The Feeling based Music Recommendation System

[1]Vegi Srinivas Naidu, [2]P.Rama Krishna
[1,2]Dept. of CSE, KIET, Kakinada, AP, India

Abstract
The internet and mobile technology have developed quickly and made it possible for us to freely access various music resources. While the music industry might lean more toward certain genres of music. But there is no particular way by which we can understand what exactly user wants to listen based on current mood or emotion. Music is a great way to express emotions and moods. For example, people like to listen to happy songs when they are feeling good, a soothing song can help us to relax when we're feeling stressed or exhausted and people tend to listen some sort of sad songs when they are feeling down. So in this project, we are going to develop a system which will capture the real time emotion of user by conversating with user or by other means and based on that emotion related songs will be recommended. We are going to categorize songs into the groups based on the categories like Happy, Sad, Neutral etc. Then according to the captured emotion from the user, the songs related to that emotion will be recommended. In this way, user can listen the songs according to the mood.

Keywords
Recommendation System, Emotion Recognition, Interactive UI, Mood based classifier.

I. Introduction

People frequently use their facial expressions to convey their feelings. It has long been recognized that music may change a person's disposition. A user's mind can be gradually calmed down and an overall nice effect can be produced by capturing and recognizing the emotion being uttered by the person and playing appropriate tunes matching the one's mood.[7] The goal of the project is to record a person's facial expressions as they exhibit emotion. The web camera interface for computer systems is used by a music player to record human emotion. The software takes a user's image and uses image segmentation and image processing techniques to extract information from a target person's face in an effort to determine what emotion they are attempting to convey. The idea attempts to uplift the user's mood by playing music that fits their needs while simultaneously taking their photograph. Facial expression recognition has been the most effective method of expression analysis known to humanity since ancient times. Facial expressions are the best way that people may deduce or evaluate the emotion, sentiment, or thoughts that another person is attempting to express. Mood modification may occasionally be useful in overcoming depressive and sad circumstances. Many health dangers can be prevented with the use of expression analysis, and actions can be done to improve a user's mood as well.

Our project focuses classification of songs based on acoustic features. Then system determines the mood of the user through its facial expressions and eventually system recommends songs to user according to mood of the user which was earlier classified into different emotions. In this way user can get recommendation of songs based on his current mood and will change dynamically based on current mood.

II. Related Work

Emotions are a basic part of human nature. Human's emotional states and present moods can be quickly seen through their facial expressions. Basic emotions (happy, sad, angry, excited, surprised, disgusted, fear, and neutral) were taken into consideration when developing this system. In this research, face identification was accomplished using a convolutional neural network.

Recommended manual playlist segregation and song annotation based on the user's emotional state as a timeconsuming and labor-intensive task. There have been many algorithms suggested to automate this process.

However, the currently used algorithms are slow, use extra hardware, raise the system's overall cost, and have substantially the overall cost of the device and have much poorer precision.

It aims to increase the accuracy of the system design. contrasting a face expression recognition module of the system with a dataset that is both user-dependent and userimpartial serves to validate it.

According to AyushGuidel [1] et al, one may quickly determine a person's emotional condition by observing their facial expressions. Basic emotions (happy, sad, angry, excited, surprised, disgusted, fear, and neutral) were taken into account when developing this system. In this research, face detection was implemented using a convolutional neural network. On the whole, people refer to music as a "language of feelings."

Sadhvika [4] et al. recommended manual playlist segregation and song annotation based on the user's current emotional state as a time-consuming and labor-intensive task. There have been many algorithms suggested to automate this process. However, the currently used algorithms are slow, use extra hardware (such EEG structures and sensors), raise the system's overall cost, and have substantially lower accuracy.

The report suggested by RamyaRamanathan [3] described an intelligent music player that used emotion recognition. A fundamental aspect of human nature is emotion. They have the most significant role in life. Human emotions are intended to be shared and understood by others. The local music library of the user is initially sorted in accordance with the feelings the album arouses. The lyrics of the music are frequently taken into consideration to ascertain this. This paper discusses in detail the methods for human emotion detection that are available for use in developing emotion-based music players, the method a music player uses to detect human emotions, and the best way to use the proposed system for emotion detection. Additionally, it gives a brief explanation of how our systems work, how to create playlists, and how to classify emotions.

Preema [2] et al. claim that creating and maintaining a large playlist requires a lot of time and work. The report claims that the "music player itself selects a song based on the user's current mood, and

the application analyses and categorizes audio files according to audio properties to construct playlists depending on moods.

III. System Overview

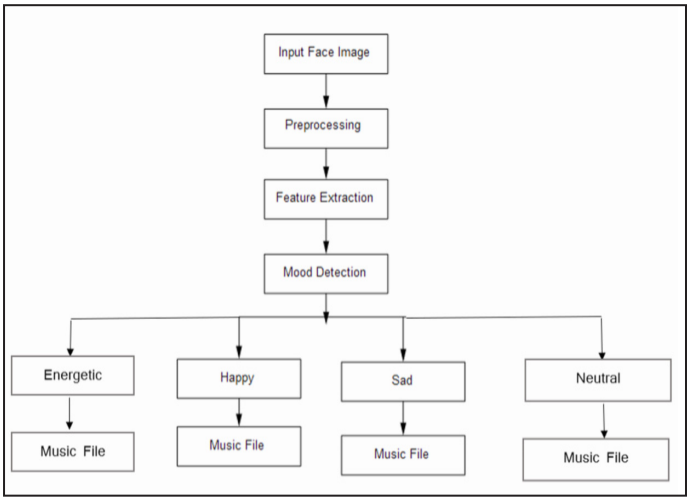

Fig. 1: System Overview

It is advantageous for us to display user and music player interaction in the proposed system. The system's main function is to ensure that the camera adequately captures the face. Convolutional Neural Network, which forecasts emotion, is given images that have been captured. After that, a playlist of songs is generated using the emotion that was recorded in the image. The major goal of our suggested method is to automatically generate a music playlist to alter the user's mood, which can be joyful, sad, natural, or astonished. When a topic has a negative feeling, the suggested system recognizes those emotions and displays a playlist of the best music to lift the person's spirits.

[5] Given diagram gives the clear understanding of how the system is going to work. This includes sequential steps in which every step has its own significance. First system classifies the songs on its acoustic feature like loudness, tempo, energy etc. and the mood of the music is determined. In this way all the songs in the dataset are classified into different moods and based on the input received from the user system recommends the songs to the user.

System mainly divided into 4 components. It includes Real Time Capture, Face Recognition, Emotion Detection, Music Recommendation. While recommending a song system goes throughout these steps. In real time capture system captures the face of the user and in the subsequent step CNN algorithm evaluates the features of that image. After detecting the mood or emotion of the image system recommends the appropriate song to the user.

IV. Methodology

A. Dataset Description:

1. Image Dataset:

The FER2013 dataset was utilized to develop the model. 48x48 pixel grayscale portraits of faces make up the data. Anger, Disgust, Fear, Happy, Sad, Surprise, and Neutral are the seven emotions that are assigned to each image. 28,709 samples make up the training set, whereas 3,589 examples make up the public test set.

2. Music Mood Dataset

A labeled dataset with the size of 686 is used.

It has 19 columns including name, album, artist, id, release_ date, popularity, length, danceability, acousticness, energy, instrumentalness, liveness, valence, loudness, speechness, tempo, key, time signature, and mood.

The moods are in four categories: Calm, Happy, Energetic, Sad. The features of Length, Danceability, Acousticness, Energy, Instrumentalness, Liveness, Valence, Loudness, Speechiness and Tempo are used for training the model because they have more influence to classify the tracks.

Fig. 2: Music Mood Dataset

B. Emotion Detection:

One application that falls under computer vision technology is face detection. Algorithms are created and trained in this procedure to correctly locate faces or objects in object detection or related systems in photos. Real-time detection from a video frame or image is possible.

By minimizing outside noises and other elements, face detection primarily aims to identify the face within the frame. This method is based on machine learning, and a collection of input files is used to train the cascade function.This makes use of machine learning techniques to extract training data with a high degree of accuracy.

We use the pre-trained network, which is a sequential model, as an arbitrary feature extractor while performing feature extraction. allowing the input image to progress to the next layer, stopping there, and using that layer's outputs as our features. Use only a few filters because the initial convolutional network layers retrieve the highest-level characteristics from the captured image. As we add deeper levels, we multiply the number of filters by two or three, depending on how big the filter was in the preceding layer.
The deeper layer filters have more features, but they require a lot.

For this, we made use of the Convolution neural network's robust, discriminative features. Feature maps, which are an intermediate representation for all layers following the first, will be the model's outputs. Load the input image that you wish to see the feature map for so you can see what features were most important for classifying the image.

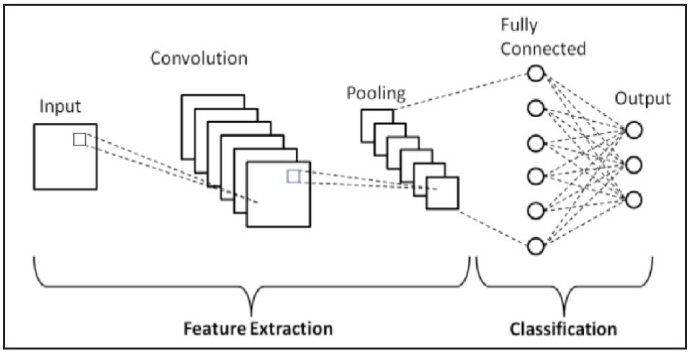

Fig. 3: Convolutional Neural Network Performance.

Using the Relu activation function in the convolution neural network architecture, filters or feature detectors are applied to the input picture to generate feature maps or activation maps. Feature detectors or filters can be used to identify edges, vertical and horizontal lines, bends, and other characteristics that are already present in the picture. To achieve translation invariance, the feature maps are then pooled. The principle behind pooling is that when we slightly alter the input, the pooled outputs remain unchanged. It is possible to utilise any pooling from the minimum, maximum, or average. However, as compared to min or average pooling, max-pooling provides better

The image's class will either be binary or multi-class to distinguish between different types of clothing or to identify digits. The learned properties in a neural network cannot be understood because neural networks are like a "black box." Consequently, the CNN model basically returns the findings after receiving an input image. The model that was learned by weights using CNN is loaded to detect emotions. When a user takes a real-time image, the image is submitted to a CNN model that has already been trained. The CNN model then predicts the emotion and adds a label to the image.

C. Music Recommendation:

Our dataset contains various features that too are also considered as acoustic features of that particular song, every feature has its magnitude which defines the intensity of the that feature. Magnitude may range in different scales. In general, less Magnitude Or negative value of the Magnitude represent that least contribution of that feature to overall mood of the song, and obviously if Magnitude has larger value then that feature dominantly contribute to the overall mood of that song.In this way, there are 4-5 main features amongst 18-19 that contribute more.[8]We get desired mood for the song after training the model based on that dataset. In this way, songs are classified into different categories like happy, sad, neutral and energetic.

When emotion is detected through facial expressions, we assign the appropriate category to that mood, for example if detected mood is happy then system will recommend songs from happy category those were already categorized. In this way, song is recommended to the user.

V. Future Scope

This system is completely operational, however there is still opportunity for advancement. A lot of changes may be made to the programme to enhance user experience generally and provide better outcomes.

Currently existing systems face difficulties while giving manual inputs but our system performs well when there is difficulty in giving manual inputs as we are dynamically determining the mood of a person from its expressions. In addition to this, this system not only used as a feature but also as a independent system. If further technical advancements done and if it improved the efficiency of the algorithm used then we can also recommend the songs on crucial and trivial expressions. [6]

Current system mainly focuses on doing the recommendation on content based features that is based on its acoustic parameters. But in future hybrid model can be created that uses both content based and collaborative(user history, feedback, liking etc.) for recommendations.

Future objectives for the system include creating a device that might support the use of music therapy in the treatment of patients who are suffering from mental stress, anxiety, acute depression, and trauma. There is a chance to add some functionality as a potential repair in the future due to the current system's poor camera resolution and performance in extremely low light conditions. If camera faces bad light issues then either based on his approximate expressions, music can be recommended or as we mentioned earlier we can make use of collaborative features to recommend appropriate song to the user. As this problem is more inclined towards hardware issue we can overcome it by using high quality camera devices and sensors having high end technical specifications.

VI. Conclusion

There are several methods to use the Music Recommender System, according to a thorough examination of the literature. The methods proposed by preceding researchers and developers were examined. So when we started studying we mainly found 2 approaches and that too independent. The first approach was like just determining the accurate emotion from the facial expression and second one was classifying the songs into the front emotions based on their acoustic features. The goals of our system were fixed based on the results. So we decided to merge these two approaches and provide complete solution for existing problem. The available technology can determine a user's emotions. The system was able to identify happy, sad, angry, neutral, or shocked emotions. The suggested approach presented the user with a playlist of music matches that corresponded to the user's emotion after identifying it. Memory and CPU usage increase as a result of processing a large dataset. Development will become more difficult and appealing as a result. The goal is to develop this application as affordably as feasible and on a common platform. Our facial emotion-based music recommendation system will make it easier for users to make and manage playlists.

References

[1] AYUSH Guidel, BiratSapkota, Krishna Sapkota, Music recommendation by facial analysis, February 17, 2020.

[2] Preema J.S, Rajashree, Sahana M, Savitri H, Review on facial expression-based music player, International Journal ofEngineering Re-search & Technology (IJERT), ISSN-2278-0181, Volume 6, Issue 15, 2018.

[3] RamyaRamanathan, RadhaKumaran, Ram Rohan R, Rajat Gupta, and VishalakshiPrabhu, an intelligent music playerbased on emo-tion recognition, 2nd IEEE International Conference on Computational Systems and Information Technology for Sustainable Solutions 2017. https://doi.

org/10.1109/CSITSS.2017.8447743
[4] CH. sadhvika, Gutta.Abigna, P. Srinivasreddy, Emotion-based music recommendation system, Sreenidhi Institute of Science and Technology, Yamnampet, Hyderabad; International Journal of Emerging Technologies and Innovative Research (JETIR) Volume 7, Is-sue 4, April 2020.
[5] MadhuriAthvale, Deepali Mudale, UpasanaShrivatsav, Megha Gupta, Music Recommendation based on Face Emotion Recognition, Department of Computer Engineering, NHITM, Thane, India.
[6] SheelaKathavate, Music Recommendation System using Content and Collaborative Filtering Methods,Department of Information Science and EngineeringBMS Institute of Technology and ManagementBangalore, India.
[7] AmeyUjlambkar, OmkarUpadhye, Akshay Deshpande, GunwantSuryawanshi, Mood Based Music Categorization System for Bollywood Music.
[8] Jose Padial. Ashish Goel, Music Mood Classification.