REVA UNIVERSITY
Bengaluru, India

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

A MINI-PROJECT REPORT

ON

"MUSIC RECOMMENDATION USING FACIAL RECOGNITION"

Submitted in partial fulfilment of the requirements for the award of the

Degree

of

BACHELOR OF TECHNOLOGY

IN

COMPUTER SCIENCE AND ENGINEERING

Submitted by

MANISH MAURYA (R21EJ019)

Under the guidance of

Prof Afifa Salsabil Fathima

Assistant Professor,
School of CSE

REVA University

Rukmini Knowledge Park, Kattigenahalli, Yelahanka,Bengaluru-560064
www.reva.edu.in

May 2024

REVA UNIVERSITY
Bengaluru, India

DECLARATION

I **MANISH MAURYA (R21EJ019)** student of B.Tech., VI Semester, School of Computer Science and Engineering(CS&IT), REVA University declare that the Mini-Project Report entitled "Music recommendation using facial recognition" done by us under the guidance of Prof Afifa Salsabil Fathima, Assistant Professor, School of Computer Science and Engineering,REVA University.

I am submitting the Mini-Project Report in partial fulfilment of the requirements for the award of the degree of Bachelor of Technology in Computer Science and Engineering(CS&IT) by the REVA University, Bengaluru during the academic year 2023-24.

I further declare that the Mini-Project or any part of it has not been submitted for award of any other Degree of REVA University or any other University / Institution.

Manish Maurya(R21EJ019)
Signature:
Date:

REVA UNIVERSITY
Bengaluru, India

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

CERTIFICATE

This is to certified that the Mini-Project entitled **"Music recommendation using facial recognition"** carried out under my guidance for **Manish Maurya (R21EJ019)** bonafide student of REVA University during the academic year 2023-24. The above-mentioned student is submitting the Mini-Project report in partial fulfilmentfor the award of Bachelor of Technology in Computer Science and Engineering during the academic year 2023-24**.** The Mini-Project report has been approved as it satisfies the academic requirements in respect of Mini-Project work prescribed for the said degree.

Signature with date

Guide

Name of the Examiner	**Signature with Date**
1.	
2.	

Abstract:

This project explores the potential of applying Convolutional Neural Networks (CNNs) for music recommendation based on facial emotion recognition. We propose a system that leverages the power of deep learning to personalize music suggestions by analyzing the user's emotional state through facial expressions captured by a webcam or camera.

The current music streaming landscape, while offering vast music libraries, often struggles to provide a truly personalized and engaging user experience. Existing recommendation algorithms primarily rely on factors like listening history and user preferences, which might not always capture the user's current emotional state or specific needs. This project aims to address this gap by developing a novel music recommendation system that incorporates the user's emotions, detected through facial recognition, into the recommendation process.

Project Overview
The proposed system utilizes two CNNs:

Facial Emotion Recognition CNN: This CNN is trained on a dataset of labeled facial expressions to recognize emotions like happiness, sadness, anger, or surprise from the user's face.

Mood Mapping to Valence-Arousal: The project employs a rule-based approach to convert the detected moods from the facial expression model into valence (positive/negative) and arousal (energy level) valueson a scale from 0 to 1. This mapping is based on a general understanding of emotional connotations:

Happy: High Valence (0.7 - 1), Medium-High Arousal (0.6 - 1)
Sad: Low Valence (0 - 0.4), Low Arousal (0 - 0.4)
Angry: Low Valence (0 - 0.4), High Arousal (0.7 - 1)
Calm: High Valence (0.7 - 1), Low Arousal (0 - 0.4)

The system integrates the outputs by mapping the user's detected emotion to the corresponding music emotion category, the system recommends songs aligned with the user's current emotional state, aiming to enhance their listening experience and potentially address emotional regulation needs.

This project acknowledges the challenges associated with facial expression interpretation and music emotion classification. We acknowledge the need for further research and development to ensure the accuracy and reliability of the system. Additionally, the project emphasizes the importance of respecting user privacy and providing them with control over their data and the recommendation process.

I. Introduction:

Today, whenever a person wants to listen to music, he has unlimited access everywhere. Be it music apps like spotify, wynk or looking up for a music video on youtube or any other platform. Music is a big part of our life, as it seems to give life to even the lifeless things. Even though with the music platforms today we can choose any music that we want to , but this is not always the best scenario because we don't always want to search the particular song that we want to listen to, sometimes we just want to hear a song depending on our mood. This is where we come with our idea.

Consider a scenario where an individual is not having a good day, he or she has worked all day long and now doesn't have the energy to do anything. Then, the individual goes to the music app and looks for a music that can uplift his mood but he fails to find a song that matches his wish.

Similarly, when an individual looks for a song in the music app they will be presented with something that they have listened to in the past but not something that matches their taste and is new in the market. Either the section will be divided into something like a "what's new" section that has all the latest songs or like a "recommended" section that has songs we listen to very often.

Motivated by such scenarios, we aim to create a platform where the individual will be recommended songs based on their real time facial expressions. By leveraging advanced technologies such as Cnn, OpenCV, Tensorflow.

Integrating facial recognition into a music recommendation system aims to enhance user experience by providing a more personalized and intuitive way to interact with technology. Recognizing facial expressions can help understand users' moods and preferences, tailoring music recommendations to match their emotional state.

II. Literature Survey:

[1] Facial emotion recognition using deep learning: review and insights,Wafa Mellouk, Wahida Handouzi

This paper discusses recent advances in facial emotion recognition (FER) using deep learning techniques from 2016 to 2019. It presents several popular databases used in FER research and reviews state-of-the-art methods that employ deep convolutional neural networks (CNNs) and recurrent neural networks (RNNs). Different deep learning architectures proposed by researchers for FER are described, including CNNs, CNN-LSTM combinations, and 3D CNNs. The paper compares the performance of these methods on different databases and finds recognition rates generally over 90%. Pre-

processing steps like data augmentation are discussed as important for training deep models

[2] Facial Emotion Recognition Using Conventional Machine Learning and Deep Learning Methods: Current Achievements, Analysis and Remaining Challenges ,Amjad Rehman Khan

The paper provides a review of facial emotion recognition techniques using both traditional machine learning and deep learning methods. It discusses conventional ML approaches that use facial detection, feature extraction and classification. Deep learning approaches like CNNs that enable end-to-end learning are also covered. Popular facial emotion datasets and evaluation metrics used to compare different approaches are described. While deep learning methods provide high accuracy, they require more resources. Remaining challenges in the field are also discussed.

[3] Emotional valence and arousal affect reading in an interactive way: Neuroimaging evidence for an approach-withdrawal framework Francesca M.M. Citron, Marcus A. Gray , Hugo D. Critchley, Brendan S. Weekes,Evelyn C. Ferstl

This study investigated how emotional valence (positive versus negative) and arousal (intensity) dimensions interact during implicit processing of emotional words. Sixteen participants underwent fMRI scanning while performing a lexical decision task with words varying on valence (positive, negative) and arousal (high, low) dimensions. The results showed greater activation in the right insular cortex for words eliciting conflicting approach-withdrawal tendencies (positive high-arousal and negative low-arousal words) compared to words eliciting congruent tendencies. Increased activation was also found in the left extrastriate cortex for emotional versus neutral words, suggesting enhanced perceptual processing of emotionally salient stimuli. These findings provide support for an interactive two-dimensional model of emotion processing and suggest the insula plays a role in integrating approach-withdrawal tendencies elicited by emotional stimuli.

[4] Emotion assessing using valence-arousal evaluation based on peripheral physiological signals and support vector machine ,Mimoun Ben Henia Wiem and Zied Lachiri

This paper presents research on recognizing human emotions using physiological signals collected from the MAHNOB-HCI database. Peripheral signals measured include ECG, GSR, skin temperature, and respiration volume. Emotional states are classified into two areas of the valence-arousal model. preprocessing, feature extraction and a support vector machine (SVM) classifier are used. Classifying individual signals and fused features are

evaluated. The ECG and respiration volume were found to be most effective for emotion recognition. Classification accuracies were calculated and compared to other works, showing improvements over previous studies using this database.

[5] Music Recommendation Based on Face Emotion Recognition ,Madhuri Athavle, Deepali Mudale, Upasana Shrivastav, Megha Gupta

We propose a new approach for playing music automatically using facial emotion. Most of the existing approaches involve playing music manually, using wearable computing devices, or classifying based on audio features. Instead, we propose to change the manual sorting and playing. We have used a Convolutional Neural Network for emotion detection. For music recommendations, Pygame & Tkinter are used. Our proposed system tends to reduce the computational time involved in obtaining the results and the overall cost of the designed system, thereby increasing the system's overall accuracy. Testing of the system is done on the FER2013 dataset. Facial expressions are captured using an inbuilt camera. Feature extraction is performed on input face images to detect emotions such as happy, angry, sad, surprise, and neutral. Automatically music playlist is generated by identifying the current emotion of the user. It yields better performance in terms of computational time, as compared to the algorithm in the existing literature

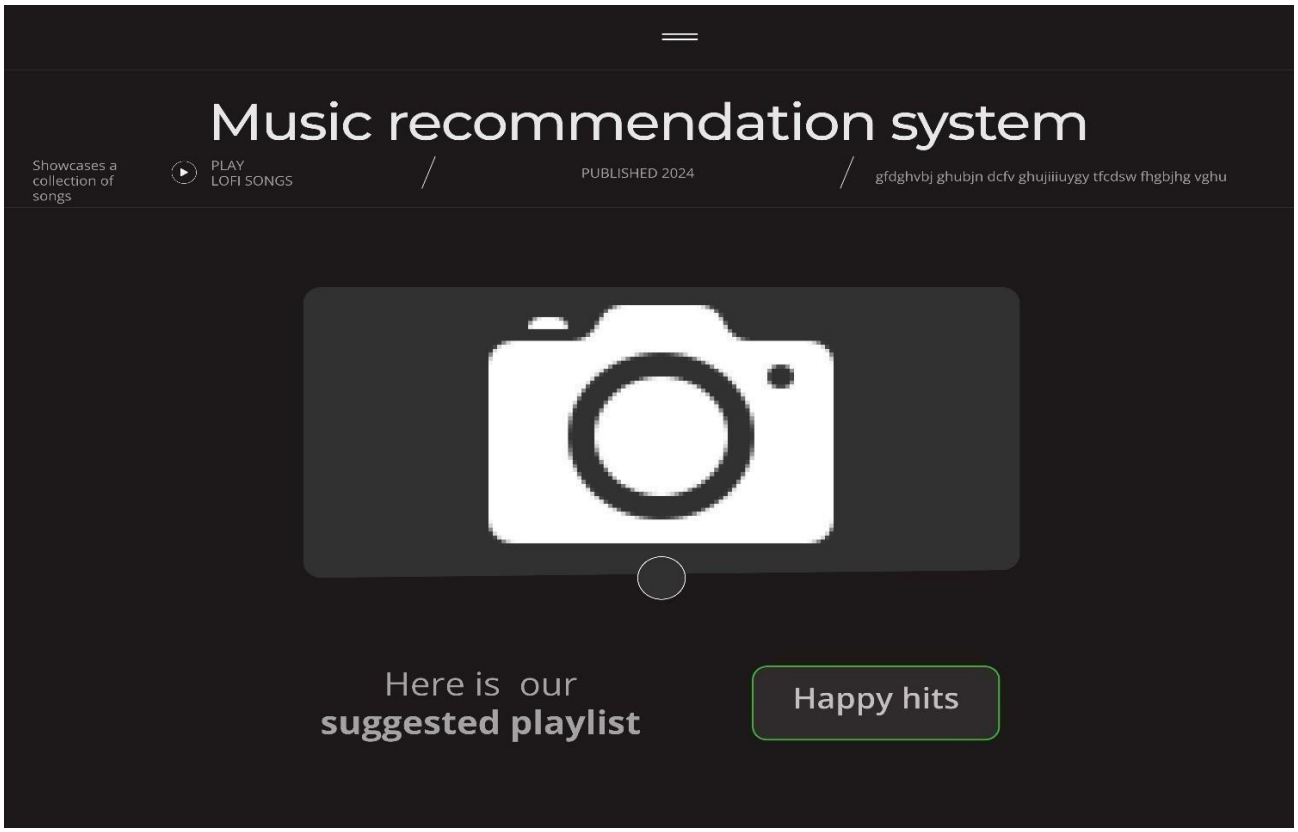
III. Objectives:

- **Develop a sentiment analysis model to accurately detect moods from user's facial expressions.**
  -- This objective focuses on building the image processing component that understands the user emotional state.

- **Create a comprehensive mapping between moods and musical genres.**
  -- This objective focuses on defining the link between emotions and music styles, which is the heart of the recommendation system.

- **Use the Spotify API to find a playlist based on the identified genre and user's language preference.**
  -- This objective focuses on integrating with an external source of music data to provide diverse and customized recommendations.

- **Design a user-friendly interface for interacting with the recommendation system.**
  -- This objective focuses on emphasizing the usability and accessibility of your system.



IV. Methodology:

This section details the steps involved in developing the mood-based music recommendation system, focusing on facial expression recognition, mood-to-valence-arousal mapping, and music recommendation using the Spotify API.

1. Facial Expression Recognition

Dataset:

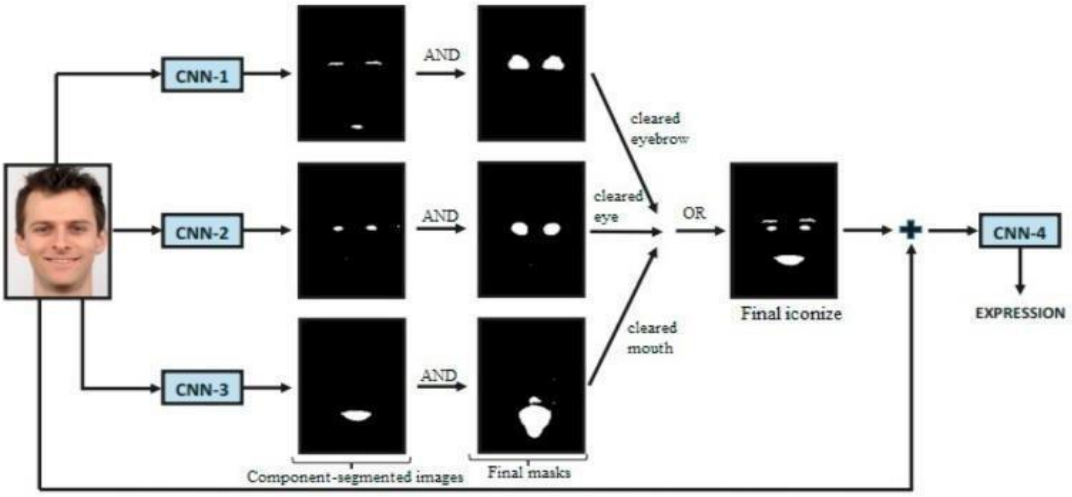
The project utilizes the FER-2013 dataset, a publicly available dataset consisting of 35,887 facial images labeled with seven emotions: neutral, anger, contempt, disgust, fear, happiness, and sadness.

Model:

A Convolutional Neural Network (CNN) model specifically designed for facial expression recognition is created. The machine learning model for the emotion-based music player project uses supervised learning. It involves training the model with pre-labeled data. The labeled data includes the mental states of people.

Evaluation:

The model's performance will be evaluated using accuracy and F1-score metrics on a held-out test set from the FER-2013 dataset.

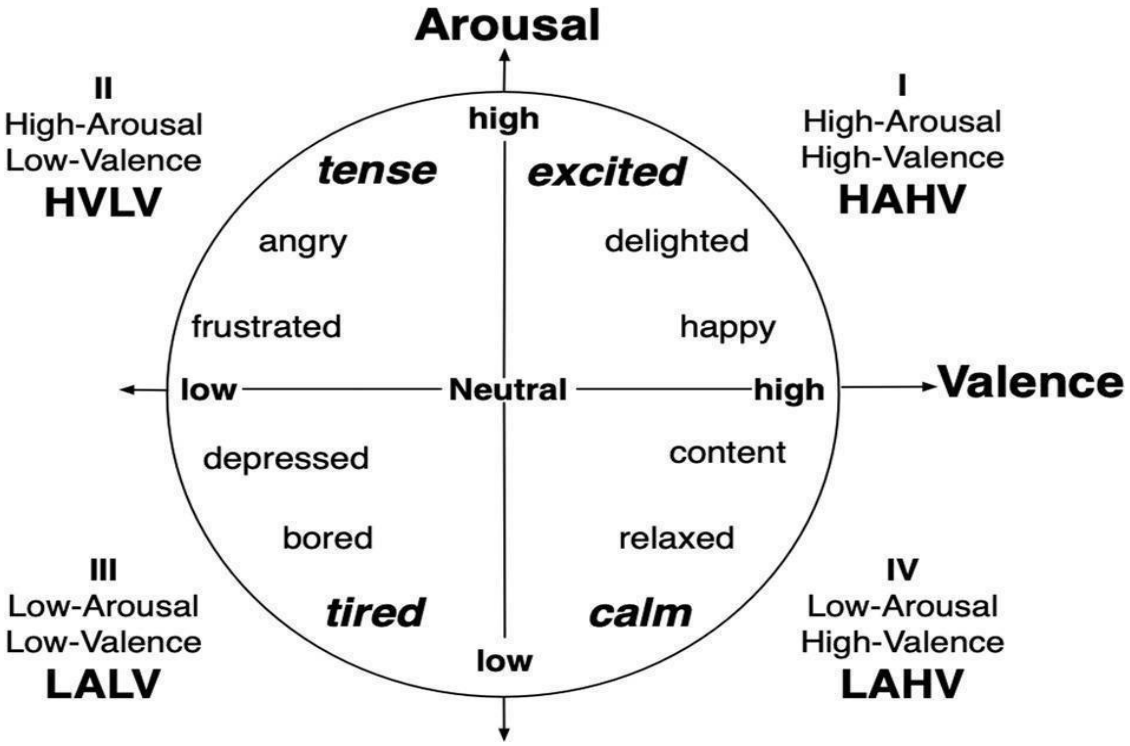


2. Mood-to-Valence-Arousal Mapping:

The project utilizes a rule-based approach to translate detected moods from facial expressions into corresponding valence (positive/negative) and arousal (energy level) values on a scale from 0 to 1.

For example:

Happy: High Valence (0.7 - 1), Medium-High Arousal (0.6- 1)
Sad: Low Valence (0 - 0.4), Low Arousal (0 - 0.4)
Angry: Low Valence (0 - 0.4), High Arousal (0.7 - 1)
Calm: High Valence (0.7 - 1), Low Arousal (0 - 0.4)
Surprise: Can be high or low valence depending on its nature. Mid-High Arousal (0.5 - 0.8)
Fear: Low Valence (0 - 0.4), High Arousal (0.7 - 1)
Disgust: Low Valence (0 - 0.4), Medium Arousal (0.5 - 0.7)
Neutral: Mid Valence (0.4 - 0.6), Low Arousal (0 - 0.4)



3. Music Recommendation

Data Source: The Spotify API will be used to access music data and retrieve playlists.

API Integration: The API will be queried directly with the calculated valence and arousal values derived from the detected mood. Spotify's search functionality allows for filtering using these parameters.

Filtering: The system will prioritize playlists that closely align with the valence and arousal range associated with the user's mood.

Recommendation Logic: The Spotify API will return a list of playlists matching the valence-arousal criteria. The system will randomly select one playlist from the returned list to present as a recommendation to the user

4. User Interface (UI)

Functionality:

The UI will allow users to upload an image of their face.

Once uploaded, the system will process the image using the facial expression recognition model to determine the user's mood.

Based on the detected mood, the system will use the mood-to-valence-arousal mapping and Spotify API to recommend a music playlist.

The recommended playlist will be displayed on the UI for the user to listen to.

Technology:

A simple web application will be developed using Javascript and a web framework like React and for styling Tailwind CSS will be used.

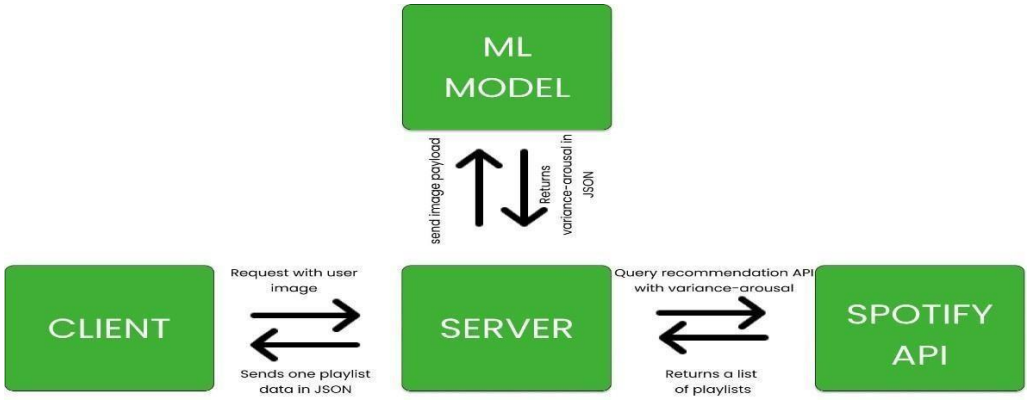


fig : architecture diagram of music recommendation system

V. Modules identified:

Frontend

React (Version: ^18.2.0)
Tailwind CSS (Version: ^3.2.4)
Typescript (Version: ^4.9.4)

Backend

Flask (Version: ^2.3.0)
TensorFlow (Version: ^2.12.0)
Keras (Version: ^2.13.1)
OpenCV (Version: ^4.6.0)

VI. Work progress / plan & Implementation

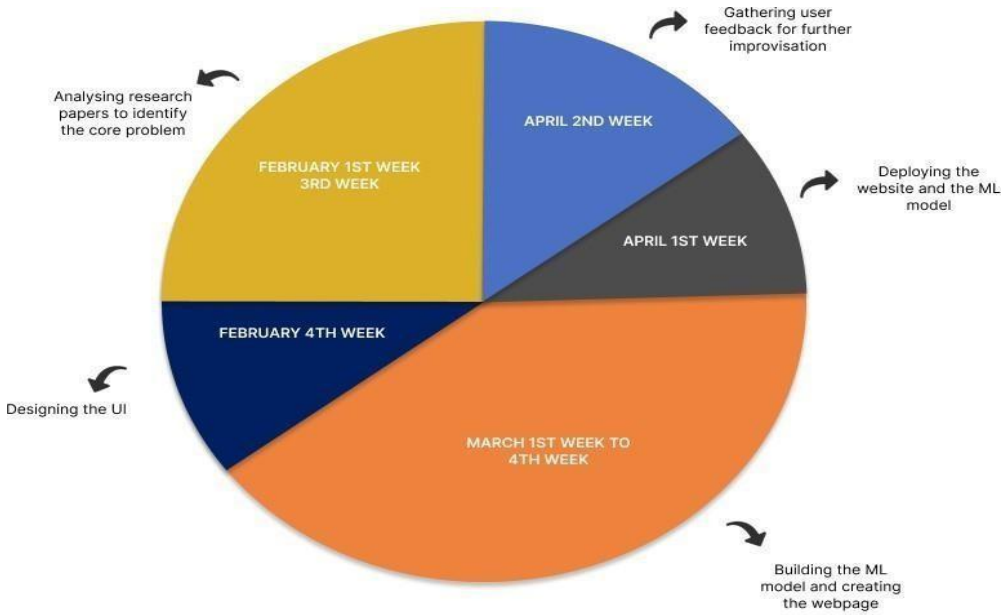


fig : work progress

VII. Sample Code

VIII. Conclusions

This project explores a mood-based music recommendation system utilizing facial expression analysis. By employing CNN, the system accurately detects moods from user images. A rule-based mapping was established to translate these moods into valence and arousal values, facilitating targeted music recommendations via the Spotify API. This approach demonstrates the potential for personalizing music discovery based on emotional state. Future directions include expanding the mood detection capabilities, refining the valence-arousal mapping through user studies, and exploring additional recommendation techniques to offer even greater diversity and personalization within the system.

IX. References:

Example for reference is given below. Kindly follow the same format for writing reference

[1] Wafa Mellouk,Wahida Handouzi, "Facial emotion recoginition using deep learning: review and insights,"August-2020,Leuven,Belgium.

[2] Mimoun Ben Henia Wiem1 , Zied Lachiri, "Emotion assessing using valence-arousal evaluation based on peripheral physiological signals and support vector machine",December-2016, Tunisia, Hammamet

[3] Francesca M.M. Citron , Marcus A. Gray , Hugo D. Critchley , Brendan S. Weekes , Evelyn C. Ferstl, "Emotional valence and arousal affect reading in an interactive way: Neuroimaging evidence for an approach-withdrawal framework," June-2013, Queensland, Australia

[4] Amjad Rehman Khan, "Facial Emotion Recognition Using Conventional Machine Learning and Deep Learning Methods: Current Achievements, Analysis and Remaining Challenges," CCIS Prince Sultan University, Riyadh 11586, Saudi Arabia.

[5] Jaladi Sam Joel, B. Ernest Thompson ," Emotion based Music Recommendation System using Deep Learning Model," Proceedings of the International Conference on Inventive Computation Technologies (ICICT 2023) IEEE Xplore Part Number: CFP23F70-ART; ISBN: 979-8-3503-9849-6.

[6] Madhuri Athavle , Deepali Mudale , Upasana Shrivastav , Megha Gupta "Music Recommendation Based on Face Emotion Recognition," Vol. 02, Iss. 02, S. No. 018, pp. 1-11 ISSN (Online): 2582-7006