Facial Expression Recognition System using Discriminant Features

Submitted in

fulfillment of the requirements for the degree of

Doctor of Philosophy

by

Abhishek Singh Kilak

ID: 2012RCP9516

Under the Supervision of

Dr. Namita Mittal

Associate Professor

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
MALAVIYA NATIONAL INSTITUTE OF TECHNOLOGY JAIPUR

September 2019

DECLARATION

I, Abhishek Singh Kilak, declare that this thesis titled, **"Facial Expression Recognition System using Discriminant Features"** and the work presented in it, are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this university.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this university or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself, jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Date **Abhishek Singh Kilak**

(2012RCP9516)

CERTIFICATE

This is to certify that the thesis entitled **"Facial Expression Recognition System using Discriminant Features"** being submitted by Abhishek Singh Kilak (2012RCP9516) is a bonafide research work carried out under my supervision and guidance in fulfillment of the requirement for the award of the degree of Doctor of Philosophy in the Department of Computer Science and Engineering, Malaviya National Institute of Technology, Jaipur, India. The matter embodied in this thesis is original and has not been submitted to any other University or Institute for the award of any other degree.

Place: Jaipur
Date:

Dr. Namita Mittal
Associate Professor
Department of Computer Science and Engineering
MNIT Jaipur

Dedication

The Thesis is dedicated to my Wife, Kids and Parents.

Acknowledgement

This thesis would not been possible without the contribution of many individuals, to whom I express my appreciation and gratitude. First of all, I am deeply indebted to my supervisor Dr. Namita Mittal, for the knowledge and experience shared with me. She was always available whenever I needed her guidance. Her invaluable suggestions and advice helped me a lot in nurturing my research work in the right direction. She is very kind and supportive. It is impossible for me to imagine having a better supervisor for my Ph.D. It was really an honour to work with her.

I am also grateful to all the members of my doctoral guidance committee: Dr. Girdhari Singh, Dr. Dinesh Goplani and Dr. Mushtaq Ahmed for their constructive suggestions helped me a lot to improve my PhD work. I would like to thank all the learned faculty members for their support and help during this period. I am also thankful to non-teaching staff of Department of Computer Engineering for all the help.

My special thanks to all my wonderful friends and fellow researchers who made the whole environment of the lab very refreshing. They celebrated my achievements and cheered me up in my difficult moments. Discussion with co-researchers helped me to remove the small mistakes in research work.

Last but not least my sincere thanks goes to my family and friends, for their endless encouragement and support during the course of my studies. Their love and care always motivated me to work harder and harder. For any inadequacies that may remain in this work, the responsibility is entirely my own.

Sign: ____

Date: ____

Contents

List of Figures

List of Tables

ABSTRACT

The emotional states can be inferred in real life situations from the traits like the way of talking, walking, sitting and other gestures. The importance of emotions is evident in our everyday chores, the routine perspective and working of people relies upon the emotional states and vice versa. They have an important role in social, business and affective sciences. There has been significant research work done in the past four decades upon emotion with intervening fields varying from history, medicine, sociology, neuroscience and psychology and computer technology.

Intelligent systems capable of cognitive capabilities require emotion inferring based on the various inputs which may be based on the data collected from voice analysis, breath analysis, text analysis, keystroke, EEG, brain mapping, ECG, gesture analysis or analysis of facial expressions. Various brain mapping methods facilitate to study the changes in the neurological functions like EEG, ECG, MRI, PET, fMRI are intrusive. They look promising, but due to the subject under consideration having knowledge of data being collected, limited mobility and use of equipment for collecting data hamper their applicability. On the other hand, voice and gesture analysis techniques are non-intrusive. Voice or speech gives an insight to emotional state as tone, pitch and volume vary in tandem. However, this technique fails to cater to input requirements in case the person does not respond to stimuli or in other words does not speak. Gesture analysis is also a way of gathering information to infer emotional state. But they have limited capability of the same type as that of speech analysis. They also fail in case of no or limited response to the stimulating emotion.

Facial expressions are a global language revealing elemental emotions. Face expressions provide an insight into the emotional state of human beings. Face expressions change in tandem with emotional states and people generally have limited or no control over their facial movements. Also if people might try to control their expressions in some specific temporal space but facial expressions are generally spontaneous and in tandem to their emotional state.

Intelligent emotion detection systems are required that respond in tandem to human feelings. There are enormous real life applications of such systems such as sentiment analysis of people arising out of business and political decisions, affective computing, robotic assistance, mob controlling, driving assistance, personalized recommenders, depression, bipolar and anxiety detection, IoT applications, Crowd behavior analysis and many more.

The capability of interpreting emotions based on expressions has been an intensive field of study for more than three decades for developing intelligent systems. Machines that can intelligently identify emotions based on human facial expressions are gaining social importance. Many techniques for the detection of emotions from facial expressions are proposed, but there is still scope of improvement in accuracy of detection.

Humans can predict the emotional state of people of different ethnicity. But such is not the case with computers as they need training for different texture, shape and appearance of the target subjects. This requires the creation of a database that is essential to train the systems, test them, validating of applicability of algorithms for emotion recognition and classification systems for building robust systems. Database creation is a toilsome and tedious task. However, it is a prerequisite to building systems that are capable of detecting emotions. Creation of a posed expression database requires proper guidance and training of the subjects by experts. Afterward, validation of database is an equally important concern which involves labeling of the images with categories that they belong to. The objective of the thesis is to develop techniques that lead to improvement in emotion detection accuracy. Also, to study the existing databases of facial images and to develop dataset that incorporates Indian facial images and experiment upon it for emotion detection accuracy.

In this thesis, validation that facial action units and expressions can be used to detect emotions is done. New approaches have been proposed for improving of accuracy of emotion classification. The usage of databases for experiments is affected by the problem generalization versus over fitting. Enhanced performance on a certain standard dataset is likely due to exploitation of biases that are specific to it. In order to have generic solution to the problem of emotion detection, algorithms should be free from bias and non-realistic assumptions. A new database is created which has facial images of Indian people. The database has posed expressions of 102 participants and has 896 images. Also it is annotated it for emotion classification. Various algorithms are applied on this database and other standard datasets.

The outcomes and results look promising and may be utilized for further experiments involving other models for the creation of hybrid vectors over different datasets that may further increase the accuracy of classification of emotions.

Chapter 1

Introduction

The emotional states can be inferred in real life situations from the traits like the way of talking, walking, sitting and other gestures. Even ages ago, many physiognomy studies suggested that the character or personality of a human can be judged based on appearance and the face [1].

Emouvoir was the French word from which "emotion" was coined back in 1579. Different terms like affection, sentiment and passion were grossed up to a single heterogeneous word called emotion. Thomas Brown coined this word in the early 17th century which led to the emergence of "No one felt emotions before about 1830. Instead, they felt other things - passions, accidents of the soul and, moral sentiments" - and explained them very differently from how we understand emotions today "[2].

Definition of emotion as per the Oxford dictionary is “A strong feeling deriving from one's circumstances, mood or relationships with others" [3]. Emotions are responses to significant internal and external events [4].

The Expression of the Emotions in Man and Animals written by Charles Darwin in 1872 commenced the study of emotions from the evolution point of view. It is asserted that emotions helped in survival and also communication in human beings

[50]. Also, it was claimed that as natural selection led to emotion expression development; therefore it has universally available complementary parts across cultures.

A minimum of five senses is concurred involving the sense of touch, taste, vision, smell and hearing in literature for humans [45]. Some literary works suggest that there are two more senses present in humans other than these. There is broad disagreement among neurologists upon the number of senses. This field of study has varied topics encompassing field theory, classification and operation.

Reality is experienced on the basis of our senses which are based upon stimulus perception physiologically from outside and inside our body. These stimulators are either exteroceptors that receive stimulations out of our bodies or interceptors that receive stimulation from inside [46]. External stimulators include the sense of taste, smell, sight etc. also pave the way to emotional reactions [47]. There are various internal stimulators which depend on internal functions of our body like pH balance, sugar levels, blood pressure etc. which changes in emotions as well [51]. The reaction of the exteroceptors and interceptors vary from person to person, which is generally based upon their likings like certain smell may be pleasant for someone and unpleasant for another. The reaction depends on how our brain responds to the stimulators [48, 49]. Heavy breathing, muscular tension, sweating or rapidly beating of heart may lead us to our body reaction of fright entitized by our nervous system.

The emotional mental state is an assimilation of feelings, thought process or behavior response. Facial expressions are not only a reflex to emotional condition but are also a part of communication [150, 151]. Motivation, personality, temperament, disposition and mood are the causes of the emotional outcome. Also, some theories suggest that emotions are related to cognition [52-54]. Varied theories have been proposed that explain origination, neuroscience and emotional functioning have been researched extensively. Emotions are activated due to psychological and physical variations and impact behavior. Alternatively, emotions are related to habitual behavior. Sociable people exhibit emotions with ease while loner people generally hide emotions [55, 56]. Emotions have been defined to be sparked as resultant of consciousness and cognition by the body. Emotions are described as "A combination of the physical properties of your body, a flexible brain that wires itself to whatever environment it develops in and your culture and upbringing, which provide that environment" [14].

Both negative and positive motivations are associated with emotions [57, 58]. Emotions are outcomes of segments including feelings, physiological differences, motivation and behavior [59, 60]. They are not seminal and none of the segments causally defines emotion. Also, these components are not resultant of emotions. The different parts attributing to emotions include physiological changes, psychological, cognition, operant conditioning, expressive behavior and conscious experience [61-64]. These are studied by academicians depending on their fields of study. There has been significant research work done in the past four decades upon emotion with intervening fields varying from history, medicine, sociology, neuroscience and psychology and computer technology [65-67].

Depending on the physiological changes produced, emotions can be categorized as positive or negative [68, 69]. Emotions can range from extreme reactions to mild reactions and they can last for a short duration to long duration [70, 71]. Like anxiety can be in the form of little worry to an extreme case [72]. ‘Anger' can be momentary or for a few minutes and ‘grief' can last for a much longer period [73-75]. Emotions involve a chain of responses including psychological, neural mechanisms, behavioral and verbal reactions [76-78].

Intelligent systems capable of cognitive capabilities require emotion inferring based on the various inputs which may be based on the data collected from voice analysis, breath analysis, text analysis, keystroke, EEG, brain mapping, ECG, gesture analysis or analysis of facial expressions [79-85]. These recent research fields in emotion recognition involve material development for stimulation and elicitation of emotions [109]. Various brain mapping methods facilitate to study the changes in the neurological functions like EEG, ECG, MRI, PET, fMRI are intrusive [86-88]. They look promising, but due to the subject under consideration having knowledge of data being collected, limited mobility and use of equipment for collecting data hamper their applicability.

On the other hand, voice and gesture analysis techniques are non-intrusive [90, 91]. Voice or speech gives an insight to emotional state as tone, pitch and volume vary in tandem. However, this technique fails to cater to input requirements in case the person does not respond to stimuli or in other words does not speak. Gesture analysis is also a way of gathering information to infer emotional state [89, 90]. But, they have limited capability of the same type as that of speech analysis [92]. They also fail in case of no or limited response to the stimulating emotion.

However, people have limited or scant control over the facial expressions and also even if they try to restrict the response to stimuli, they don't perform at all times [93]. These can be used to interpret the emotional state of people and we further explore this technique. Facial expressions are a global language revealing elemental emotions [94-98]. The importance of emotions is evident in our everyday chores; the routine perspective and working of people rely upon the emotional states and vice versa [99, 100].

The capability of interpreting emotions based on expressions has been an intensive field of study for more than three decades to develop intelligent systems. Machines that can intelligently identify emotions based on human facial expressions are gaining social importance. With the increased capability of processing and reduction in the size of the devices, this intelligence can now even be incorporated to even small handheld devices such as smartphones [101, 102].

Generic steps for detecting emotions based on facial expressions are presented in Figure 1.1.

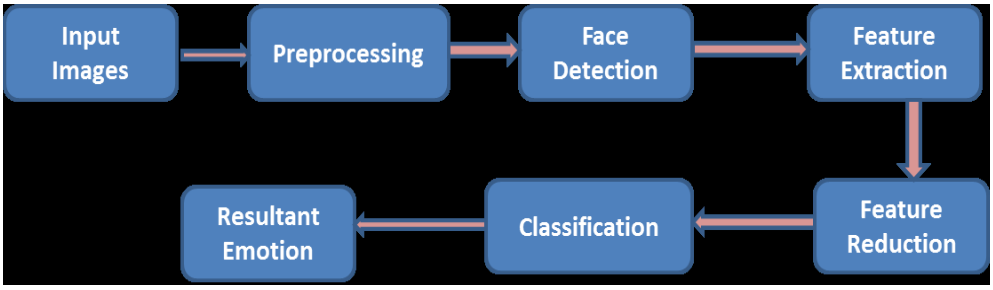

Figure 1.1: Steps required for detecting emotions from facial expressions

In recent years, exhibiting emotions on electronic media has become common and has even led to their usage in communication through SMS, Whatsapp and other messengers in the form of emoticons and lingual short forms. For example, LOL is posted for Laughing Out Loud; OMG is used for Oh My God etc.

An example of the usage of the emotion representations on electronic media is shown below:

😐 Neutral Face

🙂 Smiling Face

☹ Sad Face

😨 Fearful Face

😠 Angry Face

😲 Astonished Face

Figure 1.2: Emotion exhibition on electronic media

Facial action coding system (FACS) comprises of Action Units that describe face muscle movements. Critical and decisive information about the state of mind of a person may be acquired from the face which is an abstraction of the muscle movements leading to spatial and temporal changes [103]. Also, this information has an indispensable relation to origins, age factor and gender. It is a communication capability which is disparate to verbal communication but yet may handout in understanding the emotional state in both cognizant and unaware conditions of expression [104].

Face expressions provide an insight into the emotional state of human beings. Face expressions change in tandem with emotional states and people generally have limited or no control over their facial movements. Also if people might try to control their expressions in some specific temporal space but expressions usually are spontaneous and in tandem to their emotional state. Therefore emotion detection based on analysis of facial features is undoubtedly the best way to detect emotions as it is not limited to the lab environment, unobtrusive, coherent to the emotional state and practically applicable [105, 106].

Detection of emotions from facial expressions involves extracting mathematical information from spatial variations of the faces. This spatial information also invariably involves ethnicity, gender, scale correction as well as noise features involving occlusion sub mounting to beard, mustache, hair locks, face rotation, resolution and physical hindrances towards full face detection [107, 108].

Development of systems capable of emotion classification invariably requires training to catch the strains that effectively and categorically map them to different categories [110]. This requires data for the systems to train, which has the following two aspects. Either the training data can be too personalized to serve for the only particular subject involved [111, 112]. E.g., Personal response systems catering to specific requirements of individuals. This requires catching individual data and catering tailor-made solutions based on them which cannot be applied to generically. Or, there may be robust general systems that are capable enough to provide services to the unknown masses based on rigorous training provided that involves non-ideal conditions of illumination, gender, occlusions and other parameters [113, 114].

The wide variety of subjects and their emotion labeling may help researchers in developing robust algorithms for futuristic artificially intelligent systems. Several evaluations of accuracy are done to behave as a baseline by researchers to develop more robust algorithms.

1.1 Motivation

Facial expressions are a global language revealing elemental emotions. The importance of emotions is evident in our everyday chores, the routine perspective and working of people rely upon the emotional states and vice versa. They have an important role in social, business and affective sciences.

Intelligent emotion detection systems are required that respond in tandem to human feelings. There are enormous real-life applications of such systems such as sentiment analysis of people arising out of business and political decisions, affective computing, robotic assistance, mob controlling, driving assistance, personalized recommenders, depression, bipolar and anxiety detection, IoT applications, Crowd behavior analysis and many more.

Face expressions provide an insight into the emotional state of human beings. Face expressions change in tandem with emotional states and people generally have limited or no control over their facial movements. Also, even if people might try to control their expressions in some specific temporal space, but expressions are generally spontaneous and in tandem to their emotional state. Therefore emotion detection based on analysis of facial features is undoubtedly the best way to detect emotions.

1.2 Research Gap

Apart from being a point of deliberation, literature backs that emotional states affect the facial features. Six prevalent emotions that can be traced athwart cultures worldwide are that of happiness, surprise, fear, anger, disgust and sadness are as described by Paul Ekman [5].

The capability of interpreting emotions based on expressions has been an intensive field of study for more than three decades for developing intelligent systems. Machines that can intelligently identify emotions based on human facial expressions are gaining social importance. Many techniques for the detection of emotions from facial expressions are proposed, but there is still scope of improvement in accuracy of detection.

Humans can predict the emotional state of people of different ethnicity. But such is not the case with computers as they need training for different for the texture, shape and appearance of the target subjects. Because of the difference in shape and texture of people across the globe, headway in the field of affective computing requires different databases covering ethnicity. To our knowledge, there is no posed expression database for Indian faces that catches all the mentioned emotion classes.

This requires the creation of a database that is essential to train the systems, test them, validating of applicability of algorithms for emotion recognition and classification systems for building robust systems. Database creation is a toilsome and tedious task. However, it is a prerequisite to building systems that are capable of detecting emotions. Creation of a posed expression database requires proper guidance and training of the subjects by experts. Afterward, validation of database is an equally important concern which involves labeling of the images with categories that they belong to.

Detection of emotions based upon the facial expressions is important for the building of intelligent systems that may respond in tandem to human feelings. The various features that increase the complexity of emotion recognition systems include ethnicity, gender, pose, hair locks, beard etc. Therefore there had been an emergence of several databases covering various features that are available publicly emotion recognition. But none of them is for posed Indian origin faces.

The gap is bridged by providing Bharat Database which contains facial images of Indian people.

1.3 Objectives

The Research objectives of the work are as follows:

- To explore the feasibility of using facial expressions to detect emotions and apply techniques for the same.
- To develop new techniques for emotion detection
- To develop a dataset that incorporates Indian facial images annotated for emotion classes.
- To experiment upon Indian Facial Database as well as other databases in order to not have over-fitting but to have a generic solution to the problem of emotion detection, as algorithms should be free from bias and non-realistic assumptions.

1.4 Contributions

The contributions of the thesis are as follows:

1. Validation that facial action units and expressions can be used to detect emotions and afterward using the unexplored feature extraction technique Laplacian of Gaussian filter for emotion classification
2. A new technique of Local Binary Pattern Half used in conjunction with Histogram of oriented gradients
3. Building a Bharat Database which contains facial images of Indian people. The database has posed expressions of 102 participants and has 896 images. Also, it is annotated for emotion classification.
4. A new approach of Compact Local Binary Pattern Technique for building a hybrid feature vector used in conjunction with Histogram of Oriented Gradients. Also a novel approach of Multiple Instances Based Emotion Detection Using Discriminant Feature Tracking technique is introduced.

1.5 Thesis Structure

Chapter 1 introduces expressions, emotions, applications, motivation, research gap, objectives and contributions.

Chapter 2 is a literature review of existing techniques of emotion classification.

Chapter 3 validates the use of facial action units for emotion classification. Feature extraction technique of Laplacian of Gaussian filters which is explored for the classification of emotions. Also a new technique Local Binary Pattern Half used in conjunction with Histogram of oriented gradients.

In Chapter 4 we build a Bharat Database which contains facial images of Indian people. The database has posed expressions of 102 participants and has 896 images. We also annotate it for emotion classification. We propose a Compact Local Binary Pattern Technique for building hybrid feature vector and also we propose Multiple Instances Based Emotion Detection Using Discriminant Feature Tracking Technique.

Chapter 5 is the conclusion of the work done and future scope of work that can emanate from this thesis.

Chapter 2

Literature Survey

The book that was written in the 19th century by Charles Darwin "The Expression of Emotion in Man and Animals" is one of the pioneering works in the field of emotion studies [1]. Expression recognition is an ongoing research field. The general steps involved in this process of emotion detection from facial features include face recognition followed by preprocessing, feature extraction, feature selection and classification. For each of the step defined above, there are different techniques suggested in the literature.

The emotion detection techniques are explored and also a database is created in the work. Therefore, to maintain the information related to both of the domains, the literature survey is two-fold. First, different techniques used for dynamic facial expression detection are explored. After that, the various databases available for facial expression detection are discussed.

The Facial image was cropped into lower and upper part [15] and then fiducial points were applied to the two portions. Afterward, Gabor filters are applied to extract features. Multi-layer perceptron is used for classification. They have reported an accuracy of 81.6% and 86.6% accuracy for lower and upper face respectively.

Expression recognition is also affected by aging. The lifespan [17] database is used by Guodong Guo[16]. The fiducial points were manually labeled and Gabor filters[18] were used to extract facial features. SVM is used for classifier learning and Test is performed using 10-fold cross-validation. An average accuracy of 69.32% is reported for the cross-age group. The same setup when repeated for FACES [19] database yields an average accuracy of 64.04%. Landmark selection and tracking using Gabor filters for feature extraction followed by classifications using a combination of HMM and SVM are done by Valstar et al. [121] which has the drawback of expensive computationally. Gabor features provide better results, but they are both computational and memory intensive [122, 123].

Some approaches propose personalized classifiers as they argue that there may be a huge difference in Train pattern as compared to an unseen person. The unseen person may have heavy eyebrows, wrinkled face and such difference may lead to erroneous results. But such an approach is unrealistic. Selective Transfer machine has been proposed for a generic classifier that removes person specific biases [9]. This approach is someway between personal and generic classifier. This approach uses unsupervised learning and hence labels are not required.

Both person dependent and independent tests have been performed by Cohen et al. [115]. Piecewise B´ezier Volume Deformation tracker [116] is used and features are extracted as surface patches for Motion Units. For the experiments, they use their database as well as the Cohn-Kanade database. For the person dependent case on their database, they achieved an average accuracy of 80.5 % using Naïve Bayes, 83.1% using TAN, 80.81% using NN and 82.46% using HMM classifiers. In the person independent case, they achieved accuracy of 64.77% using NB, 66.53% using TAN, 66.44% using NN and 58.63% using Multilevel HMM. The accuracy achieved is 68.14% with NB, 73.22% with TAN and 73.81% with NN.

Principal Component Analysis [139, 140] and Linear Discriminant analysis [141-143] are widely used for dimensionality reduction and classification. Higher performance has been reported by Oh et al. [144] using a fusion of PCA and LDA. Local fisher discriminant analysis based encrypted emotion classification system is proposed by Rahulamathavan et al. [145] greater accuracy even in normal images.

Use of Viola and Jones [7] for detection of the face and afterward mouth, nose and lips are detected using SDAM [22] approach is used [23]. A classifier based on SVM is used to classify on JAFFE[8]. The same experiment is repeated on images from the web and images captured by them in their laboratory. They have reported an accuracy of 92.4%.

The difference between posed and spontaneous expressions was further explored [24]. Posed images are taken from Cohn Kanade dataset. The spontaneous expressions were captured using content images taken from the International Affective Picture system [25]. Dataset acquisition is done on a set of 52 users. For LDOS-PerAff-1[26] dataset the images from all frames are averaged to yield a neutral frame. Filtering was done using Gabor filters and standard deviation and mean were calculated. The feature vector of 240 elements was reduced to 80 using PCA for LDOS-PerAff-1 and 72 features for CK dataset describing 95% variance. KNN is used for emotion classification and 62% accuracy is reported for the spontaneous expression set.

Emotion Recognition model having subject independency is developed by Matsugu et al. [27] uses single structure Convolution Neural Network for finding the differences in the emotion and neutral faces. Two independent CNN models are used by Fasels et al. [28] for expression recognition and face identification respectively. Images of 10 subjects are used for experiments and an accuracy of 97.6% was obtained. The use of the neural network ensemble is done for classification [20] of images. An average Recognition rate of 83.7 is reported.

Registration of faces based on symmetry was used in experiments by Li et al. [29] for both 2-dimensional and 3-dimensional emotion detection wherein cubical spline interpolation is used. For the noise introduced by the hair on the subject's facial images. Principal Component Analysis followed by SVM and Linear Discriminant Analysis (LDA) is used for smile recognition through which accuracy of around 80% is obtained for 2-dimensional facial images and more than 90% for 3-dimensional facial images.

A novel approach using Curvelet features followed by Bilateral Two dimension Principle Component Analysis and Extreme Machine Learning is introduced by Mohammed et al. [30] which is used on several different databases. B2DPCA is used for feature size reduction of the mega feature size obtained from the two curvelet algorithms.

Active appearance Model uses a combination of both texture and shapes for object representation. Features are extracted using non-rigid landmarks [124-128]. AAM provides better results in comparison to Active Shape Models. It does not provide good performance in person independent cases. Landmark detection using Constrained Local Model provides better results in person independent cases [129]. Regularized Landmark Mean Shift (RLMS) proposed by Saragih et al. [130] which is a modification of CLM provides better results of accuracy in localization of landmarks. Discriminative Response Map Fitting method for Constrained Local Model that provides a generic solution in both natural and controlled imaging conditions is proposed by Asthana et al. [131]. It provides good results but is computationally intensive which hampers its applicability in real-world scenario.

Features extracted from the detected facial landmarks are required by classification algorithms. Kalman filters along with Infra-Red illumination is used for tracking landmarks which results in improvement of performance as this includes both texture and geometric features [118]. Transient features like wrinkles and relative distances were calculated applying Canny edge detector is used by Tian et al. [119], but it has limitation of feature extraction under different illumination conditions. Selecting landmarks based on matching of mouth and eye region and using image difference for emotion classification is perfomed by Uddin et al. [120].

Local Directional Number Pattern extraction method on varying compass masks (Kirsch and Gaussian derivative) for computation of feature vector under varying test conditions of illumination, time-lapse and noise is used by Rivera et al. [31]. The results are compared for different emotions and they claim that their technique is reliable and robust under different illumination conditions. The use of Median ternary pattern is also proposed and used [21] for the extraction of features from CK database. They report a recognition rate of 89.1 for a grid of 3*3 using SVM for classification.

Convolution Neural Network and Recurrent Neural Network hybrid architecture at varying frame sizes are used by Kahou et al. [32] for emotion classification. The improvement in accuracy is acclaimed to fusion at decision and feature level using Multi-Layer Perceptron. Many deep neural network architectures are used by Enrique Correa et al. [33] for emotion classification. The first is based on Krizhevsky and Hinton [34] in which image size reduction is done using max-pooling. The resulting accuracy achieved is 63%. The second approach uses three fully connected layers and normalization which lead to an accuracy of 53%. In the third approach, layers of convolution are used followed by normalization and max-pooling which lead to 63% accuracy.

CNN having seven layers with random initialization and retraining on larger dataset was done by Yu et al. [35] where the CNN were combined using minimizing of both log-likelihood loss and hinge loss. Static Facial Expressions in the Wild database [36] is used for the experiments.

Local Binary Patterns are also widely used for extraction of features [133-135]. A modification is done on LBP and termed as Local Description Patterns is proposed by Jabid et al. [136] which yields better performance lo Local Binary Patterns. Local Phase Quantization technique used for emotion classification by Dhall et al. [137] claims to have achieved even better performance. Local Directional Pattern Variance proposed by Kabir et al. which uses local variance for encoding contrast information. It is claimed that for analyzing images of low resolution, LBP features are robust [138]. It is claimed by Shan et al. [37] that Local Binary Pattern (LBP) features can be extracted at a quicker rate compared to Gabor wavelets. Adaptive boosting is used for learning of more discriminative LBP features. They use different algorithms for classification like Template matching, LDA and SVM and proved that SVM provided the best accuracy for classification using these features.

Some approaches use images of the full face while others specific facial features for extraction of features. Shan et al. [146] have proposed an approach in which they divide the image of the face into many sub-parts and then features are extracted. AdaBoost is then used on LBP histogram bins for classification. Various approaches similar to this technique have been used [147-149]. However, any fractional misalignment of sub-parts of the image would lead to erroneous classification. Also, it is not certain that the same part of the face will lie in the same sub-region for images of different people.

Detection of emotions based upon the facial expressions is important for the building of intelligent systems that may respond in tandem to human feelings. Therefore there had been an emergence of several databases covering various features that are available publicly emotion recognition. The different databases are discussed in brief in this section.

Now, the related work on databases available for emotion detection is presented. One of the most widely used databases is Cohn-Kanade Action Unit Coded Facial Expression database [132]. It comprises of 486 image sequences posed by 97 subjects is released in the year 2000. Image sequence proceeds from neutral face image to extreme expression. The peak expression images are coded using the Facial Action Coding System and annotation is provided in the form of emotion labels.

This dataset is extended to address as Extended Cohn- Kanade (CK+) database [6]. The images of 123 subjects corresponding to 593 sequences are taken which are coded using Facial Action Coding System for the last or peak frame of the sequence. The Action Units and their intensity are provided for the peak expression images. Also, the images are tracked using Active Appearance Model for 68 landmark points. Out of the total sequences, only 327 are having corresponding emotion files. This is because only these sequences are validated. The emotion labels are neutral, anger, contempt, disgust, fear, happy, sadness and surprise.

Japanese Female Facial Expression (JAFFE) database has 213 images 10 female Japanese models that posed for both neutral and six basic expressions. JAFFE database is created by Lyons et al. [8]. The images are in grayscale. Binghamton University-3D Dynamic Facial Expression database [9] has 2500 3D face expressions of 100 subjects having the six basic expressions having four intensity levels. The different aspects considered include age, race and culture.

The MMI Database initially conceived in 2002 to serve as a source that can be used across facial expression recognition community [10]. It has videos that have a sequence from neutral to apex and back to neutral expression. It has over 2900 videos of 75 subjects as well as still images. The videos are annotated for presence of Action Units.

The Belfast Database [11] has different sets of over 250 colored video clips depicting natural emotions at varying resolutions. Multimedia Understanding Group (MUG) has 1462 posed color sequences of 86 subjects which are annotated with emotion labels. The Radboud Faces Database (RaFD) has posed color images of 67 subjects at five different camera angles and three different gaze directions for eight emotion labels.

Indian Spontaneous Expression Database (ISED) [12] has 428 spontaneous color videos of 50 subjects having emotion labels of sad, happy, surprise and disgust only. Denver intensity of spontaneous facial action database (DIFSA) [13] is a color video database of 27 subjects whereby each video sequence is of 4845 frames of spontaneous reactions while viewing a video of 4-minute duration. Six intensity level annotations of Action Units are provided for the facial expressions.

The usage of databases for experiments is affected by the problem generalization versus over fitting. Enhanced performance on a certain standard dataset is likely due to exploitation of biases that are specific to it. In order to have generic solution to the problem of emotion detection, algorithms should be free from bias and non-realistic assumptions [117].

The various features that increase the complexity of emotion recognition systems include ethnicity, gender, pose, occlusion, beard, mustache etc. The type of database used for learning by systems is of crucial importance. Many databases exist for this purpose, but none of them is for posed Indian faces.

This gap is bridged by providing Bharat Database which contains facial images of Indian people. The database has posed expressions of 102 participants and has 896 images. The participants are asked to pose for different emotions by showing them images eliciting those emotions as well as with the help of expert artists. The annotation is done using polling by a panel of three experts. This database will further help the community involved in developing algorithms for emotion recognition.

Chapter 3

Emotion classification using various techniques

Emotions describe the mental state of human beings. They can be observed from their behavior as to how they are talking, responding, sitting and other gestures. Various systems have been proposed in the literature that can be used to identify human emotions.

These include but are not limited to brain mapping, voice sampling, facial expression recognition, gesture recognition, keyboard pressure information and breath information.

The association of facial features with emotions is a widely accepted fact as people respond in tandem to feelings like anger, joy and surprise. Also, they have either limited or no control over their facial expression in general. Therefore this fact has been utilized in many of the techniques offered for the detection of emotions.

Human emotions are primarily defined as Anger, Disgust, Fear, Happiness, Surprise and Sadness [5]. Changes in facial expressions are a natural reaction to situations in real life and correspond to the mental state of the people. These features can be used to analyze the emotional state of people. This technique for the detection of emotions is explored further in the chapter.

The Motion Units which had numeric values as compared to Boolean-valued Action units are used as features. The datasets used by them are Cohn-Kanade [6] and Authentic Expression Database developed by them which is labeled for Neutral, Surprise, Joy and Disgust. They left the angry, contempt, fear and sad emotion. Various classifiers are compared using these two datasets and 10-fold cross validation is performed comprising 95% confidence intervals for the determination of classification errors.

Facial action coding system (FACS) comprises of Action Units that describe face muscle movements. A subset is shown in table 3.1 and its complete description can be found in Ekman[5] which defines all the other possible movements of the facial muscles.

Table 3.1: Action Unit Representation

Action Unit	Description of the unit
1	Inner Brow Raiser
2	Outer Brow Raiser
4	Brow Loarer
5	Upper Lid Raiser
6	Cheek Raiser
7	Lid Tightener
9	Nose Wrinkler
10	Upper Lip Raiser

3.1 Proposed usage of explicit Action Units

The files having an emotion label associated with them are considered in the experiments from the Cohn-Kanade+ AU-Coded Facial Expression Database [6]. Their associated Facial action coding system files are used for the construction of the feature vector. Sample images and their associated action units are shown in figure 3.1 and 3.2 where the first number represents the action unit while the second number before semicolon corresponds to the intensity of that particular unit.

Figure 3.1: Images from database exhibiting different emotions

Figure 3.2: Images from database exhibiting different emotions

The first image of figure 3.1 shows surprise emotion and is represented by action units 1, 2, 5, 25 and 27. Similarly, the second and third image show disgust and happy emotion respectively. The images in figure 3.2 represent anger, happy and fear emotions. The classification is performed using Naïve Bayes, Support Vector Machine, Multi-Layer perceptron and K Nearest neighbor classifiers. The steps are shown in figure 3.3.

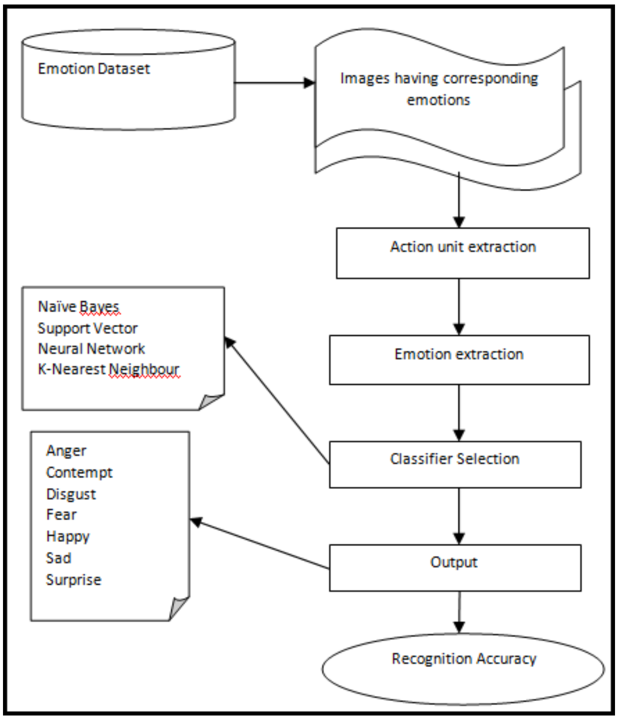

Figure 3.3: Steps involved in emotion recognition

For training and testing, the set is split at 70-30, 80-20 and 85-15 ratios. Also, the 10-fold cross validation is carried out. The test is performed using one versus all criterion. The action units are used as elements of the feature vector and the classes are those provided in the respective emotion files. Also, the contempt class combinations are used.

3.2 Results and Discussions

The results in terms of accuracy for the Naïve Bayes, Support Vector Machine, Neural Network and K Nearest Neighbor classifier are shown in table 3.2, 3.3, 3.4 and 3.5 respectively. In each table, the first column defines the emotion under consideration. The second, third and fourth column of each table corresponds to results obtained at split ratios of 70-30, 80-20 and 85-15 of the dataset for training and testing respectively. The fifth column shows the accuracy obtained at 10-fold cross-validation.

Table 3.2: Accuracy of classification using Naïve Bayes

Emotion	Train-Test Set Split			10-fold cross-validation
	70-30	80-20	85-15	
Anger	92.7	90.6	93.8	94
Contempt	95.8	93.8	93.8	94.7
Disgust	79.2	82.8	83.3	81.5
Fear	75	73.4	75	77.7
Happy	89.6	89.1	85.4	88.1
Sad	95.8	92.2	95.8	93.1
Surprise	88.5	85.9	87.5	85.9

The lowest recognition accuracy is 73.4% for fear emotion using Naïve Bayes. This is because angry and sad emotions have some common facial movements. Therefore, this probabilistic classifier fails to classify correctly because of not having strong conditional independence between the feature vectors used to represent that class.

Table 3.3: Accuracy of classification using SVM

Emotion	Train-Test Set Split			10-fold cross-validation
	70-30	**80-20**	**85-15**	
Anger	87.5	84.4	81.3	85.9
Contempt	95.8	95.3	93.8	96.8
Disgust	78.1	82.8	81.3	79.6
Fear	93.7	93.8	95.8	92.2
Happy	91.7	89.1	93.8	87.8
Sad	91.7	87.5	87.5	91.2
Surprise	90.6	89.1	89.6	90

Table 3.4: Accuracy of classification using Neural Network

Emotion	Train-Test Set Split			10-fold cross-validation
	70-30	80-20	85-15	
Anger	93.8	98.4	95.8	94.4
Contempt	99	98.4	97.9	99.7
Disgust	85.4	82.8	85.4	86.2
Fear	94.8	92.2	97.9	94.4
Happy	94.8	98.4	97.9	94.7
Sad	92.7	90.6	93.8	95.9
Surprise	99	100	100	97.2

The Neural Network performed well but the amount of time required for training of neural network is manifold of the rest of the classifiers which is because of the adaptations and approximations performed by the network. This does not undermine it's applicability as training is a cumbersome process for this classifier as opposed to testing.

Table 3.5: Accuracy of classification using KNN

Emotion	Train-Test Set Split			10-fold cross-validation
	70-30	80-20	85-15	
Anger	96.9	96.9	100	96.6
Contempt	99	98.4	97.9	99.7
Disgust	94.8	95.3	95.8	94
Fear	97.9	96.9	97.9	96.6
Happy	96.9	96.9	95.8	96.9
Sad	95.8	96.9	95.8	96.6
Surprise	97.9	100	100	98

K Nearest Neighbor (KNN) performs well as it relies on instance based learning by approximating of the functions locally and deferring the computation until classification. As learning is performed at every step of introduction of any new example to a particular cluster, this system became robust to classification. This approach provides a minimum of 94% accuracy and a maximum of 100% for classification.

3.3 Laplacian of Gaussian Filters approach for emotion classification

Cognitive science is an active multidisciplinary field of research. For emotion detection, several approaches are proposed in the literature. People have limited or scant control over the facial expressions and also even if they try to restrict the response to stimuli, they don't perform at all times. These can be used to interpret the emotional state of people and we further explore this technique.

Laplace filters are used for data collection from face images which may, in turn, are used to classify emotions. However, susceptibility to noisy feature is a disadvantage to these filters. Laplacian of Gaussian filter approach for building feature vector is proposed for facial images which are immune to noise. Binary class combinations of emotion labels are used for classification. Extended Cohn- Kanade (CK+) [6] database is used for carrying out the experiments.

3.3.1 Proposed Approach

The steps for emotion recognition from the images of the database include

(i) Taking out the Region of Interest or Pre Processing

(ii) Application of filter

(iii) Derivation of Feature Vector

(iv) Emotion Classification

Formally this can be viewed as

$$I \rightarrow I_p \rightarrow I_f \rightarrow \breve{E} \rightarrow \bar{E}$$

(i) Viola Jones [7] technique is used for detecting face part in the images of the database. The first image of any subject in the subfolders of subjects of the CK+ database exhibits neutral expression while the last image corresponds to maximal expression. Therefore these are used for detecting face part and after that cropping is done to take out the region of interest.

Cropped images I_p of three subjects who provided consent for publication is shown in the following two figures.

Figure 3.4: First image I_p of S52, S55 and S74 cropped exhibiting neutral expression.

Figure 3.5: Last image I_p of the subjects cropped exhibiting maximal expression.

(ii) Laplacian of Gaussian filters are applied to the images obtained from the previous step which are more immune to noise than Laplace filter alone. Filters are applied to the images having corresponding emotion labels. Laplacian of Gaussian filter equations are as follows:

$$f_g(x,y) = e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{3.1}$$

$$f(x,y) = \frac{(x^2+y^2-2\sigma^2)f_g(x,y)}{2\pi\sigma^2 \sum_x \sum_y f_g} \tag{3.2}$$

The sigma values are incremented ten times for obtaining different filtering values. As a result, ten filtered images I_f are obtained for each of the initial image I_p.

Sample filtered images I_f for 0.25 sigma value and 5*5 filter size are shown in the following figure.

Figure 3.6: The first image from the subfolders of subjects after application of filter

Figure 3.7: Last image from the subfolders of subjects after application of filter

(iii) Then first two statistical moments are calculated for images obtained after application of LOG filters. Mean and standard deviation are obtained μ_f, σ_f for first image and μ_l, σ_l for the last image. Six features dependent on these independent features are also computed which are $\mu_f - \mu_l$, $\sigma_f - \sigma_l$, μ_f/μ_l, σ_f/σ_l, $(\mu_f - \mu_l)/\mu_f$ and $(\sigma_f - \sigma_l)/\sigma_f$. Feature vector Ě comprising of 100 elements is obtained as a result for each image.

(iv) Binary class combinations of emotion labels are used for classification which are Anger-Disgust, Anger-Fear, Anger-Happy, Anger Sad, Anger-Surprise, Disgust-Fear, Disgust-Happy, Disgust-Sad, Disgust-Surprise, Fear-Happy, Fear-Sad, Fear-Surprise, Happy Sad, Happy-Surprise and Sad- Surprise. For classification, binary class SVM is used.

3.3.2 Results and Discussion

Table 3.6, 3.7 and 3.8 depict the results obtained for different training and test set split ratios. Table 3.9 shows the percentage of accuracy obtained in classification at 10-fold cross-validation.

Table 3.6: Accuracy of classification in percentage at 70-30 split ratio of training and test set

	Anger	Disgust	Fear	Happy	Sad	Surprise
Anger	-	82.1	58.8	78.4	75	85.3
Disgust	82.1	-	73.9	96.7	82.6	60
Fear	58.8	73.9	-	72	83.3	89.7
Happy	78.4	96.7	72	-	92	72.1
Sad	75	82.6	83.3	92	-	82.8
Surprise	85.3	60	89.7	72.1	82.8	-

Table 3.7: Accuracy of classification in percentage at 80-20 split ratio of training and test set

	Anger	Disgust	Fear	Happy	Sad	Surprise
Anger	-	84.2	81.8	75	72.7	82.6
Disgust	84.2	-	60	95	93.3	55.6
Fear	81.8	60	-	76.5	75	89.5
Happy	75	95	76.5	-	94.1	82.1
Sad	72.7	93.3	75	94.1	-	73.7
Surprise	82.6	55.6	89.5	82.1	73.7	-

Table 3.8: Accuracy of classification in percentage at 85-15 split ratio of training and test set

	Anger	Disgust	Fear	Happy	Sad	Surprise
Anger	-	85.7	87.5	72.2	75	82.4
Disgust	85.7	-	58.3	93.3	90.9	70
Fear	87.5	58.3	-	76.9	50	86.7
Happy	72.2	93.3	76.9	-	84.6	76.2
Sad	75	90.9	50	84.6	-	85.7
Surprise	82.4	70	86.7	76.2	85.7	-

Table 3.9: Accuracy of classification in percentage at 10-fold cross-validation

	Anger	Disgust	Fear	Happy	Sad	Surprise
Anger	-	74.2	69.6	74.6	77.8	83.2
Disgust	74.2	-	80.5	87.1	75	65.7
Fear	69.6	80.5	-	74.1	79.5	79.5
Happy	74.6	87.1	74.1	-	85.7	81.7
Sad	77.8	75	79.5	85.7	-	80.2
Surprise	83.2	65.7	79.5	81.7	80.2	-

The results show a high degree of accuracy for almost all pairs. This shows the effectiveness of using this technique for detecting emotions. The results confirm the usability of this technique for affective computing.

3.4 Classification of emotions from images using localized subsection information

Emotional intelligence has important social significance and literature indicates that facial features are an important factor in determining the emotional state. It has been an intense study field to build systems that are capable of recognizing emotions automatically based on facial expressions.

Various approaches are proposed, but still, there is a scope of improvement in detection accuracy because of diverse form of expressions exhibiting the same emotion. A widely used approach in the object detection field is Histogram of Oriented Gradients [38].

In this chapter extensive experiments are conducted using various subsection sizes of images of Histogram of Oriented Gradients and also along with Local Binary Pattern to extract the features for classification of emotions from facial images. Quantitative analysis of the approach in comparison with others is done to show its applicability and effectiveness.

3.4.1 Proposed Approach

The phases involved in the classification of emotions in the approach are:

(1) Preprocessing
(2) Extraction of Features
(3) Reduction of Feature Size
(4) Classification
(5) LBPh feature extraction
(6) Statistical moment and other feature calculation and appending
(7) Classification

3.4.1.1 Preprocessing

The facial part of the images is detected using Viola-Jones [7] from the images of CK+ dataset [5]. Only the first and last image are taken as they represent neutral and extreme expressions. Images are then cropped to contain only the face part. Some of the images in the dataset are colored. Therefore, to obtain uniformity, they are converted to grayscale. Example images are shown in Figure 3.8.

Figure 3.8: Cropped last frames samples

3.4.1.2 Extraction of Features

In this step Histogram of Oriented Gradients are extracted from the cropped face part of images as shown in Figure 3. The logic behind the histogram of oriented gradients descriptor is that object's local appearance and its shape can be described within an image by the edge direction distribution or gradient intensity.

The image is divided into cells which are small connected regions. For the pixels within these cells, the histogram of gradient directions is compiled. The cell size used in the experiments ranged from 10 by 10 pixels to 76 by 76 pixels. The feature vector is composed of HOG blocks where individual entries within a particular block are composed of cell histograms with orientation binning.

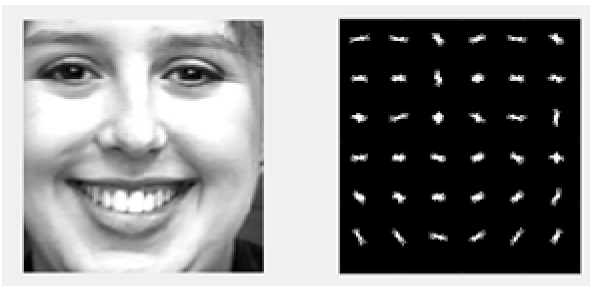

Figure 3.9: Cropped Image and its Histogram visualization for cell size 32

$$\text{Total Number of blocks} = \left(\text{floor}\left(\frac{\text{Len}}{\text{Cell}}\right) - 1\right) * \left(\text{floor}\left(\frac{\text{Bre}}{\text{Cell}}\right) - 1\right) \tag{3.3}$$

Total no of cell parts in feature vector computation = Total Number of blocks * 4 (3.4)

Total number of histograms per cell = 9 (3.5)

3.4.1.3 Reduction of Feature Size

The HOG feature extraction is done for cell sizes of 10 to 72. The resultant feature vector is bulky and therefore feature a reduction of the resultant HOGs of the images is done using principal component analysis (PCA) with variance coverage of 0.95.

3.4.1.4 Classification

The resultant vectors of the previous step are then subjected to classification. Different classifiers are used for this purpose which are linear discriminant, SVM (Linear, Quadratic and Cubic), Ensemble Subspace discriminant and Multilayer Perceptron. The results of these are shown in Table 1 to Table 4.

3.4.1.5 Local Binary Pattern Half (LBPh) feature extraction

Continuing with the experiment, local patterns are extracted for only the left half face part of the facial image (LBPh). This is done as it is assumed that the right part of the facial image would convey redundant information of the left part. The resultant image is shown in Figure 3.10.

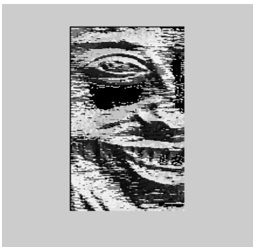

Figure 3.10: Local Pattern Image of the cropped left part of the face of S52

3.4.1.6 Statistical moment and other feature calculation and appending

For each local pattern image mean and standard deviation (first two statistical moments) are calculated. This provided μ_f, $\sigma_f$for the initial frame and μ_l,σ_l for the corresponding last frame. These two moments are used to calculate six other dependent features. These dependent features included $\mu_f - \mu_l$, $\sigma_f - \sigma_l$,μ_f/μ_l ,σ_f/σ_l ,$(\mu_f - \mu_l)/\mu_f$ and $(\sigma_f - \sigma_l)/\sigma_f$. This resulted in a feature vector ten elements. This is appended to the feature vector obtained in step (ii) i.e., by applying HOG.

3.4.1.7 Classification

The resultant vector is then subjected to PCA for feature reduction and afterward for classification. Classifiers used for this purpose are linear discriminant, SVM (Linear, Quadratic and Cubic) and Ensemble Subspace discriminant. The results of classification using (HOG + LBPh) at five-fold cross-validation are shown in following tables and figure.

3.4.2 Results

The results in terms of obtained accuracy at 5-fold cross-validation (HOG) are shown in the following tables.

Table 3.10: Results at various cell sizes

Classifier\Cell Size	**10**	**12**	**14**	**16**	**18**	**20**	**22**	**24**
Linear Discriminant	87.6	87.6	88.6	88.3	89.6	90.6	90.9	91.2
Linear SVM	59.6	61.9	67.4	69.1	70	75.2	79.8	81.8
Quadratic SVM	82.4	83.1	85	86.6	86.3	87.3	86.6	88.6
Cubic SVM	80.5	81.8	82.4	82.4	84	85.3	85.3	86.6
Ensemble Subspace Discriminant	89.6	89.3	89.9	88.9	89.3	90.9	90.6	92.8
Multilayer Perceptron	80.2	80.9	81.1	83.4	83.4	81.4	85.3	85.7

Table 3.11: Results at various cell sizes

Classifier\Cell Size	**26**	**28**	**30**	**32**	**34**	**36**	**38**	**40**
Linear Discriminant	87.6	87.6	88.6	88.3	89.6	90.6	90.9	91.2
Linear SVM	59.6	61.9	67.4	69.1	70	75.2	79.8	81.8
Quadratic SVM	82.4	83.1	85	86.6	86.3	87.3	86.6	88.6
Cubic SVM	80.5	81.8	82.4	82.4	84	85.3	85.3	86.6
Ensemble Subspace Discriminant	89.6	89.3	89.9	88.9	89.3	90.9	90.6	92.8
Multilayer Perceptron	80.2	80.9	81.1	83.4	83.4	81.4	85.3	85.7

Table 3.12: Results at various cell sizes

Classifier\Cell Size	**42**	**44**	**46**	**48**	**50**	**52**	**54**	**56**
Linear Discriminant	86.3	85.3	84.7	87	91.2	84	84.4	82.7
Linear SVM	85.7	83.7	84.4	85	87.6	82.7	84	82.7
Quadratic SVM	85	82.1	84	84.7	87.6	84.4	83.4	82.4
Cubic SVM	83.7	81.4	82.7	83.7	87.6	82.4	83.7	80.8
Ensemble Subspace Discriminant	87.6	86	85.3	86.6	89.6	84.7	85	85
Multilayer Perceptron	86.3	84.7	82.4	85.3	87	81.8	83.4	86

Table 3.13: Results at various cell sizes

Classifier\Cell Size	**58**	**60**	**62**	**64**	**66**	**68**	**70**	**72**
Linear Discriminant	84	88.3	86.6	85.3	84.4	85	70	70.4
Linear SVM	84	86.3	84.7	85	83.4	82.4	69.1	67.4
Quadratic SVM	86	86.3	85.3	84.7	80.5	82.4	67.1	67.1
Cubic SVM	83.4	84.4	83.1	83.7	78.8	79.5	66.4	66.8
Ensemble Subspace Discriminant	84.7	87.6	86.6	86.6	85	83.7	70.7	70.4
Multilayer Perceptron	87	84.4	83.7	84.4	81.1	81.8	65.1	65.1

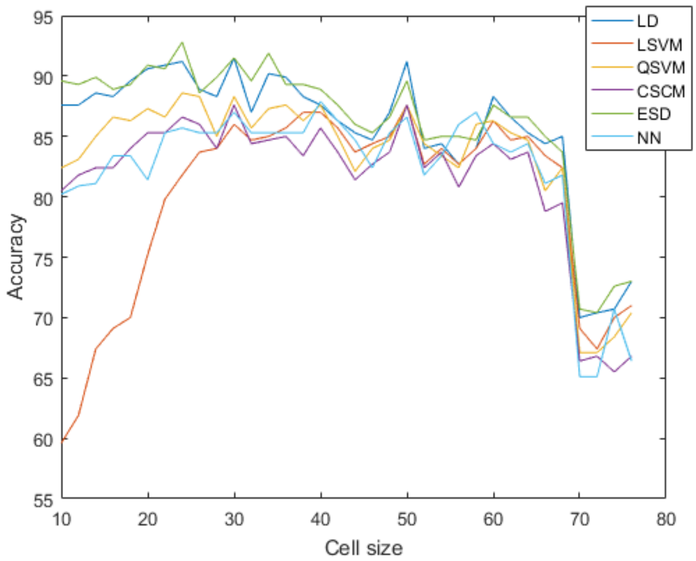

Figure 3.11: Graph of accuracy vs. cell size

By experimental results, it is found that the best results are obtained for cell size 30*30 to 50*50.

The results in terms of obtained accuracy at 5-fold cross-validation (HOG+LBP) are presented in Table 3.14.

Table 3.14: Accuracy of (HOG + LBP) at different Cell sizes

Classifier\Cell Size	**30**	**32**	**34**	**36**	**38**	**40**	**42**	**44**	**46**	**48**	**50**
Linear Discriminant	89.9	88.3	90.2	88.6	88.9	87.6	85.7	83.7	85	87	89.9
Linear SVM	86	85.7	86.6	85.3	85.3	87	85.3	84.4	83.7	86	88.6
Quadratic SVM	87.6	87.9	87	88.6	86	88.9	83.4	84.4	84	84.4	86.6
Cubic SVM	87.3	87.6	85.7	87	83.4	85.3	82.7	81.4	82.7	83.7	86.3
Ensemble Subspace Discriminant	89.9	88.6	91.2	88.9	88.9	89.6	87.6	86.6	86.6	88.3	91.9

This can be viewed graphically in figure 3.12.

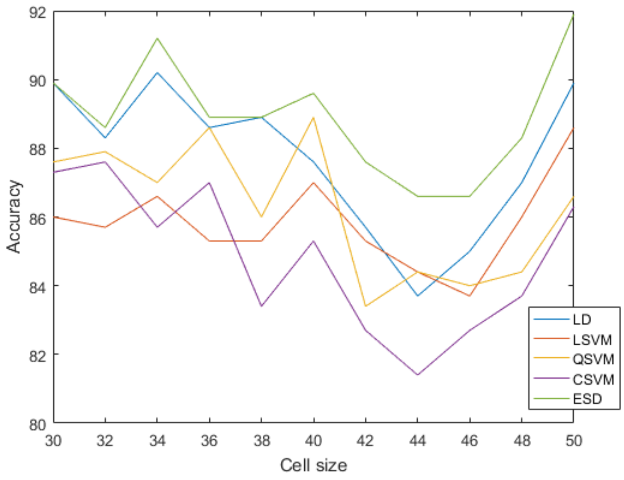

Figure 3.12: Accuracy (HOG+LBP) at various cell sizes

Comparison of performance on CK+ Database on different State of the art approaches is show in Table 3.15

Table 3.15: Quantitative analysis of method proposed with state of the art methods

Approach	Uddin [120]	Ahmad [152]	Zhong [153]	Song [149]	Zhang [154]	Our Approach
Average Accuracy	93.33	90.38	88.26	89.56	93.14	91.9

3.5 Conclusion

In this chapter, we validate the applicability of Action Units for determination of Emotions. The results show that K-NN outperforms all the other compared classifiers in terms of accurate classification of emotions.

This upholds true for classification across all split ratios of training and test sets and also for the 10-fold cross-validation. The effectiveness of using action units based on the movement of facial muscles for the detection of emotions in human beings is thus inferred based on the experimental data.

In this chapter, we also show that the LOG filters are effective in the classification of emotions. Binary classification provides fair accuracy for all emotion class pairs. Happy-Disgust class exhibits a higher percentage of accuracy for all split ratios and10 folds cross-validation. This binary pair has application in recommender systems and can be used to analyze customer reactions and recommend items accordingly. Happy-Sad binary classification can be used to detect depression, anxiety and other cognitive disorders by analyzing mood swings in the temporal domain. Fear-Surprise pair classification has elderly care homes as a potential application.

High accuracy is obtained for the multiple classifications. By experimental results, it is found that the best results are obtained for cell size 30*30 to 50*50 for HOG. This is attributed to window sizes that might actually capture features necessary for correct classification.

Adding LBP features to the vectors obtained increases accuracy in many cases. Quantitative analysis of the approach in comparison with others is done to show its applicability and effectiveness. This comparison is illustrious only and results of the approach would be even better as the other approaches have not considered Contempt as a class which has the least number of samples in CK+ dataset.

More extensive experiments can be conducted in the future which involve other models for the creation of hybrid vectors over different datasets that may further increase the accuracy of classification of emotions.

Chapter 4

Bharat Database of Indian Faces

4.1 Need of Databases

Intelligent emotion detection systems are required that respond in tandem to human feelings. Emotion detection capable computer systems is an ongoing research field that has numerous applications ranging from recommender systems, human-computer interaction, robotics and affective systems.

The keyboard and mouse approach for gathering log of user action is one of the simplest ways of predicting human emotions, but it lacks in accuracy. Also, keyboard pressure analysis for classification of emotions suffers from having a limited scope only as people are not constantly using keyboards and also all but lab equipment are not equipped with the necessary hardware to deliver the required information required for mapping them to emotional states.

On the other hand, there are intrusive approaches that are sensor based and physiological signals are gathered by way of ECG [86] and EEG [87] In between the two extremes are approaches that rely on facial feature analysis and gesture information. Over the past many years, much research has been done for developing approaches that may enable in automatic detection of emotions from expressions.

Critical and decisive information about the state of mind of a person may be acquired from the face which is an abstraction of the muscle movements leading to spatial and temporal changes. Also, this information has an indispensable relation to origins, age factor and gender. It is a communication capability which is disparate to verbal communication but yet may handout in understanding the emotional state in both cognizant and unaware conditions of expression.

Exhibition of emotions through gestures and facial expressions by humans is the most common aspect. The reverse of this, i.e., analyzing emotions from these expressions is a very common phenomenon for humans. They are capable of communicating among themselves through the exchange of these mutually dependent parameters.

Gestures may be controlled by people voluntarily and thus may not be coherent with the emotional state. Text analysis is generally a binary analysis of either a positive or negative state and therefore may provide information only about valence while emotional state comprises of not only valence but orthogonally arousal also.

Brain mapping is an important technique to predict emotional states based on surges of neurons of the human brain. This technique is more lab-based and not practical as the subjects under consideration are always aware of their examination and thus actual emotional state may be not be predicted. Also, this technique involves wearing monitoring caps which is obtrusive, limited to the lab environment and impractical in real life situations.

Breath analysis method for emotion detection is again an obtrusive technique and sufferers from the same limitations as of brain mapping technique. Humans are capable of reading gestures and facial features and understanding them near precision. But when it comes to computers, the research is novice and there is a great scope of improvement.

Face expressions provide an insight into the emotional state of human beings. Face expressions change in tandem with emotional states and people generally have limited or no control over their facial movements. Also if people might try to control their expressions in some specific temporal space but expressions are generally spontaneous and in tandem to their emotional state. Therefore emotion detection based on analysis of facial features is undoubtedly the best way to detect emotions as it is not limited to the lab environment, unobtrusive, coherent to the emotional state and practically applicable.

There are enormous real-life applications of such intelligent systems such as in artificial intelligence, sentiment analysis of people arising out of business and political decisions, affective computing, robotic assistance, mob controlling, driving assistance, personalized recommenders, depression, bipolar and anxiety detection, IoT applications, Crowd behavior analysis, surveillance systems, mob control systems, military applications, health care support systems, content delivery for social media, movie target audience and tailored support solutions for differently abled persons.

Detection of emotions from facial expressions involves extracting mathematical information from spatial variations of the faces. This spatial information also invariably comprises ethnicity, gender, scale correction as well as noise features involving occlusion sub mounting to beard, mustache, hair locks, face rotation, resolution and physical hindrances towards full face detection.

Development of systems capable of emotion classification invariably requires training to catch the strains that effectively and categorically map them to different categories. This, in turn, requires data for the systems to train, which has the following two aspects.

Either the training data can be too personalized to serve for the only particular subject involved. For example, Personal response systems based on the specific requirement of individuals. This involves catching individual data and catering tailor-made solutions based on them which cannot be applied generically.

In literature, Paul Ekman [5] reported six basic emotions that are valid across human species irrespective of gender and ethnicity. These are happy, sad, angry, fear, surprise and disgust. Humans can predict the emotional state of people of different ethnicity. But such is not the case with computers as they need training for various texture, shape and appearance of the target subjects.

Because of the different shape and texture of people across the globe, headway in the field of affective computing requires different databases covering ethnicity. To our knowledge, there is no posed expression database for Indian faces that catches all the above-mentioned emotion classes.

For capturing and building of affective systems, it is a prerequisite to obtaining these spatiotemporal displacements happening in facial features due to underlying muscles. The spatial and temporal assessment of prominent facial features may be utilized for categorization of emotions. Many facial emotion techniques have been proposed which have considered 2D images, 3D images [39], [40], expressions exhibited by infants, AAM [41] based systems and AU based systems [42].

Non-deliberate and deliberate ace expressions are the two categories of face expressions defined by Battocchi et al. [43], Expressions which are deliberate are expressed under the absence of speech. Whereas, those expressions exhibited along with speech are termed as non-deliberate. Also Valence, Arousal and Dominance multidimensional space is sometimes used for separating the emotions categorically.

This requires creating a database that is essential to train the systems, test them, validating of applicability of algorithms for emotion recognition and classification systems for building robust systems. Database creation is a toilsome and tedious task. But the less is the fact that it is a prerequisite to building systems that are capable of detecting emotions. Creation of a posed expression database requires proper guidance and training of the subjects by experts. After that, validation of database is an equally important concern. This involves labeling mages with the categories that they belong to.

4.2 Existing Databases of Facial Expressions

Detection of emotions based upon the facial expressions is important for the building of intelligent systems that may respond in tandem to human feelings. The various features that increase the complexity of emotion recognition systems include ethnicity, gender, pose, occlusion, beard, mustache etc. Therefore there had been an emergence of several databases covering various features that are available publicly emotion recognition. The different databases are discussed in brief in this section.

One of the most widely used databases is Cohn-Kanade Action Unit Coded Facial Expression database. It comprises of 486 image sequences posed by 97 subjects is released in the year 2000. Image sequence proceeds from neutral face image to extreme expression. The peak expression images are coded using the Facial Action Coding System and annotation is provided in the form of emotion labels. It had the limitation as the emotion labels are not validated.

This dataset is extended to address as Extended Cohn- Kanade (CK+) database [6]. The images of 123 subjects corresponding to 593 sequences are taken which are coded using Facial Action Coding System for the last or peak frame of the sequence. The Action Units and their intensity are provided for the peak expression images. Also, the images are tracked using Active Appearance Model for 68 landmark points. Out of the total sequences, only 327 are having corresponding emotion files. This is because only these sequences are validated. The emotion labels are neutral, anger, contempt, disgust, fear, happy, sadness and surprise.

Japanese Female Facial Expression (JAFFE) database has 213 images ten female Japanese models that posed for both neutral and six basic expressions. JAFFE database is created by Lyons et al. [8]. The images are in grayscale. Binghamton University-3D Dynamic Facial Expression database [9] has 2500 3D face expressions of 100 subjects having the six basic expressions having four intensity levels. The different aspects considered include age, race and culture.

The MMI Database is initially conceived in 2002 to serve as a source that can be used across facial expression recognition community [10]. It has videos that have a sequence from neutral to apex and back to neutral expression. It has over 2900 videos of 75 subjects as well as still images. The videos are annotated for presence of Action Units.

The Belfast Database [11] has different sets of over 250 colored video clips depicting natural emotions at different resolutions. Multimedia Understanding Group (MUG) has 1462 posed color sequences of 86 subjects which are annotated with emotion labels. The Radboud Faces Database (RaFD) has posed color images of 67 subjects at five different camera angles and three different gaze directions for eight emotion labels.

Indian Spontaneous Expression Database (ISED) [12] has 428 spontaneous color videos of 50 subjects having emotion labels of sad, happy, surprise and disgust only. Denver intensity of spontaneous facial action database (DIFSA) [13] is a color video database of 27 subjects whereby each video sequence is of 4845 frames of spontaneous reactions while viewing a video of 4-minute duration. Six intensity level annotations of Action Units are provided for the facial expressions.

The various features that increase the complexity of emotion recognition systems include ethnicity, gender, pose, occlusion, beard, mustache etc. The type of database used for learning by systems is of crucial importance. Many databases exist for this purpose, but none of them is for posed Indian faces.

We bridge this gap by providing Bharat Database which contains facial images of Indian people. The database has posed expressions of 102 participants and has 896 images. The participants are asked to pose for different emotions by showing them images eliciting those emotions as well as with the help of expert artists. The annotation is done using polling by a panel of three experts. This database will further help the community involved in developing algorithms for emotion recognition.

4.3 Creation of the Dataset

The volunteers that participated in the construction of this dataset are well informed of the purpose of its construction. They are asked to provide consent for publication for Research purposes and accordingly signed for possible usage for the same.

The BDIF has still posed facial images for various emotions. Subjects are asked to pose for all the basic as well as neutral emotion. The participants are asked to pose for different emotions by showing them images eliciting those emotions as well as with the help of expert artists. The photographs are taken in well-lit conditions.

Expression of emotions is generally involuntary through facial expressions and taming it to bring the desired effects requires effective and affective consultation. This aims to exalt the legitimacy of the cognizant expressions to requisite levels.

Depicters/ subjects need to portray emotion instances based on:

i. Presentation of key features of physiology and its imitation for the attainment of the desired features in the presence of domain experts.

ii. Any further admonish to attain the prerequisites.

The BDIF is carefully constructed by showing the subjects valid labeled emotion images from different databases and also with the help of expert artists training them how to elicit the said emotions. The effectiveness in expressing emotion is done by changing the mental state of the subjects by showing them visual cues and also narrating real life situations which pertain to those particular emotions.

The annotation of images is done using polling by three annotators who are familiar with Facial action coding system. They are shown images and told to classify them as belonging to one of the classes. Only those images are labeled in which a consensus arrived among the annotators.

4.3.1 Experimental Setup

The subjects participated voluntarily for the posing of expressions. They are made comfortable by telling them how to pose and showing visual cues as well as real-life situations which are associated with the particular emotional state. The subject is made comfortable with the environment and counseled led to the capture of expressions effectively.

These images are captured in ideal conditions which are free from noise and other distractions. The images are captured in ambient light conditions. According to the experience obtained from preliminary studies, the light conditions are as subjects are generally accustomed to and not being very bright or dull. These made the subjects comfortable with the experiment environment.

Closed rooms are used for the conduction of shoot sessions. The subjects are made well aware of the experiment and also as how it will help in future scientific studies. The subjects are asked to stand comfortably taking the support of the wall. They are allowed time to ease them for elicitation of emotions. After that images are captured by posing for neutral, anger, contempt, disgust, happy, fear, sadness and surprise emotions.

To avoid disturbance, the rooms are kept closed during the different shooting sessions. In the first part of the experiment, it is found after annotation that the sample of anger and disgust emotion is the least. The statistics of the first phase is shown in Figure 4.1.

The images are taken comprising of 4320 × 3240 pixels at 300 dpi horizontal and vertical resolution using Nikon Coolpix 120. The images are taken with compulsorily no flash. Distance maintained between the subject and the camera is around 1 meter. The emotion label categories for the complete database are shown in Figure 4.2.

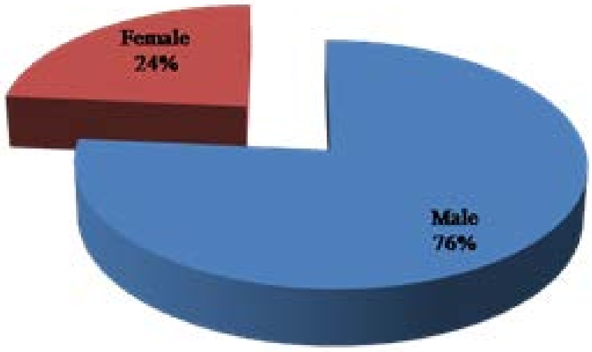

Figure 4.1: Male Female Ratio.

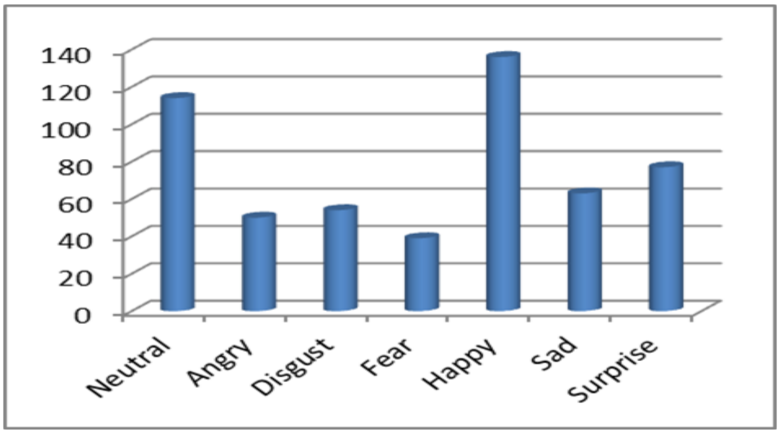

Figure 4.2: Graph showing the emotion label categories for the complete database.

4.3.2 Obstructions/ Occlusions

Gathering information from facial expressions may be hindered by the presence of glasses, mustache, beard etc. These act as noisy features in the implementation of automatic emotion detection algorithms. However, they are a regular feature in real life situations. Therefore these features are deliberately included in the formation of the database so that more robust emotion detection algorithms can be built using such images also from the database.

4.3.3 Annotation

The images obtained for the database are subjected to labeling for different emotions. This is particularly important as it serves as ground truth for comparison with results obtained by applying automated algorithms. Therefore the technique of validating emotion labels is of substantial significance. The images are labeled by a panel of three annotators who are familiar with the facial action coding system. The labeling is done for neutral, happy, angry, disgust, contempt, surprise, sad, fear and invalid emotions.

For validation, polling is done for each image by gathering the emotion labels of all the three annotators. Only those images which had a majority for particular emotion label are validated for that particular emotion. Rest of the images are not annotation validated for the particular emotion labels. The resultant database consisted of:-

The total number of subjects posing in the first phase is 59.
The total number of images collected in this phase is 436.
The Total Classes of Valid Emotions are 7.

Table 4.1: Emotion Samples from the first Phase

S.No.	Emotion Category	No of samples
1	Neutral	71
2	Angry	24
3	Disgust	14
4	Fear	3
5	Happy	91
6	Sad	40
7	Surprise	47

The results of the first phase of database creation led to a finding that even when the subjects are made comfortable with the environment, they could not provide fear and disgusting motions. This is because people do not feel comfortable to exhibit these emotions willingly and publically.

Therefore in the next phase, the subjects are given privacy and shoots are performed in confined places so that they can comfortably exhibit those emotions. Also, subjects are provided visual cues as what that expression looks like. Assistance is provided to them with the help of a professional artist to help them exhibit those emotions.

The database is extended in the next phase by including new subjects and taking care of findings of the first phase.

The extension led to:
The total number of subjects posing in the first phase is 43.
The total number of images collected in this phase is 460.
The Total Classes of Valid Emotions are 7.

Table 4.2: Emotion Samples from Second Phase

S.No.	Emotion Category	No of samples
1	Neutral	43
2	Angry	23
3	Disgust	39
4	Fear	36
5	Happy	41
6	Sad	23
7	Surprise	36

4.3.4 Consent from Subjects for Publication

The volunteers that participated in the construction of this dataset are well informed of its purpose. They are asked to provide consent for publication for Research purposes and accordingly signed for possible usage for the same. Initially, the subjects are reluctant to pose for the Database. The subjects involved in the collection of the dataset are given incentives like chocolate, juice and ice cream. The participants are informed that their images may be used for publication for research purpose. Ethically the subjects are asked to fill up their consent for publication of images for research purpose. Images of those subjects who did not provide consent are deleted from the final database.

4.3.5 Complete Bharat Database of Indian Faces

BDIF initially had expression images of 59 subjects for seven classes of emotions viz. neutral, happy, angry, fearsome, disgust, surprised and sadness. It initially had 436 images. After that, it is extended by the addition of expression images of 43 new subjects exhibiting the seven classes of emotions. Four hundred sixty new images are added in this phase.

The complete database thus has total of 102 subjects and a total of 896 images. This database has images of school going children of 11th and 12th class, subjects doing graduation, post-graduation and research scholars, staff members from different offices and random people. There are occlusions of beard, mustache and spectacles. The class-wise distribution is shown in table 4.3.

Table 4.3: Emotion Samples of Complete BDIF database

S.No.	Emotion Category	No of samples
1	Neutral	114
2	Angry	50
3	Disgust	54
4	Fear	39
5	Happy	136
6	Sad	63
7	Surprise	76

Total Male: 78

Total female: 24

Sample cropped images of the different classes of emotions can be seen in Figures 4.3 – 4. 9.

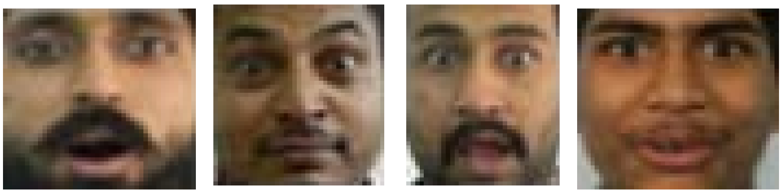

Figure 4.3: Images from Bharat Database of Indian Faces exhibiting Surprise emotion.

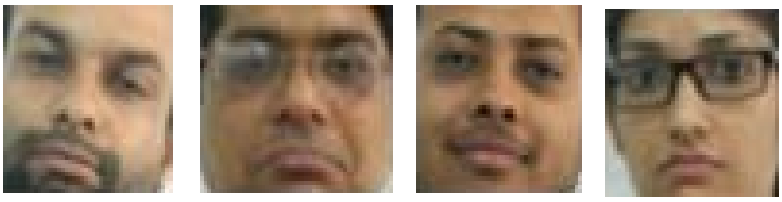

Figure 4.4: Images from Bharat Database of Indian Faces exhibiting Sad emotion.

Figure 4.5: Images from Bharat Database of Indian Faces exhibiting Neutral face.

Figure 4.6: Images from Bharat Database of Indian Faces exhibiting Anger emotion.

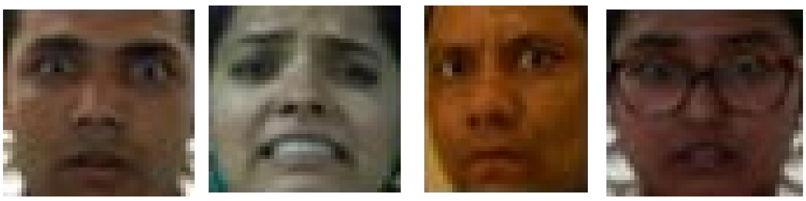

Figure 4.7: Images from Bharat Database of Indian Faces exhibiting Fear emotion.

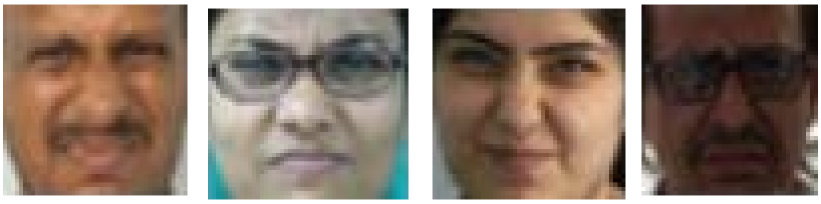

Figure 4.8: Images from Bharat Database of Indian Faces exhibiting Disgust emotion.

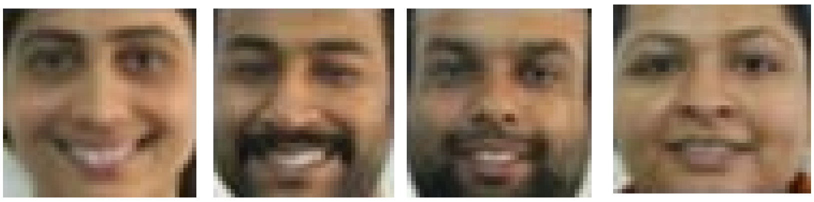

Figure 4.9: Images from Bharat Database of Indian Faces exhibiting Happy emotion.

4.4 Results and Discussion

Detection of emotions based upon the facial expressions is important for the building of intelligent systems that may respond in tandem to human feelings. The various features that increase the complexity of emotion recognition systems include ethnicity, gender, pose, occlusion, beard, mustache etc.

The type of database used for learning by systems is of crucial importance. Many databases exist for this purpose, but none of them is for posed Indian faces. We bridge this gap by providing Bharat Database which contains facial images of Indian people. The wide variety of subjects and their emotion labeling may help researchers in developing robust algorithms for futuristic artificially intelligent systems. Several evaluations of accuracy are done to behave as a baseline by researchers to develop more robust algorithms.

Gathering information from facial expressions may be hindered by the presence of glasses, mustache, beard etc. These act as noisy features in the implementation of automatic emotion detection algorithms. However, they are a regular feature in real-life situations. Therefore these features are deliberately included in the formation of the database. To show its effectiveness and applicability quantitative analysis of accuracy is done. The results show that it is robust to gender, occlusions and ethnicity.

4.5 Acknowledgment

We are grateful to all participants who have provided their time and also helped by ascertaining the advice provided for the creation of this database which is a herculean task. Also photography expert and professional artist deserve great thanks for making this project successful. Last but special thanks to the annotators who provided their time to validate the emotion labels associated with the database.

4.6 Compact Local Binary Pattern Technique and Multiple Instances Based Emotion Detection Using Discriminant Feature Tracking

Two techniques are proposed for the detection of emotions from facial images. Compact Local Binary Pattern approach is the first one which is used in constructing a hybrid feature vector and results show that accuracy of detection of emotion increases using this approach.

In the second approach, Multiple Instances Based Emotion Detection Using Discriminant Feature Tracking is proposed. This approach is applied to Bharat Database of Indian Faces (BDIF) which is indigenously developed, Karolinska Directed Emotional Faces (KDEF) and Japanese Female Facial Expression (JAFFE) Database. This approach is effective for detection of emotion which may be inferred from results for accuracy of classification.

The main contributions include:

1) Compact Local Binary Pattern (CLBP) method for feature extraction is introduced
2) The resultant vector of CLBP is integrated with the vector obtained using Histogram of Oriented Gradients technique
3) Multiple Instances Based Emotion Detection Using Discriminant Feature Tracking technique is proposed

4) Experiments are done on three databases of facial expressions including indigenously developed BDIF database, KDEF and JAFFE

Approaches Stepwise

The following steps are involved in the detection of emotions from facial images of the databases:

1) Preprocessing of images of the database
2) Feature Vector Extraction using Histogram of Oriented Gradients
3) Reduction of the size of the feature vector obtained
4) Classification of Emotions
5) Feature Vector extraction using the proposed Compact Local Binary Pattern
6) Classification

4.6.1 Image Preprocessing

The detection of part exhibiting face is done on images from Bharat Database of Indian Faces (BDIF) using the algorithm of Viola Jones. Permissible height and width are set to a minimum of 600 pixels. Face detection accuracy of 100% is obtained using these parameters on the database. Further cropping is done on the images to remove other noisy parts which may lead to erroneous results. For experiments, Database images are converted from colored to grayscale and to obtain unanimity, they are resized to the pixel size of 1024*1024. The figure shows samples of the images cropped.

4.6.2 Extraction of Features

The images cropped and grayscaled having corresponding emotional reference are subjected to the derivation of feature vector using Histogram of Oriented Gradients technique [38]. This method catches gradient intensity or edge directions distributions which enable catching local appearance and also the shape of the objects. The cells are divided into pixels of uniform breadth and length that range from the size of 10*10 to 512*512.

After that computation of directional gradient, histograms are done for such cells having an overlap of 50% for developing feature vector that compromises of histograms that are binned according to orientation. Sample visualization of the histogram is presented in the following figure.

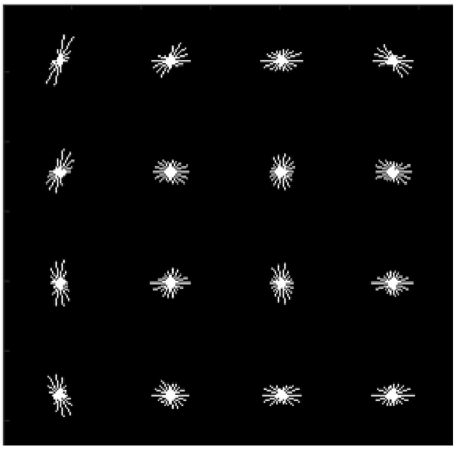

Figure 4.10: Sample Histogram of the image at the cell size of 256*256

$$T_{Blocks} = floor\left(\frac{Len}{Cell} - 1\right) \mathrm{X}\, floor\left(\frac{Bre}{Cell} - 1\right) \tag{4.1}$$

Where T_{Blocks} denotes the number of blocks in total for an image. As each block has $2 * 2$ cells, thereby:

TCells = TBlocks * 4 (4.2)

Where TCells represents the total number of cells that participate for the computation. The total number of histograms per cell used in this work is given in Equation 3.

HistCell = 9 (4.3)

Figure 4.11: Grayscale Cropped Image samples.

4.6.3 Reduction of Feature Size

Histogram features for gradient orientations [10] are extracted for sizes of cells ranging from 10*10 pixels to 512*512 pixels. Experiments are conducted on this feature vector and reduced feature vector which is done by applying Principal Component Analysis [11] covering variance of 0.95.

4.6.4 Compact Local Binary Pattern approach

Compact Local Binary Pattern technique is proposed which results in the reduction of image size by around nine times as it uses only concerned pixels derived from neighbor pixel values while discarding the neighbors. The following figures depict exemplary conversion based on neighbor pixel values in any image. The exact results after applying Compact Local Binary Pattern are shown in figure after that.

115	121	214	56	78	98
87	98	76	211	58	67
123	124	178	189	111	48
76	78	67	90	99	111
99	100	102	104	88	76
244	235	68	89	96	89

Figure 4.12: Initial image sample pixels.

1	1	1	0	1	1
0	98	0	1	58	1
1	1	1	1	1	0
0	0	0	1	1	1
0	100	1	1	88	0
1	1	0	1	1	1

Figure 4.13: Pixels and their values are depicted in red color after application of Local Binary pattern on the immediate neighbors.

0	0	0	0	0	0
0	1101110	0	0	01110111	0
0	0	0	0	0	0
0	0	0	0	0	0
0	00010110	0	0	11101111	0
0	0	0	0	0	0

Figure 4.14: The concerning values to be used for further experimentation is shown in Red.

206	119
22	239

Figure 4.15: Compact LBP

4.6.5 Feature vector extraction using statistical moments

The CLBP images are used for calculating the median and mean of each row of pixels. Length of feature vector obtained is 682 for both 341 columns and rows. Appending of a total of 51 histograms is done for all images. Standard deviation and the mean of the image are also added to the vector.

Length of the feature vector obtained is thus 735. Different classifiers are then used for obtaining resultant accuracy of classification. For the next phase of the experiment, these feature vectors are appended to the histograms of oriented gradient vector obtained previously to obtain further classification results.

4.7 Marking of points on Neutral faces Approach

Marking of points on the neutral faces of the subjects is done in the next experiment. The neutral face in the images is subjected to marking of tracking points manually. These points covered the eyebrows, eyes, nose and lips of the subjects. The reason for this allotment is that these parts of the faces convey the most information. In literature associated with physiology, these parts contemplate to Action Units. Sample image with point marked is shown in Figure 4.17.

4.7.1 Tracking of the points in corresponding images

The points that are marked on neutral images of the previous phase are then averaged for their color values in this phase. Different matrices are considered of which grid size of 5×5 pixels provided the best results. The average value is then calculated using values of color channels of these pixel points.

$$R[\,] = R(r - IR + 3, c - IC + 3) \tag{4.4}$$

$$G[\,] = G(r - Ir + 3, c - Ic + 3) \tag{4.5}$$

$$B[\,] = B(r - Ir + 3, c - Ic + 3) \tag{4.6}$$

Where $r = IR - 2$ to $IR + 2$ and $c = IC - 2$ to $IC + 2$

Where IR and IC correspond to initial Row and Column. Average Color Channel values are calculated for each color.

$$AvgR = AverageR[\,] \tag{4.7}$$

Tracing of these points is done using Color Channel average values for analogous images depicting the emotion of the corresponding subjects. Distance is varied from 10 to 25 pixels for data collection. Best results are procured at a permissible maximum distance of 15 pixels in either direction. Probable points are searched within the periphery of initial points of neutral and after that emotion images, i.e., for each pixel in the edge of 15 × 15.

$$AvgR1 = \sum_{i+2;j=2}^{i-2;j-2} \frac{R1}{25} \tag{4.8}$$

4.7.2 Vector formation based on distances

Calculation of distance between tracked points and initial positions is done using Euclidean formula for distance measurement. As the displacements imbibe useful information that may be utilized for actually relevant information that may be processed for the revelation of emotions, it is calculated. The vector of displacement is composed of the following:

$$D = \sqrt{(X_i - X_j)^2 - (Y_i - Y_j)^2} \tag{4.9}$$

Algorithm Mark Point Technique

Ensure: A − Emotion Accuracy

Require: I_i -Image i = 1......T ∗ t

1: For each image I_i do

2: Take a neutral emotion image of the subject.

3: Crop and resize

4: Mark the points(N)

5: Read coordinates of all the points

6: Get R, G and B channel values in the window of −2 to 2 x and y values around

7: Calculate Average R, G and B values for each window around N points

8: End For

9: For i = 1 to T

10: For j = 1 to all images in i folder do

11: Find face, Crop and Resize

12: Get R, G and B channel values in the window of −9 to +9 x and y values around

13: End For, End For

14: For each N (x, y) in emotion image and in the periphery of 3 to 17 pixels do

15: Calculate Average R1, G1 and B1 for widow size −2 to +2 in x and y-direction

16: Slide the widow

17: End For

18: For each Value obtained do

19: Calculate the Difference=Average R - Average R1

20: Repeat for other two color channels

21: End For

22: Calculate the point having a minimum difference out of values obtained.

23: Obtain coordinates of the pixel whose window has minimum Difference.

24: Calculate the signed difference in coordinates of Neutral Image and Emotion Image for all N points

25: Classification

The number of points marked on the neutral images are 14 to 32 at a gap of 2 points for each subject of three databases namely BDIF (Bharat Database of Indian Faces), JAFFE (The Japanese Female Facial Expression Database) [20] and KDEF (Karolinska Directed Emotional Faces) [44].

Calculation of the points having a minimum difference out of values obtained is done. After that coordinates of the pixel whose window had minimum difference are obtained. Also signed difference in coordinates of Neutral Image and Emotion Image for all N points is calculated in this step.

JAFFE has posed seven facial expressions of 10 Japanese female models and has a total of 213 images. KDEF has images of 70 amateur actors comprising of 35 males and 35 females. The total number of images in the database is 4900. The subject images have no occlusions in the form of earrings, eyeglasses, mustache, beard and also have no visible make-up. It has images taken from five different angles. However, only frontal images are used in the experiments.

The points are marked to be in conformity to action units of Facial Action Coding System as described by Paul Ekman. These points covered eyebrows, eyes, nose and mouth region of the images. The tracking of these points is done in all the emotion images of those subjects.

The accuracy of emotion detection is then calculated:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{4.10}$$

Where TP is True positive, TN is true negative, FP is False Positive and FN is False Negative.

Figure 4.16: BDIF images in Compact LBP form

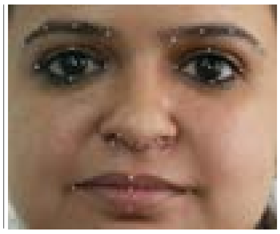

Figure 4.17: Image with points marked.

Support Vector Machines are used in machine learning as supervised learners with algorithms for data analysis for classification. Training examples are categorized into different categories based on the classification labels. SVM builds a model based on the training examples as falling into different categories. After building the model, new examples are categorized for different classes. It is a non-probabilistic classifier that represents examples as points in space. Mapping is done to categorize the examples of the different categories so that there is a clear distinction between them. After model building, the test examples are mapped into the same space. After that, the prediction is done about their class based on which side of space they come into.

4.8 Results and Discussion

The vector resultant of HOG is applied to the analysis of principal components and after that classified. Accuracy percentage is obtained for the vector obtained using the approach of Compact Local binary patterns. It carries information of the texture only and while meeting out it's matching with labeled emotions for accuracy, therefore provides lower results of accuracy.

Table 4.4 and 4.5 show results obtained after five-fold cross-validation of the HOG vector. In the next phase, the vectors obtained by both HOG and CLBP are appended and verified for accuracy for five-fold cross-validation. The results are depicted in Table and Figure.

Table 4.4: Accuracy at cell sizes ranging from 64 to 104

Classifier | Cell Size	64	72	80	88	96	104
SVM	66.2	65.5	64.9	66.2	67	64
Linear SVM	59.6	61.9	67.4	69.1	70	75.2
Quadratic SVM	82.4	83.1	85	86.6	86.3	87.3
Cubic SVM	80.5	81.8	82.4	82.4	84	85.3
Ensemble Subspace Discriminant	89.6	89.3	89.9	88.9	89.3	90.9
Multilayer Perceptron	80.2	80.9	81.1	83.4	83.4	81.4

Table 4.5: Accuracy at cell sizes ranging from 112 to 512

Classifier| Cell Size	112	120	128	256	512
SVM	64.7	62.8	61.9	50.8	39.9
Linear SVM	83.7	84	86	84.7	85
Quadratic SVM	88.3	85	88.3	85.7	87.3
Cubic SVM	86	84	87.6	84.4	84.7
Ensemble Subspace Discriminant	88.6	89.9	91.5	89.6	91.9
Multilayer Perceptron	85.3	85.3	87	85.3	85.3

As the Quadratic SVM provides the best results, therefore it is used for the classification using a hybrid vector comprising both Compact LBP and HOG which is depicted in Figure 4.18.

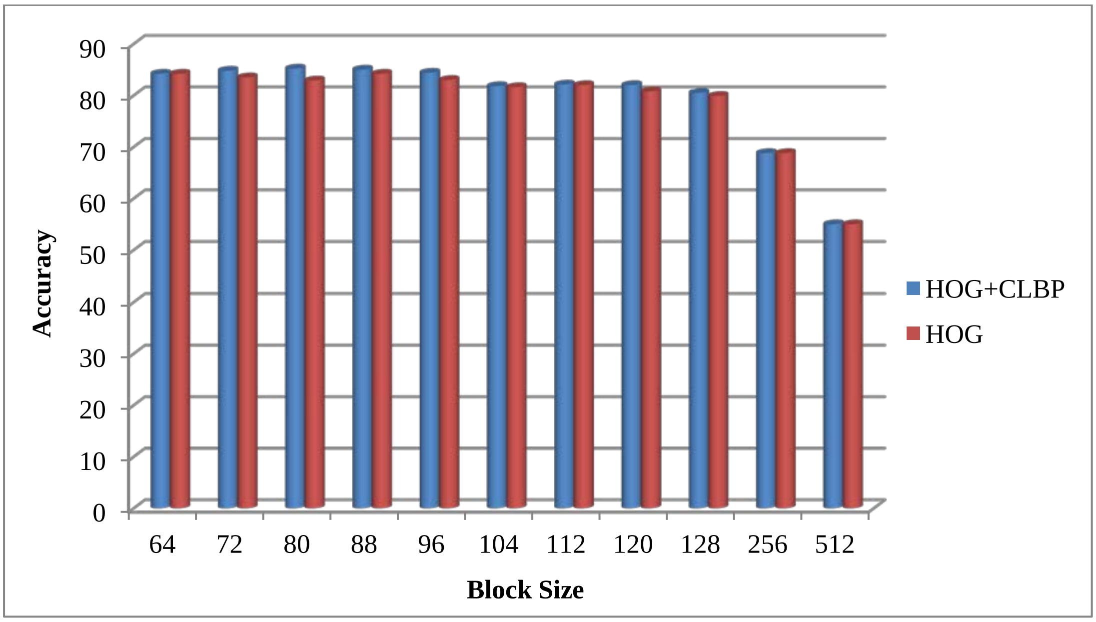

Figure 4.18: Comparison of the result achieved using HOG and HOG+CLBP.

Multilayer Perceptrons maps input data to different output categories based on feedforward artificial neural network. It has directed graph of nodes in multiple layers; the nodes of one layer are connected fully to the nodes next layer. Except for the input nodes, each node is a neuron (or processing element) with a nonlinear activation function. It has one or more hidden layers that have nonlinear activating nodes.

Navie Bayes provides the least percentage of accuracy of the three classifiers used in the experiment. This is particularly true because it is probabilistic classifier with the assumption of independence among features. But in the feature vector used in the experiments, there is a correlation among the various features associated with the points marked. SVM provides a high degree of accuracy because of its ability to construct hyperplane between the classes. Multilayer Perceptron also provides high accuracy because of a limited number of input nodes, a maximum of 34, mapped to total seven output nodes. The results show that SVM and Multilayer perceptron is best suited for this approach.

Classification is done using Neural Network (NN), Naive Bayes (NB) and Support Vector Machine (SVM). Naive Bayes is probabilistic classifiers based on applying Bayes' theorem with the assumption of independence between features. It is used for classification of examples for categorization.

The following table shows result accuracy in correspondence to the considered number of points concerning the three databases experimented upon.

Table 4.6: Emotion detection accuracy in terms of percentage (%)

Number of Points

Database	Classifier	14	16	18	20	22	24	26	28	30	32
BDIF	Naive bayes	52.1	54.1	50	55.1	56	58.6	61.3	62.8	65.2	66
	SVM	78.4	84	87.4	89.3	89.8	91.9	91.9	93	93.4	94.5
	Multilayer Perceptron	84.6	88	90.4	91.2	92.1	92.9	92.9	91.9	92.1	93
JAFFE	Naive bayes	45.2	46	47.2	47.2	47.6	48.1	54.1	56.9	59	61.8
	SVM	75.3	80.7	84.2	86	86.4	88.2	87.7	90.1	90.3	92.1
	Multilayer Perceptron	81.1	84.7	88	88.2	89.3	89.5	89.3	88.2	87.9	89.5
KDEF	Naive bayes	53.2	54.7	53	56.1	57.4	60.1	62.4	64.6	66.9	67.6
	SVM	79.1	85.3	88.9	91	91.4	93.1	93.1	94.3	94.6	95.8
	Multilayer Perceptron	85.8	89.2	92	93.1	93.5	94.1	94.3	93.1	93.5	94.5

Examining the result shows that the percentage of accuracy incurred across databases is least for JAFFE followed by Bharat Database of Indian faces or BDIF. KDEF stands on top of the other two databases.

Gathering information from facial expressions may be hindered by the presence of glasses, mustache, beard etc. These act as noisy features in the implementation of automatic emotion detection algorithms. However, they are a regular feature in real-life situations. Therefore these features are deliberately included in the formation of the database. To show its effectiveness and applicability quantitative analysis of accuracy is done. The results show that it is robust to gender, occlusions and ethnicity.

Karolinska Directed Emotional Faces (KDEF) does not have images that have occlusions like beard, mustache, eyeglasses, no make-up and earrings. The presence of such occlusions, which are common, makes BDIF stand second. JAFFE has low-resolution images which make it last in the league. Distance vector calculation in the periphery of point originally located having minimal intensity difference is therefore not that accurate for this database resulting into lesser accuracy of emotion classification.

4.9 Summary

Detection of emotions based upon the facial expressions is important for the building of intelligent systems that may respond in tandem to human feelings. The various features that increase the complexity of emotion recognition systems include ethnicity, gender, pose, occlusion, beard, mustache etc.

The type of database used for learning by systems is of crucial importance. Many databases exist for this purpose, but none of them is for posed Indian faces. We bridge this gap by providing Bharat Database which contains facial images of Indian people. The wide variety of subjects and their emotion labeling may help researchers in developing robust algorithms for futuristic artificially intelligent systems. Several evaluations of accuracy are done to behave as a baseline by researchers to develop more robust algorithms.

Accurate classification of more than 90% is achieved using the point tracking technique for emotions. The results show that initial points ranging from 14 to 34 are acceptable for characterizing of emotions. The success is because of these peculiar points attributing to forming of a captivating feature vector decisive for categorization.

Chapter 5

Conclusion and Future Work

Facial expressions are a global language revealing elemental emotions. The importance of emotions is evident in our everyday chores; the routine perspective and working of people rely upon the emotional states and vice versa. Literature backs that emotional states affect facial features. The capability of interpreting emotions based on expressions has been an intensive field of study for more than three decades for developing intelligent systems. Machines that can intelligently identify emotions based on human facial expressions are gaining social importance. With the increased capability of processing and reduction in the size of the devices, this intelligence can now even be incorporated to even small handheld devices such as smartphones.

Intelligent emotion detection systems are required that respond in tandem to human feelings. There are enormous real-life applications of such systems such as sentiment analysis of people arising out of business and political decisions, affective computing, robotic assistance, mob controlling, driving assistance, personalized recommenders, depression, bipolar and anxiety detection, IoT applications, Crowd behavior analysis and many more.

In this chapter, we present the conclusions drawn from our research work and the possible directions for future work.

5.1 Conclusions

The aim of the thesis was to study the existing techniques of emotion detection from facial images and to develop techniques that lead to an improvement in emotion detection accuracy. Also to study the existing databases of facial images and to develop a dataset that incorporates Indian facial images and experiment upon it for emotion detection accuracy.

The conclusions drawn from the thesis are as follows:

1. The association of facial features with emotions is a widely accepted fact as people respond in tandem to feelings like anger, joy and surprise. We Validate that facial action units and expressions can be used for classification of emotions using various machine learning algorithms.
2. We use Laplacian of Gaussian filter approach for building feature vector from facial images which are immune to noise. The results show a high degree of accuracy for almost all pairs. This shows the effectiveness of using this technique for detecting emotions.
3. We propose a new technique of Local Binary Pattern Half used in conjunction with Histogram of oriented gradients. By experimental results, it is found that the best results are obtained for cell size 30*30 to 50*50 for HOG. This is attributed to window sizes that might capture features necessary for correct classification for CK+ database. Adding LBP features to the vectors obtained increases accuracy in many cases because of additional texture information.
4. We build a Bharat Database which contains facial images of Indian people. The database has posed expressions of 102 participants and has 896 images. Gathering information from facial expressions may be hindered by the presence of glasses, mustache, beard etc. These act as noisy features in the implementation of automatic emotion detection algorithms. However, they are a regular feature in real life situations. Therefore these features are deliberately included in the formation of the database. We also annotate it for emotion classification.

5. We propose a Compact Local Binary Pattern Technique for building hybrid feature vector and show its usefulness in providing better accuracy when used in conjunction with Histogram of Oriented Gradients.
6. We propose Multiple Instances Based Emotion Detection Using Discriminant Feature Tracking Technique. To show its effectiveness and applicability quantitative analysis of accuracy is done using various machine learning approaches on various databases. The results show that it is robust to gender, occlusions and ethnicity.

5.2 Future Work

Further developments in the field of emotion recognition are possible. This research field is attractive for the industrialists and researchers due to extensive real-life applications. We would like to extend the following work in the future.

1. More extensive experiments can be conducted in the future which involve other models for the creation of hybrid vectors over different datasets that may further increase the accuracy of classification of emotions.
2. New techniques may be explored to modify the proposed algorithms to be even more generic and robust for emotion recognition.
3. The database developed may be augmented by annotated facial images of more subjects pan India and also may have videos comprising voluntary and spontaneous emotions.

Publications

1. Abhishek Singh Kilak and Namita Mittal, "I know when you are happy - Emotion Detection", 6th International Conference on Information Science and Application (ICISA-2015) Springer Publication, Thailand, pp 769-776.

2. Abhishek Singh Kilak and Namita Mittal, "Detection of Emotional States from Action Units of Facial Expressions", International Bulletin of Mathematical Research, Volume 2, Issue 1, March 2015, pp 281-285.

3. Abhishek Singh Kilak and Namita Mittal, "Classification of emotions from images using localized subsection information", International Conference on Advances in Computing and Data Sciences (ICACDS-2016), Springer Communications in Computer and Information Science(CCIS), vol 721, pp 562-571.

4. Abhishek Singh Kilak and Namita Mittal, "Multiple Instances Based Emotion Detection Using Discriminant Feature Tracking", Journal of Statistics and Management systems, Volume 21, Issue 4, 19 June 2018, pp 647-659.

Access to Bharat Database of Indian Faces may be obtained by mailing to 2012 rcp9516@mnit.ac.in or nmittal.cse@mnit.ac.in.

References

[1] Highfield R, Wiseman R, Jenkins R.: How your looks betray your personality. New Scientist (2009)

[2] en.wikipedia.org

[3] Schacter, D.L., Gilbert, D.T., Wegner, D.M., & Hood, B.M. (2011). Psychology (European ed.). Basingstoke: Palgrave Macmillan.

[4] "Emotion". The Stanford Encyclopedia of Philosophy. Metaphysics Research Lab, Stanford University. 2018.

[5] Ekman, Paul, "An argument for basic emotions." Cognition & emotion 6, no. 3-4 (1992): 169-200.

[6] Lucey, Patrick, Jeffrey F. Cohn, Takeo Kanade, Jason Saragih, Zara Ambadar and Iain Matthews, "The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression." In Computer Vision and Pattern Recognition Workshops (CVPRW), Computer Society Conference on, pp. 94-101, IEEE, 2010.

[7] Viola, P. and Jones, M,: Rapid object detection using a boosted cascade of simple features. In Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on (Vol. 1, pp. I-511, 2001.

[8] Lyons, Michael J., Shigeru Akamatsu, Miyuki Kamachi, Jiro Gyoba and Julien Budynek, "The Japanese female facial expression (JAFFE) database."

In Proceedings of third international conference on automatic face and gesture recognition, pp. 14-16, 1998.

[9] Yin, Lijun, Xiaozhou Wei, Yi Sun, Jun Wang and Matthew J. Rosato, "A 3D facial expression database for facial behavior research." In Au- tomatic face and gesture recognition, 7th international conference on, pp. 211-216, IEEE, 2006.

[10] Valstar, Michel and Maja Pantic, "Induced disgust, happiness and surprise: an addition to the mmi facial expression database." In Proc. 3rd Intern. Workshop on EMOTION (satellite of LREC): Corpora for Research on Emotion and Affect, pp. 65, 2010.

[11] Sneddon, Ian, Margaret McRorie, Gary McKeown and Jennifer Hanratty. "The belfast induced natural emotion database." IEEE Transactions on Affective Computing 3, no. 1, pp. 32-41, 2012.

[12] Happy, S. L., Priyadarshi Patnaik, Aurobinda Routray and Rajlakshmi Guha, "The Indian Spontaneous Expression Database for Emotion Recog- nition." IEEE Transactions on Affective Computing 8, no. 1, pp. 131-142, 2017.

[13] Mavadati, S. Mohammad, Mohammad H. Mahoor, Kevin Bartlett, Philip Trinh and Jeffrey F. Cohn, "Disfa: A spontaneous facial action intensity database," IEEE Transactions on Affective Computing 4, no. 2, pp. 151-160, 2013.

[14] Barrett, Lisa Feldman," How emotions are made" The secret life of the brain. Houghton Mifflin Harcourt, 2017.

[15] Juliano J. Bazzo Marcus V. Lamar, "Recognizing Facial Actions Using Gabor Wavelets with Neutral Face Average Difference", Proc. of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition (FGR'04) , 2004.

[16] Guodong Guo, Rui Guo and Xin Li, "Facial Expression Recognition Influenced by Human Aging", IEEE TRANSACTIONS ON AFFECTIVE COMPUTING, VOL. 4, NO. 3, JULY-SEPTEMBER 2013.

[17] M. Minear and D.C. Park, "A Lifespan Database of Adult Facial Stimuli," Behavior Research Methods, Instruments, & Computers, vol. 36, pp. 630-633, 2004.

[18] J. Daugman, “Uncertainty Relation for Resolution in Space, Spatial Frequency and Orientation Optimized by Two-Dimensional Visual Cortical Filters,” J. Optical Soc. of Am. A, vol. 2, pp. 1160-1169, 1985.

[19] N. Ebner, M. Riediger and U. Lindenberger, “Faces—A Database of Facial Expressions in Young, Middle-Aged and Older Women and Men: Development and Validation,” Behavior Research Methods, vol. 42, no. 1, pp. 351-362, 2010.

[20] Fengjun Chen, Zhiliang Wang, Zhengguang Xu, Donglin Wang, "Research on a Method of Facial Expression Recognition, "The Ninth International Conference on Electronic Measurement & Instruments (ICEMI), 2009.

[21] Farhan Bashar, Asif Khan, Faisal Ahmed and Md. Hasanul Kabir, "Robust Facial Expression Recognition Based on Median Ternary Pattern (MTP)", International Conference on Electrical Information and Communication Technology (EICT), 2013.

[22] Tong WANG, Haizhou AI, Gaofeng HUANG, “A Two-Stage Approach to Automatic Face Alignment”, in Proc. SPIE International Symposium on Multispectral Image Processing and Pattern Recognition, Beijing, China, pp. 558-563, 2003.

[23] Yubo WANG, Haizhou AI, Bo WU, Chang HUANG,"Real Time Facial Expression Recognition with Adaboost ", Proc. of the 17th International Conference on Pattern Recognition (ICPR’04), 2004

[24] Bihan Jiang, Michel Valstar, Brais Martinez and Maja Pantic,"A Dynamic Appearance Descriptor Approach to Facial Actions Temporal Modeling", IEEE Transactions On Cybernetics, Vol. 44, No. 2, February 2014.

[25] P. J. Lang, M. M. Bradley and B. N. Cuthbert, International Affective Picture System (IAPS): Affective Ratings of Pictures and Instruction Manual, University of Florida, Tech. Rep. A-8, 2005.

[26] M. Tkalčič, J. Tasič and A. Košir, “The LDOS-PerAff-1 corpus of face video clips with affective and personality metadata,” in Proc.Multimodal Corpora: Advances in Capturing, Coding and Analyzing Multimodality p. 111. (Malta, 2010), LREC, 2009,

[27] Matsugu, Masakazu, Katsuhiko Mori, Yusuke Mitari and Yuji Kaneda. "Subject independent facial expression recognition with robust face detection using a convolutional neural network." Neural Networks 16, no. 5-6, pp. 555-559. 2003.

[28] Fasel, Beat, "Robust face analysis using convolutional neural networks." In Object recognition supported by user interaction for service robots, vol. 2, pp. 40-43. IEEE, 2002.

[29] Li, Chao and Antonio Soares. "Automatic facial expression recognition using 3D faces." Int. J. Eng. Res. Innov 3 . pp. 30-34, 2011.

[30] Mohammed, Abdul Adeel, Rashid Minhas, QM Jonathan Wu and Maher A. Sid-Ahmed. "Human face recognition based on multidimensional PCA and extreme learning machine." Pattern Recognition 44, no. 10-11 . pp. 2588-2597, 2011.

[31] Rivera, Adin Ramirez, Jorge Rojas Castillo and Oksam Oksam Chae, "Local directional number pattern for face analysis: Face and expression recognition." IEEE transactions on image processing 22, no. 5 , pp. 1740-1752, 2013.

[32] Ebrahimi Kahou, Samira, Vincent Michalski, Kishore Konda, Roland Memisevic and Christopher Pal, "Recurrent neural networks for emotion recognition in video." In Proceedings of the 2015 ACM on International Conference on Multimodal Interaction, pp. 467-474. ACM, 2015.

[33] Kahou, Samira Ebrahimi, Christopher Pal, Xavier Bouthillier, Pierre Froumenty, Çaglar Gülçehre, Roland Memisevic, Pascal Vincent et al, "Combining modality specific deep neural networks for emotion recognition in video." In Proceedings of the 15th ACM on International conference on multimodal interaction, pp. 543-550, ACM, 2013.

[34] Krizhevsky, Alex and Geoffrey Hinto, “Learning multiple layers of features from tiny images.” Vol. 1, no. 4. Technical report, University of Toronto, 2009.

[35] Yu, Zhiding and Cha Zhang, "Image based static facial expression recognition with multiple deep network learning." In Proceedings of the

2015 ACM on International Conference on Multimodal Interaction, pp. 435-442, ACM, 2015.

[36] Dhall, Abhinav, Roland Goecke, Simon Lucey and Tom Gedeon, "Static facial expression analysis in tough conditions: Data, evaluation protocol and benchmark." In 2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops), pp. 2106-2112, IEEE, 2011.

[37] Shan, Caifeng, Shaogang Gong and Peter W. McOwan, "Facial expression recognition based on local binary patterns: A comprehensive study." Image and vision Computing 27, no. 6, pp. 803-816, 2009.

[38] Dalal, Navneet and Bill Triggs, "Histograms of oriented gradients for human detection." In Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, vol. 1, pp. 886-893, IEEE, 2005.

[39] Bowyer, Kevin W., Kyong Chang and Patrick Flynn, "A survey of approaches and challenges in 3D and multi-modal 3D+ 2D face recognition." Computer vision and image understanding 101, no. 1 . pp. 1-15, 2006.

[40] Jeni, LÃa˛szlÃs¸ A., Jeffrey F. Cohn and Takeo Kanade, "Dense 3D face alignment from 2D video for real-time use." Image and Vision Computing 58, pp. 13-24, 2017.

[41] Kamarol, Siti Khairuni Amalina, Mohamed Hisham Jaward, Heikki KÃd’lviÃd’inen, Jussi Parkkinen and Rajendran Parthiban, "Joint facial expression recognition and intensity estimation based on weighted votes of image sequences." Pattern Recognition Letters 92, pp. 25-32, 2017.

[42] Valstar, Michel F., Enrique SÃa˛nchez-Lozano, Jeffrey F. Cohn, Laszlo A.Jeni, Jeffrey M. Girard, Zheng Zhang, Lijun Yin and Maja Pantic, " FERA2017-Addressing Head Pose in the Third Facial Expression Recognition and Analysis Challenge”. arXiv preprint arXiv:1702. 0417, 2017.

[43] Battocchi, Alberto, Fabio Pianesi and Dina Goren-Bar. "A first evaluation study of a database of kinetic facial expressions (dafex)," In Proceedings of the 7th international conference on Multimodal interfaces, pp. 214-221, ACM, 2005.

[44] Jiang, Bihan, Michel F. Valstar and Maja Pantic. "Action unit detection using sparse appearance descriptors in space-time video volumes," In Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on, pp. 314-321, IEEE, 2011.

[45] Serres, Michel. The five senses: A philosophy of mingled bodies. Bloomsbury Publishing, 2008.

[46] Salzen, Eric A. "Emotion and self-awareness." Applied Animal Behaviour Science 57, no. 3-4 , pp. 299-313, 1998.

[47] Mahajan, Rashima, Dipali Bansal and Shweta Singh, "A real time set up for retrieval of emotional states from human neural responses." International Journal of Medical, Health, Pharmaceutical and Biomedical Engineering 8, no. 3, pp. 142-147, 2014.

[48] Upton, A. R. M., I. Amin, S. Garnett, M. Springman, C. Nahmias and I. S. Cooper, "Evoked metabolic responses in the limbicstriate system produced by stimulation of anterior thalamic nucleus in man." Pacing and clinical electrophysiology 10, no. 1 , pp. 217-225, 1987.

[49] Rigaux, Pierre, Félix Buhlmann and Pierre-Yves Müller, "Electrical neuromuscular stimulator for measuring muscle responses to electrical stimulation pulses." U.S. Patent 6,324,432, issued November 27, 2001.

[50] Darwin, Charles and Phillip Prodger. The expression of the emotions in man and animals. Oxford University Press, USA, 1998.

[51] Wagstaff, Christopher RD, "Emotion regulation and sport performance." Journal of Sport and Exercise Psychology 36, no. 4 , pp. 401-412, 2014.

[52] Eisenberg, Nancy. Altruistic emotion, cognition and behavior (PLE: Emotion). Psychology Press, 2014.

[53] Carstensen, Laura L. and Joseph A. Mikels, "At the intersection of emotion and cognition: Aging and the positivity effect." Current directions in psychological science 14, no. 3 , pp. 117-121, 2005.

[54] Gray, Jeremy R., Todd S. Braver and Marcus E. Raichle, "Integration of emotion and cognition in the lateral prefrontal cortex." Proceedings of the National Academy of Sciences 99, no. 6 , pp. 4115-4120, 2002.

[55] Norgaard, Kari Marie, “People want to protect themselves a little bit”: Emotions, denial and social movement nonparticipation." Sociological inquiry 76, no. 3, pp. 372-396, 2006.

[56] Diener, Ed and Martin EP Seligman, "Very happy people." Psychological science 13, no. 1, pp. 81-84, 2002.

[57] MacIntyre, Peter D. and Laszlo Vincze, "Positive and negative emotions underlie motivation for L2 learning," Studies in Second Language Learning and Teaching 7, no. 1, 2017.

[58] Vandercammen, Leen, Joeri Hofmans and Peter Theuns, "Relating specific emotions to intrinsic motivation: On the moderating role of positive and negative emotion differentiation." PloS one 9, no. 12, pp. e115396, 2014.

[59] Lisetti, Christine Lætitia and Fatma Nasoz, "Using noninvasive wearable computers to recognize human emotions from physiological signals." EURASIP journal on applied signal processing 2004, pp. 1672-1687, 2004.

[60] Steinbeis, Nikolaus, Stefan Koelsch and John A. Sloboda, "The role of harmonic expectancy violations in musical emotions: Evidence from subjective, physiological and neural responses." Journal of cognitive neuroscience 18, no. 8 , pp. 1380-1393,2006.

[61] Gewirtz, Jacob L. and Martha Peláez-Nogueras, "Infant emotions under the positive-reinforcer control of caregiver attention and touch." 2000.

[62] Henton, Wendon W. and Iver H. Iversen, Classical conditioning and operant conditioning: A response pattern analysis. Springer Science & Business Media, 2012.

[63] Lewon, Matthew and Linda J. Hayes, "Toward an analysis of emotions as products of motivating operations." The Psychological Record 64, no. 4 , pp. 813-825, 2014.

[64] Ingram, Richard, "Emotions, social work practice and supervision: an uneasy alliance?." Journal of social work practice 27, no. 1 , pp. 5-19, 2013.

[65] Mulligan, Kevin. "Emotions and values," In The Oxford handbook of philosophy of emotion. 2010.

[66] Burke, Peter, "Is there a Cultural History of the Emotions?." In Representing Emotions, pp. 35-48, Routledge, 2017.

[67] Plamper, Jan, The history of emotions: An introduction. OUP Oxford, 2015.

[68] Fredrickson, Barbara L. "The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions." American psychologist 56, no. 3, pp. 218, 2001.

[69] Aviezer, Hillel, Yaacov Trope and Alexander Todorov, "Body cues, not facial expressions, discriminate between intense positive and negative emotions." Science 338, no. 6111 , pp. 1225-1229, 2012.

[70] Rose, Mary R., Janice Nadler and Jim Clark, "Appropriately upset? Emotion norms and perceptions of crime victims." Law and Human Behavior 30, no. 2 , pp. 203-219, 2006.

[71] Wittig, Cynthia. "Employees' reactions to organizational change," Od practitioner 44, no. 2, pp. 23-28, 2012.

[72] Spielberger, Charles D. "State-Trait anxiety inventory," The Corsini encyclopedia of psychology, pp. 1-1, 2010.

[73] Garber, Marjorie, Vested interests: Cross-dressing and cultural anxiety. Routledge, 2012.

[74] Eysenck, Michael W., Nazanin Derakshan, Rita Santos and Manuel G. Calvo, "Anxiety and cognitive performance: attentional control theory." Emotion 7, no. 2 , pp. 336, 2007.

[75] Ammerman, Brooke A., Evan M. Kleiman, Lauren L. Uyeji, Anne C. Knorr and Michael S. McCloskey, "Suicidal and violent behavior: The role of anger, emotion dysregulation and impulsivity." Personality and Individual Differences 79 , pp. 57-62, 2015.

[76] Shortt, Joann Wu, Mike Stoolmiller, Jessica N. Smith-Shine, J. Mark Eddy and Lisa Sheeber, "Maternal emotion coaching, adolescent anger regulation and siblings' externalizing symptoms." Journal of Child Psychology and Psychiatry 51, no. 7 , pp. 799-808, 2010.

[77] Morris, Amanda Sheffield, Jennifer S. Silk, Michael DS Morris, Laurence Steinberg, Katherine J. Aucoin and Angela W. Keyes, "The influence of mother–child emotion regulation strategies on children's expression of anger and sadness." Developmental psychology 47, no. 1, pp. 213, 2011.

[78] Rieffe, Carolien, Marina Camodeca, Lucinda BC Pouw, Aurelie MC Lange and Lex Stockmann, "Don't anger me! Bullying, victimization and emotion dysregulation in young adolescents with ASD." European Journal of Developmental Psychology 9, no. 3, pp. 351-370, 2012.

[79] Bänziger, Tanja, Didier Grandjean and Klaus R. Scherer, "Emotion recognition from expressions in face, voice and body: the Multimodal Emotion Recognition Test (MERT)." Emotion 9, no. 5, pp. 691, 2009.

[80] Bahreini, Kiavash, Rob Nadolski and Wim Westera, "Towards multimodal emotion recognition in e-learning environments." Interactive Learning Environments 24, no. 3, pp. 590-605, 2016.

[81] Zhao, Mingmin, Fadel Adib and Dina Katabi. "Emotion recognition using wireless signals," In Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking, pp. 95-108, ACM, 2016.

[82] Zheng, Wei-Long, Jia-Yi Zhu and Bao-Liang Lu, "Identifying stable patterns over time for emotion recognition from EEG." IEEE Transactions on Affective Computing 2017.

[83] Deb, Suman and Samarendra Dandapat, "Fourier model based features for analysis and classification of out-of-breath speech." Speech Communication 90, pp. 1-14, 2017.

[84] Poria, Soujanya, Iti Chaturvedi, Erik Cambria and Amir Hussain, "Convolutional MKL based multimodal emotion recognition and sentiment analysis." In 2016 IEEE 16th international conference on data mining (ICDM), pp. 439-448, IEEE, 2016.

[85] Kołakowska, Agata, "Usefulness of keystroke dynamics features in user authentication and emotion recognition." In Human-Computer Systems Interaction, pp. 42-52. Springer, Cham, 2018.

[86] Favre, Pauline, Mircea Polosan, Cédric Pichat, Thierry Bougerol and Monica Baciu, "Cerebral correlates of abnormal emotion conflict processing in euthymic bipolar patients: a functional MRI study." PloS one 10, no. 8 , pp. e0134961, 2015.

[87] Li, Zhonglin, Li Tong, Min Guan, Wenjie He, Linyuan Wang, Haibin Bu, Dapeng Shi and Bin Yan, "Altered resting-state amygdala functional

connectivity after real-time fMRI emotion self-regulation training." BioMed research international 2016.

[88] Cerami, Chiara, Alessandra Dodich, Sandro Iannaccone, Alessandra Marcone, Giada Lettieri, Chiara Crespi, Luigi Gianolli, Stefano F. Cappa and Daniela Perani, "Right limbic FDG-PET hypometabolism correlates with emotion recognition and attribution in probable behavioral variant of frontotemporal dementia patients." PLoS One 10, no. 10 , pp. e0141672, 2015.

[89] Piana, Stefano, Alessandra Staglianò, Francesca Odone and Antonio Camurri, "Adaptive body gesture representation for automatic emotion recognition." ACM Transactions on Interactive Intelligent Systems (TiiS) 6, no. 1, pp. 6, 2016.

[90] Wan, Jun, Sergio Escalera, Gholamreza Anbarjafari, Hugo Jair Escalante, Xavier Baró, Isabelle Guyon, Meysam Madadi et al, "Results and analysis of chalearn lap multi-modal isolated and continuous gesture recognition and real versus fake expressed emotions challenges." In Proceedings of the IEEE International Conference on Computer Vision, pp. 3189-3197, 2017.

[91] Bahreini, Kiavash, Rob Nadolski and Wim Westera, "Towards multimodal emotion recognition in e-learning environments." Interactive Learning Environments 24, no. 3, pp. 590-605, 2016.

[92] Lee, Jinkyu and Ivan Tashev, "High-level feature representation using recurrent neural network for speech emotion recognition." In Sixteenth Annual Conference of the International Speech Communication Association, pp.1537-1540, 2015.

[93] Hogan, Candice L., Lahnna I. Catalino, Jutta Mata and Barbara L. Fredrickson, "Beyond emotional benefits, Physical activity and sedentary behaviour affect psychosocial resources through emotions." Psychology & health 30, no. 3 , pp. 354-369, 2015.

[94] Sreedharan, Ninu Preetha Nirmala, Brammya Ganesan, Ramya Raveendran, Praveena Sarala and Binu Dennis, "Grey Wolf optimisation-based feature selection and classification for facial emotion recognition." IET Biometrics 7, no. 5, pp. 490-499, 2018.

[95] Heerdink, Marc W., Gerben A. van Kleef, Astrid C. Homan and Agneta H. Fischer, "Emotional expressions as social signals of rejection and acceptance: evidence from the affect misattribution paradigm." Journal of Experimental Social Psychology 56 , pp. 60-68, 2015.

[96] Happy, S. L. and Aurobinda Routray, "Automatic facial expression recognition using features of salient facial patches." IEEE transactions on Affective Computing 6, no. 1, pp. 1-12, 2015.

[97] Wood, Adrienne, Magdalena Rychlowska, Sebastian Korb and Paula Niedenthal, "Fashioning the face: sensorimotor simulation contributes to facial expression recognition." Trends in cognitive sciences 20, no. 3 , pp. 227-240, 2016.

[98] Nummenmaa, Lauri and Manuel G. Calvo, "Dissociation between recognition and detection advantage for facial expressions: A meta-analysis." Emotion 15, no. 2 , pp. 243, 2015.

[99] Sahakian, Marlyne and Béatrice Bertho, "Exploring emotions and norms around Swiss household energy usage: When methods inform understandings of the social." Energy research & social science 45,pp. 81-90, 2018.

[100] Preece, David A., Rodrigo Becerra, Ken Robinson, Justine Dandy and Alfred Allan, "Measuring emotion regulation ability across negative and positive emotions: The Perth Emotion Regulation Competency Inventory (PERCI)." Personality and Individual Differences 135 , pp. 229-241, 2018.

[101] Piwek, Lukasz and Adam Joinson, "Automatic tracking of behavior with smartphones: Potential for behavior change interventions." In Behavior Change Research and Theory, pp. 137-165, Academic Press, 2017.

[102] Zhang, Xiao, Wenzhong Li, Xu Chen and Sanglu Lu, "MoodExplorer: Towards Compound Emotion Detection via Smartphone Sensing, " Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 1, no. 4, pp. 176, 2018.

[103] Wood, Adrienne, Magdalena Rychlowska, Sebastian Korb and Paula Niedenthal, "Fashioning the face: sensorimotor simulation contributes to

facial expression recognition." Trends in cognitive sciences 20, no. 3 , pp. 227-240, 2016.

[104] Anderson, Lisa M., Erin E. Reilly, Sasha Gorrell, Katherine Schaumberg and Drew A. Anderson, "Gender-based differential item function for the difficulties in emotion regulation scale." Personality and Individual Differences 92, pp. 87-91, 2016.

[105] Carcagnì, Pierluigi, Marco Del Coco, Marco Leo and Cosimo Distante, "Facial expression recognition and histograms of oriented gradients: a comprehensive study." SpringerPlus 4, no. 1, pp. 645, 2015.

[106] Martinez, Brais and Michel F. Valstar, "Advances, challenges and opportunities in automatic facial expression recognition," In Advances in face detection and facial image analysis, pp. 63-100, Springer, Cham, 2016.

[107] Bruce, Vicki, Recognising faces. Routledge, 2017.

[108] Chen, Jiansheng, Yu Deng, Gaocheng Bai and Guangda Su, "Face image quality assessment based on learning to rank," IEEE signal processing letters 22, no. 1 , pp. 90-94, 2015.

[109] Bandhakavi, Anil, Nirmalie Wiratunga, Stewart Massie and Deepak Padmanabhan, "Lexicon generation for emotion detection from text." IEEE intelligent systems 32, no. 1, pp. 102-108, 2017.

[110] Abdelwahab, Mohammed and Carlos Busso, "Supervised domain adaptation for emotion recognition from speech." In 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 5058-5062, IEEE, 2015.

[111] Daly, Ian, Duncan Williams, Asad Malik, James Weaver, Alexis Kirke, Faustina Hwang, Eduardo Miranda and Slawomir J. Nasuto, "Personalised, Multi-modal, Affective State Detection for Hybrid Brain-Computer Music Interfacing." IEEE Transactions on Affective Computing (2018).

[112] Ayesh, Aladdin, Miguel Arevalillo-Herráez and Francesc J. Ferri, "Towards psychologically based personalised modelling of emotions using associative classifiers," International Journal of Cognitive Informatics and Natural Intelligence (IJCINI) 10, no. 2, pp. 52-64, 2016.

[113] Chiranjeevi, Pojala, Viswanath Gopalakrishnan and Pratibha Moogi, "Neutral face classification using personalized appearance models for fast and robust emotion detection." IEEE Transactions on Image Processing 24, no. 9, pp. 2701-2711, 2015.

[114] Tahon, Marie and Laurence Devillers, "Towards a small set of robust acoustic features for emotion recognition: challenges." IEEE/ACM transactions on audio, speech and language processing 24, no. 1, pp. 16-28, 2016.

[115] Cohen, Ira, Nicu Sebe, Yafei Sun, Michael S. Lew and Thomas S. Huang, "Evaluation of expression recognition techniques." In International Conference on Image and Video Retrieval, pp. 184-195, Springer, Berlin, Heidelberg, 2003.

[116] Tao, H., Huang, ”Connected vibrations,” A modal analysis approach to non-rigid motion tracking. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR’98). pp. 735–740, 1998.

[117] Martinez, Brais and Michel F. Valstar, "Advances, challenges and opportunities in automatic facial expression recognition," In Advances in face detection and facial image analysis, pp. 63-100, Springer, Cham, 2016.

[118] Y. Zhang and Q. Ji, “Active and dynamic information fusion for facial expression understanding from image sequences,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 27, no. 5, pp. 699–714, May 2005.

[119] Y. Tian, T. Kanade and J. F. Cohn, “Recognizing lower face action units for facial expression analysis,” in Proc. IEEE Int. Conf. Autom. Face Gesture Recog., pp. 484–490, 2000.

[120] M. Z. Uddin, J. J. Lee and T.-S. Kim, “An enhanced independent component-based human facial expression recognition from video,” IEEE Trans. Consum. Electron., vol. 55, no. 4, pp. 2216– 2224, Nov. 2009.

[121] M. F. Valstar and M. Pantic, “Combined support vector machines and hidden Markov models for modeling facial action temporal dynamics,” in Proc. IEEE Int. Conf. Human–Comput. Interaction, pp. 118–127, 2007.

[122] M. Bartlett, G. Littlewort, M. Frank, C. Lainscsek, I. Fasel and J. Movellan, “Recognizing facial expression: Machine learning and application to

spontaneous behavior," in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recog., pp. 568–573, 2005.

[123] M. Lyons, J. Budynek and S. Akamatsu, "Automatic classification of single facial images," IEEE Trans. Pattern Anal. Mach. Intell., vol. 21, no. 12, pp. 1357–1362, Dec. 1999.

[124] S. Lucey, I. Matthews, C. Hu, Z. Ambadar, F. D. L. Torre and J. Cohn, "AAM derived face representations for robust facial action recognition," in Proc. 7th Int. Conf. Autom. Face Gesture Recog., pp. 155–160, 2006.

[125] A. B. Ashraf, S. Lucey, T. C. Jeffrey F. Cohn, Z. Ambadar, K. M. Prkachin and P. E. Solomon, "The painful face–pain expression recognition using active appearance models," Image Vis. Comput., vol. 27, no. 12, pp. 1788–1796, 2009.

[126] B. Abboud, F. Davoine and M. Dang, "Facial expression recogni- tion and synthesis based on an appearance model," Signal Process.: Image Commun., vol. 19, no. 8, pp. 723–740, 2004.

[127] A. Asthana, J. Saragih, M. Wagner and R. Goecke, "Evaluating aam fitting methods for facial expression recognition," in Proc. 3rd Int. Conf. Affective Comput. Intell. Interaction Workshops, pp. 1–8, 2009.

[128] C. Martin, U. Werner and H.-M. Gross, "A real-time facial expression recognition system based on active appearance models using gray images and edge images," in Proc. 8th IEEE Int. Conf. Autom. Face Gesture Recog., pp. 1–6, 2008.

[129] D. Cristinacce and T. F. Cootes, "Feature detection and tracking with constrained local models," in Proc. Brit. Mach. Vis. Conf., pp. 929–938. 2006.

[130] J. M. Saragih, S. Lucey and J. F. Cohn, "Deformable model fitting by regularized landmark mean-shift," Int. J. Comput. Vis., vol. 91, no. 2, pp. 200–215, 2011.

[131] A. Asthana, S. Zafeiriou, S. Cheng and M. Pantic, "Robust discriminative response map fitting with constrained local models," in Proc. IEEE Conf. Comput. Vis. Pattern Recog., pp. 3444–3451, 2013.

[132] Kanade, Takeo, Jeffrey F. Cohn and Yingli Tian, "Comprehensive database for facial expression analysis," In Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580), pp. 46-53, IEEE, 2000.

[133] G. Zhao and M. Pietikainen, “Dynamic texture recognition using local binary patterns with an application to facial expressions,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 29, no. 6, pp. 915–928, Jun. 2007.

[134] A. Hadid, M. Pietikainen and T. Ahonen, “A discriminative fea- ture space for detecting and recognizing faces,” in Proc. IEEE Com- put. Soc. Conf. Comput. Vis. Pattern Recog., pp. 797–804, 2004.

[135] S. L. Happy, A. George and A. Routray, “A real time facial expression classification system using local binary patterns,” in Proc. 4th Int. Conf. Intell. Human Comput. Interaction, pp. 1–5, 2012.

[136] T. Jabid, M. Kabir and O. Chae, “Robust facial expression recogni- tion based on local directional pattern,” ETRI J., vol. 32, pp. 784– 794, 2010.

[137] A. Dhall, A. Asthana, R. Goecke and T. Gedeon, “Emotion recog- nition using PHOG and LPQ features,” in Proc. IEEE Int. Conf. Autom. Face Gesture Recog. Workshops, pp. 878–883, 2011.

[138] C. Shan, S. Gong and P. W. McOwan, “Facial expression recogni- tion based on local binary patterns: A comprehensive study,” Image Vis. Comput., vol. 27, no. 6, pp. 803–816, 2009.

[139] K. I. Kim, K. Jung and H. J. Kim, “Face recognition using kernel principal component analysis,” IEEE Signal Process. Lett., vol. 9, no. 2, pp. 40–42, Feb. 2002.

[140] A. J. Calder, A. M. Burton, P. Miller, A. W. Young and S. Akamatsu, “A principal component analysis of facial expressions,” Vis. Res., vol. 41, no. 9, pp. 1179–1208, 2001.

[141] H.-B. Deng, L.-W. Jin, L.-X. Zhen and J.-C. Huang, “A new facial expression recognition method based on local Gabor filter bank and PCA plus LDA,” Int. J. Inform. Technol., vol. 11, no. 11, pp. 86– 96, 2005.

[142] Z. Zhang, Y. Yan and H. Wang, “Discriminative filter based regression learning for facial expression recognition,” in Proc. 20th IEEE Int. Conf. Image Process., pp. 1192–1196, 2013.

[143] C. Shan, S. Gong and P. W. McOwan, “A comprehensive empiri- cal study on linear subspace methods for facial expression analy- sis,” in Proc. IEEE Conf. Comput. Vis. Pattern Recog. Workshop, pp. 153–153, 2006.

[144] S.-K. Oh, S.-H. Yoo and W. Pedrycz, “Design of face recognition algorithm using PCA-LDA combined for hybrid data pre-process- ing and polynomial-based RBF neural networks: Design and its application,” Expert Syst. Appl., vol. 40, no. 5, pp. 1451–1466, 2013.

[145] Y. Rahulamathavan, R.-W. Phan, J. A. Chambers and D. J. Parish, “Facial expression recognition in the encrypted domain based on local fisher discriminant analysis,” IEEE Trans. Affective Comput., vol. 4, no. 1, pp. 83–92, Jan.-Mar. 2013.

[146] C. Shan and R. Braspenning, “Recognizing facial expressions automatically from video,” in Handbook of Ambient Intelligence and Smart Environments, New York, NY, USA: Springer-Verlag, pp. 479–509, 2010.

[147] Shan, Caifeng, Shaogang Gong, and Peter W. McOwan. "Robust facial expression recognition using local binary patterns." In IEEE International Conference on Image Processing 2005, vol. 2, pp. II-370. IEEE, 2005.

[148] C. Shan and T. Gritti, “Learning discriminative LBP-histogram bins for facial expression recognition,” in Proc. Brit. Mach. Vis. Conf., pp. 1–10, 2008.

[149] M. Song, D. Tao, Z. Liu, X. Li and M. Zhou, “Image ratio features for facial expression recognition application,” IEEE Trans. Syst., Man, Cybern., Part B: Cybern., vol. 40, no. 3, pp. 779–788, Jun. 2010.

[150] Burgoon, Judee K., Laura K. Guerrero and Kory Floyd, Nonverbal communication, Routledge, 2016.

[151] Froiland, John Mark, "Parents’ weekly descriptions of autonomy supportive communication: Promoting children’s motivation to learn and positive emotions," Journal of Child and Family Studies 24, no. 1 , pp. 117-126, 2015.

[152] P. Ahmad, N. A. Hossein, G. Marina and Y. N. Svetlana, “Gauss– Laguerre wavelet textural feature fusion with geometrical infor- mation for facial expression identification,” EURASIP J. Image Video Process., vol. 2012, no. 1, pp. 1–13, 2012.

[153] L. Zhong, Q. Liu, P. Yang, B. Liu, J. Huang and D. N. Metaxas, “Learning active facial patches for expression analysis,” in Proc. IEEE Conf. Comput. Vis. Pattern Recog., pp. 2562–2569, 2012.

[154] L. Zhang and D. Tjondronegoro, “Facial expression recognition using facial movement features,” IEEE Trans. Affective Comput., vol. 2, no. 4, pp. 219–229, Oct. 2011.