Computational Approaches for emotion detection in face and text

Meghashree C, Manasa M A, Madhuri Vijay, K B V M Madhusudan Reddy and Anooja Ali

School of CSE, REVA University, Bangalore, India

Abstract.

A method of playing music to a person's mood requires interaction with the user. Moving to computer vision technology will allow the mechanism of the software. To attain such, a design is handed down to categorize the human emotions and give a piece of music based on the emotion found. It eliminates the time and effort of searching for a song on the current state of a person. In this proposed system, the expressions of a user are recognized by pulling out the characteristics of the face using the PCA technique and fisher face classifier. A computer built-in camera is worned to grab a person's facial expressions, which decreases the design cost of the system. It also has a text emotion detector. Chatbot recognizes the emotion of a person through the users' text and recommends songs accordingly. Sentiment analysis methods are used to detect the emotions in text.

Keywords. Facial Feature Extraction, Expression Recognition, Fisher face classifier, PCA algorithm.

1. INTRODUCTION

In the present busy life, with evolution in the fields of technology many different music apps are designed with attributes such as rewinding, forwarding, and streaming music using multicast streams. In spite of the fact that these technologies only satisfy the needs of a person, the user must manually search for the track from a huge set of songs. The intention of this system is to design a software that can recognize the emotion and suggest or play a music track based on the emotion caught. Here we have two modes of recognizing the emotion. The first mode is using facial expressions and the second mode is using the text. Here we use the fisher face method to extract the facial features and sentiment analysis to detect the emotion using the users' text.

Emotion mode and Text mode are the two different ways where the user can switch to any mode in his/her comfort. Here we recognize the four basic emotions- Happy, Angry, Sad, and Neutral. We have found its applications in different fields such as Human-Computer Interaction (HCI), it is used in therapeutic approaches in healthcare, etc.

2. LITERATURE SURVEY

V.P Sharma et al. [1] proposed a system that uses a neural network approach for both emotion detection and song recommendation. They have used HAAR cascades for face detection, six layers of the CNN model for emotion detection, and another CNN model for music recommendation based on emotions.

Y. Yaslan and Shreya L [2][3] considered emotions using signals from wearable physiological sensors. The emotion is classified by the device that consists of galvanic skin response (GSR), photoplethysmography (PPG) sensors, and encephalography (EEG). Data Fusion techniques are used for combining data in PPG and GSR sensors. Emotion intensity will be shown on the arousal-non-arousal scale. The signal from PPG and GSR, sample, and features are extracted, target emotion is predicted, and the recommendation system song is sent to the player.

A. Alsaedi et al. [4] incorporated the Principal Component Analysis (PCA) and Viola Jonze algorithm for emotion recognition and song recommendation. The specific playlist is played based on the subject having positive or negative emotions that are being detected. At first user image is captured and using the Viola-Jones algorithm face is detected. they use PCA for emotion detection and based on the detected emotion a playlist will be provided to users where they can choose from music clips.

Joshi et al.,[5] considered Long Short-Term Memory (LSTM) and Convolution Neural Network (CNN) for detecting emotions. The system takes text or facial expressions as input from the user. The features are passed through CNN, LSTM, CNN-LSTM, and LSTM-CNN. They have used a third-party Music Service API. Shaikh et al., [6] proposed a system where emotions were identified by a chatbot. This chatbot identifies emotion by asking some general questions and recommends playlists accordingly. They used Singular Value Decomposition (SVD) and Nearest Neighborhood Model. For song recommendations they used Last.fm API and for emotion analysis they used IBM Tone Analyzer API.

S. Shivanand et al., [7] discussed a chatbot that recommends movies and music depending on the mood of the person. They have used HaarCascade Algorithm for detecting users' faces and CNN algorithm to detect emotion from features extracted and a Rule-based chatbot. The system has three modules such as Chatbot, Mood detection, and Music/Movie recommendation. Haar Cascade is an object detection algorithm to identify faces in a real-time video or an image. By using the algorithm face can be detected quickly. Convolutional Neural Network is used for object or image recognition and classification. Convolution is an operation performed on two functions that produce a third function that expresses what is the shape of one modified by the other. CNN consists of different layers that generate different functions which are passed on to the next layer. The same recommendations are considered in [8].

Krupa et al., [9] proposed a system that uses two input CNN and two approaches for detecting the emotion of the person, one way is by semantic analysis and the other way is through facial landmarks. They have used the CNN algorithm for classifying facial expressions. The chatbot takes text or voice commands for this it uses Google's text-to-speech (TTS) and speech-to-text API for more interaction. The music recommended is from predefined directories or YouTube API.

Deshmukh et al., [10] described the emotion recognition system in various phases. The emotion detection phase consists of normalization, training, capturing, and classification, and then the feature is selected. Once the feature is selected and emotion is detected, the song is played accordingly.

Sushmita et al.,[11] considered a system that uses the Euclidean Distance Classifier and Principal Component Analysis (PCA) to pull out the facial characteristics. At first, the image undergoes pre-processing technique and resizes the image, then the images are selected for the training set, thus finding the average face, and a matrix is created. Covariance of the matrix is found and then eigenvectors of the covariance matrix are calculated, Eigenfaces will be calculated, then Euclidean distances between image and eigenfaces are calculated and the minimum distance is found. The output will be an image of Minimum Euclidean distance and a music track is played.

A. Bhardwaj et al.,[12] incorporated a system where they have used EEG signals and Independent Component Analysis (ICP), Machine Learning techniques like Linear Discriminant Analysis (LDA), Support Vector Machine (SVM) where they classify EEG signals to various emotions. Their model has different phases such as Data Acquisition, Segmentation, Feature extraction (Energy, Power Spectral Density), Machine Learning techniques (SVM, LDA), and Emotion detection.

The literature survey of the existing web app can be summarized in Table 1

Name of existing web app	Problem Identified	Advantages	Ref. link
Spotify	No face and text emotion detection Manual browsing required Need internet	Easy to access Large songs collection	https://www.spotify.com/in-en/
Gaana	No face and text emotion detection Manual browsing required Need internet	Easy to access Large songs collection	https://gaana.com/
Hungama	No face and text emotion detection Manual browsing required Need internet	Easy to access Large songs collection	https://www.hungama.com/
Wynk	No face and text emotion detection Manual	Easy to access Large songs collection	https://wynk.in/music

	browsing required		

3. METHODOLOGY

In the proposed system, the user needs to select the mode initially. The available modes are face emotion and chat mode. Fig 1 represents the architecture. The flow chart shows an overview of the system and explains the functionality.

- For training the dataset in emotion mode we have used the fisher face algorithm. It is dependent on the reduction of face space by the PCA technique and later applies the LDA approach. Principle component analysis (PCA) is a dimensionality reduction method for large data sets. Linear discriminant analysis (LDA) helps to find a linear combination of attributes to separate two or more classes.
- API Stage: Once the image is captured, the system sends the image capture to SDK. Input is processed and the image feedback is sent to the system.
- Recognition Stage: The system receives the image information and recognizes the emotion based on the features extracted by the face. This emotion is again sent to the database to get the emotion playlist.
- Display Stage: Here, the songs are organized, and the user can also select a song manually from the list displayed. The user has the option to forward, remove, add, and modify the playlist.
- If the user is not interested in giving his/her face for emotion detection, he/she can switch to text mode.
- The chatbot asks some general questions to the user. For each response the person gives, the score related to the response is considered for knowing the entire emotion of the person, then the chatbot reacts based on the polarity of the sentence. Polarity in natural language processing is used in grasping the sentiment of a statement. The polarity score is a floating value with the range [-1.0, 1.0].

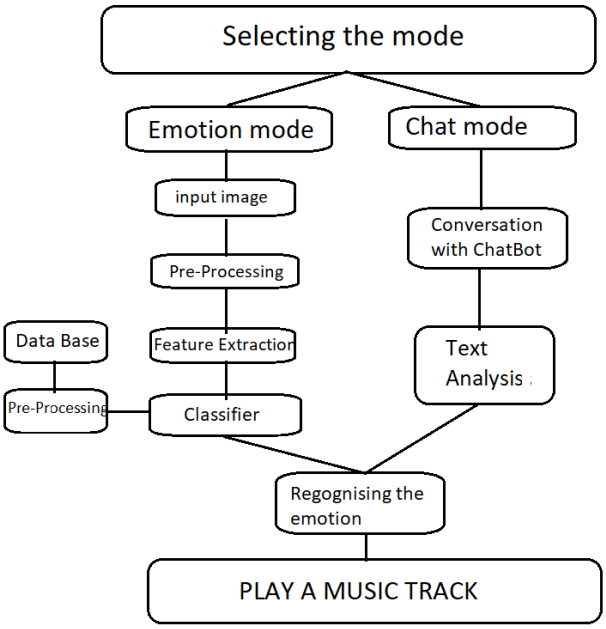

Fig. 1: System Architecture

4. DATASET

As for our study, we have used facial emotion detection in Fisher face which will work with the trained models. This will let the user choose the dataset according to their usage. If a large dataset of 26-32k is taken, then it gives good accuracy and, in this case, the users are few people. So, in these conditions, if we use a dataset of 410-460 images as input according to the user then even, in this case, we get a very good accuracy for a small quantity of dataset and occupies low memory to work. In such conditions, if we consider some dataset with around 150-200 samples as input according to the user then it also gives us a better accuracy with the advantage of low quantity dataset and low storage on memory for operation. Here we used the Extended Cohn-Kanade dataset as we need to make it to train our model.

To train the model, we used a code in Python that collects all the images classified from the folders and matches them with the emotion. This data is stored in a dictionary. We use the .train method for training the model.

The dataset we have used consists of 5,876 labeled images of 123 individuals. It has Neutral to peak expressions with similar backgrounds, in grayscale, and with 640x490 pixels. It consists of seven expressions such as anger, happiness, Neutral, sadness, disgust, surprise, and fear. Using this dataset, we observed four emotions in our study: anger, happiness, sadness, and neutral.

5. EXPERIMENTAL RESULTS

The software requirement includes an Operating system with Windows 7 or above. The front end is designed with Front End: HTML, CSS, JS. The proposed system is implemented in python with PyCharm Community Edition 2021.3.2 as the iDE. It requires a hard disk of 100 GB and RAM of 4 GB.

The following, Fig 2 shows the result screenshots of emotion mode. It shows the few song sequences selected based on facial emotion. The following shows the chat with the user in fig.3.

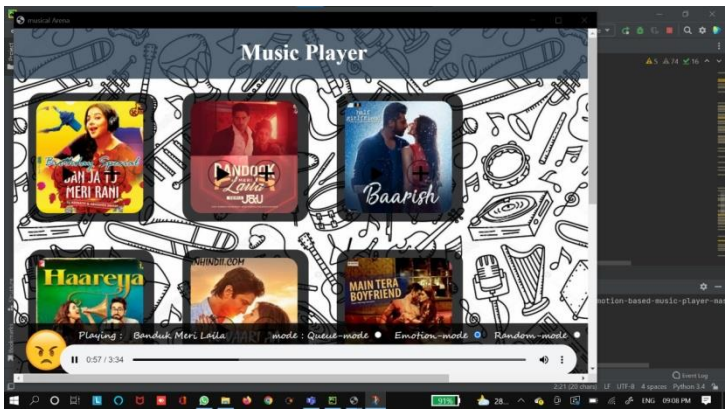

Fig 2: Sample emotion detected, and the song selected

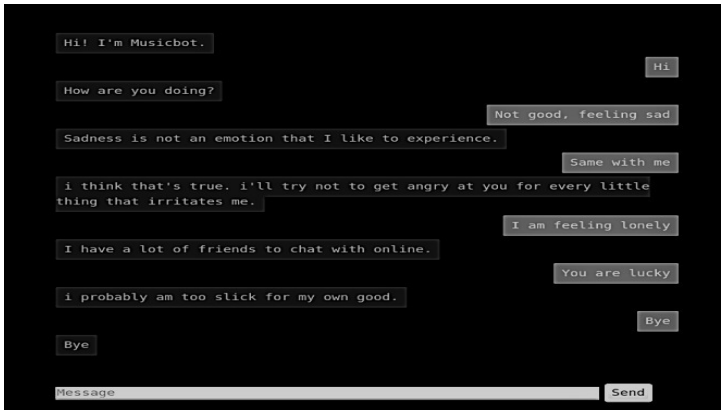

Fig 3: Chat session performed with the user

6. OBJECTIVES AND APPLICATIONS

The main aim of developing this system is to build Human-Computer Interaction. It also provides a proper and accurate Emotion Detection. Playing music is associated with improving the mood of the user. Building a website with a good user interface and functionality.

The applications include Human-Computer Interactions. It can be used by Medical – Psychiatrists, Autism Spectrum Disorder. music is the best treatment for any mental disorder. Music is the best treatment for any mental disorder. It takes advantage of music's inherent mood-lifting properties to aid individuals in improving their mental health and overall well-being. It can also be used in Autonomous Vehicles to improve the mood of the passenger

7. FUTURE SCOPE

Emotion recognition has high scope in most areas such as human-computer interaction, biometric security, etc. So, it gives awareness to artificial intelligence or machine intelligence that utilizes variously supervised and unsupervised machine-learning techniques to imitate the human brain.

The scope for the developing project would be to design it for handsets. To implement a system that helps in the music treatment to give treatment to the sufferers. who are facing mental illness, trauma, and acute depression. It can also be used in detecting the current mood of a person who is physically challenged.

8. CONCLUSION

With the growing combination of computer interfaces and computers in our daily lives, the rise in the necessity for computers to recognize and make a response to human interaction and behavioral signals of mental states and emotions. The expressions in the images and texts are significant. In certain, the features that match the cognitive and affective states of the mind which are not a slice of the emotions set. This is demanding because of behavioral cues' inherent inference of hidden mental states. In this research, facial expressions and text expressions are considered for the detection of emotions. The system is not dependent on elements like age, ethnic group, gender, background, and also birthmarks. The system hopes to be very much promising and designed for users going through mental illness during their working hours or anything by providing them music treatment.

9. REFERENCES

[1] V. P. Sharma, A.S.Gaded, D.Chaudhary, S.kumar, S.Sharma, "Emotion-Based Music Recommendation System".In the 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO) Amity University, Noida, India. Sep 3-4, 2021

[2] D. Ayata, Y. Yaslan, M.E.Kamasak, "Emotion Based Music Recommendation System Using Wearable Physiological Sensors".In IEEE TRANSACTIONS ON CONSUMER ELECTRONICS, VOL. 14, NO. 8, MAY 2018

[3] Shreya L, Nagarathna N, "Emotion Based Music Recommendation System for Specially-Abled".In 6th International Conference on Recent Trends on Electronics,Information,Communication & Technology(RTEICT), 2021

[4] A.Alrihaili, A.Alsaedi, K.Albalawi, L.Syed, "Music Recommender System for users based on Emotion Detection through Facial Features".In Developments in eSystems Engineering(DeSE) 2019

[5] S.Joshi, T.Jain, N.Nair, "Emotion Based Music Recommendation System Using LSTM - CNN Architecture".In 12th International Conference on Computing and Networking Technologies(ICCCNT), July 6-8, 2021.

[6] A.S.Shaikh, B.V.Patil, T.P.Sonawane, "Chatbot Song Recommender System", International Journal of Emerging Technologies and Innovative Research (www.jetir.org | UGC and ISSN Approved), ISSN:2349-5162, Vol.8, Issue 12, page no. ppd222-d226, December-2021

[7] Shivani Shivanand,K.S Pavan Kamini,Monika Bai.M.N,R.Ramesh, Sumathi H.R, "Chatbot with Music and Movie Recommendation based on Mood", International Journal of Engineering Research & Technology (IJERT)ISSN: 2278-0181 NCAIT - 2020 Conference Proceedings.

[8] A.Nair, S.Pillai, G.S.Nair, Anjali T, "Emotion Based Music Playlist Recommendation System using Interactive Chatbot".In 6th International Conference on Communication and Electronics Systems (ICCES-2021)IEEE Xplore Part Number: CFP21AWO-ART; ISBN: 978-0-7381-1405-7

[9] Krupa K.S, K.Rai, Ambara G, S.Choudhury, "Emotion aware Smart Music Recommender System using Two Level CNN".In 3rd International Conference on Smart Systems and Inventive Technology (ICSSIT 2020)IEEE Xplore Part Number: CFP20P17-ART; ISBN: 978-1-7281-5821-1

[10] R.S.Deshmukh, V.Jagatap, S. Paygude, "Facial Emotion Recognition System through Machine Learning approach".In International Conference on Intelligent Computing and Control Systems ICICCS 2017.

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

[12] A.Bhardwaj, A.Gupta, P.Jain, A.Rani, and J.Yadav, "Classification of Human Emotions from EEG signals using SVM and LDA Classifiers".In 2nd International conference Signal Process and Integrated Networks, 2015, pp. 180-185.

[13] Ahmed, S. T., Sreedhar Kumar, S., Anusha, B., Bhumika, P., Gunashree, M., & Ishwarya, B. (2018, November). A Generalized Study on Data Mining and Clustering Algorithms. In *International Conference On Computational Vision and Bio Inspired Computing* (pp. 1121-1129). Springer, Cham.

[14] Shalini L, Manvi SS, Gowda NC, Manasa KN, "Detection of Phishing Emails using Machine Learning and Deep Learning", 7th International Conference on Communication and Electronics Systems (ICCES), pp. 1237-1243, Jun 2022.