Marathwada Shikshan Prasarak Mandal's

Deogiri Institute of Engineering and Management Studies,

Aurangabad

Seminar Report

On

Mood Detection From Facial Expression

Submitted By

Rukar Priyanka Somnath (36172)

Dr. Babasaheb Ambedkar Technological University

Lonere (M.S.)

Department of Computer Science and Engineering

Deogiri Institute of Engineering and Management Studies,

Aurangabad

(2019- 2020)

CERTIFICATE

This is to certify that, the Seminar entitled "**Mood Detection From Facial Expression**" submitted by **Priyanka Somnath Rukar** is a bonafide work completed under my supervision and guidance in partial fulfillment for award of Bachelor of Technology (Computer Science and Engineering) Degree of Dr. Babasaheb Ambedkar Technological University, Lonere.

Place: Aurangabad
Date:

Mr/Ms/Mrs P.H.Durule
Guide

Mr. S.B. Kalyankar
Head

Dr. Ulhas D. Shiurkar
Director,
Deogiri Institute of Engineering and Management Studies,
Aurangabad

Abstract

Emotions have an extremely important role in human lives. They determine how humans think, behave and communicate with others and are the only thing which separates us from machines. tone of voice, Gestures, body posture, etc. all express some kind of information about human emotions but the facial features and expressions are one which expresses human emotions clearly and accurately during daily communication. There are many situations in real world where human and computer needs to interact with each other. The interaction between humans and we introduce seven emotions and positive and negative emotion recognition methods using facial images and the development of apps based on the method. In previous researches, they used the deep-learning technology to generate models with emotion-based facial expressions to recognized emotions.

There are existing apps that express six emotions, but not seven emotions and positive and negatives in graphs and percentages. Thus, we recognized seven emotions such as Angry, Disgust, Fear, Happy, Sad, Surprise, and Neutral and also classified the calculated emotion-recognition scores into positive, negative and neutral emotions. Then we implemented an app that provides the user with sevesssssn emotions scored and positive and negative emotion Keywords — Human machine interaction, Emotions, PCNN, Facial expression

computers will become more natural if computers can perceive and respond to non verbal communication of humans. Therefore there exist need of machines which are able to identify human mood so that a communication bridge can be established between humans and machines and a better interaction will be facilitated.

This paper proposes a system using Pulse Coupled Neural Network (PCNN) for detecting facial features which are responsible for portraying the facial expression. This information is then passed to a trained Convolution Neural Network (CNN) which is responsible for the classification of expressions in six categories as happy, sad, neutral, fear, angry and surprised.

There are existing apps that express six emotions, but not seven emotions and positive and negatives in graphs and percentages. Thus, we recognized seven emotions such as Angry, Disgust, Fear, Happy, Sad, Surprise, and Neutral and also classified the calculated emotion.

INDEX

1. INTRODUCTION

The human face is an elastic object that consists of organs, numerous muscles, skins and bones. When a muscle contracts, the transformation of the corresponding skin areas attached to the muscle results in facial expressions . These facial expressions are examined for identifying the basic human mood like anger, fear, disgust, surprised, happiness, sadness. The proposed work has many practical application in the field of security, education, medical, game, monitoring, law, marketing, entertainment etc by identifying users response to video games, commercials, or newly launched products, to identify struggling students in a classroom environment, or help autistics better interact with others, to better measure TV ratings, adding another security layer to security at malls, airports, sports arenas, and other public venues to detect malicious intent.

Mood detection is a challenging problem as it involves 3 sub problems 1) face detection 2) facial expression feature extraction and 3) expression classification to identify mood. Each sub problem has difficulties such as background details of input, illumination changes, and variable size of an input and poses variations. The approaches to facial expression recognition can be divided into two classes static image based approaches and image sequence based approaches. Static image based approaches classifies facial expressions based on a single image and the image sequence based approaches use the motion information in an image sequence. In another way, they can be classified into geometric feature based approach and appearance based approach. The geometric feature based approach relies on the geometric facial features such as locations and contours of eyebrows, nose, mouth etc. Appearance based approach uses whole face or specific region in a face image for a feature extraction via some kind of filters or transformation .

we introduce seven emotions and positive and negative emotion recognition methods using facial images and the development of apps based on the method. In previous researches, they used the deep-learning technology to generate models with emotion-based facial expressions to recognized emotions. There are existing apps that express six emotions, but not seven emotions and positive and negatives in graphs and percentages. Thus, we recognized seven emotions such as Angry, Disgust, Fear, Happy, Sad, Surprise, and Neutral and also classified the calculated

emotion-recognition scores into positive, negative and neutral emotions. Then we implemented an app that provides the user with seven emotions scored and positive and negative emotions.

we introduce a method for recognizing seven emotions such as Angry, Disgust, Fear, Happy, Sad, Surprise, and Neutral and positive and negative emotions using facial images and the development of apps based on the method. Previous research used deep-learning technology to generate models with emotion-based facial expressions to recognize emotions. The emotion-recognition Software Development Kit (SDK) made by the US company "Affectiva" extracted features from facial expressions using a Histogram of Oriented Gradient (Hog) algorithm and learned 10,000 images using a Support Vector Machine (SVM) classifier. Seven facial expressions, such as anger, disgust, fear, joy, sadness, surprise, and contempt, were used for learning to recognize emotions. The generated emotion recognizer was used to develop SDK, which provides an easy interface for other users.

The implemented app provides a total of six emotions and emoticons based on the emotion results. Although most emotion-recognition studies using deep learning proceeded with seven emotions, they did not recognize positive and negative emotions using the scores of the emotion recognition results. Also in study , they just recognized positive and negative emotions using face expression. They divided into 11 features in the face and then used the amount of movement to indicate positive and negative emotions.

Therefore, in this study, we added all seven emotions in the app and referred to the research developed to transfer emotion recognition to the mobile phone. We also provided positive and negative emotion-recognition results using the ranking and average of the scores from seven emotion-recognition results. We developed an emotion-recognition model using deep learning's Convolutional Neural Networks (CNN) to develop this app and proposed a method for recognizing emotions. Thus, in this study, we classified the calculated emotion-recognition scores into positive and negative emotions and implemented an app that provides the user with scores for seven positive and negative emotions.

Facial images in different expressions from the Cohn: Kanade database (reproduced with permission from Cohn-Kanade database.

Face emotion recognition uses support vector machine for finding the different emotions of face and also for classifying them. PCA is used to extract the facial features and to reduce the image dimensions. Face is a two dimensional image, for face analysis it is preferred to use two dimensional vector space. Therefore for dimensionality reduction also 2DPCA is best for faces under different poses. 2DPCA is used to remove the unnecessary parts of the image. Multiobjective algorithm based optimization and classifiers are used. SVMs are used to classify the image data under consideration. It finds the minimum possible separation between two or more classes of data and creates a hyper-plane with a margin. Since in general greater the margin and lesser is the generalization error. SVMs are memory efficient and effective in higher dimension spaces.

The rest of paper is organized as follows. Section II describes the face recognition and its algorithm. Section III describes the emotion detection system. Experimental results are shown in section IV. Finally in section V the conclusion of our work will be discussed.

Face recognition is a process of identification of a person's face in an image or video which includes several steps need to be processed

2. LITERATURE SURVEY

The study of face and its features is an active research area from past few decades. Pose variation, illumination conditions, bad lighting etc., are still challenging factors faced by all algorithms. Face recognition and emotion detection system are the major applications of recognition system, in which many algorithms have tried to solve these problems. The face recognition is the basic part in modern authentication/identification applications; the accuracy of this system should be high for better results. Fisherface algorithm presents high accurate approach for face recognition; it performs two classes of analyses to achieve recognition i.e. principal component analysis (PCA) and linear discriminant analysis (LDA) respectively.

- . FACE RECOGNITION

Face recognition is a process of identification of a person's face in an image or video which includes several steps need to be processed. Figure.1 shows the block diagram of face recognition system, which includes face detection, face extraction and face matching

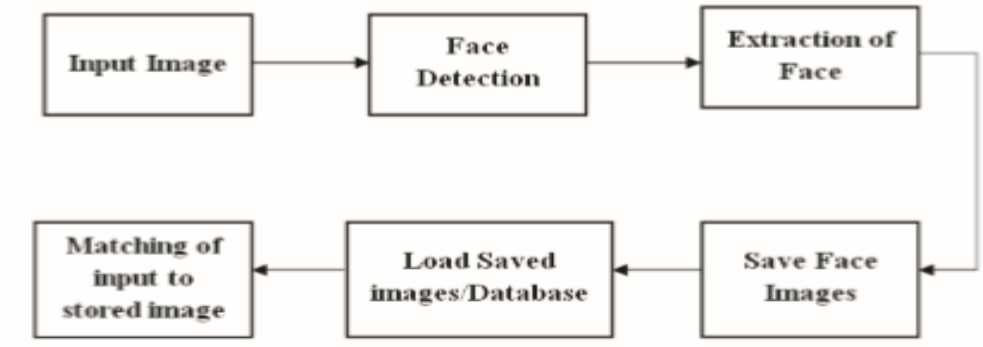

Figure 1: Block diagram of Face recognition system.

Recognition is based on the stored image data of the different group of persons. Input images are of any type can be used for recognition, 1. Still images. 2. Video frames or video stills. 3. Video. Input image is subjected for face detection to detect the face. Detected face is then extracted from the image and these images are saved as a database. Saved images are used to

compare with the input image. The matching of input image is performed to identify the user's identity. The recognition result gives identification of the person (particularly his/her name). Figure 2 depicts the step by step architecture developed for face recognition.

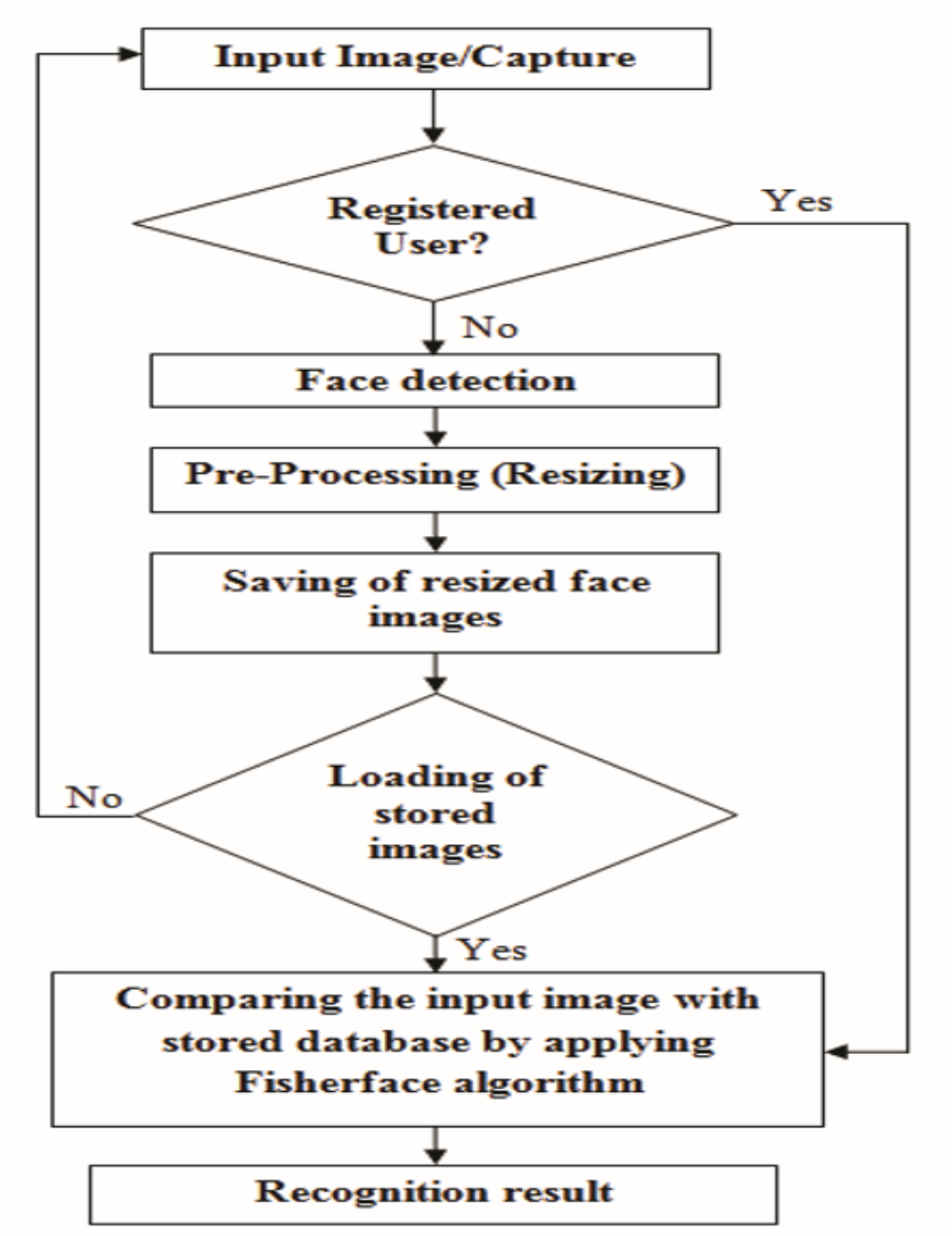

Figure 2: Real time Face recognition system.

The algorithm is designed in such a way that, if the person is recognizing for the first time then the system considers him as a new user and performs each step of operation.

But if the person/user data is already stored then it is considered as “Registered user” and it performs only matching operation to recognize the user identity. Open CV contains cascade classifiers in which Viola & Jones face detection algorithm is implemented. By using these classifiers the face region is detected from the image. It classifies the images into positive and negative images respectively. The images consists of face region is considered as positive and the images without face as negative images. These negative images are ignored for further processing. The stored images consists of face images of dimension 273x273, the more number of the images higher the recognition rate. If the stored images consists false images or wrong extensions then recognition is not possible. Therefore care should be taken while capturing input images. Fisherface algorithm is applied for classification of different users. Fisherface algorithm generates the fisherfaces of each image that are used for recognition. In Fisherface algorithm, it performs “leave-one-out” cross validation to validate the user identification.

III. EMOTION DETECTION

Face emotion detection is used to predict the emotion state of the person based on their face expressions. The overview of the emotion detection system is shown in the figure 3 as follows.

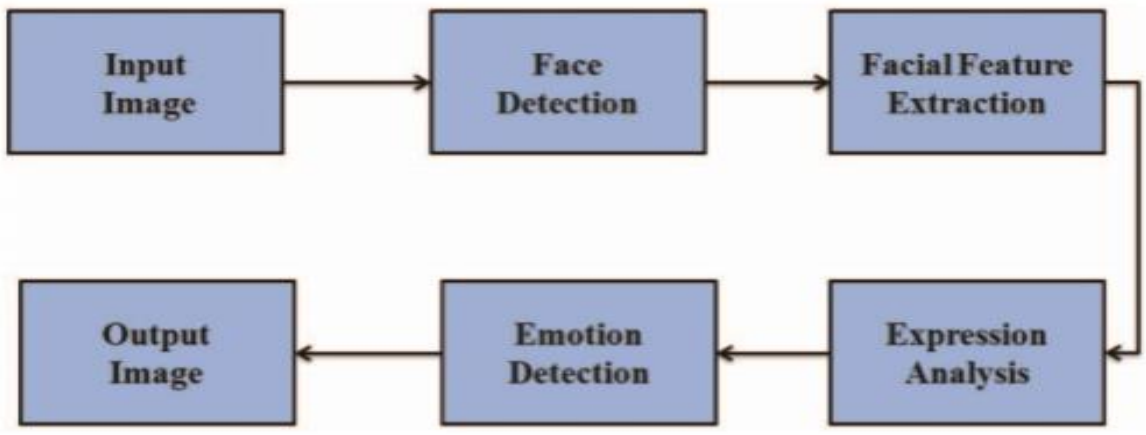

Figure 3: Block diagram of Emotion detection system.

Here input images are classified into two types, o Training images. o Testing images. Training images are used for training of classifier.

Testing images are used to verify the algorithm by predicting the different emotions of the face.

Expression analysis is the major part of the emotion detection, the schematic of expression analysis for classifying different emotions is shown in figure 4.

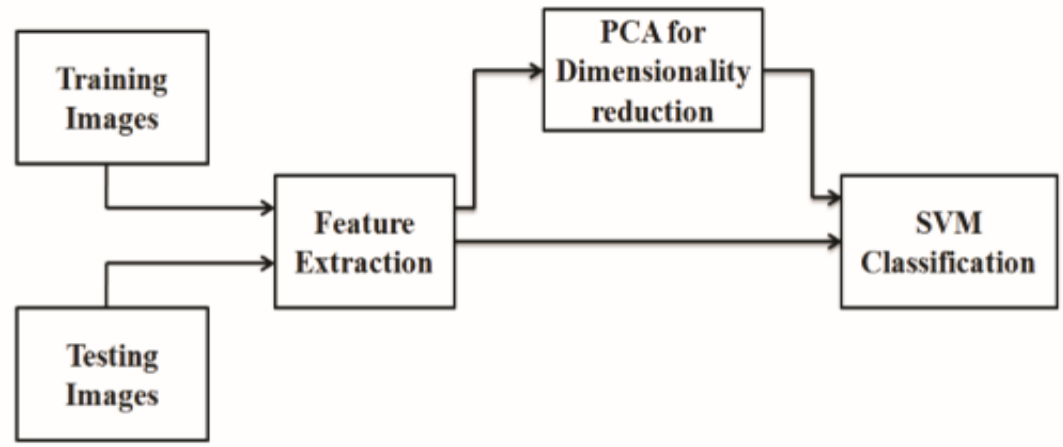

Figure 4: Block diagram of Emotion detection system.

PCA is applied to training images to reduce the dimensionality. Because training images are more compared to testing and if the dimension is high then the time taken for processing will be more. Support vector machine classification is done for classifying different emotions namely, Happy, Sad, Angry, Fear, Disgust and Surprise. The emotion detection system detailed flow diagram is shown in .

aspects: visualization, audio, and time. To solve this problem, recently many algorithms for processing spontaneously generated emotional behaviors have been proposed. Moreover, increasing research towards fusion methods is proposed for analyzing human emotion, such as feature-level fusion, decision-level fusion and multimodal fusion. Besides, the fusion methods for merging the information about facial expressions, head movements and body gestures are also proposed.

This paper mainly focuses on different learning methods. As we know, different learning methods are suitable to different prediction problems. In this paper, we have implemented several learning methods which are ones of the best methods in general cases. Besides, these prediction systems are based on the characteristics of feature set and simple fusion is also implemented. The aim is to design and construct the prediction system which is most suitable to the challenges.

3. Brief on System

Introduce a method for recognizing seven emotions such as Angry, Disgust, Fear, Happy, Sad, Surprise, and Neutral and positive and negative emotions using facial images and the development of apps based on the method. Previous research used deep-learning technology to generate models with emotion-based facial expressions to recognize emotions. The emotion-recognition Software Development Kit (SDK) made by the US company "Affectiva" extracted features from facial expressions using a Histogram of Oriented Gradient (Hog) algorithm and learned 10,000 images using a Support Vector Machine (SVM) classifier. Seven facial expressions, such as anger, disgust, fear, joy, sadness, surprise, and contempt, were used for learning to recognize emotions. The generated emotion recognizer was used to develop SDK, which provides an easy interface for other users. The implemented app provides a total of six emotions and emoticons based on the emotion results. Although most emotion-recognition studies using deep learning proceeded with seven emotions, they did not recognize positive and negative emotions using the scores of the emotion recognition results.

A. Experimental results of face recognition.

Figure 6:Real-Time Face recognition results.

Also in study , they just recognized positive and negative emotions using face expression. They divided into 11 features in the face and then used the amount of movement to indicate positive and negative emotions.

negative emotion-recognition results using the ranking and average of the scores from seven 13emotion-recognition results. We developed an emotion-recognition model using deep learning's Convolutional Neural Networks (CNN) to develop this app and proposed a method for recognizing emotions. Thus, in this study, we classified the calculated emotion-recognition scores into positive and negative emotions and implemented an app that provides the user with scores for seven positive and negative emotions.

Classifiers which groups the number of features into different classifiers. Using these classifiers, the different facial features are detected which can be used for further processing. But this method is limited to only frontal view of the face. Therefore in order to detect the facial features of the person's face under different poses, face landmark annotation is required. For annotation of facial landmarks, the popular method active shape model (ASM) or open source software dlib can be used. Emotion detection is based on different expressions of face and these expressions

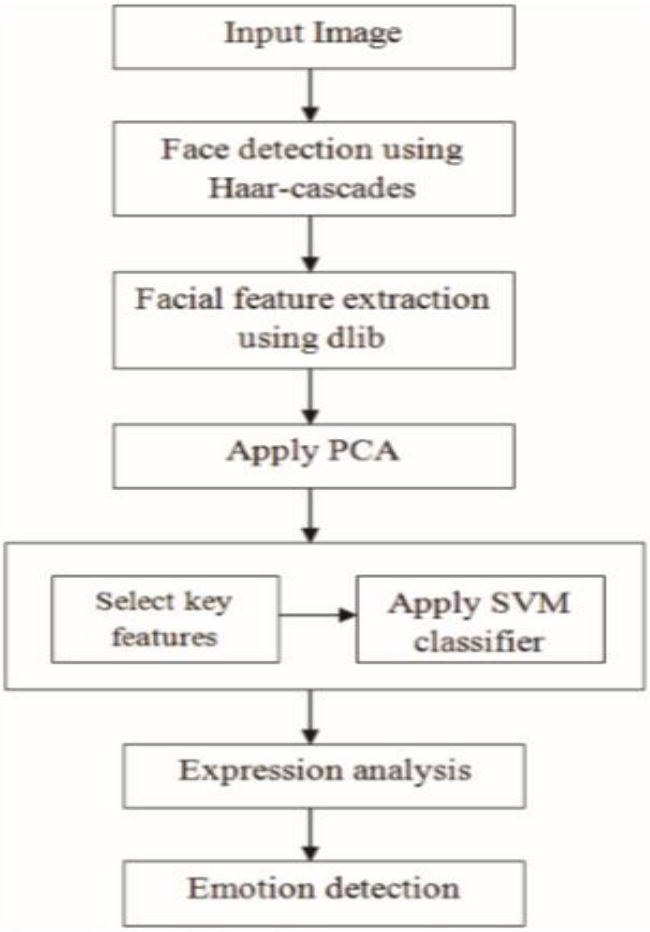

Figure 5: workflow of Emotion detection system.

PCA is applied to training images to reduce the dimensionality. Because training images are more compared to testing and if the dimension is high then the time taken for processing will be more. Support vector machine classification is done for classifying different emotions namely, Happy, Sad, Angry, Fear, Disgust and Surprise. The emotion detection system detailed flow diagram is shown in figure.

Facial features such as eyes, nose, lips and face contour are considered as the action units of face and are responsible for creation of expressions on face, are extracted using open source software called dlib. SVM classifier compares the features of training data and testing data to predict any emotion of the face. Here facial features are considered as the key points which are used for training and testing. Support vector machine is the supervised learning method of machine learning. Machine learning algorithms are advantageous over other algorithms, because of less error rate and faster results. Linear SVC which is also called as Multi SVM is used for classification. It uses “one-vs-all” strategy for training of n-class models.

TABLE I. TIME ESTIMATION

Type	Time taken (SEC)
Face detection	0.0844
Facial feature extraction	0.9216
Classification using SVM	0.1956
Emotion detection	0.1994

The emotions can be classified as positive and negative, these can be used to understand the mental condition of the person. The implementation is done using Open CV and python along with additional dependencies like dlib, scikit learn and scimage. Table 1 shows the time estimation of different detections performed. The time taken for each process is obtained using the time function of the python.

For face recognition, we have used a webcam for capturing of faces. The implemented algorithm is capable of recognizing different persons in a single window. If the recognition environment is under proper lighting condition and less background noises then the recognition rate will be high.

B. Experimental results of Emotion detection. The images of dimension 640x480 are used for testing of emotion detection system. For training, the images from Ck and Ck+ database are used. 320 images of different expressions are used for training. For testing of the algorithm, we used the images which are captured using webcam. 50 to 60 different set of images of different persons are used in testing. Multi SVM classifier is used for classification of different emotions. One –vs-All SVM classifier is used for training of different classes of expressions. Dlib is used to extract the facial features, the experimental results shows the detection of different emotions. PCA is used to reduce the dimensionality; it finds the small number of eigenfaces. These eigenfaces should span a space that is required to represent a face.

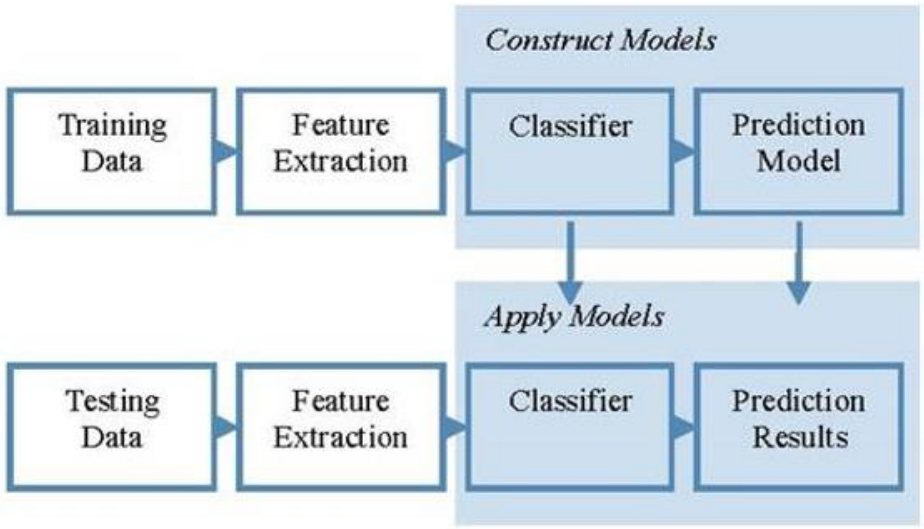

Figure 1. Emotional facial classification system

LEARNING METHODS:

A:

Support Vector Machine As we know, Support vector machines (SVM) [31] is recognized as one of the best learning methods for the databases in general cases in these years. Sequential Minimal Optimization – SMO is used to train SVM. The memory requirement for SMO is linear in the size of training data set. The computation time of SMO depends on SVM evaluation, thus SMO is fastest for linear SVM and data sets with large sparsity.

B:

Deep Boltzmann machine shallow model need to rely on artificial experience to select features of samples, the inputs of the model are these features that have been selected, the model is only responsible for classification and prediction. For the shallow model, the pros and cons of the

model are not usually the most important, but the pros and cons of selecting features are. Therefore, most researchers are devoted into the feature. extraction and feature selection, it not only need us to have a deep understanding on the domain of the thesis, but also spend a lot of time to explore repeated experiments, which also limits the performance of the shallow model.

C:

Fusion Method In order to improve the performance of system, a simple fusion method is applied. The details of simple fusion can be found in . Weights are assigned to a few sub-systems in training process, and the weighted combination of subsystems results is the final result. For classification, the probability of a sample in each class is combined from different sub-systems. And the formula of simple fusion is described as follows:

$$s(t) = \sum_{i=1}^{l} \alpha_i * s_i(t)$$

where l is the number of subsystems, si(t) is the prediction results of ith subsystem, a is the weight to be optimized, and is the final prediction results.

PREDICTION SYSTEM ON EMOTIONAL FACIAL FEATURE DATA: The emotional facial feature data is from FERA 2015 (Second Facial Expression Recognition and Analysis Challenge). And the data onto implementation is from SEMAINE database, which is used for the classification of Occurrence Detection Sub-Challenge to detect the occurrence of Action Unit (AU). The Matthews Correlation Coefficient (MCC) is used to represent inter-coder reliability. Only AUs where MCC is larger than 0.6 can be selected, as a result, AU2, AU12, AU17, AU25, AU28 and AU45 are selected for implementation from SEMAINE database. For feature extraction, both geometric features and appearance features are used.

The appearance features are extracted by LGBP, and it is generated by applying LBP to Gabor feature. Classifier using different machine learning is implemented for the classification problem Occurrence Detection, which is described in Fig. 1. Firstly, the features of training data are used to train the prediction model of SVC. Besides, test data is inputted to the prediction model. Finally, the prediction results of testing data can be obtained. The detection of occurrence status of the five AUs (AU2, AU12, AU17, AU25, AU28 and AU45) is a binary classification problem. Specifically, label 0 indicates this AU didn't occur in this frame sample, and label 1 indicates

this AU occurred in this frame sample. And this classification system is implemented several times for different AUs. Besides, the kernel function of SVC is set as polynomial kernel. For parameters of kernel function, the polynomial kernel is set as 0.001 and the cost of c-SVC is set as 100 for implementation.

IMPLEMENTATION AND RESULTS:

A. Evaluation Criteria

Here the evaluation criteria are illustrated in two-class cases. The TP, FN, TN, FP are defined in Table

TABLE I. EVALUATION CRITERIA OF TWO-CLASSES CASES

	Predicted Positive	Predicted Negative
Actual Positive	TP (True Positive)	FN (False Negative)
Actual Negative	FP (False Positive)	TN (True Negative)

Results of Emotional Facial Data

The results of classification system of Occurrence Detection of AUs with SVM, DBM and fusion is shown in Table II

TABLE II. CLASSIFICATION RESULTS OF OCCURRENCE DETECTION OF AUS WITH SVCAND DBM

	Precision	Recall	F1	UAR	Accuracy
Average of SVC	0.456	0.191	0.237	0.573	0.857
Average of DBM	0.103	0.025	0.030	0.504	0.893
Fusion	0.481	0.236	0.359	0.602	0.910
Var of SVC	0.044	0.026	0.035	0.005	0.012
Var of DBM	0.007	0.002	0.002	0.00008	0.008
Var of Fusion	0.052	0.032	0.040	0.008	0.016

The experiment performance of some AUs is much lower than others. It is because that the frame samples are randomly selected from all video frames and the distribution of dataset is extremely

unbalanced. The information selected from the whole dataset is not enough to learn the template from the video of different volunteers. For future work, the distribution of dataset need to be further improved and dimension reduction of PCA might also need to be implemented for processing more information of the dataset.

As the precision of recall computed by class of label 1, when there is extremely little label 1 in training dataset, the learning machine cannot learn the characteristics of this class. And also, we can see from the different experiment results between SVC and DBM in this challenge, SVM is more robust for extremely unbalanced data. The fusion result is better than any of classification result by any method (SVM or DBM).

The algorithm is designed in such a way that, if the person is recognizing for the first time then the system considers him as a new user and performs each step of operatio. But if the person/user data is already stored then it is considered as "Registered user" and it performs only matching operation to recognize the user identity. Open CV contains cascade classifiers in which Viola & Jones face detection algorithm is implemented. By using these classifiers the face region is detected from the image. It classifies the images into positive and negative images respectively.

The images consists of face region is considered as positive and the images without face as negative images. These negative images are ignored for further processing. The stored images consists of face images of dimension 273x273, the more number of the images higher the recognition rate. If the stored images consists false images or wrong extensions then recognition is not possible. Therefore care should be taken while capturing input images. Fisherface algorithm is applied for classification of different users. Fisherface algorithm generates the fisherfaces of each image that are used for recognition. In Fisherface algorithm, it performs "leave-one-out" cross validation to validate the user identification.

negative emotion-recognition results using the ranking and average of the scores from seven emotion-recognition results. We developed an emotion-recognition model using deep learning's Convolutional Neural Networks (CNN) to develop this app and proposed a method for recognizing emotions. Thus, in this study, we classified the calculated emotion-recognition scores into positive and negative emotions and implemented an app that provides the user with scores for seven positive and negative emotions

CONCLUSIONS

In this paper, we proposed an emotion-recognition method using facial images and implemented an app that provides seven emotion and positive and negative emotion-recognition results to users. As a result, when we applied those recognition methods into apps, application performance rate was 50.7% in seven emotions and in positive and negative was 72.3%. In the future, we will improve the recognition rate by adding more emotional databases and modified some parts of deep-learning algorithm. In addition, our research will be carried out to recognize the user's intention as well as the current user's emotion recognition.

Face recognition which is implemented in real-time helps to recognize the human faces can be used for person identification and authentication purposes.

This report has implemented several learning methods: SVM and DBM, which are all excellent methods in general and the aim is to construct the prediction system which is most suitable to the challenge.

Application :

This report has implemented several learning methods: SVM and DBM, which are all excellent methods in general and the aim is to construct the prediction system which is most suitable to the challenge facial classification system with SVM. This paper mainly focuses on the implementation and caparison of different learning methods. For future work, the prediction results of facial data and audio data should be merged, and also the prediction results of different learning methods should also be merged to enhance the performance of the prediction system. Besides, there are numerous decision fusion methods, and only the simple fusion is implemented in this paper. For future work, more fusion methods need to be learned for their principles, and more fusion methods need to be implemented and compared.

REFERENCES

1. Z. Zeng, Y. Hu, M. Liu, Y. Fu, and T.S. Huang, “Training Combination Strategy of Multi-Stream Fused Hidden Markov Model
2. G. Caridakis , L. Malatesta, L. Kessous, N. Amir, A. Paouzaiou, and K. Karpouzis, “Modeling aturalistic Affective States via Facial and Vocal Expression Recognition,” Proc. Eighth ACM Int’l Conf. Multimodal Interfaces (ICMI ’06), pp. 146-154, 2006.

ACKNOWLEDGEMENT

I would like to place on record my deep sense of gratitude to Prof. _. **S.B. Kalyankar** , HOD-Dept. of Computer Science and Engineering, Deogiri Institute of Engineering and management Studies Aurangabad, for his generous guidance, help and useful suggestions.

I express my sincere gratitude to Prof. P.H.Dulure, Dept. of Computer Science and Engineering, Deogiri Institute of Engineering and management Studies Aurangabad, for his stimulating guidance, continuous encouragement and supervision throughout the course of present work.

I am extremely thankful to **. S.B. Kalyankar** , Director, Deogiri Institute of Engineering and management Studies Aurangabad, for providing me infrastructural facilities to work in, without which this work would not have been possible.

Signature of Student

Name of Student Sign