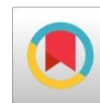# MUSIC RECOMMENDATION APPLICATION BASED ON FACIAL EXPRESSIONS

MD Arsha Sultana,
Assistant Professor-CSE Dept,
Andhra Loyola Institute of Engineering and Technology,
Jawaharlal Nehru Technological University - Kakinada
arsha.1205@gmail.com

Abdul Saherabegum, Akhila Umma, Guthula Jayasri Sai Nikitha
Department of Computer Science and Engineering,
Andhra Loyola Institute of Engineering and Technology,
Jawaharlal Nehru Technological University - Kakinada
sahera.abdul320@gmail.com, akhilaumma2544@gmail.com, jssnikitha@gmail.com

Abstract: Music is the form of art, which is known to have a greater connection with a person's emotion. It has got a unique ability to lift up one's mood. Our project focuses on building an efficient music recommendation system which determines the emotion of user using Facial Recognition techniques. It has widely attracted attention due to its enormous application value and market potential. Face recognition Technology is being implemented in various fields like security system, digital video processing, and many such technological advances. Moreover, on a larger dimension, this would render savage of time and labor invested in performing the process manually. We use Viola Jones algorithm for face recognition and Fisher Face classifier for identifying facial expressions. The overall concept of the system is to recognize facial emotion, mood and recommend songs efficiently. The proposed system will be both time and cost efficient.
Keywords: Facial Recognition, Emotion detection, Algorithms

I. INTRODUCTION

In course of history, Music is the greatest creation of mankind. In this 21st century, we see so many people attached to their headphones listening to music at any given time. People tend to listen to different types of genres in different occasions depending on their mood. Creating customized playlists for different occasions at any given time depending on the user's mood will be very much useful. By using the advanced technology, mood detection and recommendations can be possible for the better usage.

II. ABOUT THE PROPOSED WORK

A. Literature Survey

There are many algorithms and classification techniques for every problem in Machine Learning. Choosing a right method is essential to create a customized solution for different problems. This section overlooks similar existing advantages, disadvantages, similarities, measures for evaluating the algorithm, sample value of evaluations and examines the algorithms used and drawbacks. Enhancing on emotion detection and generating a playlist for the user is main task. Facial Recognition is done using Machine learning specifically Support Vector Machines. A face is initially detected by using Viola Jones algorithm. Then feature extraction on face is performed using Histogram of Oriented Gradients which essentially stores the edge of the face as well as the directionality of the edges. Training and classification of the facial databases is done using the multi class SVM where each unique face in the Facial database is a class. The system uses user's image using camera, detects the face and emotions to recommend a playlist which can enhance the user. This approach minimizes the efforts by suggesting the user a list of songs based on his current emotion.

B. Project Objective

Everyone battle with emotions due to different reasons, ultimately everyone wants to be happy. Our primary goal is to recommend music by detecting the mood of a person. As music holds the power to cheer up a person's mood, our project aims to recommend the best playlist which matches the mood of a person and ends with cheerful music. Through this project, we are creating an application which can make a person happier by detecting the mood of a person using facial recognition techniques.

C. Proposed Work

In this section, the proposed work is elaborated at a high-level scope. Here we can understand the user interface and working nature of the application

1) Design Methodology:

The proposed system recommends music that is based on automatic emotion detection. A webcam is used to capture the images that will be used as input to the proposed system, and then it goes to the expression detector to classify it to one of four classes "Happy", "Sad", "Neutral" and "Angry" and then it recommends a song as the figure below.
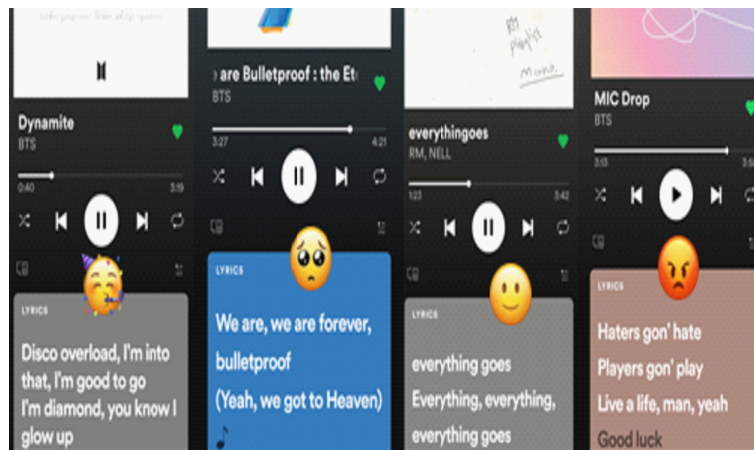

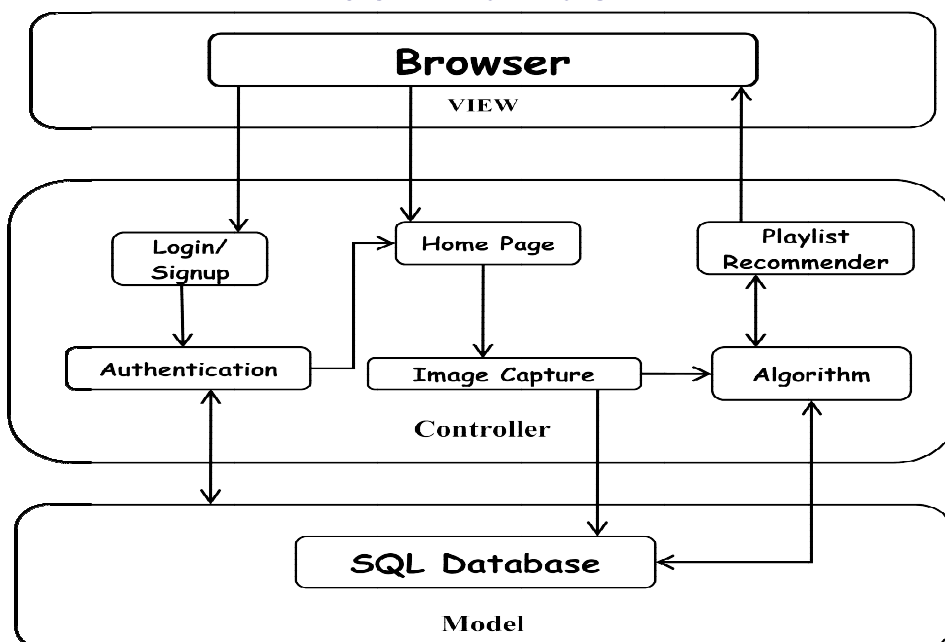Fig.1 Song Recommendations based on User's mood

II. SYSTEM ARCHITECTURE:

Fig.2 System Architecture of the Application

In System Architecture, we focus on both back-end and front-end performance.

A. Image Capturing

At first, we begin by taking the consent of the user whether they want to listen to music based on mood or not. If they want to listen to music based on their mood, the system captures the image of the user's face using a built-in laptop webcam (or any external camera can be employed). The face image to be correctly processed in the proposed system must contain one face in the frontal position in a uniformly illuminated background.

B. Face Detection

After acquiring the image, the system will start to detect the face by applying the Viola-Jones algorithm. This Algorithm is considered as one of the first frameworks that recognize the objects in real time. Simply, Viola-Jones scans the images using a sub-window to detect the features of the face in the image. When the face is determined, the image is cropped to contain the face only, to enhance the proposed system performance. Also, the Viola-Jones is reused to identify and crop the left and right eyes and mouth separately. The outcome of this step is four images, face, right eye, left eye, and mouth images. These specific outcomes will be used to later for Emotion Classification.

C. Emotion Classification

The detected human face is set to be processed further for recognizing the mood by using the Fisher Face classifier. Fisher Face is one of the popular algorithms used in face recognition, and is widely believed to be superior to other techniques, such as Eigen face because of the effort to maximize the separation between classes in the training process. Image recognition using Fisher Face method is based on the reduction of face space dimension using Principal Component Analysis (PCA) method, then apply Fisher's Linear Discriminant (FDL) method or also known as Linear Discriminant Analysis (LDA) method to obtain feature of image characteristic. After obtaining the emotion, the user is asked to confirm the emotion. For any changes, the application allows the user to recapture the image.

D. Playlist Recommendation

Based on the user's mood which is identified as one of the emotions Happy, Sad, Neutral or Angry, the respective playlist will be displayed. For happy emotion, the cheerful and party type playlist will be recommended. For neutral emotion, classy and edm type music will be recommended to savour the moment. For emotions such as sad and angry, a playlist will be recommended with r&b, pop music which will be able to enhance the user mood to lighten up and feel at peace.

III. ALGORITHMS

The algorithms used in Music Recommendation application based on Facial Expressions are
* Viola Jones Algorithm
* Fisher Face classifier
* Spotipy Library

We shall discuss the working of the algorithms in detail.

Viola Jones Algorithm

* This standard algorithm is used to recognize a human face among different objects captured by the camera in a frame.
* By selecting Haar-like features, identifying a face can be accomplished by distinguishing and grouping into these three types and interpreting the parts of a face.
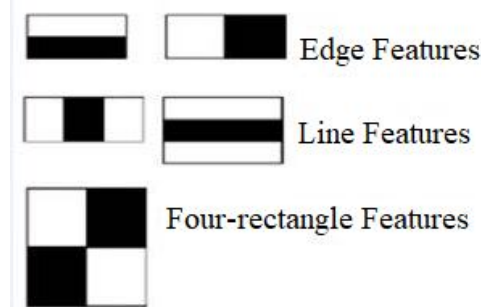

Fig.3 Haar-like Features

* Later an integral image (also known as summed area table) is created from the original image. It is used as a quick and efficient way to calculate the sum of pixel values in an image or rectangular part of an image.
* In the Viola-Jones algorithm, each Haar-like feature represents a weak learner. To decide the type and size of a feature that goes into the final classifier, Adaptive Boosting (AdaBoost) checks the performance of all classifiers that you supply to it.
* To calculate the performance of a classifier, you evaluate it on all sub regions of all the images used for training.
* Some sub regions will produce a strong response in the classifier. Those will be classified as positives, meaning the classifier thinks it contains a human face.
* Sub regions that don't provide a strong response don't contain a human face, in the classifiers opinion. They will be classified as negatives.
* The classifiers that performed well are given higher importance or weight. The final result is a strong classifier, also called a boosted classifier that contains the best performing weak classifiers.

- So when we're training the AdaBoost to identify important features, we're feeding it information in the form of training data and subsequently training it to learn from the information to predict. So ultimately, the algorithm is setting a minimum threshold to determine whether something can be classified as a useful feature or not.
- We set up a cascaded system in which we divide the process of identifying a face into multiple stages. In the first stage, we have a classifier which is made up of our best features, in other words, in the first stage, the sub region passes through the best features such as the feature which identifies the nose bridge or the one that identifies the eyes. In the next stages, we have all the remaining features.
- When an image sub region enters the cascade, it is evaluated by the first stage. If that stage evaluates the sub region as positive, meaning that it thinks it's a face, the output of the stage is maybe.
- When a sub region gets a maybe, it is sent to the next stage of the cascade and the process continues as such till we reach the last stage.
- If all classifiers approve the image, it is finally classified as a human face and is presented to the user as detection.

Fisher Face Classifier
- Fisher face classifier is one of the most widely used face mood detection technique because of it's effort to maximize the separation between classes in the training process.
- At this image processing stage, Fisher face method will be applied to generate feature vector of facial image data used by system and then to match vector of traits of training image with vector characteristic of test image using Euclidean distance formula
- Image recognition using Fisher Face method is based on the reduction of face space dimension using Principal Component Analysis (PCA) method, then apply Fisher's Linear Discriminant (FDL) method or also known as Linear Discriminant Analysis (LDA) method to obtain feature of image characteristic.
- The results show that for image recognition where the image of testing is the same as the training image, the percentage of program success is 100%, while for 73 facial test images with various expressions and various positions, 70 faces are recognized correctly and 3 faces are recognized incorrectly, so the percentage of success is 93%
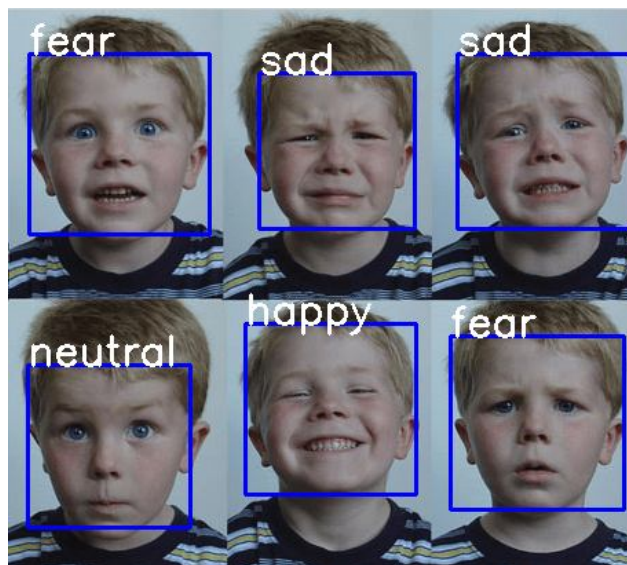

Fig.4 Emotion Detection using Fisher Face Classifier

Usage of Spotipy library
- Spotify is a streaming app with a powerful API functionality. The API is both beginner-friendly and offers a wide range of endpoints.
- Spotipy being a popular wrapper for the Spotify API enables and supports all the features of the Spotify API but allows them to be designed and run in Python. This allows for scalability and also will come in handy when we start dealing with large amounts of data.
- Spotify uses scopes to ensure programs only have permission to do what you specify they can do. We always need to specify the scope you want to use.
- Content based recommendation will be used for recommending Music playlist based on the mood detected.
- By using pre-defined clusters based on genres which are available on Spotipy, the playlist can be fetched by using nearest neighbour.

IV. RESULTS AND OBSERVATIONS

- Viola Jones algorithm was developed in 2001 by Paul Viola and Michael Jones as an object detection framework that allows the detection of image features in real time.
- Despite being an old algorithm, it's quite powerful and its applications are proven to be exceptionally notable in real-time face detection.
- The algorithm used in the process for image recognition is Fisher Face algorithm while for identification or matching face image using minimum Euclidean.
- The Fisher Face method is relatively insensitive to substantial variations in light direction, face pose and facial expressions. This is accomplished by using both Principal Component Analysis and Fisher's Linear Discriminant Analysis.
- Spotify being a well-built module, also has an active team supporting development, which keeps the environment catastrophic bugs free, but it is still better to get in the habit of creating virtual environments.
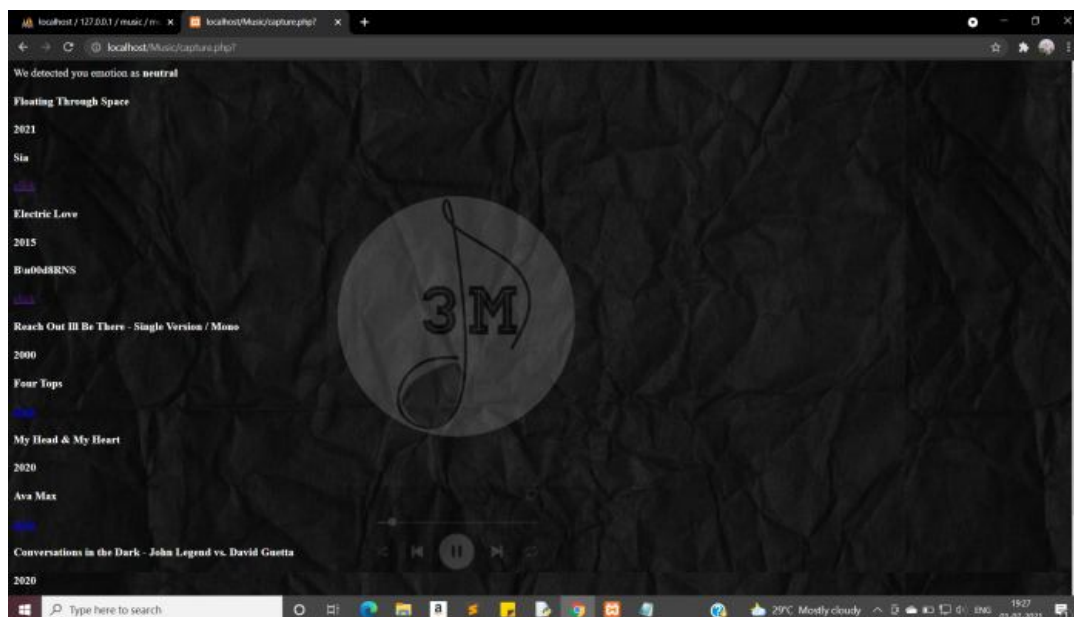

Fig.5 Music Recommendation for Neutral mood person

V. CONCLUSIONS

This study focuses on the music recommendation by the system to a user based on the user's emotions. This kind of approach will be very useful and the user can save lot of time creating playlists. As the data is collected from the internet, this approach will save storage issues and comes with cost cutting advantages.

ACKNOWLEDGMENT

REFERENCES

1. Arohi Gupta, "Emotion Detection : A machine learning project", 2019, https://towardsdatascience.com/emotion-detection-a-machine-learning-project-f7431f652b1f
2. Arun Alvappillai, Peter Neal Barrina, "Face Recognition using Machine Learning" http://noiselab.ucsd.edu/ECE285/FinalProjects/Group7.pdf
3. So-Hyun Park, Sun-Young Ihm, Wu-In Jang, Aziz Nasridinov, and Young-Ho Park, "A Music Recommendation Method with Emotion Recognition Using Ranked Attributes", 2020 https://sci-hub.se/https://doi.org/10.1007/978-3-662-45402-2_151
4. Mikhail Rumiantcev, Oleksiy Khriyenko, "Emotion based Music Recommendation System", 2020 https://fruct.org/publications/acm26/files/Rum.pdf
5. Iyer, A.V., Pasad, V., Sankhe, S. R., Prajapati, K., "Emotion based mood enhancing music recommendation", 2017 https://sci-hub.se/10.1109/RTEICT.2017.8256863
6. Mustamin Anggo1 and La Arapu, "Face Recognition using Fisher Face Method" https://iopscience.iop.org/article/10.1088/1742-6596/1028/1/012119/pdf