Automatic Music Player based on the Facial Expression

[1]Macherla Harshini

Department of InformationTechnology
MLR Institute of Technology
Hyderabad, India
harshinimacherla90@gmail.com

[2]Linga Varshith
Department of InformationTechnology
MLR Institute of Technology
Hyderabad, India
bunnyvarshith147@gmail.com

[3]Komma Rahul
Department of InformationTechnology
MLR Institute of Technology
Hyderabad, India
kommarahul7377@gmail.com

[4]Venkata Vara Prasad Kothamasu
Department of InformationTechnology
MLR Institute of Technology
Hyderabad, India
imvara7@gmail.com

[5]Yamala Gnanadeep
Department of InformationTechnology
MLR Institute of Technology
Hyderabad, India
gnanadeep.y.1@gmail.com

Abstract— **A user's emotion or mood can be detected by their facial expression. These expressions can be exported from a live stream through the system camera. Many explorations are underway in the fields of computer vision and machine learning (ML), where machines are trained to recognize the rich variety of feelings or deadly emotions. Knowledge of machines offers colorful ways to detect deadly feelings. A similar approach is to use CNN models with Keras. Mortal feelings are closely related to music. In our proposed system, an emotion-based music player is created that performs real- time emotion detection and plays songs from YouTube web based on the detected emotion. One of the main benefits of integrating emotion detection is customer satisfaction.**

Keywords-Keras, Open Computer Vision, Convolutional Neural Network.

I. INTRODUCTION

Mortal feelings can be astronomically divided into fear, nausea, anger, surprise, sadness, happiness, and neutrality. Many other feelings like joy (a variant of pleasure) and contempt (a variant of disgust) can be spread under the big tent of this feeling. These feelings are subtle. Facial muscle distortions are minimal, and detecting these differences is laborious, as small differences can indeed lead to different expressions.

Moreover, the same emotion can be expressed differently by different people or even by the same person, since feelings are very context dependent. If the focus can only be placed on the areas of the face where the sensation occurs, such as the mouth and around the eyes, how these gestures are uprooted and distributed remains an important question. Neural networks and machine literacy have been used for these tasks with good results. People often use music to regulate their mood, especially to change a bad mood, increase energy positions, or reduce stress. Additionally, listening to the right music at the right time can improve internal health. Therefore, human feelings have a strong relationship with music. In order to descry the mood of the stoner image is captured also we descry face from the image.

It is given to deep literacy model which recognizes the mood and according to the mood, song will be played.

In the existing system only, the fixed songs will be played which are present in the dataset, that are not much helpful for the user in the upcoming days. If we need to add any song into the dataset, we need to update it manually. So, still it is a manual project.

So, we came with one proposed model which are helpful for the upcoming days. Instead of using the dataset we are going to use YouTube web links, so we can get the songs based on the corresponding expression the songs may include like Latest songs, Trending songs, most viewed songs, most liked songs etc. So, it can be helpful for the user emotion song but returning the good song. For detecting the emotion, we are using Convolutional Neural Network and Keras.

A convolutional neural network is a deep learning architecture. It consists of several layers including convolutional layers, grouping layers, and fully connected layers. Convolutional layers apply filters to the input image to extract features, clustering layers down sample the image to reduce computation, and fully connected layers make the final predictions.

II. LITERATURE SURVEY

The article by Preema J.S mentioned that the face expression detected from camera by using Viola-Jones Algorithm, Support Vector Machine, Facial Expression Recognition explained the importance of emotions for human beings and it also explains about the relation between emotions and songs. That captured passed on that different stages to detect the user's emotion and it will generate the playlist according to the mood. But it has some limitations, that users need to select the songs manually and after the playlist is generated those songs may not match exact moods, it is less accurate model and time consuming [6].

The article by T V Rashma addressed the problem of existing models and this model solves some limitations of that models. This paper also explained about the importance of music for human beings in today's world. The author focused on developing

an emotion-based music system, in this face is captured by the camera and one of the following emotions are classified: sad, anger, happy, surprise, disgust, neutral and fear. This classification is done by Convolutional Neural Network [16]. Playlists are formed in a music player implemented using PyQt5 depending on the mode [7].

According to Ashish Adholiya's article, for non-verbal communication, facial expressions are very important. This research article extends convolutional neural network model development using keras to create and train deep learning models. The main focus is on detection of facial expression of the uploaded image using the classifiers, hence drawing outlines for faces. To target the emotions, used, the facial expression recognition dataset, which was split into two folders, called test and train, and then into separate folders, each containing one of seven types of sets of Face Emotion Recognition (FER) data. The image size of FER dataset is 48x48 and its size of the batch is 64, which speeds up the training process that depends on CPU memory size. The CNN model is said to be trained up to 15 epochs with an accuracy rate of 66.7% [4].

An article by Alireza Jolfaei mentions that there is a need to avoid the complicated process of explicit future extraction in traditional facial expression recognition. A facial expression recognition method using CNN [14] and image edge detection is proposed. During convolution, the edges of each layer of the facial expression image are extracted by adding filters to the image. The author used facial expression dataset and Internal Feeding Worm (IFW) dataset. In this model, their average recognition rate is 75.56%, the training speed is faster than the comparison algorithm [2].

A study by Dhruvi D. Gosai is to detect the emotion and recognition from text not from facial expression. So, this is one of the experimental models. Sentimental detection from textual sources can be performed using concepts from natural language processing. Keyword features of this model are natural language processing, keyword face detection, word processing, sentiment measurement. In sentiment analysis, the emotion like detect positive, negative or neutral sentiments from a given user input text. Then sentiment analysis can detect types of feelings like happiness, sadness, surprise, disgust, neutrality, anger, and fear from given user input [1].

III. EXISTING SYSTEM

Existing systems include various methods of emotion recognition. After detection of the emotion, the song will play from the recommended playlist from their dataset. The dataset need to updated to add the latest songs which are helpful to the user.

DISADVANTAGES OF EXISTING SYSTEM

- In this there is no overcome from the emotion like sad.
- The calculation of the results takes a long time.
- In this system it will only play the song which are present in the datasets.

IV. OBJECTIVE

The main objective of this study is to detect users' emotion based on facial expressions using Conventional Neural Network [14] and automatically plays songs from YouTube based on the corresponding emotions.

V. PROPOSED SYSTEM

The proposed system facilitates the representation of the interaction between the user and the music. The purpose of the system is to correctly capture the face using the camera. A convolutional neural network predicts the emotions from captured images or inputs. After detecting the emotion, the song will play from the YouTube web, identifying the emotion. If the same emotion repeats continuously the next song will play depending on the type of emotion. In this no datasets are used for music.
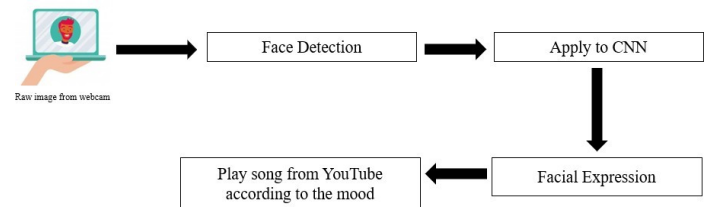

Fig 1: Architecture of proposed system

From fig 1, a webcam is used to capture live video of user's emotion.

A Conventional Neural Network is used for detection of the objects from the images extracted from the Open CV.

Once the objects are captured from images, the detected emotion can be mapped with one of the following emotions – happy, sad, surprise, neutral, fear, disgust, angry.

After mapping the emotion, it plays the song from the YouTube web.

ADVANTAGES OF PROPOSED SYSTEM

- For sad it will generate a song which overcome the emotion.
- It plays song from the YouTube web, not from the dataset.
- In this system, the user's emotion is recognized by combining Keras' convolutional neural network to recognize the user's emotion faster.

VI. RESULTS

CASE-A: A happy face emotion is detected and a song will play from YouTube web.
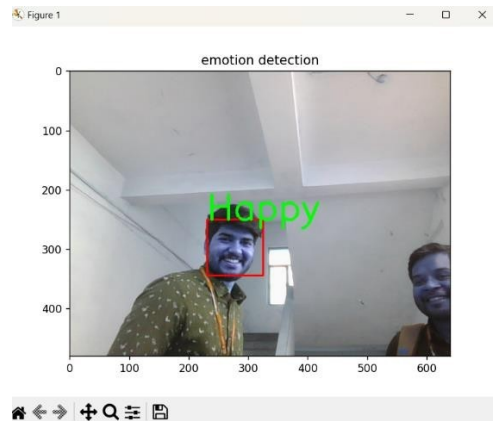
A.1: Emotion Detection

Fig 2: Detected Happy Emotion

A. 2: Played song from the YouTube corresponding to the mood.
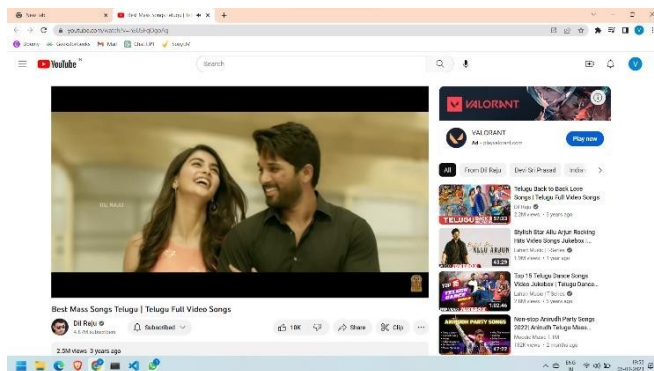

Fig 3: Happy Song

CASE-B: Detecting a Surprise Emotion and playing corresponding song from YouTube web.

B. 1: Emotion Detection

Fig 4: Detected Surprise Emotion

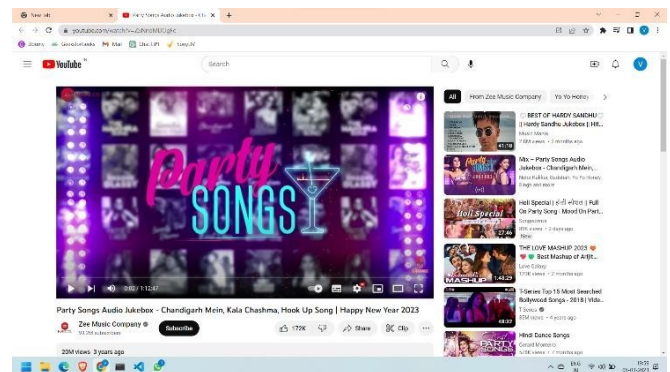**B.2:** Played song from the YouTube corresponding to the mood.

Fig. 5: Surprise Song

VII. CONCLUSION

In this research, to classify human facial expressions like happy, sad, surprised, angry, neutral, we used convolutional neural networks. Real-time expression recognition is also performed by capturing images using webcam and processing through Haar Cascade Classifier using OpenCV. Finally, a good recognition rate with an accuracy of 80%. Used web YouTube web to fetch play music from YouTube.

VIII. FUTURE EXTENSION

The everlasting target in machine learning or deep learning algorithms is to improve the accuracy as much as possible. In this project, one of the biggest challenges from the start is choosing the best CNN architecture which is a deciding factor affecting the performance of the model. Finding a suitable dataset is a major factor in predicting the correct expressions. Hence improving accuracy. Datasets can be balanced by increasing the number of minority classes using oversampling.

IX. REFERENCES

[1] Dhruvi D. Gosai, Himangini J. Gohil and Prof. Hardik S. Jayswal, "A Review on a Emotion Detection and Recognization from Text Using Natural Language Processing", IJAER, ISSN 0973-4562 Volume 13, Number 9 (2018).

[2] Alireza Jolfaei and Mamoun Alazab, "A Face Emotion Recognition Method using Convolutional Neural Network and Image Edge Computing" in IEEE Access, vol. 7, pp. 159081-159089, 2019, doi: 10.1109/ACCESS.2019.2949741. (2019).

[3] A. Jaiswal, A. Krishnama Raju and S. Deb, "Facial Emotion Detection Using Deep Learning," International Conference for Emerging Technology (INCET), Belgaum, India, 2020, pp. 1-5, doi:10.1109/INCET49848.2020.9154121.

[4] Ashish Adholiya, "Facial Expression Recognition Using CNN and Keras", Biosc.Biotech.Res.Comm. Special Issue Vol 14 No 05 (2021) Pp-47-50, DOI: http://dx.doi.org/10.21786/bbrc/14.5/10.

[5] Radhey Shyam, "Facial Emotion Detection and Recognition", IJEAST, Vol. 7, Issue 1, ISSN No. 2455-2143, Pages 176-179 (2022).

[6] Preema J. S, Rajashree, Sahana M, Savitri H, Shruthi S. J, 0, Review on Facial Expression Based Music Player, International Journal of Engineering Research & Technology (IJERT) ICRTT – 2018 (Volume 06 – Issue 15).

[7] T V Rashma and Sheena. K, "Smart Music Player Basedon Emotion Recognition from Facial Expression", IJISRT, Volume 6, Issue 3, March – 2021.

[8] Haarcascade_frontalface_default.xml file downloaded from the following link: https://github.com/opencv/opencv/tree/master/data/haarcascades.

[9] O. Lartillot, and P. Toiviainen, "A Matlab Toolbox for Musical Feature Extraction from Audio", ICDAE (DAFx-07), Bordeaux, France, September 10-15, 2007.

[10] Anukriti Dureha, "An Accurate Algorithm for Generating a Music Playlist Based on Facial Expression" 2014 by IJCA Journal, Volume 100-Number , doi: 10.5120/17557-8163.

[11] K. F. Azizan Illiana, "Facial Emotion Recognition: ABrief Review" (2018).

[12] G. Paul O. Arriaga, "Real-time Convolutional Neu-ralNetworks for Emotion and Gender Classification".

[13] M. Harshini, "Sewage Pipeline Fault Detection using Image Processing", ICECA, Dec-2022.

[14] M. Harshini, "Artificial Intelligence Based Scarecrow Prevention from Crop Destruction", ICSCSS 2023, ISBN: 979-8-3503-3359-6.

[15] A. Maurya and V. Sharma, "Facial Emotion Recognition Using Keras and CNN," 2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), Greater Noida, India, 2022, doi: 10.1109/ICACITE53722.2022.9823480.

[16] R. Jaiswal, "Facial Expression Classification Using Convolutional Neural Networking and Its Applications," 2020 IEEE 15th International Conference on Industrial and Information Systems (ICIIS), RUPNAGAR, India, 2020, pp. 437-442, doi: 10.1109/ICIIS51140.2020.9342664.