Islamic University of Gaza
Deanery of Post Graduate Studies
Faculty of Information Technology

Human Emotion Recognition Approach Based on Facial Expression, Ethnicity and Gender Using Backpropagation Artificial Neural Network

By:
AbdulRahman M. Baraka

Supervised By:
Prof. Nabil M. Hewahi

A Thesis Submitted as Partial Fulfillment of the Requirements for the Degree of Master in Information Technology

1432 H – November 2011

Abstract

The emotions of the human are create by God almighty since the creation of humankind. Human emotions are mostly represents based on the psychological situation of humans through facial expressions, speech, or through the movement of the body. Overall, the interactions between humans are using several factors including the knowledge of these emotions.

Although people differ in their veins, races and languages, the language of emotions is almost general and comprehensive which makes it easy to understand. There are six basic emotions which usually researchers consider, these are happy, sad, fear, disgust, anger and shame.

In this thesis, we study the impact of both race and gender on the accuracy of the recognition of emotion through facial expressions. We claim that knowing the gender and race would increase the accuracy of the emotion recognition. This is due to the difference between the face appearances of various races and gender. To test our claim, we developed an approach based on Artificial Neural Networks (ANN) using backpropgation algorithm to recognize the human emotion. The proposed model consists of five stages: the first stage is inputting the image. Second stage is for image preprocessing. The third is to identifying points of the face, which will help in defining the face features. The fourth stage is to extract the features and the last stage is the emotion recognition. These stages are divides into two sections: the first section consists of the first four stages and the second section consists of only the fifth stage. We have built a program to implement the first section and we used Matlab to implement the other section.

Our model has been test by using MSDEF dataset, and we found that there is a positive effect on the accuracy of the recognition of emotion if we use both the ethnic group and gender as inputs to the system. Although this effect is not significant, but considerable (Improvement rate reached 8%). In addition, we found that females have a more accurate emotion expression recognition than males. In additional, regardless of the used dataset, our approach obtained better results than some researches on emotion recognition. This could be due to various reasons such as the type of the selected features and consideration of race and gender.

Keywords: Emotion Recognition, Artificial Neural Networks, Human Computer Interaction, Features Extraction.

عنوان البحث:

نموذج اكتشاف العواطف البشرية بناءا على تعبيرات الوجه والعرق والجنس باستخدام الشبكات العصبية

ملخص:

تعتبر العواطف من الصفات البشرية التي خلقها الله عز وجل في البشر منذ خلق البشرية. وتعكس هذه العواطف الحالة النفسية الداخلية للانسان من خلال التعابير الوجهية او نبرة الصوت او من خلال حركة الجسم. وبشكل عام ويتم التفاعل بين البشر باستخدام عدة عوامل من بينها معرفة هذه العواطف.

وعلى الرغم من ان البشر يختلفون في اجناسهم وعروقهم ولغاتهم. الا ان لغة العواطف هي لغة عامة وشاملة وسهلة الفهم. تتمثل العواطف الاساسية من ستة عواطف هي الفرح والحزن والخوف والدهشة والاشمئزاز والغضب.

في رسالتنا هذه قمنا بدراسة تأثير كل من العرق والجنس على دقة اكتشاف العاطفة من خلال تعابير الوجه. ونتوقع أن معرفة الجنس والعرق يؤثر على دقة اكتشاف العاطفة وهذا يرجع الى الفارق بين تعبير الوجه الخاص بكل عرق وبين الجنس. ولاختبار فرضيتنا قمنا ببناء نموذج مقترح يستخدم الشبكات العصبية وخوارزمية الانتشار العكسي للتعرف على العاطفة. يتكون هذا النموذج من خمسة مراحل هي : مرحلة اخذ الصورة ومرحلة معالجة الصورة و مرحلة تحديد نقاط الوجه والتي تفيد في عملية استخراج المميزات. والمرحلة الرابعة هي مرحلة استخراج المميزات والمرحلة الاخيرة هي مرحلة اكتشاف العاطفة. تنقسم هذه المراحل الى قسمين : القسم الاول يتكون من اول اربع مراحل والقسم الثاني يتكون من المرحلة الخامسة فقط. ولقد قمنا ببناء برنامج لتنفيذ القسم الاول وقمنا باستخدام برنامج الماتلاب لتنفيذ القسم الاخر.

تم اختبار النموذج باستخدام قاعدة بيانات خارجية وتم التوصل الى عديد من النتائج أهمها أن هناك تأثير ايجابي على دقة اكتشاف العاطفة اذا ما اخذنا بالحسبان كلا من العرق والجنس وإن كان هذا التأثير ليس بدرجة كبيرة ولكنها معتبرة (نسبة التحسن وصلت الى 8%). كذلك توصلنا الى أن الاناث تمتلك دقة اعلى من الذكور في التعبير عن العاطفة. و كذلك بغض النظر عن قاعدة البيانات المستخدمة توصلنا الى درجة دقة أعلى لبعض العواطف من بعض الابحاث السابقة. هذا يمكن ان يكون نتيجة لأسباب مختلفة ، مثل نوع المميزات المختارة اخذ العرق والجنس في الاعتبار.

الكلمات المفتاحية: اكتشاف العاطفة والشبكات العصبية والتفاعل بين الحاسوب والبشر واستخلاص المميزات.

To my supervisor Pro.Nabil…

ACKNOWLEDGEMENT

First, I thank Allah for guiding me and taking care of me all the time. My life is so blessed because of his majesty.

I would also like to take this opportunity to thank my research supervisor, Prof. Nabil Hewahi for giving me the opportunity to work with him and guiding and helping me throughout this research and other courses. And I do not forget his commendable effort to assist me since the beginning of my master study.

I would like to thank all the academic stuff in the Faculty of Information Technology at the Islamic University of Gaza (IUG).

Very special thanks to everyone in my family…

Very special thanks to my friends…

Thanks to all my friends in IUG; whom I consider as my brothers. Specially Raed and Eyad.

Finally, I hope this thesis will be a useful addition to the research activities of Emotion Recognition.

Table of Contents

English Abstract	I
Arabic Abstract	II
Dedication	III
Acknowledgment	IV
Table of Contents	V
List of Tables	VIII
List of Figures	IX
1. Chapter 1: Introduction	1
1.1 Human Emotions	1
1.2 Facial Expression	2
1.3 Emotion Recognition	2
1.4 Face Elements	3
1.5 Gender and Emotions	3
1.6 Ethnicity	4
1.7 Supervised learning	4
1.7.1 Classification	4
1.7.2 Artificial Neural Networks	5
1.7.3 Datasets	6
1.8 Problem Statement	6
1.9 Objectives	6
1.9.1 General Objective	6
1.9.2 Specific Objective	7
1.10 Significance of the thesis	7
1.11 Scope and limitation	8

1.12 Methodology 8

1.13 Thesis Outline 9

2. Chapter 2: Literature Review and Related Works. 10

2.1 Introduction 10

2.2 Psychological Studies on emotion 10

2.3 Computerized Emotions Recognition Models 12

3. Chapter 3: The Proposed Approach 16

3.1 Input the Image 16

3.2 Pre-processing 17

3.2.1 Contrast 17

3.2.2 Resize 18

3.2.3 Change resolution 18

3.3 Points Selection 18

3.4 Features Extraction 19

3.5 ANN Classifier 26

4. Chapter 4: Implementation and Experimental Results 27

4.1 Face Features Extraction program 27

4.1.1 Implementation 27

4.1.2 User Interface 28

4.1.3 Tools and Libraries 30

4.1.4 Data Encoder 31

4.2 ANN Classifier 32

4.3 Tools and programs 33

4.4 Datasets 34

4.5 Example 35

4.6 Experiment Results and Evaluations 36

4.6.1 First group of experiments 37

4.6.2 Second group of experiments 40

4.6.2.1. Asian experiments 40

4.6.2.2. Caucasian experiments 43

4.6.2.3. African experiments 46

4.7 Discussion 50

4.7.1 The First Experimental Group 50

4.7.2 The Second Experimental Group 51

5. Chapter 5: Conclusion and Future Works 56

5.1 Conclusion 56

5.2 Future works 57

6. References 58

7.Appendix 63

List of Tables

Table(3.1)	Group one of mouth features.	21
Table(3.2)	Group two of features.	22
Table(3.3)	Group three of features.	23
Table(3.4)	Group four of features.	24
Table(3.5)	Group five of features.	25
Table(3.6)	Group six of features	25
Table(4.1)	Ethnic group codes.	31
Table(4.2)	Emotions codes.	31
Table(4.3)	Gender codes.	31
Table(4.4)	Results of male experiment.	37
Table(4.5)	Results of female experiment.	38
Table(4.6)	Results of all races with gender.	39
Table(4.7)	Results of Asian male experiment.	41
Table(4.8)	Results of Asian female experiment.	41
Table(4.9)	Results of Asian race considering the gender.	42
Table(4.10)	Results of Asian race without considering the gender.	43
Table(4.11)	Results of Caucasian male experiment.	44
Table(4.12)	Results of Caucasian female experiment.	44
Table(4.13)	Results of Caucasian race considering the gender.	45
Table(4.14)	Results of Caucasian race without considering the gender.	45
Table(4.15)	Results of African male experiment.	46
Table(4.16)	Results of African female experiment.	47
Table(4.17)	Results of African race considering the gender.	48
Table(4.18)	Results of African race without considering the gender.	49
Table(4.19)	Accuracy of emotion recognition considering the ethnic group and gender.	50

List of Figures

Figure(1.1)	General structure of ANN.	5
Figure(3.1)	The Proposed Approach Structure.	16
Figure(3.2)	The 64 Points for features extraction	19
Figure(3.3)	Six emotions from MSDEF for males Asian	20
Figure(3.4)	Points of mouth.	21
Figure(3.5)	Points of left eye.	22
Figure(3.6)	Points of right eye.	23
Figure(3.7)	Points of left eyebrow.	24
Figure(3.8)	Points of right eyebrow.	25
Figure(3.9)	Feature No.28.	25
Figure(3.10)	Structure of ANN Classifier.	26
Figure(4.1)	First window for FFE.	28
Figure(4.2)	The main window in FFE.	30
Figure(4.3)	Neural Network training.	33
Figure(4.4)	Example of MSFDE Dataset.	35
Figure(4.5)	Example of features file.	36
Figure(4.6)	Examples of female happy.	38
Figure(4.7)	Examples of female fear.	39
Figure(4.8)	Examples of Angry from different races.	40
Figure(4.9)	Examples of shame emotion for Asian race.	42
Figure(4.10)	Examples of disgust emotion for Asian race.	43
Figure(4.11)	Disgust emotion for African female.	47
Figure(4.12)	Happy emotion for African race	48
Figure(4.13)	Disgust emotion for African race	48
Figure(4.14)	Features of disgust and happy emotion for African	49
Figure(4.15)	Accuracy of emotions for male without considering the ethnic group.	51
Figure(4.16)	Accuracy of emotions for female without considering the ethnic group.	51
Figure(4.17)	Accuracy of emotions considering the ethnic group with gender as feature.	53
Figure(4.18)	Number of accuracy regardless the ethnic group with gender as feature.	54
Figure(4.19)	Number of errors for each emotion considering the ethnic group.	54
Figure(4.20)	Number of errors for each emotion regardless the ethnic group.	55

Chapter 1

Introduction

Obviously, humans communicate and interact among each other through speech and body movement. It is also obvious that there is a close linkage between human emotions expressions and his facial expressions. This implies that emotion is an important aspect in the interaction and communication between people. Since the science of Artificial Intelligence (AI) is concerned with the automation of intelligent behavior, we need to improve the interaction between computer and human. In this chapter, we discuss three main principles, emotions and human beings, supervised learning and neural networks, and finally the problem statement, objectives, and scope and limitations of the work.

1.1 Human Emotions

No aspect of our mental life is more important to the quality and meaning of our existence than emotions. They are what make life worth living, or sometimes ending. So it is not surprising that most of the great classical philosophers had recognizable theories of emotion [42]. In fact, there are wide ranges of human emotions, which can express the internal emotion; these emotions can be divided into two group's basic and secondary. The basic emotions consist of six emotions happy, sad, fear, disgust, angry and surprise, and the other emotions are secondary emotions (e.g, skeptical). In addition, each group divided into two sets, positive and negative. The basic emotions divided into happy as positive emotion and all other emotions as negative emotions [27].

Another concern is that some emotion appearances are very close to each other like skeptical and Frown emotions; this will make the recognition harder, and forced the researchers to limit the classification to small number of emotions that concentrate on the basic six emotions or in other cases on maximum of eight emotions. Recognition of the emotions based on the face appearance is very concerned with the extracted face features. This process called feature extraction [1]. The extracted features are then transformed into input data to any emotion recognizable system.

It should be note here that, human emotions occur at different levels of intensity and at different moments in time. There is a difference between "a little bit sad" and "very sad" in their role of influencing behavior, facial expression and other emotions, even though the both expressions are negative emotion.

1.2 Facial Expression

A facial expression is one or more emotions or positions of the muscles of the face. These movements convey the emotional state of the individual to observers. Also, facial expressions are a form of nonverbal communication, and their significance in the perceiver can, to some extent, vary between cultures [41].

While facial expression is the key to the knowledge of human feelings and emotions, we need to detect these emotions from facial expressions to reach the main objective and goal of developing and improving the interaction between computer and human and make it naturally specially it is the latest challenges in human-robot interaction [36].

In the computer vision community, the term ”facial expression recognition” often refers to the classification of facial features in one of the six, so called basic emotions: happiness, sadness, fear, disgust, surprise and anger[17].

1.3 Emotion Recognition

The ability to recognize emotion from facial expressions appears at least partially inborn. Newborns prefer to look at faces rather than other complex stimuli, and thus may be program to focus on information in faces [14].

During the last years, interest in emotions increased considerably in various domains of computer based systems, for example:

- Medical area: recognizing the expression of a man can help in the field of medical science where a doctor can be alert when a patient is in severe pain. It helps in taking prompt action at that time; also, expressed emotion has been use as a construct in understanding the interaction between patients and their careers and families [21].

- E-learning area: understanding the emotional reaction of a student in a complicated learning environment is a mind boggling task. This is performs by detecting intense emotional experiences being exhibited by students, and detects fluctuations in emotion as learning progresses [5].

- Entertainment area: classifying emotional states of neutral and anger, and demonstrate that a new genre of languages called Game Pidgin language (GPL) that can not only be used to capture real time speech recognition, but can also generate response from the Non-Player-Character(NPC) [33].

1.4 Face Elements

The face consists of a set of elements, eyes, lips, eyebrows, nose and Chin, where each element plays a role in the expression of the emotion. The chosen elements are usually depending on the used module or classifier, and which features are chosen.

In general, face elements can be divides into two groups based on their role, primary and secondary. The primary elements are mouth (and lips), noise, eyes and eyebrows. Other elements are the secondary elements like cheeks, chin and front face. The secondary elements of the face do not play an important role in the recognition of the basic emotions of human beings, but have an important role in the secondary emotions like skeptical and shy, and thus the primary elements sufficient to discover the basic human emotions. In addition, the secondary elements usually used in video streaming (dynamic image) [41].

1.5 Gender and Emotions

Gender is one of the most important factors in emotion detection systems. Gender differences were important in a variety of cultures and ethnic groups. The gender was even more important than ethnicity in shaping people's emotional ideology, emotional experience, and habits of emotional expression. Men tended to be persuaded that somewhat more emotional management is necessary. In addition, men and women may share roughly the same sorts of emotional experiences in their relationships, but they differ in how freely they express their feelings. Men tended to express their positive and negative emotions less frequently and less intensely than they were experienced. Women tended to be somewhat more direct in their emotional expression [13].

1.6 Ethnicity

An ethnic group (or ethnicity) is a group of people whose members identify with each other, through a common heritage, often consisting of a common language, a common culture (often including a shared religion) and an ideology that stresses common ancestry or endogamy [22].

In general, emotional communication is generally more accurate among people who share similar cultural backgrounds [14]. Appearance of emotions are universal across individuals as well as human ethnics and cultures, but the way of emotion expression is that vary from one ethnic group to another and the differences cross-cultural in the familiarity effect on recognition accuracy for different types of expressions. This is commonly known in biology, but never tests on computer systems [4].

1.7 Supervised learning

Supervised learning is a machine-learning task of inferring a function from supervised training data. The training data consists of a set of training examples (dataset). In supervised learning, each example is a pair consisting of an input object (typically a vector) and a desired output value (called the supervisory signal). A supervised learning algorithm analyzes the training data and produces an inferred function, which called a classifier or a regression function[1] [42].

1.7.1 Classification

"In machine learning and pattern recognition, classification refers to an algorithmic procedure for assigning a given piece of input data into one of a given number of categories". [2]

[1] Wikipedia Encyclopedia. http://en.wikipedia.org/wiki/Supervised_learning.
[2] Wikipedia Encyclopedia . http://en.wikipedia.org/wiki/Classification_in_machine_learning.

1.7.2 Artificial Neural Networks (ANN)

Artificial neural networks (ANN or NN) originate from the desire to mimic the human brain in a computer. ANN can be seen as a simplified model of the human brain and consists of several neurons, where some neurons are connected. The neurons are also referring to as processing elements (PE), as they process the inputs given to it, combine all values and produce one output value. ANNs are typically organized in several layers. One input layer and one output layer are always present, but one or more hidden layers can be added. The neurons from one layer connected to the neurons in the next layer, giving its output as input to the next layer neurons. The input to the input layer is the original data to be classified, in this case the feature values. The outputs from the output layer are the values on the arousal and valence scale [20].

In an ANN there is always one input layer, with as many neurons as there are features. The number of neurons in the output layer is the same as the number of desired outputs or classes. The number of hidden layers can be chosen freely. Using one or more hidden layers gives the ANN the capability of learning more complex relations between the input and output. However, it also increases the number of samples needed to train the system. Within the hidden layers, the number of neurons can also be chosen freely. More neurons in a layer can result in a better overall result, but as with the number of layers, the number of samples needed for training also increases with a higher number of neurons in a layer [20]. Figure 1.1 shows the general structure of ANN contains three layers.

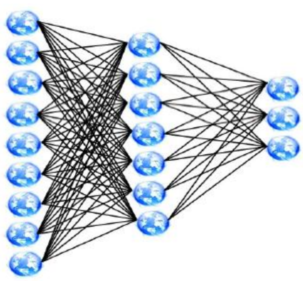


Figure 1.1 general structure of ANN. [21]

1.7.2 Datasets

One challenge of research in expression and emotion recognition is collection of suitable data for training and testing the systems. The dataset is very important to evaluate our approach. Several databases have been developed for facial expressions. Some are small and have limited accessibility, and others have one gender only [26], others have many angel face[29], other for voice signals[18], and others have used actors to portray emotions in video recordings [9][16].

Most of the previous datasets do not deal with vary races or ethnic groups in details, although there are many datasets dealing with one specific race [26]. The Montreal Set of Facial Displays of Emotion (MSFDE) is a dataset that consists six of emotional facial expressions (happy, sad, fear, disgust, angry and shame) by men and women of four ethnic groups (Caucasian, Asian, African and Hispanic), and each ethnic group in dataset has four persons (men or women).

1.8 Problem Statement

The problem can be briefly stated as: Discovering the effect of the ethnic/race group or gender on recognizing the human emotions based on face features. That leads to increase the accuracy of emotion recognition systems.

1.9 Objectives

1.9.1 General Objective

- Recognize human emotions based on facial feature considering various ethnic groups.
- Evaluate the hypothesis that's does determining ethnic group affect the accuracy of emotion recognition?
- Evaluate the hypothesis that's does the gender affect the accuracy of emotion recognition considering ethnic groups?
- Determine the emotions that may be shared between the ethnic groups expression.

- Determine the most ethnic groups that have maximum accuracy of the expression of feelings.
- Does gender alone has a better recognition performance than even with ethnic consideration?

1.9.2 Specific Objective

- Design a system for features selection and extraction from human face image**.**
- Developing the appropriate ANNs structure for classifier.
- Evaluating the approach.
- Exploring the pre-processing and selecting the proper one.
- Determine the face elements from face human image, which are required in our approach.
- Finding appropriate datasets that can be suitable to serve our goal.
- Investigating the emotions that can be considered in our work.

1.10 Significance of the thesis

1. Study the impact of the ethnic group and gender on the accuracy of the recognition of emotions. That leads to enhancement the recognition systems that used in E-learning, Entertainment, medical and other areas.
2. Finding the relations between gender, emotion and ethnic group.
3. Improve the interaction between the computer and human, which will add to the machine intelligence. This will improve the efficiency of man-machine interaction.
4. Promote and contribute to the scientific research in the area of human emotion recognition systems.

1.11 Scope and limitation

1. Consider only human face images not video.
2. Consider only stored images.
3. The human image has medium resolution and size, and not deal with low quality image. In other words, the minimum resolution of the image is 220 pixels in widths and 220 pixels in height, and the minimum depth is 92 pixels.
4. Considering the frontal face only, no bending face to any angle or rotated face to any side is accepted.
5. We use only six basic emotions, which are happy, sad, angry, fear, disgust and shame. We use the shame emotion (not surprise) because the dataset replace surprise emotion with shame emotion.

1.12 Methodology

Our approach has the following steps to achieve our objectives:

A. Literature survey: this includes reviewing the recent literature closely related to the thesis problem statement and the research question. After analyzing the existing methods, identifying the drawbacks or the lack of existing approaches, we formulate the strategies and solutions how to overcome the drawbacks.

B. Develop method of approach: to solve the research problems we build a new approach to recognition of human emotion using MSFED dataset. Chapter 3 depicts our proposed approach.

C. Implement method of approach: To apply the first four stages, we implement a program using Microsoft Visual Studio .net[1] environment

[1] Microsoft Visual Studio.net® is a complete set of development tools for building ASP Web applications, XML Web services, desktop applications, and mobile applications.

by Visual C# language[1]. After that, we use Matlab[2] environment to build an ANNs classifier and apply the other stages of approach.

D. Design experimental: to verify the developed approach we apply a various experiments on MSEDEF to test our claims.

E. Evaluate the obtained results: in this stage, we will analyze the obtained results and justify our approach feasibility by comparing it with other approaches.

1.13 Thesis Outline

The thesis is organized as follows: Chapter two is an overview of the background literature survey and related works of recognizing the emotions. Chapter three describes in detail the system structure and the proposed approach including methods, procedures, features selection, features extraction and the used classification approach. Chapter four describes the implementation, experiments and the obtained results with discussion. Chapter five presents the conclusions and future works.

---

[1] Microsoft Visual C# .net® use integrated development environment (IDE).

[2] Mathworks MATLAB® is a high-level language and interactive environment that enables you to perform computationally intensive tasks faster than with traditional programming languages such as C, C++.

Chapter 2

Literature Review and Related Works

2.1 Introduction

Most of the researches and studies that were interested in human emotions and facial expression are often physiological researches, which concentrate on studying the human emotions and feelings through facial elements. Also Cultural studies of emotions originated from anthropology, sociology and psychology. So any research interested in studying emotions and cultural must be addressed in the aspects of emotions in psychology studies [4].

Many of the research and psychological studies are interested on this subject, but few computer studies are focused on impact of race and gender with the recognition of emotion through facial expressions in detail.

This chapter surveys the background literature relating to different psychology and computerized models for emotion recognition, their uses in computing and their implications for the work described in this thesis.

2.2. Psychological Studies on Emotion

Many psychology studies describe a classic research from the 1960s demonstrating that participants around the world could judge the intended basic emotional states portrayed in posed photographs at rates better than would be expected from random guessing [30]. In the psychological literature, emotion has been defines as an individual’s response to goal-relevant stimuli that includes behavioral, physiological, and experiential components [3].

Paul Ekman[1] is a psychologist who has been a pioneer in the study of emotions and their relation to facial expressions. He has been considers one of the 100 most eminent psychologists of the twentieth century. Among the most famous research is Facial Action Coding System (FACS).

[1] Wikipedia Encyclopedia. http://en.wikipedia.org/wiki/Paul_Ekman.

FACS is the most widely used and versatile method for measuring and describing facial behaviors. The original FACS developed in 1970s by determining how the contraction of each facial muscle (singly and in combination with other muscles) changes the appearance of the face. For example, code 23 is given for lip funnel, code 4 for eye brow lower, and code 10 for chin raise …etc. And we can combine many facial codes to describe our emotion. The new version of FACS is developed in 2002 [23].

The early researchers who studied how people communicate emotion across cultures focused their efforts on establishing universality, and therefore did not pay as much attention to the cultural differences as to the cross-cultural similarities in their data. Psychological research has classified six facial expressions that correspond to distinct universal emotions: disgust, sadness, happiness, fear, anger and surprise. Many psychologists concluded that the recognition of emotion is largely universal, with the implication that this skill is not learned, but rather has an evolutionary and thus biological basis [15].

Although emotions are recognize at above chance levels across cultural boundaries, there is also evidence for an “ingroup advantage,” such that emotion recognition is more accurate when individuals judge emotions expressed by members of their same national, ethnic or regional group. While emotional communication may be a universal language, there may also be subtle differences in this language across cultural groups such that we can better understand people expressing themselves using a similar “dialect” to our own. Further, there is evidence that cross-cultural exposure can reduce the size of the ingroup advantage in emotion recognition just as familiarity with a dialect improves understanding of a spoken language [14].

Nauert .R [28], published a new study called "Are emotions universal?". This study investigates whether basic emotions are influence by the environment or are genetically hardwired into all human beings. The study, conducted from the University of London, and compared people from Britain and Namibia, and findings suggest basic emotions such as amusement, anger, fear and sadness are shared by all humans. Participants in the study listened to a short story based around a particular emotion, for example, how a person is very sad because a relative of theirs had died recently. At the end of the story they heard two sounds – such as crying and of laughter – and were asked to identify which of the two sounds reflected the emotion being expressed in the story. The British group heard sounds from the Namibia and vice versa. People from both groups seemed to find the basic emotions – anger, fear, disgust, amusement, sadness and surprise – the most easily recognizable. This suggests that these emotions are similar across all human cultures, and one positive sound was particularly well recognized by both groups of participants. Although, not all positive sounds were easily recognizable to both cultures. However, some such as the sound of pleasure or achievement appear not to be shared across cultures, but are instead specific to a particular group or region[1].

Ekman [11], presented a study that having participants in ten different countries rate expressions on the basis of seven expression types (anger, disgust, fear, happiness, sadness, surprise, contempt), and apply two rates. The first rate applied on primary expression through the first viewing of photographs, participants chose the predominantly expressed emotion. The results showed that there was high agreement across cultures supporting universality. A second rating was performed where the faces were rated in intensity for each of the seven expression types. The results show that universality in not only choice of the primary expression, but even the second most expressed emotion in each face, effectively finding cross-cultural agreement in emotion blends. There was a lack of universality found with the overall intensity ratings. The author posited that this may be due to other-culture effects; however, this study lacked sufficient evidence to support this.

We conclude that many psychological studies refer to that the ethnic group is a consider factor in emotion recognition, and may play an important role to increasing the accuracy of emotion recognition.

2.3 Computerized Emotions Recognition Models

In the early 1990's the engineering community started to construct automatic methods of recognizing emotion from facial expression in an image and videos [23].

There are many of computer studies that focused on emotions recognition. Each study uses some factors that affect the expression of the emotions like voice,

[1] Nauert PhD, R. (2010). Are Emotions Universal?. Psych Central. http://psychcentral.com/news/2010/01/27/are-emotions-universal/10999.html.

facial expressions, pulse, body movement or any factor. These factors can help in recognizing the human emotions and feeling. Usually recognition model has two stages, learning and testing stages. In learning stage some properties of factors - often called features- used to train the model. Then use the trained model to recognize the human emotion from an image or video or camera.

Raheja and Kumar [30], presented an architecture for human gesture recognition, considering color image with different gestures by using backprobagation Artificial Neural Networks (ANN or NN). Four stages applied in the approach, face detection, image reprocessing, training network and recognition module. The pre-processing stage contains three methods, histogram equalization, edge detection, thinning and token generation. The ethnic group was not considers in this model. The module was train using the three different gesture images, happy, sad and thinking expressions of faces. The model was tests with 100 images of three gestures, the results were 94.28% for happy, 85.71% for sad and 83.33% for thinking. In this research the gender is not considered, that means that the researchers put all the images all together without giving input to the gender.

Akram et.al [1], presented a fuzzy classifier system using Mamdani-style for facial expression recognition. Three modules are prepared is this classifier system, first module for Pre-Processing Module (PPM), the second for Region Extraction (REM), and the last module for Feature Extraction (FEM). PPM and REM used to detect eight facial elements eyes, eyebrows, nose, forehead, cheeks, lips, teeth and chin. FEM is use to find the facial action values for all facial action elements. Later a fuzzy module called Expression Recognition Module (ERM) is used, the inputs for this module are Membership Functions (MFs) for all facial action elements, and the outputs has seven output MFs representing the basic facial expressions (Anger, Disgust, Sad, Normal, Happy, Surprise and Fear). The best results appear in disgust and happy by 100%, 95% in surprise, and fear 90%; and other expressions give 70%. Again, the ethnic group was not used in this system.

Karthigayan et.al [23] used Genetic Algorithms (GAs) and Neural Networks (ANNs) to build human emotion classifier. This classifier detects six human emotions Neutral, Sad, Anger, Happy, Fear, Disgust (or Dislike) and Surprise. They depend on two facial elements in the classifier, eyes and lips. By applying some preprocessing methods and edge detection, they extracted the eyes and lip regions, then extracted the features from these regions. Three feature extraction methods are applied, projection profile, contour profile and moments. The GA is apply to get the optimized values of the minor axes of an irregular ellipse corresponding to the lips and the minor axis of a regular ellipse related to eye by using a set of new fitness functions. Finally, they apply the results from GA on the ANNs model. Two architectures of ANNs models are proposed with an average of 10 trials of testing. The achieved results of 3x20x7 and 3x20x3 of ANN architecture were 85.13% and 83.57% of success rate respectively. The successful classification even goes to the maximum of about 91.42% in the NN model of 3x20x7 structure. A South East Asian (SEA) subject is only considered in this work and ethnic group is not considered.

Matthew and Patterson [31], discuss a framework for the classification of emotional states, based on still images of the face, using active appearance model (AAM) and get distance using n Euclidean as features. To train and test the classifier they chose to use the facial expression database known as “FEEDTUM”[1], and seven basic emotions are used, happy, sad, angry, surprise, fear, disgust and natural state. The best results they obtained are in happy, natural and disgust emotions at the rate 93.3%, fear at 90.0%, and 79.7% in surprise, angry and sad at rate 63.9%.

In Kim and et.al [24], an emotion detection algorithm using frontal facial image presented in this paper. The algorithm is composed of three main stages: image processing stage and facial feature extraction stage, and emotion detection stage. In image processing stage, the face region and facial component extracted by using fuzzy color filter, virtual face model (VFM)2, and histogram analysis method. The features for emotion detection extracted from facial component in facial feature extraction stage. In emotion detection stage, the fuzzy classifier adopted to recognize emotion from extracted features. To evaluate the performance, 124 facial images acquired from 20 men and 20 women. Five facial images representing five emotions are acquired from each person happy, sad, disgust, surprise, and angry. The experimental results are 74% for final emotion detection accuracy, and the accuracy

[1] Facial Expressions and Emotion Database(FEED).

[2] Virtual Face Model (VFM) based histogram analysis. And it is contains position and length information of each facial component. And proposed to reduce searching space of histogram analysis method. .

of recognition in happy is 79.7%, Sad 69.9%, Angry 72.3%, Disgust 69.9% and 78.5% for Surprise.

Vogt and Andre [37] propose a framework to improve the discriminative quality of gender-dependent features. In general, the framework first predicts the gender of the speaker to determine which gender-specific emotion recognition system should be used. Then depending on the outcome, the framework uses a gender-specific emotion recognizer. A Naive Bayes classifier is used, and the basic human emotions were tested. This framework was test on two different databases, one with emotional speech produced by actors called Berlin database, and one with spontaneous emotional speech from a Wizard-of-Oz setting called SmartKom database. The classifier was tested by using three tests, first test without gender information, the second test with gender information, the last test with automatically detected gender information.

Gender detection achieved an accuracy of about 90% and the combined gender and emotion recognition system improved the overall recognition rate of a gender-independent emotion recognition system by 2–4 %. The ethnic group was not considered in this model.

From the above survey, we conclude that most of the emotion detection researches do not consider ethnic/race groups, and whether recognition of ethnic group increases the system's performance or decrease it. Also just few researches considered the gender to scale their system's performance based on known or unknown genders.

Chapter 3

The Proposed Approach

In this chapter we shall describe in details our proposed approach structure, and all the used methods. The general approach structure is shown in Figure 3.1.

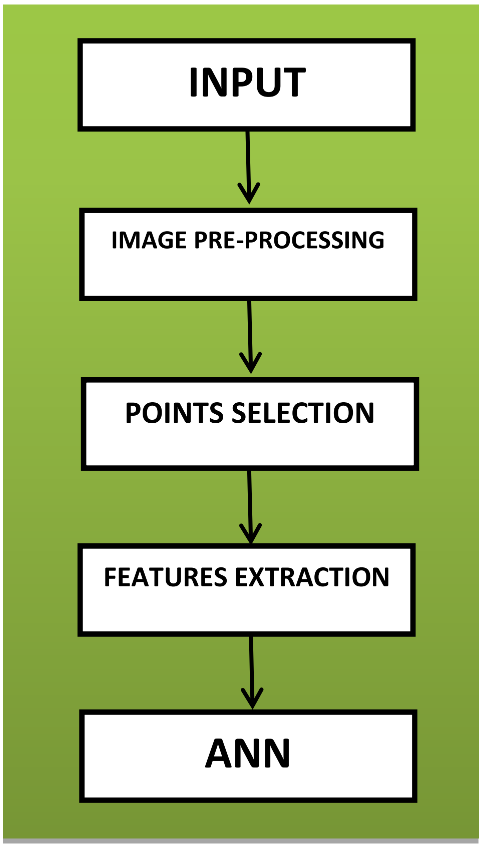


Figure 3.1 The Proposed Approach Structure.

Our proposed approach uses the face expression to detect the emotions. Six steps are applied: getting the image, applying pre-processing on the image, selecting the required points on the face, extracting the features, and finally applying ANN classifier to recognize the emotion. Each step is explained in details in the following section.

3.1 Input the Image

This step is considered to be the first input, where the image of the human being considered is inserted. The image can be either colored or gray image file of human face with most used image extension (jpg, bmp, tiff …etc.). The image must be in medium quality at least, that is, the minimum resolution of the image is 220 pixels in width and 220 pixels in height and the minimum depth is 92 pixels, because if the resolution minimum is less than what is specified, selection of the points would be very difficult.

There are also some restrictions on the image of the human face in order to ensure uniformity of characteristics and features that are taken in the next stages. The most important restrictions are:

- Human face is frontal: the face must be in front and does not rotate to any angle, and not curvature to any direction.
- Face elements are clear: this means the human does not wear a classes or any mask to ensure that the user can select the points clearly.

3.2 Pre-processing

In preprocessing we apply certain steps to achieve good images with one standard in terms of contrast and size. This will help in the next steps used in the proposed approach especially in point selection and getting the face features.

3.2.1 Contrast

Contrast is the difference in visual properties that makes an object (or its representation in an image) distinguishable from other objects and the background. In visual perception of the real world, contrast is determined by the difference in the color and brightness of the object and other objects within the same field of view. Because the human visual system is more sensitive to contrast than absolute luminance, we can perceive the world similarly regardless of the huge changes in illumination over the day or from place to place[1].

We use a simple contrast operation that increases the value of red, green and blue color by specific percentage. This process was applied by using graphics library for Visual studio.net.

[1] Wikipedia Encyclopedia. http://en.wikipedia.org/wiki/Contrast_(vision).

3.2.2 Resize

This process is to change the width and height of the image. We use a standard size for all images. The standard width is 800 pixels and height is 560 pixels with resolution 160 bit.

Simple operation to apply resizing method is that, create a new image with new size and redraw the original image into new image. This process was applied by using graphics library for Visual studio.net.

3.2.3 Change resolution

This process is to change the resolution of the image. We use a standard resolution for images. The standard resolution is 90 pixels. There are already function in Visual Studio.net to change the resolution for any image immediately.

3.3 Points Selection

At this point, the image of human face is ready to select the points that needed by using 64 points distributed over human face and use these points for features extraction. The choice of these points is to determine the shape of each element of the face (eyes, eyebrows and mouth). The number of points and the position of points are not standardized, but it is depending on the features that will be extracted, and used for the classifier. Many researches use various number of points and positions based on their view about the feature to be considered [24] [35] [39].

Each point is determined by two axis's X and Y, and after selecting all the points in the face we save the points for each image in a text file. The saved points can be loaded and shown on the face at any time.

We classified the points into seven groups; each group is concerned with one of the face elements. First group has 14 points from 1 to 14, and describe the face dimensions. Second group has 8 points from 15 to 22, and describe the layout of the right eyebrow. Third group has 8 points from 23 to 30, and describe the layout of the right eye. Forth group has 8 points from 31 to 38, and describe the layout of the left eyebrow. Fifth group has 8 points from 39 to 46, and describe the layout of the left eye. Sixth group has 4 points from 47 to 50, and describe some portions of the nose.

Last group has 14 points from 51 to 64, and describe the layout of the mouth. These points are shown in Figure 3.2.

3.4 Features Extraction

This stage is the last stage before we apply the ANN classifier. In this stage we calculate 28 features. The features are the distances between certain points explained in the previous stage. While human emotion is changing, face expression is also changing. This means that, properties of face elements do change when face expression is being changed. As a consequence of this, the distances between points are changing. The distances between certain points as will be shown later describe the human emotion. The output of this stage is the certain distances between certain points, these distances are the image features. The output of this stage will be the input for the ANN classifier.

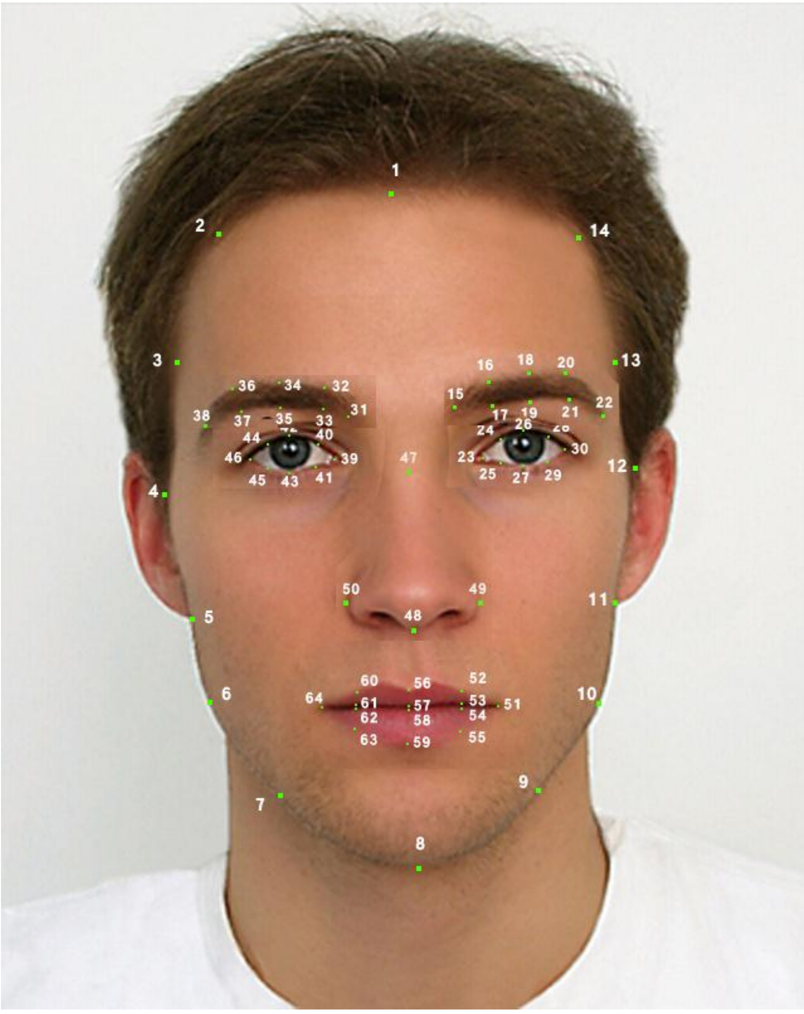


Figure 3.2 the 64 Points for features extraction.

Based on what we stated above, the features are classified into six groups. Each group describes the features of one face element. All features are a vertical distances between two points. Group one contains all features of mouth; there are 14 points on the mouth as shown in Figure 3.2. The mouth points are split into two groups, upper group and lower group. Upper group is (52, 53, 56, 57, 60 and 61) and lower group is (54, 55, 58, 59, 62 and 63).

Changing shape of the mouth has an affect on facial expression. It is to be noted that, the changes between the pointes in the mouth shape are minor at the horizontal level, whereas in the vertical level the changes are large and significant. This is illustrated in Figure 3.3. Based on this, we ignore the changing in the horizontal level which if considered may cause a recognition problem.

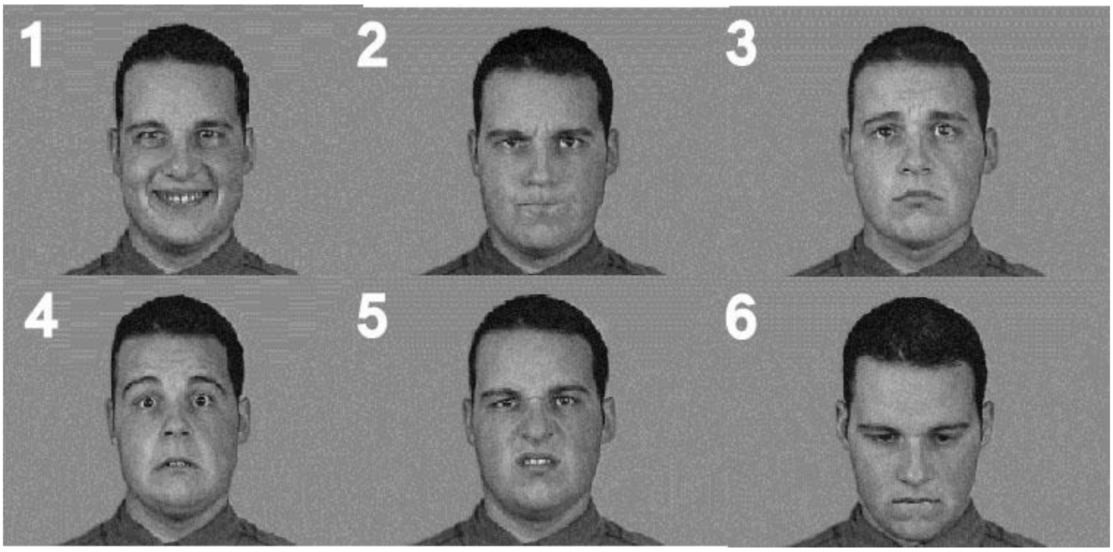


Figure 3.3 Six emotions from MSDEF for males Asian (1:Happy, 2:Angry, 3:Sad, 4:Fear, 5:Disgust and 6:Shame) .

When the emotion is happy the mouth was opening with medium distances, and the left and right sides become near from upper points and far from lower points. In sad emotion, the mouth becomes almost with no distance between middle points, and the left and right sides are near or below the lower points and far from upper points. In angry emotion, no distances in middle points and the sides become in the middle too. In fear emotion, medium distances in middle points, and the sides sometimes near or below the lower points and far from upper points, and in other times they may be in the middle. Disgust emotion has large distances in middle points, and the sides points in the middle. Final emotion is shame, and it has no distances in the middle points and the sides in middle too. Figure 3.4 shows the mouth points.

So the features of the mouth are the distances between points in upper group with their corresponding points in lower group. We ignored the upper points in upper groups (52, 56 and 60), because we want to measure the amount of mouth opening and not the overall size of the mouth. Also we ignore the lower points in lower group (55, 59 and 63) for the same previous reason.

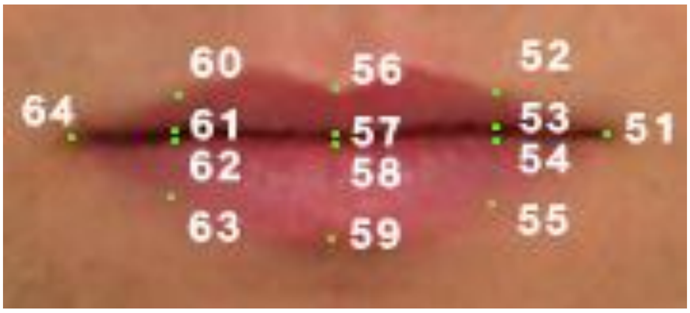


Figure.3.4 Points of mouth.

So we have seven features from these points. First three features to measure the distances between two lips. Other four features to measure the amount of deviation from the middle of the mouth (horizontal middle) on both sides (right and left side). Table 3.1 shows the group one of the features:

Feature No.	**Vertical distances between :**	
	Point 1	**Point 2**
Feature 1	**54**	**53**
Feature 2	**58**	**57**
Feature 3	**62**	**61**
Feature 4	**51**	**53**
Feature 5	**54**	**51**
Feature 6	**64**	**61**
Feature 7	**62**	**64**

Table 3.1 Group one of mouth features.

Group two contains features of the left eye. There are 8 points in the left eye, 3 points in the upper portion, 3 points in the lower portion, 1 point in the left side and 1 point in the right side. Figure 3.5 shows the points of the left eye. In this group we also focus on vertical distances because the eyes move up and down only. We also measure the deviation of eye points from the middle of the eye.

Figure 3.5 Points of left eye.

In happy emotions, the distances between upper and lower points are medium (less than sad emotion), and the left side and right side of the eye moving to the middle. In sad emotion, the distances between upper and lower points are high medium. Angry emotion has low medium distances between upper and lower points. Fear has the highest distances between upper and lower points. Disgust is as angry, the distances are low in the medium. Shame has the lowest distances between upper and lower points.

Based on the above mentioned points, we extract seven features to describe all the left eye shapes using eye points. First three features to measure the size of the eye. The other four features are to measure the amount of deviation from the middle of the eye (horizontal middle) on both sides (left and left side). Table 3.2 shows the group two of features.

Feature No.	**Vertical distances between :**	
	Point 1	**Point 2**
Feature 8	**25**	**24**
Feature 9	**27**	**26**
Feature 10	**29**	**28**
Feature 11	**23**	**24**
Feature 12	**25**	**23**
Feature 13	**30**	**28**
Feature 14	**29**	**30**

Table 3.2 Group two of features (left eye).

Group three contains features of right eye. Same position points in left eye are used, and all criteria of left eye are applied on the right eye (this does not mean that we get the same distances of the right eye). Figure 3.6 shows the points of the right eye.

Figure 3.6 Points of right eye.

Seven features are used to describe all right eye shapes. First three features are to measure the size of the eye. The other four features are to measure the amount of deviation from the middle of the eye (horizontal middle) on both sides (left and right side). Table 3.3 shows the group three of features.

Feature No.	**Vertical distances between :**	
	Point 1	**Point 2**
Feature 15	**41**	**40**
Feature 16	**43**	**42**
Feature 17	**45**	**44**
Feature 18	**39**	**40**
Feature 19	**41**	**39**
Feature 20	**46**	**44**
Feature 21	**45**	**46**

Table 3.3 Group three of features (right eye).

Group four contains features of the left eyebrow. In general, eyebrow is helpful for emotion recognition by measuring three criteria:

1- Amount of deviation.

2- Direction of deviation.

3- Distance from the eye.

Figure 3.3 shows how applicable the above criteria at various emotions, and how eyebrow shape changes depending on the face expression. The width of eyebrow is not significant in emotion detection.

We select 8 points to describe the left eyebrow and use it to get all features; these points are shown in Figure 3.7. We ignore the points (17, 19 and 21), and take features from upper points (15, 16, 18 and 20), because it is enough to measure our criteria.

Figure 3.7 Points of left eyebrow.

We choose three features from the upper points to measure the first two criteria (deviation and direction), but the third criteria is measured in group six, because it depends on one point in the mouth. Table 3.4 shows the group four of features.

Feature No.	Vertical distances between :	
	Point 1	**Point 2**
Feature 22	**15**	**16**
Feature 23	**16**	**18**
Feature 24	**18**	**20**

Table 3.4 Group four of features (left eyebrow).

Group five contains features of the left eyebrow. We use the same points and features that are used in the right eyebrow. Figure 3.8 shows the 8 points that are used in the right eyebrow.

Figure 3.8 Points of right eyebrow.

Last group has one feature only which is the distance between the beginning of the eyebrow and the beginning of the eye in same side. This is significant (from point 23 to 15) because it is used to measure the third criteria that were given in the right eyebrow features. This feature is shown in Figure 3.9 by red line.

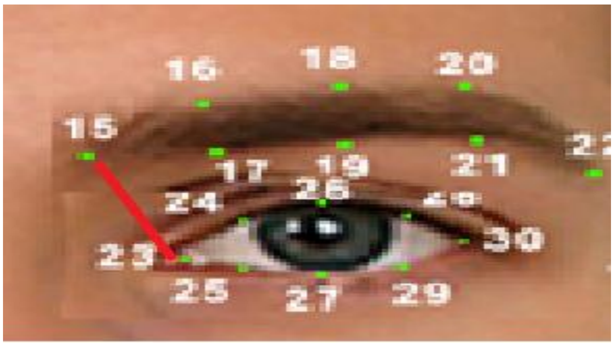


Figure 3.9 Feature No.28.

All features of group five and six are shown in Table 3.5 and Table 3.6 respectively.

Feature No.	Vertical distances between :	
	Point 1	**Point 2**
Feature 25	**15**	**16**
Feature 26	**16**	**18**
Feature 27	**18**	**20**

Table 3.5 Group five of features.

Feature No.	Vertical distances between :	
	Point 1	**Point 2**
Feature 28	**23**	**15**

Table 3.6 Group six of features.

In additional to finding the ethnic group effect of emotion recognition, we use three races from dataset Asian, Caucasian and African. Also we have six emotions for each race angry, happy, sad, fear, disgust and shame. In fact shame is not a basic emotion but we use it because we adhered by (MSFDE) dataset.

We add the race, emotion and gender for each image in dataset manually with all features and save it in text file. This file is use in ANN classifier as inputs and targets of network.

3.5 ANN Classifier

ANN has met a special attention in solving various complex classification problems. It has been widely used in face recognition and age estimation base on facial features [19]. For classification purpose of the emotions, we use ANN of supervised learning based on backprobagation algorithm (see appendix section A.1 for more details of backprbagation algorithm) .We have 28 inputs representing the extracted features and 6 outputs representing the emotions. We have also a hidden layer with 16 nodes selected after try and check to obtain the best results. The used ANN structure is shown in Figure 3.10. In some experiments as we shall see in the next chapter, we add the gender as input.

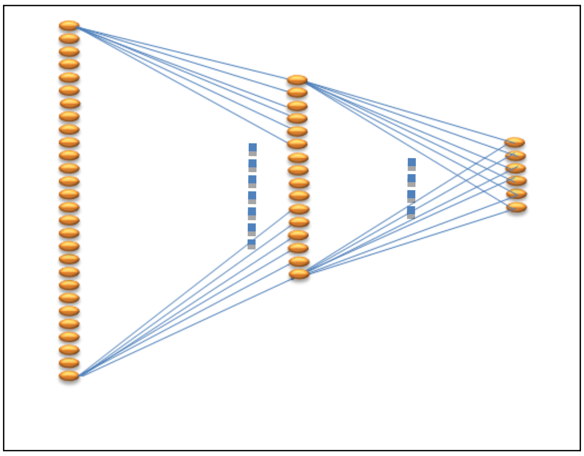


Figure 3.10 Structure of ANN Classifier.

Chapter 4

Implementation and Experimental Results

In this chapter we shall present how each stage of the four main stages of our proposed approach has been developed and implemented. We also present the tools and programming language used to achieve the thesis's target. At the last portion of this chapter, we present the experimental results.

4.1 Face Features Extraction program

We develop a desktop application to apply the first of the four stages of our approach. This application is called Face Features Extraction system (FFE), and it is used to perform four methods or functions as follow:

- Selection of Points.
- Apply some pre-processing method.
- Features Extraction.
- Save all input information for the next stage (Features, Ethnic group, Gender and emotion).

4.1.1 Implementation

We used Microsoft Visual C#.net 2010 language to implement the FFE system. Visual C#.net is used because of the following reasons:

- Widely used in image processing and computer vision field[1].
- Support for many graphics library.
- Friendly user interface.
- Easy to use.
- There is much helpful documentation.

[1] http://en.wikipedia.org/wiki/OpenCV.

The user can extract the features by do the follow steps in FFE program:

- Select an image of human face, and the program automatically applies some preprocessing operations that explained in chapter three. And the image is appears in the image box.

- The user can select 64 points on this image (the order is considerable).

- Save the features into text file. The program save two files, first file contains the coordinates of points; the second file contains the 28 features, gender, ethnic group and emotion.

After that, we combine all features that saved in files by using MS Excel.

Finally we configure a recognition classifier by using Matllab program. And use the combined features as input values and class values.

The LoadPoints is an example of function that used in FFE program. This function shows in appendix chapter in Table A.1.

4.1.2 User Interface

FFE has simply an interface with some basic functions. The interface of FFE program has two windows. First window displays human face elements and show all points for each element, this window is for display only and does not have any interaction with the user. The importance for this window is to help the user to select the points in their order. Figure 4.1 shows the first window.

Figure 4.1 First window for FFE.

Second window is the main window as shown in Figure 4.2, and it has the following parts:

- Browse button: allow the user to select the image of the human face from any physically, virtual or networks drive. Many image extensions are supported (bmp, jpg, tif and gif).
- Workspace image: It is an image control box used to show the selected image, and takes the coordinates of the mouse click that is clicked on the image by the user, and then draws a small green circle on the current coordinate. This window splits by horizontal and vertical lines, and it is useful to draw the points.
- Undo button: remove the last point that has been drawn on the image.
- Reset button: remove all points that were drawn on the image.
- Save button: save two physical files. First one is named (<ethnic-group name>-29.txt), this file contains 28 numeric features, in addition to ethnic group and emotion. The second file's name is (<ethnic-group name>-Points.txt), which contains all the coordinates of the points (64 points) that have been drawn on image.
- Load button: loads all the points from a text file, and draw it on the human face image in the workspace.
- List of Points: list the coordinates of the drawn points.
- Ethnic-Group group: select the ethnicity of the image.
- Emotions group: select the emotion of the image.
- Gender group: select the gender of the image.

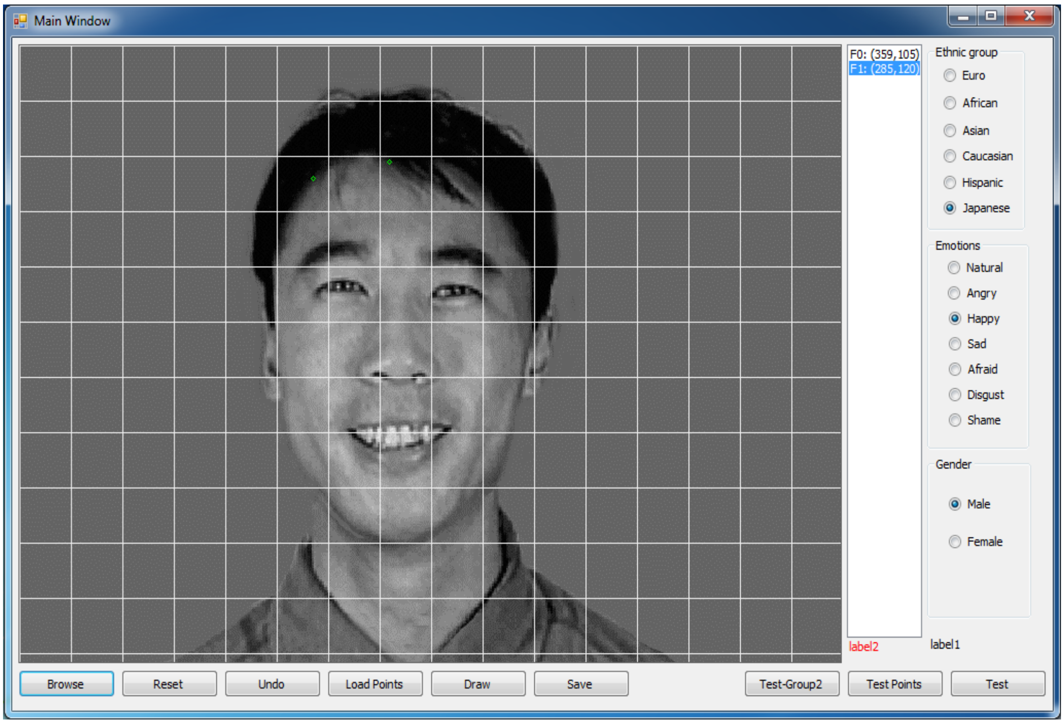


Figure 4.2. The main window in FFE.

4.1.3 Tools and Libraries

We used some tools and libraries in visual studio c#.net to help us in the implementation of FFE. We list below the used tools and libraries:

- OpenCV (Open Source Computer Vision): is a library of programming functions for real time computer vision, and it is free for both academic and commercial use.[1] We used it because it is flexible and easy to use in locating and managing the pixels, also to process some methods in face image like convert gray image or colored image.

- AForge.NET Framework: is a C# framework designed for developers and processing neural networks, genetic algorithms, machine learning, robotics … etc.[2]. We use it to apply some pre-processing on the human face image such as change the size of image.

- EmguCV: is a cross platform .Net wrapper to the Intel OpenCV image processing library. Allowing OpenCV functions to be called from .NET compatible languages such as C#, VB, VC++, IronPython etc.[3] it is powerful

[1] http://opencv.willowgarage.com/wiki/.
[2] http://www.aforgenet.com/.
[3] http://www.emgu.com/wiki/.

in many image processing methods, and we use it to draw circles and lines on the image.

4.1.4 Data Encoder

To convert the given data to numeric values, we developed a small function that converts the text data to numerical data. Three textual data elements we need to encode, ethnic group, emotions and gender.

Three ethnic groups are used African, Asian and Caucasian, and we coded them as shown in Table 4.1:

Ethnic group name	Code
African	2
Asian	3
Caucasian	4

Table 4.1 Ethnic group codes.

We have six emotions in our dataset happy, sad, angry, fear, shame and disgust, and we coded them as shown in Table 4.2:

Emotion name	Code
Happy	1
Sad	2
Angry	3
Fear	4
Shame	5
Disgust	6

Table 4.2 Emotions codes.

The gender is coded also as shown in Table 4.3:

Gender	Code
Male	0
Female	1

Table 4.3 Gender codes.

These codes are added to the features and saved in a text file to use them as inputs and outputs for ANN classifier in the training stage as well as in the testing stage.

4.2 ANN Classifier

To build our classifier we use Mathworks Matlab 2008 using NNTool tool. This tool allows building our specific ANN classifier. First we must select the inputs, targets (classes) and testing data. Second, build the classifier by determining the input, targets, number of layers, number of neurons for each hidden layer and the learning method. We use backprobagation method, and one hidden layers with 16 neurons. In general there are some necessary parameters needed to build our classifier, but we use default values as follow:

- Momentum: 1.00 e+10.
- Epochs: 1000
- Learning rate: 0.1
- Validation check: 1000

Two steps are required when ANN classifier with backpropagation is used for the recognition, training and testing. We firstly train the ANN with a training set and when the ANN gets trained based on a certain error, we use it for testing a new set of data. There are many parameters that are used in the training stage, number of epochs, initial weights, learning rate and the momentum. All these parameters are very important and affect the results of the classifier; Figure 4.3 shows a snap shot for the learning method for the ANN tool.

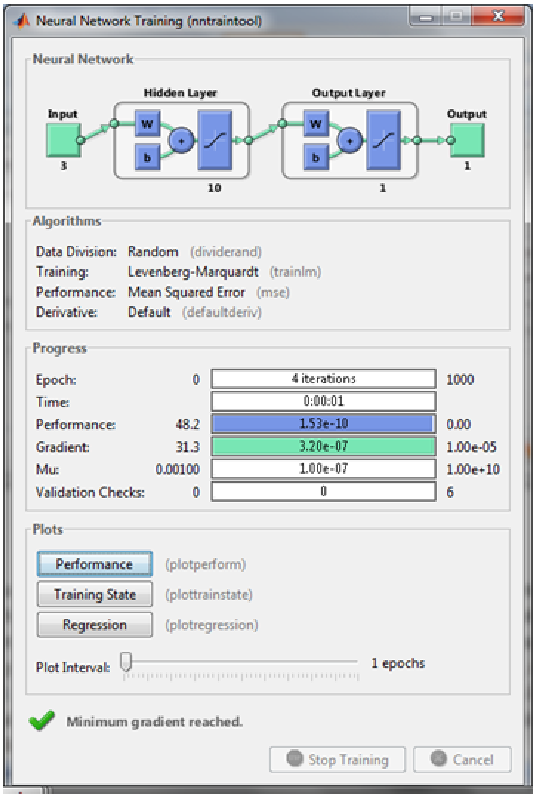


Figure 4.3 Neural Network training.

During the experimentations, we tried to reach to the best initial weights through trying several values of input parameters.

4.3 Tools and programs

Some programs used to complete the implementation and testing our approach and documentation of the thesis:

- Microsoft Visual Studio.net 2010: To develop the FFE system.
- Mathworks Matlab 2010: we use it to build our ANN classifier.
- Microsoft Excel 2010: is spreadsheet software, and it used to separate the target class from the training data, also separate the test data from training data. That is for learning and evaluating our classifier.
- Microsoft Word 2010: for thesis documentation.
- Libraries: Some C# libraries used in implementation of FFE system: OpenCV, EmguCV and AFOGE.net.

4.4 Datasets

Montreal Set of Facial Displays of Emotion (MSFDE) dataset is selected to test and evaluate our approach. This data set includes 224 gray images of females and males faces and four ethnic groups African, Asian, Hispanic and Caucasian, these images describe seven facial expressions nature, happy, sadness, angry, fear, shame and disgust. Each ethnic group contains two genders; each gender contains four persons, each person has seven emotions, but each person has a one image for nature emotion, so we ignored this emotion because we need more than one image for each emotion to be able to train well the ANN. This means we use only six emotions.

However, this dataset has many advantages:

1- There are four ethnic groups Asian, African, Caucasian and Hispanic.

2- Support female and male images.

3- The images have a good quality and resolutions: 800 x 560 with 160 DPI.

Despite the diversity of the dataset, but it has three major disadvantages:

1- Number of images relatively small. We need a large number of images for the purpose of training (learning process). After that we need another set of images to test and evaluate the learned classifier.

2- The number of people of the same race is not enough to express all the categories of race.

3- Surprise is one of the six basic emotions; this dataset ignores this emotion and replace it with shame emotion.

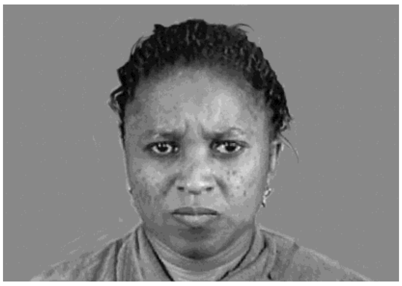
So we are training our ANN classifier by using three persons out of four for each ethnic group, and testing the classifier by the fourth one in each ethnic group.

However, we found that this dataset is one of the best datasets available to evaluate and test our approach. Figure 4.4 shows a number of examples from MSFDE dataset.

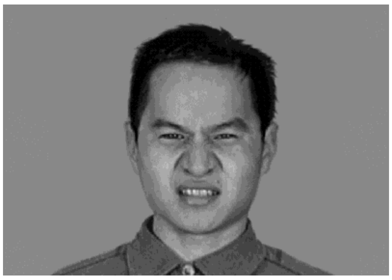
4.5 Example

In this section, we present an example of human face images to summarize all methods and steps that are applied in our approach.

First step we run FFE system and select the required image to be uses for training or testing. We then select all points on the image, and choose the emotion of the image, gender and ethnic group. After that, the system calculates all features automatically and save it in a file called features file. Figure 4.5 shows an example of the features and information that would be saves from FFE system. This file contains 31 values, the first 28 values represent the features that calculated by FFE, the 29th value represents the gender code, the 30th value represents the ethnic group code, and the final code is the emotion code.



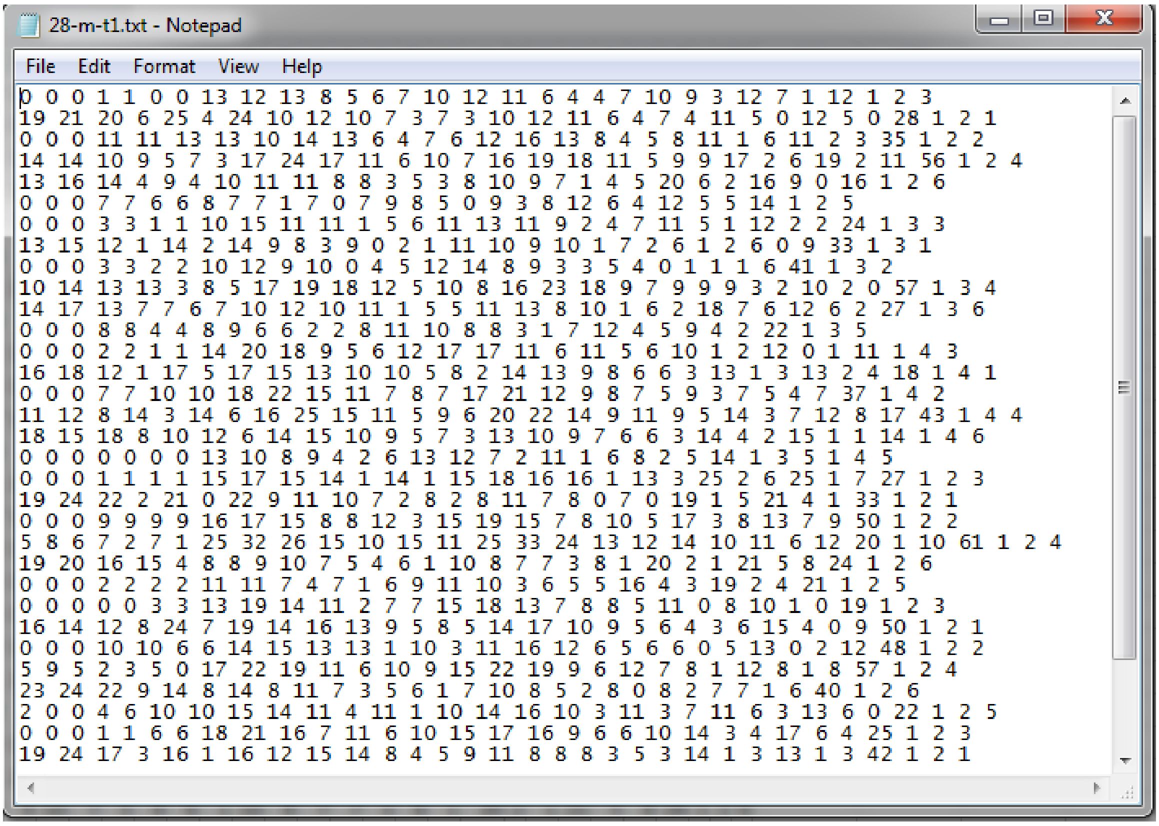
African, Female and Angry



Asian, Male and Disgust

Caucasian, Male and Happy

Figure 4.4 Example of MSFDE Dataset.



28-m-t1.txt - Notepad

File Edit Format View Help

```
0 0 0 1 1 0 0 13 12 13 8 5 6 7 10 12 11 6 4 4 7 10 9 3 12 7 1 12 1 2 3
19 21 20 6 25 4 24 10 12 10 7 3 7 3 10 12 11 6 4 7 4 11 5 0 12 5 0 28 1 2 1
0 0 0 11 11 13 13 10 14 13 6 4 7 6 12 16 13 8 4 5 8 11 1 6 11 2 3 35 1 2 2
14 14 10 9 5 7 3 17 24 17 11 6 10 7 16 19 18 11 5 9 9 17 2 6 19 2 11 56 1 2 4
13 16 14 4 9 4 10 11 11 8 8 3 5 3 8 10 9 7 1 4 5 20 6 2 16 9 0 16 1 2 6
0 0 0 7 7 6 6 8 7 7 1 7 0 7 9 8 5 0 9 3 8 12 6 4 12 5 5 14 1 2 5
0 0 0 3 3 1 1 10 15 11 11 1 5 6 11 13 11 9 2 4 7 11 5 1 12 2 2 24 1 3 3
13 15 12 1 14 2 14 9 8 3 9 0 2 1 11 10 9 10 1 7 2 6 1 2 6 0 9 33 1 3 1
0 0 0 3 3 2 2 10 12 9 10 0 4 5 12 14 8 9 3 3 5 4 0 1 1 1 6 41 1 3 2
10 14 13 13 3 8 5 17 19 18 12 5 10 8 16 23 18 9 7 9 9 9 3 2 10 2 0 57 1 3 4
14 17 13 7 7 6 7 10 12 10 11 1 5 5 11 13 8 10 1 6 2 18 7 6 12 6 2 27 1 3 6
0 0 0 8 8 4 4 8 9 6 6 2 2 8 11 10 8 8 3 1 7 12 4 5 9 4 2 22 1 3 5
0 0 0 2 2 1 1 14 20 18 9 5 6 12 17 17 11 6 11 5 6 10 1 2 12 0 1 11 1 4 3
16 18 12 1 17 5 17 15 13 10 10 5 8 2 14 13 9 8 6 6 3 13 1 3 13 2 4 18 1 4 1
0 0 0 7 7 10 10 18 22 15 11 7 8 7 17 21 12 9 8 7 5 9 3 7 5 4 7 37 1 4 2
11 12 8 14 3 14 6 16 25 15 11 5 9 6 20 22 14 9 11 9 5 14 3 7 12 8 17 43 1 4 4
18 15 18 8 10 12 6 14 15 10 9 5 7 3 13 10 9 7 6 6 3 14 4 2 15 1 1 14 1 4 6
0 0 0 0 0 0 0 13 10 8 9 4 2 6 13 12 7 2 11 1 6 8 2 5 14 1 3 5 1 4 5
0 0 0 1 1 1 1 15 17 15 14 1 14 1 15 18 16 16 1 13 3 25 2 6 25 1 7 27 1 2 3
19 24 22 2 21 0 22 9 11 10 7 2 8 2 8 11 7 8 0 7 0 19 1 5 21 4 1 33 1 2 1
0 0 0 9 9 9 9 16 17 15 8 8 12 3 15 19 15 7 8 10 5 17 3 8 13 7 9 50 1 2 2
5 8 6 7 2 7 1 25 32 26 15 10 15 11 25 33 24 13 12 14 10 11 6 12 20 1 10 61 1 2 4
19 20 16 15 4 8 8 9 10 7 5 4 6 1 10 8 7 7 3 8 1 20 2 1 21 5 8 24 1 2 6
0 0 0 2 2 2 2 11 11 7 4 7 1 6 9 11 10 3 6 5 5 16 4 3 19 2 4 21 1 2 5
0 0 0 0 0 3 3 13 19 14 11 2 7 7 15 18 13 7 8 8 5 11 0 8 10 1 0 19 1 2 3
16 14 12 8 24 7 19 14 16 13 9 5 8 5 14 17 10 9 5 6 4 3 6 15 4 0 9 50 1 2 1
0 0 0 10 10 6 6 14 15 13 13 1 10 3 11 16 12 6 5 6 6 0 5 13 0 2 12 48 1 2 2
5 9 5 2 3 5 0 17 22 19 11 6 10 9 15 22 19 9 6 12 7 8 1 12 8 1 8 57 1 2 4
23 24 22 9 14 8 14 8 11 7 3 5 6 1 7 10 8 5 2 8 0 8 2 7 7 1 6 40 1 2 6
2 0 0 4 6 10 10 15 14 11 4 11 1 10 14 16 10 3 11 3 7 11 6 3 13 6 0 22 1 2 5
0 0 0 1 1 6 6 18 21 16 7 11 6 10 15 17 16 9 6 6 10 14 3 4 17 6 4 25 1 2 3
19 24 17 3 16 1 16 12 15 14 8 4 5 9 11 8 8 8 3 5 3 14 1 3 13 1 3 42 1 2 1
```

Figure 4.5 Example of features file.

Next step is importing features file into excel sheet and split the data into three parts training data, target data and test data. Test data is the features of the image to be tested, target data is the file that has the output of each image (emotion type) and the test data contains the features for the images involved in the testing stage. After extraction of the features, run Matlab program and build ANN classifier using nntool control, then import all training data that is in excel sheet, and use it for training the classifier. Finally, we test the classifier using the data existing in the test data file.

4.6 Experiment Results and Evaluations

In this section we discuss the results of experiments and compare it with other related works, finally we evaluate our approach based on the comparisons.

We did two groups of experiments. Each group tests the impact of one factor on the accuracy of emotion recognition.

4.6.1 First group of experiments

This experiment is studying the impact of gender in the accuracy of emotion recognition. The males or females can be from any ethnic group of the three used ethnic groups. We have two neural networks, one for the males and the other one for the females. Each of the networks has three layers (inputs, one hidden layer and outputs), in the hidden layer there are 16 neurons.

The first experiment is for the male images. We trained the classifier by using 54 male images. These images representing six emotions for three races and each race contain three persons. We Test the classifier using 18 male images representing six emotions for three persons of three races. The results of this experiment are shown in Table 4.4.

Emotion	No.Correct	No.Error	Ratio
Happy	3	0	100%
Sad	2	1	66.7%
Angry	2	1	66.7%
Fear	2	1	66.7%
Disgust	3	0	100%
Shame	3	0	100%
Total	**15**	**3**	**83.3%**

Table 4.4 Results of male experiment.

We note that, the error appears in fear, sad and angry emotions, because the distances of each feature for each emotion are too close from each to another, which leads to appearance of some errors in the recognition of these emotions. So it must be accurate relatively in selecting points to avoid falling into these errors.

Second experiment is for female images. We trained the classifier by using 54 female images. These images represent six emotions for three races and each race contains three women. We test the classifier by using 18 female images representing six emotions for three women of three races. The results of this experiment are shown in Table 4.5.

Emotion	No.Correct	No.Error	Ratio
Happy	3	0	100%
Sad	3	0	100%
Angry	3	0	100%
Fear	2	1	66.7%
Disgust	3	0	100%
Shame	2	1	66.7%
Total	**16**	**2**	**88.9%**

Table 4.5 Results of female experiment.

To clarify these results let's consider a form of female expression of happy on the different races, we find that the most elements of face have a same facial expression, or at least be very close to each other. As shown in Figure 4.6. we note that, the mouth is open significantly in the Caucasian race, and less in the rest of the races.

Figure 4.6 Examples of female happy. (A) African. (B) Asian. (C) Caucasian.

Similarly, if we consider also the facial expressions of fear in the case of different races, we find that this similarity is relatively less considering the similarity of happy. The mouth is more open as possible when Africans in the expression of fear. As shown in Figure 4.7.

Figure 4.7 Examples of female fear. (A) African. (B) Asian. (C) Caucasian.

From Figure 4.6 and Figure 4.7, we can conclude here that the shape of the mouth in an expression of happy in the Asian females is very close to the shape of the mouth in an expression of fear in African females. So some of the features will be similar when there is fear and happy. This causes some errors in the recognition of fear as happy.

Last experiment is for studying the accuracy of emotion recognition for all the images without ethnicity feature but with gender feature as input. This experiment has a neural network as a classifier, Again the ANN has three layers (inputs, one hidden and outputs), and 16 neurons in the hidden layer.

We used 108 images for the training representing six emotions. The images have three males and three females. For testing, we used 36 images. The results of this experiment are shown in Table 4.6.

Emotion	No.Correct	No.Error	Ratio
Happy	6	0	100%
Sad	3	3	50%
Angry	5	1	83.3%
Fear	2	4	33.3%
Disgust	6	0	100%
Shame	5	1	83.3%
Total	**27**	**9**	**75%**

Table 4.6 Results of all races with gender.

From Figure 4.8 we can note that the expression of emotion of anger to all ethnic groups given gender is converged significantly. This can explain that the difference between anger expressions is not much for the persons of the same gender regardless of the race.

Figure 4.8 Examples of Angry from different races. (A,B) African. (C,D) Asian. (E,F) Caucasian.

4.6.2 Second group of experiments

In this stage of experimentations, we study the impact of ethnic group (race) in the accuracy of emotion recognition. With three ethnic groups, we have three experiments. Each experiment has a neural network as a classifier, and each neural network has three layers (inputs, one hidden layer and outputs), where there are 16 neurons in the hidden layer except Asian network has 17 neurons. We make variation in the number of the neurons in the hidden layer because we obtained better results with 17 neurons in the hidden layer for recognition of Asians. Again here we have the same six emotions.

4.6.2.1 Asian experiments

This experiment is for the Asian race and it has four sub experiments. The first one is to test the impact of ethnic group with male in recognizing the type of emotion. The second one is to test the impact of ethnic group with female in recognizing the type of emotion. The third experiment is to test the impact of ethnic group with gender (male or female) as additional input with the face features inputs. The fourth experiment is to test the impact of ethnic group without including the gender as input. For sure the last two experiments (the third and the fourth) include all the images of males and females.

a. Male Experiment

This experiment is to test the impact of knowing the Asian ethnic group in the accuracy of emotion recognition for Asian male. We used 18 images for training representing six emotions of three male persons, and test the classifier by using six male images representing six emotions for one Asian male person. The results of this experiment are shown in Table 4.7.

Emotion	No.Correct	No.Error	Ratio
Happy	1	0	100%
Sad	1	0	100%
Angry	1	0	100%
Fear	1	0	100%
Disgust	1	0	100%
Shame	1	0	100%
Total	**6**	**0**	100%

Table 4.7 Results of Asian male experiment.

b. Female Experiment

This experiment is to test the impact of knowing the Asian ethnic group in the accuracy of emotion recognition for Asian female. We used 18 images for training representing six emotions of three females, and test the classifier by using six females' images representing six emotions for one Asian female woman. The results of this experiment are shown in Table 4.8. The error appears in fear emotion, because the distances in fear emotion are too close from sad and angry emotions.

Emotion	No.Correct	No.Error	Ratio
Happy	1	0	100%
Sad	1	0	100%
Angry	1	0	100%
Fear	0	1	0%
Disgust	1	0	100%
Shame	1	0	100%
Total	5	1	83.3%

Table 4.8 Results of Asian female experiment.

c. Male and Female Experiment with gender

This experiment is to test the impact of ethnic group in the accuracy of emotion recognition for Asian male and female with known gender, this means we add the gender as input as any feature input. We used 36 images for training representing six emotions of three females and three males, and test the classifier by using 12 images representing six emotions for Asian males and females. The results of this experiment are shown in Table 4.9.

Emotion	No.Correct	No.Error	Ratio
Happy	2	0	100%
Sad	2	0	100%
Angry	1	1	50%
Fear	2	0	100%
Disgust	2	0	100%
Shame	1	1	50%
Total	**10**	**2**	83.3%

Table 4.9 Results of Asian race considering the gender.

Two errors were occurs. First error in angry of female, and we explained previously why this error happened. Second error in shame of male, and as we can see in Figure 4.9 and Figure 4.10. The mouth shape is totally different in cases of shame and disgust emotions, but other elements features are very closed specially eyebrows features then eyes features.

Figure 4.9 Examples of shame emotion for Asian race.(A) Female (B) Male.

Figure 4.10 Examples of disgust emotion for Asian race. (A) Female (B) Male.

d. Male and Female Experiment without gender

This experiment is to test the impact of Asian ethnic group in the accuracy of emotion recognition for Asian male and female without giving the gender as input. We used 36 images for training representing six emotions of three females and three males, and test the classifier by using 12 images representing six emotions for Asian males and females. The results of this experiment are shown in Table 4.10. The explanation of the results is similar to what has been explained before.

Emotion	No.Correct	No.Error	Ratio
Happy	2	0	100%
Sad	1	1	50%
Angry	1	1	50%
Fear	1	1	50%
Disgust	2	0	100%
Shame	2	0	100%
Total	**9**	**3**	75%

Table 4.10 Results of Asian race without considering the gender.

4.6.2.2 Caucasian experiments

This experiment is for Caucasian race and it has four sub experiments. The first one is to test the impact of ethnic group with male in recognizing the type of emotion. The second one is to test the impact of ethnic group with female in recognizing the type of emotion. The third experiment is to test the impact of ethnic group with gender as additional input with the face features inputs. The fourth experiment is to test the impact of ethnic group without including the gender as input.

A. Male Experiment

This experiment is to test the impact of knowing the Caucasian ethnic group in the accuracy of emotion recognition for the Caucasian male. We used 18 images for training representing six emotions of three male persons, and test the classifier by using six male images representing six emotions for one Caucasian male person. The results of this experiment are shown in Table 4.11.

Emotion	No.Correct	No.Error	Ratio
Happy	1	0	100%
Sad	1	0	100%
Angry	1	0	100%
Fear	1	0	100%
Disgust	1	0	100%
Shame	1	0	100%
Total	**6**	**0**	100%

Table 4.11 Results of Caucasian male experiment.

B. Female Experiment

This experiment is to test the impact of knowing the Caucasian ethnic group in the accuracy of emotion recognition for Caucasian female. We used 18 images for training representing six emotions of three females, and test the classifier by using six females images representing six emotions for one Caucasian female woman. The results of this experiment are shown in Table 4.12.

Emotion	No.Correct	No.Error	Ratio
Happy	1	0	100%
Sad	1	0	100%
Angry	1	0	100%
Fear	1	0	100%
Disgust	1	0	100%
Shame	0	1	0%
Total	**5**	**1**	83.3%

Table 4.12 Results of Caucasian female experiment.

C. Male and Female Experiment with gender

This experiment is to test the impact of ethnic group in the accuracy of emotion recognition for Caucasian male and female with known gender, this means we add the gender as input as any feature input. We used 36 images for training representing six emotions of three females and three males, and test the classifier by using 12 images representing six emotions for Caucasian males and females. The results of this experiment are shown in Table 4.13.

Emotion	No.Correct	No.Error	Ratio
Happy	2	0	100%
Sad	2	0	100%
Angry	2	0	100%
Fear	1	1	50%
Disgust	2	0	100%
Shame	2	0	100%
Total	**11**	**1**	**91.7%**

Table 4.13 Results of Caucasian race considering the gender.

D. Male and Female Experiment without gender

This experiment is to test the impact of Caucasian ethnic group in the accuracy of emotion recognition for Caucasian male and female without giving the gender as input. We used 36 images for training representing six emotions of three females and three males, and test the classifier by using 12 images representing six emotions for Caucasian males and females. The results of this experiment are shown in Table 4.14.

Emotion	No.Correct	No.Error	Ratio
Happy	2	0	100%
Sad	2	0	100%
Angry	2	0	100%
Fear	2	0	100%
Disgust	2	0	100%
Shame	1	1	50%
Total	**11**	**1**	**91.7%**

Table 4.14 Results of Caucasian race without considering the gender.

4.6.2.3 African experiments

This experiment's for African race and it has four sub experiments. The first one is to test the impact of ethnic group with male in recognizing the type of emotion. The second one is to test the impact of ethnic group with female in recognizing the type of emotion. The third experiment is to test the impact of ethnic group with gender (male or female) as additional input with the face features inputs. The fourth experiment is to test the impact of ethnic group without including the gender as input. For sure the last two experiments (the third and the fourth) include all the images of males and females.

A. Male Experiment

This experiment is to test the impact of knowing the African ethnic group in the accuracy of emotion recognition for African male. We used 18 images for training representing six emotions of three male persons, and test the classifier by using six male images representing six emotions for one African male person. The results of this experiment are shown in Table 4.15.

Emotion	No.Correct	No.Error	Ratio
Happy	1	0	100%
Sad	1	0	100%
Angry	0	1	0%
Fear	0	1	0%
Disgust	1	0	100%
Shame	1	0	100%
Total	**4**	**2**	66.67%

Table 4.15 Results of African male experiment.

B. Female Experiment

This experiment is to test the impact of knowing the African ethnic group in the accuracy of emotion recognition for African female. We used 18 images for training representing six emotions of three females, and test the classifier by using six females images representing six emotions for one African female woman. The results of this experiment are shown in Table 4.16.

B. Emotion	No.Correct	No.Error	Ratio
Happy	1	0	100%
Sad	1	0	100%
Angry	1	0	100%
Fear	0	1	0%
Disgust	0	1	0%
Shame	1	0	100%
Total	**4**	**2**	66.67%

Table 4.16 Results of African female experiment.

For the first time, there is an error in the disgust emotion. We can interpretation of this by observing the images in Figure 4.11. We find that the forms of elements of the face do not share significantly in all images. In another word the mouth in image (B) is completely different from the rest of images. In addition, the eye in image (D) is completely different from the rest of images. Feature number 28 (distance between the beginning of eyebrow and the beginning of eye) is similar between image (A) and image (B) and different in other images.

Therefore obvious that the accuracy is low in the recognition, and usually this case is occurs in the absence of images is sufficient.

Table 4.11 Disgust emotion for African female.

C. Male and Female Experiment with gender

This experiment is to test the impact of ethnic group in the accuracy of emotion recognition for African male and female with known gender, this means we add the gender as input as any feature input. We used 36 images for training representing six emotions of three females and three males, and test the classifier by using 12 images representing six emotions for African males and females. The results of this experiment are shown in Table 4.17.

C. Emotion	No.Correct	No.Error	Ratio
Happy	1	1	50%
Sad	2	0	100%
Angry	2	0	100%
Fear	2	0	100%
Disgust	2	0	100%
Shame	0	2	0%
Total	**9**	**3**	**75%**

Table 4.17 Results of African race considering the gender.

The first error is in happy emotion. If we see the happy emotion of African in Figure 4.12, and disgust emotion in Figure 4.13, we can note that the differences in elements features (eye and eyebrows) between happy and disgust are too small, and some of mouth features are only considered. Therefore, it is definite to get an error in these emotions. Also Figure 4.14 shows the values of features for happy and disgust emotions of African race, where we find the similarity of these values which are highlighted by gray lines. We can avoid the problem by increasing the number of examples of African race.

Figure 4.12 Happy emotion for African race.

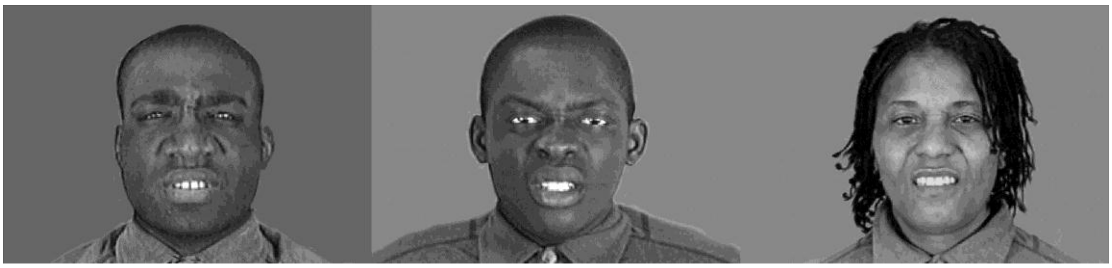


Figure 4.13 Disgust emotion for African race.

1	19	24	22	2	21	0	22	9	11	10	7	2	8	2	8	11	7	8	0	7	0	19	1	5	21	4	1	33	1	2	1
2	16	14	12	8	24	7	19	14	16	13	9	5	8	5	14	17	10	9	5	6	4	3	6	15	4	0	9	50	1	2	1
3	19	24	17	3	16	1	16	12	15	14	8	4	5	9	11	8	8	8	3	5	3	14	1	3	13	1	3	42	1	2	1
4	13	17	22	6	19	2	24	11	16	14	9	2	7	7	18	15	8	11	7	5	3	15	1	8	11	2	5	31	2	2	1
5	16	20	15	2	18	3	18	12	14	10	7	5	4	6	11	13	8	6	5	5	3	7	5	15	2	4	10	54	2	2	1
6	15	24	18	14	29	6	24	10	13	7	8	2	4	3	9	12	7	9	0	6	1	14	6	2	14	4	1	34	2	2	1
7	19	21	20	6	25	4	24	10	12	10	7	3	7	3	10	12	11	6	4	7	4	11	5	0	12	5	0	28	1	2	1
8	19	21	19	2	17	1	18	10	11	10	8	2	8	2	9	10	7	7	2	6	1	11	1	7	13	2	4	38	2	2	1
9	19	20	14	15	4	7	7	9	10	7	5	4	6	1	10	8	7	7	3	8	1	20	2	1	21	5	8	24	1	2	6
10	23	24	22	9	14	8	14	8	11	7	3	5	6	1	7	10	8	5	2	8	0	8	2	7	7	1	6	40	1	2	6
11	23	26	25	11	12	13	12	11	13	12	6	5	8	4	9	11	11	9	0	8	3	13	6	8	13	#	7	29	1	2	6
12	9	10	10	2	7	1	9	13	22	12	6	7	4	8	15	17	9	7	8	2	7	16	7	1	19	6	2	15	2	2	6
13	15	11	12	2	13	4	8	9	10	7	9	0	4	3	8	11	8	6	2	5	3	8	9	3	17	2	4	27	2	2	6
14	17	17	18	0	17	4	14	10	12	9	9	1	6	3	6	9	7	8	2	6	1	15	5	2	14	2	2	28	2	2	6
15	15	21	17	4	11	6	11	11	11	8	8	3	5	3	8	10	9	7	1	4	5	20	6	2	16	9	0	16	1	2	6
16	26	25	23	13	13	8	15	8	13	8	8	0	6	2	12	11	8	7	5	6	2	12	0	1	13	2	1	31	2	2	6

Figure 4.14. Features of disgust and happy emotion for African.(rows from 1 to 8 are for happy emotion, and others for disgust emotion)

D. Male and Female Experiment without gender

This experiment is to test the impact of African ethnic group in the accuracy of emotion recognition for African male and female without giving the gender as input. We used 36 images for training representing six emotions of three females and three males, and test the classifier by using 12 images representing six emotions for African males and females. The results of this experiment are shown in Table 4.18.

Emotion	No.Correct	No.Error	Ratio
Happy	2	0	100%
Sad	0	2	0%
Angry	1	1	50%
Fear	1	1	50%
Disgust	2	0	100%
Shame	2	0	100%
Total	**8**	**4**	**66.67%**

Table 4.18 Results of African race without considering the gender.

To summarizing the accuracy of emotion recognition considering the ethnic group and gender, we combine the results for each ethnic and calculate the percentages of accuracy. The total tests images are 36 images, and we got 83.3% accuracy in total as shown in Table 4.19.

Emotion	No.Correct	No.Error	Ratio
Happy	5	1	83.3%
Sad	6	0	100%
Angry	5	1	83.3%
Fear	5	1	83.3%
Disgust	6	0	100%
Shame	3	3	50.0%
Total	**30**	**6**	**83.3%**

Table 4.19 Accuracy of emotion recognition considering the ethnic group and gender.

Therefore, the African race affected on the accuracy of happy and shame emotion, and makes it less than the accuracy of recognition regardless the ethnic group. We can enhancement these results by increasing the number of African examples that will increase the accuracy of recognition.

4.7 Discussion

While we have two groups of experiments, we split our discussion into two parts as following:

4.7.1 The First Experimental Group

- The accuracy of the expression of emotions in women is more than men (83.3% for men and 88.9% for women). This result is consistent with some previous research [13].
- The accuracy of the expression of happy and disgust are more accurate expression than the rest of the emotions for both genders (100% accuracy). This result is also consistent with some previous research [31].
- The accuracy of the expression of fear is a less accurate expression than the rest of the emotions for both genders (66.7% accuracy). This result is also consistent with some previous research [31].
- Regardless of race, the accuracy of the expression of happy and disgust has the best accuracy.

All previous points are shown in Figure 4.15 and Figure 4.16.

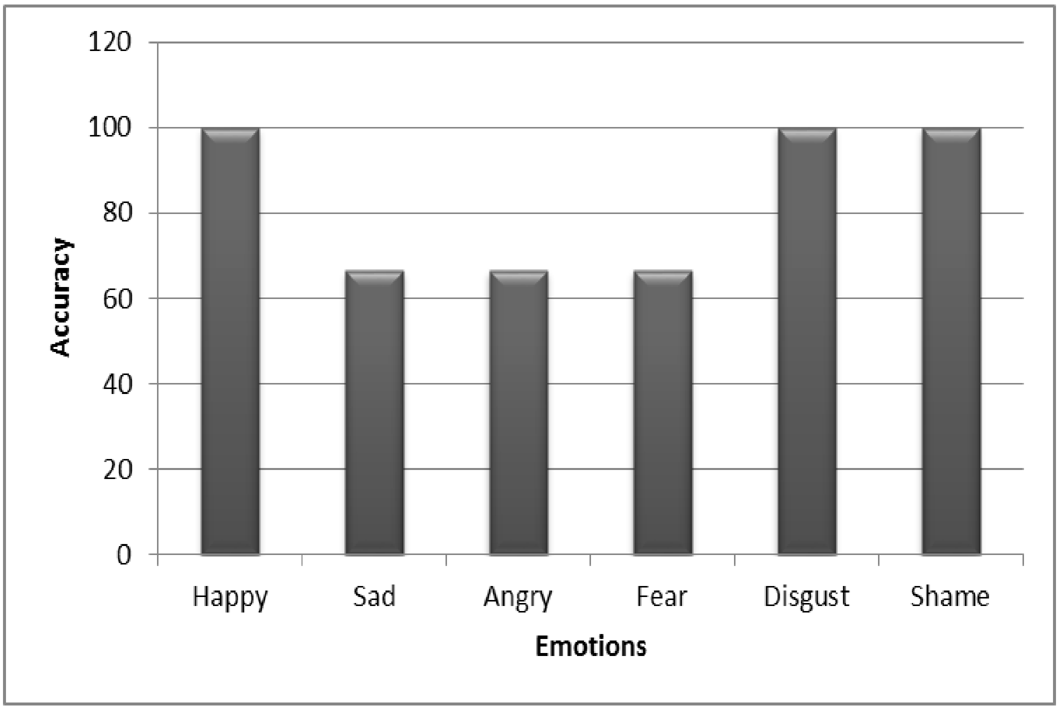


Figure 4.15 Accuracy of emotions for male without considering the ethnic group.

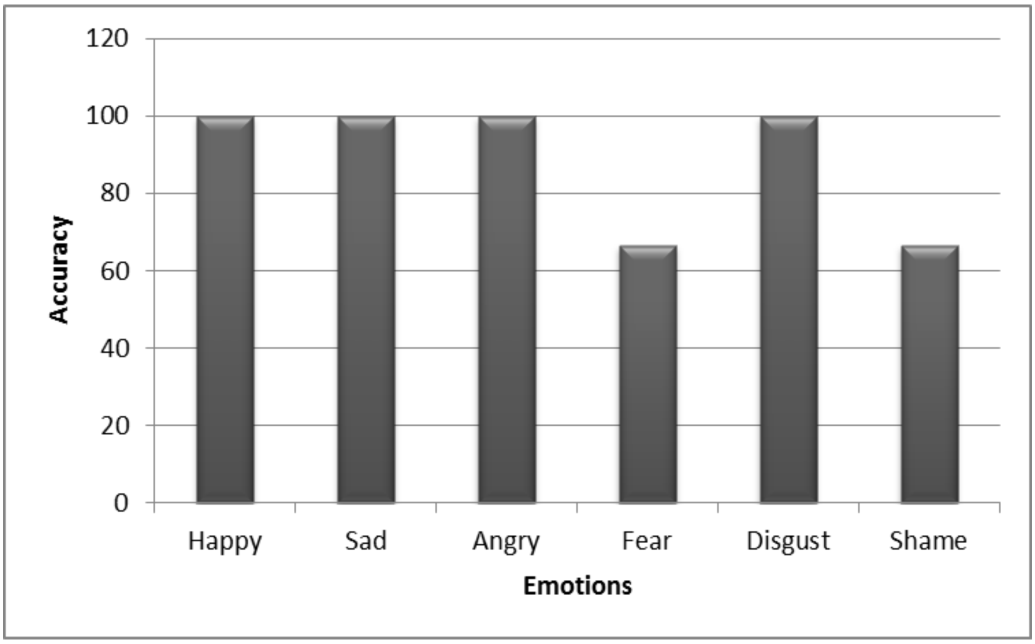


Figure 4.16 Accuracy of emotions for female without considering the ethnic group.

4.7.2 The Second Experimental Group

- The accuracy of the expression of emotions in men is more than of that in women in Asian and Caucasian races.

- The accuracy of the expression of fear is a less accurate expression than the rest of the emotions in Asian female and both gender of Africans.
- Known the gender gives more accuracy for the emotions expression in Asian (83.3% with gender and 75% without gender) and African races (75% with gender and 66.67% without gender) .
- The accuracy of the expression of happy and disgust are more accurate expressions than the rest of the emotions in Asian and Caucasian for both genders (100% accuracy).
- The accuracy of the expression of shame is a less accurate expression than the rest of the emotions in Caucasian female.
- There is no difference to know the gender or not in accuracy of emotions recognition in Caucasian (91.7% accuracy).
- Men and women in African race have the same accuracy of emotions recognition (66.67% accuracy).
- The accuracy of the expression of disgust is more accurate expression than the rest of the emotions in Africans if we know gender or not.
- The African race has the lowest accuracy, and Caucasian race has the highest accuracy of emotions recognition considering or regardless of the gender.
- In addition, there are many comparisons between the accuracy of recognizing some of the emotions between the race and another. For example, with Asians, the anger expression recognized with less accuracy than with Caucasians. Also with Africans, it is more difficult in identifying sadness expressions as shown in Figure 4.6. All of these results are consistent with some previous research [4].
- Considering the race, the accuracy of recognition for shame emotion has a lowest percentage as shown in Figure 4.17.
- Determine the gender increases the accuracy of emotion recognition.
- Impact of ethnic group on the accuracy of emotions recognition is a positive. The accuracy of emotion recognition considering ethnic group is 83.3% and 75% regardless ethnic group. This result shown in Figures from 4.17 to 4.20.
- Disgust emotion has best accuracy of recognition.
- Fear emotion has the worst accuracy of recognition.

To conclude these results, we were able to obtain a high accuracy rate in recognition of some emotions, which is the highest accuracy in disgust emotion, where we have reached 100% for both cases (with or without identification of race). As well as we got on the average percentage an accuracy of emotions recognition a 83.3% in the case of determining race, and 75% regardless of the ethnic group, thus improved the accuracy rate by almost 8%. This percentage is relatively good compared to some of the findings of some other research if we consider the diversity of races. Some of researches obtained accuracy of emotion recognition on an average 74% [24], 73.3% [6] and 82.5% [31]. These researches dealing only with one ethnic group that depending on the used dataset.

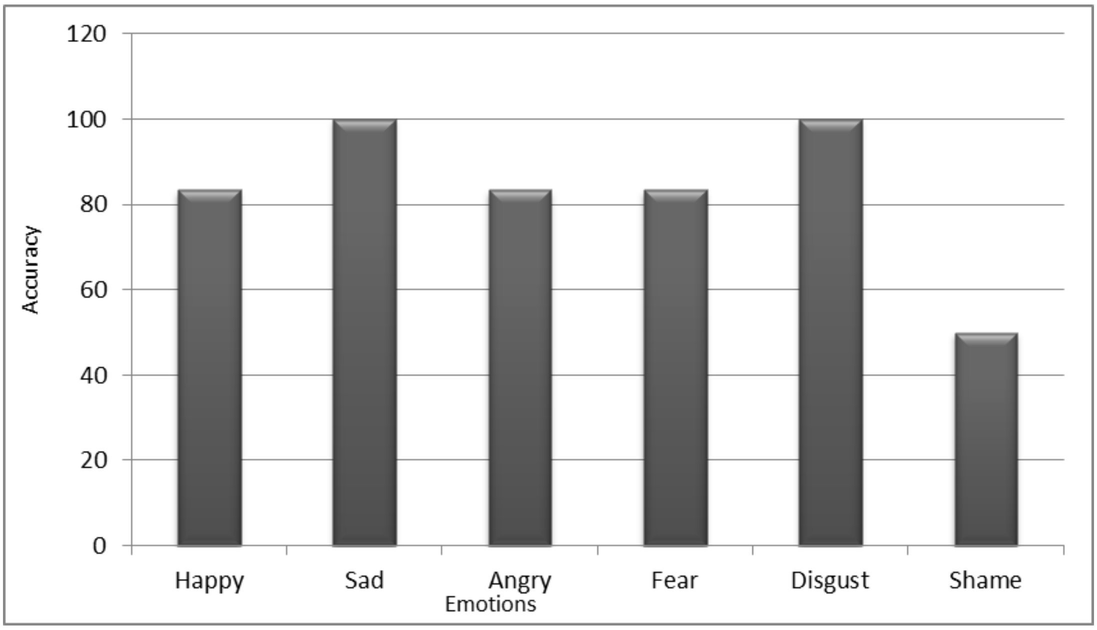


Figure 4.17 Classification accuracy of emotions considering the ethnic group with gender as feature. (Total accuracy is 83.3%)

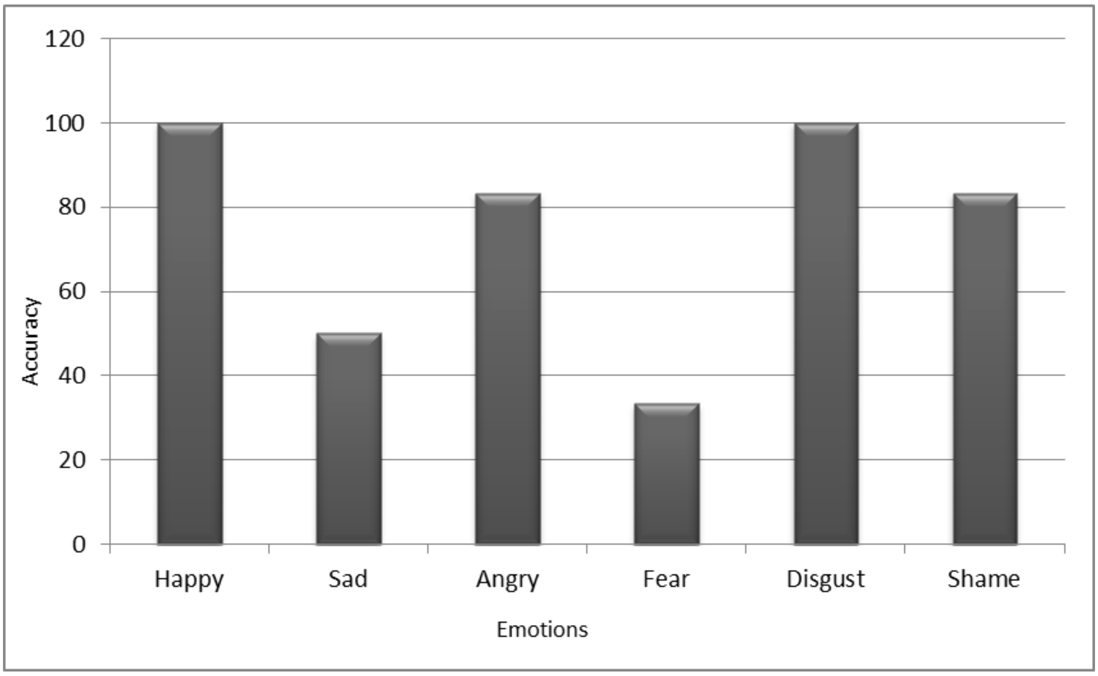


Figure 4.18 Classification accuracy regardless the ethnic group with gender as feature.

(Total accuracy 75%)

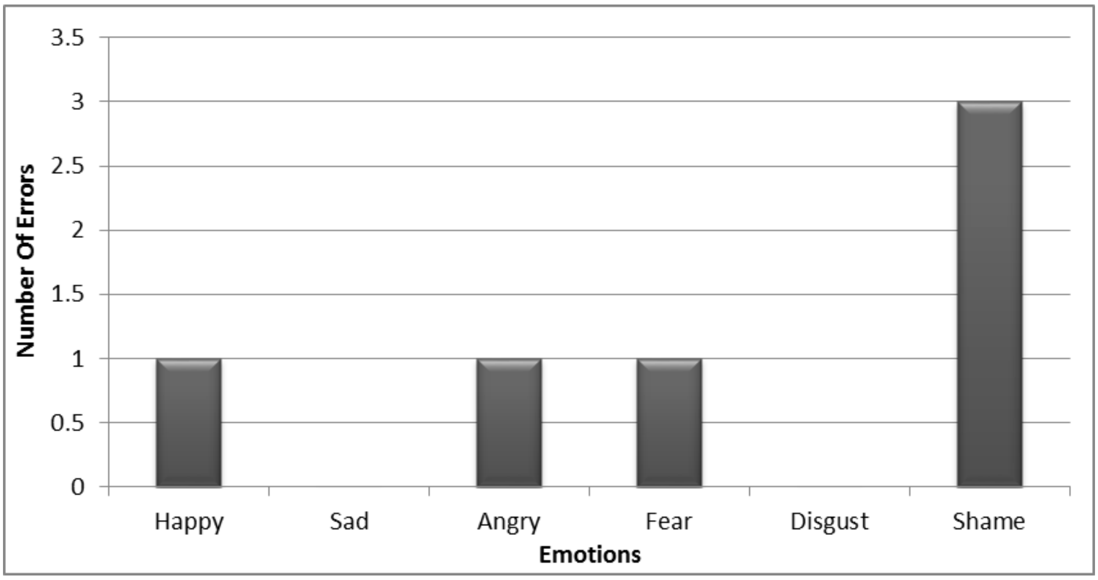


Figure 4.19 Number of errors for each emotion considering the ethnic group.

(Maximum errors are three errors)

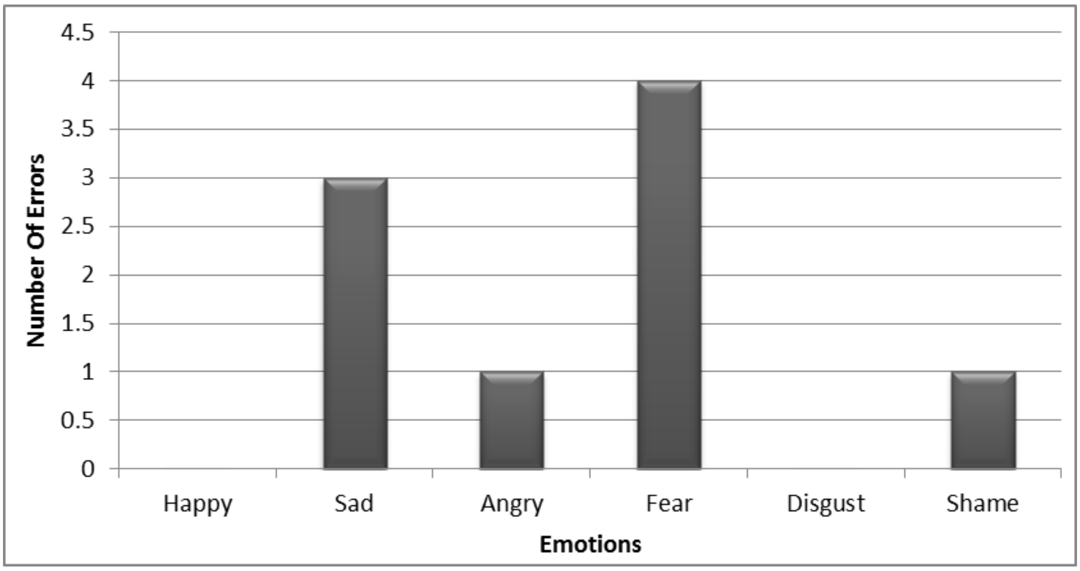


Figure 4.20 Number of errors for each emotion regardless of the ethnic group.

(Maximum errors are four errors)

Chapter 5

Conclusion and Future Works

5.1 Conclusion

In this thesis, we investigated the recognition for human emotion based on face expressions, ethnic group and gender using Backpropagation Neural Networks.

Our approach has five stages divided into two steps. First step has the first four stages, getting the image, preprocessing, features selection and features extraction. The second step has the last stage, which is concerned with building an ANN classifier and applies recognition process. To complete step one; we developed a desktop program called FFE system, and for step two we used Matlab program to build an ANN classifier. After we complete step one the results was saved in a text file , and then exported to ANN classifier and get the final results.

We use MSFDE data set to test and evaluate our approach. In addition, we used two groups of experimentations. First group is to test the impact of the gender on the accuracy of emotions recognition regardless of the ethnic group (some experiments we use each gender separately, and some other experiments we use two the genders together). Second group is to test the impact of the ethnic group on the accuracy of emotions recognition by using three ethnic groups Asian, Caucasian and African (with gender and without gender).

We were able to reach a large number of results concerning the influence of both ethnic group and gender on the accuracy of the recognition of emotions based on facial features. Perhaps the most prominent of these results is that the race has a positive impact on the accuracy of identifying human emotions based on facial features. We got 75% accuracy regardless ethnicity, and 83.3% accuracy considering ethnicity. However, this effect is not very large.

Also our results also supported some previous findings that females are more emotion expressed on faces than that of males regardless of the race.

Generally the emotion of fear has the worst recognition accuracy, whereas the disgust emotion has a best recognition accuracy, and based on the race, the Caucasian has the best accuracy of emotion recognition.

5.2 Future works

We will propose a set of suggestions, which we believe that will help in advancing the research in emotion recognition.

1- Build a complete dataset that contains groups that are more ethnic and many forms of people from the same race. Moreover, increase the number of images for each emotion for each ethnic group.

2- Finding more useful features that may lead to more accuracy on recognition process.

3- Develop a system that can extract the features from human face image automatically.

4- Develop a complete recognition system, without resorting to the use of extra software.

5- Use video sequence in addition to image in emotion recognition system.

6- Using other features rather than the face features to test the impact of ethnic group on accuracy of emotion recognition like body movements or voice signals.

References:

[1] Akram M., Zafar I., Siddique Khan W. and Mushtaq Z., ”Facial Expression Recognition Based On Fuzzy Logic”, International Conference on Computer Vision Theory and Applications, VISAPP, (2008).

[2] Arief Muhammad, Wardhana Mitra Istitiar, Sumpeno Surya, Hariadi Mochamad. "Emotion Expression of Three Dimensional Face Model Using Naive Bayes and Fuzzy Logic". IJCSNS International Journal of Computer Science and Network Security, Vol. 10 No. 5. pp. 208-214 (2010).

[3] Bailensona Jeremy N., and others. "Real-time classification of evoked emotions using facial feature tracking and physiological responses". International Journal of Human-Computer Studies. Vol. 66, No. 5. pp. 303-317.(2008).

[4] Beaupré Martin G. “Cross-Cultural Emotion Recognition Among Canadian Ethnic Groups”, Journal Of Cross-Cultural Psychology, Vol. 36, 355-370, (2005).

[5] Binali, H.H.; Chen Wu; Potdar V. “A new significant area: Emotion detection in E-learning using opinion mining techniques”, Digital Ecosystems and Technologies, 2009. DEST '09. 3rd IEEE International Conference. pp. 259 – 264. (2009).

[6] Cantelmo Nathan. "Rapid Facial Expression Classification Using Artificial Neural Networks". CS 395-22, Machine Learning Conference’04.(2004).

[7] Castelli Fulvia. "Understanding emotions from standardized facial expressions in autism and normal development". National Autistic Society, Vol 9. Issue: 4, pages 428-449 (2005).

[8] Cowie R., Douglas-Cowie E., Taylor J., Ioannou S., Wallace M. and Kollias S. ”An Intelligent System For Facial Emotion Recognition”, IEEE,7803-9332, (2005).

[9] Douglas-Cowie E., Campbell N., Cowie R. and P. Roach, “Emotional Speech: Towards a New Generation of Database” Speech Communication, Vol.40 (1-2). pp 33-60 (2003).

[10] Efford N., ”Digital Image Processing”, Personal Education Limited, (2000).

[11] Ekman, P. et al. "Universals and Cultural Differences in the Judgments of Facial Expressions of Emotion". Journal of Personality and Social Psychology. Vol.53, No.4 pp.712-7 (1987).

[12] Ekman, P., Friesen, W. V., & Hager, J. C. ”Facial Action Coding System. Salt Lake City”. Research Nexus. (2002).

[13] Elaine Hatfield, Richard L. Rapson, and Yen-Chi L. Le. “Ethnic and Gender Differences in Emotional Ideology, Experience, and Expression”. An International Journal on Personal Relationships.(2009).

[14] Elfenbein Hillary A., Marsh Abigail A. and Ambady Nalini, "Emotional Intelligence and the Recognition of Emotion from Facial Expressions" The Wisdom of. Harvard University. (1998).

[15] Elfenbein Hillary A. and Ambady Nalini, "Universals and Cultural Differences in Recognizing Emotions". American Psychological Society. Volume 12, No 5, pp. 159-164 .(2002).

[16] FG-NET Database with Facial Expressions and Emotions Database (FEED), http://cotesys.mmk.e-technik.tu-muenchen.de/isg/content/feed-database. Resources and Evaluation Conference (LREC).(2006).

[17] Fischer Robert, "Automatic Facial Expression Analysis and Emotional Classification", University Of Applied Science Darmstadt (FHD), (2004).

[18] Gunes H. and Piccardi M., “A Bimodal Face and Body Gesture Database for Automatic Analysis of Human Nonverbal Affective Behavior,” Proc. 18th Int’l Conf. Pattern Recognition (ICPR ’06), vol. 1, pp. 1148-1153, (2006).

[19] Hewahi N.,Olwan A., Tubeel N., El-Asar S. and Abu-Sultan Z., ”Age Estimation based on Neural Networks using Face features”, Journal of Emerging Trends in Computing and Information Sciences(CIS), Vol.1. No.2. pp. 61-68 (2010).

[20] Horlings Robert. "Emotion Recognition Using Brain Activity". Man-Machine Interaction Group. TU Delft. (2008).

[21] Jagdish Lal Raheja and Umesh Kumar “Human Facial Expression Detection From Detect In Captured Image Using Back Propagation Neural Network”,

International Journal of Computer Science and Information Technology (IJCSIT), Vol. 2, No.1, (2010).

[22] Happyce Giger and others. "Health Disparities in Ethnic Minorities and Other Vulnerable Populations American Academy of Nursing Expert Panel Report: Developing Cultural Competence to Eliminate". SAGE Journals Online and HighWire Press platforms. (2007).

[23] Karthigayan M., Rizon M., Nagarajan R. and Yaacob S., “Genetic Algorithm and Neural Network for Face Emotion Recognition”, University Malaysia Perlis (UNIMAP), (2008).

[24] Kim Moon Hwan, Joo Young Hoon, and Park Jin Bae. "Emotion Detection Algorithm Using Frontal Face Image". Ministry of Information and Communication (MIC). (2005).

[25] Luger, George F., and William A. Stubblefield. Artificial Intelligence: Structures and Strategies for Complex Problem Solving. Second Edition. Univ. of New Mexico: Benjamin/Cummings. (1993).

[26] Michael J. Lyons, Akamatsu Shigeru, Kamachi Miyuki and Gyoba Jiro. ”Coding Facial Expressions with Gabor Wavelets”. Third IEEE International Conference on Automatic Face and Gesture Recognition, April 14-16, Nara Japan, IEEE Computer Society, pp. 200-205.(1989).

[27] Miyamoto Yuri and Uchida Yukiko “Culture and Mixed Emotions: Co-Occurrence of Positive and Negative Emotions in Japan and the United States”. American Psychological Association. (2010).

[28] Nauert R. "Are Emotions Universal?". Psych Central. (2010).

[29] Pantic M., Valstar M.F., R. Rademaker, and Maat L., “Web-Based Database for Facial Expression Analysis” Proc. 13th ACM Intel Conf. Multimedia (Multimedia ’05), pp. 317-321, (2005).

[30] Raheja J., Kumar U., “Human Facial Expression Detection From Detected In Captured Image Using Backprobagation Neural Network”, International Journal Of Computer Science & Information technology(IJCSIT), Vol.2. No.1, (2010).

[31] Ratliff Matthew S. and Patterson Eric "Emotion Recognition Using Facial Expressions With Active Appearance Models". Third International Conference on Human Computer Interaction. ACTA Press Anaheim, CA, USA (2008).

[32] Raudys Aistis, "Interactive initialization of the multilayer perceptron" Institute of Mathematics and Informatics, Lithuania (2000).

[33] Rudra T., Kavakli, M. and Tien, D. “Emotion Detection From Male Speech In Computer Games”. TENCON , IEEE Region 10 Conference. 4428779 .pp1-4 (2008).

[34] Soto Jose´ A. and Levenson Robert W. "Emotion Recognition Across Cultures: The Influence of Ethnicity on Empathic Accuracy and Physiological Linkage". Department of Psychology, Pennsylvania State University, University Park, PA 16802, USA (2009).

[35] Suzuki Kenji, Yamada Hiroshi and Hashimoto Shuji “A similarity-based neural network for facial expression analysis”. Elsevier Science Inc. New York, NY, USA.(2007).

[36] Terrence Fong, Illah Nourbakhsh and Kerstin Dautenhahn "A Survey of Socially Interactive Robots: Concepts, Design, and Applications". Technical Report CMU-RI-TR-02-29 (2002).

[37] Thurid Vogt, Elisabeth Andr. "Improving Automatic Emotion Recognition from Speech via Gender Differentiation". Proceedings of Language Resources and Evaluation Conference (LREC)(2006).

[38] Timotheou, S. "A novel weight initialization method for the random neural network". Department of Electrical and Electronic Engineering Imperial College London. Volume 73 Issue 1-3, (2009).

[39] Tsapatsoulis Nicolas, Karpouzis Kostas, Stamou George, Piat Frederic and Kollias Stefanos. “A Fuzzy System for Emotion Classification based on the MPEG-4 Facial Definition Parameter Set”. European Signal Processing Conference (EUSIPCO'00)(2000).

[40] Vapnik, V. N. "The Nature of Statistical Learning Theory". Second Edition, Springer. (2000).

[41] Waller B., Cray .J. and Burrows A., “Selection for Universal Facial Emotion”, American Psychological Association, 1528-3542/08, (2008).

[42] William Revelle and Klaus R. Scherer. "Personality and Emotion". Oxford University Press. (2009).

Appendix

A.1 Backpropagation Algorithm

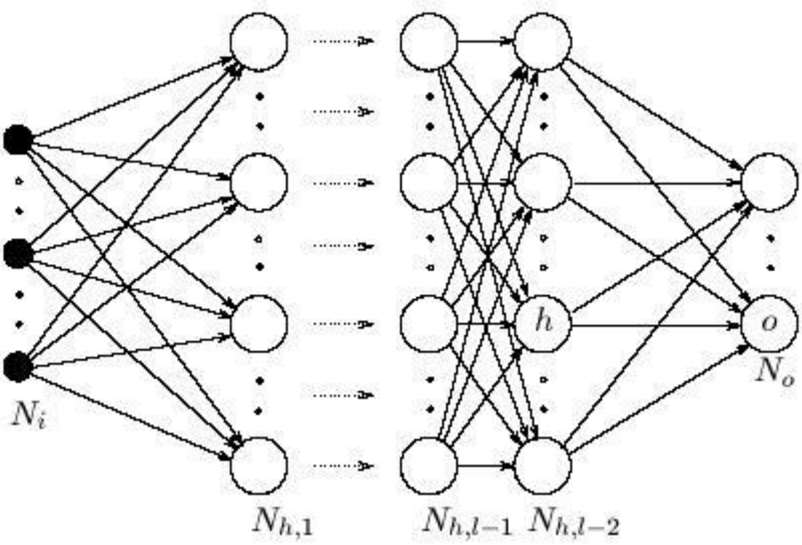
Actual algorithm for a 3-layer network (only one hidden layer):

```
Initialize the weights in the network (often randomly)
Do
       For each example e in the training set
            O = neural-net-output(network, e) ; forward pass
            T = teacher output for e
            Calculate error (T - O) at the output units
            Compute delta_wh for all weights from hidden layer to output layer ; backward pass
            Compute delta_wi for all weights from input layer to hidden layer ; backward pass continued
            Update the weights in the network
Until all examples classified correctly or stopping criterion satisfied
Return the network
```

A feed-forward network has a layered structure. Each layer consists of units which receive their input from units from a layer directly below and send their output to units in a layer directly above the unit. There are no connections within a layer. The Ni inputs are fed into the first layer of Nh;1 hidden units. The input units are merely 'fan-out' units; no processing takes place in these units. The activation of a hidden unit is a function Fi of the weighted inputs plus a bias, as given in in follow equation

$$y_k(t+1) = \mathcal{F}_k(s_k(t)) = \mathcal{F}_k\left(\sum_j w_{jk}(t)\, y_j(t) + \theta_k(t)\right),$$

The output of the hidden units is distributed over the next layer of Nh;2 hidden units, until the last layer of hidden units, of which the outputs are fed into a layer of No output units .



Although backpropagation can be applied to networks with any number of layers, just as for networks with binary units it has been shown that only one layer of hidden units success to approximate any function with finitely many discontinuities to arbitrary precision, provided the activation functions of the hidden units are non-linear (the universal approximation theorem). In most applications, a feed-forward network with a single layer of hidden units is uses with a sigmoid activation function for the units.

A.2 FFE Interface

FFE interface from Figure A.1 to Figure A.4 shown in this appendix with some explanations.

Figure A.1 show the interface of FFE program. The main window display the Asian male for shame emotion. The green points are the 64 points that selected to extract the features.

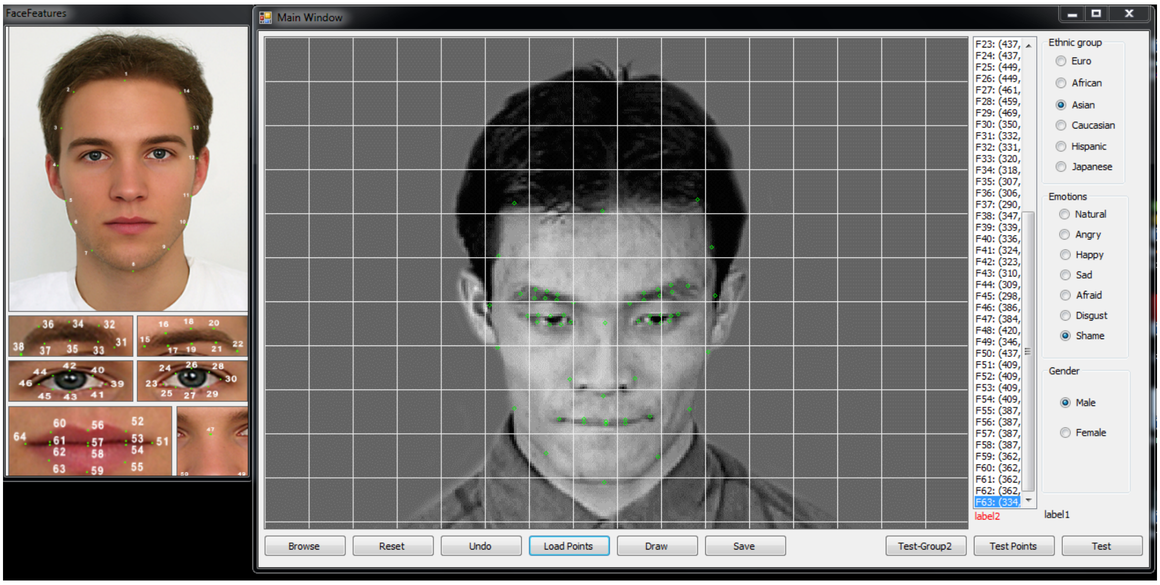


Figure A.1 interface of FFE program.

Figure A.2 shows the points controls part of FFE, which consists of six buttons. Brows button to select an image of human face with extensions (jpg, tif, bmp or png), the image is display directly on image framework with grid, and start to recive any click on image (point). The reset button for delete all points on image and start from first point. Undo button for delete last point is selected. Load Points button open dialog box for open file with extension (.txt), this file must have 31 integer values (28 for features, 1 for gender, 1 for race and 1 for emotion). Draw button using to draw the selected points into image file (this button we does not use this button in our study). Save button for save all points and features into text file)

Figure A.2 points controls.

Figure A.3 shows the coordinates of select points, that is mean if we select only two points, this list display two coordinates only. This list is very helpful to determine how many points selected.

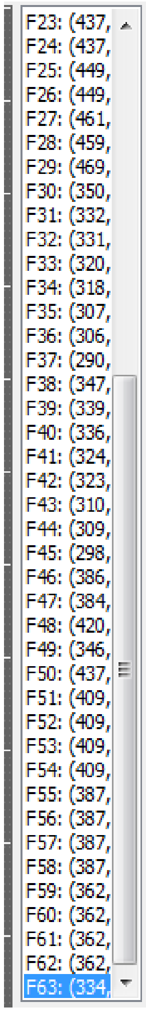


Figure A.3 List of selected points.

Figure A.4 shows the information of image for three attributes race, gender and emotion. after user select all points, he must select three attributes the race, gender and emotion, after that the user save the points and features into physical file. These attributes are codecs and saved into features file. In additional, when user loads any points from file, these attributes selected automatically depending on saved values.

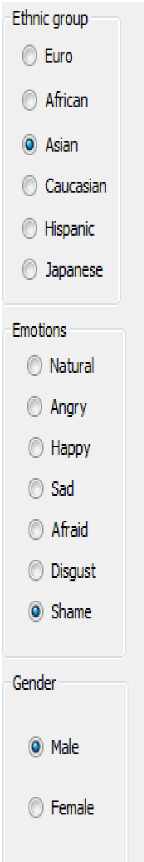


Figure A.4 race, gender and emotion values

A.3 FFE Code

Table A.1display parts of code for the FFE program.

Code	private void buttonLoadPoints ()
	{
	// Default values of some variables
	int gender=-1, emotion=-1, race=-1;
	// Choice text file that has a saved points.
	
	if (openFileDialogBrowse.ShowDialog() == DialogResult.OK)
	{
	String _file;
	_file = openFileDialogBrowse.FileName;
	
	string line;

	``
	System.IO.StreamReader
	sr=new System.IO.StreamReader(_file);
	int u = -1;
	
	// start read points coordinates from file.
	while ((line = sr.ReadLine()) != null)
	{
	    u = -1;
	    String[] tempArray = line.Split(' ');
	    for (int i = 0; i < tempArray.Length-3; i=i+2)
	    {
	        u++;
	        Features[u, 0] = Convert.ToInt32(tempArray[i]);
	        Features[u, 1] = Convert.ToInt32(tempArray[i+1]);
	        gender=Convert.ToInt32(tempArray[i+2]);
	        race=Convert.ToInt32(tempArray[i+3]);
	        emotion = Convert.ToInt32(tempArray[i+4]);
	    }
	    FeaturesNo = u+1;
	}
	// Decoding the emotions.
	if (emotion==1)RadioHappy.Checked=true ;
	else if(emotion==2) RadioSad.Checked = true ;
	else if (emotion==3)RadioAngry.Checked = true ;
	else if (emotion==4)RadioAfraid.Checked = true;
	else if (emotion==5)RadioShame.Checked = true;
	else if (emotion==6)RadioDisgust.Checked = true;
	else if(emotion==7) RadioNature.Checked = true;
	// Decoding the ethnic group.
	if (race == 1) RadioEuro.Checked=true;
	else if (race==2)RadioAfrican.Checked = true;
	else if (race==3)RadioAsian.Checked = true;
	else if (race==4)RadioCaucasian.Checked = true;
	``

	else if (race== 5)RadioHispanic.Checked = true;
	else if (race == 6)RadioJapanese.Checked = true ;
	// Decoding the gender.
	if (gender==1) RadioMale.Checked=true ;
	else RadioFemale.Checked=true ;
	// Draw points on human face image.
	FaceImage=
	IplImage.FromBitmap(RedrawPoints(OrignImage, FeaturesNo));
	
	// read features from current points
	ReadFeatures(FeaturesNo);
	pictureBoxPreview.Image =
	DrawGrid(FaceImage.ToBitmap());
	
	sr.Close();
	sr.Dispose();
	sr.DiscardBufferedData();
	}
	}

Description	This procedure for Load button. It is open dialog box of open file, and read points coordinates from selected file. Then draw the points on human face image, then determining the ethnic group, emotion and gender by decoding his values, finally calculate the features of these points.

Table A.1 example of FFE codes.