Song Recommendation Based on Facial Expression

Aishwarya Shetty[1], Apeksha Prabhu[2], Prof. Suparna[3], Dr. M. Shahina Parveen[4]. [1, 2] Department of Computer Science and Technology, Dayananda Sagar University, Bengaluru, Karnataka, India. [3] Assistant Professor of Computer Science and Technology, Dayananda Sagar University, Bengaluru, Karnataka, India. [4] Professor and Chairperson of Computer Science and Technology, Dayananda Sagar University, Bengaluru, Karnataka, India.

Abstract - A user's facial expressions can reveal his or her level of emotion. These expressions can be obtained from the system's camera's live feed. In the area of computer vision (CV) and machine learning (ML), a lot of research is being done to train machines to recognize different human emotions or moods. Machine learning offers a variety of methods for detecting human emotions.

A review of existing music systems revealed that many music applications rely on the user's past listening choices rather than recommending songs based on their current emotion. The goal of this project is to identify emotions in human faces using real-time data and to suggest songs according on those emotions.

Music is a great unifier. It binds us despite our differences in ages, backgrounds, languages, interests and levels of income. Due to its accessibility and ability to be used alongside daily activities, travel, sports, and other activities, music players and other streaming apps are in high demand. Digital music has emerged as the main form of consumer content that many young people are looking for because to the quick growth of mobile networks and digital multimedia technology.

Music is frequently used by people as a tool for mood control, specifically to improve mood, boost energy, or soothe tension. Additionally, listening to the correct music at the right moment can help with mental wellness. So, music and feelings in people are closely related.

As a result, the proposed system is an interactive platform for suggesting music depending on user's present emotional state. This also could be a great feature to be incorporated in existing music player applications

Keywords – AI,ML,CV,OpenCV,KDEF,VGG,CNN,CK,CK+,RaFD,MIREX,ANN.

I. Introduction

Numerous studies conducted over the past few years have shown that music has an impact on people's emotions, actions, and cognitive processes. One of the most crucial purposes of music, according to researchers studying the reasons people listen to music, is its relationship with our emotion. Two of music's most important applications are its ability to raise participants' levels of self-awareness and mood. It has been demonstrated that emotional states and personality traits are closely related to musical preferences.

A wide range of applications, including smart card applications, surveillance, image database investigation, criminal, video indexing, civilian applications, security, and adaptive human-computer interfaces in multimedia environments, use emotion detection as the most important technique at the moment.

Thanks to developments in digital signal processing and other effective feature extraction algorithms, automated emotion detection in multimedia attributes like music or movies is growing quickly. This system can play a significant role in many potential applications, such as music entertainment and human-computer interaction systems. We offer a technology for emotion identification that recommends music based on user moods as determined by their facial expressions. The proposed system can identify a person's emotions, and if the person is down, a playlist of the cheeriest, musically-related songs will be played.

Additionally, if the emotion is positive, a specific playlist of songs from different musical genres will be provided to reinforce the positive feelings.

Facial expression recognition is implemented using convolutional neural networks, which have a 95% accuracy rate.

II. Literature Survey

1. In a particular system [1], Anaconda and Python 3.5 softwares were used to test the functionality and Viola-Jones and haar cascade algorithms were used for face detection. Similarly, KDEF (Karolinska Directed Emotional Faces) dataset and VGG (Visual Geometry Group) 16 were used with CNN (Convolution Neural Network) model which was designed with an **accuracy of 88%**, for face recognition and classification that validated the performance measures.

2. Another system [2] used Python 2.7, OpenSource Computer Vision Library (OpenCV) & CK (Cohn Kanade) and CK+ (Extended Cohn-Kanade) database which gave approximately **83% accuracy.**

3. It was observed in a cross-database experiment [3] that raw features worked best with Logistic Regression for testing RaFD (Radboud Faces Database) database and Mobile images dataset. The **accuracy achieved was 66% and 36%** respectively for both using CK+ dataset as a training set. The additional features (distance and area) reduced the accuracy of the experiment for SVM (Support Vector Machine) from 89%.

4. There has also been research done on the Music Recommendation System. According to one such research [4], a preliminary approach to Hindi music mood classification has been described, that exploits simple features extracted from the audio. MIREX (Music Information Retrieval Evaluation eXchange) mood taxonomy gave an **average accuracy of 51.56%** using the 10-fold cross validation.

5. Renuka R Londhe et al. [5] proposed a paper which focused on the study of changes in the curvatures of the face and the intensities of the corresponding pixels. The author used Artificial Neural Networks (ANN), which was used to classify the emotions. The author also proposed various approaches for a playlist.

6. Zheng et al. [6] proposed two significant categories for facial feature extraction, which included Appearance-based feature extraction and geometric based feature extraction, which included extraction of some essential points of the face such as mouth, eyes, and eyebrows.

III. Problem Statement

Most of the existing music applications do not take user's facial expression as an input parameter to recommend songs. That drawback inspired us to develop an application that focuses on the same. The application we are developing will use **AI** to detect user's facial expression in **real-time** and recommend songs solely based on that.

Though we think that focusing on user's emotion would make our application unique, we are aware that its not the only contributing factor involved in building a software product. There are a lot of corner cases to be taken care.

One major case is that not all users are expressive. Even if they are expressive, sometimes due to bad lighting, the application might not be able to detect the user's face. In such cases the user can select the mood manually, just like how we select genre or artist in other widely used music applications.

The application works for four expressions as of now. Those are **happy**, **sad**, **angry** and **neutral**. These are the most common emotion expressed by people in real life. We are planning to add more if our application produce accurate results for the existing emotions.

With accuracy, it is also our priority to provide good user experience with attractive user interface.

IV. Problem Solution

The project being developed is a music application named **Symphony**. Symphony's aim is to play music based on user's facial expression and this feature makes symphony stand out when compared to existing music applications. The majority of the current methods include manually playing music, wearing wearable computers, or categorizing based on auditory attributes. We suggest changing the manual sorting and playing instead.

The music application is in the form of a website developed using **Flask**, a web application framework written in Python. And for database, it uses **Sqlite3**. The database is used to store user account details such as email, name and password which are required fields.

On logging in, the user will be provided with two options, either to select his or her mood manually or to let the webcam detect it. After the application detects the user's mood, it automatically plays the song based on the detected mood.

There are several techniques to detect face, but we will use the **haar cascade** algorithm because its not so complex and can run in real-time. Haar cascades were first introduced in 2001, and it was one of the most popular object detection algorithms in **OpenCV**.

Ⅴ. Methodology

Real-time mood recognition is the main goal of the programme known as the mood-based music recommendation system. It is a prototype for a brand-new product that has two primary modules:

Music suggestion and facial expression recognition/mood detection.

1)Module for detecting mood: The following two sections make up this module:

● **Face Recognition**: the capacity to locate faces in any input picture or frame. The discovered faces bounding box coordinates are the output. The Python library cv2 was taken into consideration for this purpose. **cv2 is the module import name for opencv-python**, "Unofficial pre-built CPU-only OpenCV packages for Python". The traditional OpenCV has many complicated steps involving building the module from scratch, which is unnecessary.

● **Mood Recognition:** Emotions on the face are categorized as pleased, angry, sad, neutral,
surprised, fearful, or disgusted. The typical Keras module of Python was utilized for this purpose. Haarcascade algorithm was employed for detection of the emotions. The image of the user is captured with the help of a camera/webcam. Once the picture captured, the frame of the captured image from webcam feed is converted to a grayscale image to improve the performance of the classifier, which is used to identify the face present in the picture. Once the conversion is complete, the image is sent to the classifier algorithm which, with the help of feature extraction techniques can extract the face from the frame of the web camera feed. From the extracted face, individual features are obtained and are sent to the trained network to detect the emotion expressed by the user. These images will be used to train the classifier so that when a completely new and unknown set of images is presented to the classifier, it is able to extract the position of facial landmarks from those images based on the knowledge that it had already acquired from the training set and return the coordinates of the new facial landmarks that it detected. The network is trained with the help of CK extensive data set. This is used to identify the emotion being voiced by the user

2)Music Recommendation Module

On Kaggle, a dataset of songs labelled according to mood was discovered for Hindi and English. In order to save, retrieve, and query this song's data as needed by users, research was done to find a reliable cloud storage platform. There were alternatives like AWS, Google Cloud, etc., but they were rejected since they were expensive and offered very little storage for free. After then,this project is built on recommending songs using Youtube which is online video sharing and social media platform. So when the user's emotion is detected, the songs corresponding to his facial expression is suggested so that the user can choose which song he/she wants to listen to.

3)Integration

Login and Signup Authentication modules were built using Flask and frontend using HTML, CSS and JavaScript. After the user signs up the home page was built using python, then the user can either input his emotion or can choose real time mood detection. The former is done using JavaScript button functions where when the user clicks on the button, the particular music playlist specific to that mood is displayed to the user and the user can choose the song he wants to listen to, whereas the latter is built using JavaScript using Navigator media Device which is a read-only property that returns a Media Devices object, which helps us to access the connected media input devices like camera and microphone and to detect the emotion of the user this project has been built using cv2 and haarcascade algorithm (HAAR.js image feature detection based on HaarCascades in JavaScript (Viola-Jones-Lienhart et al Algorithm) and FILTER.js video and image processing and computer vision Library in pure JavaScript (browser and node).Then keras model for detection of the emotion and the cv2 module will link the emotion to music playlist on YouTube.

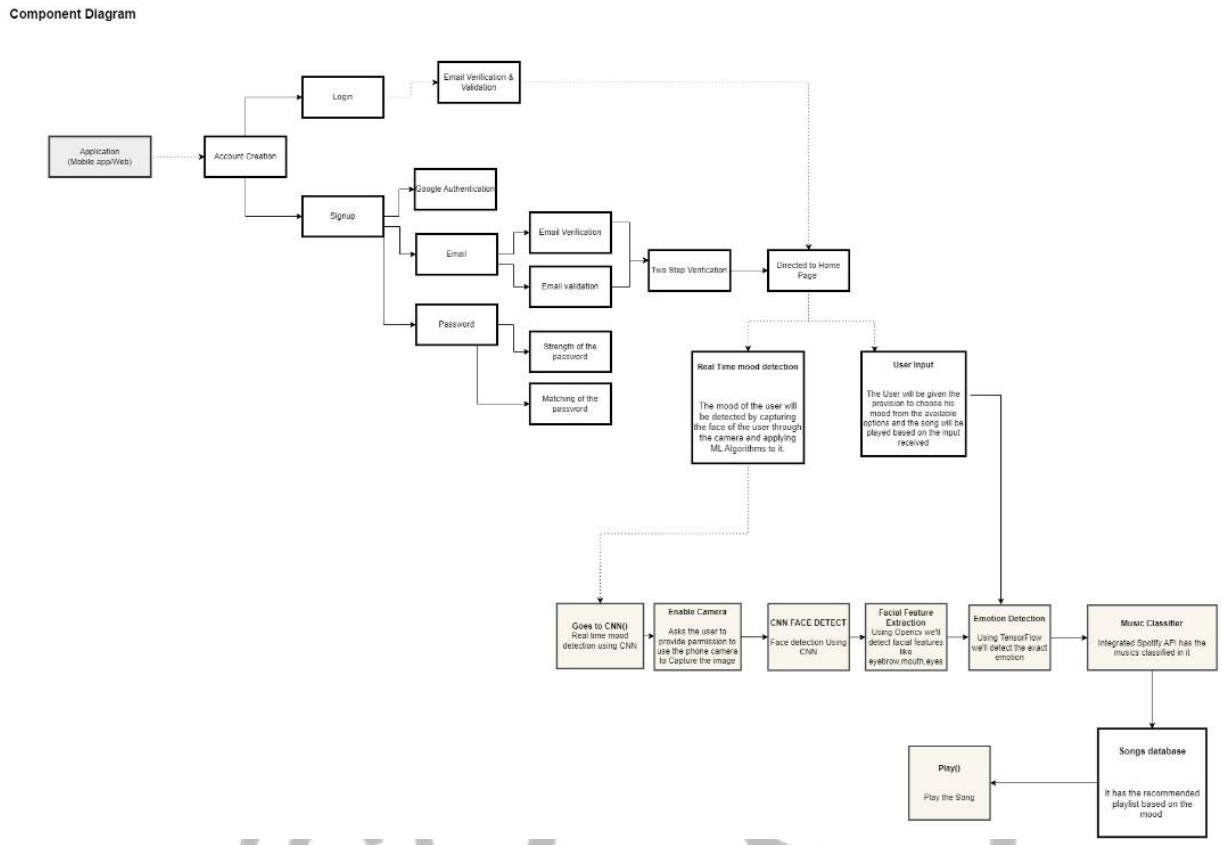


Fig1:Displays the System Architecture

The system architecture diagram shows the interactions, limitations, and boundaries between components as well as the general layout of the software system. The app will open to the home screen, which has three buttons: take pictures, provide imput, and play music. When the "take snap" button is clicked, the camera opens and the user snaps the photo. This image is used as the input for a facial recognition algorithm. An appropriate error message is presented to the user if no face is discovered or if numerous faces are detected. The image is sent as input to the mood detection module after successful single face detection.

The user is shown the detected mood, and the "play tunes" button then becomes active.

The user can choose and play a song from the playlist that is appropriate for the identified mood, as illustrated in Fig.1 screen with five emotions button will appear if the user clicks the "how do you feel" button, as seen in Fig1 as user input. Any button may be clicked by the user to access the corresponding playlist. The user only needs to touch the back button to leave the app.

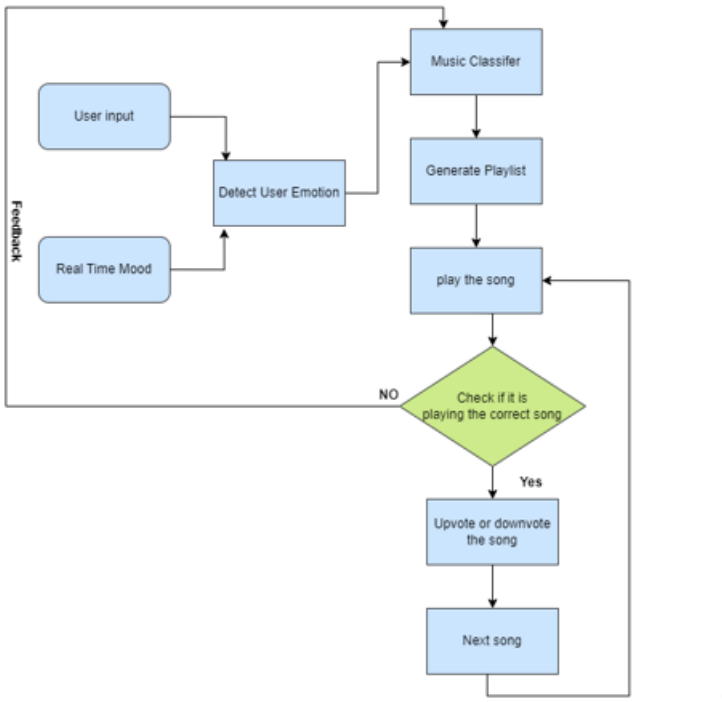


Fig2 :Shows the data flow of the system

Ⅵ. Result Analysis and Conclusion

Accurately identifying human emotion or mood is challenging since each person has distinctive facial traits. But it may be recognised to some extent with the right facial expressions.
The device's camera needs to have a greater resolution.Python-based recognition It takes a lot of work to develop this system since several technologies must be combined with the programme. The users and music enthusiasts of this system stand to benefit greatly. Even if this system may still benefit from refinement, the primary goal of this project is to play music that corresponds to a person's emotions, and as of now, that goal has been met to some extent. However, not every task in this development sector can be stated to be flawless. 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 use 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 face emotion-based music recommendation system will lessen users' playlist creation and management tasks. We have picked up a lot of knowledge about the subject of development and learnt a lot of new things. We anticipate success from this.

OUTPUT SCREENSHOTS:

Fig3:Home Page:

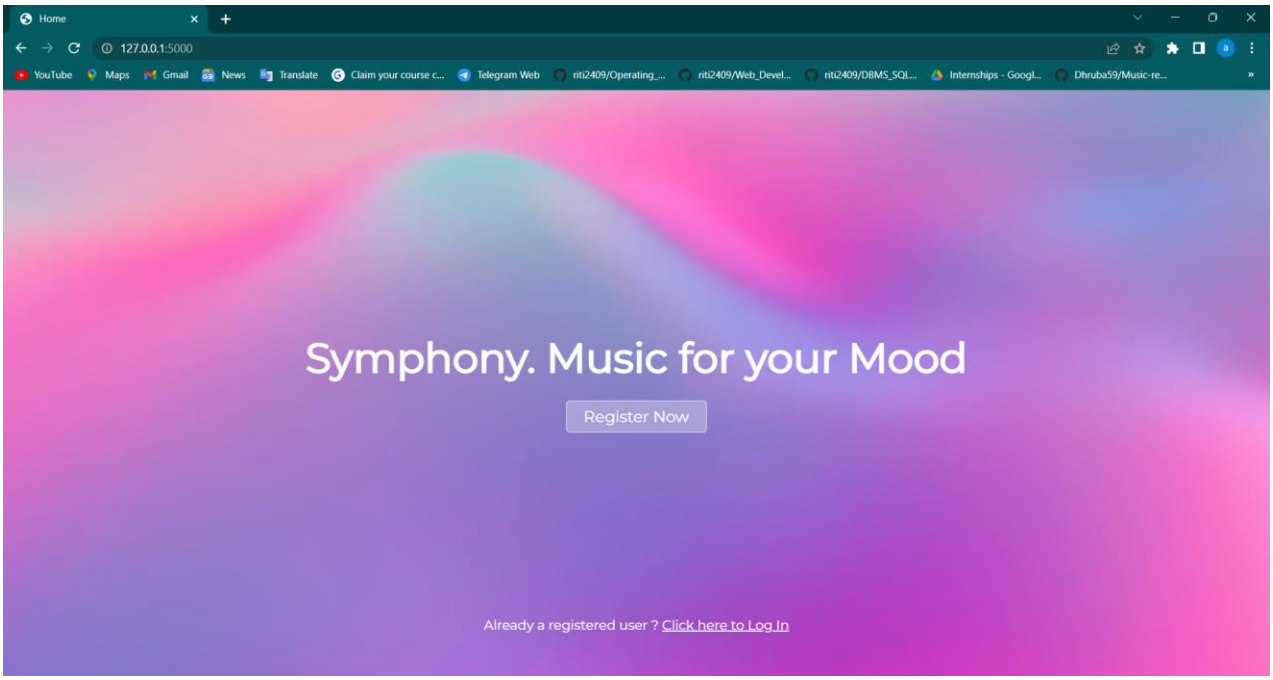


Fig 4 Signup

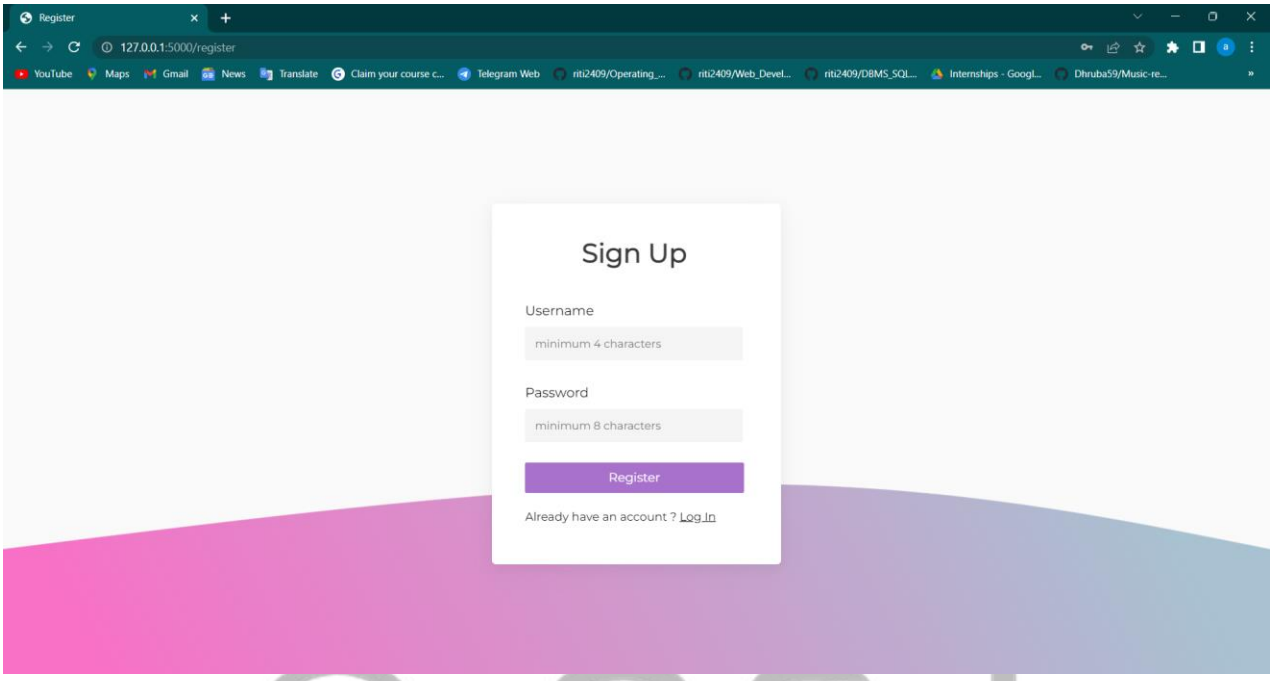


Fig 5: Login

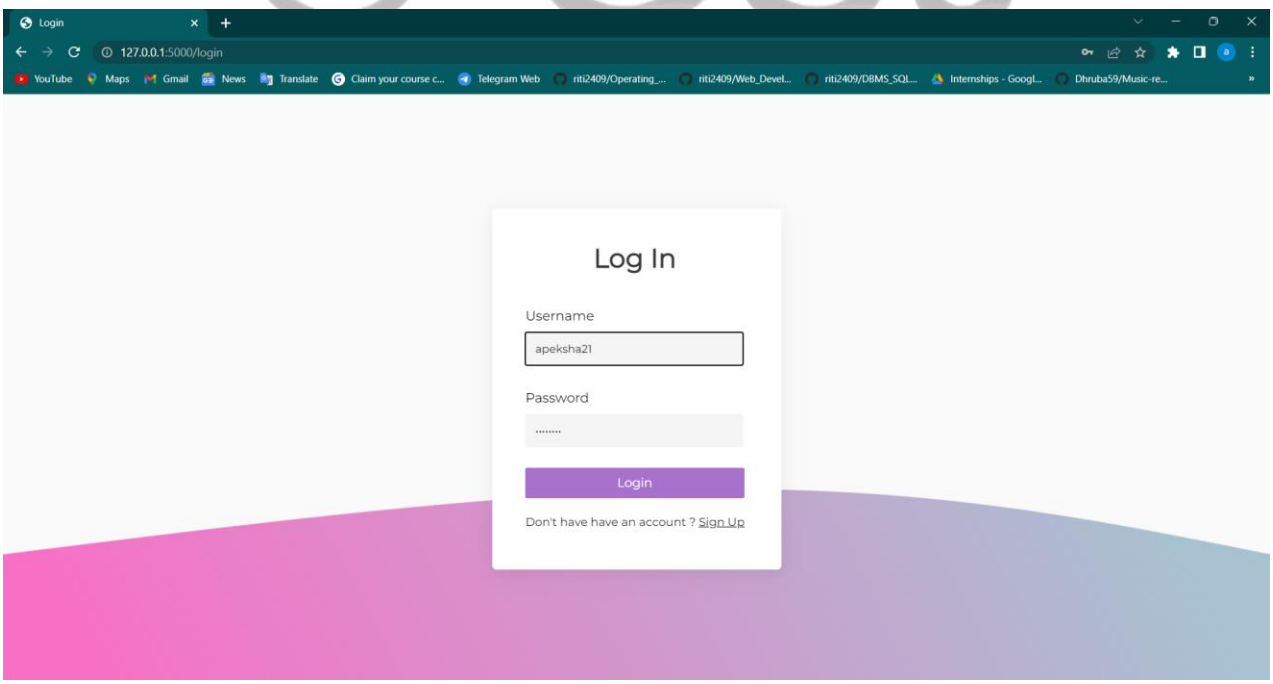


Fig6: Dashboard

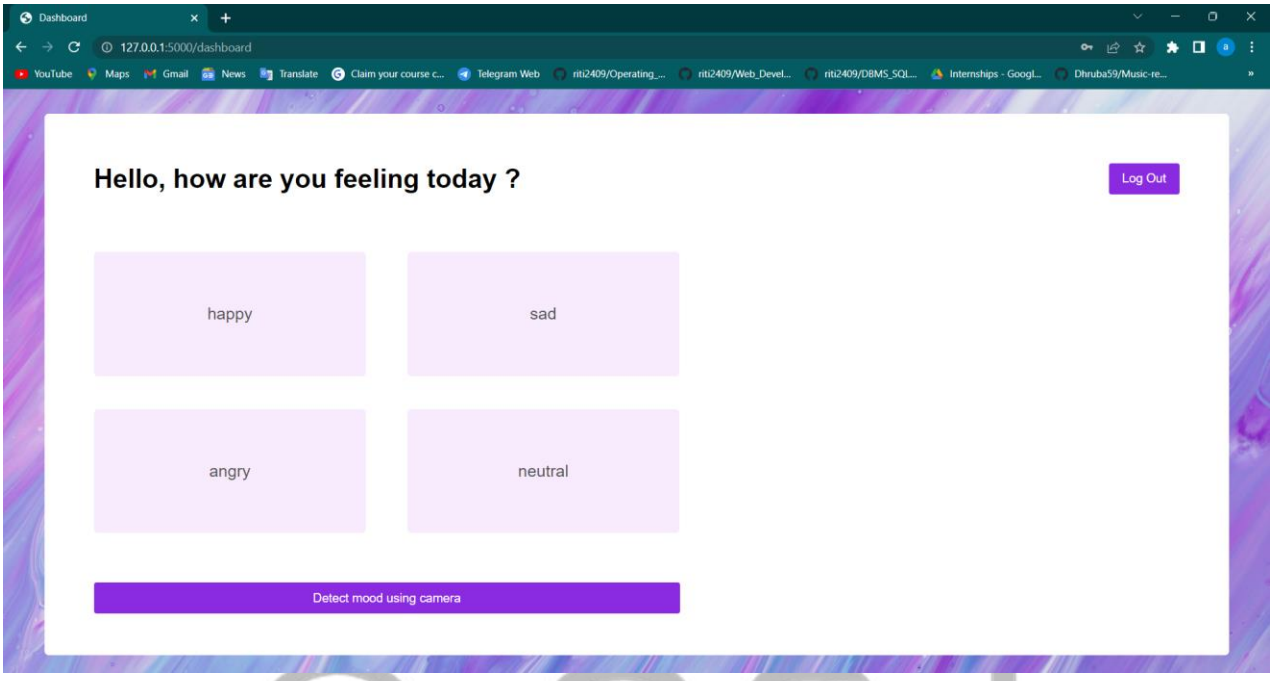


Fig 7: Webcam usage and emotion detection

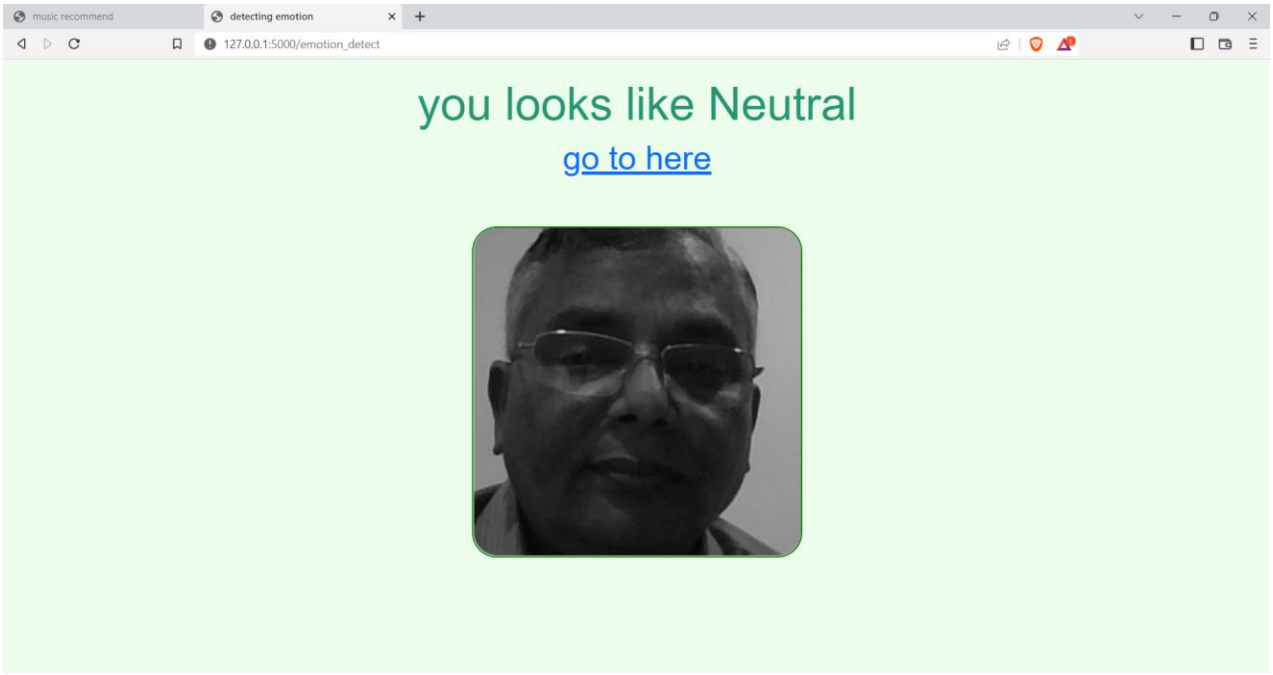
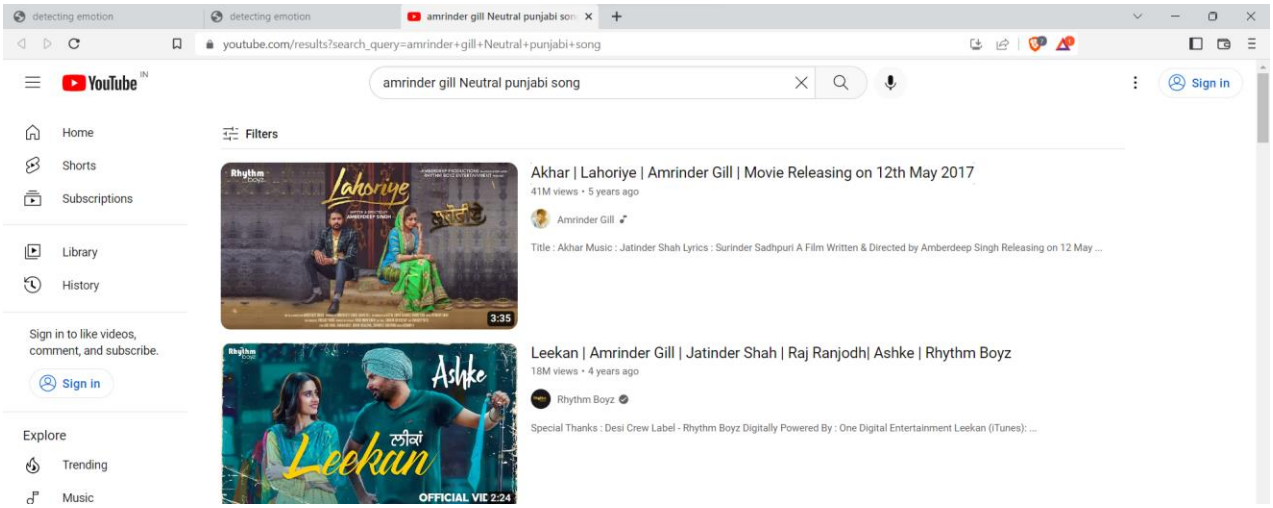


Fig 8: Music recommendation



Ⅶ. References

[1] Dr. Shaik Asif Hussain and Ahlam Salim Abdallah Al Balushi, “A real time face emotion classification and recognition using deep learning model”, 2020 Journal. of Phys.: Conf. Ser. 1432 012087.

[2] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar and I.Matthews, "The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression," 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, San Francisco, CA, USA, 2010, pp. 94-101, doi:10.1109/CVPRW.2010.5543262.

[3] Raut, Nitisha, "Facial Emotion Recognition Using Machine Learning" (2018). Master's Projects. 632. https://doi.org/10.31979/etd.w5fs-s8wd

[4] Patra, Braja & Das, Dipankar & Bandyopadhyay, Sivaji. (2013). Automatic Music Mood Classification of Hindi Songs.

[5] Londhe RR and Pawar DV 2012 Analysis of facial expression and recognition based on
statistical approach International Journal of Soft Computing and Engineering 2.

[6] Zeng Z, Pantic M, Roisman GI and Huang TS 2008 A survey of affect recognition methods Audio, visual, and spontaneous expressions IEEE transactions on pattern analysis and machine intelligence 31 39-58

[7] Ketki R. Kulkarni and Sahebrao B. Bagal ,“Facial expression recognition”, 2015 International Conference on Information Processing (ICIP).

[8] V. R. Ghule, A. B. Benke, S. S. Jadhav, and S.A. Joshi, “Emotion Based Music Player Using Facial Recognition,” vol. 5, no. 2, p. 7, 2007. International Journal of Innovative Research In Computer and Communication, 2017.

[9] Gokul Krishnan K, Parthasarathy M, Sasidhar D and Venitha E, "Emotion detection and music recommendation system using machine learning" in 2018 International Journal of Pure and Applied Mathematics, 2018, vol. 119, pp. 1487-1498.

[10]Shlok Gilda, Husain Zafar, Chintan Soni and Kshitija Waghurdekar, “Smart Music Player Integrating Facial Emotion Recognition and Music Mood Recommendation”, presented at the IEEE WiSPNET 2017 conference.