International Journal of Scientific Research in Science, Engineering and Technology
Print ISSN: 2395-1990 | Online ISSN : 2394-4099 (www.ijsrset.com)
doi : https://doi.org/10.32628/IJSRSET

# Enhanced Music Recommendation System

**Manyu Manchala[1], Abhinaya Vankadari[1], Yenugu Keerthana[1], Dr. Sunil Bhutada[2]**
[1]Student, [2]Professor
Department of Information Technology, Sreenidhi Institute of Science and Technology, Hyderabad, Telangana, India

**Article Info**
Volume 9, Issue 3

Page Number : 371-376

**Publication Issue :**
May-June-2022

**Article History**
Accepted : 01 June 2022
Published: 10 June 2022

**ABSTRACT**

Music has a great connection with a person's emotions. It is very baffling for an individual to choose which music to listen to from a huge collection which already exists. As many of us know, humans show their feelings by expressions on their faces. The purpose of this project is to suggest a playlist by capturing the facial expression from the human face and helps in analysing the emotion.

Detecting an emotion is one of the trending research topics. It culminates in understanding the current emotional state of the user. Here we use the K-means clustering algorithm for classifying songs and Haar cascade algorithm to recognize emotion. The image which is captured through video and the emotion detected by algorithms helps to display a list of songs. Based on the mood and emotion It will suggest songs to the user by matching their emotions to the song mood type so that It will save user searching time.

## I. INTRODUCTION

Emotions of a person can be identified through different means such as facial expressions, voice and tone of the person. Facial expressions play an important role in identifying the emotion of the person [9]. Music, on the other hand, could have a powerful impact on one's life and may be used as a tool to change one's mood. For example, a person who is in a sad mood can use music to uplift his mood. Music is recommended to the user in this notion by real-time recording of the user's emotions. This also recommends a list of songs to the user and also an option to change the emotion of the song if the user wants to switch to other moods.

Current systems give the user a list of songs where the user must choose one amongst those playlists. This can be tiresome because the user has to select the playlist that can alter his mood manually. In the proposed system, the music recommender system will have a front-end that captures the image and detects and classifies the emotion using Machine learning algorithms. We presented a system for categorizing music into categories such as happy, sad, neutral, or angry, among others. Based on the emotion detected, the list of songs will be displayed to the user [2]. After listing the songs to the user, we also provide the user with an option to switch the mood and language of the songs he is currently playing. The emotion of the user is detected using the algorithm that captures and compares the image with that stored in the local

database in the numerical inputs and classifies the image according to the detected emotion. For this, we have used libraries such as OpenCV, NumPy etc [2]. For the playlist of songs, we have collected the database from various sources and divided them into different categories using the K-means clustering algorithm [5]. So that, based on the emotion identified, the music recommender system gives the list of songs from the local storage to the user. If the user is bored listening to only one kind of songs, we also provide an option to change the mood of the song in real-time. This technique is primarily proposed because music has become increasingly important in recent years as a stress reliever. So, in order to detect emotion, we use the face as a primary source of data because facial expression typically defines emotion. Then, based on the mood, we play music that can alter the user's mood.

## II. LITERATURE SURVEY

By capturing the image of the person through a webcam it will extract the emotion. Here through Keras processing video is converted into an array of images. By scaling all images are converted to vectors. From OpenCV python library, the cvtColor method will convert the coloured image to grey colour space. Normalising the array of images into integers where each integer has predefined emotions like happy, sad, angry, and neutral by Keras classifier algorithm. We use the K-means algorithm for classifying songs and the Haar cascade algorithm to recognize emotion.

### Haar cascade Algorithm:

To train the classifier, this approach requires a large number of positive photos of faces and negative images of non-faces [3]. In the Initial stage, collect haar features. Haar features consist of Edge features, Line features, and Four-rectangle features. Cascade classifiers will take both positive and negative images as input [11]. Cascade classifier consists of a series of stages. These phases are designed to reject negative samples as quickly as possible in order to reduce the number of false negatives [7]. To identify objects in an image, we utilise a set of characteristics or classifiers. By identifying the number of haar features we can detect an object whether it is a face or not at a certain stage.

### K-means clustering algorithm:

K-means clustering algorithm is an unsupervised algorithm that is used to classify the unlabelled data into k clusters [4]. This algorithm groups the data such that data points with similarities are grouped together and those with dissimilarities are differentiated from each other. Here k refers to the number of clusters to be made. The optimal value of k can be found using the elbow method.

## III. PROPOSED SYSTEM

In this paper, we are proposing a music player that suggests songs based on the emotion of the user detected [6]. The proposed system aids us in facilitating user-music system interaction. The music player has different components, each having a functionality of its own. Initially, we created a database by collecting the list of songs from Spotify. To collect the information about songs, we have used the library spotipy. Using spotipy, we have collected information regarding songs such as artist, language and features of songs such as acousticness, danceability, loudness, liveness, speechiness etc. Then we have divided the songs into different groups based on their emotion using the K-means clustering algorithm [5]. When the user starts the music player, it initially captures the facial image of the user using a webcam. The major goal of this system is to produce a smart music player that can improve the users' mood, because music is one of the most effective mood changers. The captured face is then used to detect the emotion of the user. Users' emotions can be by using OpenCV. Here through Keras processing video is converted into an array of images. By scaling all images are converted to vectors. The emotion is detected by using the haar cascade

algorithm. Our music player can identify four kinds of emotions, namely, happy, sad, angry and neutral. Then the music player will display a list of songs to the user based on the language user selected and the emotion detected [12]. This will suggest the songs in such a way that it improves the mood of the user. This player also has a feature to change the emotion and language of the songs to be played. That is, if the user wants to listen to another mood song, then he can change it. Now the music player recommends a playlist of songs based on the emotion that is been changed.

## IV. SYSTEM ARCHITECTURE

In this project, by running the welcome page it will show the get started option where it will redirect to another page. There it will show a select language option after that it captures the user video by which it will detect emotion. It will open a new window with OpenCV where it shows a rectangular border around the image.

There it will detect the emotion by extracting the features from the face. We have many emotions like angry, happy, sad or neutral. According to the emotion detected from the video it will show the song's playlist.

For this project, we have created a database with the data collected from the spotify. For the database we have used the spotipy library. Initially, we need to set the spotipy variables client id and client secret to fetch the data from spotify api. These data is stored in and is been clustered into different groups using k-means clustering algorithm according to the k value specified. Here, we have classified the dataset into four different categories based on the emotions that the music player identifies.

After displaying the playlist, on the other hand it provides two options. One is the language change option and the other is the emotion change option for users. By clicking on submit it will redirect to the changed values.

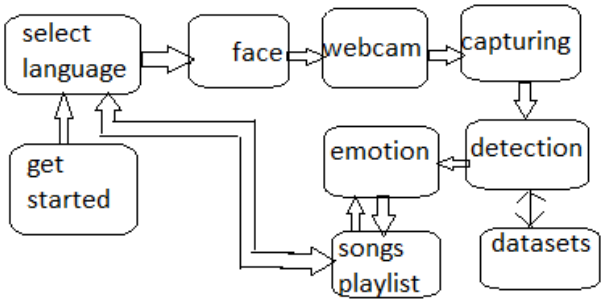
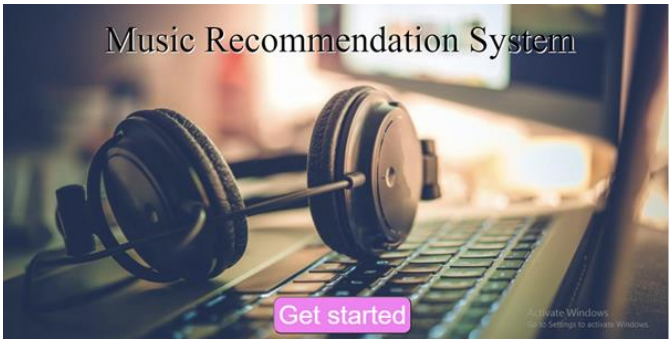



figure: Project architecture

## V. METHODOLOGY

### 5.1. Get Started:




The main objective of stated method is to start the application by clicking on displayed button which directs to another page where the intention of the theme is that the user should select the preferred language i.e. English or Telugu or Hindi which are the three options available in the displayed user interface.

### 5.2. Face Capturing

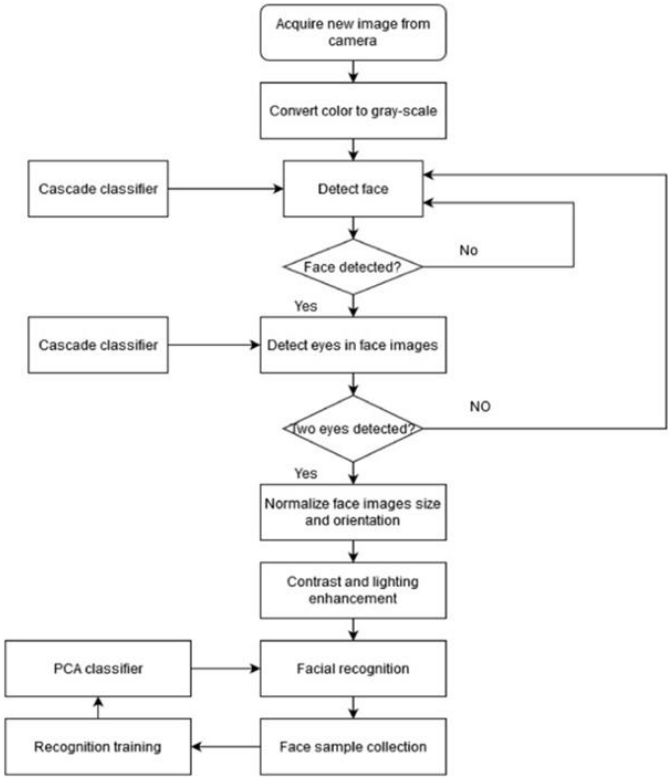
In Face Capturing session the idea is to get images of object through camera here we are using default

camera which is available in different types of devices otherwise we can use other physiological devices.

For face capturing, the openCV-python library is a collection of Python bindings aimed at resolving computer vision issues that commonly arise when utilising standard computer vision devices. In openCV there are many inbuilt-methods but for application flexibility we used cv2.VideoCapture(0) methodThis way makes it easy to combine it with other external libraries that use NumPy and Pandas, and it's mostly used for real-time computer vision applications. For application convenience video of user has been captured for 10 seconds and helps for further process. In this 10 seconds video capturing there can be positive and negative image which helps to train the classifier model.

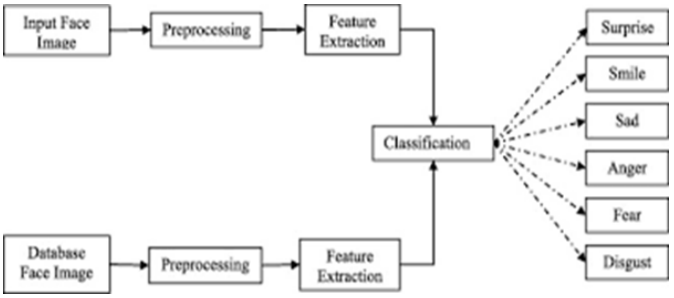
### 5.3. Face Recognition

In the application, face recognition is the main objective where emotion of the user is detected. In face capturing, as with the help of openCV, 10 seconds video of user are sent to model. As video is converted to array of images by keras processing library where each image is converted to vectors by scaling .Through cvt-color method default black and white image is colored and gray scaled and with the help of face classifier . We utilised Haar feature-based cascade classifiers since they are an excellent machine learning object detection method and cascade is a function that is taught from a large number of positive and negative images [2]. It is then used to other photos to detect things. Here scaling of each part of visioned object is done through LDA and our application needs only humans face it is vectored through scaling in X and Y-axis i.e. here eyes, nose and mouth and other visioned parts are detected through mentioned scaling.




### 5.4. Emotion Classification

Emotion classification segment is more important because it decides the scope of application. From face recognition segment as user face is detected and scaled through defined algorithms at the end a box is highlighted over the face. Now vectored array of images are processed by Keras processing classifier library and as a middleware, it uses NumPy and pandas libraries of python. Here array of images are normalized to integers where each integer is predefined emotion by Keras classifier algorithm and main method to predict emotion is classifier. predict. Here we focused on different emotions like Angry, Disgust, Fear, Happy, Neutral, Sad, and Surprise. Each emotion has scaling [10].



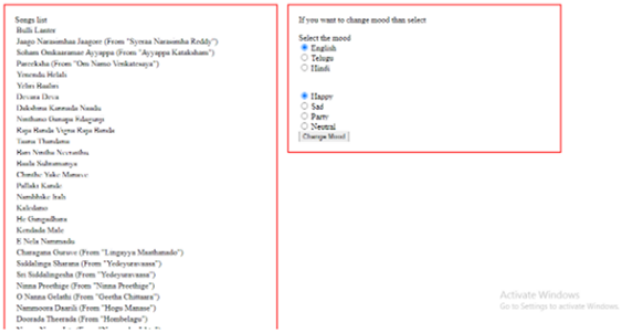
### 5.5. Music Recommendation

The image which is captured and the emotion detected by algorithms helps to display a list of song. In order to list songs, we collected a huge dataset from Spotify where each language has 500 songs from famous singers. Python is base to collect huge dataset which contains main attributes like Acousticness, Danceability, Energy, Instrumentalness, Liveness, Loudness, Valence, Tempo, Valence, and Speechiness. After collecting songs from each language we passed the whole dataset to K-means Clustering for Unsupervised Machine Learning which aims to classify the songs [1]. For classifying the emotion of songs main attributes that contributes more are Loudness, Valence, Energy, and Danceability because they have more influence to differentiate between Energetic and Relaxed songs [1]. Even though we are detecting seven different types of emotion but from the k-means algorithm we are listing four different types of list of songs we can list more by increasing value of k i.e. k=7. When it comes to Clustering, accuracy is a bit subjective when attempting to evaluate the best outcome of a Clustering Algorithm, but it's also important to see if the model is effectively separating the tracks [1]. To test the correctness of the model, we used Rstudio's Silhouette Analysis approach. The libraries "cluster" and "factoextra" in Rstudio make it easier to visualize and calculate Silhouette Analysis using the Euclidean distance.

### 5.6. Advanced feature

Here the goal of this segment is to make the application user-friendly so, in our application, there is flexibility to the user to change the language, mood and to select specific artist song list through the user interface.

## VI. RESULT




## VII. FUTURE ENHANCEMENT

Face expressions are recognized using algorithms installed on the local system. Nowadays everything is shifted to the cloud. Cloud computing has become one of the booming technologies. An alternative method, now we are using local storage for storing songs instead we can use cloud storage like AmazonS3 and deploy our model in the cloud. We can also add some additional advanced features. It can be designed in such a way that it will constantly upgrade by itself for the latest data.

## VIII. CONCLUSION

In this Project, the Music recommendation model will capture the real-time video of the user through a webcam. By identifying features of the face with the help of algorithms it will detect the emotion of the user. According to the emotion detected it will suggest a playlist. It is possible for users to switch to other songs of different emotions. Music is a stress reliever for many people. This project is designed for the purpose to reduce the user searching time for songs and creating a desired playlist according to their current emotion.

## IX. REFERENCES


[1]. https://towardsdatascience.com/

[2]. https://www.ijert.org/

[3]. Submitted to rgitkottayam Anagha S.Dhavalikar and Dr. R. K. Kulkarni, "Face Detection and Facial Expression Recognition System" 2014 Interntional Conference on Electronics and Communication System (ICECS -2014).

[4]. http://www.igi-global.com/

[5]. Intelligent Systems Technologies and Applications 2016", Springer Science and Business Media LLC, 2016

[6]. Machine Learning and Information Processing", Springer Science and Business Media LLC, 2021

[7]. Submitted to Institute of Research & Postgraduate Studies, Universiti Kuala Lumpur

[8]. sersc.org

[9]. idus.us.es

[10]. Yong-Hwan Lee , Woori Han and Youngseop Kim, "Emotional Recognition from Facial Expression Analysis using Bezier Curve Fitting" 2013 16th International Conference on Network-Based Information Systems.

[11]. Ahmed Aly Ibrahim, Zainhom Sayed Ghareeb, Omar M. Shehata, El-Sayed Imam Morgan. "A Robotic Surveillance Platform Based on an On-board Computer Vision Approach", Proceedings of the 4th International Conference on Control, Mechatronics and Automation - ICCMA '16, 2016

[12]. F. Abdat, C. Maaoui and A. Pruski, "Humancomputer interaction using emotion recognition from facial expression", 2011 UKSim 5th European Symposium on Computer T.-H. Wang and J.-J.J. Lien, "Facial Expression Recognition System Based on Rigid and Non-Rigid Motion Separation and 3D PoseEstimation," J. Pattern Recognition, vol. 42, no. 5, pp. 962-977, 2009.

[13]. Renuka R. Londhe, Dr. Vrushshen P. Pawar, "Analysis of Facial Expression and Recognition Based On Statistical Approach", International Journal of Soft Computing and Engineering (IJSCE) Volume-2, May 2012.

[14]. Anukriti Dureha "An Accurate Algorithm for Generating a Music Playlist based on Facial Expressions" : IJCA 2014. Bruce Ferwerda and Markus Schedl "Enhancing Music Recommender Systems with Personality Information and Emotional States": A Proposal: 2014.