A UNIFIED APPROACH TO EMOTION RECOGNITION USING HAAR LIKE AND GABOR WAVELET

A DISSERTATION

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE

OF

MASTER OF TECHNOLOGY

IN

SIGNAL PROCESSING AND DIGITAL DESIGN ENGINEERING

Submitted By:

SANHITA PATHAK

2K16/SPD/13

Under the supervision of

Dr. Dinesh K. Vishwakarma

Associate Professor, Department of Information Technology

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College Of Engineering)

Bawana Road, DELHI -110042

JUNE 2018

DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi -110042

CANDIDATE'S DECLARATION

I, Sanhita Pathak, 2K16/SPD/13 of M. tech. (Signal Processing and Digital Design), hereby declare that the project Dissertation titled "A unified approach to emotion recognition using Haar-like and Gabor wavelet" which is submitted by me to the Department of Electronics and Communication, Delhi Technological University, Delhi in partial fulfilment of the requirement for the award of the degree of Master of Technology , is original and not copied from any source without proper citation. This work has not previously formed the basis for the award of any Degree, Diploma Associateship, Fellowship or other similar title or recognition.

Place: Delhi Sanhita Pathak
Date

Department of Electronics and Communication Engineering
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042

CERTIFICATE

I hereby certify that the Project Dissertation titled “Framework for Human Emotion Recognition” which is submitted by Sanhita Pathak (2K16/SPD/13), Department of Electronics and Communication, Delhi Technological University, Delhi in partial fulfilment of the requirement for the award of the degree of Master of Technology, is a record of the project work carried out by the student under my supervision. To the best of my knowledge this work has not been submitted in part or full for any Degree or Diploma to this University or elsewhere.

Place: Delhi

Date:

Dr. Dinesh K. Vishwakarma
Associate Professor
Department of Information Technology
Delhi Technological University, Delhi

ACKNOWLEDGEMENT

I would like to express my immense gratitude towards **Dr. D.K. Vishwakarma** (Department of Information Technology, Delhi Technological University ,New Delhi) for his guidance and constructive suggestions during our whole thesis work. I also give special thanks to him for dedicating his valuable time whenever I needed the insights on the matter.

I would also like to thank my Head of the Department **Dr . S. Indu** (Department of Electronics and Communication) for the constant support and guidance that helped me throughout the work.

I must acknowledge all the faculties and students for providing unfailing support throughout the research process and while writing this thesis.

Finally, I would like to thank my parents for providing the continuous encouragement without which this work would not have been possible.

Sanhita Pathak

ABSTRACT

Human emotional behavior recognition has become the key topic in the research domain. A large number of techniques and tactics are applied in order to obtain the high accurate mechanism for emotion detection. The mechanisms like classification, feature extraction, fusion and others are the major tactics that can be utilized in order to enhance the accuracy and reduce the error rate of the emotion recognition system. This study is organized with an objective to develop an emotion recognition system by using Gabor and Haar-like feature extraction algorithm, PCA for image fusion, SVM, ANN (Feed Forward Neural network) and KNN classifiers for classification analysis. For the purpose of simulation, the MATLAB simulation platform is utilized and testing of the proposed work is done by using three different datasets i.e. Cohn-Kanade, Yale and Jaffe dataset. The proficiency of proposed emotion recognition technique is evaluated in the terms of Accuracy Rate. After evaluating the results, the study proves that the ANN classifier has the highest accuracy rate in comparison to the KNN and SVM classifiers.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Human Expression Recognition:

In Interpersonal type of interaction is mostly complex and nuanced. The success rate of the interpersonal interaction is dependent of large number of parameters. These parameters vary and these include the situation, frame of mind, timing of the communication, and also on the people's expectations who are participating in the communication. To be a successful participant of an interaction it is required that the participants should perceive each other's nature with the progress in the communication process also, it varies according to the participants .In humans beings this feature is inborn and the efficiency of this feature varies from person to person. Human beings have the ability to speedily assess a multitude of indicators like choosing the words, variation in the voice of the person, and body language to find out the mood of the other person. As it is known that the humans have the fundamental set of emotions and therefore it gives rise to the analytical ability, likely stems [1] .

Considerably, emotions of the human beings are represented by the expressions of the face which are constantly correspondent. This specifies that it does not matter that the person belongs to different culture and speaks different language but all the humans have a comprised a set of fundamental facial expressions through which the human interact. Various experiments have been conducted and it is revealed that seven types of facial expression that are shared by the humans mostly highlight the fundamental emotions. Various categories of the fundamental emotions are as follows: anger, contempt, disgust, fear, happiness, sadness, and surprise. To find the genuine mood of the person, one method is face expression recognition.

Expression's universal versatility says that to determine the facial expression, computers can be used. Moreover, as in various other tasks computers are used because they have the features which provide the advantage over humans during analysis. After implementation of computers for the facial expression recognition task it is observed the efficiency and automation of the procedure provides better results and can be further applied in the process such as entertainment, social media, content analysis, criminal justice, and healthcare [2]. For instance, to offer the best services to the consumer accordingly, it is firstly important to recognize the reactions of a consumer. While executing the detection process (performed by either human or computer), it is important to have a taxonomic reference to find out the seven set of emotions.

The famous facial expression coding techniques are implemented by various psychologists. Various types of action units are sued in the system which shows the variation in some of the facial muscles to find out the person's emotions. The Action Units give the detailed description of the movement in the facial muscles like inner or the outer brow raising, or nostrils dilating, or the lips pulling or puckering, as well as optional intensity information for those movements [3]. The FACS signify distinct and apparent facial muscle movements and these movements of the muscle varies in accordance with the emotion of interest, digital image processing. For the successful analysis of apparent facial characteristics it is required to provide the training to the facial expression predictors.

1.2. Introduction to facial emotions:

As the evolution in the field of image processing and cameras has been occurred then various researches had been conducted in the field of human facial expression recognition (FER). To interact and determine the emotions and genuine moods of humans the facial expressions recognition can be really helpful. In various kind of applications like virtual reality, security of driver, video conferencing, determining the personality during machine-human interaction, determining the human emotion analysis, in the field of health care, physiological, image retrieval, video analysis etc the real-time and automated facial expression plays the important role. In case of the affect determination system the facial expressions plays the important role. Various Psychologists have suggested various types of systems to determine the facial expression. Ekman et al had introduced the two widely popular and used facial action coding system (FACS). FACS gives the detailed information of different 33 Action Units (AUs) and these Action Units are capable to find out the various facial muscle variation. With the help of Action Units it is possible to explain the different human facial expression. Various techniques are used for the purpose of facial expression recognition and it can be said that large number of researches were carried out for determining the six basic emotions such as sadness, fear, anger, surprise happiness and disgust.

1.3 Physiology of facial expressions

It is the physiological fact that with the variation in the mood the facial expression changes and that move the facial muscles. Those muscles are referred as mimetic muscles or muscles of facial expressions. The mimetic muscles are the part of head muscles. The head muscles constitute muscles of the scalp, muscles of mastication, and the tongue. Facial muscles are comprised of Facial nerve. Facial nerves form the branches in the complete face and when these got activated then it results in contraction. The apparent muscle variations are due to block of skin motion such as eyebrows, lips, cheek, and wrinkles.

If the human face is studied in detail then it can be observed that it is comprised of 20 flat skeletal muscles and it is represented in Fig. 1.1 [4]. Muscles finds their position under the skin, connected with the skull bone and place in the facial skin, but these muscles and not placed in the bones or joints because these muscles causes all the body movements. These types of muscles are placed in surrounding of facial orifices such as mouth, nose, and eyes. But various other facial muscles do not show the variation with the movement in joints and bones, but usually with the skin. In result, these muscles create the deformation on the face surface that would create the changing expression of face which shows the emotions. It is noteworthy that muscles show the movement in sets instead of doing this alone and it also controls the orifices. On the basis of position, the taxonomy can be categorised into three different groups:

- Oral
- Nasal
- Orbital

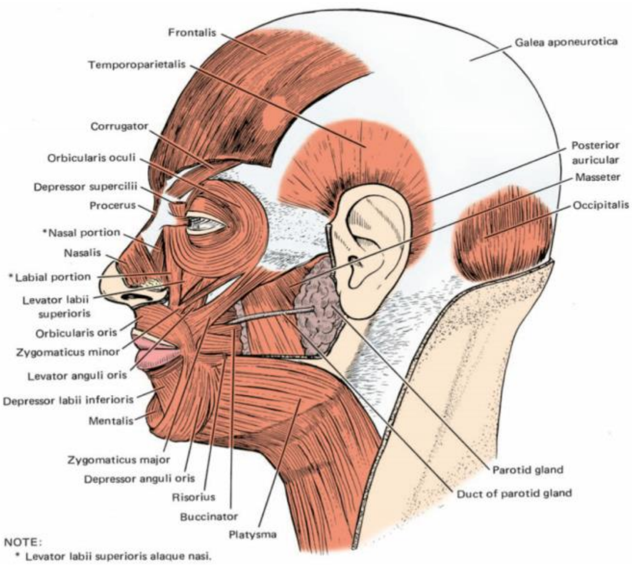

Figure 1.1 Muscles of Facial Expressions

Shape of oral orifice varies due to the variation in the oral muscles. These oral muscles creates complicated variation in mouth such as encircle the mouth, control angle of the mouth, move the lower and upper lip independently or move left and right corner and to show the variation in cheek surface.

For the movement in the nostrils, the nasal group plays the important role. The important muscle to show the facial expression can be found out in between the eyebrows. Due to these muscles eyebrows moves downwards and wrinkles are formed on the nose. For the movement in the eyelid and also to protect the eyes, the three orbital groups of muscles are playing the important role. As the insertion of muscles occur inside the skin surrounding the eyebrows then the vertical wrinkles are formed between eyebrows.

The educational tools are used to represent the specific muscles variation altogether with emotions can be determined. This educational tool helps in introducing the anatomical and biomechanical foundation of facial expression morphology to the users. This tool also provides various interactive examples which represent the muscle variation and position during a facial expression.

1.4 Psychology of emotions - basic categorization

Ekman had explained the six categories of basic emotions. In the below sections description about the emotions and the important features are explained as follow: when the various emotions are intermixed then it forms the different facial expressions, such as angrily surprised. Therefore to explain and analyze the facial expression Action Units can be used on the semantic level so that the specific action can be revealed. On the basis of Ekman's theory emotions can be grouped into 6 different categories and these groups are universal for all the people all over the world. Six different categories of emotions are as follow: joy, sadness, anger, fear, disgust and surprise.

1.4.1 Anger

Anger can be described as one of the strongest emotional reactions. Anger can also be the dangerous emotion because it may also result in the provoke violence. There can be the various reasons for the anger. Let’s consider the situation that a person is facing the hindrance on his way then it may lead to the frustration and the outburst of anger emotion. One more example for anger cause is physical threat [5].

If a person hurt some other person then it automatically leads to the anger emotion. Various other causes for the anger as follow: physical violence, verbal threats or claims. Some more examples are false accusations or when our personal values are broken by someone. Anger results in the significant affect on the complete body. Mostly it is reflected with the increased blood pressure, red face, and contracted muscles. The following types of signal are developed with the physiological response:

- Position of eyebrows gets lowered down and they will also squeeze. Vertical direction wrinkles will also appear between the eyebrows.
- Eyelids will take the tight and straight direction.
- Eyes at the time of anger focused at the anger source. Narrow pupils are created and they are focusing on the anger source.
- Position of lips is either tightly closed or normally opened (preparing for yelling).

By mixing all these features together it is appeared that the person may be in the position of physical or verbal attack.

Figure1.2 Six Universal Emotions

1.4.2 Disgust

Disgust emotion may be described as the emotion which is usually activated due the smell, taste or vision. Not like the other emotions the source of disgust emotion is not universal but it can be cultural or personal such as food. In disgust emotion the extreme physiological action is vomiting. In our face the significant parts are mouth and nose region.

- Lifting of upper lip.
- Generation of wrinkles over the nose.
- Lifting up of cheeks.
- Lifting up of eyelids but formation is not tight. Wrinkles are formed under the eyes.
- Pulling down of the eyebrows.

1.4.3 Fear

Fear emotion is created when danger or stressful situation is encountered. Fear emotion can be derived because of the future events such as fear of losing something. Due to the fear emotion the outburst of violence may occur and that may leads to the anger emotion. If the fear emotion is experienced by a person he gets ready for escape or protects himself against any kind of attack. In this case the heart rate and blood pressure increases, person’s eyes are in open position, pupils are wide and the maximum amount of light is absorbed by the eye [6].

In various adverse conditions the fear might generates loss of muscle operation such as paralysis.

- Lifting up of eyebrows and they are pulled in inward direction.
- Wrinkles are formed over the forehead.
- Lifting up of upper eyelids.
- Opening up of mouth and tightening of lips as per the emotional intensity.

1.4.4 Happiness

Happiness can be described as a type of emotion in which the positive vibes are linked and the person wears the smile on his face. This emotion is generated when a person achieve what he wants. Various features of fear emotion are as follow:

- Pulling of lips corner back and up.
- Mouth may be in open condition and teeth are apparent.

1.4.5 Sadness

Sadness emotion is visible if the person is suffering. Cause of sadness is specifically when something is lost.

In this emotion the person remains calm, not impulsive and also leads to tears from eyes. In this emotion muscles on the face are not tensed and this may leads to the specific physiological characteristics as follow:

- Pulling down of inner parts of the eyebrows.
- Pulling down of Lips corners and shaking of lips.

1.4.6 Surprise

Surprise is kind of emotion which occur suddenly. No thinking is involved in this emotion and also this emotion is for short period of time. Initiation occurs with unexpected or incorrectly expected situation. Thus it can be concluded that the surprising emotion can be positive as well as negative. In case if we have time to think for the situation going to be happens then it will not be a surprise emotion. Mostly the surprise emotion is transformed into any other kind of emotion, such as happiness or sadness [7].

Various characteristics of surprise are as follow: lifting up of eyebrows that result in creation of wrinkles on the forehead, eyes are open and dropped jaw.

- Lifting up of eyebrows and pulling it in the inward direction.
- Horizontal wrinkles appear on the forehead.
- The eyes are open widely.
- The jaw is in dropped position. Mouth is in open position and the lips are tight.

1.5 Methods considered for analysis of facial expressions

These two types of operation are as follow:

1. To use the complete front face of image and also processing the information so that the classification of six different types of frontal expression can be easily recognized such as fear, disgust, joy, surprise, anger and sadness; this give the brief introduction of first approach. In this approach it is considered that all the emotions which are mentioned above have significant facial expression and therefore it is mandatory to recognize these expressions

2. Rather than implementing the face images completely it is an efficient approach to use them by dividing the complete picture into subsections so that it can be easily processed. This forms the central idea of the 2nd approach for analysis of the facial expression. The expression of face is closely linked with slight variations of some distinct characteristics like eyes, eyebrows and lip corners. These small variations are implemented for automatic analysis. This technique was proposed for the 'Facial Action Coding System', and this system was introduced by Ekman and Friesen in order to explain the facial expressions by using 44 types of Action Units (AU's). Benefit of this approach is that by disintegration of the image the range of application for determining the face expression become wide. Because of dividing the image into sub parts, different features are processed separately instead of having six different universal facial expression prototypes. Large section of present work was executed on facial expression analysis in which the AUs are implemented. For the facial expression determination there are various other techniques in which the complete front face image and all the 44 AUs are not used but various other techniques like selecting the surface regions of facial features can be implemented for facial expression recognition [8].

1.6 System requirements

Major objective of FERS is to replicate the human visual system in the exact manner. It is really difficult in the field of computer vision as it does not only need optimum image/video analysis methods but it also required the appropriate characteristics vector implemented for the machine learning method.

In FER system first principle is that it must be optimum as well as effortless. That is connected with full automation, so that no additional manual effort is required. In this system it is also mandatory that the system must be real-time and it is important in case of human-computer interaction as well as interaction taking place in between human-robot.

Moreover, in this study the subject should have the permission to react impulsively when the information is gathered for the purpose of analysis. It is necessary that the designed system must overcome the limitation on body and variation in head position, this information can play the important role in displaying the emotion. Obstructions like facial hair, glasses or make-up should be least. Furthermore, for the system it is big task to handle the occlusions problem and there problem should be considered while designing the system.

In addition to all the features which are mentioned above there are one more features which are mandatory in FER system areas follow: it should independent of user and environment. Independence of user means that any user can use this type of system regardless of skin color, age, gender or nation. Independence of environment means that it can be implemented in any background and in different lightning situation.

1.7 Face detection

It was described earlier that the FER system is comprised of 3 different stages. In first stage the input parameters are fed to the system and after that images are processed by using different methods so that facial expressions can be recognized. This system can work on static type of face pictures and this method is known as face localization and on the other hand it can also work on the different videos for the purpose of face tracking. Big issues that will come in the way at this stage are varying scales and different orientations of face. These variations are usually occurred due to movement in the subject and also because of variation in the distance from camera.

Important body movements may result in the notable variation in the position of the face in consecutive frames that will make the tracking process complicated. Due to the complex background and varying light intensity, the tracking process will become confusing. For example, if more number of faces is seen in the image system then it important that system can differentiate between the faces so the required face can be traced. At last, it can be said that the occlusions that are visible in spontaneous reactions required to be properly handle.

It is required that the problems which are explained above should be overcome and for that different techniques were searched out so that these problems can be solved. All the techniques which were used for the face detection were categorized in two groups: first one is holistic in which the face is considered as the complete unit and analytic in which the co-occurrence of features of the facial elements is reviewed [9].

1.8 Feature extraction

If the image or video of the face has been captured then after that it can be used for analysis in terms of facial action occurrence. Generally there are two types of characteristics required to explain about the facial expression: first one is geometric features and second is appearance features. With the help of geometric features the variations in certain parts of the face can be analyzed like movement in brows or mouth corners on the other hand the visible characteristics explain the variations in face texture when the specific task is required to be executed. Along with the type of feature the FER systems can be categorized on the basis of input and these inputs can be static images or sequences of image.

The geometric feature analysis process is normally interconnected with analysis of the face area, specifically determining and tracking important points over the face. Various issues that were raised during the face decomposition process due to the occlusions and facial hair or glasses. Moreover, to explain the characteristics are not an easy task as these characteristic should be descriptive in nature and possibly not interlinked.

1.9 Expression Recognition

This part of the FER system is based on “machine learning theory”; precisely it is the classification task. The input to the classifier is a set of features which were retrieved from the facial region in the previous stage. The set of features is formed to describe the facial expression.For the purpose of classification it is required to have the supervised training therefore it is mandatory that the training set should be comprised of labelled data. When the training is given to the classifier then it can distinguish various input images by allotting then specific class label. Widely implemented facial expressions classification techniques uses Action Units which is introduced in Facial Action Coding System and in also in terms of universal emotions such as happiness, sadness, surprise, anger and fear. Various machine learning methods used for the purpose of classification are as follow: K-Nearest Neighbors (K-NN), Artificial Neural Networks (ANN), Support Vector Machines (SVM), Hidden Markov Models, Expert Systems with rule based classifier, Bayesian Networks or Boosting approach (Adaboost, Gentleboost).

Major three problems in classification process are as follow: selecting optimum characteristic set, optimum machine learning method and diverse database for training. It is required that the feature set must be comprised of characteristics which are discriminative and features for specific expression. In order to sort the feature set, Machine learning method can be implemented. In the end, the data base used for the training set must be sufficiently large and comprised of different data set. Techniques explained in the literature are shown by groups of classification output [10].

1.10 Applications

In the facial muscle movement, the large amount of data is encrypted. While the face of a person is observed that various specifications can be learnt from that as follow:

• **Affective state:** It is interlinked with different emotions such as fear, anger and happiness and moods like euphoria or irritation

• **Cognitive activity (brain activity):** It can be supposed as concentration or boredom

• **Personality**: It is comprised of features such as sociability, bashfulness or unfriendliness

• **Truthfulness**: It requires the examining of micro-expressions to show concealed emotions

• **Psychological:** In this state the information related to various disorders can be used for analysis of depression, mania or schizophrenia.

With the availability of large set of information apparent on human face it is possible to analyze the facial expression and it can be implemented in various in different fields such as science and life.

1.11 Extraction Methodology for features

The description is given about the Gabor wavelets , Haar features and also about the Haar-like coefficients. The three different kinds of coefficients which are mentioned above create the feature vector and these vectors will help in the classification technique of the FACS action units. When the features selection technique has been like Adaboost then after that the image extraction process is implemented on the normalized frontal face images [11, 12].

1.11.1 Haar Wavelet Coefficients

After comparison of the Haar Wavelet Coefficients with the Gabor Wavelets it is observed that the Haar Wavelet Coefficients has the advantage in terms of the extraction times.

The mother wavelet function of Haar Wavelet can be taken as a step function and this funmction is given below:

$$\Psi(t) = \begin{cases} 1 & 0 \leq t < \frac{1}{2} \\ -1 & \frac{1}{2} \leq t < 1 \\ 0 & otherwise \end{cases}$$

The wavelet disintegration of a picture can be explained as the grouping of disintegrated images which are calculated by using the different scales. The technique in which the Haar coefficients for an array of image samples are calculated is described : (The length Array length must be a power of two)

1. Determines an average for each sample pair.

2. Determine the differences among the calculated averages and the samples.

3. In the first half of the array the calculated averages are filled.

4. In the second half the computed results in step 2 are utilized.

5. Recurse – In this step the process is executed again from the first half of the array till then when the process of until recursion cannot be executed anymore.

Example: consider that Haar Wavelet disintegration of an eight-element array as follow:

7 9 1 1 5 3 2 -10

After that average for each pair is computed as follow [13]:

(7+9)/2 = 8; (1+1)/2 = 1; (5+3)/2 = 4; (2-10)/2 = -4

Then the difference among the averages and the sample values are computed:

(7-8) = -1; (1-1) = 0; (5-4) = 1; (2-(-4)) = 6

Then in the obtained halves of the considered array, the averages are filled up as follow:

8 1 4 -4 -1 0 1 6

The technique is implemented to the array’s first half of the array recursively.

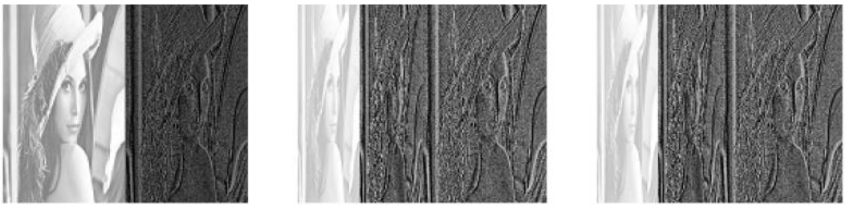

Figure1.3 Transform of Haar features for Rows

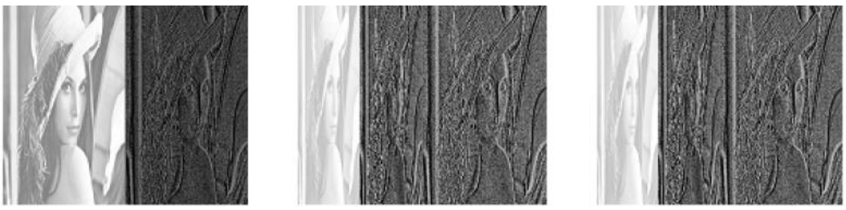

Figure1.4 Transform of Haar features for column or rows

Coefficients for Haar-Like are the characteristics which are evocative of Haar Basis functions and these are proposed by Viola and Jones. They are explained by the three different types of characteristics: “two-rectangle”, “three-rectangle” and “four-rectangle” characteristics. Here the difference among the sums of pixels of similar size pairs of rectangles are calculated out as the characteristics [14]. Drawback of these characteristics is that large amount of time is taken to obtain all the Haar-like characteristics of an input image. For more better understanding take the example: in 24*24 image there are 160000 Haar-like characteristics on the other hand in a 32*32 image there are more than 450000. Therefore the set of characteristics is more times than over-complete. On the other hand benefit of this is that any rectangular sum can be calculated by just measuring the “four array differences” by implementing the “integral image” technique. In addition to this, one more fact is that the characteristics are progressed by using the feature selection method; this creates a necessary reduction in the complication in the feature extraction technique. One more benefit of this technique is that the features [15] sensitivity to edges, boundaries and some vital data or information concealed in pixel values like the difference among the values of pixels on the regions of motion on face.

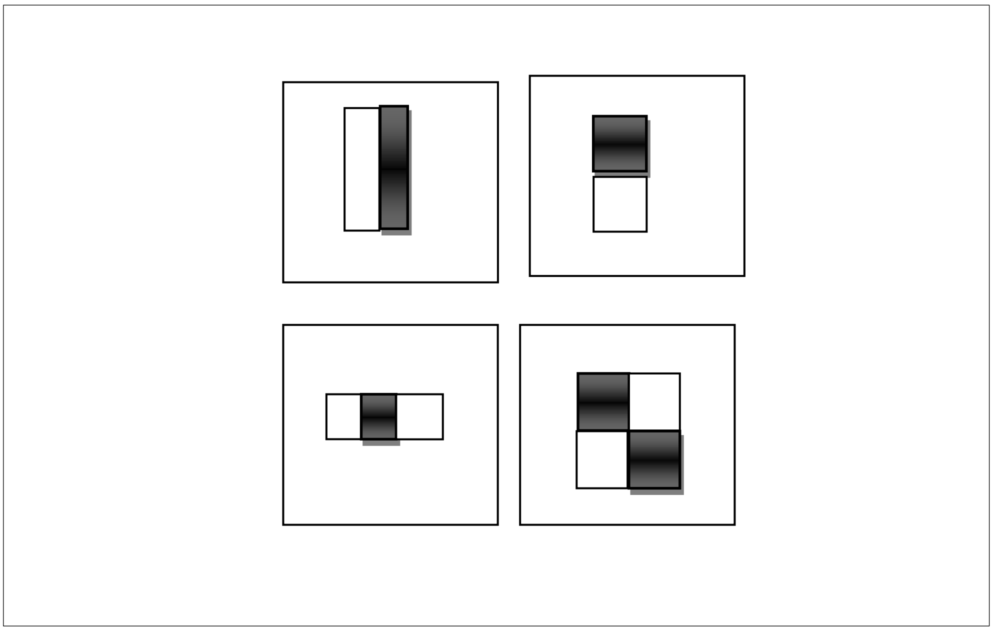

Figure 1.4 Types of rectangle features

1.11.2 Coefficients for Gabor

The “Gabor wavelet coefficients” can be implemented in different types of signals and also in pattern processing experimental regions and these are implemented in spatial as well as frequency domains and it was observed that after implementation the optimum results can be obtained in application region like texture segmentation, determination of fingerprint and also in face recognition [16]. Features of Gabor wavelets like its capability to easily adjust itself for detailed localization in both types of domains spatial as well as frequency. Due to the frequency and orientation representations similarity as compare to the human visual system this system have been widely used for specific usage regions and it is observed that optimum and satisfactory results are obtained in the application regions which are explained above.

The Gabor wavelets are derived as following [17]:

The “complex sinusoidal” can be defined as S(x, y) = exp(j(2p (a0x + b0 y + Q)) ; 0 a and 0 b are the spatial frequencies, and Q the phase. It is clear that the carrier consists of two sinusoids, the cosine wave on the real domain, and a sine wave on the imaginary domains. In order to switch from the” Cartesian coordinate system” to “polar coordinate system”; 0 *F*, the magnitude of the frequencies is defined together with 0 *w*, the direction.

$$F_0 = \sqrt{u_0^2 + v_0^2}$$

$w_0 = \arctan(v_0/u_0)$

$$s(x,y) = \exp(j(2piF_0(xcosw_0 + ysinw_0) + Q))$$

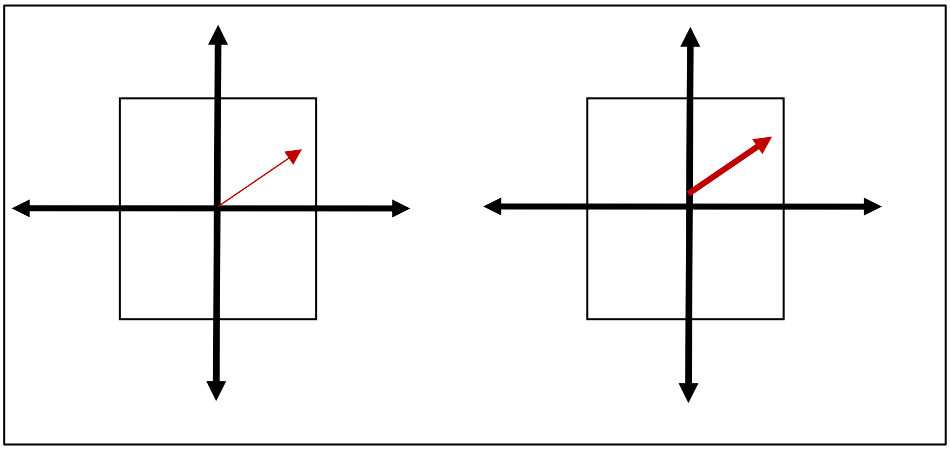

Figure1.5 The real and imaginary part of sinusoidal carriers

The real and imaginary parts of a sinusoidal carrier $u_0 = v_0 = \frac{1}{80}\ cycles\ /\ pixels$ and P=0 degrees.

In the second section of the Gabor wavelets such as the Gaussian envelope can be explained as follow [18] :

$$w_r(x, y) = Kexp(-pi(\propto^2\ (x - x_o)_r{}^2 + b^2(y - y_0)_r^2)).$$

In the above equation (x_o, y_0) shows the peak value of the function. Here a, b are showing the scaling parameters. If the scaling parameters value gets decreased then the Gaussian get increase in the spatial domain. Here$(x - x_o)_r$, $(y - y_0)_r$ are showing the clockwise rotated form of the$(x - x_o)$, $(y - y_0)$ respectively and these can be described as follow:

$$(x - x_o)_r = (x - x_o)cos\theta + (y - y_0)sin\theta$$
$$(y - y_0)_r = -(\ x - x_o)sin\theta + (y - y_0)cos\theta$$

1.12 . CLASSIFICATION TECHNIQUES

The various classification techniques involved in the proposed methodology are described as below;

1.12.1. SVM

This classification technique is a linear classification technique that was introduced by Vapnik et . al. [19] , this was introduced mainly to handle with the classification of humongous amount of data and to separate the values based on a set of features.

This is also part of a supervised learning mechanism that takes the concept of hyperplanes in consideration, the error mechanism is considered.

As the error value of the subject with respect to the present values decreases the plane transforms accordingly. The margins are increased and it can also be said that the increased amount of separation is considered [20].

1.12.2. KNN

This is an algorithm that was developed by Marcello Pellilo which was initially given as the NN rule. This algorithm is used for classification of the data values, another way of implementing it is for the predictive problems involving regression.

The involvement of the neighbouring values is considered which gives it its name.

The voting scheme involves the nearest values into one single cluster involving there feature matching, this in turn automatically gives the automated clustering of the neighbour values based on the distance.

In our methodology the features extracted and fused are in turn applied to the above algorithm to get the final results.

1.12.3. FFNN

Feed Forward Neural Network (FFNN) is the simplest and easy to implement type of Artificial Neural Network (ANN). In FFNN, the flow of information is in single direction i.e. the data is forwarded to the hidden nodes and then output nodes from input nodes. It did not comprise of loops or iterations. The example of FFNN is Single Layer Perception (SLP) and Multi Layer Perception (MLP).

CHAPTER 2

LITERATURE SURVEY

2.1 Literature Review

1. **Nwosu et al.** [1] In this paper the author projects the design of a Facial Expression Recognition (FER) system which is based on deep convolutional neural network by using facial parts. A collection of the paradigms for the feature extraction face detection, and classification was applied that was an easy resolution to the facial expression recognition was described in this work. The proposed method had used a two channel convolutional neural network that used the Facial Parts (FPs) as the input for the first convolutional layer, The results obtained show that the system offers improved classification accuracy in comparision to other methods.

2. **Tonchev et al.** [2] in this paper the author had illustrated that the “Automated expression recognition” was a reserch field that estimated the human expressions from images and videos where the computations and machine learning is combined. This work had come up with an algorithm that recognizes the expression by feature extraction that uses a log-gabor filters and a feature selection approach on the basis of sparse estimation. For the classification part SVM was used with radial basis kernel . The algorithm was tested on the posed facial expressions image

database Cohn-Kanade and provides competitive results were compared to the state of the art.

3. **Li, Zhi [3**] in this paper the author had outlined that as of late, profound learning had turned into a most looked territory. The exploration on facial acknowledgment was advancing quickly; in any case, outward appearance acknowledgment faces numerous challenges because of poor power and constant execution. The element of a few distinctive sort of outward appearance was comparative, which was anything but difficult to confound, and it turned into the key factor to influence the exactness of outward appearance acknowledgment. The Convolutional Neural Network (CNN) had been used comprehensively on a similar time in the picture classification errands through its solid limit on the circulated theoretical trademark coercion in the region of picture. A discriminative learning convolution neural system was thought and acknowledged in this work. The system consolidates the focal misfortune work and the confirmation acknowledgment show, which improve the model have qualities of the speculation and segregation capacity, and furthermore diminish the misclassification in outward appearance acknowledgment. Analysis demonstrate that the precision of the composed outward appearance acknowledgment arrange has been viably moved forward.

4. **Fen Xu and Zhe Wang** [4] in this paper the author had represented that the outward appearance acknowledgment had offered a critical part in a few applications. A novel component for the recognizable proof of the outward appearance was offered in this investigation. The technique had utilized the cubic spline interjecting coefficients of milestone focuses together with HOG (Histogram of Oriented Gradients) of chose territories as speaking to attributes and the help vector machine (SVM) was used keeping in mind the end goal to build the arrangement models for the distinguishing proof of the outward appearance. The appropriation of spline interjecting coefficients for geometrical portrayals decreases the dimensionality of highlight vectors essentially while keeping the precision. Trials additionally demonstrate that the incorporation of these two highlights at choice level once in a while performs superior to anything that of highlight level combination concerning distinguish the outward appearance.

5. **Chao Qi et al.** [5] in this paper the author had projected a novel expression recognition method on the basis of the cognition and mapped binary patterns. In order to extort the facial contours the method is based on the LBP operator firstly. Afterward the pseudo 3D model was introduced to alienate the face region into six facial expression sub-sections. The mapped LBP mechanism was utilized by the sub-regions and the global facial expression images for the characteristic extortion. A relative practice was presented on the extension of the Cohn-Kanade (CK +) facial expression data set and the test data sets gathered from ten volunteers. The simulation results had demonstrated that the confusing factors in the image were eliminated efficiently by the projected mechanism. Also by applying the circumplex emotion model the outcomes were improved than the conventional emotional models.

6. **Tsangouri et al.** [6] in this paper the author had outlined that the feelings were a staggeringly imperative part of human correspondence. People with Autism Spectrum Disorder (ASD) experience the ill effects of huge difficulties in seeing, and understanding others' feelings, and reacting candidly in reasonable way. The EmoTrain was produced as an intelligent stage to upgrade the outward appearance acknowledgment and reaction aptitudes by applying profound figuring out how to track the outward appearances from a cell phone's camera of a client. The thought of the EmoTrain stage, the profound learning model and diversion motor for the stage was clarified and the proficiency of the Emo Train's was registered in this work.

7. **Motahareh Taheri** [7] in this paper the author had projected a robust face recognition paradigm that was named as non-linear correlation filter bank (NCFB). In order to obtain the improved recognition presentation the NCFB merges the gains of a sigmoid function like non-linearities of image pixels and correlation filters (CFs). On the basis of the unconstrained minimum average correlation energy the CFs was deliberated corresponding to every sub-section of images to optimise the entire correlation outcomes. In order to effectively use the discriminative information in face sub-regions among taking into an account the fluctuations in face sub-regions the NCFB was utilized. In this paper it was

illustrated that the projected mechanism was vigorous next to illumination, pose, and facial expression fluctuations. The simulation results had demonstrated that the recognition rate of the projected mechanism was improved comparative to the conventional CFs.

8. **Zhe Sun et al** [8] in this paper the author had projected a discriminative feature learning mechanism in order to enhance the representation power of expressions. On the basis of the discriminative feature dictionary (DFD) the pixel diverse presentation was achieved at the starting. Consequently, a discriminative feature dictionary (DFD) was made by applying the total DFMs equivalent to the training samples. After that on a vertical two-dimensional linear discriminant analysis in direction (V-2DLDA) space the DFD was proposed to compute and because of the V-2DLDA operates better among the DFD in matrix presentation dispersed in the class and attained better effectiveness. Initially, to describe the labels of the query samples the nearest neighbour (NN) classifier was applied. The local characteristics alterations were presented by the DFD that were vigorous to the expression illumination. The simulation results had demonstrated that gratifying outcomes were obtained by the projected mechanism where the correctness rates were large which meant that the projected mechanism had offered improved presentation comparative to the traditional mechanisms.

9. **Xin Song and Hong Bao** [9] in this paper the writer had outlined that by investigating the progressions occuring in the element information in the constant discovery of 'outward appearances', and joining the worldly highlights and the spatial highlights on checked and set up models, the author has consequently advanced a more quick and productive approach for acknowledgment of human outward appearances. Right off the bat, the component extraction based strategy adjusting the Bezier bend was received to remove the 2D include focuses from each casing in the video stream. Which thus gives the progressions of transient trademark bend by interfacing the element purposes of each edge picture along time cut. At long last the nonlinear capacity was used to fit the progressions of the fleeting qualities bend and we set up models for every articulation to characterize and perceive. The outcomes demonstrate that the technique which consolidates the

fleeting and spatial changes in outward appearance acknowledgment had the upsides of fast and acknowledgment rate.

10. **Elizabeth Tran et al** [10] in this paper the author had anticipated an instrument for feeling acknowledgment from facial symbolism. This issue is observed to be exceptionally testing to a limited extent initially in view of the subjectivity of ground truth names and besides partially in view of the presence of similarly little marked datasets. We utilize the FER+ dataset [8], which is a dataset with various feeling marks per picture outline, so as to fabricate a feeling acknowledgment display that incorporates a total scope of feelings. Since the substance of information in the FER+ dataset is constrained, we consequently investigate the utilization of a significantly bigger face dataset, MS-Celeb-1M , in conjunction with the FER+ dataset. Particular layers inside an Inception-ResNet-v1 model prepared for facial acknowledgment were utilized for the feeling acknowledgment issue. Consequently , we use the MS-Celeb-1M dataset notwithstanding the FER+ dataset and explore different avenues regarding distinctive basic plans to evaluate the general execution of neural systems to perceive feeling utilizing facial symbolism.

11. **Iqbal et al** [11] proposed novel local descriptor called as Neighbourhood-aware Edge Directional Pattern (NEDP) to conquer the issues. On the local neighbourhood rather than relying solely in order to illustrate the characteristic around a pixel, as prepared by the conventional local descriptors. Despite of the occurrence of the slight distortion and noise in local area for the consistency of the characteristic the broader neighbourhood was investigated by the neighbouring pixels with the gradients at the target pixel that were analyzed by the NEDP. For the neighbouring pixels the template orientations were established by which in the consistent edge directions the significance was given to the gradients, preferring the particular neighbours falling in the direction of the local edge in order to present the size of the local textures, explicitly. The simulation results had demonstrated that the presentation of the NEDP was improved comparative to the conventional descriptors and also the entire presentation of the facial expression recognition was enhanced.

12. **Abubakar M. Ashir and Bayram Akdemir** [12] In this paper the author had presented a novel instrument for the recognizable proof of the outward appearance. The approach has imbedded in it both new component extraction method and classification techniques utilizing programmed auto-tuning of portion parameter enhancement in support vector machines. It for the most part starts with highlight extraction from the info vectors utilizing a blend of arithmetic means contrast and pivot invariant Local Binary Pattern. The extricated highlights are consequently anticipated into a Gaussian space to coordinate it with the "outspread premise work part" utilized as a part of help vector machines for order. Before characterization, an optimized parameter for "bolster vector machines" preparing were naturally decided in light of an approach proposed which depends on the beneficiary working qualities of the 'bolster vector machine' classifier. The outcomes got from the tests were impressive and promising. From the examinations directed on the two outward appearance databases with changed cross-approval systems, the proposed approach beats.

13. **Yacine Yaddaden et al** [13] in this paper the author had delineated that the human Computer Interaction has a huge effect in various fields of "Data and Communication Technology". It was essentially because of the significance of association between people and the advances they were utilizing. Consequently, 'outward appearance acknowledgment' has been generally used to improve this connection and make it more common. A large portion of the proposed techniques depended on pictures and even if they indicated great exhibitions, they don't coordinate the genuine cooperation model of individuals. . As the occurrence of the temporal aspect was taken into an account so the video contained a large amount of information. On the basis of the geometric characteristics a video was utilized in order to identify the facial expressions in the projected method as illustrated in the paper. We have tested it on a popular dataset and the carried experimentations showed promising results

14. **Xie et al** [14] in this paper the author had illustrated that a novel framework, named intra-class variation reduced features-based manifold regularisation dictionary pair learning model, was offered to resolve facial expression recognition (FER) jobs. The author had created intra-class variation reduced features (IVRF) from the

variation in the query face image and its equivalent approximated image of every expression class as a query face and its equivalent image with intra-class differences for instance, identity and illumination were identical in emergence. IVRF can reduce negative influence from the 'intra-class variations' and make their model robust to intra-class variations. Moreover, into the dictionary pair learning model a manifold regularisation term was incorporated that directed to an efficiently shifting sparse presentation. Their model fully takes advantage of the geometrical structure of data, which benefits the FER task. The experimental results on two public databases verify the effectiveness and superiority of their method and indicate its promising capability in expression discrimination.

15. **Revina et al** [15] in this paper the creator had outlined that the real goal of the Facial Expression Recognition (FER, depended on facial data to watch and realize human feelings. It was an energizing and urgent problem to recognize the human facial "appearance and feeling". This paper suggests a Gaussian based Edge Detection and Texture Descriptor (GEDTD) for FER. With respect to Gaussian edge descriptors GEDTD was framed. The proposed GEDTD extricate both picture surface component and edge direction. Utilizing Local XOR Coding (LXC) scheme the inside and region pixels of edge reaction headings were encoded for extraction. At last these highlights were combined and it frames the element vector.The articulations of different postures likely disturb, dismal, grin and surprise were prepared by utilizing Convolution Neural Network (CNN), which separates the outward appearances into disgust, miserable, grin and astonishment. The proposed procedure expands the acknowledgment exactness at a vital level. The under taken strategy was a suitable one for any acknowledgment necessities.

16. **Kim et al** [16] in this paper the creator had anticipated a new outward appearance acknowledgment instrument based on the attributes of the picture. There were two or three noteworthy procedures in the proposed framework, which was confront identification and facial expression recognition (FER). The face identification process had used Haar-like highlights, and the area of intrigue was reset to reduce the variable of appearance changes. The FER procedure removes histogram of arranged inclinations (HOG) highlights from every facial region, and afterward,

bolster vector machine was performed to group the last outward appearance. In the test comes about, the framework precisely recognized the outward appearance of someone in particular, and the proposed framework had the F1 score of 0.8759.

17. **Xie et al** [18] in this paper the author had projected a new mechanism in order to conquer the issue of the FER so the projected mechanism was known as Deep Comprehensive Multi-patches Aggregation Convolutional Neural Networks (DCMA-CNNs). This mechanism had majorly contained a couple of Convolutional Neural Network (CNN) branches that was deep-based framework. In this mechanism the local characteristics had illustrated the expressional details and holistics features classified the high-level semantic information of an expression. The local and holistic characteristics were aggregated by previously constructing categorization. In various scales these a couple kinds' hierarchical characteristics were offered expressions. Among the present mechanisms among single type of characteristics were measured among single type of characteristics, the projected can be presented expressions more in detail. In addition, a new pooling strategy called as Expressional Transformation-invariant pooling (ETI-pooling) was projected in order to control the nuisance fluctuations like rotations, variant illuminations, etc. in the training stage. The simulation results had demonstrated that the projected mechanism had offered improved results rather than the conventional mechanisms.

18. **Sharma et al** [20] in this paper the creator had shown that in this computer era, the most fascinating thing was to decide the human feelings with machines.With headway in the field of artificial intelligence, machines were additionally figuring out how to comprehend these human feelings. With the advancement of profound learning in the field of computer vision, PCs were utilized for tackling issues like object identification, movement acknowledgment, abnormality location, and video reconnaissance. With our work, we endeavored to utilize same deep learning strategies to handle the issue of "feeling recognition". A portion of the previous works investigated the capacities of "profound learning" and accomplished the effectiveness more than ever. Along these lines, to make the connection between human and computer more practical we prepared the machine to understand non-

verbal correspondence in type of feelings. In this paper, we utilized a novel way to deal with identify the feeling in view of the lips structure over the timeframe. The utilization of the recurrent neural system to break down the example with time gave the order of feelings into 6 distinct classes. To quantify the presentation of the anticipated component and best in class system the subjective among quantitative figuring was finished. D-FES can perform in the constant condition for precise following and arrangement of feelings.

19. **Sadeghi et al** [21] in this paper the maker had demonstrated that in this computer era, the most interesting thing was to choose the human emotions with machines.With progress in the field of artificial intelligence, machines were furthermore making sense of how to understand these human sentiments. With the headway of significant learning in the field of computer vision, PCs were used for handling issues like object distinguishing proof, development affirmation, variation from the norm area, and video surveillance. With our work, we attempted to use same deep learning methodologies to deal with the issue of "feeling acknowledgment". A segment of the previous works examined the limits of "significant learning" and achieved the viability like never before. Thusly, to make the association between human and computer more pragmatic we arranged the machine to understand non-verbal correspondence in sort of sentiments. In this paper, we used a novel method to manage recognize the inclination in perspective of the lips structure over the time period. The use of the recurrent neural framework to separate the case with time gave the request of sentiments into 6 unmistakable classes. To evaluate the presentation of the foreseen segment and best in class framework the subjective among quantitative figuring was done. D-FES can perform in the steady condition for exact after and plan of emotions.

20. **Ekweariri et al** [22] in this paper the author had projected the facial expression recognition by utilizing the local binary patterns. In this mechanism by applying the feature selection in order to present the faces the large variance pixels were chosen. The identification rates were enhanced effectively through choosing the large variance pixels on the basis of the LBPs. On the BU-3DFE database the tests were finished. The simulation results had demonstrated that the recognition rates were enhanced by utilizing the feature selection mechanism.

21. **Fathallah et al** [23] in this paper the author had illustrated that in the computer vision the Automated Facial Expression Recognition was a challenging and interesting issue. For the machine learning mechanisms the facial expressions identification is the major issue as the expressions of several persons were altered effectively. In the machine learning mechanism the Deep Learning is a novel region of research that can categorize image human images of human faces into emotion classifies by applying Deep Neural Networks (DEEP). In order to conquer the issues in the facial expression categorization the Convolution neural networks (CNN) had been broadly utilized. The author had projected a novel structural design on the basis of CNN for the identification of the facial expressions in this work. Among the Visual Geometry Group model to enhance the outcomes the projected structural design was tuned. Among various hugely public databases (CK+, MUG, and RAFD) it was verified in order to calculate the structural design. The simulation results had demonstrated that in the image expression recognition on various public databases the CNN method was much efficient that was attained an enhancement in the examination of the facial expression.

22. **Lu et al** [24] in this paper the author had projected a fresh facial expression recognition mechanism on the basis of the discrete shearlet transform that was a novel image multiscale geometric analysis mechanism. Additionally the wavelet transform had owned the multi-resolution and time-frequency localization, the anisotropy and directionality had owned by the shearlet transform. Initially, to the entire test and training images the normalization and equalization were utilized. Therefore on the basis of the discrete shearlet transform the characteristics of the facial expression were extorted and the seven expressions that were happiness, sadness, surprise, disgust, fear, anger and neural of JAFFE database were categorized by applying the Vector Support Machine. The simulation results had demonstrated that the recognition rate of the projected mechanism was enhanced comparative to other existing mechanisms.

23. **Lin et al** [25] in this paper the creator had anticipated a component with a specific end goal to learn spatiotemporal attributes for the outward appearance ID based on the video among multi-layer independent subspace investigation (ISA) worldview.

On the main layer, an arrangement of ISA channels are learned from little 3D patches of the video information, and then more conceptual and ground-breaking highlights on the second layer are learned from the element reactions of the first layer. Two public outward appearance databases, $extended Cohn-Kanade and MMI were utilized to assess our technique. Exploratory outcomes had demonstrated that the highlights learned by multi-layer engineering accomplish better recognition execution than that of single-layer display. Additionally, the celebrated hand-created qualities were beated by the anticipated instrument and the whole accuracy of the component was quantifiable to different relative systems based on the element learning

24. **Sajjanhar et al** [26] in this paper the creator had analyzed the "outward appearance acknowledgment" in view of geometric highlights, and image appearance, utilizing a scope of classifier models. To start with, we assess "demeanor acknowledgment" of face pictures in view of geometric features, in particular, facial milestones in view of the Active Appearance Model, and Action Units in view of the "Facial Action Coding System". A generalized linear display and a neural system were utilized to classify confront pictures in light of these geometric highlights. Second, the classification achieved by facial points of interest and Action Units was compared with the order accomplished by "convolutional neural system (CNN)" which extricates picture highlights from crude pixels. $Finally, we utilize exchange learning procedure to assess classification using a pre-prepared model. All experiments are performed on the best in class CK+ confront database.

25. **Varanya P V et al** [17] in this paper the author had projected a new mechanism for the investigation by applying the LBP characteristics of the chosen salient patches on CK+ video Data Base. On the basis of the location of the facial landmarks the less important active patches were extorted. Mechanically, the locations of the landmarks were achieved. Through 10 fold cross validation these active patches were additional processed in order to achieve the salient patches. From the salient patches the extorted characteristics enhance the computational cost among less fluctuation in the correctness. The extorted characteristics were

categorized by having the SVM and the ANN classifier. The outcomes were also compared in this mechanism.

2.2.SUMMARY

The existing FER(Facial Emotion Recognition) approach is realised using a number of methods that are commonly detecting the facial regions and extracting geometric features, the other approach is related to the appearance features . also the other approach comes into picture when the two above are combined to give the hybrid approach.

For geometric features the positional points or the descriptive areas are considered. The approach had been used by various as **Wang et al.** [4] that used the landmark points with for selected areas along with the HOG features with the SVM as a classifier. On the other hand **Xie et al.** [14] used an intraclass variation reduced features based manifold termed as IVRF technique that focused on the facial geometrical interpretation

The appearance model usually focuses global facial region other than local unlike geometrical features, this includes different face regions that contain different information. **Chao et al**. [5] developed an Local Binary Pattern mechanism that considers the appearance model solely. **Ekweariri et al.** [22] also considered a feature selection technique by the appearance model.

However there are models that consider both the appearance pattern and are considered as the hybrid model . **Akedemir et al.** [12] applies the hybrid model approach with the feature vector consideration and enhanced SVM , while **Yaddaden et al.** [13] developed an outward appearance system considering the video data input .

Other than the mentioned approach models a major depth is seen in the neural network models .**Xie et al.** [18] introduced the DCMA-CNN deep based approach where ETL pooling is used to control fluctuations. This method considers the local feature pattern. **Li and Zhi** [3] used the CNN for distributed abstract characteristic extortion. While a new model 'EmoTrain' was developed by **Tsangouri et al.** [6] using the concept of deep learning. **Nwosu et al.** [1] also considered the deep CNN concept where the facial parts are separately analysed ,channelled ,combined and analysed for the accuracy in performance. **Revina et al.** [15] gives a GEDTD model approach derived from the CNN concept that considers the texture ,edge and surface features.

There are a few other approaches that differ from the conventional approaches for FER system. **Sun et al.** [8] unique method 'discriminative feature learning' mechanism , consists of a DF Dictionary consisting of 2D LDA features and an NN approach for classification. Another method reinstating the facial symbolism where the images have different feeling marks per picture outline , this model is called inception model which has NN in its final stage by **Tran et al.** [10]**. Iqbal et al.** [11] proposed a local descriptor (NEDP) that considers the pixel based model for the orientation template that proves a better advancement of conventional descriptors. **Taheri et al.** [7] gave a nonlinear correlation filter bank mainly focusing the correlation of the extracted values .

Sharma et al. [20] **and Sadhegi et al.** [21] mainly focus their approach on DFES model that uses the concept for inclination of structures.**Kim et al.** [16] uses a very interesting approach using the haar like features along with the HOG finally using the bolster vector machine.

Tonchev et al. [2] developed the automated expression recognition where the feature extraction is done using log-gabor filter finally using the SVM classifier.

Hence the different approaches considered above use different schemes of processing the emotion data from the input images while each having its own limitations.

CHAPTER 3

PROBLEM FORMULATION AND PROPOSED WORK

3.1 Problem Statement

Automated facial emotion recognition is an essential step for proper Human-Machine Interaction (HMI) since much of human-human interaction occurs outside of our speech and tone of voice. Most of the researches used distinct databases for the evaluation of their techniques, so it is very challenging task to infer which one is superior. They have utilized the traditional approaches like PCA, Color shape size like feature extraction techniques.

The results obtained from those techniques are not very efficient and were not providing relevant information to achieve high accuracy, also the unnecessary information was also there in traditional approaches, that’s why it is need to be changed. On other hand the different classifiers used for processing extracted features from different techniques have different accuracy and performance. So, there is also need to check performance of different classifiers as per enhanced feature extraction module.

3.2 Objective:

1) To implement the haar like features and F-LDA feature extraction approach
2) To implement the filteration module to remove non usable feature from extracted data
3) To implement the KNN, SVM, ANN classification approach
4) To perform comparative analysis over Accuracy, ROC curve and

3.3 Proposed Work

In traditional human Emotion recognition techniques there was need to enhance a feature extraction approach. So, in the proposed work we will be using enhanced feature extraction techniques. We are using Haar like and Gabor features extraction technique. Gabor wavelet mechanism is used to extract the texture feature and Haar like features are a enhanced version of traditional Discrete wavelet transform (DWT) it is inspired from DWT. The features will be extracted using these techniques will be Filtered and only required information will be given to further processes. This is also a problem taken from traditional approach as they process whole data that is time consuming and increase a complexity in classifiers training. So, using filter before giving extracted features to classifier will reduce the time as well as complexity in training classifier and results in efficient output. Next to this as the feature extraction approaches are changed it will be hard to finalized the classifier best for this, for that we will implement it using SVM, KNN and ANN. The results will be calculated on the behalf of accuracy, ROC curve and Convenience matrix.

3.4 Methodology

The figure 5 delineates the flow work of proposed work. As per the flow chart, first most step is to select the dataset for the purpose of the simulation. After electing the dataset, next step is to extract the features from the images and for this purpose, the Gabor feature extraction and Haar-Like feature extraction is applied individually. After extracting the features, the feature selection is done by using the infinite feature selection technique. Then two different images with extracted features are fused by using the PCA mechanism. Next the different classifiers i.e. SVM, ANN and KNN is applied for the purpose of the classification. Last but not the least step is to evaluate the accuracy in order to prove the proficiency of the proposed mechanism of human emotion recognition.

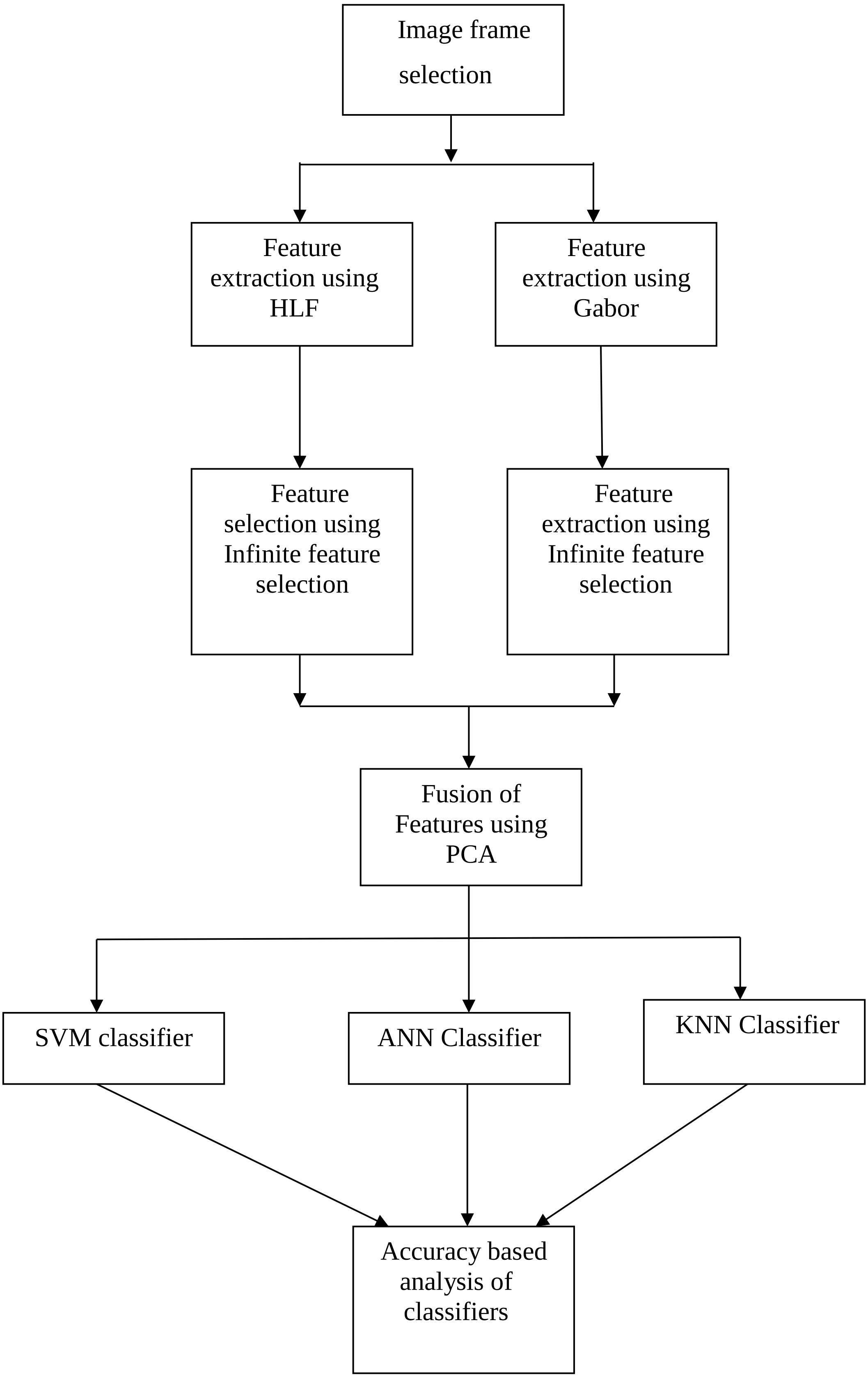

Figure 3.1. Flow diagram of proposed methodology

3.4.1 HAAR Features

Haar like feature mechanism is used to identify the difference among black and light segment of an image. In haar like features, a rectangle is created around the detected face. In this technique, on the basis of the colour shades near nose and forehead, the contouring is done. The most common haar features are:

1. Two Rectangle feature
2. Three rectangle feature
3. Four rectangle feature

Following equation is formulated to evaluate the haar like features:

$$features = \sum ie\{1, \ldots. N\}wi. RecSum(x, y, w, h) \qquad (4.1)$$

In equation (1), RecSum (x,y,w,h) is a summation of intensity corresponding to any upright or rotated rectangle that is enclosed in detection window. X,y,w,h denotes the coordinates, dimensions and rotation of the rectangle.

In present work, the six rectangle haar like feature extraction technique is used. The working of six rectangle haar like feature is as follows:

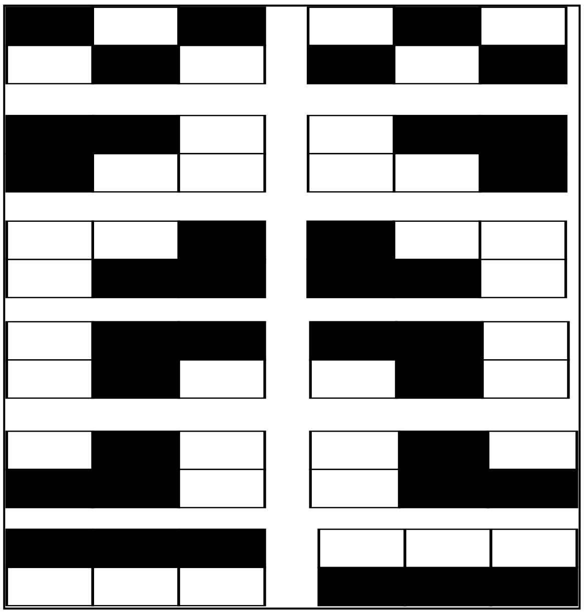

Figure 3.2 Rectangle Haar like features used in proposed work

3.4.2 Gabor Feature Extraction

Gabor feature extraction is used as a image processing or image filtration mechanism. It is alienated into two series of one- dimensional ones. This approach is essentially used to sense the edges, corners and splotches of the face image. Gabor functions assist to extort the facial appearance especially in texture-based image scrutiny. The two approaches i.e. edge detection and corner detection. The edge detection s done on the feature image and corner detection is done with the help of combination of responses to several filters with a different orientation. The equation (1) defines the Gabor filtration in spatial domain.

$$\psi(x,y) = \frac{f^2}{\pi\gamma n} e^{-\left(\frac{f^2}{\gamma^2}x'^2 + \frac{f^2}{\gamma^2}y'^2\right)} e^{-i2\pi f x'} \tag{4.2}$$

$$x' = x\cos\theta + y\sin\theta$$

$$y' = -x\sin\theta + y\cos\theta$$

f denotes the central frequency, θ denotes the rotation angle, γ is a sharpness or bandwidth along the major axis, n defines the sharpness along the minor axis. This formulation for frequency domain is implemented as follows:

$$\psi(u,v) = e^{-\frac{\pi^2}{f^2}\left(\gamma^2(u'-f)^2 + n^2 v'^2\right)} \tag{4.3}$$

$$u' = u\cos\theta + v\sin\theta$$

$$v' = -u\sin\theta + v\cos\theta$$

3.4.3 PCA

Principal component analysis i.e. PCA is a arithmetical course of action to exploits an orthogonal transformation. It is used to diminish the measurement of the data by revenue of data compression and divulges the most efficient low dimensional structure. Principal Component Analysis is a mathematical and statistical method and is useful in biometric recognition & image compression. It is a very usual technique used for searching patterns in high dimensioned data. It also highlights the differences and similarities between the dataset. The Pseudo code for PCA is as follows:

Produce the column vector from input image

Evaluate the covariance matrix of two column vector created in previous step.

$$m_X = \left(\frac{1}{M}\right) \sum_{k=1}^{M} X_k \tag{4.4}$$

m_X denotes the mean vector related to the population of vector x. For M vector samples from random population, the covariance matrix is as

$$C_X = \left(\frac{1}{M}\right)\sum_{K=1}^{M}[X_k X_k^T - m_X m_X^T] \tag{4.5}$$

Here T refers to the transposition.

Evaluate the eigen value and eigen vector of the covariance.

Normalize the column vector. Since C_X is real and symmetric matrix thus, it is easy to locate a set of n orthogonal vectors. Let e_i and $\lambda_i, i = 1,2, \dots \dots, n$ as an eigen vector corresponding to the eigen values of C_X, arrange in decreasing order so that

$$\lambda_j \geq \lambda_j + 1\ for\ j = 1,2, \dots .. n-1 \tag{4.6}$$

Normalize the vector as per the weight value and multiply it with each pixel of the input image.

Fuse both images scaled metrics will be the fused images matrix.

3.4.4 SVM

This classification technique is a linear classification technique that was introduced by Vapnik et . al. [19] , this was introduced mainly to handle with the classification of humongous amount of data and to separate the values based on a set of features.

This is also part of a supervised learning mechanism that takes the concept of hyperplanes in consideration, the error mechanism is considered.

As the error value of the subject with respect to the present values decreases the plane transforms accordingly. The margins are increased and it can also be said that the increased amount of separation is considered [20].

SVM has used to classify the face image into appropriate category. Steps followed by SVM classifier is follows as:

1. Choosing a kernel function.
2. Selecting a value for *C*.
3. Solving the quadratic programming problem.

4. Constructing the discriminate function from the support vectors.

3.4.5 KNN

To perform the classification by using KNN classifier first we need to develop a data set. The data set comprised of certain attributes. After creating the data set, the data set is divided into two parts i.e. one for training purpose and other for testing purpose. The data set for training is passed as an input to the classifier whereas the testing data set is used for testing purpose on the basis of the trained dataset. The mathematical model for KNN depicts that it only utilizes the local prior possibilities for classification purpose [12]. For given query x_t the classifier works as follows:

$$\boldsymbol{y_t} = \underset{c\in(\boldsymbol{c1},\boldsymbol{c2},\ldots\ldots\boldsymbol{cm})}{\boldsymbol{argmax}} \sum \boldsymbol{x_i} \in \boldsymbol{N}(\boldsymbol{x_t},\boldsymbol{k})^{\boldsymbol{E}(\boldsymbol{y_i},\boldsymbol{c})} \tag{4.7}$$

Where y_t denotes the predicted class corresponding to thex_t, m depicts the classes that are included in data. Also,

$$\boldsymbol{E}(\boldsymbol{a},\boldsymbol{b}) = \begin{cases} \boldsymbol{1}\ \boldsymbol{if}\ \boldsymbol{a} = \boldsymbol{b} \\ \boldsymbol{0} \quad \boldsymbol{else} \end{cases} \tag{4.8}$$

$$N(x,k) = set\ of\ k\ nearest\ neighbor\ of\ x.$$

The above defined equation (1) can also be formulated as follows:

$$argmax\left\{\sum x_i \in N(x_i,k)^{E(y_i,c1),\sum x_i \in N(x_i,k)^{E(y_i,c2),\ldots\ldots} \sum x_i \in N(x_i,k)^{E(y_i,cm)}}\right\} \tag{4.9}$$

$$y_t = argmax\left\{\sum x_i \in N(x_i,k)^{E\frac{(y_i,c1)}{k},\sum x_i \in N(x_i,k)^{E\frac{(y_i,c2)}{k},\ldots\ldots} \sum x_i \in N(x_i,k)^{E\frac{(y_i,cm)}{k}}}\right\} \tag{4.10}$$

It is known that

$$p(c_j)_{(x_t,k)} = \sum x_i \in N(x_t,k)^{\frac{E(y_i,c_j)}{k}} \tag{4.11}$$

$p(c_j)_{(x_t,k)}$ Denotes the probability of occurrence of jth class in neighbor x_t.

Thus the equation (4.9) turns to be equation (4.11)

$$y_t = argmax\{p(c_1)_{(x_t,k)}, p(c_2)_{(x_t,k)}, \ldots \ldots . p(c_m)_{(x_t,k)}\} \tag{4.12}$$

On the basis of above formulation it is defined that the KNN performs the classification on the basis of the local possibilities. Following is the pseudo code for KNN that is implemented in proposed work:

1. Evaluate $d(x,x_i)\, \mathbf{i} = \mathbf{1}, \mathbf{2}, \dots, \mathbf{n}$
2. d refers to the Euclidean Distance among the points. It is evaluated as follows:

   $$d = \sqrt{(y1 - x1)^2 + (y2 - x2)^2} \tag{4.13}$$
3. Let k a positive integer; consider the first k distance from d.
4. Evaluate k point corresponding to k distances.
5. Assume $\mathbf{k_i}$ to define the number of points related to the $\mathbf{i^{th}}$ class among k. $\mathbf{k} \geq \mathbf{0}$

   $If\ K_i > k_j \forall i \neq j\ then\ put\ x\ in\ class\ i.$

3.4.6 FFNN

Feed Forward Neural Network (FFNN) is the simplest and easy to implement type of Artificial Neural Network (ANN). In FFNN, the flow of information is in single direction i.e. the data is forwarded to the hidden nodes and then output nodes from input nodes. It did not comprise of loops or iterations. The example of FFNN is Single Layer Perception (SLP) and Multi Layer Perception (MLP). The proposed FFNN model uses 5 neurons and $traingdx$ as training function. $traingdx$ is a training function that is used to update the weights and bias values as per the gradient descend momentous and an adaptive learning rate. The function is used as follows:

$$\mathbf{traingdx}(\mathbf{net}, \mathbf{Pd}, \mathbf{Tl}, \mathbf{Ai}, \mathbf{Q}, \mathbf{TS}, \mathbf{VV}) \tag{4.14}$$

Where, net - Neural network, Pd - Delayed input vectors. Tl - Layer target vectors. Ai - Initial input delay conditions. Q - Batch size. TS - Time steps. VV - Either empty matrix [] or structure of validation vectors.

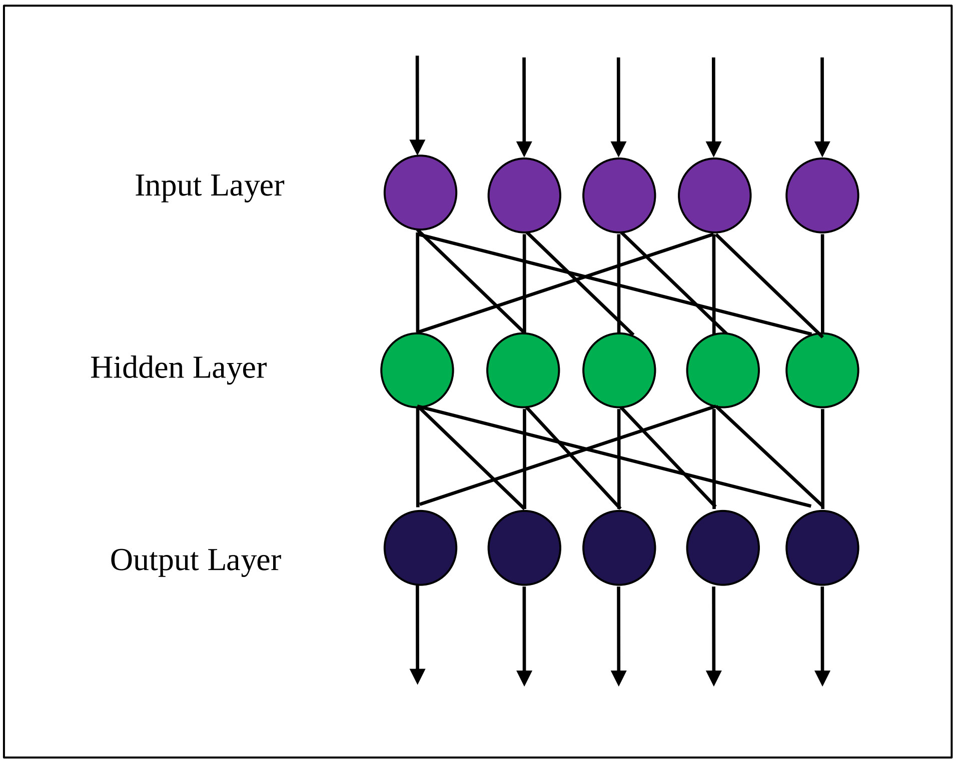

Figure 3.3 Proposed FFNN

CHAPTER 4

RESULTS AND EXPERIMENTS

4.1 Experimental Platform

MATLAB is a programming language that is specifically developed for the purpose of research work It is developed by company named ‘Mathworks’ in 1970 by Cleve Moler. The concept applied in the language development was from C++ and JAVA. As its full name suggests “Matrix Laboratory” this language mainly facilitates working with matrix with ease.

There are tools present that facilitate working in various platforms for the scientific plots, the neural network iterations ,mathematical computations etc. the environment provided by the platform is of the high level language that enables the task of suitably performing the arithmetic ,calculations , solve matrices and perform the linear algebraic solutions.

The generation of the language comes under fourth and it combines various factors in one that makes it the most preferred language in the research domain.

In the performed study the various tools facilitated by matlab were considered, a few of the objectives are mentioned as below :

- The matrix calculation and array approach regarding the value based operation while feature extraction from the still images.

- The various mathematical tools for the feature fusion for the extracted image.
- While coming on to the classification part the inbuilt mathematical tools available for the classification by using all the three SVM,KNN and ANN was made easy as matlab has inbuilt functions for the training and classification.

Apart from the mentioned functions matlab performs many other functions that can be mentioned as:

- The graphs for any function or mathematical interpretation can be easily plotted.
- Manipulation of matrix is made easy.
- Algorithmic implementation.
- Graphical user interface can be created.
- The code associated with other languages like C,C++ can be executed by using the MATLAB platform.

4.2 Dataset

In this study, different three types of datasets have been considered for the purpose of validation of algorithm. These datasets are comprised of 10 different folders and each folder comprised of 30 different images of the same person with different expressions. Thus in this way it can be said that each dataset comprised of total 300 images of 10 persons with different facial expressions. This section depicts some of the sample images from each dataset.

4.2.1 cohn-kande dataset

This dataset was developed by (Kanade, Cohn, & Tian) [27] . the class of dataset includes 486 sequences that is taken from 97 posers. Each sequence initiates with a neutral expression and proceeds upto a peak expression. The peak expression for each sequence is coded in fully FACS (Ekman, Friesen, & Hager, 2002; Ekman

& Friesen, 1979) and is given an emotion label. The assigned emotion label refers to what expression was requested rather than what may actually have been performed.

Figure 4.7 Sample images for Cohn-Kanade Dataset

4.2.2. JAFFE dataset

The JAFFE database [28] consists of 213 images consisting of 7 distinguished facial expressions (6 facial expressions + 1 neutral) posed by 10 Japanese female models. Each image had been rated on 6 emotion adjectives by 60 Japanese subjects.

Figure 4.8 Sample Images for Jaffe Dataset

4.2.3. Yale

This database [29] consists of 5760 images with single light source for 10 subjects where each subject observed under 576 viewing conditions (9 poses x 64 illumination conditions). Every particular subject with a particular pose, an image with background illumination is also captured. Hence, the total number of divided subject images is 5760+90=5850.

Figure 4.9 Sample Images for Yale Dataset

4.3 Experimental Results

This section is organized to depict the results of proposed simulation over Cohn-Kanade, Jaffe and Yale dataset.

The figure 5.4, 5.5 and 5.6 depicts the confusion matrix for Cohn-Kanade dataset by using the SVM, ANN and KNN classifiers individually. The confusion matrix is evaluated to measure the performance accuracy of the system with respect to output class and target class.

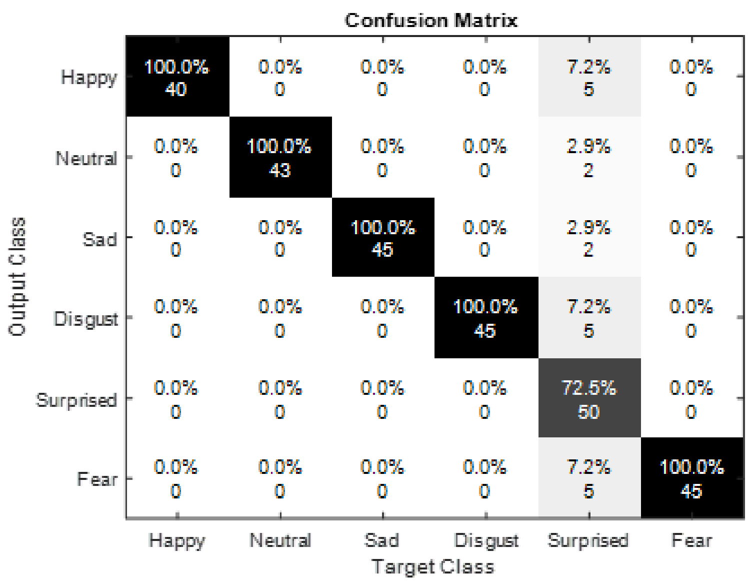

Figure 4.4 Confusion Matrix of Cohn Kanade Dataset by using SVM classifier

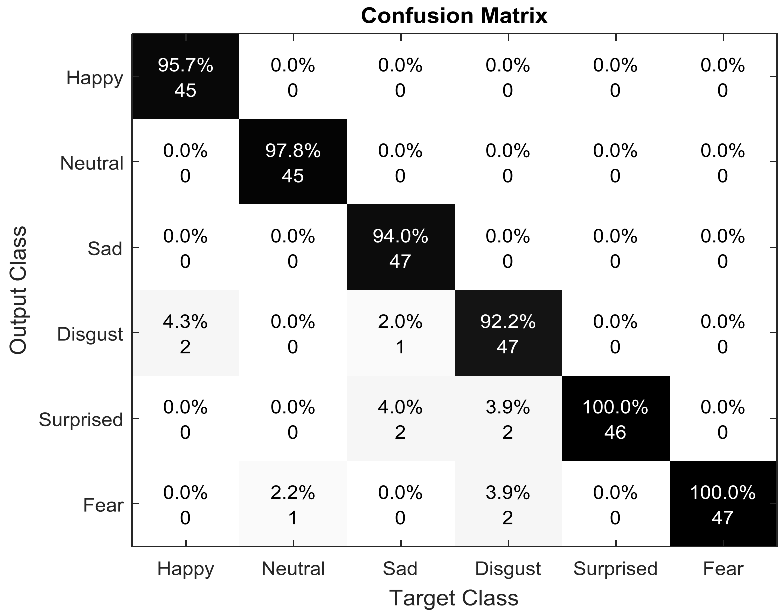

Figure 4.5 Confusion Matrix of Cohn Kanade Dataset by using ANN classifier

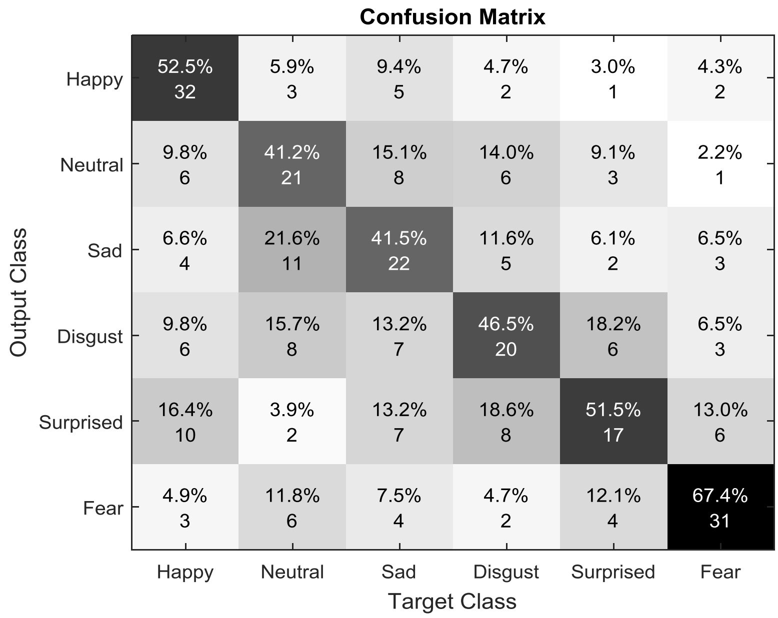

Figure 4.6 Confusion Matrix of Cohn Kanade Dataset by using KNN classifier

The graph in figure 4.7 shows the accuracy of proposed work that evaluated after implementing the Cohn-kanade dataset. The accuracy is evaluated with respect to the techniques used for the classification and represented by x axis. The y axis in the graph shows the values for accuracy and it varies from 0 to 100. As per the observations from the graph the accuracy for SVM classifier is 93.38, for ANN it is 95.47 and for KNN the accuracy is 49.83.

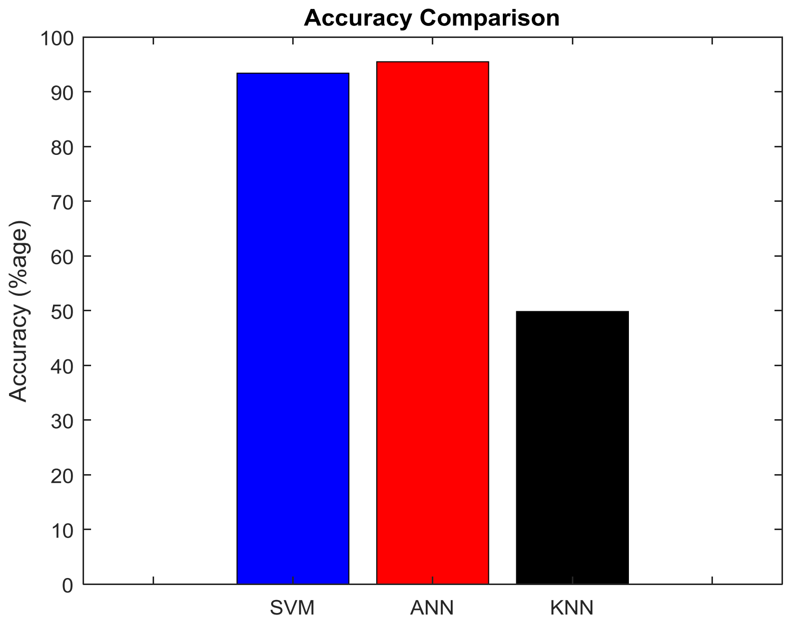

Figure 4.7 Accuracy evaluation of classifiers for Cohn-Kanade dataset

The comparison analysis of error in respective classifiers for Cohn-kanade dataset is shown in the graph of figure 5.8. The graph explains that the error rate in ANN classifier is lower than the error rate of SVM and KNN classifiers.

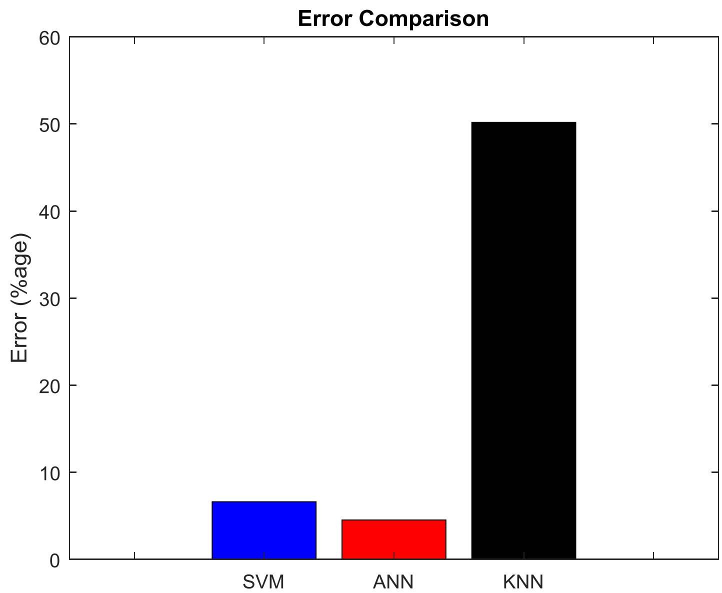

Figure 4.8 Error evaluation of classifiers for Cohn-Kanade dataset

The comparison graph of figure 4.12 shows the accuracy evaluated for Yale dataset by using different classifiers. The graph makes it clear that the accuracy of SVM classifier is 92.94

The confusion matrix for Yale dataset by using SVM, ANN and KNN classifier is shown in figure 4.9, 4.10 and 4.11 respectively.

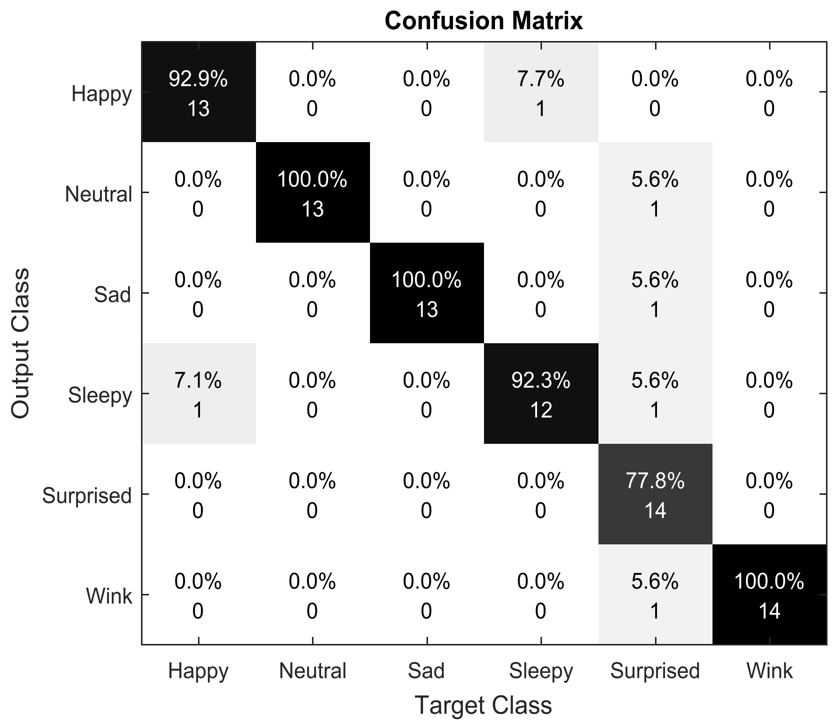

Figure 4.9 Confusion Matrix of Yale Dataset by using SVM classifier

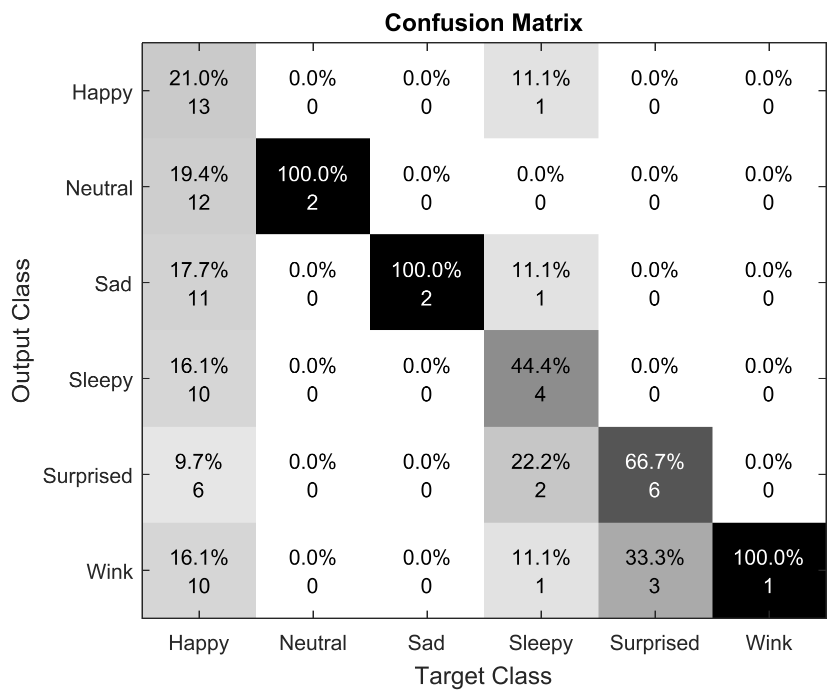

Figure 4.10 Confusion Matrix of Yale Dataset by using ANN classifier

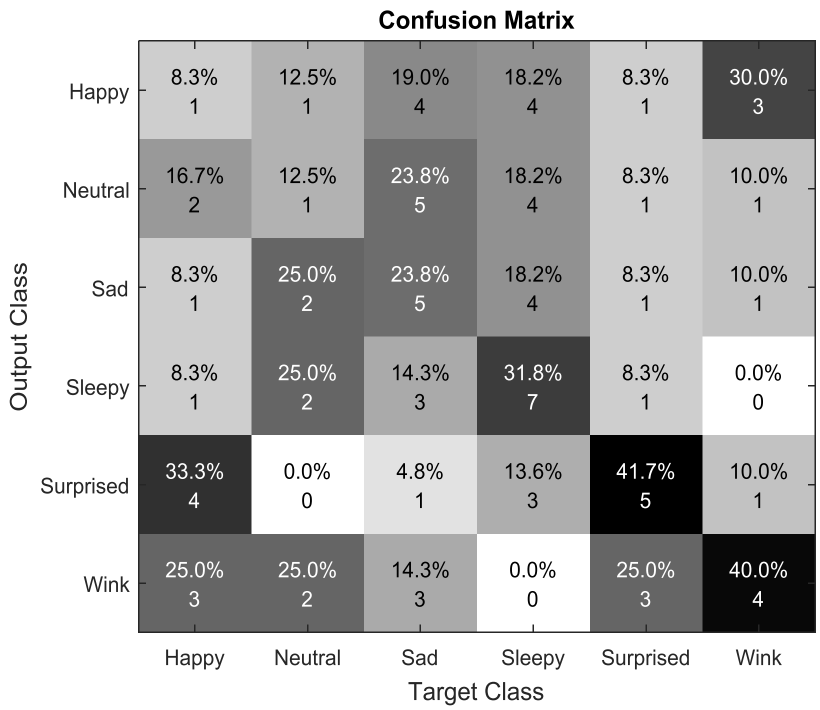

Figure 4.11 Confusion Matrix of Yale Dataset by using KNN classifier

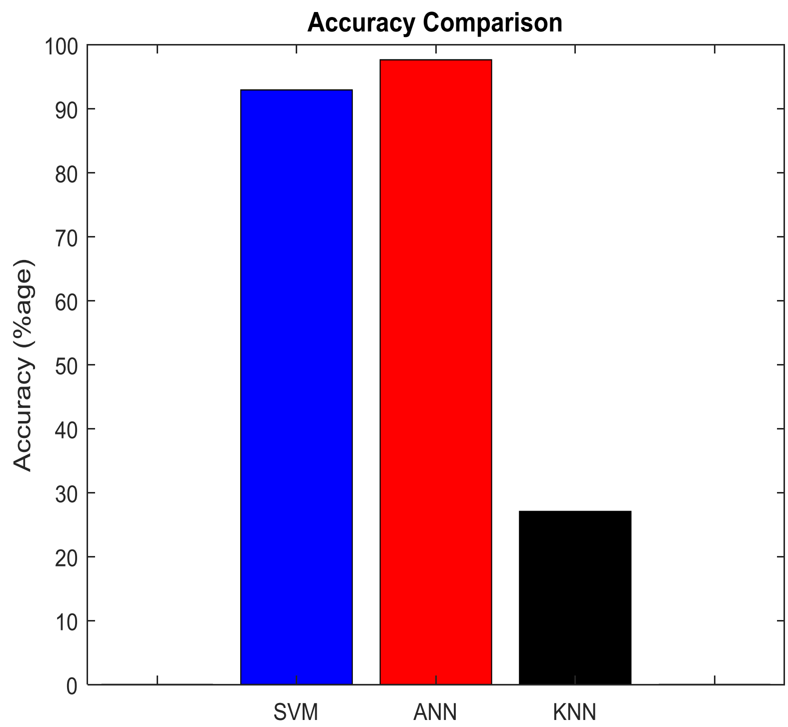

Figure 4.12 Accuracy evaluation of classifiers for Yale dataset

Similarly the graph in figure 4.13 shows the error rate that s evaluated in different classifiers for Yale dataset. The evaluated error rate for SVM is 7.059, for

ANN is 2.353 and for KNN it is 72.94. On the basis of the facts it is proved that the ANN classifier outperforms the KNN and SVM classifiers as it has comparatively low error rate.

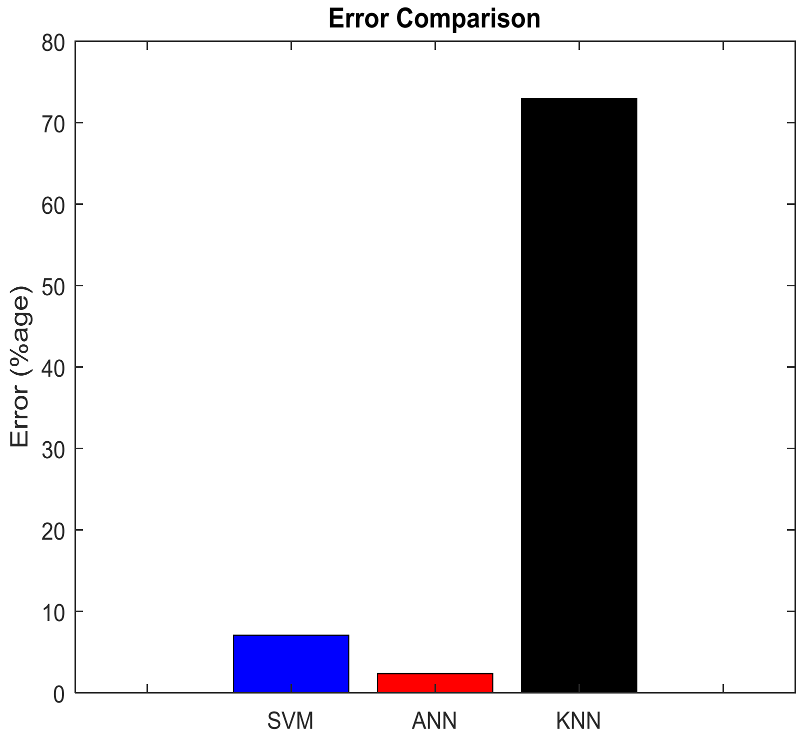

Figure 4.13 Error evaluation of classifiers for Yale dataset

The confusion matrix of figure 4.14, 4.15 and 4.16 is driven for evaluating the accuracy or performance of the SVM, ANN and KNN classifier for Jaffe dataset. The comparison graph in figure 4.17 and 4.18 delineates the accuracy rate and error rate of the SVM, ANN and KNN classifiers for Jaffe dataset. The graph depicts that the SVM classifier has the highest accuracy rate and error rate for Jaffe dataset.

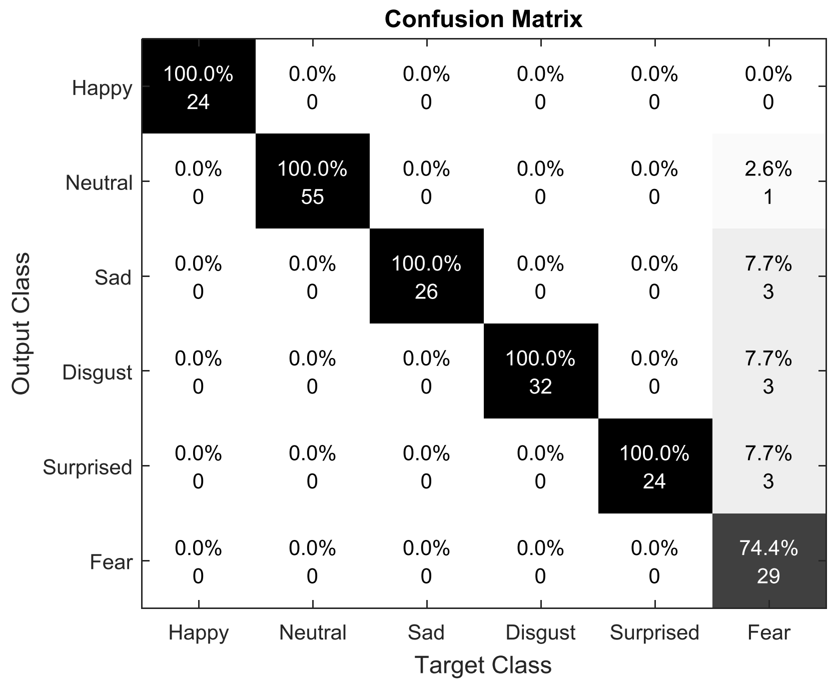

Figure 4.14 Confusion Matrix of Jaffe Dataset by using SVM classifier

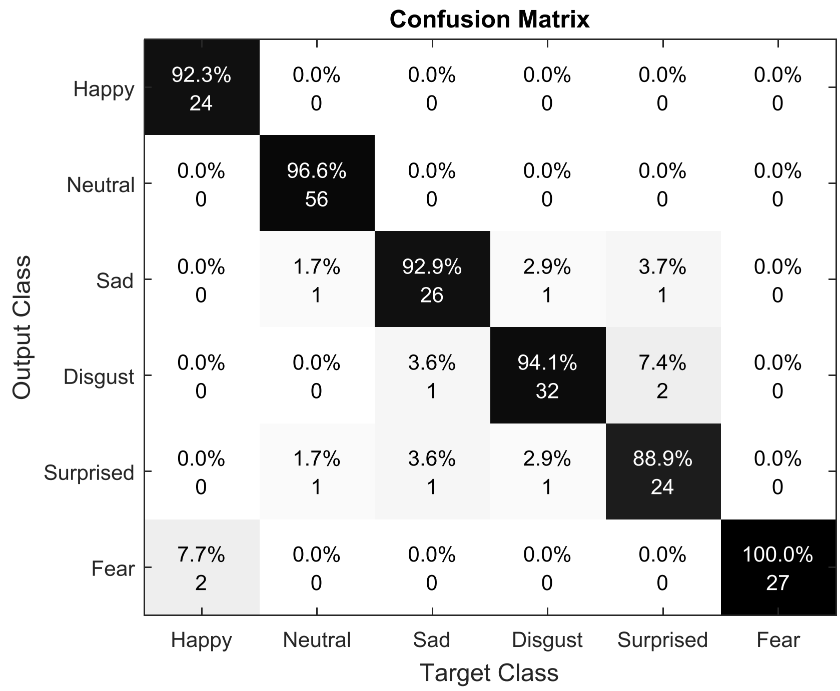

Figure 4.15 Confusion Matrix of Jaffe Dataset by using ANN classifier

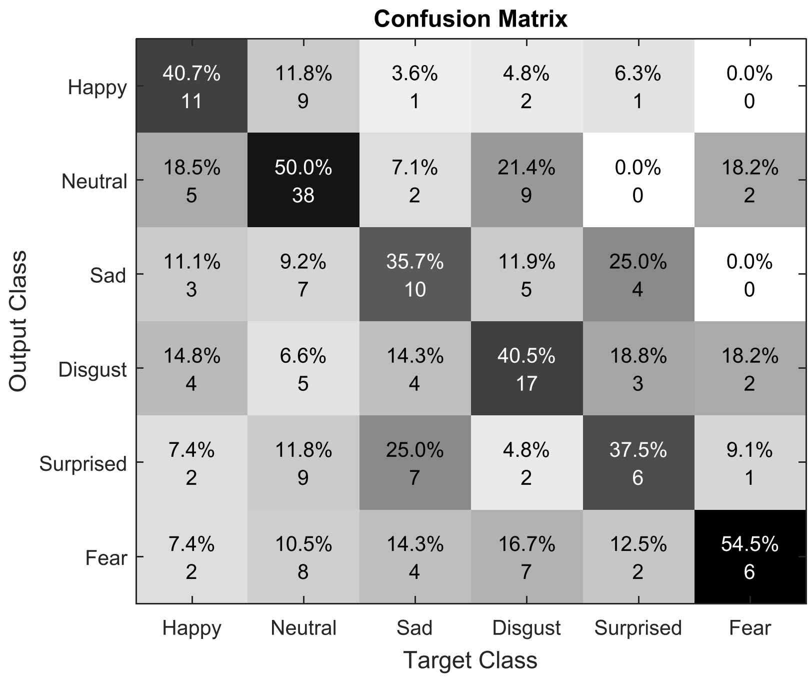

Figure 4.16 Confusion Matrix of Jaffe Dataset by using KNN classifier

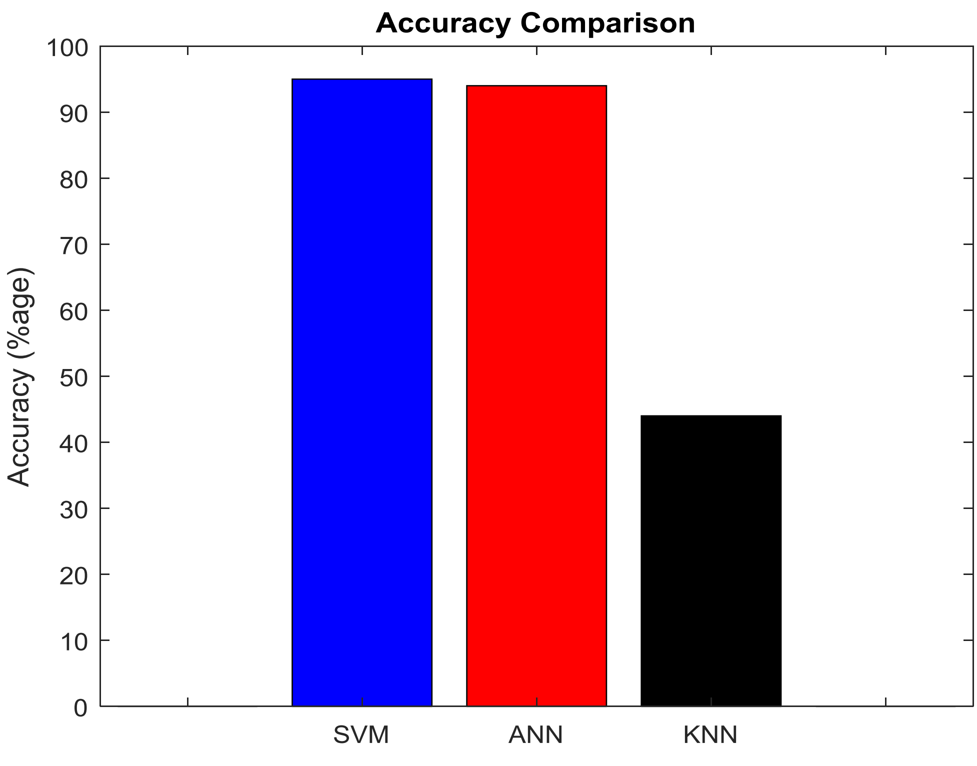

Figure 4.17 Accuracy evaluation of classifiers for Jaffe dataset

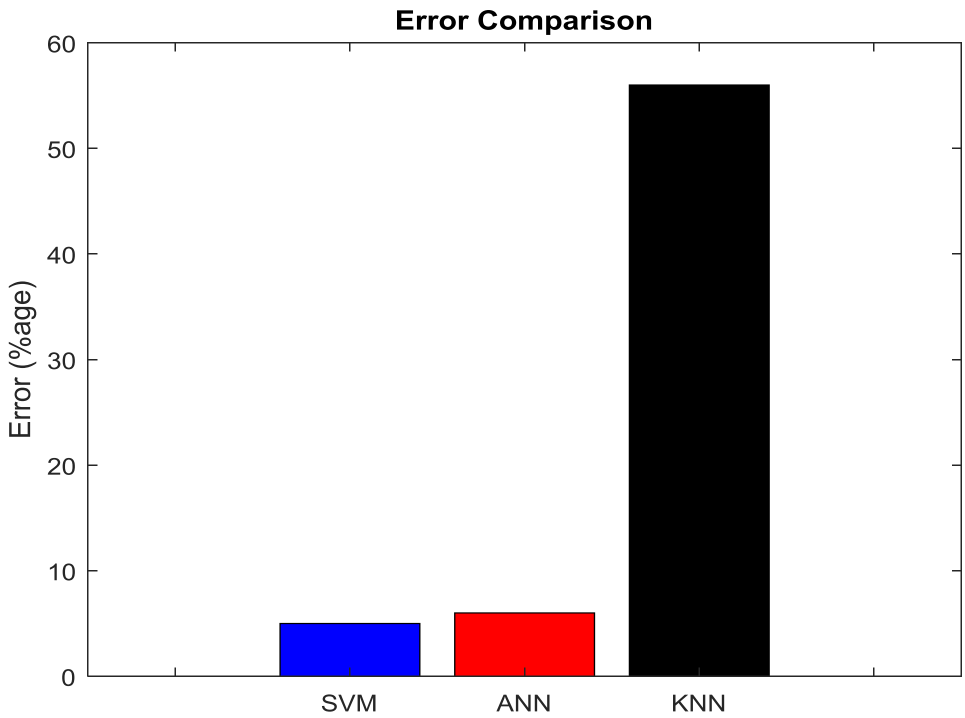

Figure 4.18 Error evaluation of classifiers for Jaffe dataset

The table 4.1 and 4.2 depicts the accuracy and error rate of the respective classifiers on the basis of the used datasets. The facts shown in the table are gathered on the basis of the results that are shown in above graphs.

Table 4.1 Accuracy Comparision

Classifiers	**Cohn-kanade**	**Jaffe**	**Yale**
SVM	93.38	95	92.94
ANN	95.47	94	97.65
KNN	49.83	44	27.06

Table 4.2 Error Analysis (%)

Classifiers	Cohn-Kanade	Jaffe	Yale
SVM	6.62	5	7.059
ANN	4.53	6	2.353
KNN	5.17	4	72.94

Table 4.3 Comparison of accuracy with previous methods on Jaffe dataset

Method	**Feature Selection Technique**	**Classifier**	**Accuracy %**
Kobayashi, et.al [30]	Facial Characteristic Points (FCP)	NN	85
Lyons, et.al [31]	Gabor,LG-PCA	LDA	92
Zhang [32]	Geometric position ,Gabor	NN-RPROP Propagation	90.1
Proposed method	Gabor+Haar-Like	ANN, SVM,KNN	97.65 , 92.9, 27

Table 4.4. Comparison of accuracy with previous methods on Cohn-Kanade dataset

Method	**Feature Selection**	**Classifier**	**Accuracy**
Pantic, et.al [33]	Encoded Action Units(AU)	NN,fuzzy logic	91
Sebe, et.al [34]	MU(Motion Units)	Naive–Bayes	88 to 95
Proposed Method	Gabor + Haar-Like	ANN,SVM,KNN	95.47,93.38,49.83

Table 4.5. Comparison of accuracy with previous methods on Yale dataset

Method	**Feature Selection**	**Classifier**	**Accuracy**
G. Edwards, et.al [35]	PCA	Mahalanobis distance	74
Proposed Method	Gabor + Haar-Like	ANN,SVM,KNN	94,95,44

CHAPTER 5

CONCLUSION AND FUTURE SCOPE

In this study, a classifiers based human expression detection mechanism is introduced. The Gabor and Haar-like feature extraction mechanism is applied to extract the important features from the selected images. The ANN, SVM and KNN classifiers are used for classification. The results are simulated by using Cohn Dataset, Yale Dataset and Jaffe Dataset. The facts in table prove that the ANN classifiers has highest accuracy rate i.e. 95.47 % for Cohn-Kanade dataset, 97.65 % for Yale dataset and 44% for Jeffe dataset. This study analyze the different dataset by extracting the texture features thus, in future more amendments are possible in near future by extracting some other important features that effects the output. Along with this, it is analyzed that the ANN classifiers outperforms the rest of the used classifiers such as SVM and KNN. In future deep learning mechanism can also be implemented in order to obtain more accurate results.

REFERENCES

[1] L. Nwosu, H. Wang, J. Lu, I. Unwala, X. Yang and T. Zhang, “Deep Convolutional Neural Network for Facial Expression Recognition using Facial Parts,” in *IEEE 15th Intl Conference on Pervasive intelligence and Computing*, 2017.

[2] K. Tonchev, N. Neshov, A. Manolova and V. Poulkov, “Expression Recognition Using Sparse Selection of log-Gabor Facial Features,” in *Fourth International Conference on Mathematics and Computers in sciences and in industry*, 2017.

[3] Z. Li, “A Discriminative Learning Convolutional Neural Network for Facial Expression Recognition,” in *International Conference on Computer and Communication*, 2017.

[4] F. Xu and z. Wang, “A facial expression recognition method based on cubic spline interpolation and HOG features,” in *IEEE International Conference on Robotics and Biomimetics (ROBIO)*, Macau, 2017.

[5] C. Qi, M. Li, Q. Wang, H. Zhang, J. Xing, Z. Gao and H. Zhang, “Facial Expressions Recognition Based on Cognition and Mapped Binary Patterns,” in *IEEE Access*, 2018.

[6] C. Tsangouri, W. Li, Z. Zhu, F. Abtahi and T. Ro, “An interactive facial-expression training platform for individuals with autism spectrum disorder,” in *IEEE MIT Undergraduate Research Technology Conference (URTC)*, Cambridge, 2016.

[7] M. Taheri, “Robust face recognition via non-linear correlation filter bank,” *IET image process,* pp. 408-415, 2017.

[8] Z. Sun, Z.-P. Hu, M. Wang and S.-H. Zhao, “Discriminative feature learning-based pixel difference representation for facial expression recognition,” *IET Computer Vision,* pp. 675-682, 2017.

[9] X. Song and H. Bao, “Facial expression recognition based on video,” in *IEEE Applied Imagery Pattern Recognition Workshop (AIPR),* Washington,DC, 2016.

[10] E. Tran, M. B. Mayhew, H. Kim, P. Karande and A. D. Kaplan, “Facial Expression Recognition Using a Large Out-of-Context Dataset,” in *IEEE Winter Applications of Computer Vision Workshops (WACVW*, Lake Tahoe, 2018.

[11] M. T. B. Iqbal, M. A. A. Wadud, B. Ryu, F. Makhmudkhujaev and O. Chae, “Md Tauhid Bin Iqbal, M. Abdullah-Facial Expression Recognition with Neighborhood-aware Edge Directional Pattern (NEDP), Journal of latex class files,” IEEE, 2015.

[12] A. M. Ashir and B. Akdemir, “Facial expression recognition with an optimized radial basis kernel,” in *6th international symposium on Digital Forensic and Security (ISDFS),* Antalya, 2018.

[13] Y. Yaddaden, M. Adda, A. Bouzouane, S. Gaboury and B. Bouchard, “Facial Expression Recognition from Video using Geometric Features,” in *8th International Conference of Pattern Recognition Systems (ICPRS 2017),* Madrid, 2017.

[14] S. Xie, H. Hu and Z. Yin, “Facial expression recognition using intra-class variation reduced features and manifold regularisation dictionary pair learning,” *IET Computer Vision,* pp. 458-465, 2018.

[15] M. I. Revina and W. S. Emmanuel, “Recognition of facial expressions using Gaussian based edge direction and texture descriptor,” in *International Conference on Inventive Computing and Informatics (ICICI),* Coimbatore, 2017.

[16] S. Kim, G. H. An and S.-J. Kang, “Facial expression recognition system using machine learning,” in *International SoC Design Conference (ISOCC),* Seoul, 2017.

[17] V. P. V and A. George, “Automatic recognition of facial expression using features of salient patches with SVM and ANN classifier,” in *International Conference on Trends in Electronics and Informatics (ICEI),* Tirunelveli, 2017.

[18] S. Xie and H. Hu, “Facial Expression Recognition Using Hierarchical Features with Deep Comprehensive Multi-Patches Aggregation Convolutional Neural Networks,” in *IEEE Transactions on Multimedia,* 2018.

[19] C. CORTES and V. Vapnik, “support vector networks,” in *machine learning,20,* netherlands.

[20] S. Sharma, K. Kumar and N. Singh, “"D-FES: Deep facial expression recognition system," 2017 Conference on Information and Communication Technology (CICT),” in *Information and Communication Technology (CICT),* Gwalior, 2017.

[21] H. Sadeghi and A.-A. Raie, “Approximated Chi-square distance for histogram matching in facial image analysis: Face and expression recognition,” in *10th Iranian Conference on Machine Vision and Image Processing (MVIP),* Isfahan, 2017.

[22] A. N. Ekweariri and K. Yurtkan, “Facial expression recognition using enhanced local binary patterns,” in *9th International Conference on Computational Intelligence and Communication Networks (CICN),* Girne, 2017.

[23] A. Fathallah, L. Abdi and A. Douik, "Facial Expression Recognition via Deep Learning," in *14th International Conference on Computer Systems and Applications (AICCSA)*, 2017, 2017.

[24] Y. Lu, S. Wang, Y. Zhao and J. Wei, "A novel approach of facial expression recognition based on shearlet transform," in *Global Conference on Signal and Information Processing (GlobalSIP)*, Montreal, 2017.

[25] C. Lin, F. Long and Y. Zhan, "Facial expression recognition by learning spatiotemporal features with multi-layer independent subspace analysis," in *10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, Shanghai, 2017.

[26] A. Sajjanhar, Z. Wu, J. Chen, Q. Wen and R. Xiamixiding, "Experimental evaluation of facial expression recognition," in *10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, Shanghai, 2017.

[27] T. Kanade, J. Cohn and Y. Tian, "ªComprehensive Database for facial expression analysis," in *international conference face and gesture recognition*.

[28] M. J. Lyons, M. Kamachi and J. Gyoba, "Japanese Female Facial Expressions," in *Database of digital images*.

[29] P. Bellhumer, J. Hespanha and D. Kriegm, "Eigen faces vs fischer faces: recognition using class specific linear projection," in *IEEE transaction on pattern analysis and machine intelligence*.

[30] H. Kobayashi and F. Hara , "Recognition of Six Basic Facial Expressions and Their Strength by Neural Network," in *Proceedings of International Workshop Robot and Human Communication, vol. 3*.

[31] M. Lyons , J. Budynek and S. Akamatsu, “Automatic Classification of Single Facial Images,” in *IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 21*.

[32] Z. Zhang, S. Akamatsu and S. Akamatsu, “Comparision between geoetry based and gabor wavelet based facial expression recognition using multi layer perceptron,” in *international conference on automatic face and gesture recognition*.

[33] M. Pantic and L. Rothkrantz , “Expert System for Automatic Analysis of Facial Expression,” in *Journal of Image and Vision Computing, vol. 18,*.

[34] N. Sebe , M. Lew , Y. Sun, L. Cohen, T. Gevers and T. Huang , “Authentic facial expression analysis,” in *image and vision computing,vol.25*.

[35] G. Edwards, T. Cootes and C. Taylor, “Face Recognition Using Active Appearance Models,” in *Proceedings of European Conference on Computer Vision, vol. 2*.

[36] M. Pantic and L. M. Rothkrantz, “Automatic analysis of facial expressions: the state of the art,” in *IEEE Transactions on Pattern Analysis and Machine Intelligence,vol.22*.

[37] C. Huang and Y. Huang , “Facial Expression Recognition Using Model-Based Feature Extraction and Action Parameters Classification.,” in *Journal of Visual Communication and Image Representation, vol. 8*.