EMOTION ESTIMATION FROM FACIAL IMAGES

A MASTER'S THESIS

in

Computer Engineering

Atilim University

By

GOMA MOHAMED SALEM NAJAH

JANUARY 2017

EMOTION ESTIMATION FROM FACIAL IMAGES

A THESIS SUBMITTED TO

THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

OF

ATILIM UNIVERSITY

BY

GOMA MOHAMED SALEM NAJAH

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

IN

THE DEPARTMENT OF COMPUTER ENGINEERING

JANUARY 2017

Approval of the Graduate School of Natural and Applied Sciences, Atılım University.

____

(Prof. Dr.İbrahim Akman)

Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

____

(Prof. Dr.İbrahim Akman)

Head of Department

This is to certify that we have read the thesis “EMOTION ESTIMATION FROM FACIAL IMAGES” submitted by “GOMA MOHAMED SALEM NAJAH” and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

____

(Asst. Prof. Dr.Gökhan Şengül)

Supervisor

Examining Committee Members:

Asst. Prof. Dr. Gökhan Şengül _____________________

Atilim University – Computer Engineering Department

Asst. Prof. Dr. Murat Karakaya _____________________

Atilim University – Computer Engineering Department

Asst. Prof. Dr. Erol Özçelik _____________________

Çankaya University – Psychology Department

Date: (January 27, 2017)

I declare and guarantee that all data, knowledge and information in this document has been obtained, processed and presented in accordance with academic rules and ethical conduct. Based on these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

GOMA MOHAMED SALEM NAJAH

Signature:

ABSTRACT

EMOTION ESTIMATION FROM FACIAL IMAGES

NAJAH, GOMA MOHAMED SALEM

M.S., Computer Engineering Department

Supervisor: Asst. Prof. Dr.Gökhan Şengül

January 2017, 73 pages

Prediction of emotions from facial images is one of the popular and active researches, and it’s implemented via many methods. In this thesis, the proposed system to predict emotions from facial expressions images contains several stages, first stage of this system is the pre-processing stage which is applied by detecting the face in images, then resizing the images, and then Histogram Equalization (HE) technique is applied to normalize the effects of illumination. The second stage is extracting features from facial expressions images using Histogram of Oriented Gradient (HOG), and Local Binary Pattern (LBP) feature extraction algorithms, which generates the training dataset and the testing dataset that contains expressions of Anger, Contempt, Disgust, Embarrass, Fear, Happy, Neutral, Pride, Sad, and Surprised. Then Support Vector Machine (SVM) and K-Nearest Neighbors (KNN) classifiers are used for the classification stage in order to predict the emotion. In addition, Confusion Matrix (CM) technique is used to evaluate the performance of these classifiers. The proposed system is tested on JAFFE, KDEF, MUG, WSEFEP, TFEID and ADFES databases. However, the proposed system achieved prediction rate of 96.13% when HOG+SVM method is used.

Keywords: Emotion estimation; Facial Expression Images; Expression Classification; Histogram of Oriented Gradient; Local Binary Pattern; K-Nearest Neighbors; Support Vector Machine.

ÖZET

YÜZ GÖRÜNTÜLERİ ÜZERİNDEN DUYGU TAHMİNİ

NAJAH, GOMA MOHAMED SALEM

Bilgisayar Mühendisliği Bölümü, Yüksek Lisans

Danışman: Yrd. Doç. Dr.Gökhan Şengül

Ocak 2017, 73 sayfa

Yüz görüntüleri üzerinden duygu tahmininde bulunma son zamanlardaki popüler ve etkin araştırmalardan biri olup bu araştırmalar birçok farklı yöntem aracılığıyla uygulanmaktadır. Bu tezdeki yüz ifadelerini tahmin edebilmek için önerilen sistem bir takım aşamalar içermektedir ve bunlardan birincisi görüntüler içinden yüzün seçilip bu görüntülerin yeniden boyutlandırılması ve sonrasında aydınlatma etkilerini normalize etmek için uygulanan Histogram Eşitlemesi- *Histogram Equalization (HE)* aracılığıyla yürütülen ön işleme aşamasıdır. İkinci aşama ise Odaklı Gradyan Histogramı- *Histogram of Oriented Gradient (HOG)* ve Yerel İkili Model- *Local Binary Pattern (LBP)* özellik çıkarma algoritmaları kullanarak yüz ifadelerinden Öfke, Kibir, İğrenme, Utanma, Korku, Mutluluk, Yansızlık, Gurur, Üzgün Olma ve Şaşkınlık gib farklı ifadelerinin özellik çıkarma aşamasıdır. Özellik çıkarma aşamasından sonra Karar Destek Makineleri - *Support Vector Machine (SVM) ve* k-En Yakın Komşuluk - *K-Nearest Neighbors (KNN)* sınıflandırıcıları kullanılarak duygu tahmininde bulunulmuştur. Buna ek olarak, Karışıklık Matrisi- *Confusion Matrix (CM)* tekniği bu sınıflandırıcıların performanslarını değerlendirmek için kullanılmıştır. Önerilen bu sistem JAFFE, KDEF, MUG, WSEFEP, TFEID ve ADFES veritabanlarında test edilmiştir ve önerilen sistemin HOG+SVM yöntemi uygulandığında 96.13% oranında bir tahmin başarısına ulaşılmıştır.

Anahtar Sözcükler: Duygu Tahmini; Yüz İfadesi Görüntüleri; İfade Sınıflandırması; Odaklı Gradyan Histogramı; Yerel İkili Model; K- En Yakın Komşular; Destek Vektör Makinesi

To My Family

ACKNOWLEDGMENTS

I express sincere appreciation to my supervisor Asst. Prof. Dr.Gökhan Şengül for his guidance and insight throughout the research. I would also like to thank my examining committee members, Asst. Prof. Dr. Murat Karakaya and Asst. Prof. Dr. Erol Özçelik for their constructive comments and feedbacks. Thanks also go to my family and my wife. I offer sincere thanks for them continuous support and patience during this period.

I would also like to thank my friends, specifically Mohamed Alhamrouni. They were always supporting me and encouraging me with their best wishes.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

HOG	Histogram of Oriented Gradients
LBP	Local Binary Patterns
KNN	K-Nearest Neighbors
SVM	Support Vector Machine
JAFFE	The Japanese Female Facial Expression Database
DCT	Discrete Cosine Transform
FFT	Fast Fourier Transform
SVD	Singular Value Decomposition
CK	Cohn-Kanade Database
PCA	Principal Component Analysis
KTFE	Kotani Thermal Facial Expression Database
MUFE	Mevlana University Facial Expression Database
SVD	Singular Value Decomposition
DWT	Discrete Wavelet Transform
LDA	Linear Discriminant Analysis
DT-CWT	Dual Tree-Complex Wavelet Transform
GWT	Gabor Wavelet Transform
NN	Neural Network
BPNN	Back-Propagation Neural Network
FER	Facial Expression Recognition Database

RFD	Radboud Faces Database
LTP	Local Ternary Patterns
WPCA	Weighted Principal Component Analysis
PPCA	Pure Principal Component Analysis
CKACFEID	Cohn-Kanade AU-Coded Facial Expression Image Database
WT	Wavelet Transform
GF	Gabor Filter
GD	Gaussian Distribution
PDM	Point Distribution Model
AP	Action Parameters
GLCM	Gray-Level Co-occurrence Matrix
PHOG	Pyramid of Histogram of Oriented Gradients
LPQ	Local Phase Quantisation
LMNN	Largest Margin Nearest Neighbor
2DPCA	2D Principal Component Analysis
IFED	Indian Facial Expression Image Database
TFEID	Taiwanese Facial Expression Image Database
MUG	Multimedia Understanding Group Database
KDEF	The Karolinska Directed Emotional Faces Database
WSEFEP	Warsaw Set of Emotional Facial Expression Pictures
ADFES	The Amsterdam Dynamic Facial Expression Set Database
TFEID	Taiwanese Facial Expression Image Database
HE	Histogram Equalization

CHAPTER 1

1. INTRODUCTION

The human face provides a number of social signals which are essentials for interpersonal communication in our everyday life. The human face also holds very important quantity of attributes and information about the person, such as facial expression, ethnic, gender, and age. Facial expression is a movement of facial muscles by a human involuntarily when they feel something like anger, happiness and fear…. etc. *(figure 1-1)*. Humans can recognize facial expressions virtually without any error or delay. But reliable and fully automated expression recognition by computers is still a challenge. Various approaches have already been attempted towards addressing this problem, but the complexities added by circumstances like inter-personal variation (ie. gender, ethnic) and inconsistency of acquisition conditions (i e. illumination, resolution) have made the task quite complicated and challenging. However, the recent advancements in the area of image analysis and pattern recognition have opened up the possibility of automated measurement of facial signals. It is believed that the automated analysis of facial expressions can facilitate machine perception of human facial behavior, and thus opens up the way of bringing facial expressions into man-machine interaction as a new modality towards making the interaction more natural and efficient. It can also enable the classification and quantification of facial expressions widely accessible for the research in behavioral science and medicine by automated psychological observation of humans. Keeping all of these in consideration, this thesis addresses the various complexities related to the classification of the encountered facial expressions present in static facial images and thus provides a solution of the problem of classifying the seven important facial expressions namely, neutral, anger disgust, fear, happiness, sadness and surprise. Six of these expressions except for the neutral" have been defined as the basic human emotions by Ekman [1].

Computer vision involves techniques and algorithms that are capable of analyzing and understanding image content, additionally to extract information from images. In other words, these techniques and algorithms make the computer able to understand human expressions. It can improve the communications between human and machines, and it is useful in human-machine interaction [2], thus in the future, robots/machines can understand human behavior. Further applications lie in security, driver safety and social sciences as a tool to analyze human affective behavior.

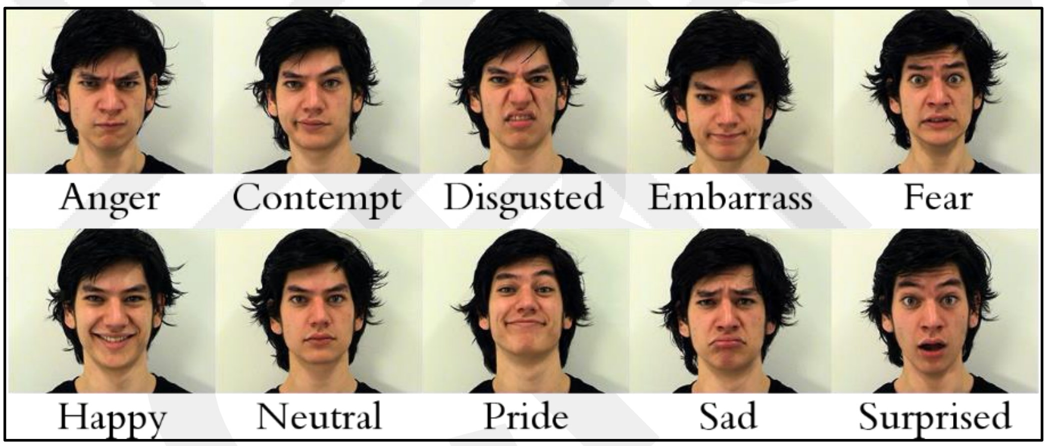

Figure 1-1: Ten universal facial expressions

The design of an emotion prediction system includes three basic stages *(figure 1-2)*, namely preprocessing stage, feature extraction stage and matching (classification) stage.

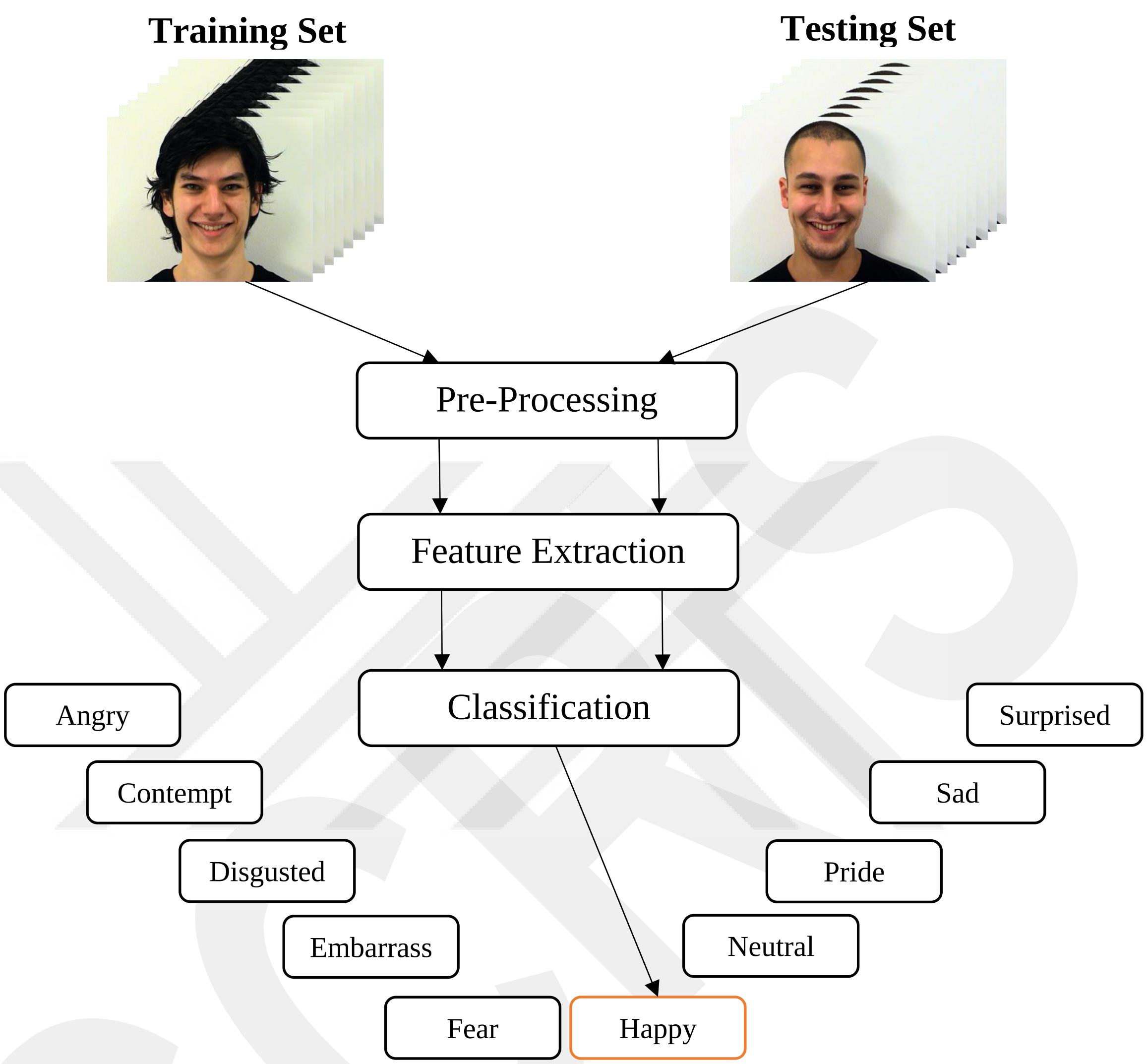

Figure 1-2: The basic stages of emotions estimation

1.1. Motivation

Facial expressions are important capability for many practical applications, it is useful for interaction between human and computer, perceptual user interfaces, distance learning and interactive computer games [3] [4]. Additionally, it is considered as an effective way of treatment for people with psycho and affective illnesses such as autism, moreover; it is a universal language, for instance; sign language to deal between people and cultures. In other words; Facial expression is a visible manifestation of the cognitive activity, intention, personality and psychopathology of a person that plays a communicative role in interpersonal relations as well as in human-to-human communication and interaction, allowing people to express themselves beyond the verbal domain [5]. Facial expressions are an important part wherever humans interact with machines or robots. So, the automated classification of facial expressions may act as a component of both natural human-machine interface and its variation which is known as the perceptual interface. Since the aggregation of the emotional information with human-computer interfaces allows much more natural and efficient interaction paradigms to be established, development of a system for the automated classification of facial expressions can play an increasing role in building effective and intelligent multimodal interfaces for next generation. This can also be a possible application domain for a diverse of disciplines including behavioral science, medicine, monitoring, communications, education, face modeling as well as face animation [6]. The main challenge in facial expression is many individuals are showing the same expression, but in a different way, some individuals' exhibit two expressions or more in the same way, therefore; it can't be known whether this person is crying or laughing. These challenges make selecting the most important features and ignoring others an important method to predict the emotions.

1.2. Problem description

In this thesis, emotion estimation from facial images is studied. Thus, the main challenge is to compare the performance of emotion estimation methods.

Many applications and researches focus on emotions prediction of individuals from facial images, and there are still many obstacles that perform to incorrect results, and low performance. For example, the facial images have a wide degree of differences in illumination, unconstrained environments, head pose, facial expression, image background, image dimensions, shadows, quality, skin color…. etc. which often can result incorrect classification. Hence, the system shall standardize and normalize the photometric characteristics of the facial images, and then it extracts the most descriptive features that allow more facial discrimination. As a result, these data are used to predict the emotions from facial expression images of individuals.

The basic aim of this thesis is to design and implement a system that can be able to increase the performance of emotion prediction of individuals through their facial expressions images.

1.3. Thesis Outlines

This thesis begins by reviewing general information about emotions, feelings and their role in human behavior, and also giving a background information about facial expressions of emotion. In Chapter 2, some literature reviews are reviewed, which contains related works in the field of emotion prediction from facial expression. In Chapter 3, the main methodology of this thesis is explained by giving general description for the two techniques that have used to extract features. These techniques are Histogram of Oriented Gradients (HOG) and Local Binary Patterns (LBP). In addition, the classification algorithms Support Vector Machine (SVM) and K-Nearest Neighbors (KNN) used in this thesis are briefly described, moreover, a brief insight into the evaluation technique (confusion matrix) is discussed, which is used to evaluate the performance of classifiers. Chapter 3 ended with providing general information about the databases that are used in this thesis. In Chapter 4, some experimentation was implemented, and various improvements were carried for the dimension alignment and illumination normalization of the facial images in order to increase the accuracy and performance of emotion prediction from facial expression. Moreover, Confusion Matrix technique has been used in order to evaluate the performance of our system. The most important results obtained when performing the proposed experiments are illustrated and discussed in this chapter. Finally, in Chapter 5, the most essential points of this thesis are summarized and concluded, and some advices are given for the future work and potential steps to improve the performance of emotion prediction from facial expression system are also presented.

CHAPTER 2

2. LITERATURE REVIEW

Papers that studied emotion estimation system from facial images have presented in the first section in this chapter. Then in the second section, a summarized table with brief details were presented. The table contains researcher name, used method, database name, images number, and the results obtained. Finally, a summary of the reviewed papers and the best results obtained are provided.

2.1. Previous studies of emotion estimation system from facial images

Chung-Lin and Yu-Ming [7] proposed Point Distribution Model (PDM) approach to analyze facial expression based on the facial feature extraction. PDM approach analysis the statistics of the coordinates of the classified or labeled points over the training set. The proposed approach is performed by using 180 images from 15 volunteers, each volunteer demonstrates six expressions, and then 12 images are chosen from each volunteer. Action Parameters (AP) Classifier is performed in order to classify and match the extracted features from facial images. The proposed approach achieved overall accuracy of 84.41%.

Support Vector Machine (SVM) algorithm is used by Philipp and Rana [8] to classify Cohn-Kanade (CK) facial expression and live video in order to identify the emotions universally recognized which are (e.g. for the basic emotions of 'anger', 'disgust', 'fear', 'joy', 'sorrow' or 'surprise') supplied during training. In this study, SVM achieved 87.90% of recognition performance.

Different feature extraction techniques were presented separately in the study performed by Tommaso, Caifeng, Vincent and Ralph [9] in order to extract features from facial images; These techniques are Histogram of Oriented Gradients (HOG), Local Binary Patterns (LBP), and Local Ternary Patterns (LTP). The used techniques were applied with various parameters of facial expression recognition. The extracted features were classified by Support Vector Machine (SVM). The Cohn-Kanade (CK) database was used, which was created from 100 persons, their ages from 18 to 30 years. 310 images were selected from CK database for these experiments. LBP achieved the best accuracy of recognition that reached to 92.9%, HOG achieved 92.7%, and finally; LTP achieved 91.7%.

Kharat and Dudul [10] investigated three various techniques for feature extraction from facial expressions for emotion recognition on six universally recognized basic emotions, namely angry, disgust, fear, happy, sad and surprise along with neutral one. These techniques are Discrete Cosine Transform (DCT), Fast Fourier Transform (FFT) and Singular Value Decomposition (SVD). Support Vector Machine (SVM) classifier is used to classify the extracted facial features. The study is performed using JAFFE database. This database contains 219 images. DCT+SVM method achieved recognition rate of 94.29%, and FFT+SVM method achieved 94.29%, and SVD+SVM method achieved 92.86%.

Caifeng, Shaogang and Peter [11] used Local Binary Patterns (LBP) approach to extract features from facial images in order to predict the emotions. These features were classified by Support Vector Machine (SVM) algorithm. This experiment has been applied on two different databases, First, MMI database which contains 96 images sequences for age from 19 to 62. The sequences come from 20 subjects, with 1–6 emotions per subject. Second, JAFFE database that contains 213 images. They achieved an accuracy of 86.70% with MMI database and accuracy of 79.80% with JAFFE database.

Murugappan, Nagarajan and Yaacob [12] extracted features from facial images by using Discrete Wavelet Transform (DWT) approach. They collected 460 image from a series of videos and used in their experiments. After extracting the features from images, they used two various classifiers to classify these features K-Nearest Neighbors (KNN) algorithm and Linear Discriminant Analysis (LDA) algorithm. LDA provides fast evaluations for input samples by calculating the distance between a new sample and training samples in each class weighed by their variability matrices. LDA tries to find an optimal hyper plane to six classes of emotions (fear, neutral, happy, disgust, surprise and sad). Recognition rate of 83.26% were achieved with KNN classifier and 75.21% with LDA classifier.

The Weighted Principal Component Analysis (WPCA) and Pure Principal Component Analysis (PPCA) were suggested by Zhiguo and Xuehong [13] as feature extraction algorithms. The proposed algorithms are considered the most two important methods that are used in order to reduce dimension and feature extraction from human facial region. The proposed algorithms were applied separately on Cohn-Kanade (CK) AU-Coded Facial Expression Image Database (CKACFEID) which includes 500 images from 100 persons. Researchers selected 400 images of this database for their experiments. Support Vector Machine is used to evaluate the proposed algorithms. They have obtained recognition rate of 88.25% by WPCA with SVM classifier, and 84.75% by PPCA with SVM classifier.

Abhinav, Akshay, Roland and Tom [14] presented Pyramid of Histogram of Oriented Gradients (PHOG) and Local Phase Quantisation (LPQ) techniques to extract and encode facial features. The features firstly extracted using PHOG technique, and then using the combination of PHOG and LPQ techniques. 289 images of GEMEP-FERA dataset are used in these experiments. Researchers applied three methods, PHOG technique with SVM classifier, (PHOG&LPQ) with SVM classifier, and (PHOG&LPQ) with Largest Margin Nearest Neighbor (LMNN) classifier. The proposed methods have achieved recognition rate of 67%, 72.40% and 73.40% respectively.

Discrete Cosine Transform (DCT), Wavelet Transform (WT), Gabor Filter (GF), and Gaussian Distribution (GD) techniques were combined by Sandeep, Shubh, Meena and Neeta [15] in order to extract features from facial images to improve the recognition rate. This experiment is applied on seven emotions (sadness, fear, surprise, anger, happiness, neutral and disgust) of JAFFE database which contains 213 images. For this experiment, only 126 images out of 213 images selected, 60% of images for training set, and 40% for testing set. ADABOOST (ADB) classifier is used to evaluate the proposed techniques performance. This method achieved recognition rate of 93.4%.

Principal Component analysis (PCA) approach is suggested by Mandeep and Rajeev [16] to extract features from facial image in order to identify or detect the person emotions (Disgust, Angry, Sad, Happy and Surprise) of his facial image. Researchers extracted features and classified these features by Singular Value Decomposition (SVD) classifier. They used 31 test images and 50 train images from both JAFFE database and real database for their experiments. The results show that using PCA approach with SVD algorithm achieved high performance, the experiments achieved recognition rate of 100%.

Patch-based Gabor features is a various approach applied by Ligang and Dian [17] to extract features from facial images. This approach is characterized in terms of extracting regional features and keeping the facial region information. The proposed is approach applied on the JAFFE database which consists of 213 gray images for seven emotions (six basic and one neutral). The Cohn-Kanade (CK) database is also used in this experiment. Researchers classified the extracted features using Support Vector Machine (SVM). They have obtained an accuracy of 92.93% of prediction with JAFFE database, and 94.48% of prediction with CK database.

Punithaand Geetha [18] used a various technique to identify the facial expression from facial image. Gray-Level Co-occurrence Matrix (GLCM) technique is used to extract features from facial image. The extracted features are highly efficient and require less computation time. Here Support Vector Machine (SVM) is used to train the extracted features using various kernels to recognize the main emotions Happy, Surprise, Disgust, Neutral and Sad. Researchers applied their experiment on Facial Expression Database, and achieved recognition rate of about 90%.

2D Principal Component Analysis technique (2DPCA) is applied by Bagga et.al [19] as a feature extractor. In this study, firstly, 2DPCA approach is used to extract features. Secondly, 2DPCA technique is combined with LBP approach to improve the performance of proposed method which is applied on Cohn-Kanade (CK) facial expression dataset. The database contains 2000 images. Euclidian Distance (ED) Classifier is used to evaluate the performance of proposed system. Moreover, 2DPCA+ HD method achieved 95.13% and (2DPCA&LBP) + HD method achieved 95.83% of recognition.

Junkai *et al* [20] proposed an effective method to extract features from facial expression image. In research, the proposed technique detects and extracts the facial components from the facial image instead of using a whole facial image. Histogram of Oriented Gradients (HOG) technique is used for this task. Support Vector Machine (SVM) classifier is applied to perform the feature classification step. The proposed method has been evaluated on two databases. Firstly; the JAFFE database which contains 213 images from 10 persons, and secondly; Cohn-Kanade (CK) database which include 593 images from 123 persons, but they used 327 images. The overall average of recognition was 94.30% on JAFFE database and 88.70% on Cohn-Kanade (CK) database.

In feature extraction stage, Manar, Aliaa and Atallah [21] used Histogram of oriented of gradients (HOG) to extract facial features in order to identify emotions through these features. Support Vector Machine (SVM) classifier was applied to classify and match these features. This work is implemented through using the Cohn-Kanade (CK) database version 2, which include 593 images from 123 persons, and random database which is collected from the internet. Here the proposed method is applied on static images and videos from both databases. This method achieved accuracy of 95% of recognition on static images, and 80% of recognition on videos.

Two different approaches of feature extraction were combined by Suja, Tripathi, and Deepthy [22] in order predict the emotion from facial expression. These approaches are Dual Tree Complex Wavelet Transform (DT-CWT), and Gabor Wavelet Transform (GWT). JAFFE database and Cohn-Kanade (CK) database were used in this work. JAFFE database contains 219 images taken from 10 Japanese people, each person has six emotions and neutral. CK database consists of 720 images taken from 30 persons, each emotion consists of 120 images. To classify the extracted features, they used KNN classifier and Neural Network (NN) classifier. The overall accuracy was 93% for NN, and 80% for KNN.

Muzammil and Alaa [23] proposed facial expression recognition method which is based on Principal Component Analysis (PCA) to extract features from facial image in order to achieve the emotion prediction. This experiment has been applied on two databases, Mevlana University Facial Expression (MUFE) database which contains 630 images of 15 persons and Japanese Female Facial Expression (JAFFE) database which contains 213 images of 10 persons. Support Vector Machine (SVM) classifier is used to classify and match the extracted features. PCA + SVM method has achieved recognition rate of 87% with JAFFE database and recognition rate of 77% with MUFE database.

Histogram of Oriented Gradients (HOG) approach were proposed by Carcagnì, Coco, Leo and Distante [24] to be used in feature extraction step in order to identify the emotions of facial expression from facial image. Support Vector Machine (SVM) classifier is used to evaluate the performance of feature extraction approach (HOG). In this experiments, researchers applied three different databases, Cohn-Kanade (CK2) database Version 2, Facial Expression Recognition database (FER), and Radboud Faces Database (RFD). Researchers have obtained an accuracy of 95.9% of recognition on Cohn-Kanade (CK2) database, 94.9 % of recognition on FER database, and 94.15% of recognition on RFD database.

Local Binary Pattern approach is used by Talele Kiran and Tuckley Kushal [25] to extract features from facial image. Here features are calculated by tracking the pixels of each block in clockwise and anticlockwise direction, these features calculated in binary form, then collected with each other to get the features of complete face image. Researchers applied their experiments on Indian Facial Expression Image Database (IFED), Taiwanese Facial Expression Image Database (TFEID), and Japanese Female Facial Expression Database (JAFFE). Support Vector Machine (SVM) is used to classify the obtained features. The proposed method achieved efficiency of recognition 96.47% on (TFEID) database, 95.08% on (IFED) database, and 97.10 % on (JAFFE) database.

Wasista, Setiawardhana and Oktavia [26] performed facial emotions prediction by using Principal Component Analysis (PCA) for features extraction from facial image. The aim of using PCA is to take a critical component of the facial image (dimension reduction) in order to make smaller dimensions for processed data. To classify the extracted features, Support Vector Machine (SVM) is used for facial emotions classification process, and then this work is applied on real time images. They have achieved an accuracy of 64.95%.

The histogram-based data analysis is used by Basu, Routray and Deb [27] for feature extraction. It is considered the one of most popular techniques in the feature extraction field. In addition, it is basically statistics base features where it is used to represent the probability distributions of the intensity levels. The researchers used Multi-class Support Vector Machine (SVM) to classify the extracted features. This method was applied on Kotani Thermal Facial Expression Database (KTFE) which contains 264 images of 22 subjects for just 4 expressions for people from Japanese, Vietnamese and Thai, whose ages range from 11 years old to 32 years old. They used 40% of images for testing and

60% for training. The obtained results show high accuracy of prediction reached to 81.95%.

Nikunja, Korra and Sanjay [28] used three techniques together in feature extraction stage, Histograms of Oriented Gradients (HOG), Principal Component Analysis (PCA), and Linear discriminant analysis (LDA). HOG is used to extract features while PCA is used to reduce features. Moreover, LDA is used to select the most important discriminant features. Researchers have applied their experiments on Cohn-Kanade (CK) dataset. 414 images were selected from CK database for this work, 105 neutral images, and 309 peak expressive images were used to perform this experiments. Finally, Back-Propagation Neural Network (BPNN) classifier and Support Vector Machine (SVM) classifier have been selected to classify the obtained features. The obtained accuracy is 99.51% with BPNN, and 99.27% with SVM.

2.2. Previous studies result

Table 2-1 shows a brief detail of previous studies which reviewed emotion estimation from facial images. This table includes researchers' names, methods, databases, the total number of images, number of expressions, number of train images, number of test images and the accuracy achieved in each study.

SN	Researcher name	Methods	Databases	Total Number of Images	Number of Expressions	Number of Train Images	Number of Test Images	Accuracy
1)	Chung-Lin & Yu-Ming [7]	PDM+AP	Real DB	180	6	90	90	84.41%
2)	Philipp & Rana [8]	SVM	(CK&CK+) and live video	-	6	-	-	87.90%
3)	Tommaso, Caifeng, Vincent & Ralph [9]	HOG+SVM	CK	310	6	-	-	92.70%
		LBP+SVM	CK	310	6	-	-	92.90%
		LTD+SVM	CK	310	6	-	-	91.70%
4)	KHARAT &DUDUL [10]	DCT+SVM	JAFFE	219	7	-	-	94.29 %
		FFT +SVM	JAFFE	219	7	-	-	94.29 %
		SVD +SVM	JAFFE	219	7	-	-	92.86 %
5)	Caifeng, Shaogang and Peter. [11]	LBP+SVM	MMI	79	6	-	-	86.70%
			JAFFE	213	7	-	-	79.80%
6)	Murugappan, Nagarajan, & Yaacob [12]	DWT+KNN	Real Database	460	5	-	-	83.26%
		DWT+LDA	Real Database	460	5	-	-	75.21 %

7)	Zhiguo & Xuehong [13]	WPCA+SVM	CKACFEID	500	6	100	400	88.25%
		PPCA+SVM	CKACFEID	500	6	100	400	84.75%
8)	Abhinav, Akshay, Yogesh &Tom [14]	PHOG+SVM	GEMEP-FERA	289	5	155	134	67%
		(PHOG&LPQ) + SVM	GEMEP-FERA	289	5	155	134	72.40%
		(PHOG&LPQ) + LMNN	GEMEP-FERA	289	5	155	134	73.40%
9)	Sandeep, Shubh, Yogesh &Neeta [15]	(DCT+WT+G F+GT) +ADB	JAFFE	213	7	126	87	93.40%
10)	Mandeep& Rajeev [16]	PCA+SVD	JAFFE	14	6	7	7	100%
			Real Database	81	6	50	31	100 %
11)	Ligang & Dian [17]	Patch-based Gabor+SVM	JAFFE	213	7	143	70	92.93%
			CK	327	6	247	80	94.48%
12)	Punitha & Geetha [18]	GLCM+SVM	Facial expression DB	-	-	-	-	90%
13)	Bagga, S., *et al* [19]	2DPCA+ ED	CK	2000	6	1700	300	95.13%
		2DPCA&LBP + ED			6	1700	300	95.83%
14)	Junkai, Zenghai, Zheru & Hong [20]	HOG+SVM	JAFFE	213	7	190	23	94.30%
		HOG+SVM	CK	327	7	268	59	88.70 %
15)	Manar , Aliaa & Atallah [21]	HOG+SVM	CK2 & random DB	Static images	6	-	-	95%
			CK2& random DB	videos	6	-	-	80%
16)	Suja, Tripathi, & Deepthy [22]	(DT-CWT+GWT) + KNN	JAFFE	140	7	70	70	100%
			CK	720	6	540	180	80%
		(DT-CWT+GWT) + NN	JAFFE	140	7	70	70	95.71%
			CK	720	6	540	180	99.86%
17)	Muzammil&Alaa [23]	PCA + SVM	JAFFE	213	7	137	76	87%
			MUFE	630	7	315	315	77%
18)	Carcagnì, Coco, Leo & Distante [24]	HOG+SVM	CK2	347	6	-	-	95.9%
			FER	-	-	-	-	94.9 %
			RFD	469	7	-	-	94.15%
19)	Kiran, T. & T.Kushal [25]	LBP+SVM	IFED	-	6	-	-	96.47%
			TFEID	328	6	272	56	94.77%
			JAFFE	213	6	144	69	97.10%
20)	Wasista,Setiawardhana& Oktavia [26]	PCA+SVM	Real Time Images	-	4	-	-	64.95%
21)	Basu, Routray&Deb [27]	Histogram Features Extraction	KTFE	264	4	159	105	81.95%
22)	Nikunja, Korra& Sanjay [28]	(HOG+PCA+ LDA) +BPNN	CK	414	7	331	83	99.51%
		(HOG+PCA+ LDA) +SVM	CK	414	7	331	83	99.27%

Table 2-1: Common Emotion prediction methods

2.3. Summary

Based on the literature review and the brief table of the researches and papers which studied the topic mentioned above, it can be concluded that PCA, DCT, SVD and FFT techniques achieved high performance on various databases. On other hand, SVM classifier is the best algorithm that is used to classify the extracted features from facial images in these studies. Moreover, JAFFE and CK databases were more widely used because the images in these databases are very clear, therefore, these databases helped to achieve good results.

CHAPTER 3

3. METHODOLOGY

In this chapter, a general overview of the databases used in this thesis is provided, followed by detail description of the four main stages *(figure 3-1)* used to implement emotion estimation system from facial images. Firstly, image pre-processing; in this stage the image goes through several processes, such as face detection in order to determine target region, then image resizing to control an image size, and then Histogram Equalization applied on the image to control the illumination degree of the image. Secondly, feature extraction; in this stage two approaches are applied to extract features from the image, these approaches are (HOG and LBP). Thirdly, SVM and KNN classifiers are used to classify the extracted features. Finally, the performance of the classifiers is evaluated by confusion matrix technique.

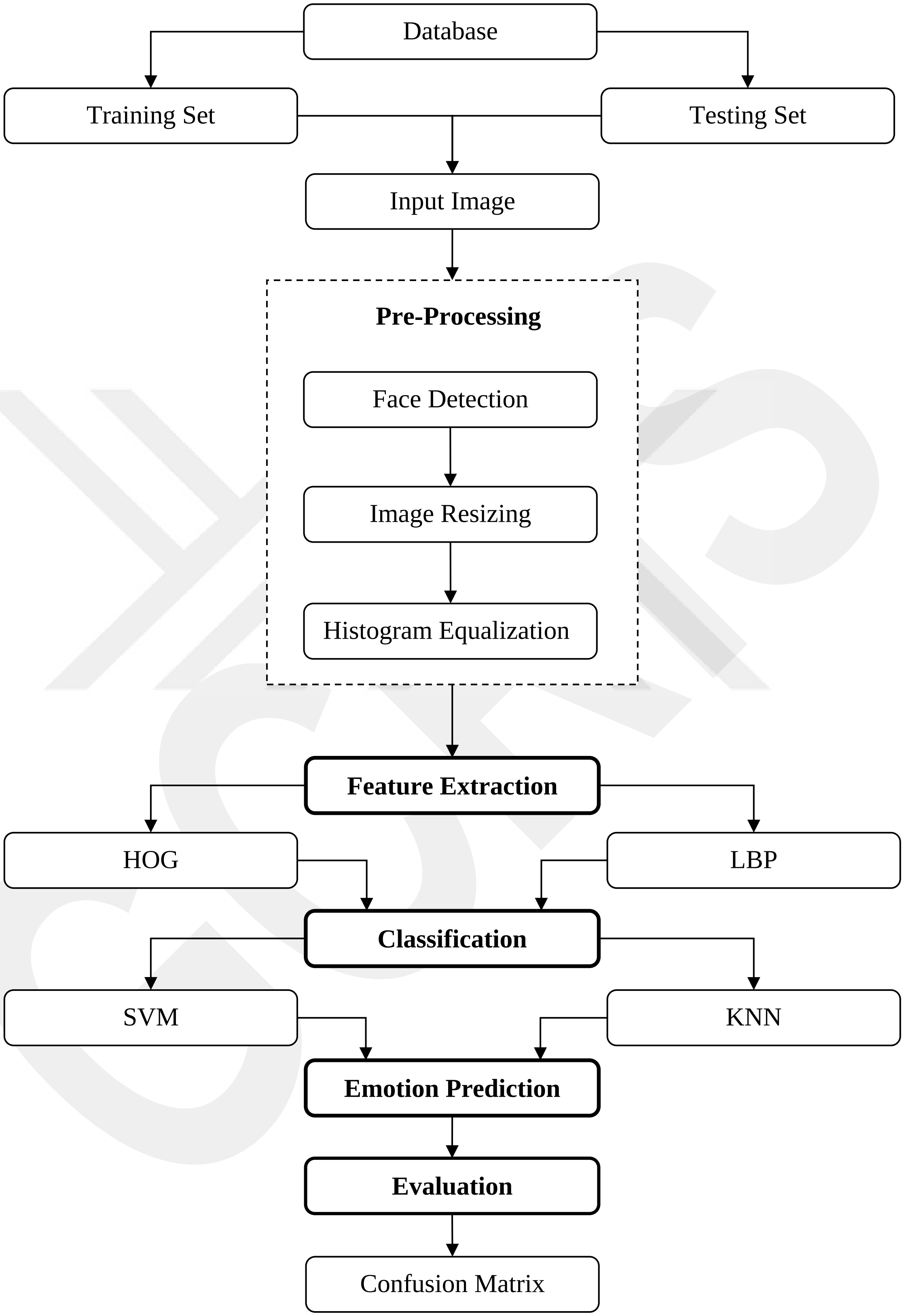

Figure 3-1: Emotion Prediction System

3.1. Databases

Six databases are used in this thesis. These are TFEID, ADFES, WSEFEP, MUG, KDEF and JAFFE. First database has ten facial expressions and second one has eight facial expressions, and the others has seven basic facial expressions. *(Table 3-1)* illustrates all databases used in the experiments, and the number of images in each one. Every facial expression is considered as a class.

Expressions	**Angry**	**Contempt**	**Disgusted**	**Embarrass**	**Fear**	**Happy**	**Neutral**	**Pride**	**Sad**	**Surprised**	
Databases	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8	Class 9	Class 10	**Total**
ADFES	22	21	22	21	22	22	22	20	22	21	**215**
TFEID	34	68	40	/	40	40	39	/	39	36	**336**
WSEFEP	30	/	30	/	30	30	30	/	30	30	**210**
MUG	260	/	255	/	240	260	260	/	245	260	**1780**
KDEF	140	/	140	/	140	140	140	/	140	140	**980**
JAFFE	30	/	29	/	32	31	30	/	31	30	**213**
Total	516	89	516	21	504	523	521	20	507	517	**3734**

Table 3-1: Total number of images in all databases

Moreover, the databases will be presented & explained separately in the following sections.

3.1.1. TFEID Database

Taiwanese Facial Expression Image Database (TFEID) [29] is an interesting Facial Expression Image database, which is established by the Brain Mapping Laboratory (National Yang-Ming University), and Integrated Brain Research Unit (Taipei Veterans General Hospital). This database consists of 7200 stimuli captured from 40 models (20 males and 20 female), each model with eight facial expressions: anger, contempt, disgust, fear, happiness, neutral, sadness and surprise, which were based on the operative definition from Ekman's intervention (2003). All images in this database are captured by two CCD-cameras in different viewing angels (0° and 45°). Each expression includes two kinds of intensities (high and slight). Sample images from the database can be seen in (Figure 3-2 and Figure 3-3).

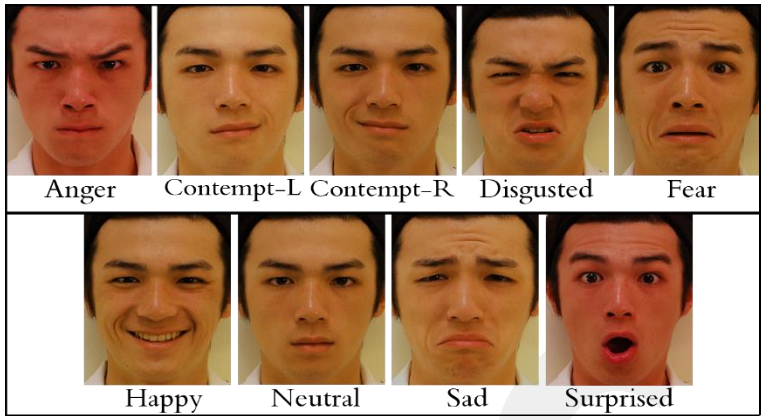

Figure 3-2: Examples from color TFEID Database

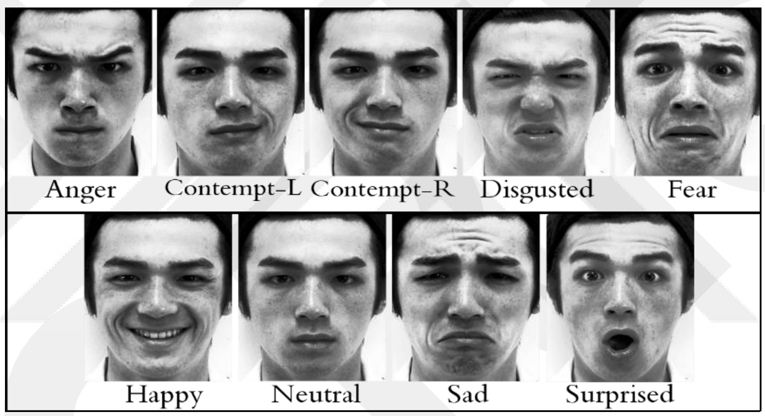

Figure 3-3: Examples from Grey Scale TFEID Database

3.1.2. ADFES Database

The Amsterdam Dynamic Facial Expression Set (ADFES) [30] has an advantage of containing ten facial expressions, which are anger, contempt, disgust, embarrassment, fear, joy, neutral, pride, sadness and surprise. This database contains of full videos set in MPEG-2 formats and high dynamic (HD) recording and still pictures. 22 models (10 female, 12 male) are between the ages 18 and 25. All images in this database have been captured from different viewing angels (0° and 45°). In total this database have 216 images *(Figure 1-1)*.

3.1.3. WSEFEP Database

Warsaw Set of Emotional Facial Expression Pictures (WSEFEP) [31] contains 210 images taken from 30 models (14 male and 16 female) of each having basic expression (anger, disgust, fear, happy, neutral, sadness, surprise). 100 people are chosen from over 1000 actors and then this number is decreased to 65 people depending to some session of work with them. After that, some exercises are given to them during hours to develop their expression skills. Finally, all images are evaluated and the best of them are selected *(Figure 3-4)*.

Figure 3-4: One sample image for each basic expression of WSEFEP Database

3.1.4. MUG Database

Multimedia Understanding Group Database (MUG) [32] is created by Multimedia Understanding Group. Additionally, many people participated to collect this database. It consists of 86 subjects (51 men and 35 women). All subjects are between ages of 20 to 35 years. Only 52 subjects are available to internet researchers. The images are captured by one camera and two light sources of 300W each, the subject sits on a chair in front of blue background. The image capturing rate was 19 frame per second with size 896X896 pixels as JPEG format. Each subject has seven expressions and each expression is saved in a few image sequences (usually three to five), and each sequence contains 50 to 160 images. In total, each subject has more than 1462 images. *(Figure 3-5)*.

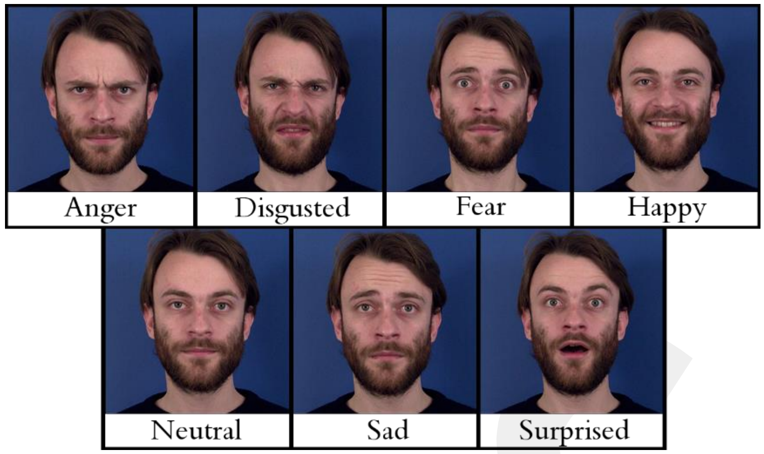

Figure 3-5: One sample image for each basic expression of MUG Database

3.1.5. KDEF Database

The Karolinska Directed Emotional Faces (KDEF) [33] is a standard and non-commercial large database of facial expression images. This database is developed in 1998 by Anders Flykt, Daniel Lundqvist and Professor Arne Öhman at Karolinska Institutet, Stockholm, Sweden. This database contains 4900 images captured by (Pentax LX) Camera, and taken from 70 individuals (35 male and 35 female), their age ranging from 20 to 30 years. Each individual displays 7 expressions (angry, disgusted, afraid, happy, neutral, sad, surprised), which are captured twice from 5 different angles (-90, -45, 0, +45, +90 degrees) and saved as JPEG format with size 562X762 pixels. *(Figure 3-6).*

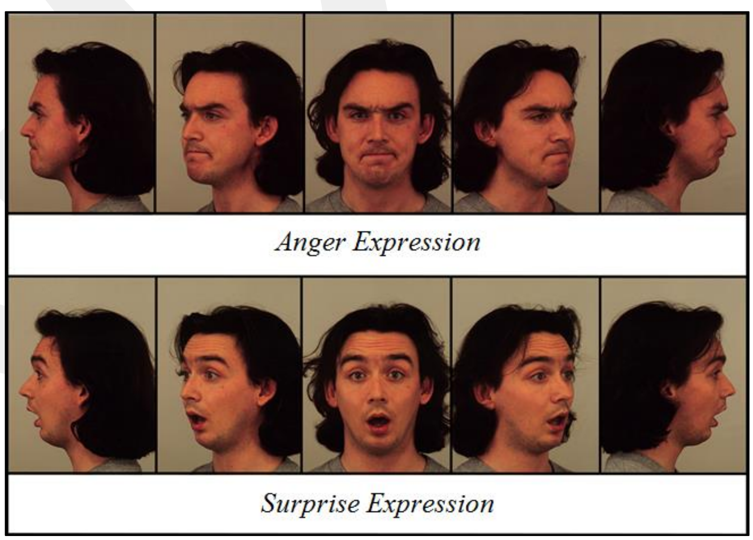

Figure 3-6: Sample images for Anger & Surprise expressions of KDEF Database from Five different angles

3.1.6. JAFFE Database

The Japanese Female Facial Expression Database is collected by Kamachi and Gyoba at Kyushu University, Japan [34] ,and it's for non-commercial research. This database contains 213 images, and the number of images corresponding to each of the 7 categories of expression (anger, disgust, fear, happiness, neutral, sadness and surprise). All images are taken from 10 female, and each female has different number of images in each facial expression. All images saved in TIFF format with a resolution of: 256 pixels X 256 pixels. *(Figure 3-7).*

Figure 3-7: One sample image for each basic expression of JAFFE Database

3.2. Pre-Processing

3.2.1. Face detection

Face detection is the most important first step in the proposed system in order to determine and extract the face region from background [35]. However, this step helps to ignore the regions undesirable in the image, thus decrease image size and target region appears clearly. Here face region is determined and extracted as shown in *(figure 3-8)* from image using Viola–Jones algorithm [36] [37].

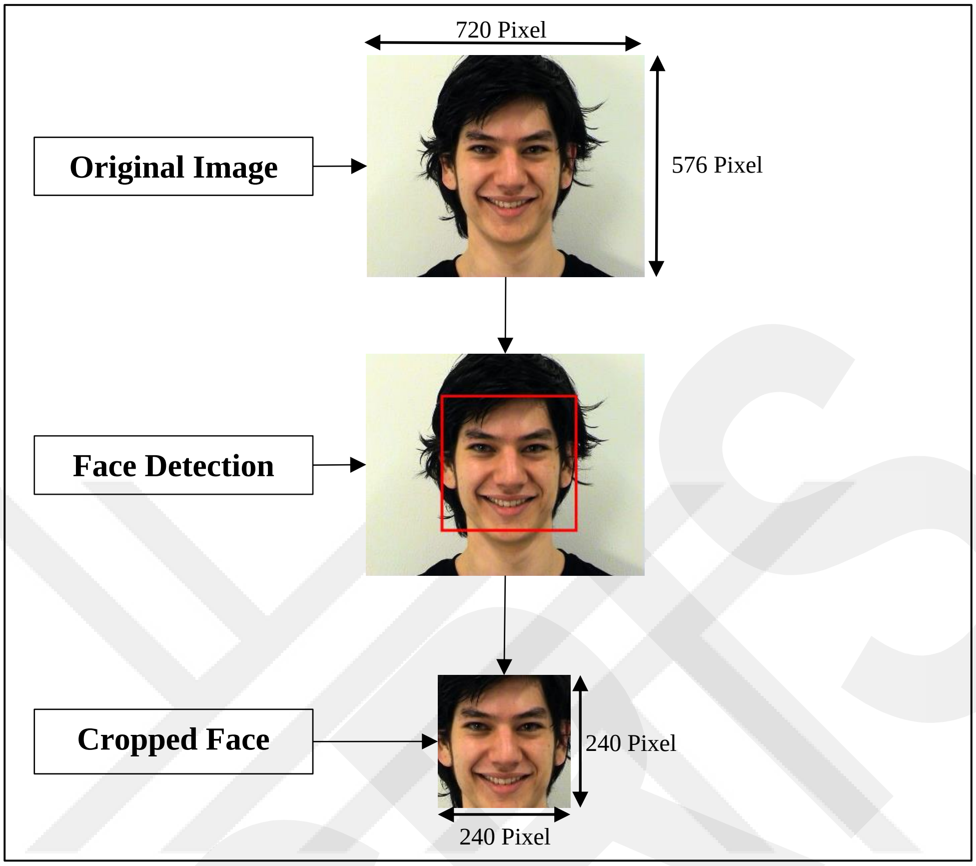

Figure 3-8: Face Detection

3.2.2. Dimensions Alignment

Dimensions alignment technique means making the images sizes bigger or smaller than original size *(figure 3-8)*. However, in this thesis various databases are used, and each database includes many images, which have different dimensions, therefore; it makes the feature extraction process impossible. Thus, applying image resizing stage become necessary to decrease the images sizes in order to unify the image sizes. As mentioned earlier face detection process ignores the regions undesirable, thus decrease images size which leads to high performance and less processing time, then image resizing technique is applied to obtain the same dimensions for all images in each database. The result of dimension alignment of all databases is 240X240 pixel except JAFFE database which is 160X160 pixel.

3.2.3. Histogram Equalization

Histogram Equalization (HE) [38] is an important technique for improving image contrast. Moreover, the main idea of HE is reducing the effect of light and unify luminosity. On other hand, HE technique gives inclusive contrast improvement, and it has been applied in various applications such as medical image processing [39]. Furthermore, it is used in the famous programs in image processing such as Adobe Photoshop and Lispix. Here, histogram equalization is used to reduce the effect of light and unify luminosity of all images in the databases, this allows to get good accuracy and high performance of the proposed system.

(Figure 3-9), shows the histogram for facial expression image before and after Histogram Equalization. Moreover, it can be observed how the Histogram Equalization cleaned the image from undesirable artifacts and improves the image's contrast. In addition, it tries to redistribute the intensity value equally on the image because an intensity value can be affected by faraway pixels.

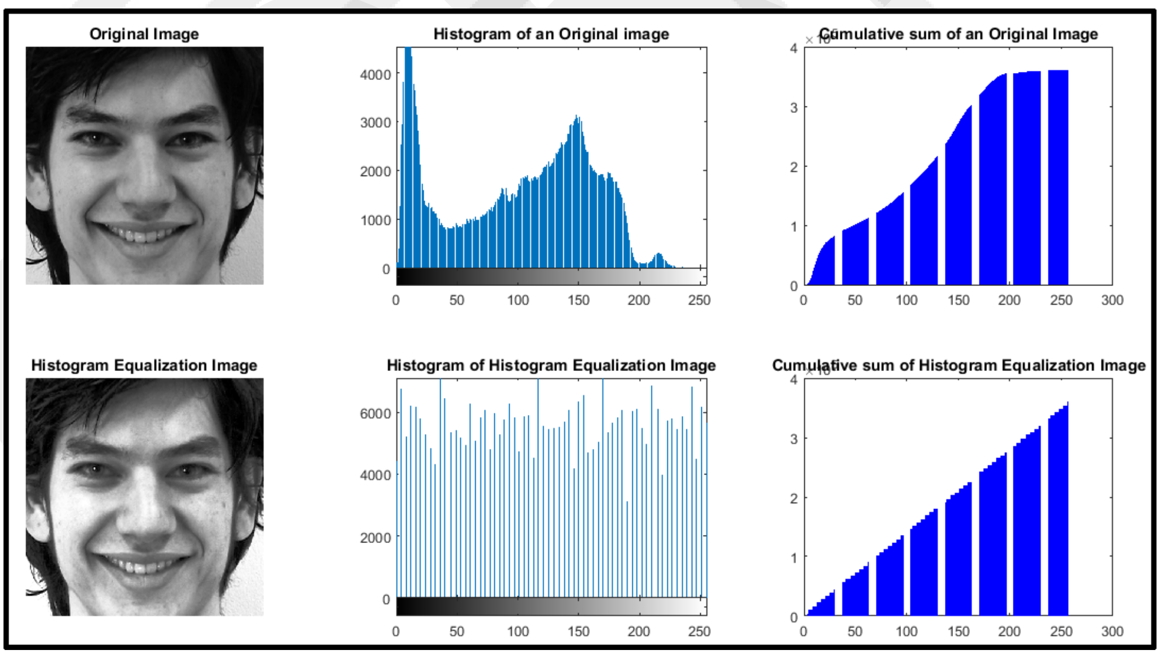

Figure 3-9: Histogram Equalization (HE)

3.3. Feature Extraction

In order to extract features from facial image, two approaches (HOG & LBP) are applied. These approaches are different from each other in a method of extracting features, and the method of collecting these features. These approaches will be explained in details.

3.3.1. Histograms of Oriented Gradients (HOG)

Histograms of Oriented Gradients is introduced by Navneet Dalal and Bill Triggs in 2005 [40], it is a popular technique to extract a dense feature from all locations in the image. It has achieved high performance in computer vision through finding radical solutions for a variety of problems related to the object detections, extracting the features from these objects, and achieving the recognition [41]. Furthermore, it succeeded in object identification from noisy background without using any segmentation techniques [42].

HOG tries to capture the pattern by capturing information about gradients in this pattern by dividing the image into small cells (usually 8x8 pixels). Each cell has a several of gradient orientation directions, and each pixel in this cell is voting for that direction with a vote commensurate to the gradient size for that pixel, then it normalizes the group of cells (block) histograms, which represent a one-dimensional array of histograms called the descriptor *(Figure 3-10).*

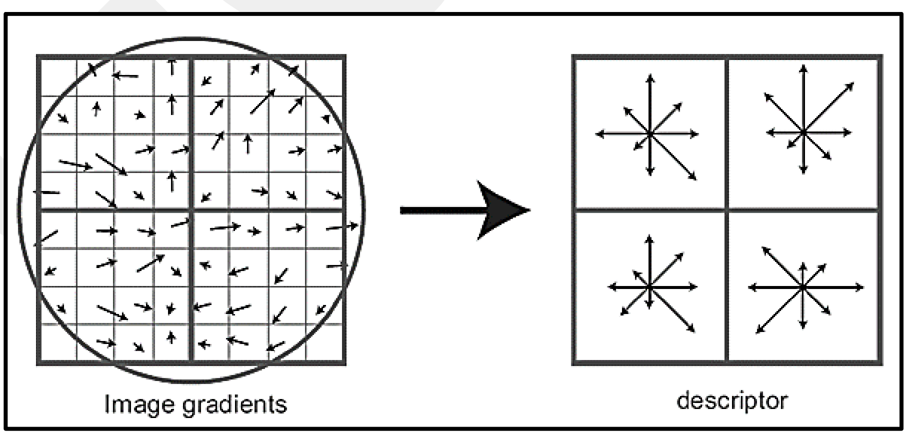

Figure 3-10: Image Gradients and Orientation Histogram [43].

(Figure 3-11) shows how the HOG approach captures the facial image and divides the image to blocks, each block is divided in to cells based on the given value for cell size. In here, various values for cell size are allocated *(figure 3-11)*. Beginning of Cell size= 4, 8, 16, 32 and 64 pixels, each value reflects the amount of information that has been encoded. For instance, cell size 4x4 shows a lot of information that are encoded in the facial image, but increases the dimensionality of HOG feature vector, while cell size 64x64 shows less amount of information, which are encoded in the facial image.

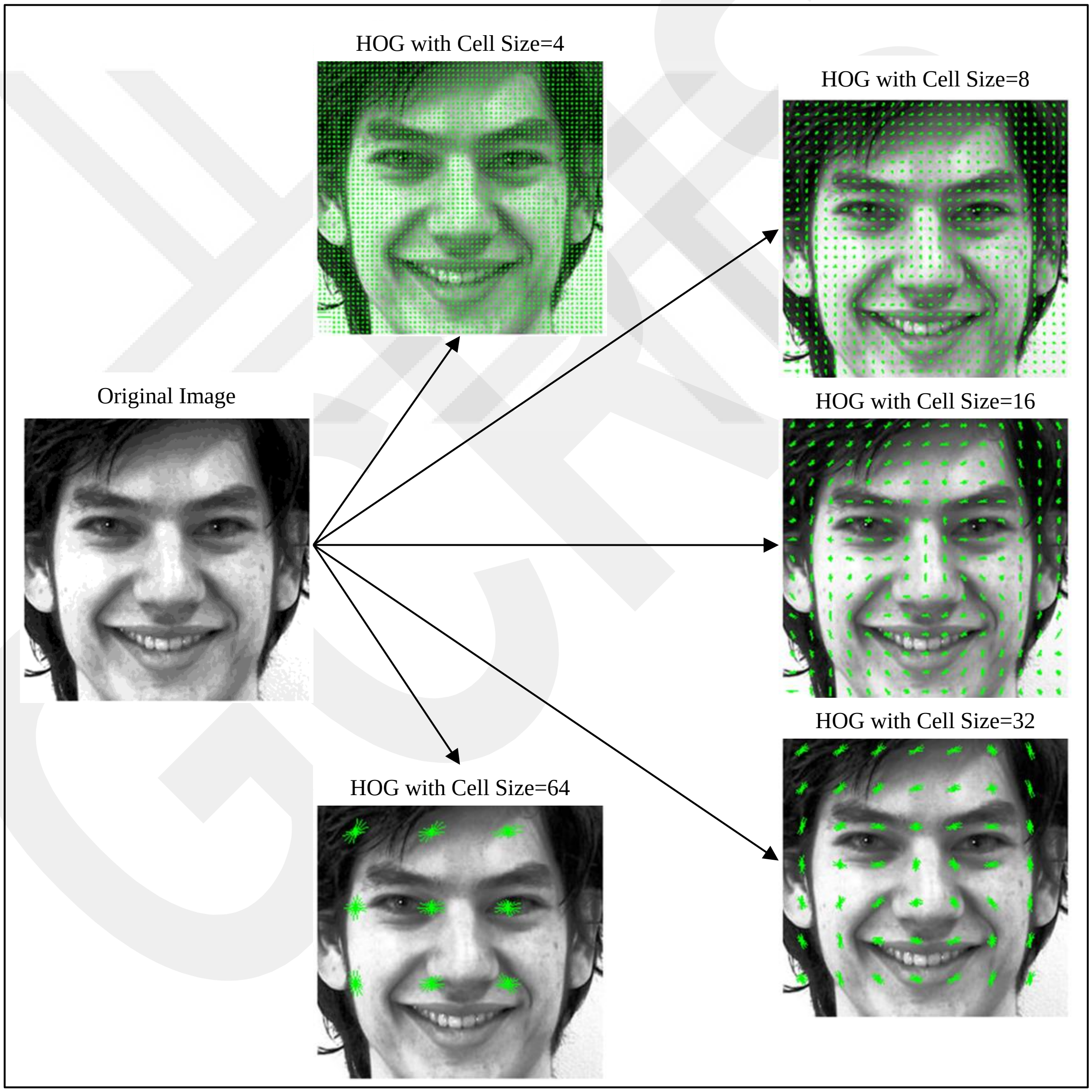

Figure 3-11: HOG with Different Cell Size

3.3.2. Local binary pattern (LBP)

Local Binary Pattern (LBP) was first introduced in 1990 as the texture spectrum model [44] [45]. It has become a powerful approach in feature extraction [46]. Moreover, it proved to achieve high performance on some databases, Especially when it is combined with HOG approach [47]. In addition, it presented as a complementary measure for image contrast, and it is developed in order to capture the pattern and extract features from it by comparing each pixel in the image with its neighbors [48]. In other words, LBP is considered a suitable technique for applications that require feature extraction. Due to its computational simplicity and distinguishing power, it has become a popular approach in many applications such as biomedical image analysis, image retrieval, visual inspection, motion analysis, and remote sensing [49].

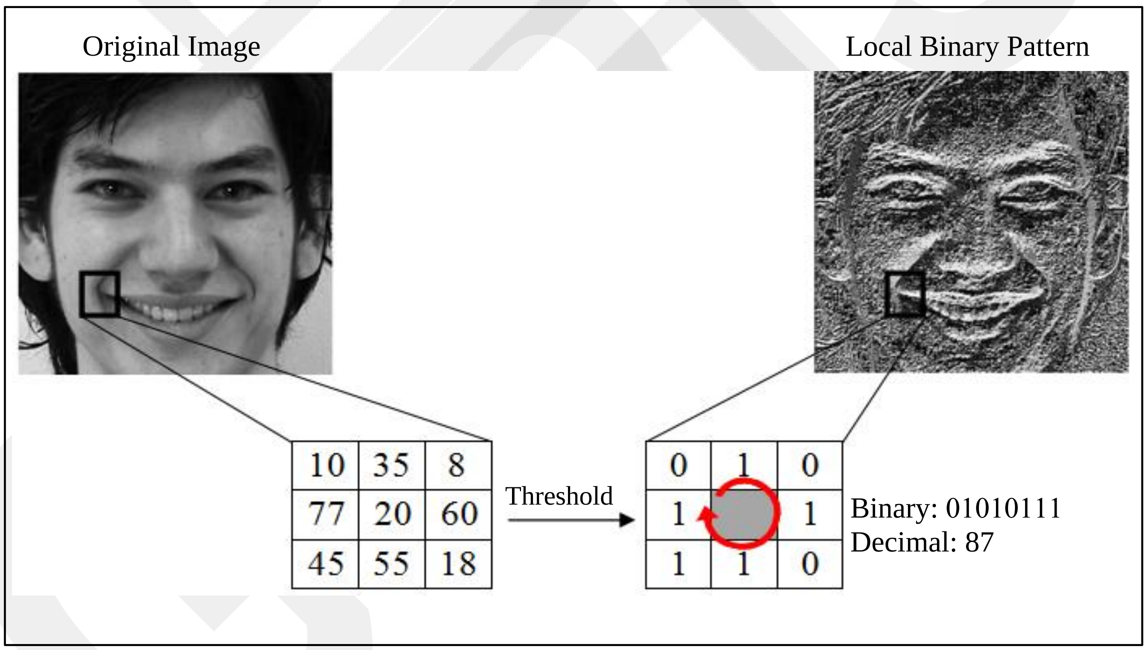

Figure 3-12: The basic LBP operator

LBP approach divides the input image to several blocks. For instance, in *(figure 3-12),* the image is divided into 80 blocks, and then determine the cell size of the neighborhood (e.g. 3x3). The center pixel will be compared to all neighbor points. In other words, the center pixel will be encoded by its density value based on the relationship between the center pixel and its neighbor points. After then threshold all values using central pixel, which is 20 in above figure, each value bigger than 20 is equals 1 and each value smaller than 20 is equals 0. All binary values in each block are arranged as shown by the red arrow, then an 8-bit binary number obtained, which is (01010111) in this example. This number is converted to decimal number to become $(01010111)^2 = (87)^{10}$. Finally, all obtained decimal numbers have been represented in one-dimensional matrix.

3.4. Classification

In this stage, two different classifiers are used to classify the features, which are extracted using previous approaches. Moreover, each classifier is trained by several randomly images, then tested by another image. These classifiers are K-Nearest Neighbors (KNN), and Support Vector Machine (SVM). Now, we will present in details these algorithms.

3.4.1. K-Nearest Neighbors (KNN)

KNN is one of the most important classifiers that are used in machine learning to predict the class of a test sample. The K-NN classifier is a nonparametric technique that is used to classify unknown objects by finding its closest neighbors [50]. Moreover, KNN is widely used in pattern recognition for classification [51], because it is an efficient and simple method in Pattern recognition. KNN classifier *(figure 3-13)* makes predictions by computing the distance between training samples and test sample, then collects the training samples that are closest to the test sample. After then it computes the average of these samples. However, the KNN classifier performance varies depending on the value of ***K***.

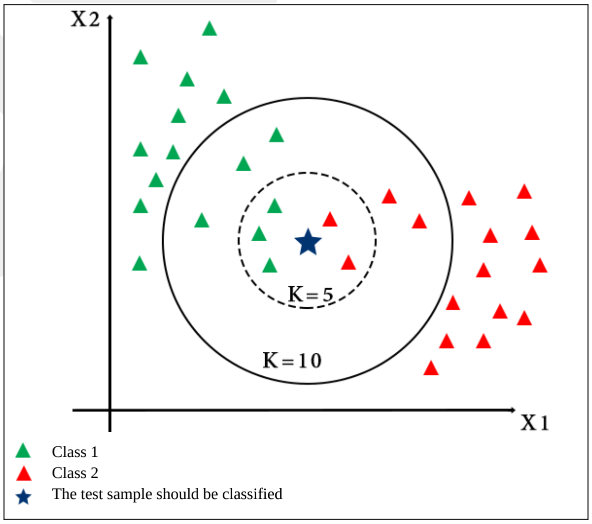

Figure 3-13: K-Nearest Neighbors

In *(Figure 3-13),* there are two groups, green triangles represent the first Class and the red triangles represent the second Class, and these Classes are represented on a feature space. For example, the data that is represented by a blue star is a new data and we want to add it to one of these classes, this step is called the classification. In this figure, when ***K***=5, there are two red triangles and three green triangles, the new data (blue star) is the closest to the red triangle neighbor therefore; KNN will classify the closest 5 training objects to the test object (blue star), and then calculates the average of them. In this example, the classification process relies on the nearest neighbor, therefore; this technique is called the Nearest Neighbor. Moreover, when ***K***=10, there are four red triangles and six green triangles, thus, in this case: KNN will classify the closest 10 training objects to the test object (blue star), and then calculates the average of them.

3.4.2. Support Vector Machine (SVM)

SVM is a machine learning algorithm used for classification and regression analysis. The current SVM standard was proposed by Vapnik and Cortes in 1993, and released in 1995 [52]. Moreover, it is considered one of important hyper-plane classification techniques that depends on results from statistical learning theory in order to ensure high performance. In addition, SVM achieves a better classification even if the available training data is simple amount, making it specially suitable for classification [8] [23] .

SVM algorithm is characterized by many advantages, that make it one of the most important classifiers in computer vision such as; images classification can be implemented by SVM. Results of experiments shown that the SVM achieved higher performance than traditional techniques in images classification fields. Additionally, SVM is able to recognize on the characters which have written by hand [53]. Moreover, SVM is used widely in various biological sciences, and proved its efficiency.

In SVM mechanism *(figure 3-14)*, the closest point [54] between the two classes of data in training set is determined, which is called "Optimal Separating Hyper-plane". By increasing the space between these classes, SVM can capture more objects from these classes, which are found in a hyper-plane. Moreover, SVM is able to reduce both structural and empirical risk that leads to reduce the number of predictable errors even though the samples in the training set are a few. On other hand, SVM classifier is able to train both two class classifier, and multi-class classifier [55].

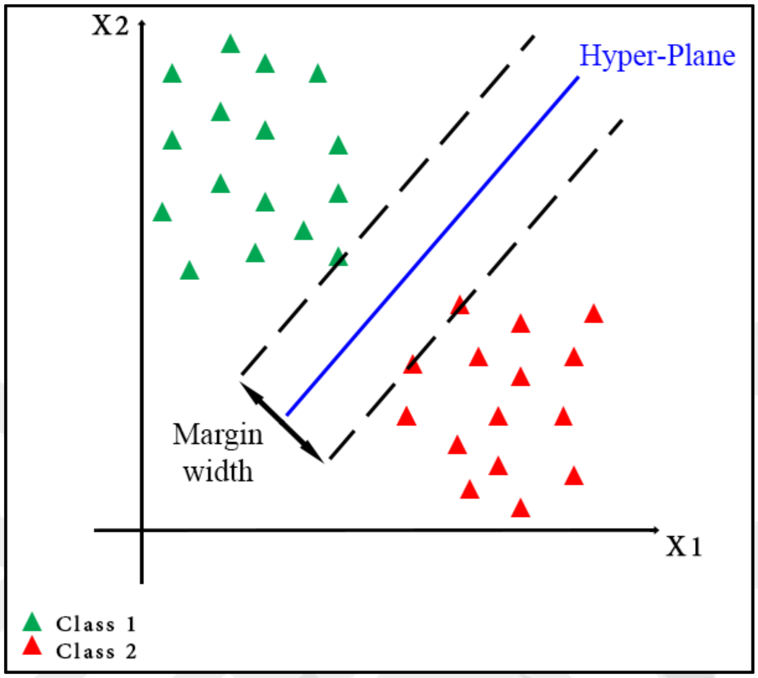

Figure 3-14: SVM mechanism

In *Figure 3-14*, Class1 is represented by the green triangles, which has 16 objects and Class2 is represented by the red triangles, which has 15 objects. Here, Class1 is considered with the green triangles features and class2 is considered with the red triangles features. SVM performs classification by finding the hyper-plane that maximizes the margin between the two classes, it captures two objects from class2 and one object from class1, then draws the widest channel between the two classes, and then SVM finds the closest two points from the two classes that support or define the best separating line which is in the middle between the two points. When the margin of SVM is increased, more objects will be captured by SVM from both Classes.

3.4.2.1. Multi-Class SVM problem

Since many classes are used in this thesis, binary SVM classifier is not suitable for our proposed system. Therefore, multi class SVM is applied to solve the problem. On other hand, many approaches are used to solve multi classes problem such as One-vs-One approach [56], One-vs-All approach [57], and Weston and Watkins' method [58]. In here, we used One-vs.-All approach to classify the classes (expressions) in our system. However, the databases which are used in this thesis contain different number of expressions, there are 10 expressions in ADFES database, 8 expressions in TFEID database, and 7 expressions in MUG, KDEF, WSEFEP, and JAFFE databases.

The main idea of One-vs-All approach is using each class against all other classes. In other words, using Class_1 vs. not Class_1, Class_2 vs. not Class_2; Class_3 vs. not Class_3;… until Class_10 vs. not Class_10, and then class-1 represents the positive objects and all other classes represent the negative objects. Then the algorithm continues applying this on all classes to the last class. Finally, the classifier chooses the suitable class that is related to the tested sample.

3.5. Evaluation

After extracting the features from facial expression images by HOG & LBP approaches, and classifying these features by KNN & SVM algorithms, the performance of these classifiers will be evaluated by Confusion Matrix technique that will be presented in the next section.

3.5.1. Confusion Matrix

Confusion matrix (CM) is a specific table that shows the performance of the classifier. In confusion matrix each column represents the situations in a predicted class and each row represents the situations in actual class [59]. Confusion matrix is also called an error matrix [60]. On other hand, confusion matrix shows the correct and incorrect results of the classifier. The classifiers are evaluated using the mentioned details in confusion matrix is shown in *(figure 3-15)*.

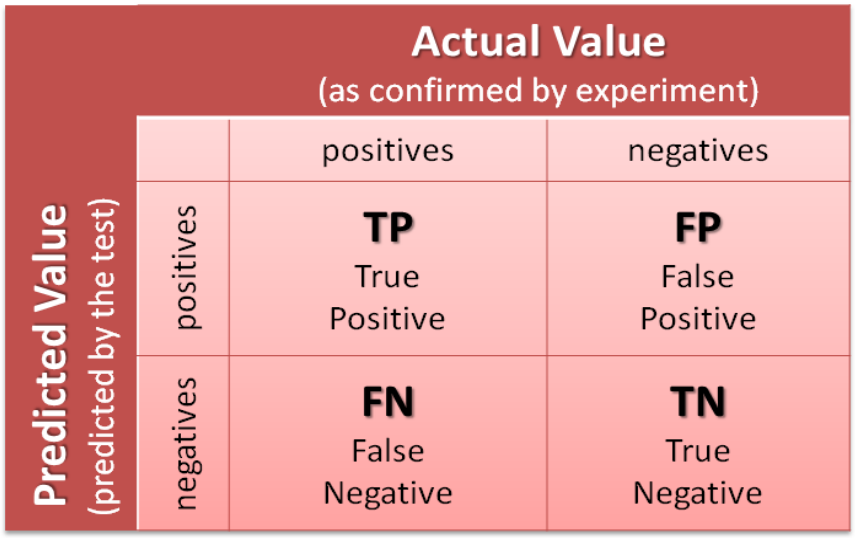

Figure 3-15: Confusion matrix [61]

In other words:

- True Positive is the number of true or correct predictions that an example is positive.
- False Negative is the number of false or incorrect predictions that an example is negative.
- False Positive is the number of false or incorrect predictions that an example is positive.
- True Negative is the number of true or correct predictions that an example is negative.

3.6. Summary

In this chapter, a full detail about pre-processing steps of face detection, Dimensions Alignment, and Histogram Equalization are presented, also the techniques that are used in each step. Moreover, two of the main approaches used in feature extraction which are HOG and LBP, then detailed explanation of the classifiers used to classify the extracted features in this thesis are introduced, which are KNN and SVM classifiers. Finally, the performance of these classifiers is evaluated using confusion matrix technique. Also, a general information about the databases that are used in our experiments are given at the beginning of this chapter.

CHAPTER 4

4. IMPLEMETATION AND RESULTS

In this thesis, different experiments are implemented using various techniques. However, two different algorithms are used to extract the features from facial image. These algorithms are Histogram Oriented gradients (HOG), and Local Binary Pattern (LBP). Moreover, these algorithms differ of each other in the method of extracting features and in the method of calculating these features. On other hand, the extracted features are classified using two different classifiers, Support Vector Machine (SVM), and K-Nearest Neighbors, also these classifiers are differing of each other in the method of classification of these features. After that, the evaluation of the performance of these classifiers are carried out using confusion matrix technique. Our experiments are applied on several databases that differ in type and number of images. In here, the experiments of the proposed system, and the obtained results will be illustrated.

In the pre-processing stage, the face detection technique is applied in order to determine only face region and to ignore the unwanted parts. This helps to disregard the useless information. Therefore; decrease the implementation time in feature extraction stage and classification stage. Moreover, dimension alignment technique helps to increase or decrease the image sizes as needed. On other hand, histogram equalization technique is used to control the level of illumination in the image and describes the distribution of image's density value.

Briefly, the databases used in our experiments can be introduced as:

- ADFES database contains 215 images for 10 expressions, these images are collected from 22 persons (10 female and 12 male), their ages from 13 to 25 years.
- TFEID database includes 336 images for 8 expressions, these images are collected from 40 persons (20 female and 20 male).
- WSEFEP database consists of 210 images for 7 expressions, these images are collected from 30 persons (16 female and 14 male).

- MUG database contains 1780 images for 7 expressions, these images are collected from 52 persons (22 female and 30 male), and their ages from 20 to 35 years.
- KDEF database consists of 980 images for 7 expressions, these images are collected from 70 persons (35 female and 35 male), and their ages from 20 to 30 years.
- JAFFE database includes 213 images for 7 expressions, these images are collected from 10 Female.

In our experiments, the testing set contains only one person and the training set contains all other persons in each turn, which means this operation repeated (N-1) times, where N is number of persons in each database. Besides, the experiments are divided based on the proposed methods. Moreover, each method is applied on six databases, therefore; all results in each method are documented separately based on the used databases.

Additionally, the "Cell Size" means the number of shape information that will be encoded in specific dimensions of the extract features algorithms. For instance, Cell Size of [8 8] means that there are a lot of shape information that will be encoded, while Cell Size of [64 64] means that there is less information that will be encoded.

4.1. Results based on Histograms of Oriented Gradients (HOG) Algorithm.

In this section, HOG is used to extract the features from facial image. On other hand, KNN and SVM classifiers are used to classify these features. In here, the experiments are applied on six different databases, different cell sizes are used in each database to show the cell sizes effectiveness on classifier performance.

4.1.1. ADFES Database Results

In ADFES database, HOG+KNN method is applied with different cell sizes *(Table 4-1)*, when cell size=8, the best prediction achieved was 95.45% with Surprised expression, and for the cell size=16, the achieved prediction was 95.45% with Pride expression, moreover; for the cell size=32, the attained prediction was 95.45% with Happy expression, and when the cell size=64, a prediction rate of 85.71% is achieved with Disgusted expression. On the other hand, the results obtained by HOG+SVM method are presented in *(Table 4-2)*. When cell size =8, and cell size=16; the achieved prediction accuracy was the best and reached to 100% with Surprised expression, moreover; for the cell size=32, the achieved accuracy was 95.45% with Sad expression, and for the cell size=64 was 90.91% with Disgusted expression.

However, HOG+KNN method achieved overall accuracy 86.05% by using cell size=32, while HOG+SVM method achieved overall accuracy 91.16% by using cell size=32.

Expressions / Cell Size	Angry	Contempt	Disgusted	Embarrass	Fear	Happy	Neutral	Pride	Sad	Surprised	**Overall Accuracy**
8	86.36%	85.71%	85.71%	76.19%	86.36%	63.64%	81.82%	85%	90.91%	95.45%	83.72%
16	90.91%	85.71%	86.36%	85.71%	68.18%	95%	81.82%	95.45%	72.73%	80.95%	84.19%
32	63.64%	95.24%	90.91%	80.95%	81.82%	95.45%	81.82%	95%	81.82%	95.24%	86.05%
64	77.27%	77.27%	85.71%	66.76%	72.73%	63.64%	81.82%	65%	72.73%	61.90%	72.56%

Table 4-1: HOG + KNN Results on ADFES Database

Expressions / Cell Size	Angry	Contempt	Disgusted	Embarrass	Fear	Happy	Neutral	Pride	Sad	Surprised	**Overall Accuracy**
8	86.36%	95.24%	86.36%	98%	86.36%	81.82%	63.64%	90%	72.73	100%	86.05%
16	86.36%	85.71%	90.91%	76.19%	81.82%	90.91%	81.82%	95%	90.91%	100%	87.91%
32	90.91%	95.24%	90.91%	95.24%	86.36%	90.91%	90.91%	85%	95.45%	90.48%	91.16%
64	81.82%	66.67%	90.91%	61.90%	77.27%	86.36%	77.27%	75%	59.09%	85.71%	76.28%

Table 4-2: HOG + SVM Results on ADFES Database

4.1.2. TFEID Database Results

In this database, HOG+KNN method is applied with different cell sizes *(Table 4-3)*, when cell size=8, the best predicted accuracy was 100% with Happy expression, and the same accuracy attained for the cell size=16, and the cell size =32 with Happy expression, on the other hand, an accuracy of only 97.50% achieved when the cell size=64 with Happy expression. Moreover, the results obtained by HOG+SVM method are presented in *(Table 4-5)*. When cell size=8, the best prediction accuracy achieved was 100% with Surprised expression, and the same accuracy attained for the cell size=16. On the other hand, for cell size=32, the best prediction achieved accuracy was 100% with Surprised expression, and an accuracy of only 86.11% achieved when the cell size=64 with Surprised expression.

Here, HOG+KNN method achieved overall accuracy 87.20% by using cell size=32, while HOG+SVM method achieved overall accuracy 96.13% by using cell size=32.

Cell Size \ Expressions	Angry	Contempt	Disgusted	Fear	Happy	Neutral	Sad	Surprised	**Overall Accuracy**
8	94.12%	61.76%	82.50%	72.50%	100%	94.87%	66.67%	94.44%	81.25%
16	94.12%	85.29%	82.50%	70%	100%	87.18%	71.79%	97.22%	85.71%
32	91.18%	88.24%	80%	75%	100%	89.74%	76.92%	97.22%	87.20%
64	76.47%	64.71%	62.50%	52.50%	97.50%	66.67%	53.85%	94.44%	70.24%

Table 4-3: HOG + KNN Results on TFEID Database

As shown in *(Table 4-3)*, the overall accuracy value of 87.20% is obtained by using HOG+KNN method when cell size=32. In addition, KNN classifier is evaluated using Confusion Matrix (CM), which provides details and visualization about predicted and actual classes. The Confusion Matrix of the best results are shown in *Table 4-4*.

Predicted Classes \ Actual Classes	Angry	Contempt	Disgusted	Fear	Happy	Neutral	Sad	Surprised	**Accuracy**
Angry	31	1		2					91.18%
Contempt		60		4		2	2		88.24%
Disgusted	2		32	4		2			80%
Fear	2			30		2	6		75%
Happy					40				100%
Neutral	1	2	1			35			89.74%
Sad	2	2		5			30		76.92%
Surprised					1			35	97.22%
Total No of Images	34	68	40	40	40	39	39	36	

Table 4-4: CM Evaluation of HOG + KNN for TFEID DB using Cell Size=32

Cell Size \ Expressions	Angry	Contempt	Disgusted	Fear	Happy	Neutral	Sad	Surprised	**Overall Accuracy**
8	88.24%	72.06%	90%	80%	97.50%	84.62%	69.23%	100%	83.93%
16	88.24%	82.35%	95%	85%	95%	79.49%	74.36%	100%	86.90%
32	97.06%	95.59%	95%	95%	97.50%	94.87%	94.87%	100%	96.13%
64	76.47%	79.41%	80%	77.50%	85%	74.36%	61.54%	86.11%	77.68%

Table 4-5: HOG + SVM Results on TFEID Database

As shown in *(Table 4-5)*, the overall accuracy value of 96.13% is obtained by using HOG+SVM method when cell size=32. In addition, SVM classifier is evaluated using Confusion Matrix (CM), which provides details and visualization about predicted and actual classes. The Confusion Matrix of the best results are shown in *Table 4-6*.

Predicted Classes \ Actual Classes	Angry	Contempt	Disgusted	Fear	Happy	Neutral	Sad	Surprised	**Accuracy**
Angry	33						1		97.06%
Contempt		65		3					95.59%
Disgusted			38	2					95%
Fear				38			2		95%
Happy					39			1	97.50%
Neutral	1	1				37			94.87%
Sad				2			37		94.87%
Surprised								36	100%
Total No of Images	34	68	40	40	40	39	39	36	

Table 4-6: CM Evaluation of HOG + SVM for TFEID DB using Cell Size=32

4.1.3. WSEFEP Database Results

In this database, HOG+KNN method is applied with different cell sizes *(Table 4-7)*, when cell size=8, the achieved prediction reached to 100% with Happy expression, and it was the same in the case of cell size=16, and 32 with Happy expression, but when the cell size=64, the attained prediction rate was 90% with Surprised expression. Moreover, the results obtained by HOG+SVM method are presented in *(Table 4-8)*, for the cell size =8 the prediction accuracy was the best, and reached to 100% with Happy expression, but for the cell size=16, the accuracy achieved was only 96.67% with both Disgusted and Happy expressions. Moreover; for the cell size=32, the achieved accuracy was 100% with Surprised expressions, and for the cell size=64 the obtained accuracy is only 86.67% with Disgusted expression.

However, HOG+KNN method achieved overall accuracy 82.38% by using cell size=32, while HOG+SVM method achieved overall accuracy 89.05% by using cell size=32.

Expressions / Cell Size	Angry	Disgusted	Fear	Happy	Neutral	Sad	Surprised	**Overall Accuracy**
8	70%	70%	53.33%	100%	86.67%	60%	73.33	73.33%
16	66.67%	73.33%	63.33%	100%	90%	50%	86.67%	75.71%
32	83.33%	76.67%	73.33%	100%	93.33%	56.67%	93.33%	82.38%
64	66.67%	60%	60%	86.67%	83.33%	56.67%	90%	71.90%

Table 4-7: HOG + KNN Results on WSEFEP Database

Expressions / Cell Size	Angry	Disgusted	Fear	Happy	Neutral	Sad	Surprised	**Overall Accuracy**
8	86.67%	86.67%	70%	100%	60%	63.33%	83.33%	78.57%
16	90%	96.67%	76.67%	96.67%	73.33%	76.67%	83.33%	84.76%
32	90%	96.67%	80%	100%	76.67%	80%	100%	89.05%
64	83.33%	86.67%	66.67%	70%	56.67%	73.33%	76.67%	73.33%

Table 4-8: HOG + SVM Results on WSEFEP Database

4.1.4. MUG Database Results

In this database, HOG+KNN method is applied with different cell sizes *(Table 4-9)*, when cell size=8, the prediction accuracy achieved was 91.92% with Happy expression, and when cell size=16, the achieved accuracy was 95.38% with Happy expression, moreover; when cell size=32, the attained prediction rate was about 94.23% with Happy expression, and for the cell size=64, the accuracy was 96.54% with Surprised expression. Moreover, *(Table 4-10)* illustrates the obtained results when applying HOG+SVM method on MUG database, where the best accuracy achieved was 95% with Happy expression when the cell size=8, on the other hand; when cell size=16 the accuracy achieved to 93.85% with Happy expression, moreover; for the cell size=32, the achieved accuracy was 92.31% with Surprised expressions, and for the cell size=64 was 88.16% with Sad expression.

Here, HOG+KNN method achieved overall accuracy 76.35% by using cell size=32, while HOG+SVM method achieved overall accuracy 85.51% by using cell size=32.

Expressions / Cell Size	Angry	Disgusted	Fear	Happy	Neutral	Sad	Surprised	**Overall Accuracy**
8	54.23%	68.24%	76.25%	91.92%	59.23%	81.63%	69.62%	71.46%
16	73.85%	71.76%	87.50%	95.38%	66.15%	62.45%	65.77%	74.66%
32	71.54%	74.51%	78.33%	94.23%	73.08%	69.80%	72.69%	76.35%
64	57.31%	68.63%	57.92%	95%	64.62%	60%	96.54%	71.69%

Table 4-9: HOG + KNN Results on MUG Database

Expressions / Cell Size	Angry	Disgusted	Fear	Happy	Neutral	Sad	Surprised	**Overall Accuracy**
8	82.31%	87.84%	49.58%	95%	65.38%	68.98%	92.31%	77.70%
16	84.23%	83.53%	72.50%	93.85%	79.23%	68.16%	79.23%	80.28%
32	86.54%	78.43%	90.42%	97.69%	71.15%	82.04%	92.31%	85.51%
64	82.69%	70.98%	70.42%	84.62%	66.15%	88.16%	79.23%	77.47%

Table 4-10: HOG + SVM Results on MUG Database

4.1.5. KDEF Database Results

In KDEF database, HOG+KNN method is applied with different cell sizes *(Table 4-11)*, the best prediction achieved was 92.14% with Happy expression when the cell size =8, and when the cell size=16, the achieved accuracy was 94.29% with Happy expression, additionally; for the cell size=32, the attained prediction rate was 93.57% with Happy expression, and 87.14% for the cell size=64 with the same expression. Moreover, the results obtained by HOG+SVM method are presented in *(Table 4-12)*, when cell size =8 the achieved results gave the best prediction and reached to 91.43% with Surprised expression, and when cell size=16, the accuracy achieved only 90% with Happy expression. Moreover, for the cell size=32, the achieved accuracy was 92.86% with Angry expression, and for the cell size=64 was 86.67% with Disgusted expression.

However, HOG+KNN method achieved overall accuracy 76.94% by using cell size=32, while HOG+SVM method achieved overall accuracy 85% by using cell size=32.

Expressions / Cell Size	Angry	Disgusted	Fear	Happy	Neutral	Sad	Surprised	**Overall Accuracy**
8	57.86%	53.57%	60.71%	92.14%	85%	60.71%	80.71%	70.10%
16	75.71%	65%	55%	94.29%	82.14%	57.14%	77.14%	72.35%
32	60.71%	68.57%	72.14%	93.57%	86.43%	64.29%	92.86%	76.94%
64	64.29%	67.14%	72.86%	87.14%	65%	67.86%	73.57%	71.12%

Table 4-11: HOG + KNN Results on KDEF Database

Expressions / Cell Size	Angry	Disgusted	Fear	Happy	Neutral	Sad	Surprised	**Overall Accuracy**
8	76.43%	75.71%	57.14%	90.71%	80.71%	60.71%	91.43%	76.12%
16	78.57%	78.57%	56.43%	90%	82.14%	66.43%	87.14%	77.04%
32	92.86%	78.57%	82.14%	90.71%	82.86%	78.57%	89.29%	85%
64	83.33%	86.67%	66.67%	70%	56.67%	73.33%	76.67%	73.33%

Table 4-12: HOG + SVM Results on KDEF Database

4.1.6. JAFFE Database Results

In this database, HOG+KNN method is applied with different cell sizes *(Table 4-13)*, when cell size =8 the achieved accuracy was 93.33% with Surprised expression, and when cell size=16, the accuracy achieved to 96.67% with Surprised expression, moreover; for the cell size=32, the achieved accuracy was 100% with Surprised expression, and for the cell size=64 was 96.67% with Surprised expression. Moreover, the results obtained by HOG+SVM method are presented in *(Table 4-14)*, the best prediction result reached to 90.32% with Sad expression for cell size =8, and when cell size=16, the accuracy achieved to 90.32% with Sad expression. Furthermore, for the cell size=32, the achieved accuracy was 93.55% with Sad expressions, and for the cell size=64 was only 83.87% with Happy expression.

Here, HOG+KNN method achieved overall accuracy 77.46% by using cell size=32, while HOG+SVM method achieved overall accuracy 82.16% by using cell size=32.

Expressions / Cell Size	Angry	Disgusted	Fear	Happy	Neutral	Sad	Surprised	**Overall Accuracy**
8	60%	68.97%	59.38%	74.19%	60%	80.65%	93.33%	70.89%
16	73.33%	51.72%	56.25%	77.42%	66.67%	87.10%	96.67%	72.77%
32	76.67%	55.17%	62.50%	80.65%	73.33%	93.55%	100%	77.46%
64	66.67%	48.28%	46.88%	74.19%	66.67%	83.87%	96.67%	69.01%

Table 4-13: HOG + KNN Results on JAFFE Database

Expressions / Cell Size	Angry	Disgusted	Fear	Happy	Neutral	Sad	Surprised	**Overall Accuracy**
8	66.67%	79.31%	65.63%	80.65%	66.67%	90.32%	86.67%	76.53%
16	70%	82.76%	71.88%	87.10%	66.67%	90.32%	90%	79.81%
32	70%	86.21%	75%	90.32%	66.67%	93.55%	93.33%	82.16%
64	66.67%	79.31%	62.50%	83.87%	60%	80.65%	76.67%	72.77%

Table 4-14: HOG + SVM Results on JAFFE Database

However, Overall Accuracies of HOG Approach when cell size=32 are presented in *(Table 4-15)*.

Method	Database	Overall Accuracy
HOG + KNN	ADFES	86.05%
HOG + SVM		91.16%
HOG + KNN	TFEID	87.20%
HOG + SVM		96.13%
HOG + KNN	WSEFEP	82.38%
HOG + SVM		89.05%
HOG + KNN	MUG	76.35%
HOG + SVM		85.51%
HOG + KNN	KDEF	76.94%
HOG + SVM		85%
HOG + KNN	JAFFE	77.46%
HOG + SVM		82.16%

Table 4-15: Overall Accuracies of HOG Approach when cell size=32.

4.2. Results based on Local Binary Pattern (LBP) Algorithm.

In this method, LBP technique used to extract the features from facial image. KNN and SVM classifiers are used to classify these features. In here, the experiments are applied on six different databases, and different cell sizes are used in each database to show the cell sizes effectiveness on classifier performance.

4.2.1. ADFES Database Results

In ADFES database, LBP+KNN method is also applied with different cell sizes *(Table 4-16)*, when cell size=8, the best prediction achieved was 90.48% with Surprised expression, and for the cell size=16, the prediction accuracy reached to 90.91% with Sad expression. Moreover, for the cell size=32, the attained prediction was 95.45% with both Angry and Happy expressions, and when the cell size=64, a prediction rate of 80.95% achieved with Surprised expression. Moreover, the results obtained by LBP+SVM method are presented in *(Table 4-18)*, the best prediction achieved was 100% with Surprised expression for cell size=8 and cell size=16, the achieved prediction was 100% with Surprised expression. Moreover, for the cell size=32, the attained prediction was 95.45% with Angry expression, and when the cell size=64, a prediction rate of 86.36% achieved with Disgusted expression.

However, LBP+KNN method achieved overall accuracy 87.44% by using cell size=32, while LBP+SVM method achieved overall accuracy 90.23% by using cell size=32.

Expressions / Cell Size	Angry	Contempt	Disgusted	Embarrass	Fear	Happy	Neutral	Pride	Sad	Surprised	**Overall Accuracy**
8	86.36%	66.67%	86.36%	76.19%	68.18%	81.82%	77.27%	85%	81.82%	90.48%	80%
16	81.82%	80.95%	77.27%	85.71%	72.73%	81.82%	77.27%	90%	90.91%	90.48%	82.79%
32	95.45%	76.19%	90.91%	76.19%	72.73%	95.45%	90.91%	95%	86.36%	95.24%	87.44%
64	72.73%	71.43%	68.18%	66.67%	72.73%	77.27%	68.18%	75%	72.73%	80.95%	72.56%

Table 4-16: LBP + KNN Results on ADFES Database

As shown in *(Table 4-16)*, the overall accuracy value of 87.44% is obtained when Cell Size=32 with LBP+KNN method. In addition, KNN classifier is evaluated using Confusion Matrix (CM), which provides details and visualization about predicted and actual classes. The Confusion Matrix of the best results are shown in *Table 4-17*.

		Actual Classes										
		Angry	Contempt	Disgusted	Embarrass	Fear	Happy	Neutral	Pride	Sad	Surprised	**Accuracy**
Predicted Classes	Angry	21								1		95.45%
	Contempt		16	2		3						76.19%
	Disgusted			20		1				1		90.91%
	Embarrass	3	1	1	16							76.19%
	Fear					16			1	2	3	72.73%
	Happy		1				21					95.45%
	Neutral							20		2		90.91%
	Pride		1						19			95%
	Sad	1				1		1		19		86.36%
	Surprised					1					20	95.24%
	Total No of Images	22	21	22	21	22	22	22	20	22	21	

Table 4-17: CM Evaluation of LBP + KNN for ADFES DB using Cell Size=32

Expressions / Cell Size	Angry	Contempt	Disgusted	Embarrass	Fear	Happy	Neutral	Pride	Sad	Surprised	**Overall Accuracy**
8	77.27%	80.95%	77.27%	90.48%	81.82%	72.73%	77.27%	90%	68.18%	100%	81.40%
16	77.27%	76.19%	86.36%	80.95%	72.73%	90.91%	68.18%	90%	90.91%	100%	83.26%
32	95.45%	80.95%	90.91%	85.71%	90.91%	90.91%	90.91%	95%	86.36%	95.24%	90.23%
64	68.18%	61.90%	86.36%	57.14%	63.64%	81.82%	63.64%	70%	50%	76.19%	67.91%

Table 4-18: LBP + SVM Results on ADFES Database

As shown in *(Table 4-18)*, the overall accuracy value of 90.23% is obtained when cell size=32 with LBP+SVM method. In addition, SVM classifier is evaluated using Confusion Matrix (CM), which provides details and visualization about predicted and actual classes. The Confusion Matrix of the best results are shown in *Table 4-19*.

		Actual Classes										
		Angry	Contempt	Disgusted	Embarrass	Fear	Happy	Neutral	Pride	Sad	Surprised	**Accuracy**
Predicted Classes	Angry	21			1							95.45%
	Contempt		17		2	1		1				80.95%
	Disgusted			20		2						90.91%
	Embarrass	2		1	18							85.71%
	Fear					20					2	90.91%
	Happy		2				20					90.91%
	Neutral				1	1		20				90.91%
	Pride		1						19			95%
	Sad	1				2				19		86.36%
	Surprised					1					20	95.24%
	Total No of Images	22	21	22	21	22	22	22	20	22	21	

Table 4-19: CM Evaluation of LBP + SVM for ADFES DB using Cell Size=32

4.2.2. TFEID Database Results

In TFEID database, LBP+KNN method *(Table 4-20)*, when cell size=8, the best prediction achieved was 94.44% with Surprised expression, and for the cell size=16, the achieved prediction was 97.22% with Surprised expression. Additionally, for the cell size=32, the attained prediction was 97.5% with Happy expressions, and for the cell size=64, a prediction rate of 100% achieved with Surprised expression. Moreover, the results obtained by LBP+SVM method are presented in *(Table 4-21)*, the best prediction accuracy reached to 100% for both cell size=8 with Happy expression, and cell size=32 with Surprised expression. On other hand, for the cell size=16, the achieved prediction was 97.50% with Happy expression, and when the cell size=64, a prediction rate of 90% achieved with Happy expression.

Here, LBP+KNN method achieved overall accuracy 85.12% by using cell size=32, while LBP+SVM method achieved overall accuracy 83.33% by using cell size=32.

Expressions / Cell Size	Angry	Contempt	Disgusted	Fear	Happy	Neutral	Sad	Surprised	**Overall Accuracy**
8	64.71%	79.41%	67.50%	70%	92.50%	84.62%	78.18%	94.44%	80.06%
16	88.24%	79.41%	80%	70%	95%	84.62%	69.23%	97.22%	82.44%
32	88.24%	80.88%	85%	75%	97.50%	84.62%	76.92%	97.22%	85.12%
64	70.59%	69.12%	77.50%	55%	92.50%	71.79%	53.85%	100%	73.21%

Table 4-20: LBP + KNN Results on TFEID Database

Expressions / Cell Size	Angry	Contempt	Disgusted	Fear	Happy	Neutral	Sad	Surprised	**Overall Accuracy**
8	82.35%	73.53%	85%	67.50%	100%	64.10%	58.97%	97.22%	77.98%
16	85.29%	73.53%	92.50%	77.50%	97.50%	76.92%	66.67%	97.22%	82.44%
32	82.35%	82.35%	87.50%	82.50%	97.50%	69.23%	66.67%	100%	83.33%
64	88.24%	73.53%	80%	67.50%	90%	53.85%	56.41%	88.89%	74.40%

Table 4-21: LBP + SVM Results on TFEID Database

4.2.3. WSEFEP Database Results

As shown in *(Table 4-22)*, when LBP+KNN method is applied on WSEFEP database, the best prediction achieved was 93.33% with Happy expression for cell size=8, and for the cell size=16, the achieved prediction was 90% with Happy expression. Also for the cell size=32, the attained prediction was 93.33% with both Angry and Happy expressions, and when the cell size=64, a prediction rate of 83.33% achieved with Angry expression. Moreover, the results obtained by LBP+SVM method are presented in *(Table 4-23)*, when cell size=8, 16 and 32 the best prediction achieved was 100% with Happy expression, furthermore; for the cell size=64, the attained prediction reached only to 93.33% with Happy expressions.

However, LBP+KNN method achieved overall accuracy 80% by using cell size=32, while LBP+SVM method achieved overall accuracy 78.10% by using cell size=32.

Cell Size \ Expressions	Angry	Disgusted	Fear	Happy	Neutral	Sad	Surprised	**Overall Accuracy**
8	86.67%	66.67%	60%	93.33%	76.67%	43.33%	50%	68.10%
16	86.67%	60%	60%	90%	80%	46.67%	66.67%	70%
32	93.33%	73.33%	70%	93.33%	86.67%	60%	83.33%	80%
64	83.33%	60%	63.33%	73.33%	73.33%	46.67%	66.67%	66.67%

Table 4-22: LBP + KNN Results on WSEFEP Database

Cell Size \ Expressions	Angry	Disgusted	Fear	Happy	Neutral	Sad	Surprised	**Overall Accuracy**
8	73.33%	60%	46.67%	100%	36.67%	53.33%	96.67%	66.67%
16	80%	86.67%	73.33%	100%	56.67%	46.67%	93.33%	76.67%
32	83.33%	90%	66.67%	100%	56.67%	56.67%	93.33%	78.10%
64	73.33%	76.67%	53.33%	93.33%	43.33%	33.33%	63.33%	62.38%

Table 4-23: LBP + SVM Results on WSEFEP Database

4.2.4. MUG Database Results

In MUG database, LBP+KNN method is applied with different cell sizes *(Table 4-24)*, when cell size=8, the best prediction achieved was 91.54% with Happy expression, and for the cell size=16, the achieved prediction was 98.46% with Happy expression, furthermore; for the cell size=32, the attained prediction was 96.54% with Surprised expression, but for the cell size=64, a prediction rate 83.85% achieved with Happy expression. Moreover, the results obtained by LBP+SVM method are presented in *(Table 4-25)*, when cell size=8, the best prediction achieved was 93.46% with Happy expression, and for the cell size=16, the achieved prediction reached to 96.15% with Happy expression. Moreover, for the cell size=32, the attained prediction was 90.77% with Happy expression, and when the cell size=64, a prediction rate of 86.92% achieved with Happy expression.

Here, LBP+KNN method achieved overall accuracy 81.91% by using cell size=32, while LBP+SVM method achieved overall accuracy 77.81% by using cell size=32.

Expressions / Cell Size	Angry	Disgusted	Fear	Happy	Neutral	Sad	Surprised	**Overall Accuracy**
8	78.85%	71.37%	65.42%	91.54%	63.85%	56.33%	75.77%	72.08%
16	75.77%	67.45%	77.92%	98.46%	74.62%	68.98%	69.23%	76.12%
32	86.15%	75.69%	67.08%	93.85%	67.69%	85.31%	96.54%	81.91%
64	83.46%	66.27%	72.92%	83.85%	63.08%	64.08%	75.38%	72.81%

Table 4-24: LBP + KNN Results on MUG Database

Expressions / Cell Size	Angry	Disgusted	Fear	Happy	Neutral	Sad	Surprised	**Overall Accuracy**
8	81.92%	83.14%	56.67%	93.46%	63.08%	65.71%	72.31%	72.47%
16	79.23%	79.22%	49.17%	96.15%	67.69%	72.65%	88.08%	76.35%
32	77.31%	72.94%	87.50%	90.77%	72.31%	60%	83.46%	77.81%
64	68.08%	72.94%	60.83%	86.92%	66.15%	76.33%	76.69%	72.08%

Table 4-25: LBP + SVM Results on MUG Database

4.2.5. KDEF Database Results

Applying the LBP+KNN method on KDEF database gives the results illustrated in *(Table 4-26)*, which achieved a prediction rate of 91.43% for the cell size=8 and 16 with Happy expression. On the other hand, the perdition accuracy reached only 88.57% for the cell size =32 with Happy expression, and 88.57% for the cell size=64 with Surprised expression. Moreover, the results obtained by LBP+SVM method are presented in *(Table 4-27)*, the best prediction achieved when cell size=8 was 96.43% with Surprised expression, and for the cell size=16, the achieved prediction was 90.71% with both Happy and Surprised expressions, furthermore; for the cell size=32, the attained prediction was 90% with Happy expression, and when the cell size=64, a prediction rate of 82.86% achieved with Happy expression.

However, LBP+KNN method achieved overall accuracy 78.67% by using cell size=32, while LBP+SVM method achieved overall accuracy 75.51% by using cell size=32.

Expressions / Cell Size	Angry	Disgusted	Fear	Happy	Neutral	Sad	Surprised	**Overall Accuracy**
8	45.71%	60%	55.71%	91.43%	85%	55%	81.43%	67.76%
16	64.29%	76.43%	60%	91.43%	73.57%	60%	75.71%	71.63%
32	85.71%	75%	70%	88.57%	72.14%	73.57%	85.71%	78.67%
64	69.29%	72.14%	63.57%	85.71%	60.71%	72.14%	87.86%	73.06%

Table 4-26: LBP + KNN Results on KDEF Database

Expressions / Cell Size	Angry	Disgusted	Fear	Happy	Neutral	Sad	Surprised	**Overall Accuracy**
8	73.57%	73.57%	35.71%	91.43%	72.14%	47.86%	96.43%	70.10%
16	77.14%	75.71%	52.86%	90.71%	79.29%	56.43%	90.71%	74.69%
32	75%	73.57%	64.29%	90%	70%	71.43%	84.29%	75.51%
64	67.86%	70%	60.71%	82.86%	75.71%	64.29%	75%	70.92%

Table 4-27: LBP + SVM Results on KDEF Database

4.2.6. JAFFE Database Results

In here, LBP+KNN method is applied with different cell sizes *(Table 4-28)*, when cell size=8, the best prediction achieved was 86.67% with Surprised expression, and for the cell size=16, the achieved prediction was 93.33% with Surprised expression, additionally; for the cell size=32, the attained prediction was 96.67% with Surprised expression, and for the cell size=64, a prediction rate of 86.67% achieved with Surprised expression. Moreover, the results obtained by LBP+SVM method are presented in *(Table 4-29)*, the best prediction achieved was 96.67%, 93.33%, 96.67% and 93.33% for the cell size=8, 16, 32 and 64 respectively with Surprised expression.

Here, LBP+KNN method achieved overall accuracy 75.12% by using cell size=32, while LBP+SVM method achieved overall accuracy 77.46% by using cell size=32.

Expressions / Cell Size	Angry	Disgusted	Fear	Happy	Neutral	Sad	Surprised	**Overall Accuracy**
8	60%	58.62%	46.88%	64.52%	60%	77.42%	86.67%	64.79%
16	63.33%	62.07%	50%	70.97%	66.67%	80.65%	93.33%	69.48%
32	73.33%	68.97%	56.25%	70.97%	73.33%	87.10%	96.67%	75.12%
64	70%	65.52%	62.50%	67.74%	60%	80.64%	86.67%	70.42%

Table 4-28: LBP + KNN Results on JAFFE Database

Expressions / Cell Size	Angry	Disgusted	Fear	Happy	Neutral	Sad	Surprised	**Overall Accuracy**
8	63.33%	62.07%	62.50%	74.19%	63.33%	83.87%	96.67%	72.30%
16	66.67%	62.07%	65.63%	77.42%	70%	90.32%	93.33%	75.12%
32	83.33%	65.52%	65.63%	70.97%	70%	90.32%	96.67%	77.46%
64	76.67%	51.72%	62.50%	61.29%	60%	83.87%	93.33%	69.95%

Table 4-29: LBP + SVM Results on JAFFE Database

However, Overall Accuracies of LBP Approach when cell size=32 are presented in *(Table 4-30)*.

Method	Database	Overall Accuracy
LBP + KNN	ADFES	87.44%
LBP + SVM		90.23%
LBP + KNN	TFEID	85.12%
LBP + SVM		83.33%
LBP + KNN	WSEFEP	80%
LBP + SVM		78.10%
LBP + KNN	MUG	81.91%
LBP + SVM		77.81%
LBP + KNN	KDEF	78.67%
LBP + SVM		75.51%
LBP + KNN	JAFFE	75.12%
LBP + SVM		77.46%

Table 4-30: Overall Accuracies of LBP Approach when cell size=32.

4.3. Summary

In order to predict the emotion of any person from his/her face image, four different methods, HOG+KNN, HOG+SVM, LBP+KNN, LBP+SVM are applied on six different databases. These experiments are applied based on ten expressions. The proposed system achieved high performance through controlling cell size. Observing the results obtained in the experiments, some important points were noticed such as the performance of HOG algorithm gave better performance than LBP algorithm. Moreover, SVM classifier achieved better performance than KNN classifier. On other hand, from the different cell sizes used in these experiments which are 8, 16, 32, and 64, the best performance achieved by using cell size=32, while the cell size=64 achieved lower performance because of the increasing of cell size leads to extracting fewer features from the facial image.

Finally, Table 4-31 illustrates the final results of the proposed system using cell size=32.

Method	Overall Accuracy
HOG+KNN	87.20%
HOG+SVM	96.13%
LBP+KNN	87.44%
LBP+SVM	90.23%

Table 4-31: Results of all methods using cell size=32.

CHAPTER 5

5. CONCLUSION AND DISCUSSION.

5.1. Conclusion

In this thesis, an emotion estimation system from facial images is investigated by using different techniques. However, emotion estimation from facial images is not an easy task because analyzing data still suffer from difficulties, therefore classifying data is difficult operation too, especially when some data are almost similar. In this thesis, we began by describing our challenges and discussed the motivations that encouraged us to study this field. Moreover, we described the main problems related to these tasks in Chapter 1. Next, in Chapter 2 we presented some literature reviews, which contains related works in the field of emotion prediction from facial expression.

In chapter 3, we explain the methodology of this thesis, we began by explaining pre-processing steps of face detection, Dimensions Alignment, and Histogram Equalization, also the techniques that are used in each step. Moreover, the features are extracted from facial image using HOG & LBP algorithms and then these features are classified using KNN and SVM classifiers. The experiments achieved different performances, and the overall accuracy was 96.13% which is achieved by HOG+SVM method. Generally, from the performance of proposed methods, we noticed that the performance of SVM classifier is better than KNN classifier, on other hand, the performance of the proposed system is better when using cell size=32. In addition, HOG algorithm is better than LBP algorithm in feature extraction stage. Moreover, the expressions happy and surprise achieved high performance if compared with other expressions.

5.2. Discussion

The comparison between the previous studies and the proposed study is provided in *(Table 5-1)*. The previous methods and the proposed methods are applied on two databases (JAFFE & TFEID). However, the difference between the performance of the proposed methods and the previous methods depends on some factors such as; in the previous studies, the accuracy is chosen from the highest performance of expressions, while in the proposed system the overall accuracy is calculated by dividing the number of correct predictions on the total number of images. On other hand, in the previous studies, they didn't use all images in the databases, while in the proposed system, all images in the databases are used.

Researcher name	Database	Methods in the previous studies	Accuracy in previous studies	Methods in the proposed system	Accuracy in the proposed system
Caifeng, Shaogang and Peter. [11]	JAFFE	LBP+SVM	79.80%	LBP+SVM	77.46%
Junkai, Zenghai, Zheru & Hong [20]	JAFFE	HOG+SVM	94.30%	HOG+SVM	82.16%
Kiran, T. & T.Kushal [25]	TFEID	LBP+SVM	94.77%	LBP+SVM	83.33%
	JAFFE		97.10%		77.46%

Table 5-1: Comparison between the performance of previous studies and proposed study

5.3. Future Work

In this thesis, the proposed methods proved their effectiveness through achieving high performance. However, we documented some suggestions that may lead to improve the proposed system and its performance, and proving its quality.

These suggestions are:

- Instead of using the proposed system to predict emotions from facial image, a possible suggestion is to use this system to improve face recognition system which can be used in different security models such as criminal detection.
- Investigating the possibility of using other feature extraction algorithms instead of HOG and LBP. Thus, the effectiveness of other algorithms can be compared with

those used in this thesis. Similarly, using other classifiers instead of SVM and KNN is also another area of study and evaluation.

- Applying the proposed system to improve the performance of gender and age prediction system which could predict the gender and age of people from facial images.
- Using the databases that include facial images from different ethnic and applying the proposed system on these databases in order to ethnic prediction.

References

1. Ekman, P. and W.V. Friesen, Unmasking the face: A guide to recognizing emotions from facial cues. 1975, Englewood Cliffs, NJ: Prentice Hall.

2. Bartlett, M.S., et al. Real Time Face Detection and Facial Expression Recognition: Development and Applications to Human Computer Interaction. in Computer Vision and Pattern Recognition Workshop, 2003. CVPRW'03. Conference on. 2003. IEEE.

3. Pantic, M. and L.J.M. Rothkrantz, Automatic analysis of facial expressions: The state of the art. IEEE Transactions on pattern analysis and machine intelligence, 2000. 22(12): p. 1424-1445.

4. Martinez, A. and S. Du, A model of the perception of facial expressions of emotion by humans: Research overview and perspectives. Journal of Machine Learning Research, 2012. 13(May): p. 1589-1608.

5. Donato, G., et al., Classifying facial actions. IEEE Transactions on pattern analysis and machine intelligence, 1999. 21(10): p. 974-989.

6. Mehrabian, A., Communication without words. Communication Theory, 2008: p. 193-200.

7. Huang, C.-L. and Y.-M. Huang, Facial expression recognition using model-based feature extraction and action parameters classification. Journal of Visual Communication and Image Representation, 1997. 8(3): p. 278-290.

8. Michel, P. and R. El Kaliouby. Facial expression recognition using support vector machines. in The 10th International Conference on Human-Computer Interaction, Crete, Greece. 2005. Citeseer.

9. Gritti, T., et al. Local features based facial expression recognition with face registration errors. in Automatic Face & Gesture Recognition, 2008. FG'08. 8th IEEE International Conference on. 2008. IEEE.

10. Kharat, G. and S. Dudul, Human emotion recognition system using optimally designed SVM with different facial feature extraction techniques. WSEAS Trans. Comput, 2008. 7(6): p. 650-659.

11. Shan, C., S. Gong, and P.W. McOwan, Facial expression recognition based on local binary patterns: A comprehensive study. Image and Vision Computing, 2009. 27(6): p. 803-816.

12. Murugappan, M., N. Ramachandran, and Y. Sazali, Classification of human emotion from EEG using discrete wavelet transform. Journal of Biomedical Science and Engineering, 2010. 3(04): p. 390.

13. Niu, Z. and X. Qiu. Facial expression recognition based on weighted principal component analysis and support vector machines. in 2010 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE). 2010. IEEE.

14. Dhall, A., et al. Emotion recognition using PHOG and LPQ features. in Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on. 2011. IEEE.

15. Gupta, S.K., et al. A hybrid method of feature extraction for facial expression recognition. in Signal-Image Technology and Internet-Based Systems (SITIS), 2011 Seventh International Conference on. 2011. IEEE.

16. Kaur, M. and R. Vashisht, Comparative study of facial expression recognition techniques. International Journal of Computer Applications, 2011. 13(1): p. 43-50.

17. Zhang, L. and D. Tjondronegoro, Facial expression recognition using facial movement features. IEEE Transactions on Affective Computing, 2011. 2(4): p. 219-229.

18. Punitha, A. and M.K. Geetha, Texture based emotion recognition from facial expressions using Support Vector Machine. International journal of computer applications, 2013. 80(5).

19. Bagga, S., et al. Facial expression representation and classification using LBP, 2DPCA and their combination. in IT in Business, Industry and Government (CSIBIG), 2014 Conference on. 2014. IEEE.

20. Chen, J., et al. Facial expression recognition based on facial components detection and hog features. in International Workshops on Electrical and Computer Engineering Subfields. 2014.

21. Donia, M.M., A.A. Youssif, and A. Hashad, Spontaneous facial expression recognition based on histogram of oriented gradients descriptor. Computer and Information Science, 2014. 7(3): p. 31.

22. Suja, P., S. Tripathi, and J. Deepthy, Emotion recognition from facial expressions using frequency domain techniques, in Advances in signal processing and intelligent recognition systems. 2014, Springer. p. 299-310.

23. Abdulrahman, M. and A. Eleyan. Facial expression recognition using Support Vector Machines. in 2015 23nd Signal Processing and Communications Applications Conference (SIU). 2015. IEEE.

24. Carcagnì, P., et al., Facial expression recognition and histograms of oriented gradients: a comprehensive study. SpringerPlus, 2015. 4(1): p. 1.

25. Kiran, T. and T. Kushal. Facial expression classification using Support Vector Machine based on bidirectional Local Binary Pattern Histogram feature descriptor. in Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), 2016 17th IEEE/ACIS International Conference on. 2016. IEEE.

26. Wasista, S. and A. Oktavia, HUMAN FACIAL EXPRESSION DETECTION SYSTEM USING PCA AND SVM 2014.

27. Basu, A., A. Routray, and A. Deb, Human Emotion Recognition from Facial Thermal Image using Histogram based Features and Multi-Class Support Vector Machine. http://dx.doi.org/10.21611/qirt.2015.0055.

28. Kar, N.B., K.S. Babu, and S.K. Jena, Face Expression Recognition using Histograms of Oriented Gradients with Reduced Features. 2016.

29. Biehl, M., et al., Matsumoto and Ekman's Japanese and Caucasian Facial Expressions of Emotion (JACFEE): Reliability data and cross-national differences. Journal of Nonverbal behavior, 1997. 21(1): p. 3-21.

30. Van Der Schalk, J., et al., Moving faces, looking places: validation of the Amsterdam Dynamic Facial Expression Set (ADFES). Emotion, 2011. 11(4): p. 907.

31. Olszanowski, M., et al., Warsaw set of emotional facial expression pictures: a validation study of facial display photographs. Frontiers in psychology, 2015. 5: p. 1516.

32. Aifanti, N., C. Papachristou, and A. Delopoulos. The MUG facial expression database. in 11th International Workshop on Image Analysis for Multimedia Interactive Services WIAMIS 10. 2010. IEEE.

33. Lundqvist, D., A. Flykt, and A. Öhman, The Karolinska directed emotional faces (KDEF). CD ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet, 1998: p. 91-630.

34. Lyons, M., et al. Coding facial expressions with gabor wavelets. in Automatic Face and Gesture Recognition, 1998. Proceedings. Third IEEE International Conference on. 1998. IEEE.

35. Hjelmås, E. and B.K. Low, Face detection: A survey. Computer vision and image understanding, 2001. 83(3): p. 236-274.

36. Viola, P. and M. Jones. Rapid object detection using a boosted cascade of simple features. in Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on. 2001. IEEE.

37. Viola, P. and M.J. Jones, Robust real-time face detection. International journal of computer vision, 2004. 57(2): p. 137-154.

38. Sim, K., C. Tso, and Y. Tan, Recursive sub-image histogram equalization applied to gray scale images. Pattern Recognition Letters, 2007. 28(10): p. 1209-1221.

39. Adam, M.Y., M.M. Saeed, and A. Al Samani, Medical Image Enhancement Application Using Histogram Equalization in Computational Libraries. [Volume 6, Issue 1, January 2015].

40. Dalal, N. and B. Triggs. Histograms of oriented gradients for human detection. in 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). 2005. IEEE.

41. Mary, P.F.G., P.S.K. Paul, and J. Dheeba, Human identification using periocular biometrics. International Journal of Science, Engineering and Technology Research (IJSETR) Volume, 2013. 2.

42. Rybski, P.E., et al. Visual classification of coarse vehicle orientation using histogram of oriented gradients features. in Intelligent Vehicles Symposium (IV), 2010 IEEE. 2010. IEEE.

43. Lowe, D.G., Distinctive image features from scale-invariant keypoints. International journal of computer vision, 2004. 60(2): p. 91-110.

44. He, D.-C. and L. Wang, Texture unit, texture spectrum, and texture analysis. IEEE transactions on Geoscience and Remote Sensing, 1990. 28(4): p. 509-512.

45. Wang, L. and D.-C. He, Texture classification using texture spectrum. Pattern Recognition, 1990. 23(8): p. 905-910.

46. Ojala, T., M. Pietikäinen, and D. Harwood, A comparative study of texture measures with classification based on featured distributions. Pattern recognition, 1996. 29(1): p. 51-59.

47. Wang, X., T.X. Han, and S. Yan. An HOG-LBP human detector with partial occlusion handling. in 2009 IEEE 12th International Conference on Computer Vision. 2009. IEEE.

48. Savithiri, G. and A. Murugan, Performance analysis on half iris feature extraction using GW, LBP and HOG. International journal of computer applications, 2011. 22(2): p. 27-32.

49. Sarode, N.S. and A. Patil, Iris Recognition using LBP with Classifiers-KNN and NB. Index Copernicus Value (2013).

50. Tapkın, S., et al., Estimation of polypropylene concentration of modified bitumen images by using k-NN and SVM classifiers. Journal of Computing in Civil Engineering, 2013. 29(5): p. 04014055.

51. Boser, B.E., I.M. Guyon, and V.N. Vapnik. A training algorithm for optimal margin classifiers. in Proceedings of the fifth annual workshop on Computational learning theory. 1992. ACM.

52. Cortes, C. and V. Vapnik, Support-vector networks. Machine learning, 1995. 20(3): p. 273-297.

53. Sharma, N., B. Kumar, and V. Singh. Recognition of off-line hand printed English Characters, Numerals and Special Symbols. in Confluence The Next Generation Information Technology Summit (Confluence), 2014 5th International Conference-. 2014. IEEE.

54. Barkana, B.D. and J. Zhou, A new pitch-range based feature set for a speaker's age and gender classification. Applied Acoustics, 2015. 98: p. 52-61.

55. Sudha, L.R. and R. Bhavani, Gait based gender identification using statistical pattern classifiers. International Journal of Computer Applications, 2012. 40(8): p. 0975-8887.

56. Platt, J.C., N. Cristianini, and J. Shawe-Taylor. Large Margin DAGs for Multiclass Classification. in nips. 1999.

57. Vapnik, V.N. and V. Vapnik, Statistical learning theory. Vol. 1. 1998: Wiley New York.

58. Weston, J. and C. Watkins, Multi-class support vector machines. 1998, Citeseer.

59. Powers, D.M., Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. 2011.

60. Stehman, S.V., Selecting and interpreting measures of thematic classification accuracy. Remote sensing of Environment, 1997. 62(1): p. 77-89.

61. http://alearningaday.com/2016/09/confusion-matrix/ . 27.12.2016 - 11:12PM