International Conference on Recent Advances in Computer Science, June 2021
Dept. of Computer Science & Engineering
Saveetha Engineering College ,Chennai, India.

FACIAL RECOGNIZATION

[1] Wuriti.Kavya, [2] Prabhakar Marry , [3] M.Swathi
[1] [2] [3]Assistant professor, Vignan Institute of Technology & Science,
[1] kavyawuriti@gmail.com [2] marryprabhakar@gmail.com [3] swathi.mandava599@gmail.com

ABSTRACT

Facial expressions convey non-verbal cues, which play an important role in inter-personal relations. Automatic recognition on Facial expressions can be an important component of natural human machine interfaces; it may also be used in behavioral science and clinical practices. It can also be improvised to be used in applications like a dynamic music to play songs in ones' playlist based on ones' emotion. The proposal is to detect an emotion based on facial expressions. It is done by taking an image as an input and classifying player the emotion of the face in the image. It categorizes the image into one of the basic prototype expressions namely, Happiness, Sadness, Anger, Disgust and Neutral (Class Labels).

The working involves three major steps: Face Detection, Feature Extraction and Classification. The dataset used for the whole process is Cohn-Kandace (CK) Dataset. OpenCV Library in Python is used for Face Detection which has four pre-trained classifiers. The vital facial parts like corners of mouth, cheeks, eyes and the area around are considered for Feature Extraction.

INTRODUCTION

A Facial expression is the visible manifestation of the affective state, cognitive activity, intention, personality and psychopathology of a person and plays a communicative role in interpersonal relations. It has been studied for a long period of time and obtaining the progress in recent decades. Though much progress has been made, recognizing facial expression with a high accuracy remains to be difficult due to the complexity and varieties of facial expressions. Generally human beings can convey intentions and emotions through nonverbal ways such as gestures, facial expressions and involuntary languages. This system can be significantly useful, nonverbal way for people to communicate with each other. The important thing is how fluently the system detects or extracts the facial expression from image. The system is growing attention because this could be widely used in many fields like lie detection, medical assessment and human computer interface. The Facial Action Coding System (FACS), which was proposed in 1978 by Ekman and refined in 2002, is a very popular facial expression analysis tool. On a day today basics humans commonly recognize emotions by characteristic features, displayed as a part of a facial expression. For instance happiness is undeniably associated with a smile or an upward movement of the corners of the lips. Similarly other emotions are characterized by other deformations typical to a particular expression.

Research into automatic recognition of facial expressions addresses the problems surrounding the representation and categorization of static or dynamic characteristics of these deformations of face pigmentation [8]. The system classifies facial expression of the same person into the basic emotions namely anger, disgust, fear, happiness, sadness and surprise. The main purpose of this system is efficient interaction between human beings and machines using eye gaze, facial expressions, cognitive modeling etc. Here, detection and classification of facial expressions can be used as a natural way for the interaction between man and machine. And the system intensity vary from person to person and also varies along with age, gender, size and shape of face, and further, even the expressions of the same person do not remain constant with time. However, the inherent variability of facial images caused by different factors like variations in illumination, pose, alignment, occlusions makes expression recognition a challenging task. Some surveys on facial feature representations for face recognition and expression analysis addressed these challenges and possible solutions in detail.

There are three major steps in a facial expression recognition system

1. To detect the face from the given input image or video.

2. To extract the facial features like eyes, nose, mouth from the detected face and

3. To classify the facial expressions into different classes like Happy, Angry, Sad, Fear, Disgust, and Surprise.

OBJECTIVE

The objective of the project is to recognize the facial expressions with the help of machine learning algorithms. This project makes use of fisher face classification algorithm to build the model and then classifies the new object into one of the five emotions: anger, disgust, happiness, neutral and sadness difficult.

In countries like Nepal the rate of crimes are increasing day by day. No automatic systems are there that can track person's activity. If we will be able to track Facial expressions of persons automatically then we can find the criminal easily since facial expressions changes doing different activities. So we decided to make a Facial Expression Recognition System. We are interested in this project after we went through few papers in this area. The papers were published as per their system creation and way of creating the system for accurate and reliable facial expression recognition system. As a result we are highly motivated to develop a system that recognizes facial expression and track one person's activity.

SCOPE AND APPLICATIONS

The scope of this system is to tackle with the problems that can arise in day to day life. Some of the scopes are:
1. The system can be used to detect and track a user's state of mind.
2. The system can be used in mini-marts, shopping center to view the feedback of the customers to enhance the business
3.The system can be installed at busy places like airport, railway station or bus station for detecting human faces and facial expressions of each person. If there are any faces that appeared suspicious like angry or fearful, the system might set an internal alarm.
4.The system can also be used for educational purpose such as one can get feedback on how the student is reacting during the class.
5.This system can be used for lie detection amongst criminal suspects during interrogation.
6.This system can help people in emotion related-research to improve the processing of emotion data.
7.Clever marketing is feasible using emotional knowledge of a person which can be identified by this system.

LITERATURE SURVEY

Recent work in the computer vision field has paid attention to the recognition and classification of facial expressions. Methods by which a computer can recognize visually communicated facial actions-expressions can contribute to the effective human-computer interaction ant to the applications such as very low bit rate visual communication and face recognition from dynamic imagery. Thorough research denoted that at least five emotions conveyed by human faces are universally associated with distinct facial expressions. In static approaches, recognition of a facial expression is performed using a single face image. The process of facial expression recognition is classified into three stages: 1. Pre-processing of input images 2. Face detection 3. Facial expression recognition and classification This literature survey section gives an outline of the techniques and algorithms used to implement the above mentioned stages of facial expression recognition.

FACE DETECTION

For facial feature extraction in static images first a static image is taken. Then, skin region is extracted from that image using Hue Saturation Value (HSV). After extraction, the eyes and the mouth part are isolated and the unwanted skin region is removed. Viola-Jones face detection

algorithm popularly known as cascade face detector is used for face detection. Another face recognition method "expression removal vector" is proposed by [4] that can be integrated into any face and expression recognition system to improve the overall recognition accuracy even under limited training samples. An efficient algorithm for accurate detection and extraction of facial features has been proposed. The color image is converted into a gray scale image. Then pre-processing is performed by applying contrast equalization. To find the eye, the dark regions are found using two methods: 1) rule detection method and 2) morphological method. The face region is divided into a number of windows by applying the geometric relationship of features where each window contains a facial feature such as eyes, nostrils and mouth etc.

FACIAL EXPRESSION RECOGNITION

There are two types of face representations for face expression recognition i.e. holistic template based methods and geometric feature based method. In holistic method the whole face image is processed to obtain a template which is either a pixel image or a feature vector. In the geometric feature based method, the shape and location of facial components are used to obtain feature vector.The Local binary pattern (LBP) summarizes local structures of images efficiently by comparing each pixel with its neighbouring pixels. Another method named "Thermal Image Processing and Neural network" for recognition of facial expression is based on 2D detection of temperature distribution of face, using Infrared Rays. The front view face is first normalized in terms of size and location and the local temperature difference between averaged neutral and unknown expression faces are measured.

In Fuzzy logic based emotion recognition system the system compares an image of interest to a neutral expression baseline image to determine the expressed emotion. The system consists of two stages: 1) Image processing and 2) emotion recognition. In first stage the subjects face and facial features are extracted and then the relevant identifying points are extracted from each feature using which relevant information for each AU is calculated. In the second stage the identifying points are used to fuzzify and determine the strength of different facial actions. These strengths are then used to determine the subject's emotion

FACIAL EXPRESSION CLASSIFICATION

Six categories of emotions are used in the classification stage: neutral, joy, anger, surprise, sadness and disgust. Facial expressions can be classified by using some face parts such as eyes and mouth. Different classification methods are used to classify facial expressions in still images. A classifier also known as distance metric or matching criteria, is used to retrieve the similar face image from a large dataset. Some types of classifiers are Euclidean distance, histogram intersection distance, chi square distance, Support Vector Machine (SVM) etc. Euclidean Distance measures the summation of difference among the paired values of the feature set. After taking the square root of the summation the closest distance measure is taken as the final result for that particular image. The histogram intersection distance is a very simple method and is very useful in similarity measure where a large database is involved and quick replies are required.

The classification algorithm is divided into three phases of processing: training, validation and testing. In the training phase, the important characteristic properties of the image features are taken. The input images and the database images can be trained by using a classifier. They are then compared to find out the final expression. Using Support Vector Machine (SVM) classifier, images are trained by finding the Eigen faces which is detected by calculating the Eigen values.

DATABASES

Databases:The databases used in the above mentioned methods are
Cohn-Kanade database: This database consists of 593 images and 7 expressions: neutral, sadness, surprise, happy, anger, fear and disgust. The images present in this database are mostly gray and it is a database consisting of American faces.123 subjects were used to create this database

JAFFE database: This database consists of 213 images and 7 expressions: neutral, sadness, surprise, happy ,anger ,fear and disgust. The images present in this database are all gray and it is a database consisting of Japanese faces. 10 subjects were used to create this database

FISHER'S CLASSIFIER

Principal component analysis (PCA) is a statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables (entities each of which takes on various numerical values) into a set of values of linearly uncorrelated variables called principal components. If there are n observations with p variables, then the number of distinct principal components is min(n-1,p). This transformation is defined in such a way that the first principal component has the largest possible variance (that is, accounts for as much of the variability in the data as possible), and each succeeding component in turn has the highest variance possible under the constraint that it is orthogonal to the preceding components. The resulting vectors (each being a linear combination of the variables and containing n observations) are an uncorrelated orthogonal basis set. PCA is sensitive to the relative scaling of the original variables.

Linear Discriminant Analysis, a statistical method often used for dimensionality reduction and classification. It was invented by the great statistician Sir R. A. Fisher, who successfully used it for classifying flowers in his 1936 paper"The use of multiple measurements in taxonomic problems". The PCA finds a linear combination of features that maximizes the total variance in data. While this is clearly a powerful way to represent data, it doesn't consider any classes and so a lot of discriminative information may be lost when throwing some components away.

This can yield bad results, especially when it comes to classification. In order to find a combination of features that separates best between classes the Linear Discriminant Analysis instead maximizes the ration of between-classes to within-classes scatter. The idea is that same classes should cluster tightly together.

The Fisher faces algorithm basically goes like this:

- Construct the Image matrix X with each column representing an image. Each image is a assigned to a class in the corresponding class vector C.
- Project X into the (N-c)-dimensional subspace as P with the rotation matrix WPca identified by a PrincipalComponentAnalysis, where
- N is the number of samples in X
- c is unique number of classes (length(unique(C)))
- Calculate the between-classes scatter of the projection P as

  Sb =\sum_{i=1}^{c} N_i*(mean_i - mean)*(mean_i - mean)^T, where
- mean is the total mean of P
- mean_i is the mean of class I in P
- N_i is the number of samples for class i
- Calculate the within-classes scatter of P as Sw = \sum_{i=1}^{c} \sum_{x_k

  \in X_i} (x_k - mean_i) * (x_k - mean_i)^T, where
- X_i are the samples of class i
- x_k is a sample of X_i
- mean_i is the mean of class i in P
- Apply a standard Linear Discriminant Analysis and maximize the ratio of the determinant of between-class scatter and within-class scatter. The solution is given by the set of generalized eigenvectors Wfld of Sb and Sw corresponding to their eigenvalue. The rank of Sb is atmost(c-1), so there are only(c-1) non-zero eigen values, cut off the rest.
- Finally obtain the Fisher faces by W = WPca * Wfld.

SOFTWARE AND HARDWARE REQUIREMENTS

The following are the software and hardware requirements for the smooth execution and running of the project

SOFTWARE REQUIREMENTS

The project is implemented using python and its libraries and the data set is also arranged using python.

PYTHON

Python version higher than 2.7 can be used.The version used is 3.7.3.

LIBRARIES

- Numpy
- OpenCV
- Matplotlib
- Glob

HARDWARE REQUIREMENTS

Recommended Operating Systems

- Windows 7 or later
- Linux Ubuntu
- MacOS X orlater

PROCESSOR POWER

Processor: Minimum 2 GHz; Recommended 2.3GHz or more.

Storage: Minimum Storage of 32 GB is required. Free storage available should be 1GB.

Memory: Memory of 2GB should be present.4 GB recommended.

IMPLEMENTATION

This project uses python scripting language and Anaconda prompt for implementation.

DATASET

Itmakesuse ofCohn-Kanndedataset(CK-dataset),alsoknownasCMU- Pittsburgh AU-Coded facial expression database consisting of 486 images of 97 adults with ages between 18 and 50 years. 69% are female and 31% are male. The subjects posed for the basic expressions like anger, fear, disgust, contempt, neutral, surprise, sadness and happiness.

PRE-PROCESSING AND FACE DETECTION

OpenCV is used for image pre-processing and face detection. It is the leading open source library for computer vision, image processing and machine learning, and now features GPU acceleration for real-time operation. The pre-processing includes converting the images into grey scale and then resizing all the images to a single specified size.In order to detect faces in the images, cascade classifiers of OpenCV are used. The word “cascade” in the classifier name means that the resultant classifier consists of several simpler classifiers (stages) that are applied subsequently to a region of interest until at some stage the candidate is rejected or all the stages are passed.

The basic classifiers are decision-tree classifiers with at least 2 leaves.

Haar-like features are the input to the basic classifier. The feature used in a particular

classifier is specified by its shape, position within the region of interest and the scale (this scale is not the same as the scale used at the detection stage, though these two scales are multiplied).Cascading is a particular case of ensemble learning based on the concatenation of several classifier, using all information collected from the output from a given classifier as additional information for the next classifier in the cascade. Unlike voting or stacking ensembles, which are multi expert systems, cascading is a multistage one.

Cascading classifiers are trained with several hundred "positive" sample views of a particular object and arbitrary "negative" images of the same size. After the classifier is trained it can be applied to a region of an image and detect the object in question. To search for the object in the entire frame, the search window can be moved across the image and check every location for the classifier.

TRAINING AND CLASSIFICATION

For training and classification, Fisher face algorithm is used. Fisher face is one of the popular algorithms used in face recognition, and is widely believed to be superior to other techniques, such as eigenface because of the effort to maximize the separation between classes in the training process. Image recognition using fisherface method is based on the reduction of face space dimension using Principal Component Analysis (PCA) method, then apply Fisher's Linear Discriminant (FDL) method or also known as Linear Discriminant Analysis (LDA) method to obtain feature of image characteristic.

FEATURE GENERATION PROCESS WITH FISHER FACE METHOD

Assumed-Size of rectangular face image with height = N and width = N and consists of h samples image, (each image is represented by a vector) and C classes.

PCA ALGORITHM

- Conversion training image with size N x N into vector form with length size N
- Calculate the average of all face images
- Calculate A = [a-m,b-m,...,h-m]
- Compute vector Eigen (eigVecs) and value eigen (eigenVals) by using the method svd of the matrix A. Sort eigvecs then reduction with the pca method, pe . pe is Eigen faces.

LDA ALGORITHM

- Calculate the average of each person / class
- Construct the scatter matrix S1, S2, S3, S4
- The construct of also matrix between class scatter, (ScatB)
- Compute the multiplication of matrices tranpose of pe , (PeT), with ScatW and ScatB until obtain :
  Sww = PeT * ScatW *

Pe Sbb=

PET

*ScatB*Pe

- Find eigenvector (VeSbb) and generalized eigenvalues (NeSww) of (Sbb, Sww) and then sort in ascending order.
- Output as Fisherfase◊Projection back VeSbb with Pe eigenfaces then formed (Pe *VeSbb)
- Normalization Fisherface Pe*VeSbb*N
- Find the transpose of the normalized Fisherfaces, Pe*VeSbb*Nt
- Calculate Weights for each training image into a normalized fisher face, U = Pe * VeSbb * Nt * A.

The result of the above process is the weight of each training image in the form of eigen vector which will be used to find similarity with face image which will be recognized by using Euclidean distance formula.

IDENTIFICATION OR RECOGNITION ALGORITHM

The identity classification steps are as follows

- Conversion of the face image tested by the size of N*N into the column vector form.
- Normalization of facial image input to the image of training by finding the value of different matrix phiinp by subtracting the average value of training image.

- Calculates the weight of the test image by multiplying the eigenval transpose matrix V^T with the matrix phiinp matrix ;
- Calculate the distance of the difference between the image testing with training face image using euclidean distance. The result of the identification is the image that has the smallest distance with the test image displayed by the system

PROCEDURE

The CK dataset is a combination of image files with different emotions. It also consists of emotion label files represented in floating point numbers. Therefore we need to organize the source images into their respective emotions. Once organized, equal proportion of images from each emotion is used to train the model.

The training step includes pre-processing the source images, detecting faces using cascade classifiers of OpenCV and then extracting the features for different emotions. These extracted features are associated with their respective emotion labels. For a given new image, the image is pre- processed , face is detected and then the emotion is identified. The output is the emotion label of the given image.

Datasets: The training dataset for different emotions is

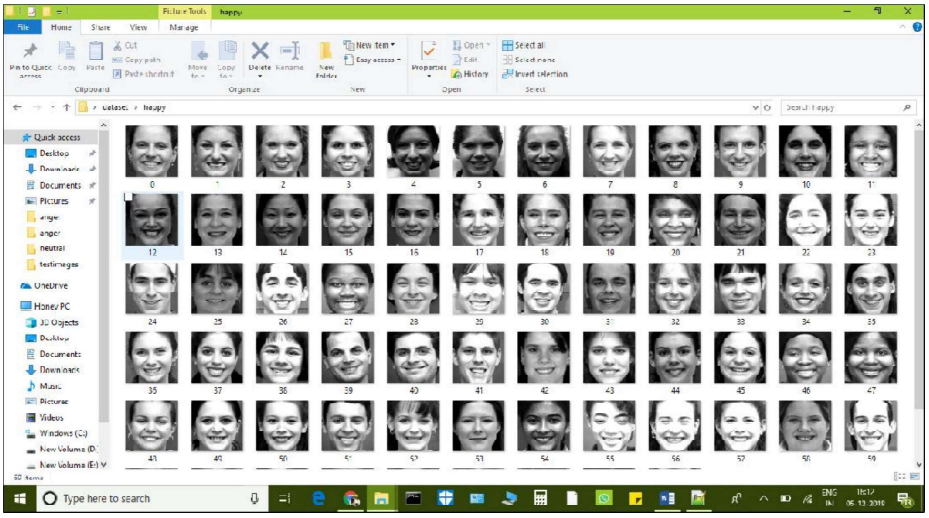

FIG:1 TRAINING DATASET 1

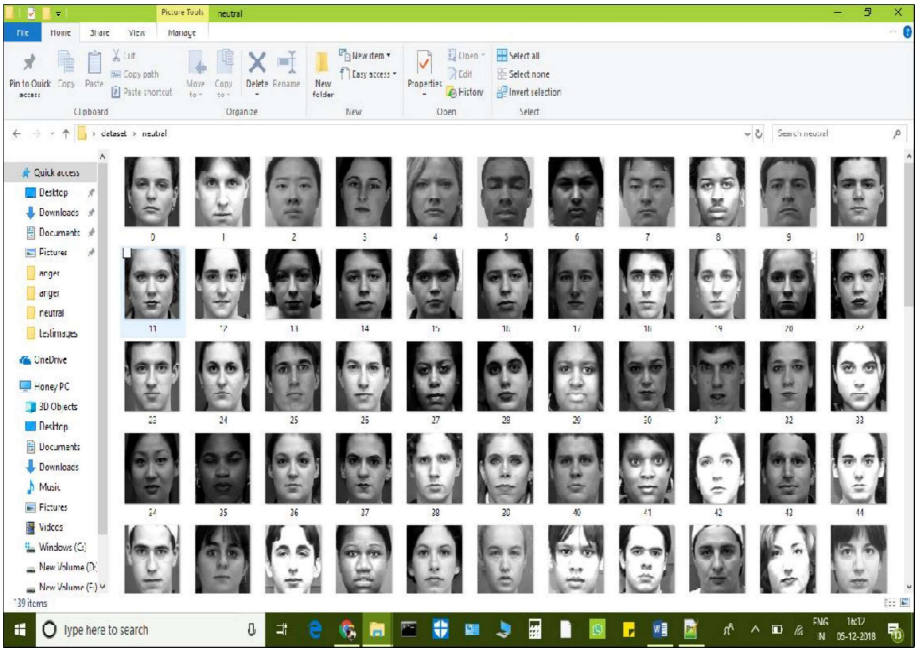

FIG 2. TRAINING DATASET 2

THE TESTING DATASET

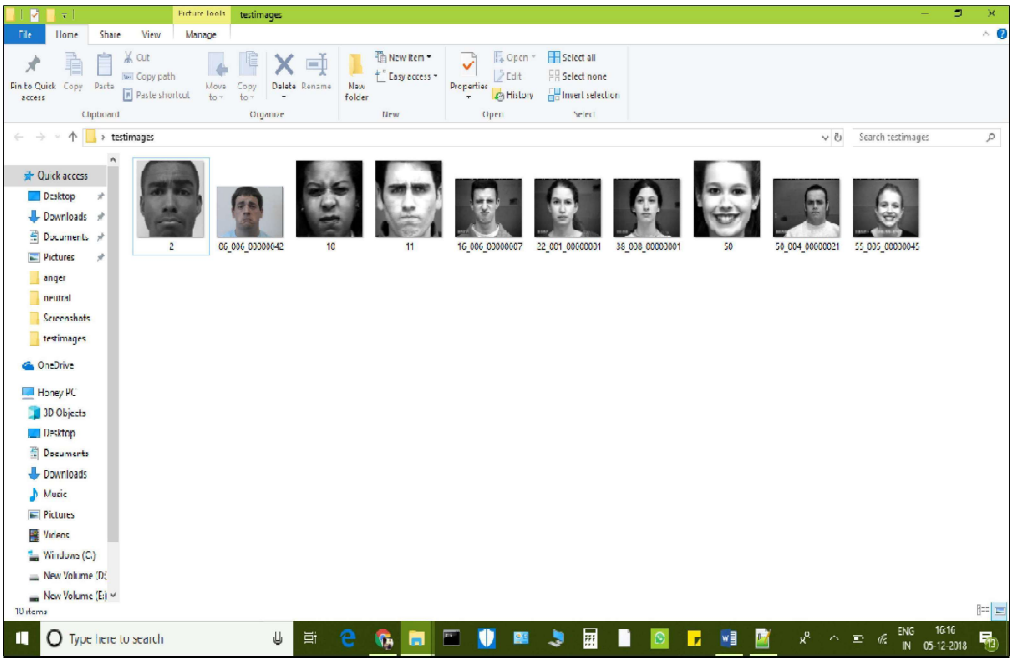

FIG:3 TRAINING DATASET 3

EXTRACTING FACES

The classifier will work best if the training and classification images are all of the same size and have (almost) only a face on them (no clutter). We need to find the face on each image, convert to grayscale, crop it and save the image to the dataset. We can use a HAAR filter from OpenCV to automate face finding. Actually, OpenCV provides 4 pre-trained classifiers, so to be sure we detect as many faces as possible let's use all of them in sequence, and abort the face search once we

have found one. Create another folder called "dataset", and in it create subfolders for each emotion ("neutral", "anger", etc.). The dataset we can use will live in these folders. Then, detect, crop and save faces as such.

The last step is to clean up the "neutral" folder. Because most participants have expressed more than one emotion, we have more than one neutral image of the same person. This could (not sure if it will, but let's be conservative) bias the classifier accuracy unfairly, it may recognize the same person on another picture or be triggered by other characteristics rather than the emotion displayed. Do this by hand: get in the folder and delete all multiples of the same face you see, so that only one image of each person remains.

CREATING THE TRAINING AND CLASSIFICATION SET

The dataset has been organised and is ready to be recognized, but first we need to actually teach the classifier what certain emotions look like. The usual approach is to split the complete dataset into a training set and a classification set. We use the training set to teach the classifier to recognize the to-be-predicted labels, and use the classification set to estimate the classifier performance. The reason for splitting the dataset: estimating the classifier performance on the same set as it has been trained is unfair, because we are not interested in how well the classifier memorizes the training set.

Rather, we are interested in how well the classifier generalizes its recognition capability to never-seen-before data. In any classification problem; the sizes of both sets depend on what you're trying to classify, the size of the total dataset, the number of features, the number of classification targets (categories). we randomly sample and train on 80% of the data and classify the remaining 20%, and repeat the process 10 times.

RESULTS

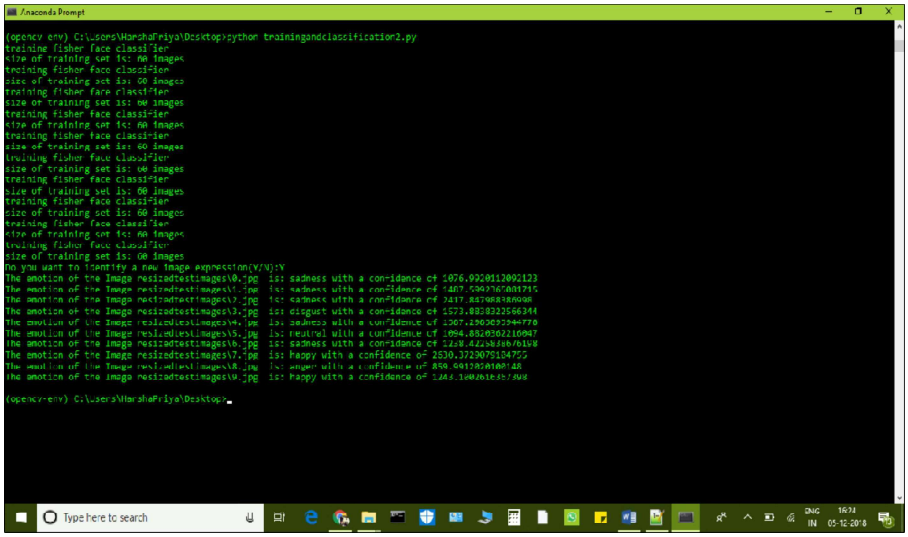

FIG: 4. OUTPUT SCREEN1

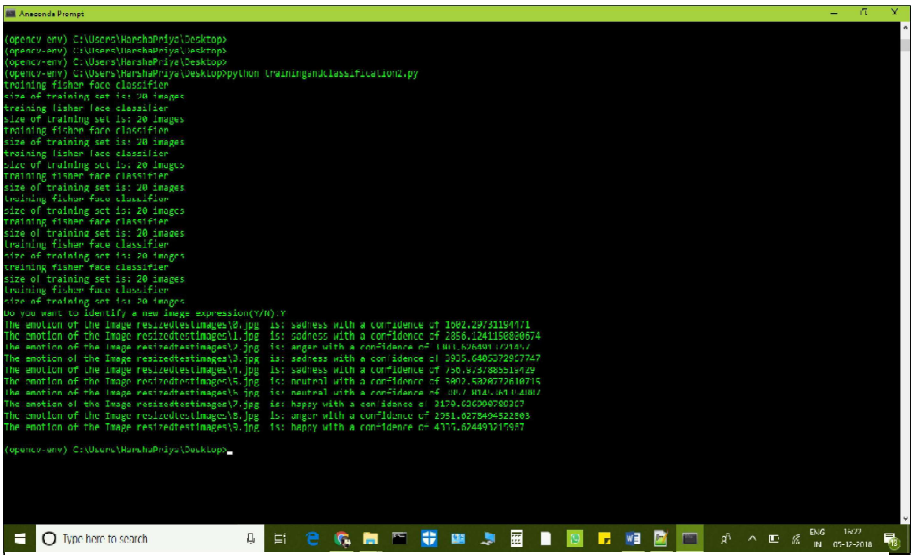

FIG: 5. OUTPUT SCREEN 2

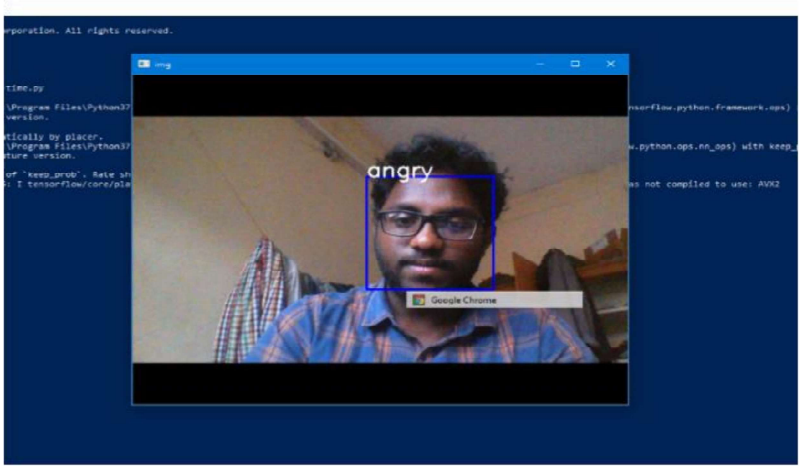

FIG: 6. OUTPUT SCREEN 3

CONCLUSION

This project proposes an approach for recognizing the category of facial expressions. Face Detection and Extraction of expressions from facial images is useful in many applications, such as robotics vision, video surveillance, digital cameras, security and human-computer interaction. This project's objective was to develop a facial expression recognition system implementing the computer visions and enhancing the advanced feature extraction and classification in face expression recognition.

In this project, we classify human emotions into 7 categories (5 for optimal results) by first detecting the face, feature extraction using fisher face, haar cascade filters.

The facial expression recognition plays a very important role in intelligence between human and computer. With the help of Fisher face algorithm, we are able to recognize the facial expressions such as anger, disgust, neutral, happiness and sadness. This project addressed the problems of how to detect a human face in static images and how to represent and recognize facial expressions present in those faces. More efforts should be made to improve the classification performance for important applications.

FURTHER ENHANCEMENTS

PERFORMANCE FOR A REAL-WORLD APPLICATION

This project might not be very reflective of a real-world application. The data set we use is very standardized. All faces are exactly pointed at the camera and the emotional expressions are actually pretty exaggerated and even comical in some situations.

It's clear that emotion recognition is a complex task, more so when only using images. Even for us humans this is difficult because the correct recognition of a facial emotion often depends on the context within which the emotion originates and is expressed.

FUTURE WORK

In this work, the proposed system deals with the static images. The future work will be dealing with other types of images such as image sequences and 3D images. Dealing with image sequence may require approaches with very low execution time. Also dealing with 3D images will need to start with creating 3D face models and require special pre-processing techniques.

REFERENCES

1. Andrew Ryan, Jeffery F. Cohn, Simon Lucey," Automated Facial Expression Recognition System", 50th Annual 2018 International Carnahan Conference on Security Technology.

2. Marcelo J. Cossetin, Julio C. Nievola, Alessandro L. Koerich,"Facial Expression Recognition using a Pairwise Feature Selection and Classification Approach",2017 International Joint Conference in Neural Networks,pp-5149-5155.
3. OpenCV libraries. Retrieved from https://opencv.org/releases.html

4. http://www.pace.ac.in/documents/ece/FACE%20RECOGNITION%20SYSTEM%20WITH%20FACE%20DETECTION

5. http://www.paulvangent.com/2016/04/01/emotion-recognition-with-python-opencv-and-a-face-dataset/