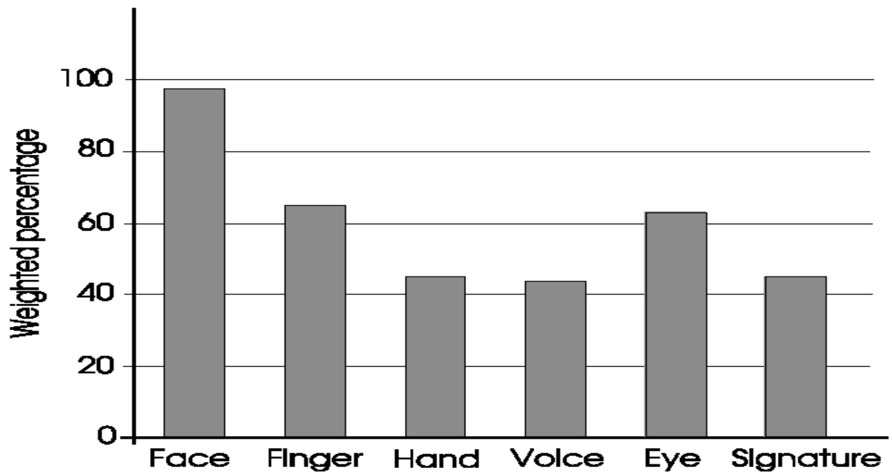


Figure (1.2) different biometric traits comparison based on MRTD [9]

1.4 Literature survey

Disguise and makeup is significant to face identification problem. Facial recognition has been actively studied since its inception in the 1990s, owing to its wide range of practical applications. Processing a huge volume of data in real-world environments has become extremely complex in recent years. There had been a lot of research accomplished in that area. The most common research in this field is as follows:

1- S. Haji and A. Valor (2016) [10], this research proposed a real-time application, based on the windows based real time system, for face recognition. The system is depending on the Eigen and Local Binary Patterns to measure the similarity between face images, to authenticate users, to eliminate the effect of multiple illumination conditions. Despite the high recognition accuracy of 90%, the study shows that the accuracy is dramatically affected by any changes to the conditions of the environment, such as distance between the individual and the camera or ambient lighting, or when the images are collected using different cameras.

2- M Shujah Islam et al. (2017) [11], this research proposed a real-time face recognition system based on computer vision techniques. The system extracts face images based on the use of viola jones for faces detection, which are then cropped out of the image before the features extraction, using the SURF method. The extracted features are matched using the M-estimator Sample Consensus (MSAC) in order to authenticate users. The method has been able to achieve 95.9% recognition accuracy, which shows the importance of using face detection prior to matching stage, to eliminate any additional features that may exist in the image.

3- Abhila et al. (2018) [12], this research Presented for disguised face detection, the Viola Jones approach was utilized. The algorithm works by looking for particular haar-like characteristics. A pre-trained CNN (Alex-Net Model) is used to extract facial features before a multi-class SVM method is used to perform the classification task There has been a new dataset created. It includes images of people wearing various scarves in various ways. All of the images are of the front of the face, with the eye region clearly visible. The goal of this article is to recognize a disguised face.

4- Sarapakdi et al. (2019) [13], this research studies the impacts of two-dimensional principal component analysis (2DPCA) in the initialization stage to image classification using (CNN) to lessen the complexity of a facial occluded recognition system. The experiments suggested that applying 2DPCA instead of the whole image for training reduces the dimension of the image while retaining the accuracy, which is 81% for AR and 99% for GTAV databases.

5- N Christou et al. (2019) [14], this research tried to identify a efficient solution for enhancing recognition accuracy. Convolutional neural networks could be used to solve the problem (CNNs). On the FER2013 data, which includes genuine face images assigned to the seven facial expressions categories (Anger, Disgust, Fear, Happy, Sad, Surprise, and Neutral), the experimental results show that the proposed solution can enhance the accuracy to 91.12 %.

6- Vidya P et al. (2020) [15], this research been explore image processing IP along with machine learning algorithms ML and techniques ensuring efficiency and high accuracy. The work of the proposed system is the automatic marking of the student’s attendance established on the detection and recognition of students utilized face image features. Using Viola-Jones algorithm to face detection. In face recognition, LDA is implemented with two classifiers KNN and SVM. As Final result, LDA along with KNN hands high accuracy on the used database uses a camera installed at a proper place for taking students' face images in a constrained environment. One of the main reasons for the increase in accuracy is more facial expressions are taken for each student.

7- Pranav et al. (2020) [16], this research implementation and evaluation of a real time facial recognition system based on CNN are presented in this research. The proposed architecture is first evaluated uses standard AT&T datasets. and then the same methodology is applied to a real-time system. The tuning of the model CNN parameters to analyse and enhance the proposed system's recognition accuracy is also described. To improve the system's performance, a systematic method for tuning the parameters is also proposed. The standard datasets with real-time inputs, the proposed system achieves maximum recognition accuracy (98.75 % and 98.00 %).

8- Hung et al. (2021) [17], this research present Face detection and identification. HOG features + SVM classifier are used in the face detection approach. The suggested face recognition model is based on a convolutional neural network (CNN). On the FEI, LFW, and UOF datasets, the model's efficiency is tested, and the results reveal that the suggested model achieves good accuracy.

Table (1.1) Literature Survey Summary

No.	Ref	Feature extraction methods	Classification methods	Result	Datasets
1	[10]	Eigenface PCA-LBP	Matching method	90%	Webcame
2	[11]	SURF	MSAC	95.5%	Graz01
3	[12]	AlexNet	SVM	88%	Disguise face

4	[13]	2DPCA	CNN	81.91% 99%	AR GTAV
5	[14]	CNN	CNN	91.12%	FER2013
6	[15]	LDA	KNN,SVM	97% LDA+KNN	Webcam
7	[16]	CNN	CNN	98%	AT&T
8	[17]	CNN	CNN	98.7% 98% 95.2%	FET-P1,P2,P3 Faces96 LFW
9	This Thesis	LDA	hybrid- deep learning (CNN+DNN)	99%	DMFD

1.5 Problem Statement

Due to changing the appearance of the face under the effect of Disguise and makeup, the face identification system meets some difficulties that make the system unable to verification the face with reasonable performance exist. So the main problem that is addressed in this thesis is "Recognition human Face covered with Disguise (like glasses, beards, hairstyles, and other facial accessories) and makeup variations". The overall efficiency and accuracy of a facial recognition system is largely dependent on the techniques used to extract facial features, though facial feature extraction has a number of issues that must be considered and resolved, including:

1) Variation in appearance this variation involves the facial image is with or without (makeup, different disguises like as glasses, beards, hairstyles, and other facial accessories), as well as facial expression, whether (mouth and eyes) is open or closed.
2) Pose variations represented in the camera-face pose (frontal, profile, upside down), and distance from the camera.
3) In this thesis, facial image captured in a real-world environment was used to reflect pose variations in the camera-face pose and distance of the face from the camera.
4) Changes in the light source and Texture, in particular, can dramatically alter the appearance of a face, particularly the face's surface properties and the light sources, where this thesis employed the feature extraction based on (LDA) linear discriminant analysis and classification based on hybrid- deep learning.

1.6 Aim of thesis

There is a growing interest in human identification in managed environments such as airports, banks, and parking lots, and as a result, ground-breaking biometric recognition methods for human identification at a distance have become a pressing need and have sparked a lot of interest among computer vision researchers. The aim of this thesis are:

1- Building a model with highly effective Face recognition under Disguise (like glasses, beards, hairstyles, and other facial accessories) and makeup variation, and the accuracy of identification not affected by changing facial features during these variations.

2- This thesis addressed establishing a facial recognition system depending on features of human faces, with the following phases: (image Features extraction and classification).

3- Linear discriminant analysis LDA is used for features extraction and proposed hybrid-deep neural network with 12 layers used for classification, and its comparison with the pre-trained model (Alex net) and four classification algorithms are also used from machine learning (NB, KNN, RF and SVM) to show the better one.

1.7 Outlines of Thesis

Addition on the **First Chapter**, which serves as the thesis' introduction, there are four additional chapters, which are arranged as follows:

1. The **second chapter** discusses face pre-processing, feature extraction and classification methods.
2. The **third chapter** presents a detailed overview of the proposed face recognition system's architecture and implementation.
3. The **fourth chapter** explains the proposed system's implementation and discusses the results.
4. The **fifth chapter** describes the results reached as well as suggestions for future work.

Chapter Two

Theoretical Background

Chapter Two

Theoretical Background

2.1 Introduction

This chapter provides a Theory Background of all of the mechanisms and techniques used in the proposed face recognition system, include a brief overview of the properties and challenges of facial images, as well as feature extraction methods and machine learning. These terms do the key outlines that will be presented and explored in this chapter, and they play an important role in thesis motivation.

2.2 Image Processing

Image processing is a method of improving or extracting usable information from an image by performing an operation on it. It's a type of signal processing in which an image is an input and the output is an image or the image's features [18]. A digital image is a numerical representation of a real image that a computer can store and process. The image is split into tiny areas called pixels in order to convert it into numbers (picture elements). For each pixel, the imaging device maintains a number, or a small set of numbers, that describes a pixel's characteristics, such as brightness (light intensity) or color that are arranged in an (array of rows and columns) that match the image's vertical and horizontal pixel positions.[19]

Depending on the color model used, such as the RGB model, HIS and HSV model, and the grey Color Model, each part of this vector corresponds to various aspects of color . Image processing is a technique for enhancing images obtained from cameras/transducers mounted on satellites, space ships, and aircraft, as well as images captured in everyday life for a range of applications. In image processing, variation techniques have advanced over the last four to five decades. The majority of technology was developed to improve image quality. Because of the easy availability of powerful personal computers, big-size memory instruments, graphics applications, and so on, image processing is becoming more popular. [20] Image processing is used in a wide range of applications, including:

1. Face recognition.
2. Character recognition.
3. Age estimation and classification.
4. Iris recognition.

2.3 Facial Recognition Approaches

Face recognition is a technique that is utilized for recognizing the individuals from only his/her face and it is one of the uttermost efficient biometric techniques for a person's identification procedure. Biometric recognition became an integral part of our living and especially in authenticating and verification of individuals and there is plenty of biometrics that exists and can be used such as clothing, the shape of the body, voice or gait might be establishing identities with regard to conditions in which the facial details might not be provided. Recently, human identification based on face recognition was of high importance.

The face images of the humans show a lot of information, it might specify attentiveness, intention, and mood, and it might be utilized for identifying humans since the human face contain much significant biometrics information. Facial recognition aims for identifying individuals based on some facial characteristics like lines and wrinkles in the face, spaces between the cheekbones, the position of eyes, and so on. In addition, face recognition is of high importance due to its many applications in computer entertainment, surveillance, access control, security, law enforcement, and internet communication. Automatic face recognition systems through computers might be classified into three methods [21]: As shown in Figure (2.1)

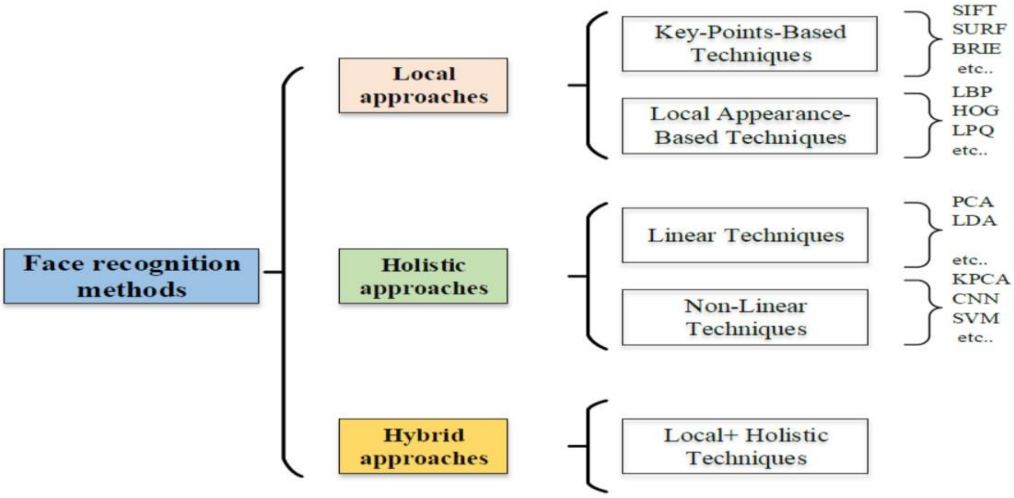


Figure (2.1): Facial recognition approaches [21]

1- Local Approaches: Simply few facial features are addressed by local approaches. They are higher sensitive to facial pose, occlusions, and expressions than other people. The basic goal of these approaches is to find distinguishing characteristics. In general, these approaches can be categorized into (2) groups: (a) To extract local features, local appearance-based approaches are applied, and the facial image is separated into little regions or patches. (b) To extract the features centered on these places, key point based algorithms are utilized to identify the points of interest in the facial images.

2- Holistic Approaches: Holistic approaches or subspace are intended to process the entire faces, without the need to extract faces areas or feature positions (mouth, eyes, noses, and etc). The fundamental purpose of these methods is to describe the face image as a pixel matrix, which is later converted into feature vectors to make handling easier. Then, in low-dimensional space, these feature vectors are implemented. Holistic or subspace techniques, on the other hand, are sensitive to differences (facial positions, illumination, and expressions), and those advantages have made them popular. Furthermore, those methods can be classified into kinds, such as (linear and non-linear) techniques, based upon the representation way for the subspace.

3- Hybrid Approach: To take advantage of both subspace and local strategies, the hybrid approaches are based on local and subspace characteristics. which have the potential to improve the accuracy of face recognition.

2.4 Challenges of facial recognition systems

Facial recognition is one of the most important areas of research in pattern recognition. Additionally, Authentication based on facial recognition was also widely used. Face recognition in unregulated environments is still difficult, despite the fact that face images were considered in unconstrained conditions and showed significant variations in the expression, complex context, occlusions, pose, and lighting, as shown in figure (2.2). Typically, face verification evaluation techniques assumed that the individual's identification in testing and training sets was unique, requiring predictions about faces that had never been seen before.[22]

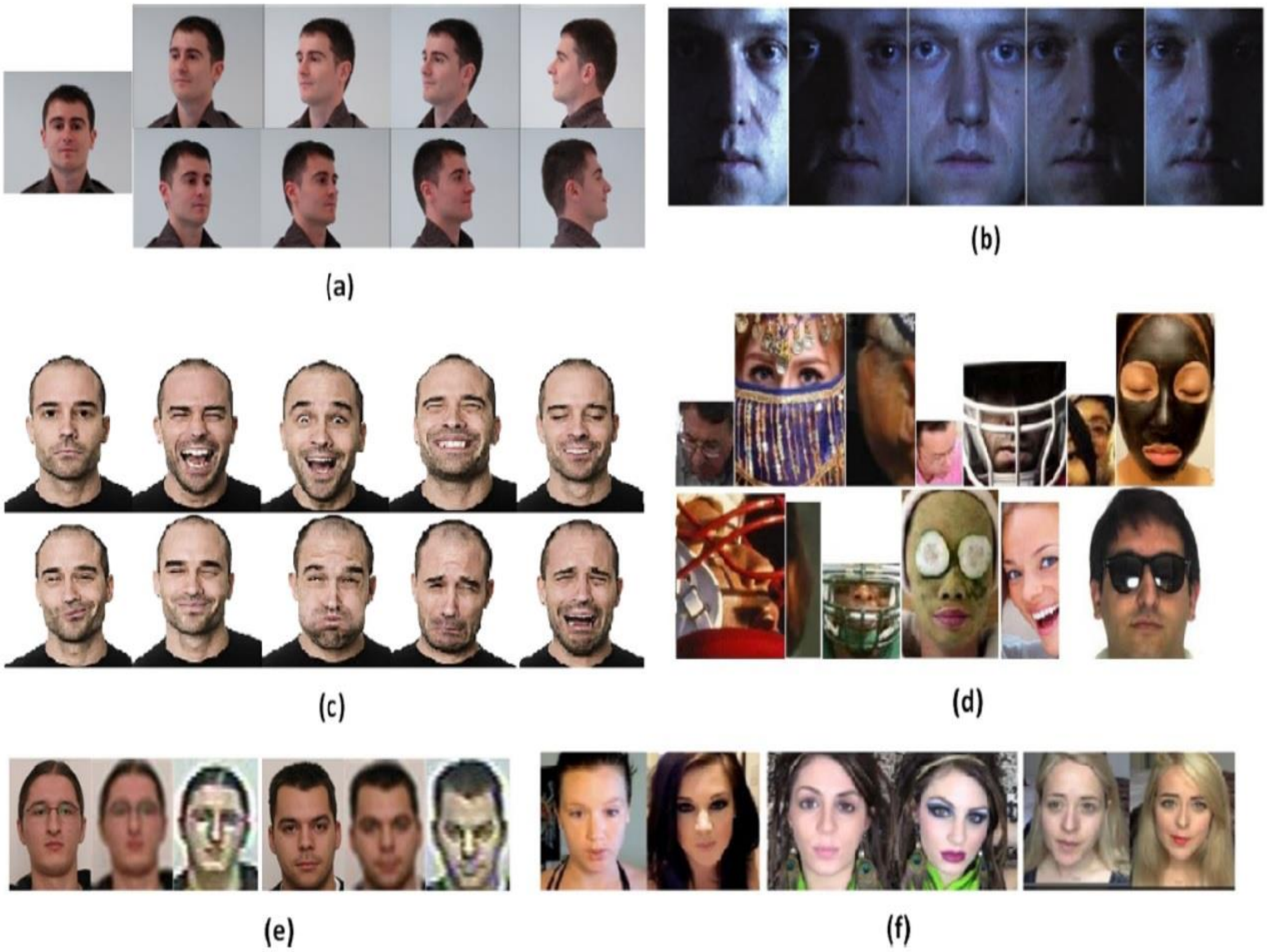


Figure (2.2): Faces with variations (a) pose, (b) illumination, (c) expression, (d) occlusions, (e) low resolution, and (f) makeup. [22]

In reality, there were several key factors and challenges that could have a significant effect on face recognition performance as well as other factors that could affect matching scores, such as the following:

a) **Illumination variations:** When a picture is taken, variables such as lens and sensor reaction, as well as lighting (intensity, distribution, and spectra), have an effect on the appearance of the human face. Because of the internal camera control and skin reflectance properties, as well as the use of a low-cost camera, this can cause illumination variations. One of the main technical issues confronting the designers of face recognition systems is the problem of illumination/lighting variations, in which a person's face can be extremely different [23]. Despite the fact that image pre-processing techniques were quick and easy, they overlooked the effect of lighting direction shifts on the extracted local features, which can be used to differentiate people from their physical appearance.

b) **Pose variations**: The pictures related to facial changes because of the relative camera face pose, certain facial features, such as the nose and eyes, can be completely or partially occluded. Actually, due to self-occlusion and the projective deformations, pose changes have an effect on the recognition process. As a result, pose tolerance becomes even more important for face recognition systems that rely on the subject's single view. The most noticeable differences in appearance are caused by poses, which result in extreme performance degradation. [24]

c) **Wrinkles and Ageing:** Aging can be both artificial (using makeup tools) and normal (aging naturally due to the progression of age). Wrinkles and aging can have a significant effect on the efficiency of face recognition methods in all of these cases. Aging is a normal part of life that happens to everyone but is slow. Individual aging habits vary and are influenced by a variety of factors like environment, culture, lifestyle, gender, illness, alcohol use, and antiaging medications, among others. The extracted features change as the appearance of the individual's face changes. The features saved in the face's database must be modified to reflect these improvements, thus enhancing the usabillity of the facial recognition system. [25]

d) Facial expression or style: In the field of computer vision, analysing facial expressions is a hot topic. When the intelligent system's input is a facial image, Facial Expression Recognition is an effective visual recognition technology for detecting emotions. Expression Recognition is widely applied in Human-Computer Interaction (HCI), autonomous driving driver monitoring, education, healthcare, and psychological therapies. Several algorithms are used to recognize facial expressions, but they do not have the drawbacks of imperfect facial expression recognition [26]. The individual's facial expression had a significant effect on the appearance of his or her face. Facial hair, such as a moustache or beard, can also alter the facial appearance and features in the lower half of the face, especially near the chin and mouth. Hairstyles may be altered to modify the look of the face or to conceal facial features.

e) **Occlusion**: Faces might be partially occluded by other objects such as a variety of glasses, masks, and accessories, as occlusion, as illuminated in figure (2.2) set (d). A few objects of faces in an image related to a group of individuals may partially occlude other faces, resulting in only a small part of the face being provided in different circumstances, making features difficult to recognize. [27] The recognizer conditioned by these images does not perform well in real-life situations. When a person's face is obscured in an unregulated scene, the recognition accuracy plummets. To summarize, most existing methods for synthesizing occlusion images cover some abnormal occlusion masks to the normal picture, deviating from the actual situation.

f) **Low Resolution:** Surveillance video cameras provide images with small faces, resulting in low resolution. Comparing the low-resolution query image to the high-resolution gallery image is a challenging task. Frame from surveillance video with such a low-resolution image. The data in a video of a captured face is extremely limited because the majority of the points of interest are lost. This has the potential to significantly lower the recognition rate. [28]

2.5 Color Models for Digital Images

Digital images are described as a numerical representation of a real image that can be handled and processed by digital computers. Images will be separated into small areas known as pixels (picture elements) for the purpose of converting them to numbers. In an imaging device, a number or small collection of numbers is registered for each pixel, which describes some pixel property, such as color or brightness (light intensity) [29]. The numbers have been organized in an array of (columns and rows) matching to horizontal and vertical locations of image pixels, each of the vector components matches to a distinct color aspect depending on the color model used, such as HSV, RGB, HIS, and Gray Color Model.

1) **RGB:** Any color can be represented in this model as a linear mixture of three basic colors (red, green, and blue). Colors are shown on TVs and computers using this model. The digital image was defined using three 2D matrices, one for each of the primary colors in the RGB model, with equivalent sizes, and the values from these three matrices were mixtured to produce an image to display. Each element in the three matrices was typically represented using 8 bits. As shown in Fig. (2.3), the color pixel was defined using 3×8 = 24 bits. As a result, there are (2^{24} = 16777216) possible colors [30].

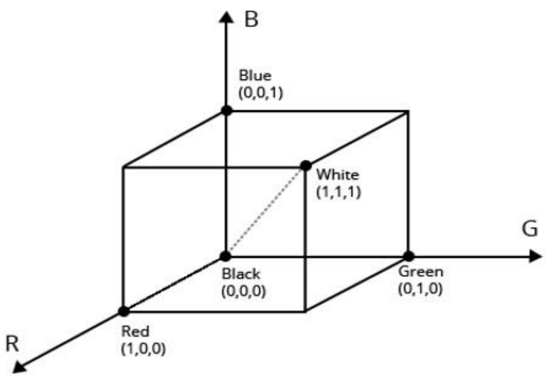


Figure (2.3): the RGB color model [30]

2) **Gray color model:** A picture with just brightness information and not color information is referred to as a gray image. Is a 1D (channel) that contains only gray shades of color (256 gray colors) and uses an 8-bit representation. Grayscale images are distinguished by the equality between red, green, and blue color levels. The color code would be RGB (R,R,R), RGB (G,G,G), or RGB (B,B,B), with "R,G,B" each being a number between (0 - 255). In some applications, converting the color images to a grayscale representation is needed, and most of now show and picture capture hardware can just support (8-bit) images. Furthermore, for several tasks, grayscale images are appropriate, therefore there is none need to employ more complicated and difficult to process color images. The graying processing on graphics RGB is usually done using the weighted average method, in which that the brightness's values of pixels (R, G, B) in all layer are then weighted and summed to get brightness's value of a the gray image. In general, the human being eye is more sensitive to colours in the order (green > red > blue), so the weight coefficient must be (G > R > B). If the coefficients from (R, G, and B) are (0.299 , 0.587 , and 0.114) sequentially, according to the human eye's sensitivity to visible range, a gray scale image suitable to the human being eye can be achieved [31]. As a result, can use the following equation to convert the original image to grayscale:

$$GRAY = 0.30\,R + 0.59\,G + 0.11\,B \qquad (2.1)$$

2.6 Face Image Pre-process

The operation of preprocessing includes two-step, improving the face image via modifying the contrast using histogram equalization and the enhanced face images are resized. The first step, processing of an image in order to improve a certain feature of the image known as image enhancement. Image enhancement is the process of enhancing the interpretability or perception of information in images for human viewers while also giving better input for other automated image processing approaches. Its main goal is to change the characteristics of an image to make it more suited for a specific task and observer. In the following situations, image enhancement is used: Noise removal from an image, the dark image is enhanced, and the edges of the items in the image are highlighted, making the image more smother. For some uses, the output is more suitable than the original image [32].

Image enhancement is the process of changing a picture f into an image g utilizing T. (Where T is the transformation). Pixel values in images f and g are indicated by the letters r and s, respectively. To improve images in some way, many diverse, often simple and heuristic methods are applied. Of fact, the problem is not adequately defined because there is no objective measure of image quality.

A grey level slicing function can either emphasize a set of intensities while diminishing all others, or it can emphasize a group of grey levels while ignoring the others. Histogram Processing is a popular and important image enhancement technique in which histograms are usually normalized by the total number of pixels in the image. A discrete function is the histogram of a digital image with intensity levels in the range [0, L-1]. The histogram enhancement method is applied in the digital image to improve the intensity of facial images [33].

Histogram equalization is a spatial domain method that creates an output image with a uniform pixel intensity distribution. This means that the output image's histogram is flattened and extended consistently. Due to its ease of use and efficacy, histogram equalization is commonly utilized for contrast enhancement in a range of applications. Medical image processing and radar signal processing are two examples. One disadvantage of histogram equalization is that the brightness of an image can be modified after it has been equalized, which is primarily due to the histogram equalization's flattening property.

This strategy increased the total contrast of many images, especially when the image's relevant data is represented by an adjacent value of contrast, as seen in figure (2.4). The modification allows the intensity distribution on the histogram to be finer, allowing areas with low local contrast to have a superior contrast. Histogram equalization is achieved by diffusing out the most frequent intensity values in an effective manner. As demonstrated in the following equation, the histogram cumulative distribution function is useful in computing histogram equalization [34].

$$cdf(X) = \sum_{i=1}^{x} \mathrm{h}(i) \qquad (2.2)$$

Where (X) is the gray value and (h) denotes the histogram of the image.

$$\mathrm{T[pixel]} = \mathrm{round}\left(\left(\frac{cdf(x) - cdf(x)_{min}}{E*F - cdf(x)_{min}}\right) * (\mathrm{L} - 1)\right) \qquad (2.3)$$

$cdf(x)_{min}$: is the cumulative distribution function's minimum value.

E * F: Columns and rows of images, L: Gray levels. (used =256).

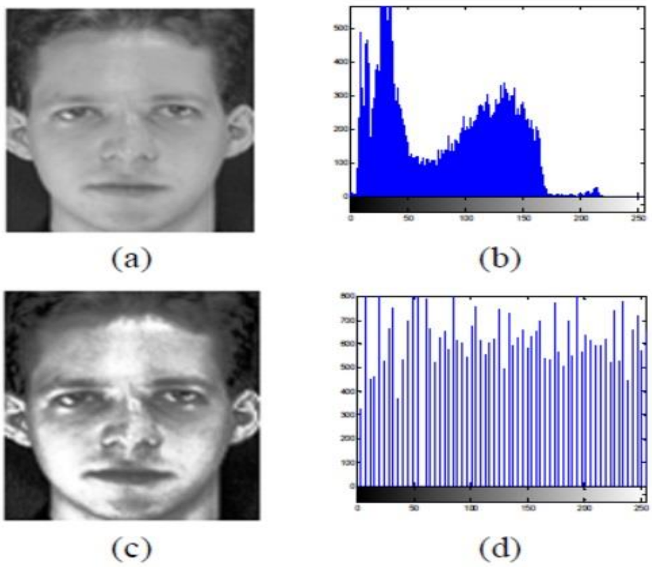


Figure (2.4): Facial image after applying histogram equalization [33].

The detailed steps in the algorithm (2.1)

Algorithm (2.1): The Histogram Equalization Algorithm
Input: Face image $\boldsymbol{F_{Gray}}$ with gray levels from 0 to 255 **Output:** Face image in enhanced intensity $\boldsymbol{F}_{\text{Enhanced}}$
BEGIN 1 Get size of $\boldsymbol{F_{Gray}}$: E* F, gray level (0 - 255) , Build a (G1 matrix) with size (256) and start with a 0. 2 Scanning each pixel in the image and increasing the related element of the matrix to create the image's histogram. G1 [gray-val (pixel)] = G1 [gray-val (pixel)] + 1 3 Get the function of cumulative distribution, as shown in the equation (2.2). 4 Use the formula of general histogram equalization to calculate the new value. as shown in the equation (2.3). 5 Replace (original) gray values with (new) gray values to create a new Image: Img_{New}[E][F] = T1 [Img_{old} [E][F]] **END**

The second step, Bilinear interpolation is considered to be a significant approach applied for image resizing (zoom in/out). It does take a weighted average of four neighborhood pixels for calculating the final interpolated value, which will result an incredibly smoother image in comparison to the original one. This technique gives better result than nearest neighbor interpolation [35]. The bilinear interpolation process was implementation in this thesis by algorithm (2.2).

Algorithm (2.2): The Bilinear interpolation
Input: Face image $\boldsymbol{F}_{\text{Enhanced}}$ $\boldsymbol{RH_{in}}$: refers to the height of a facial image that has been entered. $\boldsymbol{CH_{in}}$: refers to the width of a facial image that has been entered. $\boldsymbol{MH_R}$: describe the row's minimization ratio $\boldsymbol{MH_C}$: describe the column's minimization ratio **Δrh, Δch**: a fractional part of the (row,column) location rh, ch **Output:** Minimized face image $\boldsymbol{F_{newsize}}$
BEGIN **1** $RH_{OUT} = MH_R * RH_{in}$ $CH_{OUT} = MH_C * CH_{in}$ **2** *for* $rh = 1$ *to* RH_{OUT} $rh_f = rh.\ MH_R$ **3** *for* $ch = 1$ *to* CH_{OUT} $ch_f = ch.\ MH_C$ **4** $rh = \lfloor rh_f \rfloor$ *and* $ch = \lfloor rh_f \rfloor$ **5** $\Delta rh = rh_f - rh,\ \Delta ch = ch_f - ch$ **6** $\boldsymbol{F_{newsize}}(rh, ch) = \boldsymbol{F}_{\text{Enhanced}}(rh, ch).\ (1-\Delta rh).\ (1-\Delta ch)$ $+\boldsymbol{F}_{\text{Enhanced}}(rh+1, ch).\ \Delta rh.\ (1-\Delta ch)$ $+\boldsymbol{F}_{\text{Enhanced}}(rh, ch+1).\ (1-\Delta rh).\ \Delta ch$ $+\boldsymbol{F}_{\text{Enhanced}}(rh+1, ch+1).\ \Delta rh.\ \Delta ch$ **7** Return $\boldsymbol{F_{newsize}}$ **END**

2.7 Approaches of the Feature Extraction

Feature extraction has become a clear requirement in numerous processes that have frequently to do in computer vision, image processing, object detection, image retrieval, pattern recognition, bioinformatics, machine learning. Feature extraction is utilized to extract features that special existing in a dataset (text, image, voice) which are utilized to perform and describe the data [36]. The feature extraction step of the face recognition system comprises the decreasing of the number of resources that describe large data amounts. Feature extraction is a technique for reducing the size of the original face dataset by extracting properties that can be used to identify and extract patterns from the input facial images [37].

When the algorithm's input data is very large for processing and expected to be redundant (a lot of data, but little information), the data is transform into a reduced representation set related to the features referred to as (features vector). Furthermore, feature extraction (FE) is the process that transforming the input data in to a collection of (features). It is assumed that when the extracted features are correctly chosen, the features collection will extract essential information from input data in order to complete the desired task using a reduced representation rather than full-size input.

Pattern recognition is regarded as one of the most promising areas of image processing science. Face recognition, postal address reading, text reading, extracting information from cheques, document authentication, health insurance, character recognition, credit card applications, recording bank deposit slips, loan, check sorting, data entry, and script recognition were among the applications it was used in Pattern recognition. Pattern recognition's main goal is to take input patterns and properly allocate them to a possible output class. The excellent feature set comprises discriminating information that can be distinguished the objects. It requirement be robust enough to prevent the generation of several feature codes for the same object class. Moreover, the features can be split into two separate groups: [38]

1) **Local features:** generally geometric (joints, number of endpoints, branches, branches, convex or concave parts).
2) **Global features:** these were usually statistical (invariant moments) or topological (number of holes, connectivity, projection profiles, and so on). Focusing on the feature extraction process is critical because it has a discernible impact on the efficacy of the recognition system.

Linear Discriminant Analysis (LDA) is studied in machine learning, statistics, and pattern recognition in widely. LDA can be considered as a Fisher’s linear discriminant (FLD), which is designed to find an optimal conversion to extract discriminant features that differentiate two or more classes. Its application of LDA has been kept in a small image database that used to overcome the limitations of PCA, was accomplished by the projection of an image onto Eigen-face space over PCA, then perform pure LDA through it for classifying Eigen-face projected data. LDA looks for vectors in implicit space that are better at discriminating between classes [39]. In contrast to the LDA group's images that are linked to the same class and the images that separate distinct class images.

Furthermore, LDA was compared to LDA feature space to determine eigen-decomposition relating to the appropriate matrix, given that all data were provided in advance. However, there were circumstances in which data was delivered in order and LDA features should be updated incrementally by identifying new incoming samples. In addition, the LDA approach's common implementation requires that all samples be provided ahead of time. In certain cases, however, the entire data set was not provided, and the input data was classified as a stream. In these circumstances, LDA feature extraction must be able to update the evaluated LDA features by detecting new samples instead of performing the algorithm on the complete data set

[40]. Furthermore, LDA variations could be used in a variety of study domains, including text recognition, bioinformatics, and face recognition. LDA determining the orientation decreases high-dimensional feature vectors from many classes to low-dimensional feature space, allowing feature vectors from other classes to be successfully separated from feature vectors from other classes, As can be seen below [41][42]: The scatter matrices S_W (within-class) and S_B (between-class) are generated for a set of C-classes where each class C_j has N_j face images:

$$\mu_{Cj} = \frac{1}{N_j}\sum_{k=1}^{N_j} x_k \qquad (2.4)$$

$$\mu = \frac{1}{M}\sum_{i=1}^{M} x_K \qquad (2.5)$$

$$S_B = \sum_{j=1}^{c}\left(\mu_{C_j} - \mu\right)\left(\mu_{C_j} - \mu\right)^T \qquad (2.6)$$

$$S_W = \sum_{j=1}^{c}\sum_{k=1}^{N_j}(x_{kj} - \mu_{C_j})(x_{kj} - \mu_{C_j)})^T \qquad (2.7)$$

Where (μ) is overall mean, (M) is the set of all face images, and μ_{Cj} is the mean of class Cj.

By increasing the between-class scatter S_B and decreasing the within-class scatter S_W, the goal of LDA is to discover (W_{opt})) an optimal projection.

W is the optimal projection which satisfies the following equation:

$$W = S_W^{-1} S_B \qquad (2.8)$$

$$W_{opt} = \arg\max_W \frac{|WTSBW|}{|WTSWW|} \qquad (2.9)$$

The eigenvectors and eigenvalues of $W = S_W^{-1} S_B$ S B must be solved to complete this optimization. For face projection, we employed the first K eigenvectors in the collection, similar to PCA. All of the LDA steps are detailed in Algorithm (2.3).

Algorithm (2.3): Linear Discriminant Analysis LDA
Input: Face image $\boldsymbol{F_{newsize}}$ **Output:** Feature Vectors
BEGIN 1 Read Face images 2 Using Equation (2.4), find the average value of each of the classes. 3 As in Equation (2.5), calculate the total mean of all databases. 4 Using Equation (2.6) to get the between-class matrix S_B (M×M). 5 Using Equation (2.7) to get the within -class matrix S_W (M×M). 6 In Equation (2.8), also known as Fisher's criterion, is used to calculate the transformation matrix (W) of the LDA approach. 7 The Eigen values (λ) are calculated, as well as the Eigen vectors (V) of W. 8 Sorting eigenvalues by their corresponding Eigen values in decrement order. As a lower-dimensional space V_k, the first (k) Eigen vectors are used. 9 As in Equation (2.9), apply each original sample (X) to the LDA's lower dimensional space. 10 Return the feature vectors **END**

2.8 Machine learning

One of the most widely used technologies is machine learning. The internet has altered people's lives and the way they conduct business over the previous century and a half, producing many petabytes of data in the process. Predictive analytics and machine learning have once again altered society by transforming data into positive predictions [43]. Machine learning is the study of automated systems for learning to make exact predictions based on previous observations, and it is broadly defined to include any automated computing process based on binary or logical operations that learns a task from a series of instances. The goal of Machine Learning (ML) is to create classification expressions that are basic enough for a human to understand. They must be able to sufficiently simulate humans reasoning in order to provide insight in to the decision-making processing.

ML is a method for designing computer systems that learn and improve automatically as they gain experience. [44] Data mining is one of the most essential applications for machine learning. During research or, in some situations, when attempting to build relationships between multiple features, people are prone to making mistakes. This makes it harder for them to find solutions to specific issues. In most cases, machine learning may be successfully applied to these issues, resulting in improved system efficiency and machine design. It's also capable of automatically extracting relevant information from a set of data by constructing suitable probabilistic models, and it's best suited to fields with a lot of data [45]. The same set of features has been utilized to represent each instance in each dataset used by machine learning techniques. There are three types of features: categorical, continuous, and binary. ML techniques can be classified into three basic types based on how they handle a collection of provided cases and features: Supervised, Unsupervised, and Reinforced [46]

1- Supervised ML

Supervised machine learning (supervised ML) is a method for learning a function of the training data, which consist of pairs of the input objects (usually vectors) and the desired outputs. The function output could be a continuous value known (regression) or could be predicting an input object class label (which has been referred to as the classification). After seeing numerous examples of training, the supervised learner must predict the function value for any of the valid input objects. In classification and regression, supervised learning uses labeled instances (i.e. X & Y) for the prediction of their relationships.

2- Unsupervised ML

Unsupervised learning algorithms, unlike supervised learning algorithms, do not require the data to be labeled by a human expert. Unsupervised approaches take unlabeled features from the input data and classify it using self-taught rules. As a result, models are frequently used to identify unknown or hidden relationships in data. Unsupervised machine learning is a more difficult process in which the examples are left unlabeled in order to learn about their distributions.

3- Reinforced learning

Unlike supervised learning, which uses labeled data, reinforced learning algorithms just use training data to determine whether or not they are correct. They learn "good" behavior by interacting with their environment iteratively. They train using supervised learning principles, but instead of having a large volume of labeled data, the model must “interact” with the environment, which results in a positive reward or negative penalty. This feedback reinforces the model's behavior, providing it the name. Exploration and exploitation are phrases used frequently in reinforcement learning algorithms to refer to performing action that produces the largest potential reward exploitation and doing action that has not been conducted before exploration.

2.8.1 Machine Learning Classification Algorithms

When it comes to pattern recognition systems, classification is crucial in determining which form belongs to which class. Classification's main goal is to assign instances (forms) to unknown pre-defined classes based on their descriptions as parameters. Crossover was used in the classification process. One partition is utilized for testing and K-1 partition is used for training in each iteration. There are a variety of machine learning classification and prediction algorithms as show in figure (2.5), and this section will cover a few of them in detail.

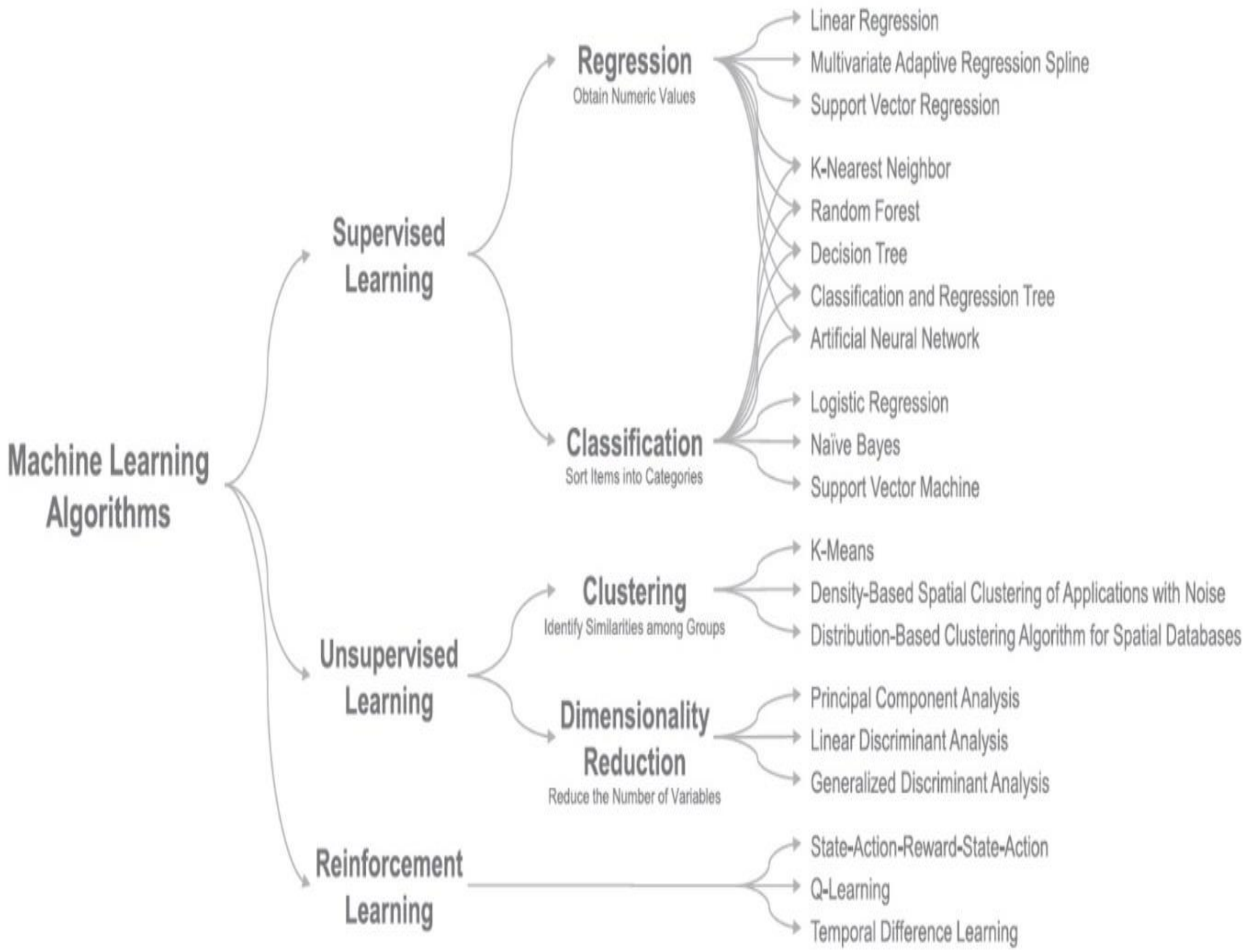


Figure (2.5): Various divisions of ML algorithms according to learning techniques, with examples**.** [47]

1. **Naïve Bayes (NB):** The Naive Bayesian classifier is a simple probabilistic classifier that creates a set of probabilities by computing the frequency and collections of values in a given data set. The naive Bayes (NB) supervised classifier is a probabilistic model that uses the joint probabilities of terms and categories to estimate the probabilities of categories The naive Bayes learner combines Bayesian reasoning with the assumption that the measurable features are independent. It estimates the class probability and the conditional probability distribution of a class given the feature values using standard probability distribution methods to learn relative frequencies of different classes and feature values in the training data [48].

   The beauty of the nave Bayes strategy is that it separates the estimation of one feature distribution from the estimation of others. Bayesian classifiers assign the most likely class to a given example defined by its feature vector and have been shown to be effective in a variety of applications, including text classification, medical diagnosis, and system performance management. The following is how the Nave Bayesian works: Given C_n classes and each one of these classes has its own probability $P(C_n)$ evaluated from the training dataset and show the prior probability of classifying an attribute V_j into C_n. For attribute value, V_j , so the classification utilized is to find this probability is illustrated in the equation bellow:

$$\frac{P(v_1 \wedge v_2 \ldots\ldots v_j \mid c_n)P(c_n)}{P(v_1 \wedge v_2 \ldots\ldots v_j)} \qquad (2.10)$$

   The purpose of image classification is to determine the best class for the image. As demonstrated in equation (2.10), the Naive Bayes classifier predicts the class with the highest posterior probability [49]. The procedure detailed in the algorithm (2.4).

Algorithm (2.4): Naive Bays (NB)
Input: Read the data set (Feature Vectors) **Output:** the accuracy of NB (evaluation matrices)
BEGIN 1 Split the data set into two parts: training and research. 2 $P(C_i)$ = For every class in the training set, calculate frequency and probability. 3 $P(F_i)$ = Using classlable, calculate frequency and likelihood for every feature in the entire dataset. 4 For every Fi in the Feature Set 5 BF_i= get the probability value for a feature from the training set (F_i) 6 BF_i (C) = calculate the probability (C) 7 Apply equation (2.10) 8 Next F_i 9 Class = class that has the maximal posterior probabilities.. 10 Return evaluation matrices **END**

2- **K_Nearest Neighbour:** KNN it's a supervised learning algorithm and is considered one of the most common instance-based learning approaches in pattern recognition. Due to the lack of a training phase, K-NN is a lazy learner that performs well when all of the data has the same scale. The KNN's ease of use has made it one of the most popular tools for classification in a number of applications. When classifying a sample Si, for example, the algorithm searches feature space for the K nearest neighbors using feature vectors and a given distance. The algorithm then votes on those neighbors based on their labels. The sample of the object will be classified into a group with the most neighbors that have the same label. [50]. The K-NN method has used a variety of distance metrics, the best of which are:

Euclidean $$D(p,q) = \sqrt{\sum_{i=1}^{k}((p_i - q_i)^2} \qquad (2.11)$$

Manhattan $$D(p,q) = \sum_{i=1}^{k} |p_i - q_i\,| \qquad (2.12)$$

Which (p,q) vectors space with fixed with Cartesian coordinate system. The Algorithm (2.5) show the steps of KNN classifier.

Algorithm (2.5): K-Nearest Neighbour KNN
Input: initialize the (K), Feature Vectors **Output:** the accuracy of the KNN (evaluation matrices)
BEGIN 1 Chose K nearest-neighbour sample. 2 Using the training set Equation (2.11), calculate the distance between each sample in the testing set. 3 Based on the distance, choose the K samples that are nearest. 4 Calculate the maximum number of samples based on their class. 5 Determine class label. 6 Return evaluation matrices **END**

3- **Random Forest:** RF is a supervised ensemble learning method for classification that is fairly easy. A forest is built by assembling several trees for prediction and decision-making. For dividing a node in the forest, the best of the random subsets of features is chosen. This algorithm creates a large number of decision trees that work together. [51] In this method, decision trees serve as pillars. The term "random forest" refers to a collection of decision trees whose nodes are defined during the pre-processing stage. The best feature is chosen from a random selection of features after numerous trees have been constructed.

Another notion that is produced utilizing a decision tree technique is to generate a decision tree. As a result, a random forest is made up of these trees that are used to classify new objects based on the input vectors. Each decision tree constructed is used to classify data. If we assign tree votes to that class, the random forest will select the classification with the most votes from all the trees in the forest. There are numerous advantages to adopting the Random Forest algorithm, including the following: (1) It can be used for classification as well as regression tasks, (2) can deal with missing values, and (3) it can handle large data with high dimensionality.[52] The main idea behind this algorithm is illustrate in Figure(2.6):

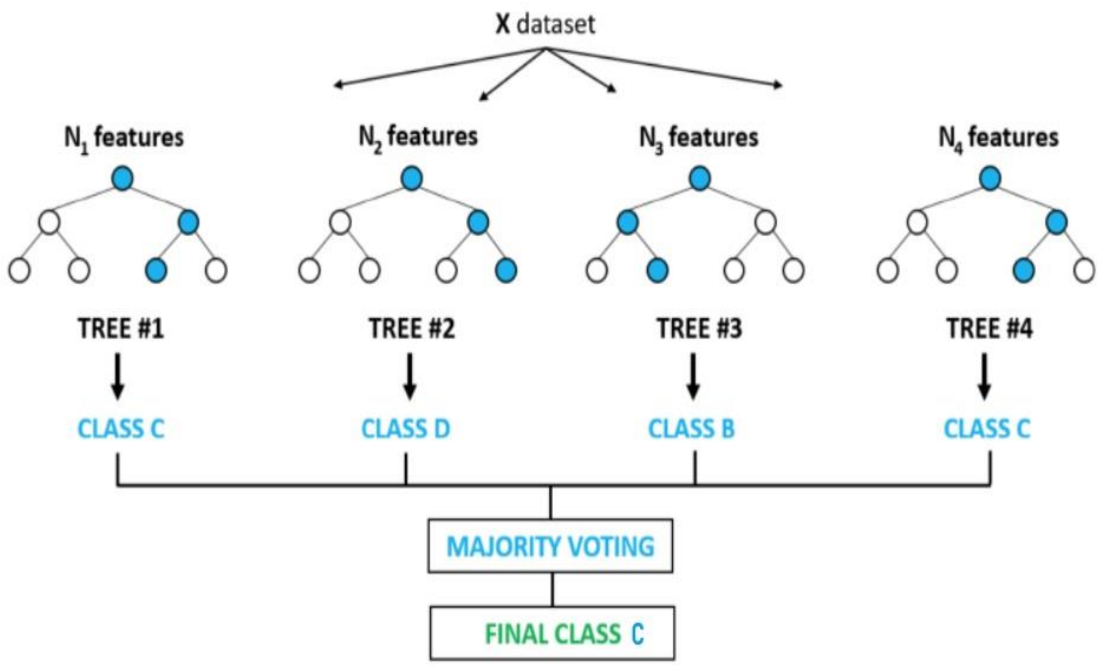


Figure 2.6: Random Forest's Main Idea [53]

The Algorithm (2.6) show the steps of RF classifier.

Algorithm (2.6): Random Forest RF
Input: Feature Vectors **Output:** the accuracy of the RF (evaluation matrices)
BEGIN 1 (k) Randomly selected features from a total of (m) features. K<m 2 Using the best-split point, compute the node (d) among (k) features. 3 Using the best-split, split the node into daughter forms. 4 Repeat steps 1–3 until (I) the number of nodes reaches . 5 Create a forest by repeating steps 1–4 for a (n) number of times, resulting in a (n) number of trees. 6 Determine the class label. 7 Return evaluation matrices **END**

4- **Support Vector Machine (SVM):** is a supervised machine learning method that uses a set of training examples, each of which is labeled as belonging to one of several classes, to try to build a model that can predict the class of a new example. SVM is frequently used to classify patterns. This means that the method is mostly used to classify different sorts of patterns. Linear and non-linear patterns are the two types of patterns that exist in general. Patterns that are easily distinguishable are known as linear patterns. In other words, patterns can be divided simply into low dimensions. Non_linear patterns, in contrast to linear patterns, are patterns that are difficult to distinguish or separate, necessitating additional manipulations to separate them [54]. In the SVM classifier, each data item is plotting as a point in (n-dimensional) space, with random hyper-plans defined. The primary concept of the SVM classifier is to find the best hyper-plane (a linear decision surface which divides space into two regions) for classifying linearly separable patterns [55]. The space of SVM is made up of (two) different patterns, and the purpose of the (SVM) is to separate them as illustrate in Figure (2.7).

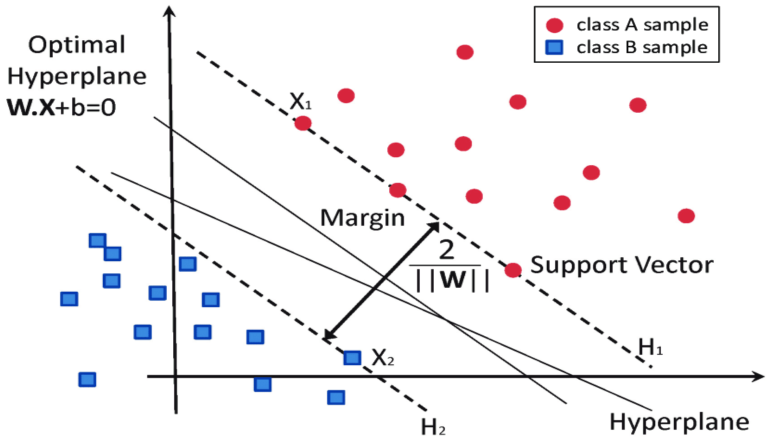


Figure 2.7: Space representation in SVM.[55]

The SVM model typically consists of four lines. The chosen hyper-plane is represented by the first line (w.x+b=0), The second line is referred to as the separation margin or marginal line. The limiting hyper-planes (w.x + b = 1) and (w.x + b = -1) are the two other lines on either side of the marginal line. These lines come together to generate the best hyper-plane for separating the provided patterns and patterns on the hyper-boundary, planes known as support vectors. The vertical distance between the marginal line and the hyper-boundary is the margin. plane's One of the SVM's purposes for accurate classification is to maximize the margin of the training data, which is one of the SVM's goals [56]. The data points must be stated as given in the Equations to achieve the "maximum-margin hyper-plane" that is found in the (middle) between the two patterns:

If $(w.x_i + \mathrm{b} \geq 1)$ Then $(Y_i = +1)$ (2.13)

If $(w.x_i + \mathrm{b} < 1)$ Then $(Y_i = -1)$ (2.14)

If $(w.x_i + \mathrm{b} = 0)$ Then $(Y_i = 0)$ (2.15)

Where x is a point vector, w is the weight of the margin vector, and b is the bias. The data is labeled with one of two labels that correspond to the two classes: +1 (positive patterns) or -1 (negative patterns). Equation (2.13) should continually be larger than zero in order to separate the data. SVM chooses the hyper-plane with the largest distance among all potential hyper-planes. If the chosen hyper-plane is placed as far away from data as possible. As shown in Figure (2.8), this hyper-plane maximizes the margin and separates the lines between the closest points on the convex hulls of the two patterns.

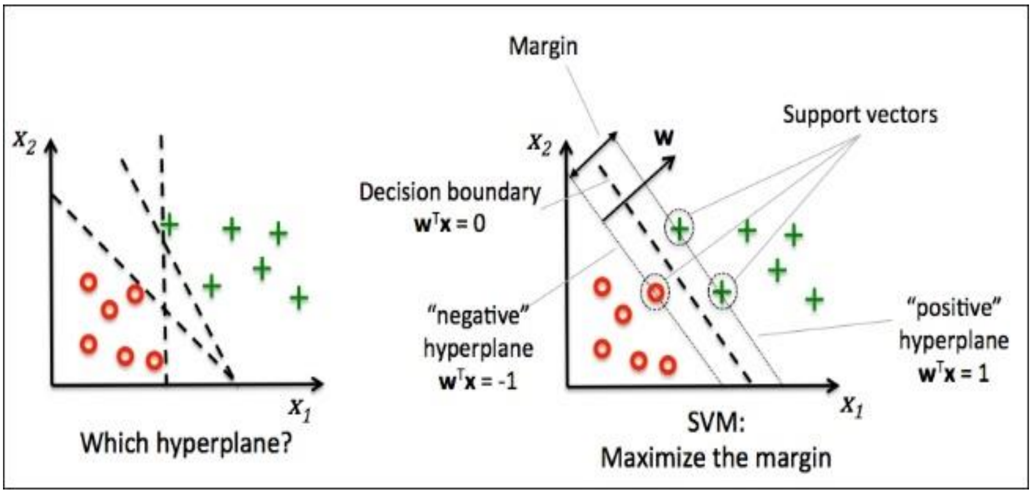


Figure 2.8: hyperplane with greatest margin, Hyperplane representation (left), hyperplane set (right) [57]

The distance within these two hyperplanes is (2 / ||w||) in mathematics. As demonstrated in Equation (2.16), ||w|| should be minimized beside the condition that not data points exist between margins to maximize the distance between the planes.

Minimize $||w|| = \frac{1}{2}\sum_{i=1}^{n} w^2 i$ Subject to: $Y_i\ (w.x_i\ +\ b) >=\ 1$ (2.16)

Where (i=1 to n) total number of patterns in the dataset.

Algorithm (2.7): SVM
Input: Feature Vectors **Output:** the accuracy of the SVM (evaluation matrices)

BEGIN

1 Create an SVM space (S) that has a set of Feature Vectors for all facial images in training set, every of which is representing as a point.

3 In (S), create a random set of hype-planes and choose optimal one as:

- Choose two limiting hyperplanes for each hyperplane that split the set of vectors of features with none points in between, as determined in equations (2.14) and (2.15).
- Minimize the width of the hyperplanes as described in equation (2.17). to get the maximum distance between the two limiting hyperplanes.
- Place the resulting hyperplane (H) as optimal one.

4 Classify set of Features in the (S) used (H) Optimal as:

- Calculate whether the Features vector is above or below the the (H) Optimal.
- If the return value of the (H) Optimal equation is greater than 0, the point is above the (H) Optimal and belongs to the specific class.
- If the return value of the (H) Optimal equation is less than 0, the point is below the (H) Optimal and belongs to the other class

5 Return evaluation matrices

END

5- Neural Networks (NNs): The standard NN consists of multiple simple, connected processors known as neurons, each of which generates a set of real-value activation types. The input neurons are activated by sensors that detect changes in the environment, while the other neurons are stimulated by weighted connections from previously active neurons. Through trigger actions, any neurons would be able to influence the environment. Finding the weights that make NN to present desired actions, as driving a car, is the focus of credit assignment or learning. NN is capable of completing many classifications and/or regression tasks simultaneously, despite the fact that each network normally only performs one [58].

The ANN is based on three basic factors: the network's architecture, unit input, and activation functions, and the weight of each input connection. Given that the first two factors are fixed, the ANN's behaviour has been described using the current weight values. The weight values of a network that needs to be trained are set to arbitrary values at first, and then training set instances are frequently exposed to this network. The input values for each instance have been assigned to input units, and the network output is compared to the instance's required output. Following that, each weight value in the network is slightly updated in the direction of bringing the network output values closer to desire output values [59]. Neural networks typically contain three fundamental layers, each of which is made up of nodes. A node is just a place where computation takes place, as seen in figure (2.9), which depicts each node's diagram. The input, hidden layer, and output layer are the three main layers of a neural network.

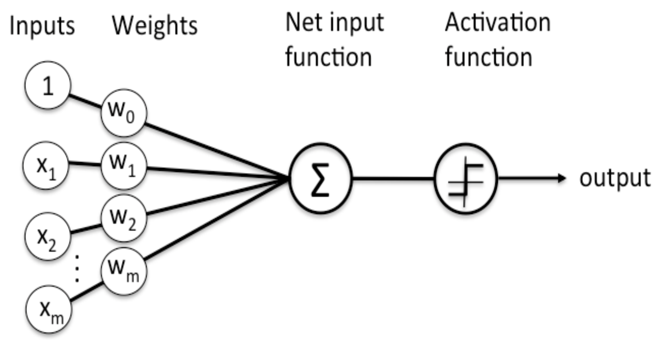


Figure (2.9): Block diagram of the (NN) nodes [60]

Figure (2.10) shows an ANN architecture example with different layers, weights, and different types of activation functions that can be applied to the weighted sum.

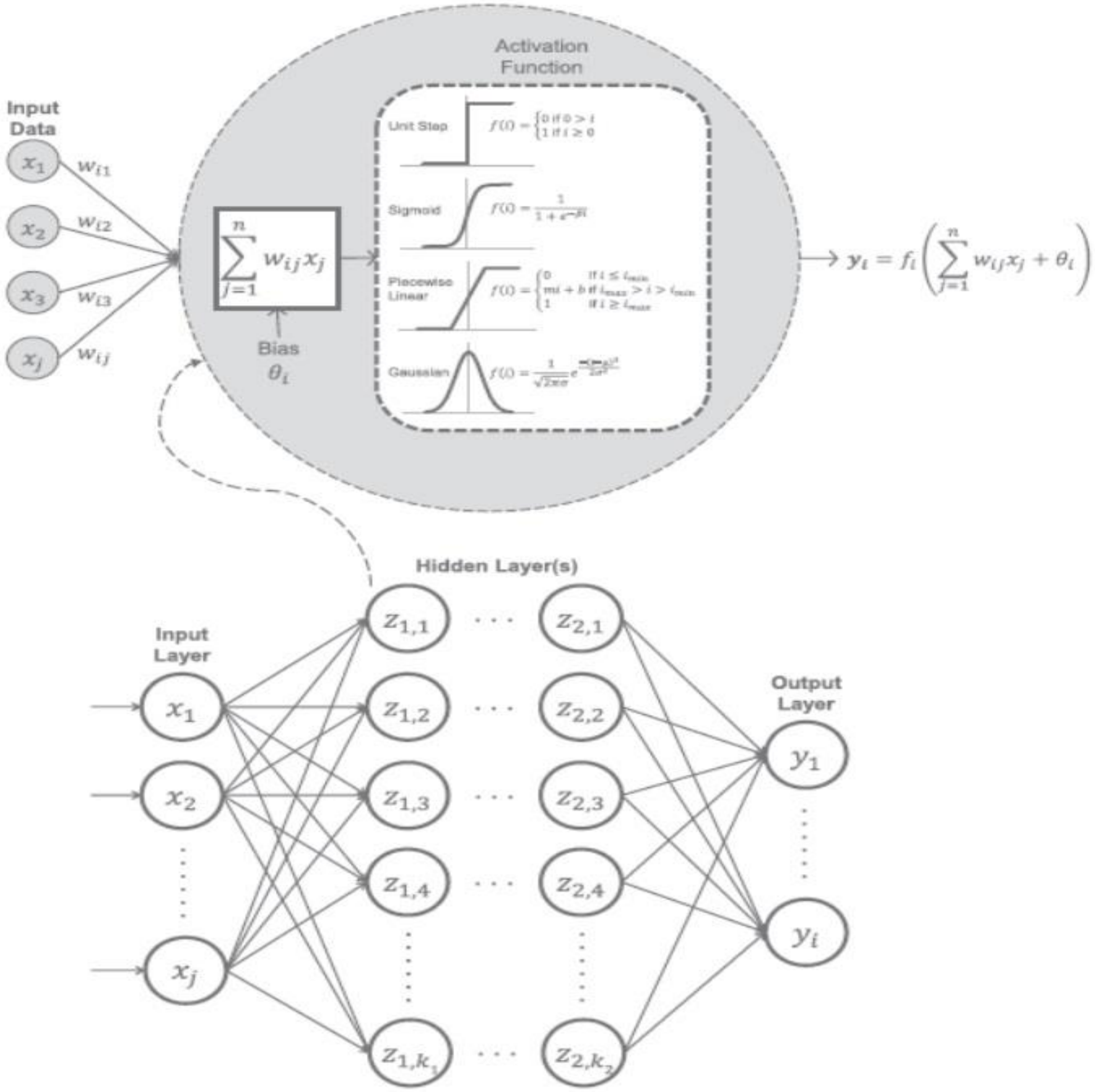


Figure (2.10): ANN architecture example [47]

2.8.2 Deep Learning (DL) in Neural Networks

DL comprises a wide variety of machine learning architectures and methodologies, with are characterized by the usage of multiple layers from nonlinear information processing steps that hierarchical in nature. Most architectures might be classified into three collections based on (how) the methodologies and architectures were suggested to be employed. [61] The architecture types are illustrated in Figure (2.11).

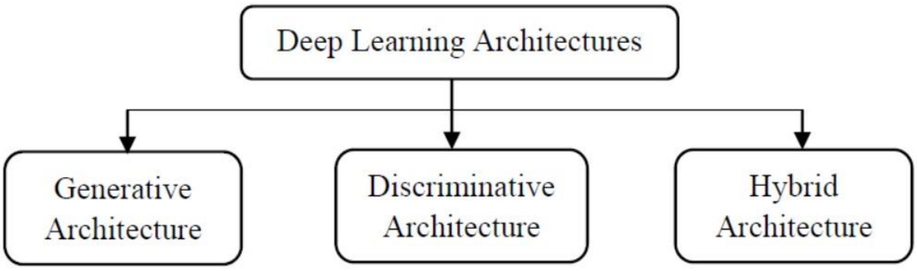


Figure (2.11): Types of DL Architecture [61]

1) **Generative architectures:** Because data labels are not taken into consideration in this strategy, Unsupervised feature learning models are generative architecture-based deep learning models. The basic notion in generative deep learning architectures is supervised learning and unsupervised pretraining. This form of architecture evolved when there was little data to train a complicated network. Without relying on the other layers, these models learn the lowest level of input and provide the necessary solutions.
2) **Discriminative architectures:** As a shallow architecture, they are extensively used in information and signal processing, alongside Hidden Markov Models and Conditional Random Fields. Deep structures with conditional random fields have recently been constructed, in which the output of each lower layer of random field is stacked with the original input data on the higher layer. Discriminative architectures are widely used in language processing and recognition applications, and they have generated plenty of research papers on the subject.
3) **Hybrid architectures:** Both generating and discriminative processes are included in hybrid systems. In most hybrid architectures, generative and discriminative components are used to arrive at the final solution. Initialization issues are reduced since generative models are used to handle nonlinear parametric challenges. Furthermore, generative models provide regularized control characteristics, reducing the system's complexity.

Hybrid deep architectures, in which the goal is discrimination but the results of generative structures are aided (typically significantly) by greater better optimization or/and regularization.

The neural network community came to the conclusion in the early 1990s that training multi-layered networks with backpropagation, or really any gradient following algorithm, was nearly impossible, and that solutions obtained with deeper neural networks starting from random initialization performed worse than networks with (1) or (2) hidden layers. Weights in multi-layer networks tend to decline to zero or increase without bound, and the networks have a large ratio of saddle points to local minima, as proven. The phrase "deep learning" refers to "stacked neural networks," or networks with multiple layers, which is the primary distinction between the original neural network and the one utilized in deep learning (as depicted in Figure (2.12)). Deep learning algorithms have provided useful tools for dealing with big data analysis because of their capacity to handle large amounts of unlabeled data. [58] The conventional neural network is one of the most important types of neural networks used in deep learning.

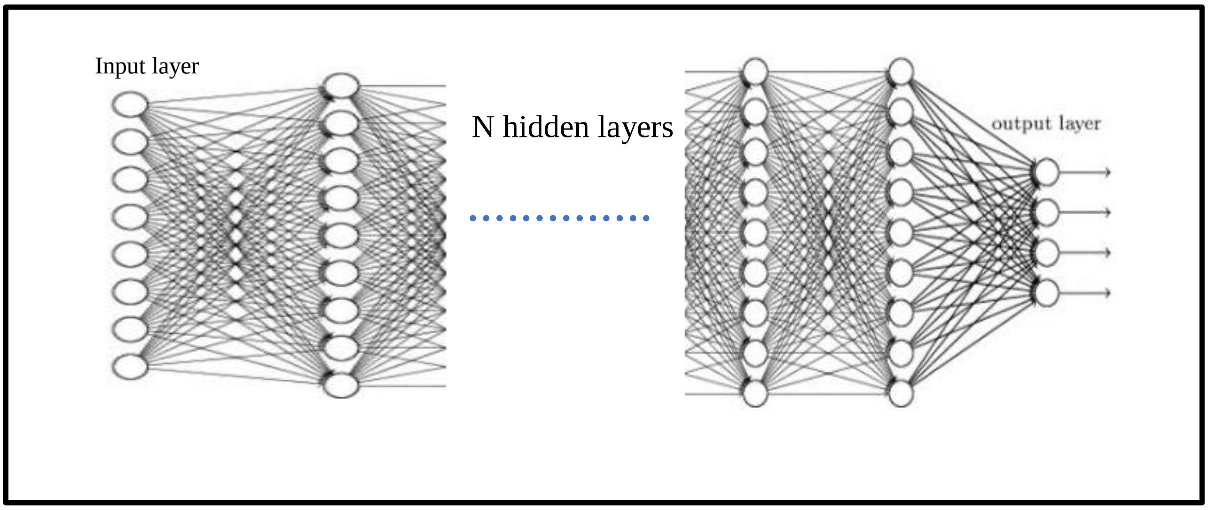


Figure (2.12): layers of deep neural network [58]

2.9 Convolutional Neural Network (CNN)

CNNs are a kind of discriminative architecture that has shown to be effective at processing 2-D data beside a grid-like topology, such as videos and images. Deep 2-D Convolutional NNs with millions of parameters and multiple hidden layers are capable of learning complex patterns and objects, allowing them to be trained on a large visual database with ground-truth labels. Such distinctive capability has made them the key tool for a number of 2-D signal engineering applications such as video frames and images with appropriate training. The discovery from a visual mechanism, visual cortex inside the human brain, which includes multiple cells, responsible for the detection of light in smaller, overlapping sub-areas of the visual range referred to as (receptive) fields, inspired the CNN design. These cells serve as local filters across input space, with larger receptive fields in the higher complexity cells. [62]

2.9.1 The Architecture of CNN

CNN is a multi-layer NN that consists of two distinct layer types: convolutional layers (i.e. c-layers) and sub-sampling layers (i.e. s-layers) that are alternately connected and constitute the middle net part as illustrated in Figure (2.13). In general, in the state when each neuron's input is connected to the prior layer's local receptive range, the c_layers have been used to extract the features. After obtaining each local feature, the relationship between their positions may be determined.

The s-layer is essentially a feature mapping layer. Those layers combine their weight values to form a plane. The pooling process, which has also been referred to as down-sampling or sub-sampling, has been presented for reducing the total size of the signal. In fact, sub-sampling has been successfully used to reduce the size of data in audio compression. Down-sampling was also used in a 2D filter to increase the invariance to the positions. A function of pooling is used to replace the network output at a given location, and the max-pooling strategy has been demonstrated through the summary of neighboring output statistics.

The usage of max-pooling allows to get the most output in the rectangle area around you. The pooling process may also result in a representation that is invariant to input translations. The insertion of a max-pooling layer between convolution layers currently results in an increase in spatial abstractness as the abstractness of the features increases. In convolutional NNs, rather than establishing parameters like in traditional NNs, all that is required is to train filters. Furthermore, the CNNs are not reliant on human intervention or prior knowledge for feature extraction [63]. Computer vision, speech recognition, Object recognition, image classification, face detection, behaviour recognition and handwriting recognitions are just a few of the applications where CNNs are used.

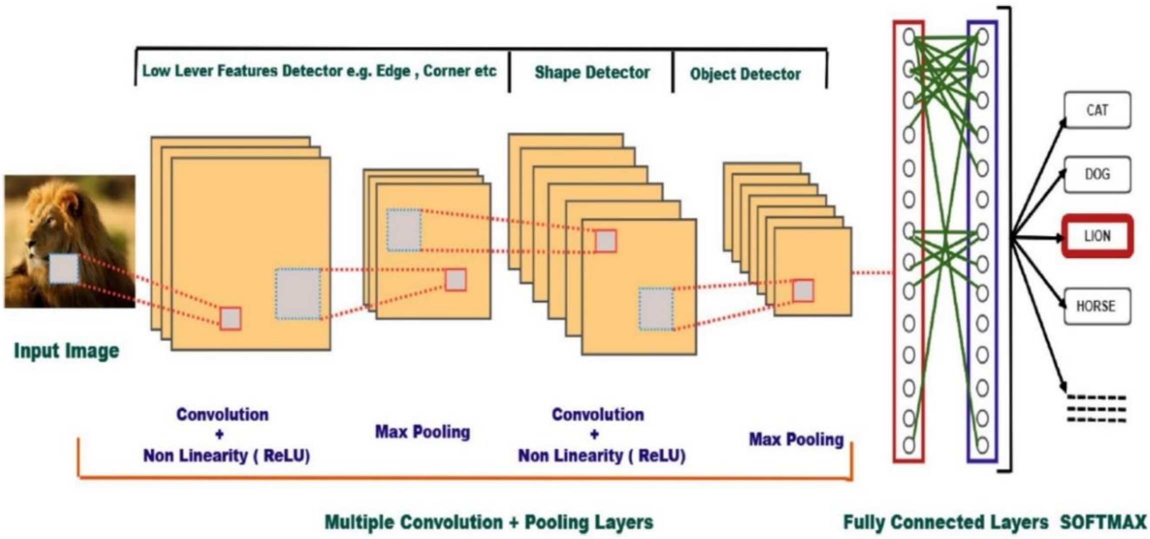


Figure (2.13): Architecture of CNN [63]

The activation function works as a mathematically “gate” among the current neuron's input and output, which is given to next layer. A simple step function that toggles neuron output based on a rule or threshold could suffice. It could also be a transition that changes the neural network's input signals into the output signals it requires to function. In neural networks, non-linear activation functions are increasingly being employed to allow the network to learn complex data, compute and learn nearly any function representing a question, and generate accurate predictions. Sigmoid / Logistic, Tan, ReLU, Parametric ReLU, Leaky ReLU, Softmax, and Swish are examples of nonlinear activation functions.

2.9.2 The CNN Advantages

Because of the advantages it brings, CNN has recently piqued the interest of many academics, particularly in the field of image processing. The following are some of the advantages: [64]

1- **Ruggedness to shift and distortion:** The CNN-based identification is resistant to distortions as changes in shape caused by the variable lighting conditions, camera lens, partial occlusions, horizontal and vertical shifts, different poses, and so on.

2- **Low Memory Requirements:** Fully connected layers can theoretically be used to extract all of the features. For example, if a picture of size 32×32 has 1000 hidden features, then 106 coefficients order is required, using a considerable amount of memory. However, in the convolutional layer, these coefficients will be employed in various locations across space, resulting in a significant reduction in memory use.

3- **Better Training:** The number of parameters in CNN is considerably reduced. As a result, when compared to traditional neural networks, CNN takes less time. Furthermore, if a neural network is created to be similar to CNN, the normal neural network may have more noise during training because it has more parameters than the CNN, and its performance is lower.

2.10 Performance Measures

Several metrics have been used in the system and are designed for evaluating system performance [65].

1) **Precision:** indicates the number of true positives divided via number of true positive states, as well as number of false positives, which are instances that the model accurately identifies as positive but are in fact negative, such as individuals classified as terrorists but who are not. The ratio of (true positive) to total predictions is known as precision. This can be mathematically expressed as

$$\textbf{Precision}=\frac{TP}{TP+FP} \qquad (2.17)$$

TP: True Positives

FP: False Positives

2) **Recall:** The ability to locate every relevant example in a dataset; precision describes the percentage of data points that this model claims are relevant that are actually relevant. The ratio of the number of correct predictions and all correct observation in the sample space. Mathematically:

$$\boldsymbol{Recall} = \frac{\boldsymbol{TP}}{\boldsymbol{TP} + \boldsymbol{FN}} \qquad (2.18)$$

TP: True Positives

FN: False Negatives

3) **F-measure or F1-score (F1) :** Is the weighted harmonic mean of precision and recall. When the dataset is unbalanced, this evaluation metric is normally used. F-measure it is used to evaluate the model's performance when the class distribution is unbalanced. On a more serious note, the greater the F-measure, the better the results. Mathematically:

$$\mathbf{F_1} = \frac{\mathbf{1}}{\frac{\mathbf{1}}{\mathbf{recall}} + \frac{\mathbf{1}}{\mathbf{precision}}} = \mathbf{2} * \frac{\mathbf{precision} * \mathbf{recall}}{\mathbf{precision} + \mathbf{recall}} \qquad (2.19)$$

4) **Accuracy:** The ratio of the (correct predictions) to the total observations is known as accuracy. For the two-class problem, the accuracy of a model is deemed best if and only if we have a symmetric dataset in which the values of FP and FN are almost equal. In many and imbalanced data sets, accuracy is not the best option; therefore, other evaluation measures, such as the F1-score, may be evaluated. Mathematically:

$$\boldsymbol{Accuracy} = \frac{\boldsymbol{TP+TN}}{\boldsymbol{TP+TN+FP+FN}} \qquad (2.20)$$

TP: True Positives

FN: False Negatives

2.11 K-fold Cross-validation

It is a resampling approach used to evaluate machine learning models on a small set of data. That is, a small sample size is used to test how the model would perform in general when used to make predictions on data that was not used during training. The algorithm has only one parameter, k, which determines how many groups a given data sample should be divided into. As a result, k-fold cross-validation is a common name for the procedure. When an exact number for k is stated, it can be used as the model's reference value, such as k=10 for 10-fold cross-validation [66]. It's an useful method since it's easy and provides a less biased or optimistic evaluation of model capability than other methods, as a basic train/test split. Algorithm (2.8):

Algorithm (2.8): K_Fold cross validation
Input : data set, value (k)
Output: split data to training dataset, testing dataset
Begin : 1 Shuffle the dataset in a random order.. 2 Divide the dataset into k groups. 3 For each distinct group: ✓ Select the group being a holdout or test dataset. ✓ Select the remaining groups being a training data set. ✓ Fit a model upon the training set and estimate it upon the test set. ✓ Keep the evaluation score but discard out the model. 4 Summarized the model's skill using model evaluation scores **End**

.

Chapter Three

The Proposed System Design

Chapter Three

The Proposed System Design

3.1 Introduction

The proposed system for person face recognition under disguise and makeup methods and all followed processes are described in this chapter. A comprehensive description of all the steps that were used in the proposed recognition scheme, including the algorithms, a description of each step, and its aid in achieving the thesis's goal will be presented.

3.2 Proposed System Design

Human recognition based on facial pictures has recently gained a lot of attention and is becoming increasingly relevant in people's social lives. The design of every system is critical because it demonstrates how the system works and explains the exact steps and processes that will be followed to achieve the desired result.

The proposed recognition system classifies people based on their facial images, which are significant and well-known biometric traits. Figure 3.1 explains a block diagram of the recognition system that describes the phases of the proposed recognition system implementation for recognizing people based on its input facial images. Figure 3.2 shows the classifiers building steps.

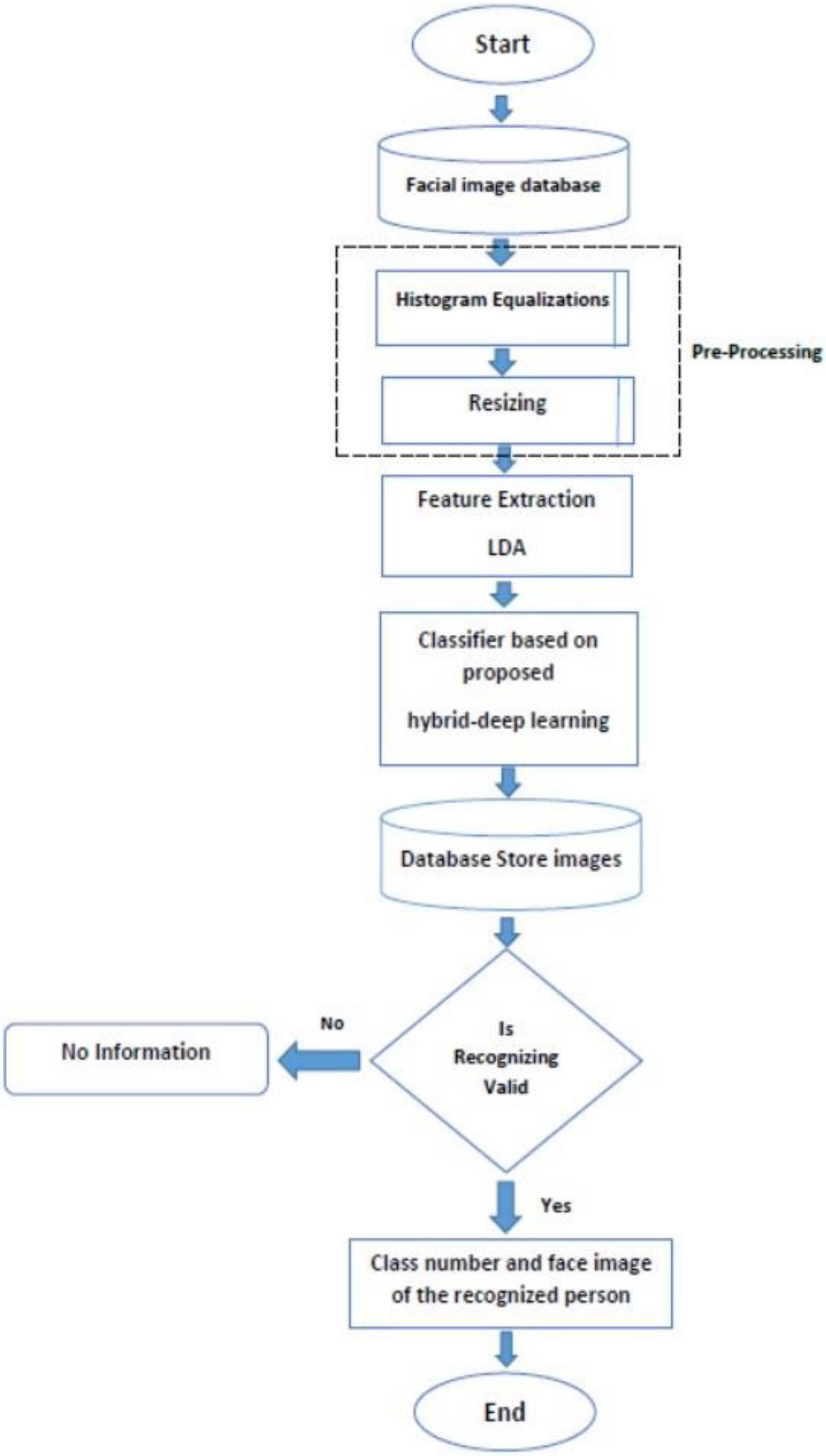


Figure (3.1): Flow chart of the proposed recognition system

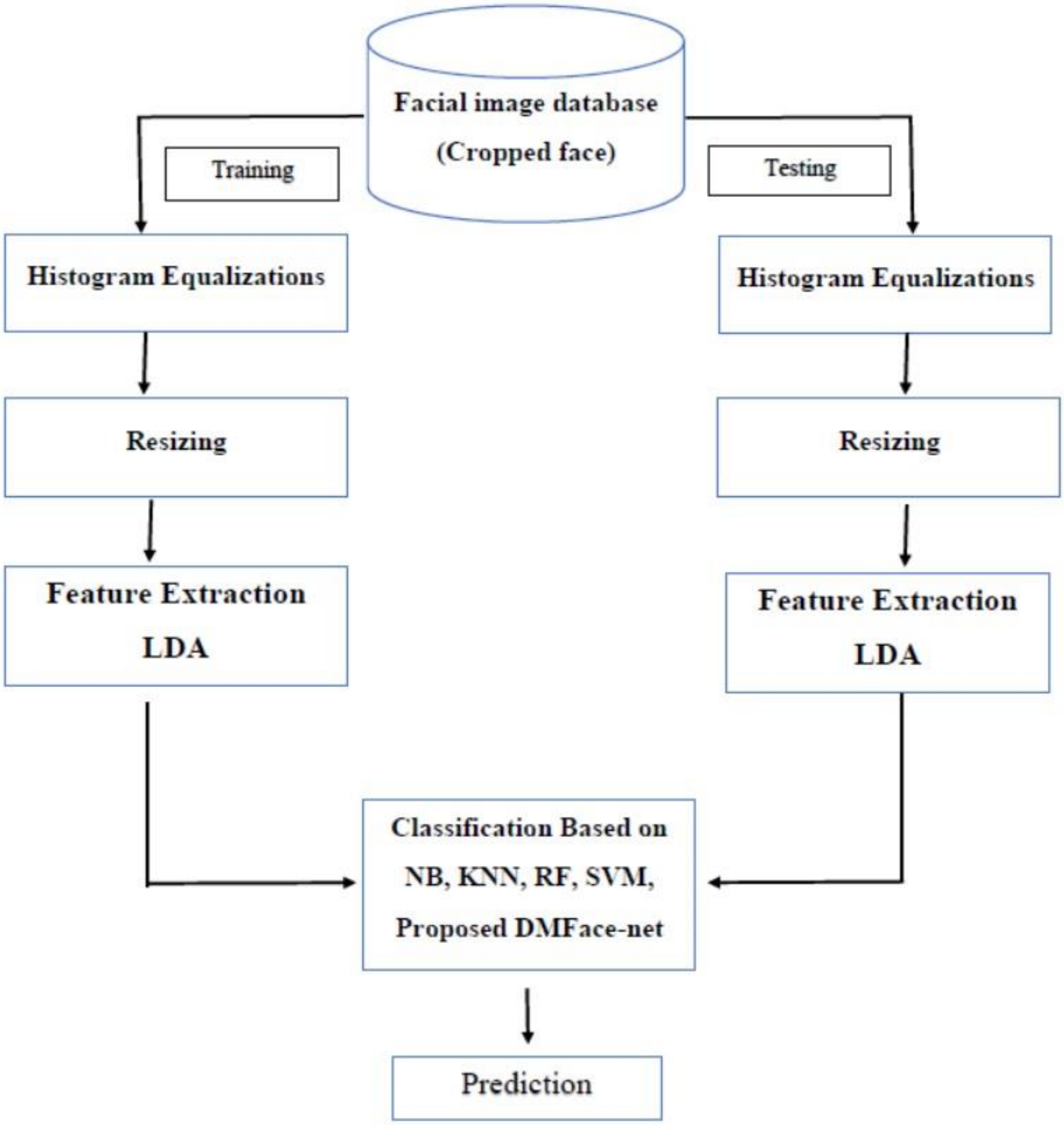


Figure (3.2): Framework of classification system

3.3 Proposed Methodology

The proposed methodology involves four main parts, as follows: Preparing Dataset, Preprocessing, Features Extraction, and Classification and identification. The first part, which is the Preparing Dataset, which includes the process of collecting and organizing data and creating training and testing sets. The second part is data pre-processing. It includes all data processing before starting the testing process and training. Includes Image Enhancement (Employ the Histogram equalization technique when the digital face images suffer from either have a lower contrast value like as a lack of illumination or a non-formal distribution of image illumination, therefore, the input gray image contrast will be enhanced by the usage of the cumulative histogram equalization method). As well as resizing face images (facial images are resized of various dimensions pixel are decreased to a smaller size of (30×30 pixels) applying bilinear interpolation approach).

The third part is extracting images features from a facial image are obtained by used Linear Discriminant Analysis (LDA) which represents images as a vector of features. LDA looks for the best vectors that discriminating among classes in the underlying space. LDA groups images belonging to the same class together and separates images belonging to different classes.

The fourth part in this thesis is classification. Its methods are used in two stages (1- four machine learning classifiers and pre-trained model (AlexNet), 2- one proposed deep learning classifier). In terms of machine learning, the extracted features will be fed into a variety of machine learning algorithms in order to generate predictions and pick the best one. Experiments on Navie Bays (NB) call algorithm (2.4), K-nearest Neighbor

(KNN) call algorithm (2.5), Random Forest (RF) call algorithm (2.6), and Support Vector Machine (SVM) call algorithm (2.7). To construct a classifier model (trained) and obtain the best precision was used K-fold cross-validation (call algorithm (2.8)). The database will be split into (k) sets; this thesis used, k=10, dividing the database into (10) equivalent sets. The model uses 9 (k-1) sets as the training set and (1 set) as the testing set for each of the ten folds.

The second stage in classification is by using the proposed hyper-deep learning classifier Model. This model is trained to recognize face images for a specific face group. This model is described in detail in the following sections. The final part is the identification Model. This model performs the matching between the vectors obtained from the proposed classification Model with the actual vector image obtained from the Features extraction Method. Generally, this model decides and recognizes the class and face of the image. The overall system structure was illustrated in the Figure 3.1.

3.4 Classification based on the hybrid deep learning

Convolutional neural networks are one of the most common forms of neural networks used to perform image recognition. The hybrid CNN receives an input image, processes it, and then it classifies using a specific denomination. Each image that input will be processed through a sequence of convolution layers including filters (Kernels), pooling, Dense layers and fully connected layers (FC) in order to classifed an object with probabilistic values ranging from 0 to 1. The CNNs work in the similarly way whether they hold one, two, or three dimensions. Structure of the data that is inputted, as well as how this filter known as (convolution kernel) or (feature detector) flows within the data, demonstrate the differences in processing. In this thesis, the hybrid-deep learning classification technique is employed for more features learning capability. The obtained hybrid deep learning layers and parameters are one-dimensional and will be described and explained as follows:

1. **Convolution layer 1D:** Its Modified version from 2D Convolutional Neural Networks named 1D, that developed recently. The benefit of 1D CNNs is that they have a low computational requirement, and since they only have 1D convolutions, they are easy to configure, have a compact shape, and need low-cost hardware. To generate a tensor of outputs, the 1D conv layer obtains a convolution kernel that travels across just a single spatial dimension.
2. **Max Pooling:** Pooling layer is the new layer that is added after the convolutional layer, and it is added after the convolutional layer has performed nonlinearity on the feature maps it has obtained. Maximum output can be achieved by using the Max Pooling method. The representation can also be done invariant to the input translations utilizing the pooling technique. By placing a (max-pooling) layer between the convolutional layers, spatial abstractness rises in tandem with feature abstractness, and the highest value for each patch in the feature map is computed. For 1-dimensional temporal data, the maxpooling approach minimizes the representation of input data by achieving the highest value through a window known as pool-size. Each window is shifted by a (strides).
3. **Dense layer:** It is the most widely used and obtained layer, and also a connected neural network layer in deep learning. In general, the Dense layer performs the operations described above on the data inputted,

and the shape of the Dense layer's output is influenced by a community of neuron/units inserted in the Dense layer.

4. **Activation Function (ReLU) :** The activation functions are simply a mathematical equation that determines the output of a neural network. This role is related to each neuron in a network and determines if it needs to be activated (fired) or not, based on whether the feedback of each neuron is relevant to the model's prediction. It's also assists to normalize each neuron's output to a value between 1 and 0, or even -1 and 1.

   Rectified Linear Activation Function (ReLU): It is a form of non-linear activation feature that can be attained in multi-layer or deep neural networks. This function is expressed as:

$$f(x) = \begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases}$$

   Due to the **ReLU** derivative, which has a value of 1 for positive input, the **ReLU** function can make deep neural networks training faster than general activation functions.

5. **Softmax Function :** The softmax function converts real-valued numbers so that their sum equals 1 in a given vector. The softmax converts the input values, which can be negative, positive, zero, or larger than one, to values between(0 and 1), permitting them to be interpreted as probabilities. The softmax converts a small or negative input into a small probability, and a large input into a large probability, but it still stays between 0 and 1. Because several multilayer neural networks end in a penultimate layer that outputs realvalued scores that are not easily scaled and can be difficult to work with, it is usually applied to the performance of the very last layer. The softmax is particularly useful in this situation because it converts the scores to a

normalized probability distribution that can be presented to users or used as feedback to other systems. Consequently, it's common to append a softmax feature as the neural network the last layer, which can be formulated as:

$$\sigma(z)_i = \frac{e^{z_i}}{\sum_{j=0}^{N} e^{z_i}}$$

N= classes.

z = input vector.

$\sigma(z)_i$= class probability of that product.

6. **Stride :** It is a filter parameter in an NN that controls the number of movements through an image or video. This filter can shift one pixel (or unit) on a time if the stride =1. Since the size of the filter holds an effect upon encoded output length, the stride is often fixed to a whole number rather than a fraction or decimal.
7. **Padding:** itis a concept used in convolutional neural networks to define the number of pixels that added to an image as it processed by the CNN kernel. If the padding is set to zero, for example, all pixel values added will be zero. If the padding (zero) option is set to one, a one-pixel border with a pixel value of zero will be added to the image. Padding functions by enlarging the processing region of the convolutional neural network. The kernel is a neural network filter that scans each pixel in an image and transmits the data into a smaller or larger format. Padding is implemented to the image frame to aid the kernel is processing the image by providing more room for the kernel to cover the image. Adding padding to a CNN-processed image allows for more accurate image analysis.

8. **Flatten :** Takes the previous layers' output and "flattens" it into a single vector that can be used as an input for the next phase.

The classification steps of the proposed model are illustrating in algorithm (3.1) as follows:

Algorithm (3.1): Implement the proposed 1-D CNN classification model's main steps
Input: feature vector for face image, Epochs, learning rats. **Output:** 1-D Model
BEGIN 1. I $_{train}$ ←Train (1-D hybrid deep learning model, training $_{set}$) 1.1 Initialize the weights at random values. 1.2 Take the input from training data and pass it through the proposed model layers figure (3.3). 1.3 Calculate the output's error according the equation $\boldsymbol{Error} = \mathbf{1/2}\ (\boldsymbol{actual} - \boldsymbol{predicted})^2$ 1.4 Update the weights according to Adam optimizer. 1.5 Perform steps 1.2-1.4 to overall training set for each epoch 1.6 Repeat steps 1.2-1.5 until the error decreased to acceptable tolerance. 2. Finish training when reached acceptable performance. 3. Saved proposed model and weights for predications. **END**

The proposed hybrid deep learning model (that we called DMFace-net) will be described in detail and illustrated in figure (3.3), which will display the (12) layers that make up the model:

1) **Four** convolutional layers (conv 1D) for Feature Extraction.
2) **Four** MaxPooling 1D layers.
3) **Two** Features Collecting layers represented by the (Dense).
4) **One** Flattens layer.
5) **One** Fully Connected layer represented by the (Dense).

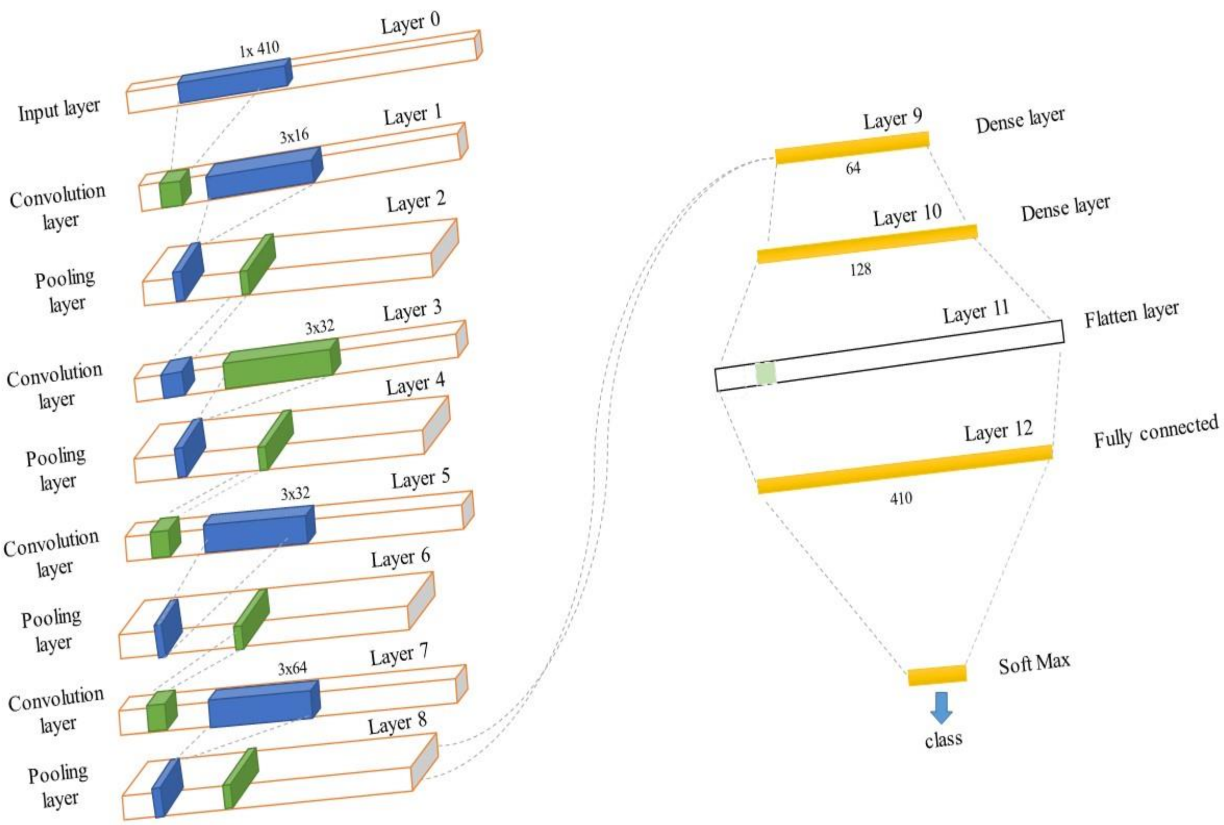


Figure (3.3): Network Architecture for the proposed model (DMFace-net)

The whole system stages of the proposed system are explaining in Algorithm 3.2 as follows:

Algorithm (3.2): The overall proposed methodology algorithm
Input: Face image **Output:** Decision of identification
BEGIN **Step1:** preparing data (face dataset). - splitting Face image Dataset into : Training Dataset (training $_{set}$), Test Dataset (testing $_{set}$) - Return (training $_{set}$, testing $_{set}$). **Step2:** Call pre-processing algorithm. - Enhance face image (call algorithm 2.1). - Resize the face image (call algorithm 2.2). **Step3:** Call Feature extraction algorithm. - Call algorithm (2.3). **Step4:** Construct proposed model 1-D using (training $_{set}$) - call algorithm (3.1) **Step5 :** Evaluate proposed model and saved it and weights for Predication. **Step6 :** Select query face image from the testing $_{set}$ (unseen data) - Perform step 2, step3 in query face image. - Use saved proposed model for predictation. **Return** decision // (" class number " and " matched image face ") **END**

The proposed hybrid deep learning model (DMFace-net) for facial image recognition are listed in the table (3.1), which includes extensive details on each layer of the classification algorithm.

Layer	Type	Filters	Stride	Kernel size	Activation function	Parameters
1	Conv.1D	16	1	3	RelU	64
2	Maxpooling1D	-	1	1	-	-
3	Conv.1D	32	1	3	RelU	1568
4	Maxpooling1D	-	1	1	-	-
5	Conv.1D	32	1	3	RelU	3104
6	Maxpooling1D	-	1	1	-	-
7	Conv.1D	64	1	3	RelU	6208
8	Maxpooling1D	-	1	1	-	-
9	Dense	64	-	-	RelU	4160
10	Dense	128	-	-	RelU	8320
11	Flatten	-	-	-	-	-
12	Fully connected	-	-	410	Softmax	21044890

Table (3-1): The details of the proposed model (DMFace-net)

In addition to the previous table that describes the details, the information about the attained hybrid deep learning: the total number of the parameters is 21,068,314, the batch size is equal to 64, number of the epoch is 100 each epoch includes 32 steps, Categorical Cross-Entropy loss function, and also adopts the Adam optimization algorithm with a learning rate of 0.0001.

Chapter Four

The Experimental Results and Evaluation

Chapter Four

The Experimental Results and Evaluation

4.1 Introduction

The experimental results obtained from the proposed system procedures implementation on facial images are presented and discussed in this chapter. This chapter will include the algorithms and procedures shown in chapter three that are focused on facial image implementation data. First and foremost, an overview of the environment in which the system is implemented is presented, including all of the programming languages and platforms used. Next, the findings of each step of the proposed face recognition system will be presented and addressed to demonstrate the system's effectiveness.

4.2 Description of the Proposed System Implementation Environment

The implementation environment description is significant in evaluating the proposed system behaviour and performance and the way it works. The implementation of the proposed model will be discussed in this section, along with its maps. The following programming languages were used to implement the proposed system function:

- Python the version 3.6 under IDLE.
- New java 8 under the NETBEANS IDE the version 8.2.
- Everything in proposed system procedures are performed on a LENOVO laptop with Intel(R) Core(TM) i7-7700HQ with 16 GB RAM size and the screen card is NVIDIA GEFORCE GTX 6 GB, 250 GB SSD, and 64 bit Operating System is windows 10.

4.3 Facial Image Dataset

This section will described the Disguise and makeup database (DMFD) that was used in this study for personal identification and verification dependent on his or her face. The Dataset from (Hong Kong Polytechnic University) coming with cropped face images that created in 2016 [67][68]. The dataset is composed of (2460) images of (410) different persons and each one has a different number of sample images (between two at least and six images at most) . The majority of the images are from celebrities (are movie/TV stars, politicians, or athletes). For all of the subjects, the first picture is a frontal shot with no disguise or minimal makeup. There's at least one to two clean (no disguises or makeups, pure face) face images in each subject, and the rest of the images have different disguises like as glasses, beards, hairstyles, and other facial accessories. The images that acquisition on real environments as shown in figure 4-1. The main challenge in DMFD dataset is matching face images captured in a real world environment with many covariates such as pose, occlusion, distance and illumination. As a result, using this database to test the efficiency of face recognition algorithms for close-range face images with disguise and/or make-up as a primary covariate is extremely difficult [68].

Figure (4.1): illustrates an example of modifications in the appearance of a person by using different faces disguises and makeup

4.4 Implementation results of the proposed system phases

A detailed explanation of the recognition mechanism based on a person's facial images is outline in this section of the thesis. A number of measures will be described and implemented, as well as a detailed overview of each step of the system's operation, including the pre-processing phase, feature, and classification using four machine-learning algorithms, pre-traind model (AlexNet), and proposed a hybrid deep learning model (that we called

DMFace-net). The face images that inserted into the proposed system shown in Figure 4.2.

Figure (4.2): DMFD Facial Images Examples

4.4.1 Results of the Pre-processing Phase of Facial Images

The proposed recognition system's first step is the pre-processing of facial images. Pre-processing is significant, since the images will be processed in several stages, each of which will result in a different image and have a different effect on the image. Pre-processing is a set of steps that gives us a precise result, a perfect facial image to work with in feature extraction and classification. The pre-processing step divided into two parts (Histogram Equalization and Resizing). Because of it's the simple function and efficacy, histogram equalization commonly used for contrast enhancement in a number of applications. In this thesis, the histogram equalization applied to the facial images to increase contrast and make the image information more visible, which will aid in the feature extraction process. After applying histogram equalization, the resulted facial images shown in Figure (4.3).

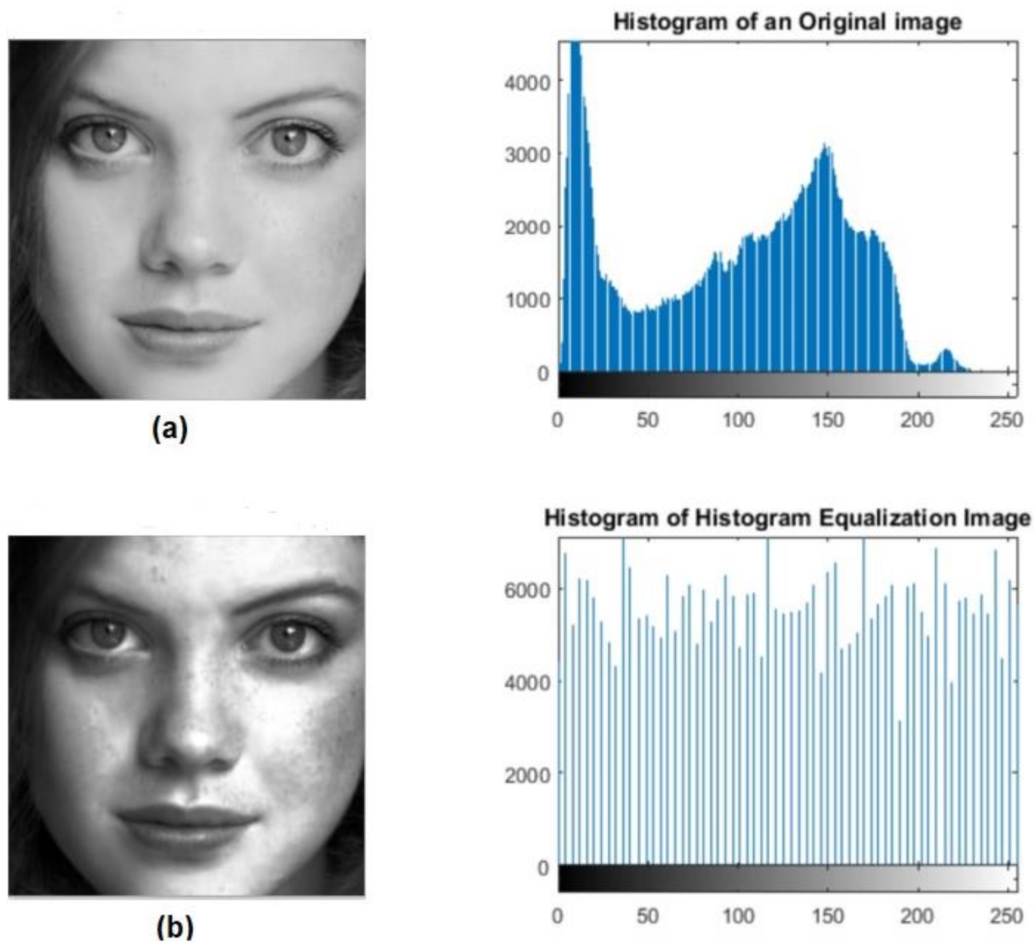


Figure (4.3): Example of an image (a) before/ (b) after applying histogram equalization with its histogram

4.4.2 Results of the Feature Extraction Process

The next step in this thesis is feature extraction. The resulting facial image after pre-processing is the input in this thesis. The extraction of features is done using LDA. The following is a description of the results obtained by using the feature extraction algorithm, as shown in figures (4.4).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	class	a1	a2	a3	a4	a5	a6	a7	a8	a9	a10	a11	a12	a13	a14	a15	a16	a17
2	1	47.72801	-1.91342	13.36512	5.142873	27.89963	-21.1517	-3.88738	10.47881	0.109151	-12.4797	-15.779	0.044998	-3.48836	8.067162	-22.9536	-6.31569	31.59115
3	1	17.93974	19.5278	-14.9186	9.090442	-0.16322	-6.84462	-3.01184	-17.5126	-2.00652	-11.8013	18.83947	0.254438	-4.27913	-1.22905	-5.67475	-5.04405	17.75323
4	1	24.26641	-5.34514	-1.32068	16.94898	2.607776	13.51304	-7.17605	11.12037	-15.477	-9.92993	13.48626	-0.81771	-12.6981	-0.77704	-30.3154	-0.64838	14.06256
5	1	52.31266	12.81955	6.568484	10.97203	12.1631	-10.2227	-7.42522	8.806529	-1.29538	-13.7993	9.600871	11.73126	-9.72498	-3.54512	-14.9883	-14.0675	18.7439
6	1	8.584141	4.314341	9.575032	12.17246	25.74436	-22.8092	-10.9925	-4.69677	-19.9932	-23.2921	4.563958	4.224476	-12.2817	0.239979	-20.9203	-28.4352	13.81598
7	10	31.61632	35.05082	-37.2487	-15.7351	-43.8034	26.59349	15.48438	-20.6244	-17.1329	-11.3308	21.03093	-9.47798	-9.37009	-17.5064	-25.7577	-6.64653	-5.13501
8	10	15.3806	24.96516	-30.9756	-12.8097	-12.1545	-0.51098	14.08664	-14.6243	-12.6608	-15.1427	8.55651	-16.4728	-11.5601	-19.079	-19.9322	-13.743	-10.1293
9	10	23.07711	40.38358	-29.1525	-20.4425	-42.7166	13.66163	15.85491	-22.2713	-10.2642	-17.1718	14.07758	-9.51882	-3.7735	-29.8471	-23.242	-7.90967	-6.71234
10	10	18.29653	23.94065	-19.6476	-27.5354	-29.367	18.5075	9.95634	-3.30875	-11.7926	6.379815	10.95092	6.240865	15.22452	-9.62527	-2.0156	-11.2792	-0.8068
11	10	21.29346	30.54261	-17.4155	-14.3027	-33.7474	18.27087	9.812014	-18.8508	-11.4066	-1.13428	12.55848	-6.34568	-10.5456	-16.2016	-12.0645	-10.0975	-10.3154
12	10	10.53504	41.68284	-31.8853	1.228399	-20.5499	10.28603	-3.61569	-3.35333	-30.6122	-6.29855	12.31269	-5.69454	-5.15928	-16.0247	-17.8909	8.470718	4.78262
13	100	-31.8562	-1.78533	-14.0473	4.505598	2.629806	11.94542	-34.6147	16.82943	-5.00009	-13.6353	15.18488	0.852988	-8.63568	-22.8839	-2.24848	-16.4385	30.30308
14	100	-13.5596	-1.73825	-13.3044	1.520631	-11.0804	2.119027	-41.226	36.90455	-2.24315	6.856938	-4.06002	-1.74823	-11.8555	-15.7609	19.92404	-27.8072	30.3636
15	100	-48.1035	-3.40731	-15.3955	-5.16811	2.308913	-8.46127	-33.2884	22.87776	-15.5594	8.462786	21.29586	7.732565	-16.599	-12.976	-1.53446	-16.1999	14.8049
16	100	-37.189	-6.13786	-5.00406	11.22442	-5.85088	8.033874	-24.5507	40.00346	-9.49462	6.980918	24.12861	1.711231	-8.84807	-17.2962	18.29873	-16.1888	27.68022
17	100	-23.8601	-11.8769	-26.8171	20.59103	-19.2952	10.05464	-38.9733	27.30969	-22.272	11.70709	12.625	5.00378	-18.7744	-12.9781	20.19604	-17.2306	20.27409
18	100	-31.9729	-9.52884	-6.24022	16.58918	-8.40253	10.77933	-22.7984	25.96532	-8.50987	3.802959	18.77521	-8.50289	-20.4889	-18.2869	1.787409	-25.4499	30.99789
19	101	41.97269	18.66869	13.14869	-1.58567	28.20931	-18.8556	2.037203	-5.55253	-4.97607	-11.6835	7.309503	23.62776	-5.55492	1.826187	23.58441	5.835926	-15.3825
20	101	24.70751	13.29806	1.900622	1.428612	0.958946	3.48025	5.094098	-9.82329	-12.6034	-5.38339	0.409946	11.55592	-4.51357	9.004058	20.74211	10.81035	-12.1421
21	101	22.97548	18.9958	12.96427	-0.80542	19.098	-26.0016	8.177434	-16.1432	-23.8766	-5.284	11.84305	-8.58795	-11.3122	-8.03988	20.04144	22.5544	-12.7615
22	101	26.40175	12.29846	19.18651	3.062258	18.58537	-5.18247	-8.07618	-16.5358	-7.50362	-14.0781	11.92138	1.827734	-1.77455	8.383882	10.30906	9.593164	-2.98627
23	101	-1.25923	19.31806	11.0344	11.45236	30.62946	-16.6071	1.322657	-13.7716	-5.05076	-4.31783	6.351523	-6.43049	1.300504	17.08339	11.07319	-3.558	4.871367
24	101	36.05875	22.65909	7.255982	7.041841	7.817606	-20.3565	8.742051	-6.8743	-15.4194	-6.52855	13.35791	4.595732	-7.08361	-4.38038	14.57146	10.53026	-1.40964
25	102	22.45333	-12.2491	-17.4231	-21.0057	28.20192	-2.26556	35.16746	-20.3125	1.702013	-32.2512	13.96821	4.574145	-7.50413	-5.97046	-18.3445	3.91907	-7.30744
26	102	3.052701	-16.9961	-15.955	-8.20617	25.43625	19.94678	42.06065	-20.8207	6.035778	-21.6499	-1.75962	-2.52267	-6.0983	0.500329	-14.7915	5.298754	-1.867
27	102	13.77956	-7.92372	-5.47338	-14.5036	19.17854	3.797475	43.87987	-26.5917	1.261847	-18.8748	-8.31974	1.603373	-1.04728	0.91639	-8.2693	19.71183	1.595841

Figure (4.4): Samples of resulted features extracted using LDA

4.4.3 Results and measurements from the classification Based on NB, KNN, RF and SVM

The classification is a key feature of the system presented in this thesis, and four machine learning algorithms were used to train and evaluate the features, with the results compared to those obtained using the hybrid deep learning model (DMFace-net). The results of applying NB, KNN, RF and SVM machine learning algorithms for classifying the features obtained from the previously mentioned feature extraction algorithms, described in this section of the thesis. Table (4.1) will show the result of implemented NB, KNN, RF and SVM with LDA.

Table (4.1): The experimental results of implemented the **four machine learning algorithms** with **LDA**.

Methods	**Precision**	**Recall**	**F-measure**
Naïve Bayes + LDA	74.24	66.52	65.69
KNN + LDA	80.97	51.17	59.59
RF + LDA	85.97	84.69	84.17
SVM + LDA	97.95	96.56	96.84

4.4.4 Results and measurements from the classification Based on pre-trained model AlexNet and proposed hybrid deep learning

In comparison to traditional approaches that existed before AlexNet, this architecture that created in 2012 that developed to Neural Network as deep as “CNN”, is the first deep network that successfully classified specific objects in the (ImageNet) dataset. There are five (conv) layers, three fully connected layers and they use Relu activation in each of these layers except the output layer in this model. They also used the dropout layers.

When using the LDA with the pre-trained model AlexNet and the proposed hybrid deep learning model the following general results obtained, as shown in table (4.2).

Table (4.2): The experimental results of implemented the AlexNet and **proposed hybrid deep learning model** with LDA.			
Methods	**Precision %**	**Recall %**	**F-measure %**
LDA+ AlexNet	70	65	63
LDA+ DMFace-net	**99**	**98**	**98**

4.5 Comparison Results Charts

In this part, three comparisons have been made among main performance Metrics precision, recall and F-measure, which have been used in the system.

4.5.1 Precision Measurements

It determines the percentage of how accurate the classifier is in its correct predictions.

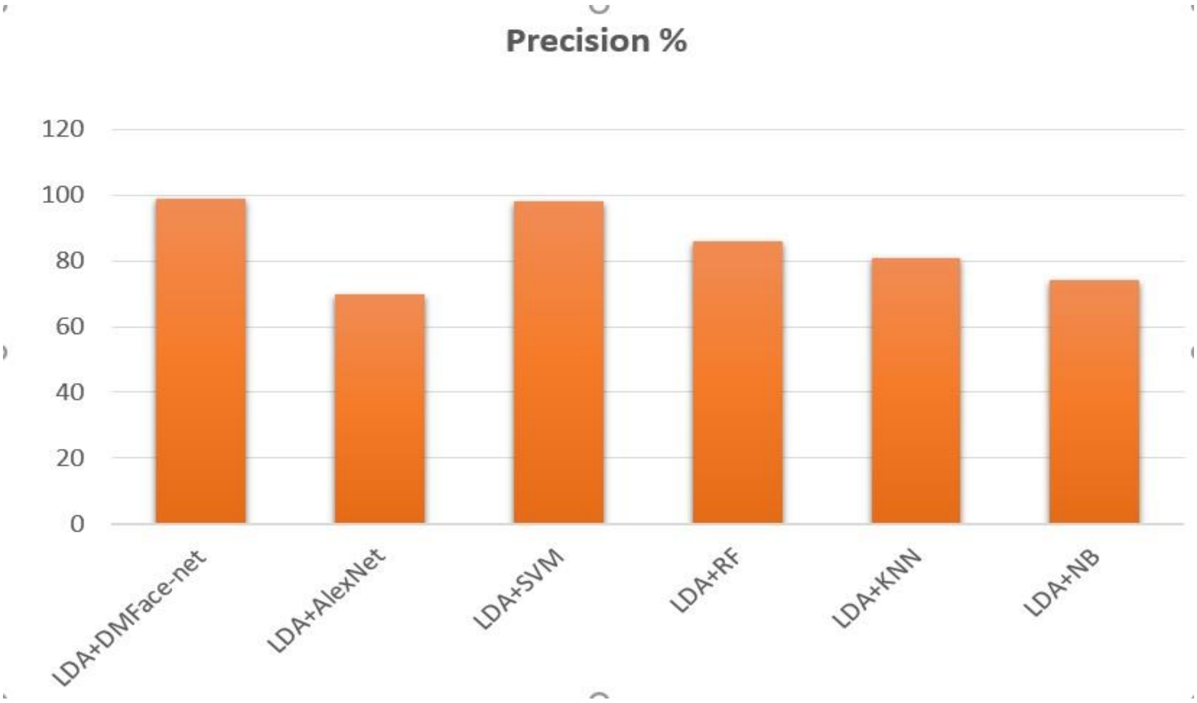


Figure (4.5): Precision of implemented NB, KNN, RF, SVM, AlexNet and Proposed DMFace-net with LDA

This charts show that the best result for precision obtained when applying the LDA with the proposed algorithm is 99 %.

4.5.2 Recall Measurement

It calculates the data that we expected to be correct from among all the data that we expected to be correct.

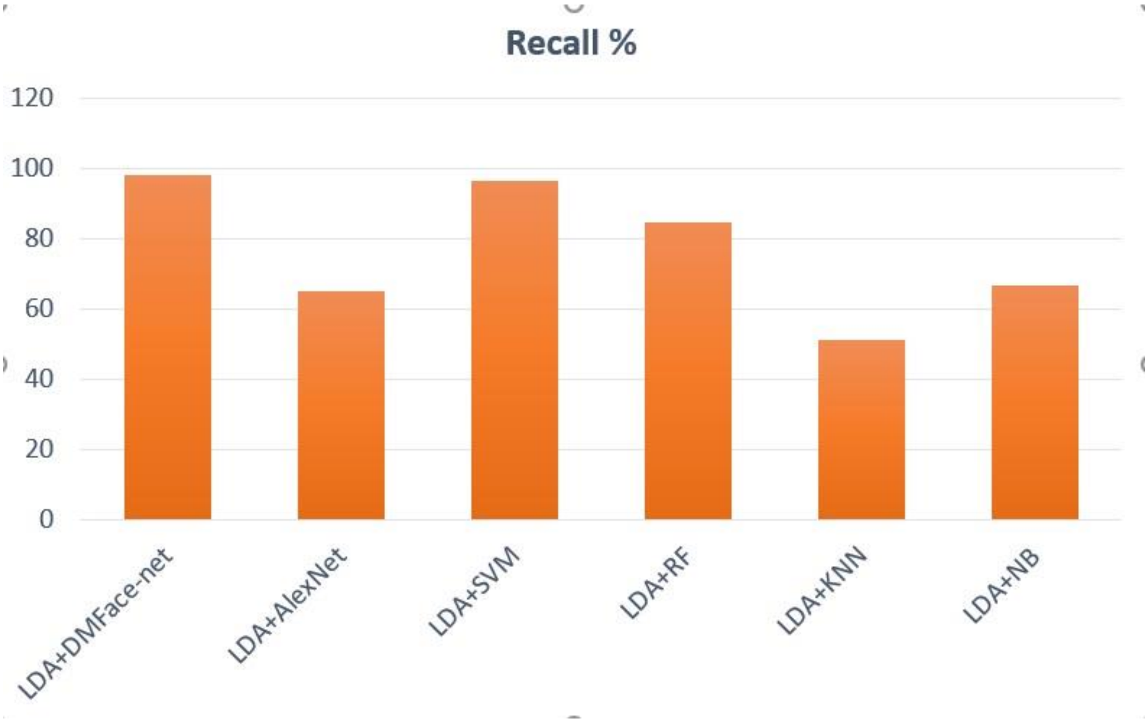


Figure (4.6): Recall of implemented NB, KNN, RF, SVM, AlexNet and Proposed DMFace-net with LDA

This charts show that the best result for recall obtained when applying the LDA with the proposed algorithm is 98 %.

4.5.3 F-measure Measurement

It is the full result of the efficiency and accuracy of the model.

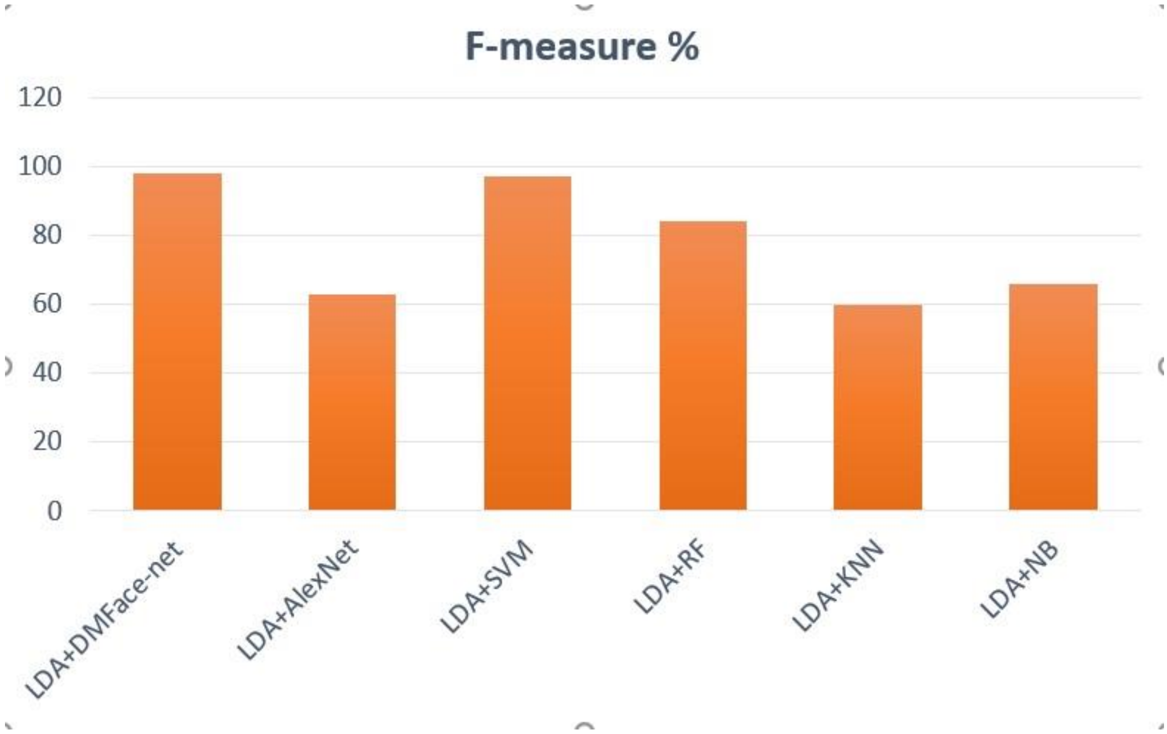
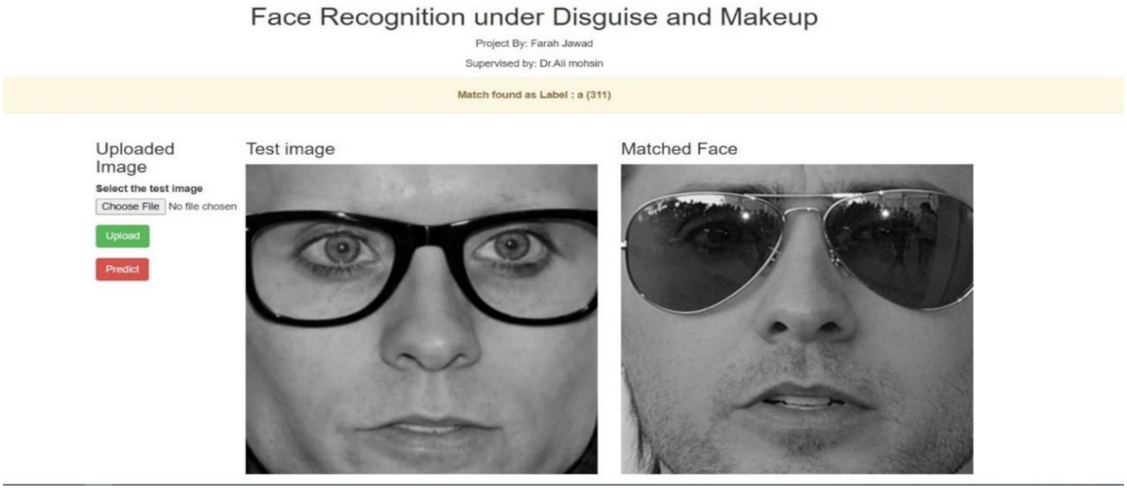


Figure (4.7): F-Measure of implemented NB, KNN, RF, SVM, AlexNet and Proposed DMFace-net with LDA

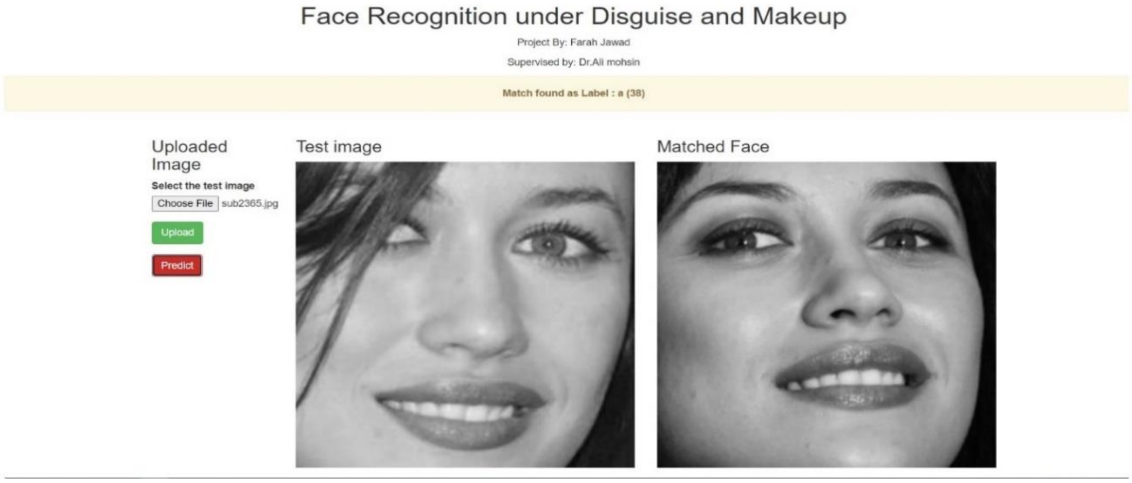
This charts show that the best result for F-Measure obtained when applying the LDA with the proposed algorithm is 98 %.

4.6 Face recognition System Implementation

A graphical interface is designed to display the resulted images and class number to which the entered image belongs obtained after applying each procedure of the proposed recognition system as shown in figure (4.8), which illustrate the proposed facial recognition system's main interface."Choose File" to select the test image. "Upload" button to load test image and show it. "Predict" button to running the proposed system.



(a)



(b)

(c)

Figure (4.8): (a,b and c) the main interface of proposed recognition system

4.7 The results of this thesis discussed as follows:

Preparation of dataset is one of the prime aspects for training and testing the deep learning algorithm to give the desired results. The facial images have passed through several steps in pre-processing phase in order to enter the hybrid deep-learning classification algorithm layers and make it gives an accurate result in the recognition procedure. Some aspects related to the classification are important, must illustrate, and discussed as follow:

1- **Training time:** the time that required training the deep learning algorithm is very important to give us a look at the accuracy obtained from the classification algorithm. In this proposed recognition system, the hybrid deep learning recommended that to run for 100 epochs. The training time using the attained feature extraction algorithms with the hybrid deep learning in which it is 15 minutes.

2- **Accuracy:** classification accuracy is an important aspect of the proposed facial image recognition system in this thesis. The use of hybrid deep-learning conventional neural network algorithms gives 99% precision.

3- **Testing time:** the time require to test the proposed system for recognizing human, testing time for hybrid-deep learning with LDA feature extraction methods it only takes a few seconds with accuracy of prediction.

4- **Comparison with previous studies:** to compare with the previous works on (DMFD) dataset, the research that studied and present this dataset is Wang and Kumar [68]. LBP was utilized in conjunction with biometric and non-biometric blocks and two commercial matchers (VeriFace, Face++). To improve accuracies, just the biometric blocks are used for matching. The given experimental results suggest that the examined matchers' face recognition capability for recognizing disguised and makeup faces are rather weak. After this study, this data was not used in researchs for the purpose of facial recognition until the year 2021, when this database was used for the purpose of identifying disguised faces using transfer learning techniques and pre-trained models, and the accuracy ranged between (78% - 99%) [69]. Table (4.3) show the comparison details:

Table (4.3): The comparison details with previous work

No.	Ref	Using DMFD dataset	The method	Results
1	[69]	• Using all color image in data base ((410) different persons and each one has a six sample images). • Using balanced dataset. • The train-test ratio: 80% images are used for training and the remaining 20% images are used for testing.	• Using 2D transfer learning Model incorporating simple noise-based data augmentation. • The proposed method detects face in an image using Viola Jones face detector and classify it using a pre-trained Convolutional Neural Network (CNN) ne-tuned for DIFR. During transfer learning, a pre-trained CNN learns. generalized disguise-invariant features from facial images of several subjects to correctly identify them under varying facial disguises. • Use models (Resnet-18, Resnet-50, Squeezenet, Inception-v3).	99% 97% 94% 78%

2	This thesis	• Using cropped grayscale image in data set ((410) different persons and each one has a different number of sample images (between two at least and six images at most). • Using imbalanced dataset. • The train-test ratio: 70% images are used for training and the remaining 30% images are used for testing.	• Using 1D hyper-deep learning Model • The proposed method, The features of facial images are extracted using Linear Discriminant Analysis (LDA) method that gives the minimum feature length and minimum extraction time. Classification model is done by using the proposed hybrid-deep learning Classifier for more precise feature learning.	99%

.

Chapter Five

Conclusions and Future work Suggestions

Chapter Five

Conclusions and Future work Suggestions

5.1 Conclusions

Through the implementation period of the proposed work, a number of conclusions have been achieved depending on the actual and practical results were the following statements summarize the most important ones:

1. The efficiency of the recognition based on the preprocessing stage in order to dispose of the defects associated with the process of image acquisition.

2. The enhanced images will have the best contrast than the original image since more embedded colors may be found in the original image are appeared after the process of enhancement. Sharp edges and clear extended regions are very suitable to describe the face in terms of the numerical feature.

3. Also, the features vector (the feature space is not complex) give more speed to make decision and classification. This is what has been relied on in this research, where selecting good features based on the LDA algorithm led to the best results.

4. Extraction of features from facial images with great accuracy and precision is critical to achieving the desired results. in this thesis. On face images, the LDA algorithm is used for feature extraction.

5- The proposed hybrid-deep learning classifier with the LDA achieved a 99% precision score. The LDA produces the smallest number of stable features under a variety of working conditions, such as lighting and contrast variance.

6- The training time using the attained feature extraction algorithm with proposed hybrid deep learning when using LDA is 15 minutes.

7- When implemented proposed system, the class number is recognized for the person whose image has been entered into the system, with the corresponding image displayed.

5.2 Future Work Suggestions

Several recommendations would be made and may be implemented, and some of them are presented in this section:

1. The proposed system could be used another type of facial images with lower qualities and more than different environment conditions.
2. The proposed system could be used the proposed system to analyse three-dimensional facial images in order to identify individuals.
3. The proposed system could be used in gender identification to distinguish between male and female people.
4. The proposed system could be used facial images in conjunction with another form of biometric to identify an individual, such as iris, eye, fingerprint, and so on.
5. The proposed system could be created a mobile application that can link to the cloud using the proposed identification system.
6. The proposed system could be used the proposed system in the real-time face recognition system will deal with features derived from a series of frames (i.e., for video).
7. Using a variety of image datasets and the same techniques to compare the results..

References

[1] Abdelwhab, Abdelgader, and Serestina Viriri, 2018, "A survey on soft biometrics for human identification." Machine Learning and Biometrics", p (37).

[2] Sabdar Aman, India, Available: URL: http://sabdarac.blogspot.com/2018/01.

[3] Dargan, Shaveta, and Munish Kumar. "A comprehensive survey on the biometric recognition systems based on physiological and behavioral modalities." Expert Systems with Applications 143 (2020): 113114.

[4] Shaheed, Kashif, Aihua Mao, Imran Qureshi, Munish Kumar, Qaisar Abbas, Inam Ullah, and Xingming Zhang. "A Systematic Review on Physiological-Based Biometric Recognition Systems: Current and Future Trends." Archives of Computational Methods in Engineering (2021): 1-44.

[5] Rai, Vinayak, Kapil Mehta, Jatin Jatin, Dheeraj Tiwari, and Rohit Chaurasia. "Automated Biometric Personal Identification-Techniques and Applications." In 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), pp. 1023-1030. IEEE, 2020.

[6] Jain, Anil K., and Stan Z. Li. Handbook of face recognition. Vol. 1. New York: springer, 2011.

[7] Shetagar, Poornima, Sehraantaj Sadarbhai, Vaibhavi Kulkarni, and Vijayalaxmi Patil. "Survey on Feature Extraction Methods." International Research Journal of Engineering and Technology (IRJET), p-ISSN: 2395-0072, vol.08 (2021).

[8] Xu, Jingxiao, Gongping Yang, Kuikui Wang, Yuwen Huang, Haiying Liu, and Yilong Yin. "Structural sparse

representation with class-specific dictionary for ECG biometric recognition." Pattern Recognition Letters 135 (2020): 44-49.

[9] Gh, Mohammad Basman. "A novel Face Recognition System based on Jetson Nano developer kit." In IOP Conference Series: Materials Science and Engineering, vol. 928, no. 3, p. 032051. IOP Publishing, 2020.

[10] Haji, Suad, and Asaf Varol. "Real time face recognition system (RTFRS)." In 2016 4th International Symposium on Digital Forensic and Security (ISDFS), pp. 107-111. IEEE, 2016.

[11] Sameem, M. Shujah Islam, Tehreem Qasim, and Khush Bakhat. "Real time recognition of human faces." In 2016 International Conference on Open Source Systems & Technologies (ICOSST), pp. 62-65. IEEE, 2016.

[12] Abhila, A. G., and S. H. Sreeletha. "A deep learning method for identifying disguised faces using AlexNet and multiclass SVM." Int Res J Eng Technol 5, no. 07 (2018).

[13] Sarapakdi, Sittiphan, Phaderm Nangsue, and Charnchai Pluempitiwirivawej. "Occluded facial recognition with 2DPCA based convolutional neural network." In 2019 IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia), pp. 135-138. IEEE, 2019.

[14] Christou, Nikolaos, and Nilam Kanojiya. "Human facial expression recognition with convolution neural networks." In Third International Congress on Information and Communication Technology, pp. 539-545. Springer, Singapore, 2019.

[15] Patil, Vidya, Anushka Narayan, Vaishnavi Ausekar, and Anahita Dinesh. "Automatic Students Attendance Marking System Using Image Processing And Machine Learning." In 2020 International Conference on Smart Electronics and Communication (ICOSEC), pp. 542-546. IEEE, 2020.

[16] Pranav, K. B., and J. Manikandan. "Design and Evaluation of a Real-Time Face Recognition System using Convolutional Neural Networks." Procedia Computer Science 171 (2020): 1651-1659.

[17] Hung, Bui Thanh. "Face recognition using hybrid HOG-CNN approach." In Research in Intelligent and Computing in Engineering, pp. 715-723. Springer, Singapore, 2021.

[18] Plataniotis, Konstantinos N., and Anastasios N. Venetsanopoulos. Color image processing and applications. Springer Science & Business Media, 2013.

[19] Ruslau, M. F. V., R. A. Pratama, and E. Meirista. "Edge detection of digital image with different edge types." In Journal of Physics: Conference Series, vol. 1569, no. 4, p. 042069. IOP Publishing, 2020.

[20] Gonzalez, Rafael C., and Richard E. Woods. "Digital image processing 4th edition, global edition." (2018): 133-153.

[21] Kortli, Yassin, Maher Jridi, Ayman Al Falou, and Mohamed Atri. "Face recognition systems: A survey." Sensors 20, no. 2 (2020): 342.

[22] Guo, Guodong, and Na Zhang. "A survey on deep learning based face recognition." Computer vision and image understanding 189 (2019): 102805.

[23] Dewantara, Bima Sena Bayu, Mochamad Mobed Bachtiar, and Syah Embo Lantang. "Door Access Control based on Illumination Invariant Face Recognition in Embedded System." 2020 10th Electrical Power, Electronics, Communications, Controls and Informatics Seminar (EECCIS). IEEE, 2020.

[24] He, Mingjie, Jie Zhang, Shiguang Shan, Meina Kan, and Xilin Chen. "Deformable face net for pose invariant face recognition." Pattern Recognition 100 (2020): 107113.

[25] Jayaraman, Umarani, Phalguni Gupta, Sandesh Gupta, Geetika Arora, and Kamlesh Tiwari. "Recent development in face recognition." Neurocomputing 408 (2020): 231-245.

[26] Farzaneh, Amir Hossein, and Xiaojun Qi. "Facial expression recognition in the wild via deep attentive center loss." Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. 2021.

[27] Huang, Baojin, et al. "When Face Recognition Meets Occlusion: A New Benchmark." ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2021.

[28] Anwarul, Shahina, and Susheela Dahiya. "A comprehensive review on face recognition methods and factors affecting facial recognition accuracy." Proceedings of ICRIC 2019 (2020): 495-514.

[29] Jahne, B. "Digital Image Processing. 6th revised and extended edn, Vol. 583." SpringerVerlag, Berlin Heidelberg (2005).

[30] Azad, Mir Mohammad, and Md Mahedi Hasan. "Color image processing in digital image." International Journal of New Technology and Research 3.3 (2017).

[31] Dan, Danhui, and Qiang Dan. "Automatic recognition of surface cracks in bridges based on 2D-APES and mobile machine vision." Measurement 168 (2021): 108429.

[32] Seema, Gaurav Bansal. "A Review of image contrast enhancement techniques." International Research Journal of Engineering and Technology (IRJET), e-ISSN (2017): 2395-0056.

[33] Yoo, Sung-Hoon, Sung-Kwun Oh, and Witold Pedrycz. "Optimized face recognition algorithm using radial basis function neural networks and its practical applications." Neural Networks 69 (2015): 111-125.

[34] Ali, Hussam, M. I. Lali, Muhammad Zohaib Nawaz, Muhammad Sharif, and B. A. Saleem. "Symptom based automated detection of citrus diseases using color histogram and textural descriptors." Computers and Electronics in agriculture 138 (2017): 92-104.

[35] Pankaj S. Parsania, Paresh V. Virparia, “A Review: Image Interpolation Techniques for Image Scaling,” Int. Journal of Innovative Research in Computer and Communication Engineering, Vol. 2, No. 12, PP. 7409-7414, 2014.

[36] Salau, Ayodeji Olalekan, and Shruti Jain. "Feature Extraction: A Survey of the Types, Techniques, Applications." In 2019 International Conference on Signal Processing and Communication (ICSC), pp. 158-164. IEEE, 2019.

[37] Oloyede, Muhtahir O., Gerhard P. Hancke, and Hermanus C. Myburgh. "A review on face recognition systems: recent approaches and challenges." Multimedia Tools and Applications 79, no. 37 (2020): 27891-27922.

[38] Kumar, Gaurav, and Pradeep Kumar Bhatia. "A detailed review of feature extraction in image processing systems." 2014 Fourth international conference on advanced computing & communication technologies. IEEE, 2014.

[39] Singh, Aruni, Sanjay Kumar Singh, and Shrikant Tiwari. "Comparison of face recognition algorithms on dummy faces." The International Journal of Multimedia & Its Applications 4, no. 4 (2012): 121.

[40] Ghassabeh, Youness Aliyari, Frank Rudzicz, and Hamid Abrishami Moghaddam. "Fast incremental LDA feature extraction." Pattern Recognition 48, no. 6 (2015): 1999-2012.

[41] Li, Chun-Na, Yuan-Hai Shao, Wei-Jie Chen, Zhen Wang, and Nai-Yang Deng. "Generalized two-dimensional linear discriminant analysis with regularization." Neural Networks 142 (2021): 73-91.

[42] Lahaw, Zied Bannour, Dhekra Essaidani, and Hassene Seddik. "Robust Face Recognition Approaches Using PCA, ICA, LDA Based on DWT, and SVM Algorithms." In 2018 41st International Conference on Telecommunications and Signal Processing (TSP), pp. 1-5. IEEE, 2018.

[43] Bhavitha, B. K., Anisha P. Rodrigues, and Niranjan N. Chiplunkar. "Comparative study of machine learning techniques in sentimental analysis." In 2017 International

conference on inventive communication and computational technologies (ICICCT), pp. 216-221. IEEE, 2017.

[44] Dhall, Devanshi, Ravinder Kaur, and Mamta Juneja. "Machine learning: a review of the algorithms and its applications." Proceedings of ICRIC 2019 (2020): 47-63.

[45] Meng, Tong, Xuyang Jing, Zheng Yan, and Witold Pedrycz. "A survey on machine learning for data fusion." Information Fusion 57 (2020): 115-129.

[46] Goh, Guo Dong, Swee Leong Sing, and Wai Yee Yeong. "A review on machine learning in 3D printing: applications, potential, and challenges." Artificial Intelligence Review 54, no. 1 (2021): 63-94.

[47] Ayoub, Mohammed. "A review on machine learning algorithms to predict daylighting inside buildings." Solar Energy 202 (2020): 249-275.

[48] Sen, Pratap Chandra, Mahimarnab Hajra, and Mitadru Ghosh. "Supervised classification algorithms in machine learning: A survey and review." In Emerging technology in modelling and graphics, pp. 99-111. Springer, Singapore, 2020.

[49] Alloghani, Mohamed, Dhiya Al-Jumeily, Jamila Mustafina, Abir Hussain, and Ahmed J. Aljaaf. "A systematic review on supervised and unsupervised machine learning algorithms for data science." Supervised and unsupervised learning for data science (2020): 3-21.

[50] Aziz, Sumair, Syed Zohaib Hassan Naqvi, Muhammad Umar Khan, and Taimoor Aslam. "Electricity theft detection using empirical mode decomposition and K-

Nearest neighbors." In 2020 International Conference on Emerging Trends in Smart Technologies (ICETST), pp. 1-5. IEEE, 2020.

[51] Kabir, Monika, Mir Md Jahangir Kabir, Shuxiang Xu, and Bodrunnessa Badhon. "An empirical research on sentiment analysis using machine learning approaches." International Journal of Computers and Applications (2019): 1-9.

[52] Shah, Kanish, Henil Patel, Devanshi Sanghvi, and Manan Shah. "A comparative analysis of logistic regression, random forest and KNN models for the text classification." Augmented Human Research 5, no. 1 (2020): 1-16.

[53] Kirasich, Kaitlin, Trace Smith, and Bivin Sadler. "Random forest vs logistic regression: binary classification for heterogeneous datasets." SMU Data Science Review 1, no. 3 (2018): 9

[54] Cervantes, Jair, Farid Garcia-Lamont, Lisbeth Rodríguez-Mazahua, and Asdrubal Lopez. "A comprehensive survey on support vector machine classification: Applications, challenges and trends." Neurocomputing 408 (2020): 189-215.

[55] García-Gonzalo, Esperanza, Zulima Fernández-Muñiz, Paulino José García Nieto, Antonio Bernardo Sánchez, and Marta Menéndez Fernández. "Hard-rock stability analysis for span design in entry-type excavations with learning classifiers." Materials 9, no. 7 (2016): 531.

[56] Parikh, Krupal S., and Trupti P. Shah. "Support vector machine–a large margin classifier to diagnose skin illnesses." Procedia Technology 23 (2016): 369-375.

[57] Web link : http://getdrawings.com/vector-machine#vector-machine-14.jpg.
[58] Panda, Manaswinee Madhumita, Surya Narayan Panda, and Prasant Kumar Pattnaik. "Exchange rate prediction using ANN and deep learning methodologies: A systematic review." In 2020 Indo–Taiwan 2nd International Conference on Computing, Analytics and Networks (Indo-Taiwan ICAN), pp. 86-90. IEEE, 2020.
[59] Osisanwo, F. Y., J. E. T. Akinsola, O. Awodele, J. O. Hinmikaiye, O. Olakanmi, and J. Akinjobi. "Supervised machine learning algorithms: classification and comparison." International Journal of Computer Trends and Technology (IJCTT) 48, no. 3 (2017): 128-138.
[60] https://wiki.pathmind.com/neural-network.
[61] Smys, S., Joy Iong Zong Chen, and Subarna Shakya. "Survey on Neural Network Architectures with Deep Learning." Journal of Soft Computing Paradigm (JSCP) 2, no. 03 (2020): 186-194.
[62] Liu, Weibo, Zidong Wang, Xiaohui Liu, Nianyin Zeng, Yurong Liu, and Fuad E. Alsaadi. "A survey of deep neural network architectures and their applications." Neurocomputing 234 (2017): 11-26.
[63] Singh, Navdeep, and Hiteshwari Sabrol. "Convolutional Neural Networks-An Extensive arena of Deep Learning. A Comprehensive Study." Archives of Computational Methods in Engineering (2021): 1-26.
[64] Kruthika, S., K. V. Anushree, P. Nagesh, and H. B. Niteesh. "Glucoma Disease Detection Using CNN In Deep

Learning." PhD diss., CMR Institute of Technology. Bangalore, 2020.

[65] Mullah, Nanlir Sallau, and Wan Mohd Nazmee Wan Zainon. "Advances in Machine Learning Algorithms for Hate Speech Detection in Social Media: A Review." IEEE Access (2021).

[66] Zhang, Pin, Zhen-Yu Yin, Yin-Fu Jin, Tommy HT Chan, and Fu-Ping Gao. "Intelligent modelling of clay compressibility using hybrid meta-heuristic and machine learning algorithms." Geoscience Frontiers 12, no. 1 (2021): 441-452.

[67] Web link for downloading The Hong Kong Polytechnic University Disguise and Makeup Faces Database, 2016, available http://www.comp.polyu.edu.hk/~csajaykr/DMFaces.htm Download for this thesis in 2020.

[68] Wang, T. Y., & Kumar, A. (2016, February). Recognizing human faces under disguise and makeup. In 2016 IEEE International Conference on Identity, Security and Behavior Analysis (ISBA) (pp. 1-7). IEEE.

[69] Khan, Muhammad Junaid, Muhammad Jaleed Khan, Adil Masood Siddiqui, and Khurram Khurshid. "An automated and efficient convolutional architecture for disguise-invariant face recognition using noise-based data augmentation and deep transfer learning." The Visual Computer (2021): 1-15.

Publication

Journal of Al-Qadisiyah for Computer Science and Mathematics Vol. 13(3) 2021 , pp Comp. 68–77 68

University Of AL-Qadisiyah

Available online at www.qu.edu.iq/journalcm

JOURNAL OF AL-QADISIYAH FOR COMPUTER SCIENCE AND MATHEMATICS

ISSN:2521-3504(online) ISSN:2074-0204(print)

Detection of Human Faces Covered with disguise and Makeup

Farah Jawad Al-ghanim[1], Ali mohsin Al-juboori [2]

[1,2] College of Computer Science and Information Technology, Al-Qadisiyah University, Iraq.

Email: com.post07@qu.edu.iq, Ali.mohsin@qu.edu.iq

ARTICLE INFO

Article history:
Received: 15 /05/2021
Rrevised form: 10 /06/2021
Accepted : 30 /06/2021
Available online: 02 /09/2021

Keywords:

Disguised/Makeup face
Face detection
Facial parts
Machine learning
landmark points
HOG + SVM detector.

ABSTRACT

Face detection is kind of the identification. When we look at someone's face, we can get information like his or her gender and age. Face detection research has exploded in popularity during the last few decades. Starting with algorithms that can detect faces in constrained environments, today's face detection systems can attain extremely great accuracies at the large scale unconstrained facial datasets. While new algorithms continue to increase performance, the majority of face detection systems are vulnerable to failure when subjected to disguise and cosmetics alterations, which is one of the most difficult covariates to overcome. In this article, the database of disguised and makeup faces (DMFD) is employed. In order to address this issue, we detected the location and size of the facial in the image by using Histogram of Oriented Gradients (HOG) + Linear SVM Machine Learning detector on the Disguise and makeup face database (DMFD).This approach is effective and can detect any disguise and makeup faces in the complex background and illumination variation. The results shows the effectiveness of the face detection system on a database (DMFD) and it provided better results of (99.3%).

MSC. 41A25; 41A35; 41A36

DOI : https://doi.org/10.29304/jqcm.2021.13.3.839

1. Introduction

The process that detecting faces in an image is known as face detection. The object is to locate each face and draw a rectangle around it.[1] Face detection is a critical first step in tasks including face recognition, face editing, facial attribute classification, and face tracking, and its accuracy has a direct impact on the tasks' effectiveness. Face detection is extremely important in the surveillance and security paradigms. [2,3] Face recognition is a technique for determining a person's identification after they have been detected. On these subjects, extensive research has been conducted. Another major research issue is detecting hidden faces, particularly in high-security areas such as airports or busy areas such as concerts and retail malls, where they may pose a security danger.[4] Also, People should wear masks during the pandemic to assist successfully prevent the transmission of Coronavirus, particularly

Face Recognition with Disguise and Makeup Variations Using Image Processing and Machine Learning

Farah Jawad Al-ghanim(✉) and Ali mohsin Al-juboori(✉)

College of Computer Science and Information Technology,
Al-Qadisiyah University, Al Diwaniyah, Iraq
{com.post07,Ali.mohsin}@qu.edu.iq

AQ1

Abstract. Face recognition is a research challenge continuous and has seen colossal development during the last two decades. While coming algorithms keep on achieving improved the performance, a greater part of the face recognition systems are receptive to failure under disguise and makeup variations that is one of the common challenging covariates of facial recognition. In past researches, some algorithms show promising results on the existing disguise datasets, still, most of the disguise datasets include images with limited variations (oftentimes captured in controlled settings). This not simulate a real-world scenario, wherever both the intended/ unintended unconstrained disguises and makeup are encountered by a face recognition systems. In this paper, the disguised and makeup faces database (DMFD) is used. In order to handle this problem, One of simple, yet efficient ways for extracting face image features is (LBPH), Principal Component Analysis (PCA) that was majorly utilized in pattern recognition. Also, the technique of Linear Discriminant Analysis (LDA) employed for overcoming PCA limitations was efficiently used in face recognition. Further, classification is employed following the feature extraction. The Naïve Bayes, KNN and Random forest RF algorithms are used. The results paper show the effectiveness and generalization of the proposed system on the Disguise and makeup face database (DMFD) and the features which are extracted by means of (LDA) with (RF) provided the better results of (F-measure, Recall, and Precision).

Keywords: Face recognition · Disguised and makeup faces · Machine learning · Classification

1 Introduction

With fast advancement of the computer and networks technology, data security shows remarkable significance. Personality Identity is a fundamental essential to guarantee the security of the system. The precise recognizable is needed in the fields, national security, finance, e-commerce, justice, and so on. The personal ID system dependent on biometric acknowledgment innovation is getting the most attention for its superior security, validity, and reliability, and has begun to go on into all scopes of our lives. today and in the

M. Singh et al. (Eds.): ICACDS 2021, CCIS 1440, pp. 1–15, 2021.
https://doi.org/10.1007/978-3-030-81462-5_35

ICACDS/2021/Author/058

CONSILIO Intelligence Research Lab

KBTCOE

ICACDS 2021

Springer

CCIS

GISR FOUNDATION

Certificate of Participation

This is to certify that Prof./Dr./Mr./Ms. Farah Jawad Al-ghanim *from* College of Computer Science and Information Technology, Al-Qadisiyah University, Iraq *for virtually participating and presenting a paper entitled* Face Recognition with Disguise and makeup Variations Using Image Processing and Machine Learning *in 2021 5th International Conference on Advances in Computing and Data Sciences (ICACDS-2021) organized by the MVPS'S Karmaveer Adv. Baburao Ganpatrao Thakare College of Engineering, Nashik, Maharashtra, India on April 23-24, 2021.*

Prof. (Dr.) Satish. R. Devane	*Dr. Mayank Singh*	*Prof. (Dr.) P. K. Gupta*	*Prof. (Dr.) Vipin Tyagi*
Principal	General Chair	Advisory Board Chair	Advisory Board Chair
KBTCOET, Nashik	ICACDS 2021	ICACDS 2021	ICACDS 2021

MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH
UNIVERSITY OF MOSUL
COLLEGE OF COMPUTER SCIENCE AND MATHEMATICS

7th International Conference on Contemporary Information Technology and Mathematics (ICCITM)

Certificate of Acceptance

This certificate is granted to

Farah Al-ghanim and Ali mohsin Al-juboori

to crtifie the acceptance of the research paper entitled:

Face Identification Under Disguise and Makeup Based on Hybrid Deep Learning

in ICCITM 2021

which will be held on 25-26 August 2021, in Mosul, IRAQ and is organized by the College of Computer Science and Mathematics/ University of Mosul.
Sponsored by IEEE represented by IEEE Iraq Section.

Prof. Dr. Dhuha Basher Abdullah
Dean of Computer Science and Mathmematics
Conference Chair

Prof. Dr. Sattar B. Sadkhan
IT College, University of Babylon,
IEEE Iraq Section

المستخلـــــص

تم استخدام التعرف على الوجه على نطاق واسع في الأنظمة الذكية المتقدمة (على سبيل المثال: المراقبة الذكية بالفيديو ونظام التحكم الذكي في الوصول للأنظمة المختلفة والدفع عبر الإنترنت). يتم إعاقة أداء الخوارزميات الحالية للتعرف التلقائي على الوجه من خلال العديد من المتغيرات المشتركة مثل اختلافات الوضع، وشيخوخة الوجه، والأقنعة او الكمامات، التنكر، والماكياج. يتم استخدام التنكر والماكياج بشكل خاص لتغيير مظهر الوجه بشكل مقصود أو غير مقصود لإخفاء الهوية الشخصية أو انتحال هوية مختلفة لشخص ما. إنه أيضًا أحد المتغيرات التقليدية الشائعة التي يضعها الناس في حياتهم اليومية. بينما تستمر الخوارزميات الجديدة في تحسين الأداء، فإن معظم أنظمة التعرف على الوجوه عرضة للفشل عند التنكر أو المكياج، وهو أحد أكثر العوامل صعوبة في التغلب عليها. بفضل القدرات الهائلة والنتائج الواعدة، تجتذب تقنية التعلم العميق أكبر قدر من الاهتمام للبحث في مجموعة متنوعة من مهام الرؤية الحاسوبية. للتغلب على هذه المشكلة تم استخدام قاعدة بيانات الوجوه المقنعة والمكياج (DMFD) والتي تحتوي على صور تحت هذين المتغيرين المذكورين. يتم الحصول على صور الوجه من مجموعة بيانات (DMFD) من (Hong Kong Polytechnic) والتي تأتي مع صور تم اقتصاصها وتمر عبر خطوات المعالجة المسبقة قبل إجراء التصنيف. تتضمن مرحلة المعالجة المسبقة تنفيذ معادلة الرسم البياني لتحسين قيمة التباين لصورة الوجه. بعد ذلك، يتم تغيير حجم الصورة لتجهيزها لمرحلة استخراج الميزة. يتم استخراج ملامح صور الوجه في هذه الرسالة باستخدام طريقة التحليل الخطي التمييزي (LDA) التي تعطي الحد الأدنى لطول الميزة والحد الأدنى لوقت الاستخراج. يتم التعرف على الوجوه عن طريق استخدام مصنف التعلم العميق المختلط المقترح لتعلم أكثر دقة للميزات. قمنا أيضًا بمقارنة الطريقة المقترحة مع شبكة مدربة مسبقا (AlexNet) وأربعة خوارزميات للتعلم الآلي ((Naïve Bayes (NB) و K- Nearest Neighbor (KNN) و Random Forest (RF) وآلات ناقلات الدعم (SVM)). تظهر النتائج التجريبية أداء عالي الجودة وقيمة دقة مثالية تساوي 99٪ عند استخدام استخراج ميزة LDA مع المصنف المقترح على صور الوجه. كما كان النظام المقترح سريعًا في مفهوم السرعة مع الوقت الذي يساوي أجزاء من الثانية في تنبؤ الشخص.

جمهورية العراق
وزارة التعليم العالي والبحث العلمي العراقية
جامعة القادسية
كلية علوم الحاسوب وتكنولوجيا المعلومات
قسم علوم الحاسوب

تمييز الوجه تحت التنكر والمكياج على أساس تقنية التعلم العميق

رسالة ماجستير
مقدمة الى مجلس كلية علوم الحاسوب وتكنولوجيا المعلومات في
جامعة القادسية كجزء من متطلبات نيل شهادة الماجستير في
تخصص علوم الحاسوب

مقدمة بواسطة

فرح جواد عبد الكاظم

تحت اشراف

الأستاذ المساعد الدكتور علي محسن محمد

2021 A.D. **1443 A.H.**