International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 05 Issue: 06 | June - 2021 ISSN: 2582-3930

Facial Expression Based Music Player

Anshu Mariyam Sam[1], Harshitha M[2], Rithik Sah[3], Himani Singh[4], Mr. Janardhan Singh[5]

[1, 2, 3, 4]UG student, Department of Computer Science and Engineering, Cambridge Institute of Technology, Bangalore, Karnataka, India
[5]Professor, Department of Computer Science and Engineering, Cambridge Institute of Technology, Bangalore, Karnataka, India

ABSTRACT

The human face is an important part of an individual's body and extracting the required input from the human face can now be done directly using a camera. One of the applications of this input can be for extracting the information to deduce the mood of an individual. This data can then be used to get a list of songs that comply with the mood derived from the input provided earlier. This eliminates the time-consuming and tedious task of manually segregating or grouping songs into different lists and helps in generating an appropriate playlist based on an individual's emotional features. Various algorithms have been developed and proposed for automating the playlist generation process. The expressions of a person are detected by extracting the facial features using the HAAR classifier and Fisherface algorithm. The results show that the proposed system achieves up to 88.82% of accuracy level in recognizing the expressions.

***Key Words*:** mood, HAAR classifier, Fisherface, emotional feature, expression.

1. INTRODUCTION

Listening to music nowadays has become a day to day activity. There are a huge number of categories of music a person to listen to. Human emotions are related to music so much since we choose to listen to a song that relates to our mood at a particular time. Facial expressions are a great indicator of the state of mind of an individual. It is indeed the most natural and basic way to express emotions.

The main objective of this work is to develop an intelligent system that can easily recognize the facial expression and accordingly play a music track based on that particular expression/emotion recognized. This model predicts emotions like Happy, Sad, Anger and Neutral.

The algorithm that is used in detecting the expressions is HAAR classifier and Fisherface Algorithm. The designed algorithm is very efficient due to less computational time taken hereby increasing the performance of the system. Songs are stored in the computer memory and combine the two to generate a user customized music playlist.

The following segment gives a review of a portion of the present work done for music and mood arrangement.

1.1 Related Works

Several solutions have been proposed in previous works given in [1], [2], [3], [4], and [5] for improving the facial expression process. The summary of each solution is discussed as follows.

In paper [1], **Facial Expression Based Music Player** by Prof. Jayshree Jha, Akshay Mangaonkar, Deep Mistry, Nipun Jambaulikar, Prathamesh Kolhatkar. One of the applications of this input can be for extracting the information to deduce the mood of an individual. This data can then be used to get a list of songs that comply with the „mood" derived from the input provided earlier. This eliminates the time-consuming and tedious task of manually segregating or grouping songs into different lists and helps in generating an appropriate playlist based on an individual's emotional features. Various algorithms have been developed and proposed for automating the playlist generation process. Facial Expression Based Music Player aims at scanning and interpreting the data and accordingly creating a playlist based the parameters provided. Extracting the required input from the human face can now be done directly using a camera.

In paper [2], **Facial Expression Based Music Player** by Sushmita G. Kamble and Associate. Prof. A. H. Kulkarni.To achieve this goal, an algorithm is used to classify the human expressions and play a music track as according to the present emotion detected. It reduces the effort and time required in manually searching a song from the list based on the present state of mind of a person. An inbuilt camera is used to capture the facial expressions of a person which reduces the designing cost of the system as compared to other methods. The results show that the proposed system achieves upto 84.82% of accuracy level in recognizing the

expressions.

In paper [3], **Face Player: Facial Emotion Based Music Player** by Pushkar Mahajan. Pratik Khurad, Prateek Chaudhari. Computer vision techniques are used in many fields such as traffic control, event monitoring, marketing, healthcare field, qu ality control, military technology, etc. One of the sub-areas of computer vision is facial expression recognition. Human face acts as the main indicator for the behavioral and the emotional state of the individual. Facial expressions which can be classified as fear, happiness, joy, sadness, aggressiveness are recognizable with computer vision techniques. Human reacts and responds to music and this music has a high impact on a person's brain activity.

In paper [4], **EMO PLAYER: Emotion Based Music Player** by Hemanth P,Adarsh, Aswani C.B,Ajith P, Veena A Kumar. This project Emo player (an emotion based music player) is a novel approach that helps the user to automatically play songs based on the emotions of the user. Evaluation allows the testing of the model against data that has never been seen and used for training and is meant to be representative of how the model might perform when in the real world. According to the emotion ,the music will be played from the predefined directories.

In paper [5], **Emotion Based Music Player** by Vinayak Bali, Shubham Haval , Snehal Patil, R. Priyambiga. istening to music affects the human brain activities. Emotion based music player with automated playlist can help users to maintain a particular emotional state. This research proposes an emotion based music player that creates a playlists based on captured photos of the user. The most important goal is to make change the mood of person if it is a negative one such as sad, depressed. This model is validated by testing the system against user dependent and user independent dataset.

1.2 Proposed System

The proposed system goes through various stages in order to get the desired result:

Initially an image is uploaded as an input of size NxN. Then it undergoes pre-processing technique that is conversion of RGB images into grayscale images to enhance the appearance quality of the input sample, it also resizes the dimensions of these sample images in order to create a lower dimensional space. Here the images are resized to (N^2)x1. Also in this step, the noise is removed from the input image either by using some filters or by applying median filtering. With this an image will preserve the edges which help in localizing the faces.

Then these images are selected for the training set with dimensions (N^2)xM where M is the number of sample images.

Next an average face is found from this input image. Average face is a standard statistical mean calculated for every single pixel.

$$\varphi = \frac{1}{M}\sum_{N=1}^{M} r$$

Then, subtraction of these average faces is done from the faces in the training set and finally a matrix say B is created. Now, the covariance matrix of B is calculated that is BB'. The next step after calculating the covariance matrix is to calculate the eigenvectors of the matrix obtained.

$$\emptyset = r - \varphi$$

$$C = \frac{1}{M}\sum_{N=1}^{M} \emptyset * \emptyset'$$

= BB'

$$Bv = \mu v \qquad \text{(eigen Vector)}$$

The next step is to find the eigenfaces which is a weighted combination of some components or basis faces. These basis faces are differently weighted to represent different faces. And then the reduced eigenface space (subspace) is created by selecting a set of vectors that are multiplied by B.

$$u = \sum_{k=1}^{M} v * \emptyset$$

With this study, the calculations are significantly reduced from the order of number of pixels (N^2) to the number of images (M) in the training set.

All these steps are carried out for training phase and testing phase. Now in testing phase, the eigenfaces is calculated for the test image.

2. METHODOLOGY

There are different techniques which can be used for classification so as to fit properties of the database and which then leads to different outcomes with respect to actual ground truth data.

2.1 Model

In the following approach we convert audio files of different moods to numerical values using features of these files in the python language. Feature extraction method is applied to retrieve different kind of audio features such as harmonic features, spectral, rhythm, energy and chroma vectors of the audio files, and this paper explores through many algorithms based on classification so as to suggest a new approach to classify and detect moods.

2.2 Feature Selection

Not all features affect musical mood in the same way and hence we need to carefully select the features we intend to use for the purpose of mood detection from audio tracks. Audio features are put into 4 suitable dimensions namely: Dynamic, Harmony, Rhythm and Spectral. Using Feed Forward selection it is clear that spectral, dynamic and harmony features used together help achieve the best accuracy.

2.3 Classification

For classification the paper presents two approaches. First, to test and predict directly by using audio feature. Second, use the selected audio features to predict the valence and arousal values using SVM regression with 3 different kernels- Linear, Poly and rbf. These are then mapped into 2D Space to identify the mood.

2.4 Facial Mood Detection

For detecting a user's mood this paper makes use of facial expressions. The pre-trained HAAR frontal-face classifier and fisherface method is used for detecting a user's face on screen. Before training the model, it is necessary we preprocess and standardize the images by keeping only the face portion of the image and by turning it into black and white. Once the model is trained successfully it can be used to detect a user's mood. After detecting the user's mood confusion matrix is plotted and precision was then calculated using it.

This paper will study various algorithms based on classification to provide a clear methodology to i) classify songs into 4 mood categories and ii) detect users mood through his facial expressions and then combine the two to generate user customized music playlist.

3. MODELLING AND ANALYSIS

3.1 Hardware and Software Requirements

Table-1: Specification of the Hardware/Software used.

Software	Specification
Processor speed	2.30 GHz
RAM	8.0 GB
Hard Disk	1 TB
Processor	intel® Core i5-8300H
Software	Python

3.2 System Design

The steps involved in facial expression based music player is explained below: -

Step 1: User opens the app in desktop.

Step 2: The device gets access to the webcam.

Step 3: The desktop camera is invoked and a real time image of the user is captured.

Step 4: The Fisherface and HAAR classifier algorithm checks the presence of a face in the input and then generates the emotion and displays it on the screen.

Step 5: After the expression is classified successfully the appropriate playlist is selected and the song from the playlist are played.

Step 6: If no face is detected, expression not detected message is shown on the screen.

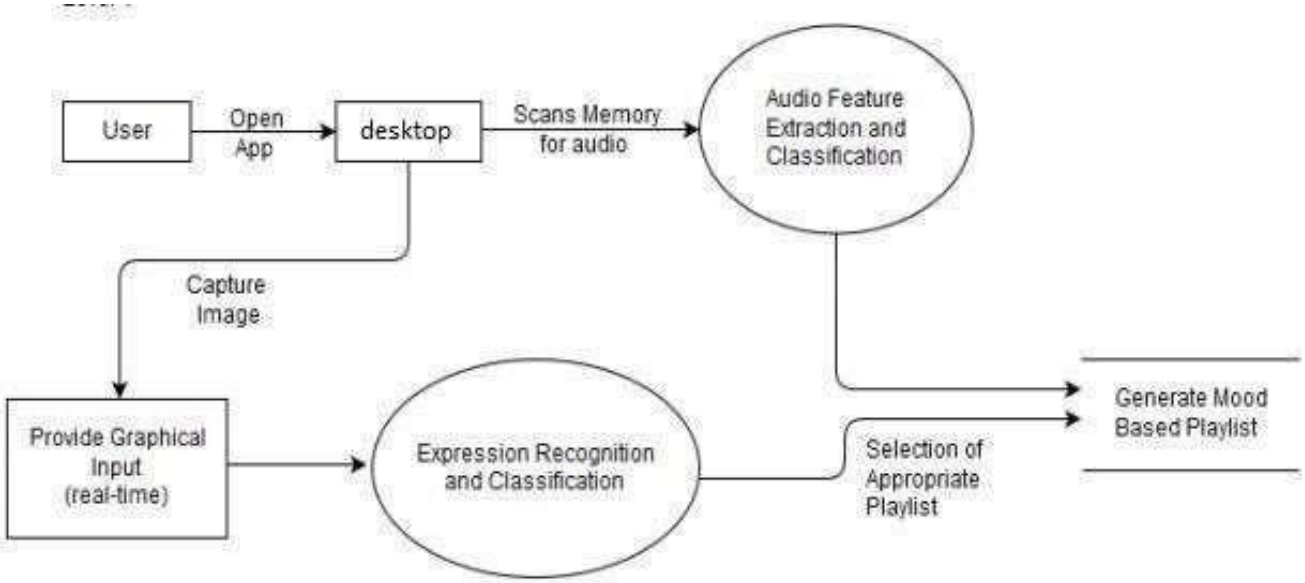

Figure-1: System Design

3.3 Implementation

Collecting data: Facial expression detection in Fisherface works with the help of trained models. Reason behind this is to allow user to take dataset according to their use. Suppose if we take a huge amount of dataset of around 25-30k it will give nice accuracy no doubt but if the situation is like that the user of the devices are a few people. Now in such condition if we take some precise dataset with around 400-450 images as input releted to the user then it will also give good accuracy with the benefit of less amount of dataset and less storage on memory to operate. As well as small memory of data give output fast which result in quick response time. Here we first tried with Cohn-Kanade dataset then we made some classification in the as our need make it to train our model.

Loading and saving trained model: For training, we have used Fisherface method of cv2 library. For training data model we have make a python code which grab all the classified images from folders and map it with it's emotion. These data we at an instance stored in dictionary and then use .train method of cv2 to train model. To save the model for later use we have implemented .save method of cv2. Now at the detection time first we have load model in memory using .read method. Prediction of result is based on the prediction and confidence value which .predict method return.

Haarcascade Model: Haarcascade model is precise face detection trained model which is provided by Open-cv. It return the co-ordinates in terms of (x, y) at (left, bottom) of face frame and it's width and height from those co-ordinates. The .detectMultiScale() method it is capable of detect multiple faces and it return an array of all the faces(co-ordinates) as an element. The arguments has set according to the threshold what we need for our checking purpose. We have set it such like it doesn't affect our model accuracy. Gray-scaling of images has been performed so as for better accuracy for the algorithm being used to detect the emotions.

Detecting Emotions: We have implemented the linking of python with javascript through eel library. Which provide us the privilege to access python methods from Javascript as well as vice versa. Here the striating flow will be in python code as the library is implemented in python then it transfer the control to html, JS. And according to the result we show emoticons.

4. RESULTS AND DISCUSSION

4.1 Web Application:

The user clicks on the python application hosting this web application and web application pops up in the preferred browser. In this case, Google Chrome. These actions are controlled by libraries used in Javascript. Users can select the mode he wants to play in. The three modes in this application are:

Queue based mode: Based on user's preferences, the user can select the songs the user wishes to play and add it to the queue. Based on the queue basis, the song is played. Done based on FIFO (First In First Out) Method.

Emotion based mode: In this mode, the user's camera's permissions are checked. The python code of the application is invoked. Next the application captures a image of the user and labels it as the test image. Next based on the emotion made by the user, the python code interprets the emotion and plays the song according to the emotion labelled based on input.

Random based method: The web application plays songs on a random basis. Random songs are selected from the backend of the application. All the methods are dynamic it can handle as change in number of songs accordingly.

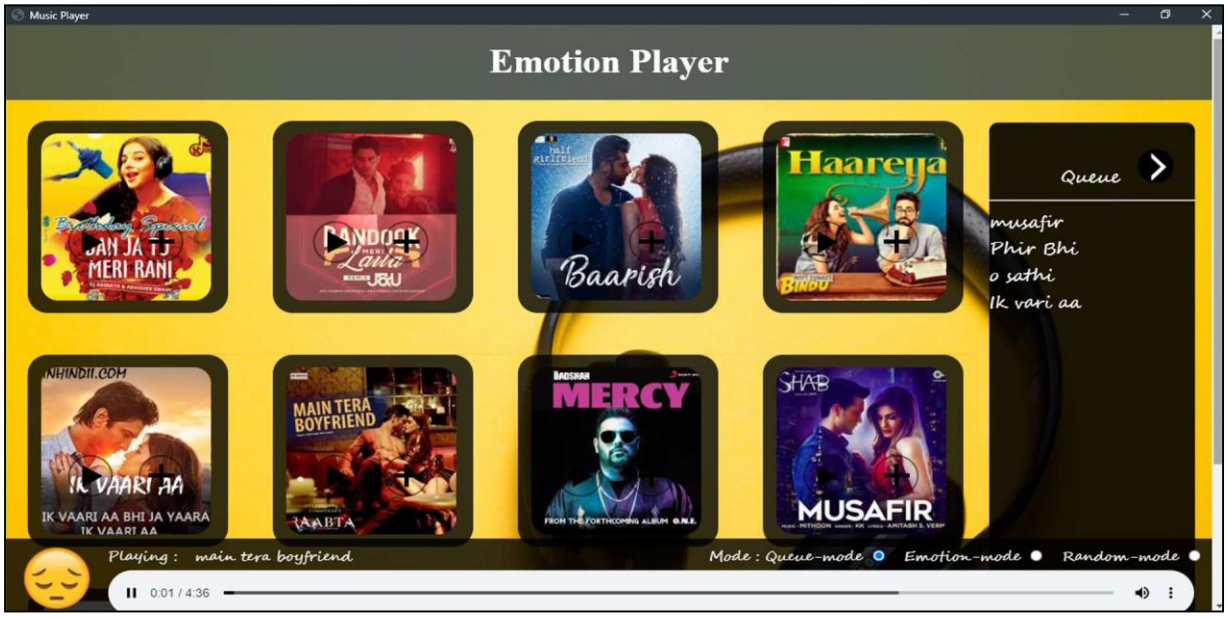

Figure-2: Queue based mode

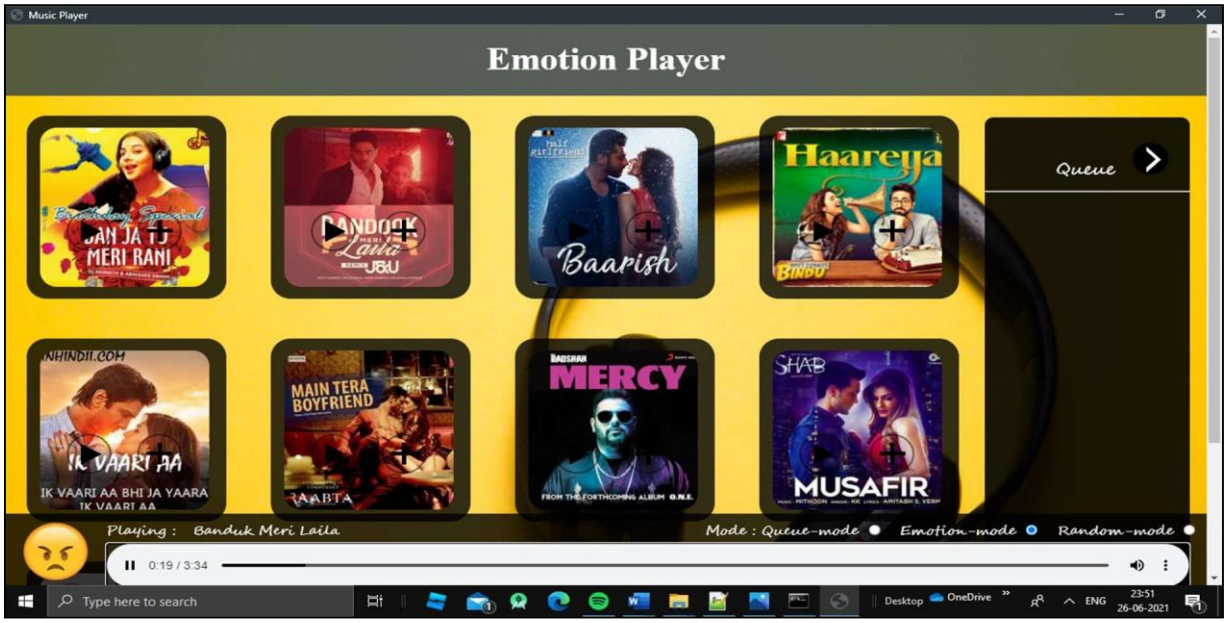

Figure-3: Emotion based mode

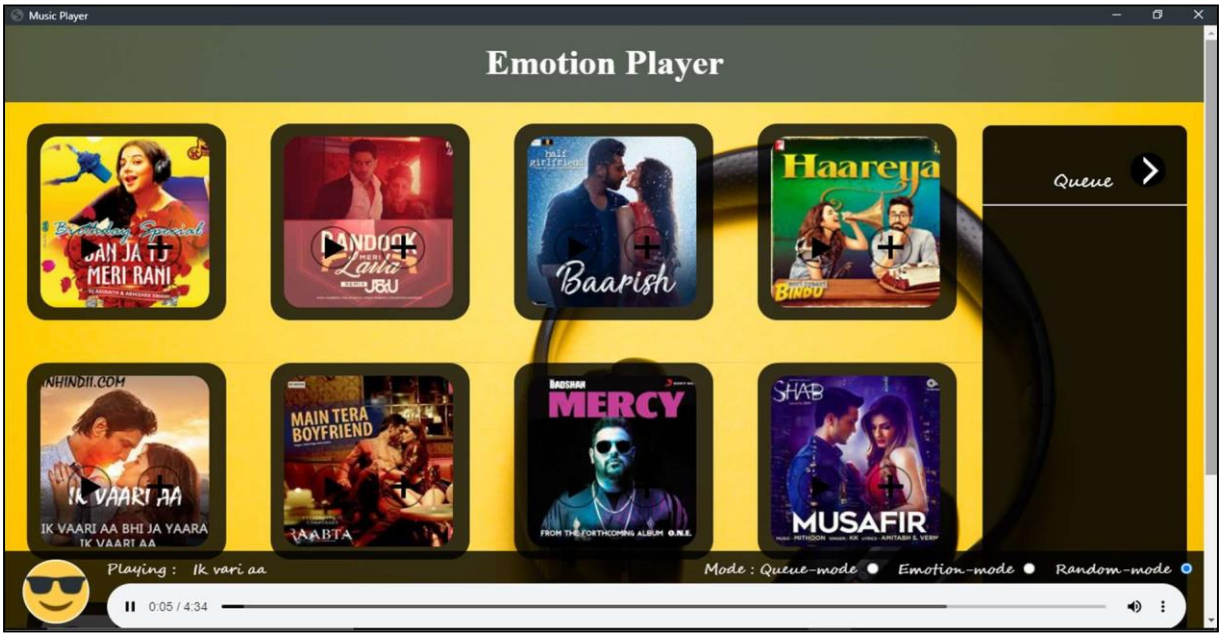

Figure-4: Random based mode

4.2 Command prompt

When the application is run from the command prompt, and the chosen mode is emotion based mode. The emotion displayed by the user is displayed in the terminal, based on which the python code is invoked and the song to the corresponding emotion is displayed.

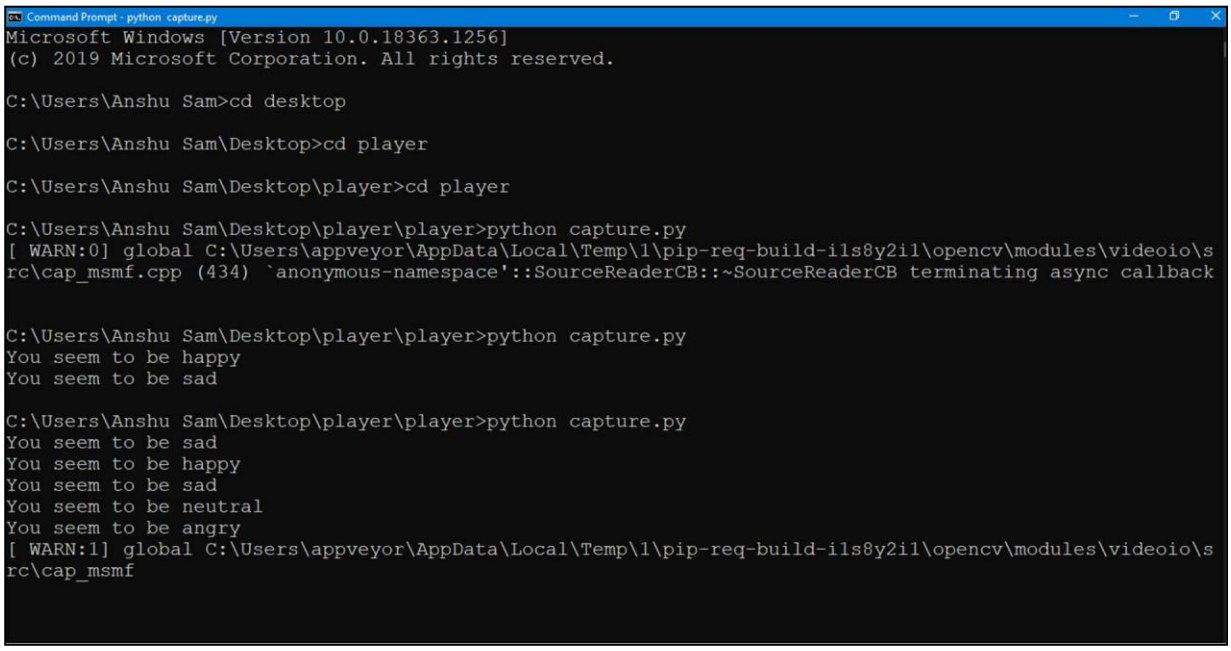

Figure-5: Terminal

4.3 Discussion

Facial expressions are one of the natural means to communicate the emotions and these emotions can be used in entertainment and Human Machine Interface (HMI) fields. In today's world, with the advancements in the areas of technology various music players are deployed with features like reversing the media, fast forwarding it, streaming playback with multicast streams. Although these features satisfy the basic requirements of the user, yet one has to manually surf for the song from a large set of songs, according to the current circumstance and mood. This is a time-consuming task that needs some effort and patience. The main objective of this work is to develop an intelligent system that can easily recognize the facial expression and accordingly play a music track based on that particular expression/emotion recognized. The future scope for the proposed system would be to implement it on mobiles. To design a mechanism that would help in the music therapy treatment for the music therapists to treat the patients suffering from mental stress, acute depression and trauma. It can also be used to determine the mood of a physically challenged person.

4. CONCLUSION

The proposed work presents facial expression recognition system to play a song according to the expression detected. It uses PCA approach to extract features, and classifier classifies these expressions. In this work, real images i.e. user dependent images are captured utilizing the in-built camera. For the task of user's mood detection, the model was trained with HAAR frontal-face classifier and fisherface algorithm. The final result shows the accuracy level obtained is upto 88.82%.

ACKNOWLEDGEMENT

We would like to express our sincere gratitude to all our friends and family members for their support and encouragement throughout the completion of the project. We would also like to thank our guide Mr. Janardhan Singh for his presence and his guidance throughout the entire project.

REFERENCES

[1] Prof. Jayshree Jha , Akshay Mangaonkar, Deep Mistry, Nipun Jambaulikar and Prathamesh Kolhatkar, "Facial Expression Based Music Player", International Journal of Advanced Research in Computer and Communication Engineering Vol. 4, Issue 10, October 2015, pp. ISSN (Online) 2278-1021 ISSN (Print) 2319 5940

[2] Sushmita G. Kamble and Asso. Prof. A. H. Kulkarni , "Facial Expression Based Music Player", pp. 2016 Intl. Conference on Advances in Computing, Communications and Informatics (ICACCI), Sept. 21-24, 2016.

[3] Pushkar Mahajan, Pratik Khurad and Prateek Chaudhari "Face Player: Facial Emotion Based Music Player", pp. 2020 IJRAR March 2020, Volume 7, Issue 1.

[4] Hemanth P,Adarsh,Aswani C.B,Ajith P and Veena A Kumar "EMO PLAYER: Emotion Based Music Player", International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 05 Issue: 04 | Apr-2018.

[5] Vinayak Bali, Shubham Haval, Snehal Patil, R. Priyambiga "Emotion Based Music Player", Journal of Software Engineering & Software Testing Volume 4 Issue 1 ISSN: 2457-0516.

[6] Anuja Arora, Aastha Kaul, Vatsala Mittal "Mood based Music Player",Computer Science Institute of Electrical and Electronics Engineers(IEEE) 978-1-5386-9436-7/19/$31.00 2019 IEEE.

[7] A. Lehtiniemi and J. Holm, "Using Animated Mood Pictures in Music Recommendation", In 16th International Conference on Information Visualisation.(2012).

[8] A. S.Dhavalikar and Dr. R. K. Kulkarni, "Face Detection and Facial Expression Recognition System" ,International Conference on Electronics and Communication System (ICECS -2014)