FACIAL IMAGE BASED EXPRESSION CLASSIFICATION SYSTEM USING
COMMITTEE NEURAL NETWORKS

A Thesis
Presented to
The Graduate Faculty of The University of Akron

In Partial Fulfillment
of the Requirements for the Degree
Master of Science

Gayatri Paknikar
August, 2008
FACIAL IMAGE BASED EXPRESSION CLASSIFICATION SYSTEM USING COMMITTEE NEURAL NETWORKS

Gayatri Paknikar

Thesis

Approved:                     Accepted:

Advisor                        Department Chair
Dr. Narender P. Reddy          Dr. Daniel B. Sheffer

Committee Member              Dean of the College
Dr. Dale H. Mugler             Dr. George K. Haritos

Committee Member              Dean of the Graduate School
Dr. Yingcai Xiao               Dr. George R. Newkome

Date
ABSTRACT

Human communication has two main aspects: verbal (auditory) and non-verbal (visual). Facial expressions, body movements and physiological reactions are the basic units of the non-verbal communication. Facial expressions and related changes in facial patterns give us information about the emotional state of the person. Psychopathology, stress detection, human-computer interface, and terror deterrence etc., are some of the applications of facial expression detection.

The goal of this study was to classify different facial expressions of individuals from static facial images from a large database with an improved accuracy over previously presented systems. Two classification approaches were used. First, a parameter based classification system was developed which classified the expressions based on the actual parameter values directly. Evaluation of the parameter based system revealed that it could accurately classify only one expression. In the second approach, the committee neural network system was used to classify seven basic emotion types from facial images. Two types of committees, viz. primary committee and secondary committee were trained and evaluated. Committee performance was better than performance of individual networks. The integrated committee system, which incorporated both primary and secondary committees, accurately classified the expressions in 94.73% of the cases.
ACKNOWLEDGEMENTS

I would like to express my sincere thanks to Dr. Reddy for providing me the opportunity to work on this extremely interesting and important topic. His guidance and support have been a constant source of encouragement throughout the work of this thesis. It has been a great honor to have worked under his supervision. His valuable suggestions and feedback at every critical phase throughout the work were of utmost importance for timely completion of the thesis. His tremendous knowledge about the subject has gone a long way in ensuring the successful completion of this thesis.

I would like to express my gratitude towards Dr. Mugler and Dr. Xiao for supporting this work by agreeing to be part of the committee and providing me with critical inputs throughout the work of this thesis.

A special thanks to all the staff of the Department of Biomedical Engineering without whose help this work would not have been possible.

Finally, I would like to thank my family and friends for their continued support which has helped me stay strong and focused on the thesis work.
TABLE OF CONTENTS

LIST OF TABLES............................................................................................................. ix

LIST OF FIGURES......................................................................................................... x

CHAPTERS

I - INTRODUCTION........................................................................................................ 1

1.1 Background ....................................................................................................... 1

1.2 Concept ......................................................................................................... 2

1.3 Test of hypothesis ....................................................................................... 5

1.4 Objectives of the study ............................................................................. 5

II - LITERATURE REVIEW........................................................................................... 7

2.1 Facial expressions ....................................................................................... 8

2.3 Variability of facial expressions ............................................................. 9

2.4 Facial expression analysis ........................................................................ 11

2.4.1 Facial image acquisition / detection ................................................ 11

2.4.2 Facial Feature Data Extraction .......................................................... 11

2.4.3 Facial Expression Classification ......................................................... 12

2.5 Overview of the existing research .......................................................... 13

2.5.1 Facial Action Coding System (FACS) ............................................. 13

2.5.2 EMG based analysis ................................................................. 14
2.5.3 Thermal methods ........................................................................................................... 14
2.5.4 Neural network based analysis .................................................................................... 15
2.5.5 Facial Expression recognition using Committee Neural Network .............................. 15
2.6 Neutral networks ............................................................................................................. 16
   2.6.1 Analogy with biological nervous system ................................................................. 16
   2.6.2 Neural Network Architectures ............................................................................... 18
   2.6.3 Transfer Functions ................................................................................................. 20
   2.6.4 Training functions ................................................................................................. 22
   2.6.5 Learning Algorithms ............................................................................................. 24
   2.6.6 Learning functions ................................................................................................. 26
2.7 Applications of neural networks ..................................................................................... 27
2.8 Summary ....................................................................................................................... 29

III - METHODOLOGY .......................................................................................................... 31
   3.1 Introduction .................................................................................................................. 31
   3.2 Facial image database ................................................................................................. 34
   3.3 Image processing and feature extraction ...................................................................... 35
      3.3.1 Parameters used in the previous system .............................................................. 35
      3.3.2 Additional parameters in new system ................................................................. 37
      3.3.3 Parameter extraction method ............................................................................. 39
   3.4 Normalization ............................................................................................................. 40
   3.5 Statistical analysis of parameters and parameter based classification ...................... 41
   3.6 Neural network classification ..................................................................................... 41
   3.7 Primary neural network classification ........................................................................ 42
3.7.1 Training of primary neural networks ............................................................... 42
3.7.2 Initial evaluation of primary networks and recruitment of the primary committee ............................................................... 45
3.8 Secondary neural networks ..................................................................................... 46
3.8.1 Training of secondary neural networks ............................................................ 47
3.8.2 Initial evaluation and the secondary committee recruitment ......................... 47
3.9 Integrated committee neural network system ......................................................... 47
3.10 Final testing ........................................................................................................... 48
3.11 Statistical analysis ................................................................................................. 50

IV - RESULTS .................................................................................................................. 51
4.1 Parameter extraction ............................................................................................... 52
4.2 Statistical analysis of individual parameters ........................................................... 64
4.3 Parameter based classification ................................................................................. 69
4.4 Results from neural networks ................................................................................. 72
4.5 Summary of evaluation of the integrated committee system .................................. 78

V - DISCUSSION ............................................................................................................. 82
5.1 Analysis of results ................................................................................................... 83
5.1.1 Parameter based system: .................................................................................. 83
5.1.2 Neural network based system: ......................................................................... 84
5.2 Comparison with the previous systems ................................................................. 86
5.3 Limitations of the study ......................................................................................... 87
5.4 Significance of the study ......................................................................................... 88
5.5 Clinical significance of the study ............................................................................ 88
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1: Neural Network output configurations and interpretations</td>
<td>45</td>
</tr>
<tr>
<td>4.1: SNK grouping for upper lip entropy parameter</td>
<td>64</td>
</tr>
<tr>
<td>4.2: SNK grouping for lower lip entropy parameter</td>
<td>64</td>
</tr>
<tr>
<td>4.3: SNK grouping for eyebrow entropy parameter</td>
<td>65</td>
</tr>
<tr>
<td>4.4: SNK grouping for upper lip fractal dimension parameter</td>
<td>65</td>
</tr>
<tr>
<td>4.5: SNK grouping for lower lip fractal dimension parameter</td>
<td>66</td>
</tr>
<tr>
<td>4.6: SNK grouping for eyebrow fractal dimension parameter</td>
<td>66</td>
</tr>
<tr>
<td>4.7: SNK grouping for center of nose to left corner of lip distance parameter</td>
<td>67</td>
</tr>
<tr>
<td>4.8: SNK grouping for center of nose to right corner of lip distance parameter</td>
<td>67</td>
</tr>
<tr>
<td>4.9: SNK grouping for center of nose to center of chin distance parameter</td>
<td>68</td>
</tr>
<tr>
<td>4.10: SNK grouping for center of nose to left corner of lip distance parameter</td>
<td>68</td>
</tr>
<tr>
<td>4.11: SNK grouping for center of nose to right corner of lip distance parameter</td>
<td>69</td>
</tr>
<tr>
<td>4.12: Example of parameter based classification system</td>
<td>72</td>
</tr>
<tr>
<td>4.13: Sample results: individual networks vs committee for the primary network</td>
<td>72</td>
</tr>
<tr>
<td>4.14: Sample results: individual networks vs committee secondary network</td>
<td>73</td>
</tr>
<tr>
<td>4.15: Sample results: individual networks vs integrated committee</td>
<td>74</td>
</tr>
<tr>
<td>4.16: Correct classifications by individual networks vs committee system</td>
<td>75</td>
</tr>
<tr>
<td>4.17: Confusion matrix</td>
<td>81</td>
</tr>
<tr>
<td>Figure</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
</tr>
<tr>
<td>1.1: A typical facial expression analysis system</td>
<td>2</td>
</tr>
<tr>
<td>2.1: Synapse in the nervous system</td>
<td>17</td>
</tr>
<tr>
<td>2.2: A typical model of artificial neuron</td>
<td>18</td>
</tr>
<tr>
<td>2.3: Typical neural network architecture</td>
<td>19</td>
</tr>
<tr>
<td>2.4: Hard limit transfer function</td>
<td>21</td>
</tr>
<tr>
<td>2.5: Piecewise linear transfer function</td>
<td>21</td>
</tr>
<tr>
<td>2.6: Output representation: Logsig transfer function and Tansig transfer function</td>
<td>22</td>
</tr>
<tr>
<td>3.1: Overall block diagram of the methodology</td>
<td>33</td>
</tr>
<tr>
<td>3.2: Examples of facial images</td>
<td>34</td>
</tr>
<tr>
<td>3.3: Real-valued measures from a sample neutral expression image</td>
<td>35</td>
</tr>
<tr>
<td>3.4: Binary measures from sample expression images</td>
<td>37</td>
</tr>
<tr>
<td>3.5: Real-valued measures from a sample neutral expression image</td>
<td>38</td>
</tr>
<tr>
<td>3.6: Binary measures from sample expression images</td>
<td>39</td>
</tr>
<tr>
<td>3.7: Seven-network committee neural network architecture</td>
<td>46</td>
</tr>
<tr>
<td>3.8: Flow chart of overall classification system</td>
<td>49</td>
</tr>
<tr>
<td>4.1: Average percentage deviation from neutral: eyebrow entropy</td>
<td>53</td>
</tr>
<tr>
<td>4.2: Average percentage deviation from neutral: upper lip entropy</td>
<td>54</td>
</tr>
<tr>
<td>4.3: Average percentage deviation from neutral: lower lip entropy</td>
<td>55</td>
</tr>
</tbody>
</table>
4.4: Average percentage deviation from neutral: eyebrow fractal dimension .............. 56
4.5: Average percentage deviation from neutral: upper lip fractal dimension .......... 57
4.6: Average percentage deviation from neutral: lower lip fractal dimension .......... 58
4.7: Average percentage deviation from neutral: center of nose to left corner of the lip distance ........................................................................................................... 59
4.8: Average percentage deviation from neutral: center of nose to right corner of the lip distance ........................................................................................................... 60
4.9: Average percentage deviation from neutral: center of nose to center of chin distance ................................................................................................................. 61
4.10: Average percentage deviation from neutral: inner corner of left eye to left corner of the lip distance ......................................................................................... 62
4.11: Average percentage deviation from neutral: inner corner of right eye to right corner of the lip distance ......................................................................................... 63
4.12: Comparative plot of individual neural network output vs primary and integrated committee neural network outputs ................................................................. 76
4.13: Plot of correct expression classifications for primary and secondary committee classification systems compared to total expressions ................................................. 77
4.14: Plot of percentage correct classifications v/s expressions .................................. 78
4.15: Plot of output summary ......................................................................................... 79
1.1 BACKGROUND

A person’s face is considered as the mirror of the mind. Facial expressions and the changes in facial expressions provide important information about affective state of the person, his temperament and personality, psychopathological diagnostic information, information related to stress levels, truthfulness etc. [1]. With growing terrorist activities all over the world, detection of potential troublemakers continues to be a major problem. Body language and facial expressions are the best ways to know the personality of a person and the response of a person in various situations. The facial expressions tell us about concealed emotions which can be used to verify if the information provided verbally is true. These expressions representing the emotional state of a person can serve an important role in the field of terrorism control and forensics.

Facial expression analysis can also be used effectively in psychopathology. It may give us information related to the diagnostic information relevant to depression, mania, schizophrenia and other disorders. The information relevant to the patient’s response to the treatment could also be monitored with the facial expression analysis [2, 3]. Thus, expression analysis can be effectively used for behavioral studies and medical rehabilitation. Stress detection through facial expression analysis can be useful in cases
like monitoring stress levels in astronauts since other methods may not work in that environment [4]. Thus, there is an increasing need to understand the human facial expressions in a better way and to develop a system to accurately classify the emotions of subjects based on their facial images.

1.2 CONCEPT

Facial expression analysis deals with analysis of different facial motion changes by extraction of facial parameters. A typical system extracts number of facial parameters from an image, and classifies the image into the set of defined expressions. There are six universally recognized expressions viz. Angry, Disgust, Fear, Happy, Sad and Surprised. Figure 1.1 shows components of a typical expression classification system.

![Figure 1.1: A typical facial expression analysis system](image)

Over the last few years, active research is being done to correlate movements of the face with emotional states. Darwin [5] first published “The Expression of the emotions in Man and Animals” in which he stated the three basic principles related to
expressions and gestures in humans, viz., “(a) Principle of Serviceable associated habits, 
(b) Principle of Antithesis, and (c) Principle of actions due to constitution of nervous 
system, independently from the will, and independently to a certain extent habit” [5].

Ekman and Friesen [6] developed the Facial Action Coding System (FACS) to 
measure the facial behavior. In FACS, they used Action Units (AUs) based on the 
muscular activity that produces momentary changes in the facial expression. The system 
further classified an expression by correctly identifying the action unit or combination of 
action units related to a particular expression. A FACS database was created to determine 
the category or categories in which to fit each facial behavior. This database is available 
in the form of a FACS manual. Based on the action units, the researcher has to interpret 
the actual emotion.

Padgent [7], Hara and Kobayashi [8, 9], Zhang [10] and Zhao [11] used neural 
network approach for expression classification. They classified images into six or seven 
emotional categories. Padgett et al., [7] trained neural networks from the data of 11 
subjects and tested with the data from one subject. The training and testing dataset was 
interchanged and new networks were trained and tested. A classification accuracy of 86% 
was achieved in this study. Hara and Kobayashi [8, 9] also used neural networks 
approach. The training dataset consisted of from data of 15 subjects (90 images) and 
these networks were tested using data from another 15 subjects. The classification 
accuracy achieved was 85 %. Zhang et al., [12] used the JAFFE data base which consists 
of 10 Japanese female subjects. Although an accuracy of 90.1% was achieved; same data 
was used for training and testing. A 100 % recognition rate was achieved by Zhao et al., 
[11] who used the Ekman and Friesen database [13], but they used the same data for
training and testing. Khan et al [14] used thermal methods to quantify the facial expressions. He could achieve an accuracy of 56 %

As mentioned above, most of the facial expression analysis systems have been based on FACS and identifying AUs coded from static facial images, and further classifying the expressions based on combinations of these AUs identified. These systems also required sequential video images (from a video) for expression classification. Some of the systems used neural network approach, but these systems either used a very small dataset or the training and testing datasets were not well differentiated. Thus, a method for recognizing expressions needs to be developed for fast, easy, and accurate classification of facial expressions and tested on a larger dataset of facial expressions like the Cohn-Kanade database [15]. A classification system consisting of a committee of neural networks can give better results than a classification system consisting of a single network. Reddy and Buch [16], Das et al., [17] and Reddy et al. [18] observed such an enhanced classification performance in their studies.

Kulkarni et al [19] developed a system to classify expressions from static facial images using committee neural networks. They used eight real valued and seven binary parameters extracted from the static images to perform the study. This system could correctly classify the moods in only 90.426 % cases. However, parameters related to the complexity of the image like fractal dimension [20, 21] and the entropy [22, 23] were not included in the system. Thus, the question remains whether addition of these and other parameters will improve the accuracy of the classification system. The purpose of the current study was to address this question.
1.3 TEST OF HYPOTHESIS

**Null Hypothesis:**

The probability of the committee neural network system making a correct decision \((p1)\) is the same as the probability of the committee neural network system making an incorrect decision \((p2)\). \((p1 = p2)\)

*That is,*

Committee Neural Networks cannot be used to effectively classify the different facial expression based on input facial parameters.

**Alternate Hypothesis:**

The probability of the committee neural network system making a correct decision \((p1)\) is greater than the probability of the committee neural network system making an incorrect decision \((p2)\). \((p1 > p2)\)

*That is,*

Committee Neural Networks can be used to effectively classify the different facial expression based on input facial parameters.

1.4 OBJECTIVES OF THE STUDY

Specific objectives of the study were:-

- To identify and include more number of parameters in addition to the parameters used in the previous study so as to improve the accuracy of classification.
• To train several neural networks and to compare them to find the optimum input parameters to give better accuracy.

• To develop a classification system based on committee neural networks which would classify the facial expressions based on the derived parameters.
CHAPTER II
LITERATURE REVIEW

Darwin [5, 24], based on his observations, demonstrated the universality of facial expression by defining the three principles. He proposed that, these principles account for most of the expressions and gestures used by man.

1. *The principle of serviceable associated habits*: It emphasizes the importance of habits in formulating certain expressions in man.

2. *The principle of antithesis*: This principle gives the explanation of certain expressions which are not formulated by habits but are shown when an exact opposite state of mind to the habits is induced.

3. *The principle of actions due to the constitution of the Nervous System, independently from the will, and independently to a certain extent of Habit*: This principle gives an account of those expressions that are a direct result of the action of nervous system and are independent of will or habits.

The relationship between facial expression and emotion is based on muscular movements and not on measurements of specific muscular actions. In addition, certain prototype facial expressions have been attributed to specific emotion categories, but the entire set of emotions, felt and expressed by humans is not well known. The validity of an
expression being genuine or simulated also decides whether the emotion is being actually felt by the expresser or not.

2.1 FACIAL EXPRESSIONS

The face is the site for the major sensory inputs and the major communicative outputs [25]. There are four general classes in which the facial signals can be defined [25]:

- **Static facial signals:** Attributed to relatively permanent features of the face, such as the bony structure and soft tissues masses, which contribute to an individual's appearance.

- **Slow facial signals:** Attributed to changes in the appearance of the face, such as the development of permanent wrinkles and changes in skin texture, which occur gradually over time.

- **Artificial signals:** Attributed to the external factors such as eyeglasses and cosmetics.

- **Rapid facial signals:** Attributed to temporary changes in neuromuscular activity that may lead to visually detectable changes in facial appearance.

The facial expressions for emotion mainly result from the rapid facial signals. These temporary movements of the facial muscles pull the skin, temporarily changing the shape of the eyes, eyebrows, and lips, and the appearance of folds, furrows and bulges in different patches of skin. These changes last for just a few seconds.
2.2 ANATOMY OF FACIAL MOVEMENT

The rapid facial signals result into facial expression changes caused by movements of facial skin and connective tissue due to contraction of single or combination of 44 bilaterally symmetrical facial striated muscles [26]. Out of these 44 facial muscles, four are innervated by the 5th cranial nerve (trigeminal), and are involved in movement of skeletal muscles like jaw muscles. The second group of 40 muscles is innervated by 7th cranial nerve (facial), and they are not directly involved in the movement of any skeletal muscles, but just arrange the facial muscles in a meaningful configuration. When the striated muscles are activated by some neural activity, it results in release of acetylcholine at the motor end plates [26]. This leads to muscle action potential which travels along the muscle fiber and causes the muscle to contract. The activating neurotransmitter acetylcholine is nullified by the enzyme acetylcholinesterase resulting in a continued propagation of muscle potential and fiber contraction. The low amplitude neural activity signals inhibit small motor-neurons, which are further responsible for innervating few and small facial muscle fibers. Thus, dynamic changes take place in these facial muscles. Depending on the structure, arrangement of the facial muscles, facial skin elasticity, size of adipose tissue layer, the changes in facial movements take place.

2.3 VARIABILITY OF FACIAL EXPRESSIONS

Universality of facial expressions is the most debated topic in the facial expression analysis. Darwin [5] in his studies demonstrated that the basic facial expressions are universal whereas another view proposes that the facial expressions are
not universal but are different in individuals based on various factors. The variability of facial expressions could be attributed to the following factors:

- **Anatomical variation**
  
  Different facial muscles produce different types of movements, and they are most likely heterogeneous in their structure and innervations [27]. Muscle fiber types, shapes, and sizes may vary in different individuals.

- **Variation due to difference in neural control**
  
  Neurobiological, facial expressions are dually controlled by extra pyramidal and pyramidal tracts, providing for automatic and voluntary control of facial expression [27].

- **Sexual variation**
  
  The expressions may vary between men and women.

- **Cultural variation**
  
  This is the most debated type of variation based on the subject’s ethnic background.

Thus, an ideal expression classification system should take into consideration these probable causes of variation in the database. It should be universal in order to classify all the expressions correctly irrespective of all these variations.
2.4 FACIAL EXPRESSION ANALYSIS

A typical facial expression analysis system consists of three subsystems.

(1) Facial image acquisition / detection

(2) Facial feature data extraction

(3) Facial expression recognition / classification

2.4.1 Facial image acquisition / detection

The image resolution, camera type, digitizer used, size of image, image illumination etc. play an important role in the facial expression analysis. Usually, the facial image in the frontal or near frontal view is used to recognize facial expressions. Pantic and Rothkrantz [28] first used two cameras, with one capturing the frontal portion of the face and one on the right hand side capturing the profile image of the face. Image resolution of 69 x 93 is considered as optimum for any automated face processing related techniques. Corners of the mouth and eye become hard to detect at a resolution of 48 x 64. Facial expression cannot be recognized at an image resolution of 24 x 32 [29].

2.4.2 Facial Feature Data Extraction

The facial data extraction from static images involves mainly two types of features: geometric features and appearance features. Geometric feature measures the variations in shape, location, distance of facial components like mouth, eyes, eyebrows, nose, etc. in different expressions. The appearance features present the appearance (skin texture) variations of the face, such as wrinkles, furrows, etc. The appearance feature can be extracted on either the whole face or specific regions in a facial image. To recognize facial expressions, an automatic facial expression analysis system can use geometrical
features only [28, 30], appearance features only [31, 32, 33], or hybrid features (both geometric and appearance based) [34, 35, 36]. Cohn, Kanade and Tian [35, 36] used normalized feature measurements by using a reference neutral face image to remove the variations in face style, motion, and other factors.

2.4.3 Facial Expression Classification

The expression classification is the most critical part of the system. The classification system should be robust enough to handle the variability of the facial expressions. It should be capable of correct classification of expression irrespective of the subject’s age, sex and ethnicity. Ekman [7, 37] defined the six basic emotion classification categories as happiness, sadness, surprise, anger, fear and disgust. Most classification systems classify the emotions into these six universal categories. In reality the expression could be a combination of any of these six expression types. Thus, the classification system should be able to identify and point out the blended expressions appropriately. Most of the classification systems follow a template-based approach, neural network based approach or rule based approach for the classification. Template based classification systems compare the template of each image with a fixed prototype template prepared for each expression. Neural network based classification systems identify the expression based on learning data presented to it in the training phase. Quantified classification and blended classification can be obtained in neural network based systems. Rule based classification systems classify the expression based on its previously encoded facial action.
2.5 OVERVIEW OF THE EXISTING RESEARCH

Since the work of Darwin, over the last few years there has been active research going on to correlate the facial movements with the expression of emotion. Following are a few significant contributions.

2.5.1 Facial Action Coding System (FACS)

FACS is the most widely used system for the facial movement measurement and describing facial behavior. Ekman et al [7] developed the original FACS by determining how the contraction of each facial muscle changes the appearance of the face. They examined videotapes of facial behavior to identify the specific changes that occurred with muscular contractions and to differentiate one from another. A FACS database was created to determine the category or categories in which to fit each facial behavior. This database is available in the form of the FACS manual. The second version of FACS was released in 2002.

FACS determined how the contraction of each facial muscle (singly and in combination with other muscles) changed the appearance of the face. Forty six Action Units were defined that accounted for changes in facial expression, and twelve AUs which described changes in gaze direction and head orientation. The facial expressions were categorized into six basic emotions defined by Ekman [7].

Since FACS deals with movement, not with other visible facial signs, it limits a full understanding of the psychology of facial behavior. Secondly, the scoring system used in case of FACS is descriptive, and does not directly infer the emotion involved.
Thus, the person performing the classification has to be trained to interpret the expression from the action units obtained.

2.5.2 EMG based analysis

Facial EMG measures the electrical activity of the facial muscles. Dimberg et al. [38] used EMG based measurement for facial expressions. Facial expression analysis using EMG based techniques requires invasive insertion of electrodes into the facial muscle fiber for accurate results. Surface EMG based methods are noninvasive and relatively inexpensive, but have inaccurate results due to artifacts and interference of other muscles. Surface EMG measures muscular activity in general area and does not concentrate on specific muscles. EMG units for measuring muscle activity are smaller and enable precise measurement of facial actions. Intensities of facial action can also be measured using EMG based methods. In addition, invisible muscle activities can also be measured and quantified using surface EMG. The major disadvantage of using surface EMG based methods is that it may alter the normal behavior of the subjects due to attachment of electrodes and confuse the subject.

2.5.3 Thermal methods

Khan et al [14] presented a system for automated facial expression classification using infrared measurement of facial skin temperature variations. They captured Infrared thermal images with participants’ normal expression and intentional expressions of happiness, sadness, disgust, and fear. Based on these thermal images, facial points that undergo significant thermal changes with a change in expression termed as Facial
Thermal Feature Points (FTFPs) were identified. A principle component analysis was carried out for the Thermal Intensity Values (TIVs) recorded at the FTFPs. Linear Discriminant Analysis (LDA) was used for classification. The classification gave 56.0% success rate.

2.5.4 Neural network based analysis

Neural networks are one of the widely used methods of classification. Hara and Kobayashi [8, 9], Padgent [7], Zhang [10] and Zhao [11] all used a neural network approach to classify images into six or seven emotional categories. Padgett et al., [7] trained neural networks from the data of 11 subjects and tested with the data from one subject. The training and testing dataset was interchanged and new networks were trained and tested. The average recognition rate achieved was 86 %. Hara and Kobayashi [8, 9] trained neural networks from data of 15 subjects (90 images) and tested them by data from another 15 subjects. The average recognition rate achieved was 85 %. Zhang et al., [12] used the JAFFE data base consisting of 10 Japanese female subjects. Same data was used for training and testing of the system. They achieved an accuracy of 90.1%. Zhao et al., [11] used the Ekman and Friesen database [13] to achieve a 100 % recognition rate, but they used the same data for training and testing.

2.5.5 Facial Expression recognition using Committee Neural Network

Kulkarni et al [19] proposed a system to classify moods from static facial images using committee neural networks. They used eight real valued and seven binary parameters extracted from the static images to perform the study. The values obtained from the measure of these parameters corresponding to each subject input were fed to the
neural networks. Number of such neural networks were trained. Based on the classification accuracies of individual networks, five top performing networks were selected and recruited in the committee. The final output was decided based on the majority opinion. The proposed system correctly classified the moods in only 90.426\% cases.

2.6 NEURAL NETWORKS

An artificial neural network is a non linear and adaptive mathematical module inspired by the working of a human brain. It consists of simple neuron elements operating in parallel and communicating with each other through weighted interconnections.

2.6.1 Analogy with biological nervous system

The human brain consists of thousands of neurons, through which the communication in the nervous system takes place. A typical nervous cell consists of the cell body, axon and the dendrites. The Axon further divides into number of branches at the synapse. Figure 2.1 shows a typical synapse. At synapse, the communication between two nervous cells takes place through electro-chemical activity. The neuron sends out spikes of electrical activity called “Action potential" through the axon. At the synapse, the signal from the axon is converted into electrical effects that inhibit or excite activity in the connected neurons. When a neuron receives excitatory input that is sufficiently large compared with its inhibitory input, it sends a spike of electrical activity down its axon. Learning occurs by changing the effectiveness of the synapses so that the influence of one neuron on another changes. [26]
Keeping the biological analogy in mind, in case of artificial neural networks, the synapses of the neuron were modeled as weights. The strength of the connection between an input and a neuron is defined as the value of the weight. Negative weight values correspond to inhibitory connections, while positive values correspond to excitatory connections. The adder sums up all the inputs modified by their respective weights. Finally, a transfer function controls the amplitude of the output of the neuron. An acceptable range of output is usually between 0 and 1, or -1 and 1 depending on the transfer function selected. Figure 2.2 shows a typical model of an artificial neuron.
2.6.2 Neural Network Architectures

The typical components of a neural network are,

- One input layer of neurons
- One or more hidden layers of neurons
- One output layer

The way in which these components of a neural network are arranged defines the architecture of that neural network. Figure 2.3 shows a typical architecture of a neural network. In the network, information is propagated through the network by the inputs and outputs of each unit. Neural networks can be broadly classified into two categories viz. feed forward and feed back networks. The networks where signals are only propagated forward through the layers are known as a feed forward networks. The networks where signals are propagated both forward and backward are known as a feedback networks.
The various subcategories of the neural network architecture are as follows:

1. Single layer feed forward networks (Perceptron)

The Perceptron is the simplest form of a neural network. It is a single layer network consisting of number of inputs, weights corresponding to them, a bias and the output. The weights and biases could be trained to produce a correct output when presented with the input. When we say “single layer” we refer to the output layer. The input layer is not considered as no computation is performed at this layer.
2. Multilayer feed forward networks

A typical multilayered network consists of an input layer, one output layer and one or more hidden layers. The input layer receives the input data from the environment, the hidden layer contains most of the trained intelligence of the network and the output layer generates the final output of the network. Each neuron is unidirectional and connected to the other neuron of the next layer through synapses. A fully connected network is the one in which every neuron in one layer is connected to every neuron in the subsequent layer. The output of each neuron is the weighted function of the inputs. The output is subjected to a nonlinear transfer function, which is usually a threshold function or bias, and the output is generated only when the weighted sum exceeds the bias value.

3. Recurrent networks

A recurrent network consists of at least one feedback loop. The presence of feedback loop has an effect on the learning capability of the network and its performance.

2.6.3 Transfer Functions

The transfer function defines the output of the neurons. The choice of the transfer function strongly influences the complexity and performance of neural networks. There are three basic types of transfer functions viz. Threshold function, pair wise linear function and sigmoid function.
1. Threshold function (Hard limit transfer function)

In this type of transfer function, the output is forced to either zero or one depending on the net input strength.

![Figure 2.4: Hard limit transfer function](image1)

2. Piecewise linear function

In the case of this transfer function, the amplification factor inside the linear region is assumed to be one. This function approximates a linear function.

![Figure 2.5: Piecewise linear transfer function](image2)
3. Sigmoid transfer function

A sigmoid transfer function is a smooth and continuous thresholding function of the type:

\[ S(x) = \frac{1}{1 + e^{-ax}} \]

It takes any input value between plus and minus infinity and forces the output in the range of -1 to 1 in case of tan sigmoid and 0 to 1 in case of log sigmoid function.

![Graphs of Logsig and Tansig transfer functions](image)

Figure 2.6: Output representation of (a) Logsig transfer function (b) Tansig transfer function.

2.6.4 Training functions

During the training sessions, the weights and biases of the network are iteratively adjusted to minimize the network performance error. Some of the training functions used are as follows,
1. **Trainlm (Levenberg-Marquardt back propagation):**

Trainlm is a network training function that updates weight and bias values according to Levenberg-Marquardt optimization. It assumes that the network has the mean square error (mse) performance function. Back propagation is used to calculate the Jacobian \( jX \) of performance perf with respect to the weight and bias variables \( X \) [39]. Each variable is adjusted according to Levenberg-Marquardt,

\[
jj = jX \cdot jX \\
je = jX \cdot E \\
dX = -(jj+I*mu) / je
\]

Where, \( E \) is all errors and \( I \) is the identity matrix.

2. **Traingd (Gradient descent back propagation):**

Traingd is a network training function that updates weight and bias values according to gradient descent. Back propagation is used to calculate derivatives of performance perf with respect to the weight and bias variables \( X \) [39]. Each variable is adjusted according to gradient descent:

\[
dX = lr \cdot d\text{perf}/dX
\]

3. **Traingdm (Gradient descent with momentum back propagation)**

Traingdm is a network training function that updates weight and bias values according to gradient descent with momentum. Back propagation is used to calculate derivatives of performance perf with respect to the weight and bias variables \( X \) [39]. Each variable is adjusted according to gradient descent with momentum,

\[
dX = mc*dX\text{prev} + lr*(1-mc)*d\text{perf}/dX
\]

Where, \( dX\text{prev} \) is the previous change to the weight or bias.
2.6.5 Learning Algorithms

The operation of a neural network involves two stages, training and recall. “Learning is a process by which the free parameters of a Neural Network are adapted through a process of stimulation by the environment in which the network is embedded. The type of learning is determined by the manner in which the parameter changes take place” [40]. The learning process is governed by a learning algorithm. A neural network forms an output based on all its inputs and transfer function. The network undergoes a process of learning where it becomes better aware of the nature of its inputs and outputs. The network uses this process to better predict and classify the data it processes in the future. Different learning algorithms differ from each other in a way in which the adjustment to a synaptic weight of a neuron is formulated [40].

- Unsupervised Learning

The kind of neural network training where no target output data is provided is known as unsupervised learning. The goal of these networks is not to achieve high accuracy but to apply induction principles to organize data. The learning occurs through supplying substantial data so that the network can observe similarities and therefore use clusters to generalize its input data. These networks are also known as self-organizing networks.

--- Hebbian Learning

Hebb [ ] designed a learning algorithm, which was designed to readjust the weights of the connections based on the probability of firing between units. If the probability of a unit exciting another unit is relatively high (or over a
certain set threshold), then the strength of the connection between those units is increased. An increase in its connection denotes the learning process.

- **Supervised Learning**

During the process of supervised training of neural networks, both inputs and expected outputs are fed into the system and the resulting outputs are then compared to the expected outputs to calculate the approximate error margins. Supervised training methods generally give high accuracy. Different models are used to readjust the weights and connections between the units in the network.

— **Error-correction Learning**

A type of supervised training based on system accuracy is known as the error correction learning. It is a closed loop feedback system. This training model uses the error value, which depends on the difference between the output of the network and the target output. The model then propagates the error back to adjust the weight values of the units of the system.

— **Back propagation algorithm**

The back propagation algorithm is a methodology to perform supervised training on a neural network using certain error calculations. It follows the principles of Hebbian learning. The back propagation algorithm incorporates the Hebbian principles of readjusting and the error correction-training model. It is a three-step process for training the network. Firstly, a forward sweep is done to take all the
inputs and process it through all the layers, from input to hidden to output. Next step involves the calculation of the error value, based on the difference of the output value and the target value. If the error is within the acceptable range, then the network is trained. Otherwise, if the error is greater than the acceptable range, the third step includes propagating back a calculated value used to adjust the connection weights in order to produce better and more accurate outputs during the future forward sweeps.

• Competitive Learning

In competitive learning, the output neurons of a neural network compete among themselves to become active or get fired. As opposed to Hebbian learning, in competitive learning, only a single output neuron is active at a time. Due to this feature of competitive learning, it is highly suited to discover statistically salient features that may be used to classify a set of input parameters [40].

2.6.6 Learning functions

These functions are triggered during the weight adaptation phase of the learning process. They modify the weights on the input side of each processing element based on certain optimization algorithms. Some of the learning functions are:

• Learngd (Gradient descent weight and bias learning function): This function calculates the change of weight $\Delta w$ for a given processing element from the input $i$, the error $e$ and the learning rate $lr$ according to the gradient descent function:
\[ \Delta w = lr \cdot gw \]

- **Learngdm (Gradient descent with momentum weight and bias learning function):**

  This function calculates the change of weight \( \Delta w \) for a given processing element from the input \( i \), the error \( p \), the weight \( w \), learning rate \( lr \), and the momentum constant \( mc \).

  Thus:

  \[ \Delta w = mc \cdot \Delta w_{prev} + (1-mc) \cdot lr \cdot gw \]

  \( \Delta w_{prev} \) is the previous weight which is obtained from the previous learning state.

2.7 APPLICATIONS OF NEURAL NETWORKS

1. Medicine

   Neural networks are ideal in recognizing diseases using scans since there is no need to provide a specific algorithm on how to identify the disease. Neural networks learn by example so the details of how to recognize the disease are not needed. What is needed is a set of examples that are representative of all the variations of the disease. The examples need to be selected very carefully if the system is to perform reliably and efficiently. Some of the applications in medical field are as follows,

   - **Diagnosing the Cardiovascular System:** Neural Networks are used to diagnose the human cardiovascular system. Diagnosis can be achieved by training the neural network with patient data for normal and diseased patients. These networks can then be used to classify the real time physiological measurements taken from the patient [41].
• Diagnosis of Dysphagia: Reddy et al.[42] demonstrated the use of committee neural networks and implemented them in the diagnosis of dysphagia.

• Diagnosis of breast implant: In 1997, Salchenberger et al., [43] used back propagation and radial basis functions for the diagnosis of breast implant rupture using ultrasound.

• Speaker Verification: Reddy and Buch in 2003 [16] used committee neural networks for speaker verification using speech signals.

• Instant Physician: In this application, neural network can be trained to store a large number of medical records, each of which includes information on symptoms, diagnosis, and treatment for a particular case. After training, the net can be presented with input consisting of a set of symptoms; it will then find the full stored pattern that represents the "best" diagnosis and treatment.

2. Language processing: The language processing application include text-to-speech conversion, auditory input for machines, automatic language translation, secure voice keyed locks, automatic transcription, aids for the deaf, aids for the physically disabled which respond to voice commands, and natural language processing.

4. Image processing, segmentation and compression: Because neural networks can accept a vast array of input at once, and process it quickly, they are useful in
image compression. Image compression based techniques reduce eight bits of data to three and then reverse that process upon restructuring to eight bits again.

5. Forecasting: Neural networks are used to predict the weather, movement of stocks currencies etc., from previous data [44].

2.8 SUMMARY

The studies presented in this chapter show that there have been various approaches used to classify the facial expressions. Most of these approaches used the Facial Action Coding System (FACS) developed by Ekman and Fraise [6] to quantify the facial movements. The EMG based methods were used to measure the muscle movements. These methods were to be invasive for accurate results and thus, limited to lab or hospital environment. Various classifiers could be used for the classification of the facial feature data like neural networks, hidden-markov models, linear discriminant analysis, self organizing maps etc. Amongst these, neural networks show a significant performance in the expression classification systems. Although some of the neural network based systems showed a good classification performance (86% by Padgett, 90.1% by Zhang et al.), they either used a small database or the datasets used for training and testing were not well differentiated. In some cases the dataset was not diverse in terms of sex and ethnicity. Also, in most of the cases, individual neural networks were used for classification. It was observed by Reddy et al [16, 18] and Das et al [17] that a committee neural network system performs better than the individual networks. Kulkarni et al [19] used the committee neural network approach using static facial parameters.
They did not consider the image complexity measures in the study. Though various systems were developed to classify the facial expressions, all of the existing systems have several drawbacks discussed above. Thus, there was a significant need for a study which can classify a large and diverse database with an improved accuracy.
CHAPTER III
METHODOLOGY

3.1 INTRODUCTION

The objective of the present study was to develop a system using committee neural networks to improve the classification accuracy. The facial images were obtained from the Cohn-Kanade database [15]. The database contained images from 97 subjects expressing six basic emotions along with the neutral image for each expression. Several features were extracted and used to train numerous networks with different training parameters, different initial weights, etc. In addition to the 15 parameters including eight real valued and seven binary parameters, used in the previous system (Kulkarni et. al), 12 new parameters were extracted and fed to the neural networks. In addition to the existing 15 parameters, five more real valued distance measures were included along with three entropy and three fractal dimension parameters and one binary valued parameter. All the parameters except the binary parameter were normalized to a value between ‘0’ and ‘1’. A post priory test with SNK grouping was performed to check if the parameter values show significant difference for different expressions. Based on the results of SNK grouping, it was checked if the parameters could classify the expressions independently. A classification scheme was developed based on the individual parameters. The parameters showing a significant difference for a particular expression over other
parameters were grouped to develop the classification scheme. The next step taken was training and testing of neural networks. Several neural networks were trained while varying the number of input parameters and corresponding target values in each set of networks. The best performing neural networks were selected based on initial evaluation. These networks were then recruited in a committee for primary expression classification. During the initial testing, number of all zero output and ambiguous cases were observed specially for angry, disgust, fear and sad expressions. Secondary neural networks were then trained to evaluate these expressions. The best performing neural networks were then recruited in a separate committee to perform secondary classifications. An integrated committee neural network classification system was designed, that included both primary committee networks and the secondary committee networks for final evaluation. The integrated system was then evaluated with an independent expression dataset not used in training or initial testing. Using the obtained results, a binomial test was performed to validate the hypotheses statistically.
Figure 3.1: The System Architecture
3.2 FACIAL IMAGE DATABASE

The images used for the study were obtained from the Cohn-Kanade database [15]. The database is diverse in terms of age, sex ethnicity etc. It consisted of facial images taken from 97 subjects in the age from 18 to 30 years. The database had 65 percent female subjects. 15 percent of the subjects were African-American and three percent were Asian or Latino. The Panasonic WV3230 camera was used for image capturing. Frontal images were taken using the camera located directly in front of the subject. The subjects performed different facial displays starting and ending with a neutral face. The displays were based on descriptions of prototypic emotions (i.e., happiness, surprise, anger, fear, disgust, and sadness). The image sequences were digitized into 640 by 480 pixel arrays with 8-bit precision for grayscale values. The images were saved in PNG format. Figure 3.2 presents a few examples from the database.

Figure 3.2: Examples of facial images

(Reproduced with permission from Cohn-Kanade database [15]. Further reproduction is prohibited)
3.3 IMAGE PROCESSING AND FEATURE EXTRACTION

Twelve new parameters were extracted from the facial images in addition to the fifteen existing parameters. In the current system, five new real valued distance measures were added to the group of real valued parameters. One binary parameter was extracted. The unique feature of the new system was addition of image complexity measures. Image entropy and fractal dimension parameters for eyebrow and lips were extracted from the facial images. Total three entropy and three fractal dimension parameters were obtained. Thus, the integrated system consisted of total 27 parameters with 15 parameters from the previous study and 12 new parameters.

3.3.1 Parameters used in the previous system

Real valued parameters

Figure 3.3 shows the real valued distance measures.

Figure 3.3: Real-valued measures from a sample neutral expression image.

(Reproduced with permission, from Cohn-Kanade database [15]. Further reproduction is prohibited)
1. Eyebrow Raise distance – The distance between common point of upper and lower eyelid and the lower central tip of the eyebrow.

2. Upper eyelid to eyebrow distance – The distance between the upper eyelid and eyebrow surface.

3. Inter-eyebrow distance – The distance between the lower central tip of both the eyebrows.

4. Upper eyelid – lower eyelid distance – The distance between the upper eyelid and lower eyelid.

5. Upper lip thickness – The measure of the thickness of the upper lip.


7. Mouth width – The distance between the tips of the lip corner.

8. Mouth opening – The distance between the lower surface of upper lip and upper surface of lower lip.

Binary parameters

1. Upper teeth visible – Presence or absence of visibility of upper teeth.

2. Lower teeth - Presence or absence of visibility of upper teeth.

3. Forehead lines - Presence or absence of wrinkles in the upper part of the forehead.

4. Eyebrow Lines – Presence or absence of wrinkles in the region above the eyebrows.

5. Nose Lines – Presence or absence of wrinkles in the region between the eyebrows extending over the nose.

Figure 3.4 shows the various binary parameters
6. Chin Lines – Presence or absence of wrinkles or lines on the chin region just below the lower lip.

7. Nasolabial lines – Presence or absence of thick lines on both sides of the nose extending until the upper lip.

3.3.2 Additional parameters in new system

Real Valued parameters:

1. Center of nose to center of chin distance: The distance between the center of nose to the center of chin

2. Center of nose to left tip of lips: The distance between the center of nose to left tip of the lip
3. Center of nose to right tip of lips: The distance between the center of nose to right tip of the lip

![Image with measurements marked]

Figure 3.5: Real-valued measures from a sample neutral expression image.

(Reproduced with permission, from Cohn-Kanade database [15]. Further reproduction is prohibited)

4. Inner corner of left eye to tip of left lip: The distance between inner corner of left eye to tip of left lip.

5. Inner corner of right eye to tip of right lip: The distance between inner corner of right eye to tip of right lip.

Image complexity parameters:

6. Fractal dimension of eyebrow

7. Fractal dimension of upper lip

8. Fractal dimension of lower lips

9. Entropy of eyebrow
10. Entropy of upper lip

11. Entropy of lower lip

Binary Parameter:

Figure 3.6: Binary measures from sample expression images.

(Relroduced with permission, from Cohn-Kanade database [15]. Further reproduction is prohibited)

1. Wrinkles in lips

3.3.3 Parameter extraction method

Real valued parameters:

The real-valued parameters were the distances measured between the specified facial features. The distances were measured using the distance tool in GIMP software. The distance was measured in term of the number of pixels.

Image complexity parameters:

1. Fractal Dimension: This can be used to measure the complexity of an Image. It describes how the object fills the space. Fractal dimension is generally expressed by a noninteger - that is, by a fraction rather than by a whole number. The fractal dimension is calculated as [20,21]
2. Entropy: This can be used to measure the information content of the image. It is calculated in terms of probabilities [22, 23, and 45].

\[
\text{Entropy} = - \sum (p \log p)
\]

For calculation of fractal dimension and entropy of eyebrows and upper and lower lips, the images of eyebrow and lips were extracted from the face image using segmentation methods. The values were then calculated using Matlab for the extracted images.

For extraction of features from images and measurement software like Photoshop and GIMP were used. Matlab image processing functions were used for image segmentation in some cases and for the analysis of these extracted images.

3.4 NORMALIZATION

All the real valued measures including the image complexity measures were derived from both the neutral image as and for each expression image. Normalization was performed to convert it to the form suitable as input to the neural networks. The normalization was performed in the following manner:

\[
\text{Normalised Value} = \frac{\text{Measured Value} - \text{Neutral Value}}{\text{Neutral Value}}
\]
3.5 STATISTICAL ANALYSIS OF PARAMETERS AND PARAMETER BASED CLASSIFICATION

Statistical analysis of individual parameters was conducted to check if any particular parameter can be used as stand-alone criteria for correctly classifying the moods. The post priory test with SNK grouping was carried out for checking the statistical significance. Based on the SNK grouping, the parameters showing significantly different values for various expressions were selected and tested with the testing data. All the data points were compared with the mean value for the expressions and if the value lied within one standard deviation from the mean, it was classified as that particular expression.

3.6 NEURAL NETWORK CLASSIFICATION

Several neural networks were trained based on the input parameters extracted. Two types of neural networks were trained. Primary networks were trained using the input parameter values for all the seven expressions. The outputs of these networks classified the data into seven different expressions viz. neutral, angry, disgust, fear, happy, sad, and surprised. The secondary networks were trained to classify the data into five expressions viz. neutral, angry, disgust, fear and sad. The classification system consisted of three stages.

1. Training
2. Initial testing
3. Final testing
3.7 PRIMARY NEURAL NETWORK CLASSIFICATION

Primary neural networks were used to classify the expressions into seven different classes viz. neutral, angry, disgust, fear, happy, sad and surprised.

3.7.1 Training of primary neural networks

Several multi layered, fully connected, feed forward neural networks were trained to classify different expressions based on the extracted parameters as inputs to each network. Targets were the corresponding expressions each subject expressed. Data from 25 subjects was used to train the neural networks. In all, 139 datasets were used in the training phase. Different sets of neural networks were trained by varying the inputs in each case. A total of 120 networks were trained. The types of neural networks trained are as listed below.

1. Networks with all 26 parameters (15 from previous study and 11 from new study as listed in section 3.3): 20 such networks were trained. Each network had 26 input nodes, each corresponding to the 26 input parameters. Each of these networks had seven output nodes each corresponding to one of the seven expressions (neutral, angry, disgust, fear, happy, sad and surprised). The number of hidden layer neurons were varied in the range of 10 to 30. The transfer function was also varied (Tansig and Logsig)

2. Networks with only the new parameters derived in the current study: 20 such networks were trained. Each network had 11 input nodes, each corresponding to the 11 input parameters. Each of these networks had seven output nodes each
corresponding to one of the seven expressions (neutral, angry, disgust, fear, happy, sad and surprised). The number of hidden layer neurons were varied in the range of 10 to 30. The transfer function was also varied (Tansig and Logsig).

3. Networks with only the parameters from previous study: 20 such networks were trained. Each network had 15 input nodes, each corresponding to the 15 input parameters. Each of these networks had seven output nodes each corresponding to one of the seven expressions (neutral, angry, disgust, fear, happy, sad and surprised). The number of hidden layer neurons were varied in the range of 10 to 30. The transfer function was also varied (Tansig and Logsig).

4. Networks with only the addition of complexity parameters to the previous data: 20 such networks were trained. Each network had 21 input nodes, each corresponding to the 21 input parameters. Each of these networks had seven output nodes each corresponding to one of the seven expressions (neutral, angry, disgust, fear, happy, sad and surprised). The number of hidden layer neurons were varied in the range of 10 to 30. The transfer function was also varied (Tansig and Logsig).

5. Networks with only the addition of entropy parameters to the previous data: 20 such networks were trained. Each network had 18 input nodes, each corresponding to the 18 input parameters. Each of these networks had seven output nodes each corresponding to one of the seven expressions (neutral, angry, disgust, fear, happy, sad and surprised). The number of hidden layer neurons were varied in the range of 10 to 30. The transfer function was also varied (Tansig and Logsig).

6. Networks with only considering the significantly different parameters: 20 such networks were trained. Each network had 20 input nodes, each corresponding to the
input parameters. Each of these networks had seven output nodes each corresponding to one of the seven expressions (neutral, angry, disgust, fear, happy, sad and surprised). The number of hidden layer neurons were varied in the range of 10 to 30. The transfer function was also varied (Tansig and Logsig).

In all the neural networks, since the normalized input data was in the range of -1 to 1, Tansig function was used for the hidden layer neurons. The output of the neural network was to be in the 0 to 1 range. Hence, the “Logsig” function was used as the transfer function for the output layer neurons. The output of each node was converted to a binary number (either 0 or 1). An output of 0.6 or more was forced to 1 and an output of less than 0.6 was forced to 0. An output of 1 indicated that particular expression was present and an output of 0 indicated that particular expression was absent. Table 3.1 shows the different configurations of the outputs at the seven nodes and its corresponding interpretations. The networks were trained using the Levenberg-Marquardt (trainlm) learning algorithm using MATLAB. The goal was to make the error less than $1 \times 10^{-6}$ and the maximum number of epochs used for training was 100.
Table 3.1: Neural Network output configurations and interpretations

<table>
<thead>
<tr>
<th>Node 1 Neutral</th>
<th>Node 2 Angry</th>
<th>Node 3 Disgust</th>
<th>Node 4 Fear</th>
<th>Node 5 Happy</th>
<th>Node 6 Sad</th>
<th>Node 7 Surprised</th>
<th>Neural Network Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Neutral</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Angry</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Disgust</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Fear</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>Happy</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>Sad</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>Surprised</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>Ambiguous (Angry/Disgust)</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>Ambiguous (Fear/Sad/Surprised)</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>No Classification</td>
</tr>
</tbody>
</table>

3.7.2 Initial evaluation of primary networks and recruitment of the primary committee

All the 120 trained neural networks were subject to initial testing using data from 30 subjects not used in the training. Fifteen networks were selected on the basis of their performance during the initial testing. Two criteria for selection were decided. (a) The
percentage accuracy achieved was to be greater than 75% and (b) Networks which were able to correctly classify the expressions that were misclassified by other networks, were selected. These networks not necessarily had a high overall accuracy. Committees of size 3, 5, 7, 9 and 11 networks were formed with different combinations of the selected fifteen networks. These committees were then evaluated with the fresh data not used for initial testing. The best performing committee was retained.

Figure 3.7: Seven-network committee neural network architecture

3.8 SECONDARY NEURAL NETWORKS

The initial evaluation of the primary committee classification system presented a few all-zero or ambiguous cases. To obtain a better classification performance for the negative expressions viz. angry, disgust, fear and sad, a secondary neural network system was developed.
3.8.1 Training of secondary neural networks

Secondary networks were trained to classify among angry, disgust, fear and sad. Data from the same 25 subjects related to angry, disgust, fear and sad expressions was used for training. An additional binary parameter corresponding to presence or absence of wrinkles on the lips was incorporated in the secondary networks. There were a total of 27 input parameters used for training. A total of forty secondary networks were trained with varying number of hidden layer neurons and transfer functions and initial weights.

3.8.2 Initial evaluation and the secondary committee recruitment

All the 40 secondary trained neural networks were subject to initial testing using data from 30 subjects not used in the training. These networks classified the data into five expression classes viz. neutral, angry, disgust, fear and sad. The best performance networks were selected based on a cutoff accuracy of 80% and recruited into the committee. Committees of size 3, 5 and 7 networks were formed using the best performing secondary networks. The best performing committee was retained.

3.9 INTEGRATED COMMITTEE NEURAL NETWORK SYSTEM

An integrated committee neural network system was developed incorporating the primary and secondary networks in combination. Figure 3.8 shows the flowchart of the integrated committee neural network system classification process. The integrated committee neural network system consisted of a seven network primary committee neural network classification system in combination with a five-member secondary committee neural network classification system. If the primary committee output was observed as
ambiguous, then those input images were fed to the secondary committee for further classification.

3.10 FINAL TESTING

Data from forty subjects (190 images) were used for final evaluation of the integrated system. These forty subjects were independent subjects not used in training or initial evaluations. Input data was first fed to the primary committee neural network classification system. If the primary classification system evaluated a test case as an all-zero or no classification case, then the corresponding test case was fed to the secondary networks. The secondary networks further classified the expression between angry, disgust, fear and sad expressions. Finally, the primary and the secondary committee network outputs were combined to present the final expression classification.
Figure 3.8: Flow chart of overall classification system
3.11 STATISTICAL ANALYSIS

A Binomial test was carried out to compare the output of the integrated committee with the actual expected output. The expected outputs are provided with the database in which the outputs were obtained by translating the FACS codes into corresponding emotion prototypes. Statistical analysis was performed between the two classifications by the Binomial test for $\alpha = 0.01$. A correct classification was termed as a positive result and an incorrect classification was termed as a negative result.
CHAPTER IV

RESULTS

In the current study, the goal was to test whether the addition of new parameters to the existing system (Kulkarni et al) improves the performance accuracy. The first task in the process was to derive the new parameters. Twelve additional parameters were extracted in addition to the existing fifteen parameters from ninety-seven subjects for seven different expressions constituting 467 facial images.

All these parameter values were statistically analyzed using post priory test using SNK grouping to check if they can classify the expressions independently. The results showed that parameters could classify the expressions in groups but cannot classify the expressions in seven different classes. Based on the results of the statistical analysis, a parameter based classification method was developed. This method could classify surprised expression in 95.79% cases but was not effective for other expressions. Thus, neural networks were trained with combinations of these 27 parameters as inputs. Two types of neural networks were trained; (a) Primary networks in which data from 25 subjects consisting of 139 images were used to train 120 neural networks. These networks were trained to classify image into seven expressions viz. angry, disgust, fear, sad, surprised, happy and neutral. (b) Secondary networks in which data from 25 subjects consisting of 71 facial images were used to train 40 networks with five output nodes.
Secondary networks were trained to classify image into five expressions viz. angry, sad, fear and disgust and neutral. All these trained networks were then tested a fresh dataset for initial testing. The primary networks were initially tested with data from 30 subjects constituting of 138 images and secondary networks were initially tested with data from 30 subjects constituting of 74 images. Committees of neural networks of different sizes and combinations were recruited for both primary networks and secondary networks. The best performing committee of primary networks consisting of seven members (Appendix C) and the best performing committee of secondary networks consisting of five members (Appendix C) were recruited to form an integrated committee neural network system. The integrated committee neural network system was further evaluated with data from subjects constituting 190 facial images. This data was obtained from 40 subjects not used in training or initial testing. The integrated system correctly identified 180 out of 190 expressions presented to it, thus giving an overall accuracy of 94.73 %.

4.1 PARAMETER EXTRACTION

Parameter extraction was the first step in the classification process. Twelve new parameters were extracted from the facial images of the dataset. The values of the extracted parameters were then normalized with respect to the neutral value for that particular subject.
Figure 4.1 below shows the average trend line for the parameter eyebrow entropy in different expressions. It can be observed from the trend line that, a negative percent deviation was observed for the disgust expression, a positive percent deviation was observed for angry, fear, happy and sad expressions. No substantial deviation was observed for Surprised.

Figure 4.1: Plot of average percentage deviation from neutral value, for the eyebrow entropy parameter, for different expressions with the corresponding standard deviations.
Figure 4.2 below shows the average trend line for the parameter Upper lip entropy in different expressions. It can be observed from the trend line that, a negative percent deviation was observed for the angry and sad expression, a positive percent deviation was observed for disgust, fear, happy and surprised expressions.

Figure 4.2: Plot of average percentage deviation from neutral value, for the upper lip entropy parameter for different expressions with the corresponding standard deviations.
Figure 4.3 below shows the average trend line for the parameter lower lip entropy in different expressions. It can be observed from the trend line that, a negative percent deviation was observed for the angry expression, a positive percent deviation was observed for disgust, fear, happy and surprised expressions. No substantial deviation was observed for sad expression.

![Figure 4.3: Plot of average percentage deviation from neutral value, for the Lower lip entropy parameter for different expressions with the corresponding standard deviations.](image)

Figure 4.3: Plot of average percentage deviation from neutral value, for the Lower lip entropy parameter for different expressions with the corresponding standard deviations.
Figure 4.4 below shows the average trend line for the parameter eyebrow fractal dimension in different expressions. It can be observed from the trend line that, all the expressions show a positive deviation from the neutral. The Surprised expression shows a substantial deviation from neutral.

![Graph showing average percentage deviation from neutral value for eyebrow fractal dimension](image)

Figure 4.4: Plot of average percentage deviation from neutral value, for the eyebrow fractal dimension parameter, for different expressions with the corresponding standard deviations.
Figure 4.5 below shows the average trend line for the parameter upper lip fractal dimension in different expressions. It can be observed from the trend line that, a negative percent deviation was observed for the fear and happy expression, a positive percent deviation was observed for angry, disgust, sad and surprised expressions.

Figure 4.5: Plot of average percentage deviation from neutral value, for the Upper lip fractal dimension parameter for different expressions with the corresponding standard deviations.
Figure 4.6 below shows the average trend line for the parameter lower lip fractal dimension in different expressions. It can be observed from the trend line that, a negative percent deviation was observed for the happy expression, a positive percent deviation was observed for angry, disgust, fear, sad and surprised expressions.

Figure 4.6: Plot of average percentage deviation from neutral value, for the Lower lip fractal dimension parameter for different expressions with the corresponding standard deviations.
Figure 4.7 below shows the average trend line for the parameter center of nose to left corner of the lip distance in different expressions. It can be observed from the trend line that, a negative percent deviation was observed for the angry and disgust expression, a positive percent deviation was observed for fear, happy and sad and surprised expressions.

![Graph showing percentage deviation from neutral value for different expressions](image-url)

Figure 4.7: Plot of average percentage deviation from neutral value, for the Center of nose to left corner of the lip distance parameter for different expressions with the corresponding standard deviations.
Figure 4.8 below shows the average trend line for the parameter center of nose to right corner of the lip distance in different expressions. It can be observed from the trend line that, a negative percent deviation was observed for the angry and disgust expressions, a positive percent deviation was observed for fear, happy and sad and surprised expressions.

![Figure 4.8](image)

Figure 4.8: Plot of average percentage deviation from neutral value, for the Center of nose to right corner of the lip parameter, for different expressions with the corresponding standard deviations.
Figure 4.9 below shows the average trend line for the parameter center of nose to center of chin distance in different expressions. It can be observed from the trend line that, a negative percent deviation was observed for the angry and sad expression, a positive percent deviation was observed for disgust, fear, happy and surprised expressions.

Figure 4.9: Plot of average percentage deviation from neutral value, for the Center of nose to center of chin parameter for different expressions with the corresponding standard deviations.
Figure 4.10 below shows the average trend line for the parameter inner corner of left eye to left corner of the lip distance in different expressions. It can be observed from the trend line that, a negative percent deviation was observed for the disgust and happy expressions, a positive percent deviation was observed for sad and surprised expressions. No substantial deviation was observed for fear expression.

Figure 4.10: Plot of average percentage deviation from neutral value, for the Inner corner of left eye to left corner of the lip parameter for different expressions with the corresponding standard deviations.
Figure 4.11 below shows the average trend line for the parameter inner corner of right eye to right corner of the lip distance in different expressions. It can be observed from the trend line that, a negative percent deviation was observed for the angry, disgust and happy expressions, a positive percent deviation was observed for sad and surprised expressions. No substantial deviation was observed for fear expression.

Figure 4.11: Plot of average percentage deviation from neutral value, for the inner corner of right eye to right corner of the lip parameter for different expressions with the corresponding standard deviations.
4.2 STATISTICAL ANALYSIS OF INDIVIDUAL PARAMETERS

Upper lip entropy

The table below shows the SNK grouping for the upper lip entropy parameter. It can be observed that fear, surprised, happy and disgust can be grouped in group A. Disgust, neutral and sad can be grouped in group B and angry can be grouped in group C.

Table 4.1: SNK grouping for upper lip entropy parameter

<table>
<thead>
<tr>
<th>Mean</th>
<th>N</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.16812</td>
<td>15</td>
</tr>
<tr>
<td>A</td>
<td>0.16155</td>
<td>21</td>
</tr>
<tr>
<td>A</td>
<td>0.14665</td>
<td>22</td>
</tr>
<tr>
<td>B</td>
<td>0.05528</td>
<td>17</td>
</tr>
<tr>
<td>B</td>
<td>0.00000</td>
<td>24</td>
</tr>
<tr>
<td>B</td>
<td>-0.04461</td>
<td>20</td>
</tr>
<tr>
<td>C</td>
<td>-0.18526</td>
<td>18</td>
</tr>
</tbody>
</table>

Lower lip entropy

The table below shows the SNK grouping for the upper lip entropy parameter. It can be observed that happy and fear can be grouped in group A. Fear and surprise can be grouped in group B and sad, angry, disgust and neutral can be grouped in group C.

Table 4.2: SNK grouping for lower lip entropy parameter

<table>
<thead>
<tr>
<th>Mean</th>
<th>N</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.40201</td>
<td>22</td>
</tr>
<tr>
<td>B</td>
<td>0.32365</td>
<td>15</td>
</tr>
<tr>
<td>B</td>
<td>0.20881</td>
<td>21</td>
</tr>
<tr>
<td>C</td>
<td>0.02080</td>
<td>20</td>
</tr>
<tr>
<td>C</td>
<td>0.00000</td>
<td>24</td>
</tr>
<tr>
<td>C</td>
<td>-0.06629</td>
<td>16</td>
</tr>
<tr>
<td>D</td>
<td>-0.19803</td>
<td>19</td>
</tr>
</tbody>
</table>
Eyebrow entropy

The table below shows the SNK grouping for the eyebrow entropy parameter. It can be observed that none of the values are significantly different than other values.

Table 4.3: SNK grouping for eyebrow entropy parameter

<table>
<thead>
<tr>
<th>Mean</th>
<th>N</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>A 0.01384</td>
<td>22</td>
<td>Happy</td>
</tr>
<tr>
<td>A 0.00000</td>
<td>24</td>
<td>Neutral</td>
</tr>
<tr>
<td>A -0.00814</td>
<td>20</td>
<td>Sad</td>
</tr>
<tr>
<td>A -0.03263</td>
<td>21</td>
<td>Surprise</td>
</tr>
<tr>
<td>A -0.03570</td>
<td>18</td>
<td>Angry</td>
</tr>
<tr>
<td>A -0.03961</td>
<td>15</td>
<td>Fear</td>
</tr>
<tr>
<td>A -0.05175</td>
<td>17</td>
<td>Disgust</td>
</tr>
</tbody>
</table>

Upper lip fractal dimension

The table below shows the SNK grouping for the upper lip fractal dimension parameter. It can be observed that surprised can be grouped in group A. Disgust, neutral, angry and sad can be grouped in group B and sad, neutral, angry, fear and happy can be grouped in group C.

Table 4.4: SNK grouping for upper lip fractal dimension parameter

<table>
<thead>
<tr>
<th>Mean</th>
<th>N</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>A 0.047911</td>
<td>21</td>
<td>Surprise</td>
</tr>
<tr>
<td>B 0.015395</td>
<td>17</td>
<td>Disgust</td>
</tr>
<tr>
<td>C B 0.005107</td>
<td>20</td>
<td>Sad</td>
</tr>
<tr>
<td>C B 0.000000</td>
<td>24</td>
<td>Neutral</td>
</tr>
<tr>
<td>C B -0.000337</td>
<td>18</td>
<td>Angry</td>
</tr>
<tr>
<td>C -0.004701</td>
<td>15</td>
<td>Fear</td>
</tr>
<tr>
<td>C -0.006933</td>
<td>22</td>
<td>Happy</td>
</tr>
</tbody>
</table>
Lower lip fractal Dimension

The table below shows the SNK grouping for the lower lip fractal dimension parameter. It can be observed that surprised can be grouped in group A. All other expressions i.e. fear, sad, neutral, disgust, angry and happy can be grouped in group B

Table 4.5: SNK grouping for lower lip fractal dimension parameter

<table>
<thead>
<tr>
<th>Mean</th>
<th>N</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>A 0.06257</td>
<td>21</td>
<td>Surprised</td>
</tr>
<tr>
<td>B 0.00638</td>
<td>14</td>
<td>Fear</td>
</tr>
<tr>
<td>B 0.00473</td>
<td>20</td>
<td>Sad</td>
</tr>
<tr>
<td>B 0.00000</td>
<td>24</td>
<td>Neutral</td>
</tr>
<tr>
<td>B -0.00514</td>
<td>17</td>
<td>Disgust</td>
</tr>
<tr>
<td>B -0.00758</td>
<td>19</td>
<td>Angry</td>
</tr>
<tr>
<td>B -0.01112</td>
<td>22</td>
<td>Happy</td>
</tr>
</tbody>
</table>

Eyebrow fractal dimension

The table below shows the SNK grouping for the eyebrow fractal dimension parameter. It can be observed that surprised can be grouped in group A. All other expressions i.e. fear, sad, neutral, disgust, angry and happy can be grouped in group B

Table 4.6: SNK grouping for eyebrow fractal dimension parameter

<table>
<thead>
<tr>
<th>Mean</th>
<th>N</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>A 0.024007</td>
<td>21</td>
<td>Surprised</td>
</tr>
<tr>
<td>B 0.001215</td>
<td>20</td>
<td>Sad</td>
</tr>
<tr>
<td>B 0.000000</td>
<td>24</td>
<td>Neutral</td>
</tr>
<tr>
<td>B -0.001378</td>
<td>15</td>
<td>Fear</td>
</tr>
<tr>
<td>B -0.001722</td>
<td>22</td>
<td>Happy</td>
</tr>
<tr>
<td>B -0.002611</td>
<td>17</td>
<td>Disgust</td>
</tr>
<tr>
<td>B -0.007830</td>
<td>18</td>
<td>Angry</td>
</tr>
</tbody>
</table>
Center of nose to left corner of lip distance

The table below shows the SNK grouping for the center of nose to left corner of lip distance parameter. It can be observed that surprised can be grouped in group A. Fear, happy and sad can be grouped in group B. Happy, sad and neutral can be grouped in group C. Angry and disgust can be grouped in group D.

Table 4.7: SNK grouping for center of nose to left corner of lip distance parameter

<table>
<thead>
<tr>
<th>Mean</th>
<th>N</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.16359</td>
<td>21</td>
</tr>
<tr>
<td>B</td>
<td>0.07323</td>
<td>15</td>
</tr>
<tr>
<td>C  B</td>
<td>0.03710</td>
<td>22</td>
</tr>
<tr>
<td>C  B</td>
<td>0.03310</td>
<td>20</td>
</tr>
<tr>
<td>C</td>
<td>0.00000</td>
<td>24</td>
</tr>
<tr>
<td>D</td>
<td>-0.04462</td>
<td>18</td>
</tr>
<tr>
<td>D</td>
<td>-0.07129</td>
<td>17</td>
</tr>
</tbody>
</table>

Center of nose to right corner of lip distance

The table below shows the SNK grouping for the center of nose to right corner of lip distance parameter. It can be observed that surprised can be grouped in group A. Fear, happy and sad can be grouped in group B. Happy, sad and neutral can be grouped in group C. Angry and disgust can be grouped in group D.

Table 4.8: SNK grouping for center of nose to right corner of lip distance parameter

<table>
<thead>
<tr>
<th>Mean</th>
<th>N</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.16359</td>
<td>21</td>
</tr>
<tr>
<td>B</td>
<td>0.07323</td>
<td>15</td>
</tr>
<tr>
<td>C  B</td>
<td>0.03710</td>
<td>22</td>
</tr>
<tr>
<td>C  B</td>
<td>0.03310</td>
<td>20</td>
</tr>
<tr>
<td>C</td>
<td>0.00000</td>
<td>24</td>
</tr>
<tr>
<td>D</td>
<td>-0.04462</td>
<td>18</td>
</tr>
<tr>
<td>D</td>
<td>-0.07129</td>
<td>17</td>
</tr>
</tbody>
</table>
Center of nose to center of chin distance

The table below shows the SNK grouping for the center of nose to center of chin distance parameter. It can be observed that surprised can be grouped in group A. Disgust and happy can be grouped in group B. Happy, sad, neutral, fear and angry can be grouped in group C.

Table 4.9: SNK grouping for center of nose to center of chin distance parameter

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>N</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.32302</td>
<td>21</td>
<td>Surprised</td>
</tr>
<tr>
<td>B</td>
<td>0.06331</td>
<td>17</td>
<td>Disgust</td>
</tr>
<tr>
<td>C B</td>
<td>0.02725</td>
<td>22</td>
<td>Happy</td>
</tr>
<tr>
<td>C</td>
<td>0.00000</td>
<td>24</td>
<td>Neutral</td>
</tr>
<tr>
<td>C</td>
<td>-0.00704</td>
<td>14</td>
<td>Fear</td>
</tr>
<tr>
<td>C</td>
<td>-0.00858</td>
<td>20</td>
<td>Sad</td>
</tr>
<tr>
<td>C</td>
<td>-0.01384</td>
<td>19</td>
<td>Angry</td>
</tr>
</tbody>
</table>

Inner corner of left eye to left corner of lip distance

The table below shows the SNK grouping for the inner corner of left eye to left corner of lip distance parameter. It can be observed that surprised can be grouped in group A. Sad, neutral, fear and angry can be grouped in group B, and Happy and disgust in group C.

Table 4.10: SNK grouping for center of nose to left corner of lip distance parameter

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>N</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.16395</td>
<td>21</td>
<td>Surprise</td>
</tr>
<tr>
<td>B</td>
<td>0.02632</td>
<td>20</td>
<td>Sad</td>
</tr>
<tr>
<td>B</td>
<td>0.00000</td>
<td>24</td>
<td>Neutral</td>
</tr>
<tr>
<td>B</td>
<td>-0.01704</td>
<td>15</td>
<td>Fear</td>
</tr>
<tr>
<td>B</td>
<td>-0.03039</td>
<td>18</td>
<td>Angry</td>
</tr>
<tr>
<td>C</td>
<td>-0.11752</td>
<td>22</td>
<td>Happy</td>
</tr>
<tr>
<td>C</td>
<td>-0.12669</td>
<td>17</td>
<td>Disgust</td>
</tr>
</tbody>
</table>
Inner corner of right eye to right corner of lip distance

The table below shows the SNK grouping for the inner corner of right eye to right corner of lip distance parameter. It can be observed that surprised can be grouped in group A. Sad, neutral, fear and angry can be grouped in group B. Happy and disgust can be grouped in group C.

Table 4.11: SNK grouping for center of nose to right corner of lip distance parameter

<table>
<thead>
<tr>
<th>Mean</th>
<th>N</th>
<th>expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>A 0.16395</td>
<td>21</td>
<td>Surprise</td>
</tr>
<tr>
<td>B 0.02632</td>
<td>20</td>
<td>Sad</td>
</tr>
<tr>
<td>B 0.00000</td>
<td>24</td>
<td>Neutral</td>
</tr>
<tr>
<td>B -0.01704</td>
<td>15</td>
<td>Fear</td>
</tr>
<tr>
<td>B -0.03039</td>
<td>18</td>
<td>Angry</td>
</tr>
<tr>
<td>C -0.11752</td>
<td>22</td>
<td>Happy</td>
</tr>
<tr>
<td>C -0.12669</td>
<td>17</td>
<td>Disgust</td>
</tr>
</tbody>
</table>

4.3 PARAMETER BASED CLASSIFICATION

Based on the observations from the trend lines for different parameters and the statistical analysis, a parameter based system was developed in which all the points in the testing data were compared with the mean value of a particular expression. When the value lied within one standard deviation from the mean, the image was classified under that expression. Maximum numbers of correct classifications were observed for the surprised expression. Thus, a system to classify the surprised expression using a committee of seven individual parameters was developed. The committee consisted of parameters which showed significant difference for surprised expression in SNK.
grouping. These parameters were fractal dimension of upper lip, fractal dimension of lower lip, Center of nose to left corner of lip distance, Center of nose to right corner of lip distance, Center of nose to center of chin distance, Inner corner of left eye to left corner of lip and Inner corner of right eye to right corner of lip. Table 4.12 shows an example of the classification system. A ‘1’ signifies that the value lies within one standard deviation from the mean. A ‘0’ signifies that the value lies within one standard deviation from the mean. This parameter based system could correctly classify 30 out of 34 surprised cases. The system correctly classified 152 out of remaining 156 cases of non-surprised expressions. The overall accuracy achieved was 95.79% with 182 out of 190 correct classifications.
Table 4.12: Example of parameter based classification system

<table>
<thead>
<tr>
<th></th>
<th>Fractal dimension of upper lip</th>
<th>Fractal dimension of lower lip</th>
<th>Center of nose to left corner of lip distance</th>
<th>Center of nose to right corner of lip distance</th>
<th>Center of nose to center of chin distance</th>
<th>Inner corner of left eye to left corner of lip</th>
<th>Inner corner of right eye to right corner of lip</th>
<th>Committee Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.0405</td>
<td>0.0612</td>
<td>0.1754</td>
<td>0.1719</td>
<td>0.3167</td>
<td>0.1647</td>
<td>0.1695</td>
<td></td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.0343</td>
<td>0.0437</td>
<td>0.1041</td>
<td>0.1063</td>
<td>0.121</td>
<td>0.0703</td>
<td>0.0711</td>
<td></td>
</tr>
<tr>
<td>Fear</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Surprised</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
4.4 RESULTS FROM NEURAL NETWORKS

The integrated committee classification system correctly identified 180 out of 190 cases from 40 different subjects. Table 4.13 shows an example of a primary committee classification output for the sad expression for a particular subject. In this example, six out of seven members classified the expression as sad while in one case the output was all 0. The committee decision was based on majority opinion and thus, the committee output was sad.

Table 4.13: Sample results for individual networks and committee results in the primary network case for sad expression for subject 85 (S-085).

<table>
<thead>
<tr>
<th>Network</th>
<th>Expression</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Neutral</td>
<td>Angry</td>
<td>Disgust</td>
<td>Fear</td>
<td>Happy</td>
<td>Sad</td>
<td>Surprised</td>
</tr>
<tr>
<td>NN - 1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>NN - 2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NN - 3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>NN - 4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>NN - 5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>NN - 6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>NN - 7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Committee Result</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
Table 4.14 shows an example of a secondary committee classification output for the sad expression for a particular subject. In this example, four out of five members classified the expression as disgust while one of the networks classified it as angry. The committee decision was based on majority opinion and thus, the committee output was disgust.

Table 4.14: Sample results for individual networks and committee results in the secondary network case for sad expression for subject 82 (S-082).

<table>
<thead>
<tr>
<th>Network</th>
<th>Expression</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Neutral</td>
<td>Angry</td>
<td>Disgust</td>
<td>Fear</td>
<td>Sad</td>
</tr>
<tr>
<td>NN - 1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NN - 2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NN - 3</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NN - 4</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NN - 5</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Committee Result</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 4.15 shows an example of an integrated committee neural network system. In case of the primary committee, three of the networks had fear as output, three of the networks had all 0 outputs and one network had disgust as the output. As a result, the classification output of the primary committee was ambiguous. Since the decision of the
primary committee was ambiguous, the data was then fed to the secondary committee consisting of 5 networks. The committee output for the secondary committee was observed as fear. Thus, the output of the integrated committee neural network system was fear.

Table 4.15: Sample results for individual networks and committee results in the integrated committee case for fear expression for subject 88 (S-088).

<table>
<thead>
<tr>
<th>Network</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Neutral</td>
</tr>
<tr>
<td>NN - 1</td>
<td>0</td>
</tr>
<tr>
<td>NN - 2</td>
<td>0</td>
</tr>
<tr>
<td>NN - 3</td>
<td>0</td>
</tr>
<tr>
<td>NN - 4</td>
<td>0</td>
</tr>
<tr>
<td>NN - 5</td>
<td>0</td>
</tr>
<tr>
<td>NN - 6</td>
<td>0</td>
</tr>
<tr>
<td>NN - 7</td>
<td>0</td>
</tr>
<tr>
<td>Primary Committee</td>
<td>Ambiguous</td>
</tr>
<tr>
<td>Secondary Committee</td>
<td>0</td>
</tr>
<tr>
<td>Integrated Committee</td>
<td>0</td>
</tr>
</tbody>
</table>
Table 4.16 shows the performance of all the seven member networks along with the committee results. Network NN-2 correctly classified 160 out of 190 expressions and was the best performance network amongst the recruited 7 networks with an accuracy of 84.21%.

Table 4.16: Correct classifications by individual networks and committee system

<table>
<thead>
<tr>
<th>Network</th>
<th>Number of correct classifications</th>
<th>% Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>NN - 1</td>
<td>151</td>
<td>79.47%</td>
</tr>
<tr>
<td>NN - 2</td>
<td>160</td>
<td>84.21%</td>
</tr>
<tr>
<td>NN - 3</td>
<td>152</td>
<td>80.00%</td>
</tr>
<tr>
<td>NN - 4</td>
<td>155</td>
<td>81.58%</td>
</tr>
<tr>
<td>NN - 5</td>
<td>144</td>
<td>75.79%</td>
</tr>
<tr>
<td>NN - 6</td>
<td>143</td>
<td>75.26%</td>
</tr>
<tr>
<td>NN - 7</td>
<td>149</td>
<td>78.42%</td>
</tr>
<tr>
<td>Committee result</td>
<td>162</td>
<td>85.26%</td>
</tr>
<tr>
<td>Integrated committee result</td>
<td>180</td>
<td>94.74%</td>
</tr>
</tbody>
</table>
A comparison plot of integrated committee performance and the seven individual neural network performance is shown in figure 4.12. The committee showed a higher rate of correct identification than individual neural networks.

![Comparison between committee performance and individual network performance](image)

Figure 4.12: Comparative plot of individual neural network output v/s primary and integrated committee neural network outputs.

The comparative expression classification performance by the integrated committee neural network system is shown in figure 4.13. It gives a plot of expression wise performance of the primary and integrated committee neural network system. The angry, disgust and fear expressions showed low classification accuracy while neutral,
happy, and sad and surprised showed high classification accuracy in the primary neural networks. The performance was improved after integration of the primary and secondary committee systems. The integrated committee neural network system showed more than 90 % correct classifications except for disgust which showed only 78% correct classifications.

Figure 4.13: Plot of correct expression classifications for primary and secondary committee classification systems compared to total expressions.
An expression wise accuracy performance of the integrated committee neural network classification system is shown in figure 4.14. The angry, disgust and fear expressions showed low classification accuracy in the range of 50 % to 70 %, while neutral, happy, sad and surprised showed high classification accuracy of more than 85 % for the primary neural network classification system. For the integrated system, all expressions except disgust showed more than 90 % accuracy.

<table>
<thead>
<tr>
<th>Expressions</th>
<th>Primary neural networks</th>
<th>Integrated neural networks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angry</td>
<td>61.54</td>
<td>92.31</td>
</tr>
<tr>
<td>Disgust</td>
<td>68.18</td>
<td>94.12</td>
</tr>
<tr>
<td>Fear</td>
<td>77.27</td>
<td>94.12</td>
</tr>
<tr>
<td>Happy</td>
<td>95.65</td>
<td>97.06</td>
</tr>
<tr>
<td>Sad</td>
<td>87.50</td>
<td>100.00</td>
</tr>
<tr>
<td>Surprised</td>
<td>87.08</td>
<td>100.00</td>
</tr>
<tr>
<td>Neutral</td>
<td>100.00</td>
<td>100.00</td>
</tr>
</tbody>
</table>

Figure 4.14: Plot of percentage correct classifications v/s expressions

4.5 SUMMARY OF EVALUATION OF THE INTEGRATED COMMITTEE SYSTEM

The integrated committee system correctly classified 180 out of 190 expressions. There were 10 incorrect classifications. The incorrect classifications were either misclassifications, or ambiguous classification. A misclassification accounts for an expressions being classified as an expression, which is actually not being posed by the
subject. An ambiguous classification accounts for two or more than two expressions being identified for a classification output. There were seven misclassification cases, and three ambiguous classification cases amongst the 190 expressions evaluated. A comparative summary of the total expression classification analysis is shown in figure 4.15. Table 4.17 shows the confusion matrix.

![Output Summary](image)

Figure 4.15: Plot of output summary in terms of percent accuracy for different classification types.
Table 4.17: Confusion matrix

<table>
<thead>
<tr>
<th>Expressions presented</th>
<th>System Classification</th>
<th>Neutral</th>
<th>Angry</th>
<th>Disgust</th>
<th>Fear</th>
<th>Happy</th>
<th>Sad</th>
<th>Surprised</th>
<th>all 0</th>
<th>Ambiguous</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutral</td>
<td>Neutral</td>
<td>40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>40</td>
</tr>
<tr>
<td>Angry</td>
<td></td>
<td>12</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>13</td>
</tr>
<tr>
<td>Disgust</td>
<td></td>
<td>2</td>
<td>17</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>22</td>
</tr>
<tr>
<td>Fear</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>23</td>
</tr>
<tr>
<td>Happy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>34</td>
</tr>
<tr>
<td>Sad</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>24</td>
</tr>
<tr>
<td>Surprised</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>34</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>40</td>
<td>14</td>
<td>17</td>
<td>23</td>
<td>33</td>
<td>26</td>
<td>34</td>
<td></td>
<td>3</td>
<td>190</td>
</tr>
</tbody>
</table>
The seven misclassification cases were observed when angry, disgust, fear or sad expression data were presented as input to the system. Three angry expressions were misclassified as sad and one angry expression was misclassified as disgust. Two disgust expressions were misclassified as angry, two were misclassified as fear and one disgust expression was misclassified as sad. Seven fear expressions were misclassified as happy and one fear expression was misclassified as disgust. One sad expression was misclassified as angry.
CHAPTER V
DISCUSSION

The present study demonstrated an improvement in performance accuracy over the previously presented system with addition of new parameters. The study is rather unique as several new parameters such as entropy and fractal dimension (of the eyebrow and lips) were developed and used in the classification system. A total of twelve new parameters were quantified (Section 3.3) and were used for classification in addition to the existing fifteen parameters to classify emotions. Two classification approaches were used. First, a parameter based classification system was developed which classified the expressions based on the actual parameter values directly. In the second approach, the committee neural network system was used to classify seven basic emotion types from facial images. The integrated committee neural network system consisting of primary and secondary networks, could classify the emotion in the given facial image into either of the following expression groups: neutral, angry, disgust, fear, sad, surprised or happy with an accuracy of 94.73%. (Table 4.16) Binomial test ($\alpha=0.01$) showed that the probability of the integrated committee neural network system, giving a correct decision is greater than its probability of giving an incorrect decision ($p_1 > p_2$) and thus, the results of the study failed to support the null hypothesis and the alternate hypothesis was accepted (Appendix A).
5.1 ANALYSIS OF RESULTS

Two types of classification approaches were used viz. parameter based system and neural network based system. The results are analyzed in the following paragraphs.

5.1.1 Parameter based system:

The parameter-based system was used to classify the expressions based on the actual values for the various parameters used in the study. In the case of current study, a total of 27 parameters were used as inputs to the neural network system. These parameters were analyzed statistically to check their individual performance in the classification of the expressions. It was observed that none of the parameters were able to classify the expressions into seven different groups. SNK grouping could classify the expressions in two, three or four groups (Tables 4.1 to 4.11). It was observed that in case of seven out of eleven real valued parameters (Section 3.3.2), values for surprised expression showed significant difference from other expressions. Thus, a committee of these seven parameters was made and then all the testing data was fed to the committee. The values were checked for the range of ± standard deviations from the mean of all the data for the surprised expression. If the value was found to be in that range, the expression was classified as surprised. If the value was outside the range, then it was classified into other group of the remaining six expressions. (Table 4.12) The results from the parameter based system showed that; it could correctly classify the surprised expression with a greater accuracy. The committee of parameters used for surprised expression could classify the data with 95.79% accuracy. Such parameter based system could only be applied to the surprised expression and thus, neural network system was used to classify the expressions.
5.1.2 Neural network based system:

The results of the present study showed that the accuracy was improved with addition of new parameters. Networks with different sets of inputs were recruited in the committee. This further increased the accuracy of classification. The integrated committee neural network system further increased the performance of classification due to incorporation of secondary neural network committees. The integrated committee neural network decision provided accurate and reliable classification in 94.73% of the cases. The system could classify 180 out of 190 cases correctly (Table 4.16). Seven cases were misclassifications and three cases were ambiguous classifications.

Various parameters like variability in the database, parameters extracted, neural network training and testing etc could affect the overall performance of the system. The database used for the study consisted of subjects performing series of expressions. Most of these expressions were assumed to be performed deliberately and the database was classified into six basic facial expressions (angry, disgust, fear, happy, sad and surprised) [14]. There could be a difference between deliberate and spontaneous expressions. Expressions performed deliberately could be singular showing the specified expression whereas in reality, a spontaneous expression could be combination of two or more expressions (e.g. fear and surprise). Thus, the variability and reliability of these expressions can introduce variability in the overall dataset. Additional variability due to the environmental conditions like illumination etc was avoided by using the same lighting conditions and same equipment for all the images.

Other factors affecting the performance of the system included the parameters fed to it, the training function and the transfer function used etc. The performance was also
affected by the preprocessing done before feeding it to the neural network. The data was normalized to ensure efficient training of the networks. The database was divided into distinct training, initial testing and final testing groups ensuring that the classification system performed true classification.

The unique feature of the system was addition of the complexity measure parameters (Section 3.3.2). Another new feature was that, six groups of networks were trained where each group had different number of training input parameters used. This helped in obtaining optimum performance from different combinations of the input parameters. 120 such networks were trained. It was observed that the networks with combination of all the inputs from the previous study and the current study performed better than the other combinations consisting of only the parameters from previous study or only the new parameters derived. The best performing networks were then recruited in the committee.

A committee neural network system gave better performance than individual network classifiers. No single network classification results were as good as committee classification results. (Table 4.16) Since the primary network outputs resulted in misclassifications and ambiguous classifications in case of the angry, disgust, sad and fear expressions. (Table 4.17), secondary networks were used to classify these expressions. An integrated committee neural network classification system, constituting of a combination of primary committee networks and secondary committee networks, further increased the performance of the classification system. (Figure 4.13 and Figure 4.14)
5.2 COMPARISON WITH THE PREVIOUS SYSTEMS

Ekman and Friesen [6] developed a Facial Action Coding System (FACS) to measure the facial behavior. In FACS, they used Action Units (AUs) based on the muscular activity that produces momentary changes in the facial expression. They developed action unit scores which do not directly give the meaning of the movement. The investigator has to interpret the meaning from the FACS scores. Hara and Kobayashi [8, 9], Padgent [7], Zhang [10] and Zhao [11] have all used a neural network approach using back propagation learning technique to classify images into six or seven emotional categories. Padgett et al., [7] trained neural networks from the data of 11 subjects and tested with the data from one subject. The training and testing dataset was interchanged and new networks were trained and tested. The average recognition rate achieved was 86 \%.

Hara and Kobayashi [8, 9] trained neural networks from data of 15 subjects (90 images) and tested them by data from another 15 subjects. The average recognition rate achieved was 85 \%. Zhang et al., [12] used the JAFFE data base consisting of 10 Japanese female subjects. Although they achieved results with 90.1\% accuracy; they used the same data for training and testing. Zhao et al., [11] used the Ekman and Friesen database [13] to achieve a 100 \% recognition rate, but they used the same data for training and testing. Khan et al used thermal methods to quantify the facial expressions [14]. He could achieve an accuracy of only 56 \%. Thus, the present system provides significant improvement in accuracy (94.73\%) when compared to all previous studies discussed above.

The system presented by Kulkarni et al [19] could correctly classify the expressions in 90.4\% of cases. They used a committee neural network system with fifteen
parameters including eight real valued parameters and seven binary parameters. All the fifteen inputs were used to train the primary and secondary networks. The parameters selected did not contain any image complexity measures.

In comparison, the present study used a large dataset consisting of 97 subjects. The integrated committee neural network system was developed with a combination of primary and secondary committees. The present study included total of 27 parameters including fifteen parameters from previous study. The new parameters extracted included real valued measures including distance measured as well as the image complexity measured such as entropy and fractal dimension. Neural networks were trained using different combinations of the available 27 input parameters. This further helped in optimizing the performance of the system. With addition of new parameters and the new approach of training the networks, the new system could classify the expression data in 94.73 % of cases.

5.3 LIMITATIONS OF THE STUDY

Classification techniques based on the present study require neutral images of the individual. For terror deterrent applications, neutral images may not be available and neutral image independent techniques may be needed to classify the correct expression of the person. The images used in the study consisted of only the frontal views of the face. In reality, it is hard to achieve an exact frontal image of the patient or a suspect. Also, the parameters, in the present study, were manually extracted from the image. An automated process will make the system faster and more useful for real-time applications. The accuracy could be further improved by first classifying the image into a neutral, positive
(happy and surprised) and negative (angry, disgust, fear, and sad). Then the image could be sub-classified by developing specialist committee networks for each expression. Another limitation of the study was it did not consider the gesture of laugh which may have similar values as for surprised expression for some of the parameters. This limitation could be attributed to the database as it did not contain any images related to laughing.

5.4 SIGNIFICANCE OF THE STUDY

The study was conducted on a database consisting of total 467 images from 97 subjects. 190 images from 40 subjects were used for the final evaluation of the system, out of which, 180 were correctly classified by the system giving an accuracy of 94.73%. These numbers are fairly large to support the reliability of the system. The database used was diverse in terms of ethnicity, age and sex. In most of the investigations reviewed in the previous chapters, the systems were based on or they used smaller datasets for the study. The current system could successfully classify a large dataset in seven different expression groups with an accuracy of 94.73%. Thus, the system developed in the present study can be considered as a significant step forward in correctly identifying the facial expression involving large data bases involving images from a diversified ethnic group.

5.5 CLINICAL SIGNIFICANCE OF THE STUDY

The presented system of classification can be efficiently used in the field of psychology, stress detection and in the field of terror deterrence. The use of facial
expression analysis in psychopathology may give us information related to the diagnostic information relevant to depression, mania, schizophrenia and other disorders, and also information relevant to monitoring response to treatment [2,3]. Thus, expression analysis can be effectively used as a tool for behavioral studies and medical rehabilitation. Stress detection through facial expression analysis can be proved useful in cases like stress detection for astronauts [4].
CHAPTER VI
CONCLUSIONS

The presented system was able to correctly classify the expressions in seven different expression groups (neutral, angry, disgust, fear, sad, happy and surprised). The specific conclusions drawn from the study are as follows:

1. Twelve new parameters were successfully derived in addition to the existing parameters in the previous study from 97 subjects (467 facial images) for seven different expressions.
2. A parameter based classification system was developed which was used effectively to classify the surprised expression.
3. Sets of networks were trained with different combinations of the input parameters from all 27 parameters.
4. Primary neural networks were trained to classify the image into seven different expressions.
5. Secondary neural networks were trained to classify the image into four different expressions (angry, disgust, fear and sad).
6. Committee neural networks were recruited for both primary and secondary neural networks.
7. An integrated committee neural network system was developed incorporating primary neural network committee and secondary neural network committee.

8. Performance of the integrated committee neural network system was evaluated by testing the system with an independent dataset of 40 subjects (190 facial images). The system correctly classified 180 cases, seven cases were misclassifications and three cases were ambiguous classifications. The overall system achieved an accuracy of 94.73 % which was 4.31 % more than the previous study.

9. The result of the binomial test showed that the z value lies in the critical region and thus, we failed to accept the null hypotheses of the study and thus we can say that the addition of new parameters increased the overall performance of the system.
CHAPTER VII
FUTURE WORK

1. The current data consisted of facial images taken from 97 subjects in the age from 18 to 30 years. The database had 65 percent female subjects. 15 percent of the subjects were African-American and three percent were Asian or Latino. A study could be conducted on a more diverse database which will make the classification more reliable.

2. The parameters were extracted manually and one at a time in the current system. An automated parameter extraction system could be developed to extract all the expressions at a time.

3. A study could be conducted in which the contribution of each parameter in the expression could be analyzed.

4. A study focusing on the clinical applications of the expression analysis could be undertaken.

5. A study focusing on the persuasive computers could be undertaken.
REFERENCES


39. MATLAB Neural Networks Toolbox, Users Guide


APPENDICES
APPENDIX A

STATISTICAL ANALYSIS

Null Hypothesis: The probability of the committee neural network system making a correct decision \( (p_1) \) is the same as the probability of the committee neural network system making an incorrect decision \( (p_2) \). \( (p_1 = p_2) \)

\[ H_0: p_1 = p_2 = 0.5 \]

That is, Committee Neural Networks cannot be used to effectively classify the different facial expression based on input facial parameters.

Alternate Hypothesis: The probability of the committee neural network system making a correct decision \( (p_1) \) is greater than the probability of the committee neural network system making an incorrect decision \( (p_2) \). \( (p_1 > p_2) \)

\[ H_1: p_1 > p_2 \]

Committee Neural Networks can be used to effectively classify the different facial expression based on input facial parameters.
BINOMIAL TEST:

For sample size $N > 25$,

$$z = \frac{( ( x \pm 0.5 ) - N * p_1 )}{( N * p_1 * p_2 )^{1/2}}$$

Where,

$N =$ sample size

$p_1 =$ probability of the committee neural network system making a correct decision.

$p_2 =$ probability of the committee neural network system making an incorrect decision.

$x =$ number of incorrect classifications.

For $x = 10$, $P = 0.5$, $Q = 0.5$, $N = 190$,

$$z = -12.26$$

The probability associated with the occurrence under $H_0$ for the observed value of $z$, is $p < 0.0003$. This probability $p$ is smaller than $\alpha = 0.01$. Thus we reject $H_0$ in favor of $H_1$. We conclude that $p_1 > p_2$, that is, the probability of the committee neural network system making a correct decision ($p_1$) is greater than the probability of the committee neural network system making an incorrect decision ($p_2$).
APPENDIX B

DATABASE AGREEMENT

AGREEMENT ON USE OF IMAGE DATA
Cohn-Kanade Facial Expression Database

I agree

- to cite Kanade, Cohn, & Tian (2000) in any paper of mine or my collaborators that makes any use of the database. The reference is:
- to use the images for research purposes only.
- not to provide the images to second parties.
- if I reproduce images in electronic or print media, to use only those from the following subjects and include notice of copyright (©Jeffrey Cohn).

<table>
<thead>
<tr>
<th>Size</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>S32</td>
<td>$55</td>
</tr>
<tr>
<td>S121</td>
<td>$124</td>
</tr>
</tbody>
</table>

Signature: [Signature]

Name: Gayatri Patnkar and NARENDER REDDY

Title: Graduate Student

Institution: University of Akron

Date: 02/11/08

Email address: gpat@uakron.edu

ngreddy@uakron.edu

Nicole Taubman
Project Coordinator
Affect Analysis Group
Univ. of Pittsburgh
ntrb@pitt.edu
APPENDIX C
DETAILS OF COMMITTEE MEMBERS

The primary committee consisted of seven members while the secondary committee consisted of five members. The members of the committees were trained using different combinations of input parameters. The various input parameters and the detailed specifications of the neural networks are listed in this appendix.

Inputs Parameters:

1. Eyebrow Raise distance
2. Upper eyelid to eyebrow distance
3. Inter-eyebrow distance
4. Upper eyelid – lower eyelid distance
5. Upper lip thickness
6. Lower lip thickness
7. Mouth width
8. Mouth opening
9. Upper teeth visible
10. Lower teeth visible
11. Forehead lines
12. Eyebrow Lines
13. Nose Lines
14. Chin Lines
15. Nasolabial lines
16. Center of nose to center of chin distance
17. Center of nose to left corner of lips
18. Center of nose to right corner of lip
19. Inner corner of left eye to left corner of lip
20. Inner corner of right eye to right corner of lip
21. Fractal dimension of eyebrow
22. Fractal dimension of upper lip
23. Fractal dimension of lower lip
24. Entropy of eyebrow
25. Entropy of upper lip
26. Entropy of lower lip
27. Wrinkles on lips

Primary Committee:

NN1:

- Number of Inputs: 26 (Inputs 1-26)
- Number of outputs: 7
- Number of layers: 3 (1 hidden layer)
- Number of Hidden layer neurons: 15
• Training function: TRAINLIM (Levenberg-Marquardt back propagation
• Transfer function for the hidden layer :TANSIG
• Transfer function for the output layer : LOGSIG
• Number of epochs: 100

NN2:
• Number of Inputs: 26 (Inputs 1-26)
• Number of outputs :7
• Number of layers: 3 ( 1 hidden layer)
• Number of Hidden layer neurons: 21
• Training function: TRAINLIM (Levenberg-Marquardt back propagation
• Transfer function for the hidden layer :TANSIG
• Transfer function for the output layer : LOGSIG
• Number of epochs: 100

NN3:
• Number of Inputs: 26 (Inputs 1-26)
• Number of outputs :7
• Number of layers: 3 ( 1 hidden layer)
• Number of Hidden layer neurons: 17
• Training function: TRAINLIM (Levenberg-Marquardt back propagation
• Transfer function for the hidden layer :TANSIG
• Transfer function for the output layer : LOGSIG
• Number of epochs: 100
NN4:

- Number of Inputs: 11 (Inputs 16-26)
- Number of outputs: 7
- Number of layers: 3 (1 hidden layer)
- Number of Hidden layer neurons: 17
- Training function: TRAINLIM (Levenberg-Marquardt back propagation)
- Transfer function for the hidden layer: TANSIG
- Transfer function for the output layer: LOGSIG
- Number of epochs: 100

NN5:

- Number of Inputs: 11 (Inputs 16-26)
- Number of outputs: 7
- Number of layers: 3 (1 hidden layer)
- Number of Hidden layer neurons: 30
- Training function: TRAINLIM (Levenberg-Marquardt back propagation)
- Transfer function for the hidden layer: TANSIG
- Transfer function for the output layer: LOGSIG
- Number of epochs: 100
NN6:

- Number of Inputs: 18 (Inputs 1-15, 24-26)
- Number of outputs: 7
- Number of layers: 3 (1 hidden layer)
- Number of Hidden layer neurons: 15
- Training function: TRAINLIM (Levenberg-Marquardt back propagation)
- Transfer function for the hidden layer: TANSIG
- Transfer function for the output layer: LOGSIG
- Number of epochs: 100

NN7:

- Number of Inputs: 24 (Inputs 1-20, 24-26)
- Number of outputs: 7
- Number of layers: 3 (1 hidden layer)
- Number of Hidden layer neurons: 20
- Training function: TRAINLIM (Levenberg-Marquardt back propagation)
- Transfer function for the hidden layer: TANSIG
- Transfer function for the output layer: LOGSIG
- Number of epochs: 100
The Secondary Committee:

NN1:

- Number of Inputs: 27 (Input 1-27)
- Number of outputs :5
- Number of layers: 3 ( 1 hidden layer)
- Number of Hidden layer neurons: 15
- Training function: TRAINLIM (Levenberg-Marquardt back propagation)
- Transfer function for the hidden layer :TANSIG
- Transfer function for the output layer : LOGSIG
- Number of epochs: 100

NN2:

- Number of Inputs: 27 (Input 1-27)
- Number of outputs :5
- Number of layers: 3 ( 1 hidden layer)
- Number of Hidden layer neurons: 17
- Training function: TRAINLIM (Levenberg-Marquardt back propagation)
- Transfer function for the hidden layer :TANSIG
- Transfer function for the output layer : LOGSIG
- Number of epochs: 100
NN3:

- Number of Inputs: 27 (Input 1-27)
- Number of outputs: 5
- Number of layers: 3 (1 hidden layer)
- Number of Hidden layer neurons: 20
- Training function: TRAINLIM (Levenberg-Marquardt back propagation)
- Transfer function for the hidden layer: TANSIG
- Transfer function for the output layer: LOGSIG
- Number of epochs: 100

NN4:

- Number of Inputs: 26 (Input 1-26)
- Number of outputs: 5
- Number of layers: 3 (1 hidden layer)
- Number of Hidden layer neurons: 15
- Training function: TRAINLIM (Levenberg-Marquardt back propagation)
- Transfer function for the hidden layer: TANSIG
- Transfer function for the output layer: LOGSIG
- Number of epochs: 100
NN5:

- Number of Inputs: 26 (Input 1-26)
- Number of outputs: 5
- Number of layers: 3 (1 hidden layer)
- Number of Hidden layer neurons: 25
- Training function: TRAINLIM (Levenberg-Marquardt back propagation)
- Transfer function for the hidden layer: TANSIG
- Transfer function for the output layer: LOGSIG
- Number of epochs: 100