International Journal of Research in Engineering, Science and Management
Volume-2, Issue-4, April-2019
www.ijresm.com | ISSN (Online): 2581-5792

Implementation of Emotion based Music Player from Facial Image using SVM

Pratiksha Yadav[1], Jyoti Jangid[2], Purini Vani[3], Purnima Chandrasekar[4]
[1,2,3]*Student, Department of Electronics and Telecommunication Engineering, Thakur College Of Engineering and Technology, Mumbai, India*
[4]*Assistant Professor, Department of Electronics and Telecommunication Engineering, Thakur College Of Engineering and Technology, Mumbai, India*

Abstract*: The most effective media is music as it shows feeling of a person. Music plays a vital role in detecting the mood of a person. The kind of music they listen to will show the mood of a person. Facial expression of a human tells a lot about a person's behavior/mood. Now when a person wants to listen to music, he may want to play a track according to his current mood, but shuffling or searching through 500/600 list of songs is a tedious and upheld task and as we know that the person's mood changes frequently so to change or search according to his requirement it becomes a time consuming task and when the user is not getting the playlist according to his mood then the user might get irritate too. We know that traditional music system requires human interference. So the main objective of this work titled as "Implementation of Emotion Based Music Player from Facial Image using SVM (Support Vector Machine)" is to develop an intelligent system that can easily recognize the facial expression from an input image. Our project consists of three stages namely pre-processing stage, Feature Extraction stage, Classification stage for recognition of emotion. We will be using Facial Landmark for feature extraction. These landmark describes the position of all moving part of the detected face. We will extract the features of an image and then classify it using SVM (Support Vector Machine) classifier into different kinds of emotions like happy, sad, Angry, Neutral. Here we will be using two databases namely JAFEE database and Cohn-Kanade (CK+) database. For training the system the input image will be from both the databases. At the end we will conclude which database give the more accuracy with our implemented system. On implementing the system, we have achieved an accuracy of 71% for JAFEE database and 91% for Cohn-Kanade (CK+) database.

***Keywords*: Facial Landmark Detection, SVM classifier**

1. Introduction

Face is an organ of human body which plays a vital role in extracting the behavior and emotional state of a person. To identify emotion of a person, facial expression gives important clues about it. Music plays a vital role in relieving stress as it plays with our emotions. When we listen to music we feel relaxed and calm. It is considered to be the most effective medium as it can induce deep feelings with some kind of message in it. In short it influences our emotion which in turn affects our mood. When we listen to sad songs we feel low and there is mood decline. When we listen to happy songs, we feel excited. Manual segregation of a song based on a person's mood is time consuming and upheld task.

We know that a person is able to express emotion through different feature of face .The project describes implementation of emotion based music player which is an application of computer where users will be able to manage large playlists which will require less efforts. In general, we know that playlists of a person contain N number of songs. So to avoid manually selecting a song generally people will randomly select a song and that will not be appropriate according to the current mood of the user and it might irritate the user. So, the songs sometimes will not be matching to the current emotion of the user which he/she is going through. Therefore, the model which has been proposed by us will extract the feature and will determine the current mood of the person. Now that the emotion has been detected, song playlist will be given to the user according to the emotion identified which will enlighten the mood of a person. The model will include following moods: happy, sad, anger, disgust, joy, fear and neutral. The input will be taken from the stored database or it can be taken from the webcam.

After this we will extract the feature from the recognized face and on the basis of Facial landmark detection algorithm facial expression is detected and afterwards it is classified using Support Vector Machine (SVM). The classifier uses two methods to classify the extracted feature which are supervised and unsupervised classifying methods. Now that SVM has classified the images, a playlist consisting of different songs will be generated on the basis of the expression detected. Then the person will play a song according to his/her choice. The existing music players have features such as party shuffling, manual selection and the limitation of such systems are that the randomly played song will not be matched according to the current emotion of the user and user will have to classify the songs according to their emotion and while playing they have to manually select the song.

2. Literature survey

In [1], they proposed a system in which they used PCA

(Principal component approach) for feature extraction. To classify and recognize the expression Euclidean distance classifier was used. After the classification of the feature, the user's corresponding emotional state is recognized. For that music database was made which had a number of song from different domains pertaining to a number of emotions is collected and put up in the list. When the user's expression is recognized, songs belonging to that category are then played. They used the database with 7 expressions of 4 individual's persons that results into 112 trained images.

In [2], they proposed an android application which gives user a list of song based on user's current emotion. At the start of the application if the user has internet connection it will provide the song from online or otherwise it played song from the device memory. If user want emotion based music it will immediately start camera activity and capture image. Next it will upload the image to server and detect face using Viola Jones algorithm and the will give to FisherFace algorithm for recognition of emotion. Detected emotions are send to the user's device and are also send to the server for fetching the list of song. At last it will generate the playlist based on emotion recognized. They used 450 images to train the classifier and tested it's accuracy. The images used are from CK+ database.

In [3], image from database is passed to the facial landmark detection stage to remove noise by applying Gaussian Filter or mask. Here itself they used Viola Jones technique of Haar-like features with Adaboost learning for face detection. The feature detection stage consists of Eyebrow corners detector, Eye detector, Noise detector, Lip corner detector. After this active facial patches are extracted. The classification of features is done by SVM (Support Vector Machine). While testing it will take the hundreds of images from the database and extract the features and classifies accordingly. They used CK+ (Cohn-Kanade) dataset and JAFEE dataset for training and testing the database. The training database consist of 329 images in total.

In [4], they developed a robust system which can detect and recognize the human emotion from live feed consisting mainly of two stages- facial detection which is done by extraction of Haar Cascade features of a face using viola Jones algorithm and the second stage is emotion recognition which is done by Deep Neural Network. They used FER2013 dataset having thousands of images with all types of emotion for their system. For training their system they used the 9000 images from the FER2013 dataset.

In [5], the input images are selected from the training set. After this Landmark detection & Local representation will be done. By using LBP (Local Binary Pattern) algorithm, Local regions LBP features and Local Regions NCM (Normalized Centre Movement) features are extracted. Both the extracted features are added together and all these are passed to the SVM (Support Vector Machine) classifier. CK+ dataset consisting of 593 sequences of different emotions from 123 subjects. Only 327 out-off 593 sequences were given label for the human facial expression. They used at least two peak expression frames for anger, fear, sadness and one peak expression frame for disgust, happy, surprise. The system was trained with 6 types of facial expression.

3. Proposed work

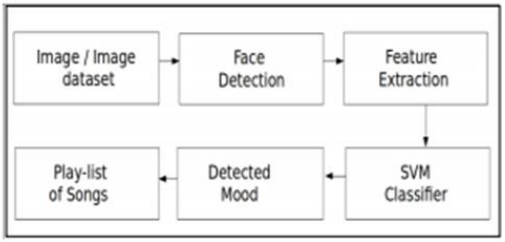

Fig. 1. Block diagram

A. *Database*

We used the JAFFE and CK database in our project. JAFFE contains 213 images of 7 facial expressions (6 different facial expression and 1 neutral) posed by 10 Japanese female models. All these images are in grayscale. CK database contains 1159 images of 7 facial expressions (6 different facial expression and 1 neutral). All these images are in grayscale. The number of images in Anger, Fear, Disgust, Happy , Neutral , Sad and Surprise are 120 , 141 , 68 , 251 , 142 , 240 and 197 respectively.

B. *Feature Extraction [6]*

Emotion detection is a very important step after face detection which can be obtained by extracting the feature of the face for example the lips are stretched then it may imply that person is happy etc. Here we will be using Facial landmark detection algorithm for feature extraction. Facial landmark is defined as localization and detection of certain points which are the keypoints on face which have an impact on particular task like face recognition, tracking of face, recognition of expression etc. Facial landmarks are eyes corners, nose tip, mouth corners, chin, nostril corners, eyebrow arcs etc. In our project we will be using Facial Landmark for feature extraction. These landmark describes the position of all moving part of the detected face. In this we will extract all the coordinates of facial landmarks. This all the coordinates are the first collection of features.

C. *Classification [7]*

A Support Vector Machine (SVM) is a classifier which is discriminative defined formally by a hyperplane separating. The objective of SVM is to create a model (in light of the preparation information) which predicts the objective estimations of the test information given just the test information properties. Proposed Procedure: In this the model is trained by taking the training dataset. Here we used the SVM classification Technique for classification. Here we used the linear kernel mode by defining the various parameter. After the model has trained we tested the model by giving the testing dataset and check the predicted output of the system. This also helps us to check the accuracy of the system for each emotion.

We also generated the confusion matrix of the testing part of the model. Also generated the classification report. Both the above data will give us the accuracy of the system.

4. Results

The implementation of the system is done on Python 3. Here, for the facial emotion recognition purpose testing of the database was carried out on static image. The input image is taken from the set of images kept aside for testing. Here we used 2 databases (JAFEE database and Conh-Kanade database) for testing our system. Cohn-Kanade database contain total of 890 images and JAFEE database contain total of 213 images out of which we separate the database in two parts: training & testing set. For training and testing purpose we organize the database in 80/20 ratio. Both sets contain 7 expression.

Cohn-Kanade Database: For the testing the model we gave one single input image to the system.

(a) (b)

Fig. 2. (a) Input Image (CK+) and (b) Detection of facial Landmarks for Feature Extraction. (CK+) database.

The system first detect the position of the landmark and then locate those landmark. As a feature it will return the position value (coordinates) of located landmarks.

For training and testing the model, we gave the database (total 890 images) to the model. After testing the confusion matrix and classification report is generated. The confusion matrix is a matrix which describe the performance of classification model or classifier. The confusion matrix is shown in Fig 3 and classification report is shown in Table 1.

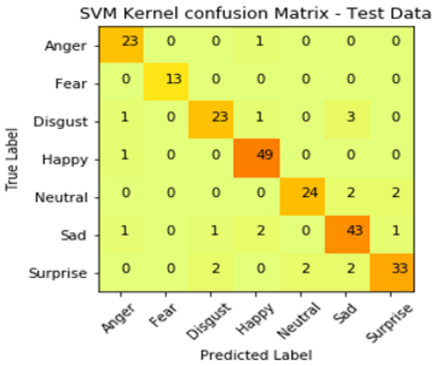

Fig. 3. Confusion Matrix of CK+ database

Table 1
Classification Report of CK+ database

Emotion	precision	recall	f1-score
Anger	0.88	0.96	0.92
Fear	1.00	1.00	1.00
Disgust	0.88	0.82	0.85
Happy	0.92	0.98	0.95
Neutral	0.92	0.86	0.89
Sad	0.86	0.90	0.88
Surprise	0.92	0.85	0.88

Table 2
Accuracy Table of CK+ database.

Emotions	Accuracy
Anger	95.83 %
Fear	100 %
Disgust	85.2 %
Happy	98 %
Neutral	85.71 %
Sad	89.58 %
Surprise	84.61 %

After the training and testing of the model the input image is recognized and appropriate music is played according to the recognized emotion.

The detected emotion is shown as follows:

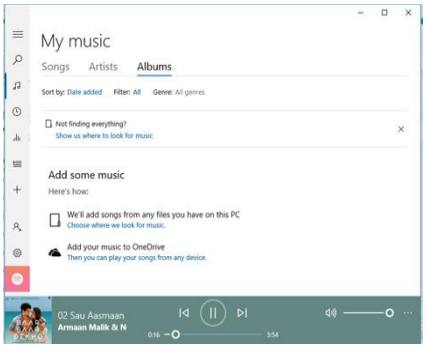

Fig. 4. Emotion is Recognized (Cohn-Kanade)

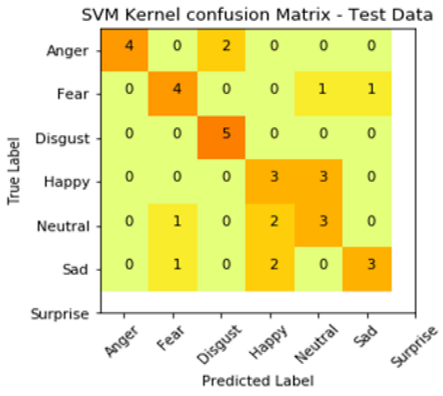

Fig. 5. For playing this window this window will open and appropriate music is played

JAFEE Database:

After the training and testing the model the confusion matrix is generated for JAFEE dataset.

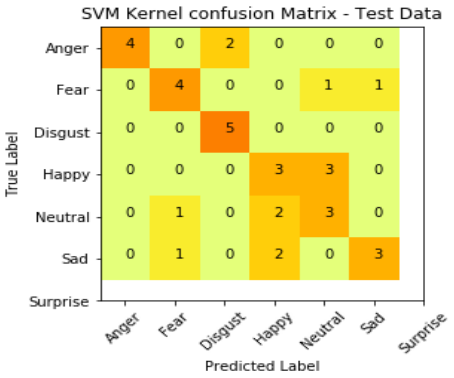

Fig. 6. Confusion Matrix of JAFEE database

Table 3
Classification Report of JAFEE database

Emotion	precision	recall	f1-score
Anger	1.00	0.67	0.80
Fear	0.67	0.67	0.67
Disgust	0.71	1.00	0.83
Happy	0.43	0.50	0.46
Sad	0.43	0.50	0.46
Surprise	0.75	0.50	0.60

Table 4
Accuracy Table of JAFEE database

Emotions	Accuracy
Anger	67 %
Fear	67 %
Disgust	100 %
Happy	50 %
Neutral	50%
Sad	50 %

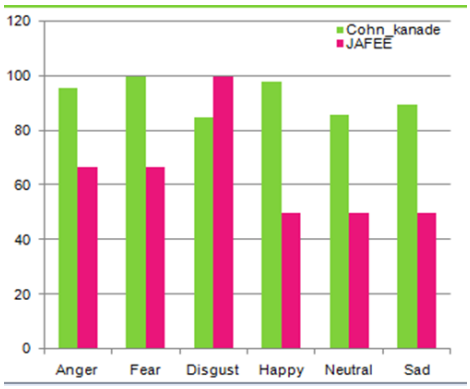

Fig. 7. Bar Graph of both the database with accuracy level of each emotion

From the above bar graph we observed the accuracy level for both the database. We observed the accuracy level of Cohn-kanade database is always higher than the JAFEE database for each emotion except the Disgust emotion.

5. Conclusion

For analysis of facial behavior, we have studied this field thoroughly. This project aimed at making Emotion based music player so the proposed work presents the facial expression recognition system and play the song according to the emotion detected. Here we used facial landmark detection as feature extraction algorithm and SVM as classifier. We used two databases i.e JAFFE and Cohn Kanade database and compared their accuracy level .The overall accuracy of the system obtained by using JAFFE database was 71% and the overall accuracy obtained by using the CK+ database is 90%. The individual accuracy of different emotions obtained by using JAFEE database is: anger (67 %), fear (67%), disgust (100%), happy (50%), neutral (50%), sad (50%). The individual accuracy of different emotions obtained by using Cohn-Kanade (CK+) database is: anger (95.83%), fear (100%), disgust (85.2%), happy (98%), neutral (85.71%), sad (89.58%), surprise (84.61%). If we test the system by using JAFFEE database, the system not give the appropriate results for the happy emotion this is because may be the JAFEE database has less number of images for training the model. On the other hand, CK+ database gives the best results for all the emotions.

6. Future Scope of the project

As of now we are taking the music from the database available online. But in future we can even take directly take songs from online where different kind of options will come like languages, newest released song, old songs from 80's and 90's etc. The future scope of this proposed method is that it can be used for music therapy session which will help music therapists to help patients suffering from depression, mental illness etc. It can also be implemented as a mobile app in future which will be convenient for the user.

References

[1] Sushmita G. Kamble and A. H. Kulkarni, "Facial Expression Based Music Player ", in 2016 Intl. Conference on Advances in Computing, Communications and Informatics (ICACCI), Jaipur, India, pp. 561-566, Sept. 21-24, 2016.

[2] Aurobind V. Iyer, and Viral Prasad," Emotion Based Mood Enhancing Music Recommendation" in 2017 2nd IEEE International Conference On Recent Trends in Electronics Information & Communication Technology (RTEICT), India, pp. 1573-1577, May 19-20, 2017.

[3] S. L. Happy and A. Routray, "Automatic facial expression recognition using features of salient facial patches," in *IEEE Transactions on Affective Computing*, vol. 6, no. 1, pp. 1-12, 1 Jan.-March 2015.

[4] J. Jayapradha Soumya Sharma and Yash Dugar," Detection and Recognition of Human Emotion using Neural Network", in International Journal of Applied Engineering Research, Vol. 13, No. 8 (2018) , pp. 6472-6477.

[5] Deepak Ghimire, Sung Wan Jeong, Joonwhoan Lee, Sang Hyun Park,"Facial Expression Recognition based on Local Region Specific Features and SVM" in Multimedia Tools and Applications, Vol. 76, Issue 6, pp. 7803–7821,March 2017.

[6] Dhananjay Rathod, Vinay A, Shylaja SS and S Natarajan," Facial Landmark Localization – A Literature Survey",in 2014 International Journal of Current Engineering and Technology E-ISSN 2277 – 4106, , P-ISSN 2347 – 5161,2014.

[7] "Machine Learning 101" https://medium.com/machine-learning-101/chapter-2-svm-support-vector-machine-theory-f0812effc72.