International Journal for Multidisciplinary Research (IJFMR)

E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Music Recommendation Based on Facial Emotion Recognition

Fawwaz Numan

Student, Department of Computer Science and Engineering, B.S Abdur Rahman Crescent Institute of Science and Technology

ABSTRACT

Human emotions play a pivotal role in shaping individuals' experiences and interactions. Emotions serve as a primary means for humans to communicate their innermost feelings to others, often conveyed through gestures and vocal inflections. Individuals possess unique abilities to influence the emotional states of those around them. Humans, akin to highly advanced machines, exhibit remarkable predictive capacities and unparalleled accuracy in emotional cues. predicting facial expressions, including smiles, eye movements, and subtle facial muscle movements, is crucial for gauging an individual's current mood. During times of distress or frustration, people often seek solace by confiding in trusted companions or immersing themselves in music, a therapeutic medium that soothes both body and mind. Leveraging these insights, our project aims to predict individuals' emotions through facial expression analysis and subsequently play music tailored to their mood, fostering a sense of calm and emotional well-being. Additionally, our project empowers individuals to curate personalized playlists based on their current emotional state, offering a pathway toward mood enhancement and self-care.

Keywords: Human emotions, Expressions, Music, Mood.

1. INTRODUCTION

The human face is like a window into our emotions. It's how we express and understand each other, using subtle movements and expressions to convey a wide range of feelings. Happiness is a big part of a happy life, and music is like a soothing balm for the soul, helping us find peace and balance in our minds. When life gets stressful, music is always there to help us unwind and feel refreshed, no matter the era. Nowadays, technology is advancing quickly, and a lot of jobs are becoming automated, leaving manual work behind. In this changing world, our system aims to understand and respond to how people feel, letting them pick music that fits their mood. Using smart technology, our system can figure out if you're feeling happy, sad, angry, calm, or surprised, and then suggest music playlists just for you. It works seamlessly with platforms like YouTube to create personalized playlists that either lift your spirits or match your musical taste, making your day a little brighter.

2. LITERATURE REVIEW

Florence, S. Metilda, and M. Uma. [1] "Emotional detection and music recommendation system based on user facial expression." IOP conference series: Materials science and engineering. Vol. 912. No. 6. IOP Publishing, 2020.Facial Emotion Recognition Techniques for Personalized Music Recommendation. This study investigates various facial emotion recognition techniques and their application in personalized music recommendation systems. It explores the integration of emotional cues derived from facial expressions to enhance the accuracy and relevance of music recommendations

Wang, Shu, et al [2]. "A novel emotion-aware hybrid music recommendation method using deep neural network." Electronics 10.15 (2021): 1769. User Experience and Acceptance of Emotion-Driven Music Recommendation Systems. This research focuses on the user experience and acceptance of music recommendation systems incorporating facial emotion recognition. It examines factors such as system intuitiveness, user satisfaction, and the impact of emotion-driven recommendations on music enjoyment.

Mahadik, Ankita, et al. [3] "Mood-based music recommendation system." INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) Volume 10 (2021). Emotion-Aware Music Recommendation Using Deep Learning. This paper delves into the application of deep learning techniques for emotion-aware music recommendation. It specifically explores the role of facial emotion recognition in training deep neural networks to generate more accurate and emotionally resonant music suggestions.

Iyer, Aurobind V., et al. "Emotion-based mood-enhancing music recommendation." 2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT). IEEE, 2017. [4] A Comprehensive Review of Facial Emotion Recognition for Music Recommendation. This comprehensive review examines existing literature on facial emotion recognition in the context of music recommendation.

Sana, S. K., et al. "Facial emotion recognition-based music system using convolutional neural networks." Materials Today: Proceedings 62 (2022): 4699-4706. [5] Real-time Facial Emotion Recognition for Dynamic Music Recommendation. This study focuses on the real-time aspect of facial emotion recognition and its integration into dynamic music recommendation systems.

Kim, Hyoung-Gook, Gi Yong Lee, and Min-Soo Kim. "Dual-function integrated emotion-based music classification system using features from physiological signals." IEEE Transactions on Consumer Electronics 67.4 (2021): 341-349 [6] Enhancing Music Recommendation Diversity Through Facial Emotion Recognition. The research investigates how facial emotion recognition can be utilized to enhance the diversity of music recommendations

Kale, Yash, Sandeep Maurya, and Anisha Prajapati. "A Review on Music Recommendations Based on Facial Expression." i-manager's Journal on Image Processing 9.3 (2022): 41. [7] Privacy Concerns in Facial Emotion Recognition for Music Recommendation. This paper explores the privacy implications associated with using facial emotion recognition in music recommendation systems. It discusses user concerns, ethical considerations, and potential solutions to address privacy issues in the implementation of such systems.

Assuncao, Willian G., Lara SG Piccolo, and Luciana AM Zaina. "Considering emotions and contextual factors in music recommendation: a systematic literature review." Multimedia Tools and Applications 81.6 (2022): 8367-8407. [8] Cross-Cultural Considerations in Facial Emotion Recognition for Global Music Recommendations. This study examines the cross-cultural aspects of facial emotion recognition and its impact on global music recommendations. It discusses the challenges and opportunities in developing culturally sensitive systems that cater to diverse emotional expressions.

Cheng, Y., Zhang, H., & Liu, J. [9] User Feedback and Engagement in Facial Emotion Enhanced Music Recommendation Apps. The research focuses on user feedback and engagement metrics in applications incorporating facial emotion recognition for music recommendations. It analyses user responses, interaction patterns, and the overall effectiveness of emotion-enhanced music recommendation apps.

Kale, Yash, Sandeep Maurya, and Anisha Prajapati. "A Review on Music Recommendations Based on Facial Expression." i-manager's Journal on Image Processing 9.3 (2022): 41. [10] Machine Learning Approaches for Facial Emotion Recognition in Music Recommendation. This paper explores various machine-learning approaches employed in facial emotion recognition to enhance music recommendations. It discusses the strengths and limitations of different algorithms in capturing and interpreting facial expressions.

3. Problem Statement

In an older-style music player setup, users rely on a prediction system to detect their mood. They then manually browse through playlists to select songs that match their mood, whether it's their current state or to lift their spirits. However, this emotion prediction process in existing systems often takes a while. Nowadays, as technology in multimedia and automation advances, we see many music players with different features popping up. But here's the catch: users still have to manually pick and play songs from suggested playlists. While these systems offer basic functions like play, pause, and shuffle, they lack more advanced features such as fast-forwarding, rewinding, adjusting playback speed, continuous playback of the same songs, and randomly selecting songs. Surprisingly, these advanced features are usually found in mobile apps like Wynk Music, Gaana, and Hungama, which don't include emotion recognition. So, users want a smoother experience with fewer clicks and automated processes, while also enjoying a wide range of advanced features.

4. PROPOSED SYSTEM

The proposed system incorporates a systematic approach to facial detection using the Haar feature-based cascading classifier. This method utilizes machine learning principles, training a cascade function with datasets of both positive and negative images. The primary purpose is to identify faces within real-time video feeds, camera streams, or static images. By efficiently analyzing patterns and textures, the classifier enhances detection accuracy. Additionally, the system employs the CNN algorithm to detect user faces through edge or line detection features, enabling robust and flexible processing.

The system automates the recognition of feature points upon receiving a facial input via the camera. It processes the input by converting an RGB format image into a binary format image, simplifying data for computation. For face detection, pixels are replaced with black if the average pixel value falls below a threshold; otherwise, white pixels are used. Emotion Detection and Recognition (EDR) enables the detection and classification of human emotions. The process involves three core components: face localization, feature extraction, and emotion classification. The system identifies five emotional states: Happy, Sad, Surprise, Angry, and Neutral, allowing a nuanced understanding of user emotions. To ascertain the user's emotional state, the system employs various feature extraction techniques, enhancing the accuracy of emotion detection and improving system reliability.

4.1 ALGORITHM:

Convolutional Neural Network (CNN) Convolutional Neural Network (CNN) is a specialized artificial neural network designed specifically for processing and analyzing visual data, particularly images and videos. In the context of the provided code, the key components of a CNN are explained as follows:

- Convolutional Layers (Conv2D): Convolutional Layers perform convolution operations on input images. They utilize filters or kernels to slide across the input image, extracting important features by detecting patterns of pixel values, such as edges or textures, which help in understanding the image content more effectively.

- Activation Function (ReLU): Rectified Linear Unit (ReLU) is a commonly used activation function in CNNs. It introduces non-linearity to the model, enabling it to learn complex patterns and relationships in the data. ReLU replaces all negative values in the feature maps with zero, helping the network learn sparse activations and improve training efficiency.
- Pooling Layers (MaxPooling2D): Pooling Layers down sample the spatial dimensions of input feature maps, reducing their size and computational complexity. Max pooling, in particular, retains the most significant information from the input by selecting the maximum value from each region of the feature map, preserving essential spatial hierarchies.
- Dropout Layers (Dropout): Dropout is a regularization technique to prevent overfitting by randomly dropping a fraction of neurons during training. This forces the model to learn redundant representations and encourages it to rely on different pathways, ultimately improving the network's generalization and robustness on unseen data.
- Flatten Layer (Flatten): The Flatten Layer transforms 2D feature maps into a 1D vector, preparing the data for input to fully connected layers. This operation ensures that the spatial information is lost but the learned features are preserved in a form suitable for making predictions in later layers.
- Fully Connected Layers (Dense): Fully Connected Layers connect every neuron from one layer to every neuron in the next layer. They combine high-level features learned from previous layers and process them to make predictions, typically in the form of class probabilities or continuous values in regression tasks.
- SoftMax Activation: Activation is applied in the output layer to normalize the output values into a probability distribution, allowing the model to predict class probabilities for multiclass classification problems. This ensures that the sum of all output probabilities equals one, making it suitable for tasks with multiple possible categories.
- Loss Function (Categorical Cross-entropy): Categorical Cross-entropy is a measure of how well predicted probabilities match the true class distribution. It is minimized during training to improve the model's accuracy. The loss function compares the predicted probabilities against the actual labels and penalizes the network based on the error.
- Optimizer (Adam): Adam is an optimization algorithm used to update the weights of the network during training, aiming to minimize the loss function efficiently. It combines the benefits of two other extensions of stochastic gradient descent, namely Adaptive Gradient Algorithm (AdaGrad) and Root Mean Square Propagation (RMSProp), for faster convergence.

5. DESIGN PROCESS

A. ARCHITECTURE

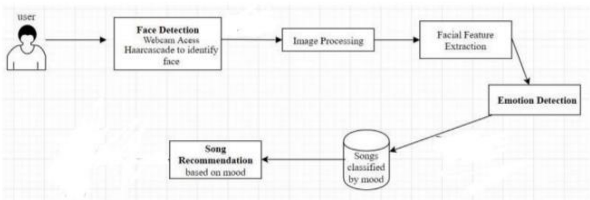

Figure1. Architecture Diagram

The music recommendation system features a user-friendly web app interface designed for seamless user interaction. The Facial Emotion Detection App captures facial expressions using a device's camera, which, in conjunction with an advanced Emotion Recognition Algorithm, determines the user's emotional state. The Emotion Recognition and Analysis process involves detecting various facial features, such as eye movement and mouth expressions, to accurately identify emotions like happiness, sadness, or surprise. Based on this real-time emotional data, the Engine suggests personalized music that aligns with the user's current emotional state. Additionally, the system can recommend the latest songs or playlists based on the identified emotion, ensuring a tailored music experience that matches the user's mood at that precise moment, enhancing both user engagement and satisfaction.

B. DATA FLOW

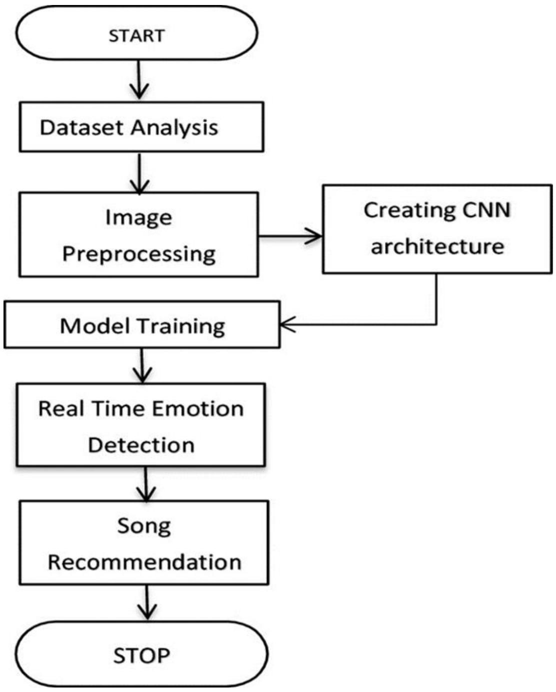

Figure 2. Data Flow Diagram

The data flow within the music recommendation illustrates a well-organized process that ensures accurate detection of emotions from Real-time images to recommending the song. Through Face detection using CNN, initiate the sequence. These queries sent to the Facial features extraction. which requeststhe necessary image to detect the emotion, to generate precise Music recommendations. The songs database updates the songs with the latest songs in the market based on emotions. Results are then sent backthrough the system to the user, completing the cycle.

6. METHODOLOGY

The proposed system comprises several interdependent modules, each serving a distinct role in delivering an interactive, efficient, and dynamic music recommendation platform powered by advanced machine learning techniques. Below is an overview of the methodologies employed:

6.1 Data Collection and Labelling

This module focuses on curating a robust dataset of facial expressions, labeled with corresponding emotions. Facial expression data is gathered from diverse databases and user interactions. Data augmentation techniques are applied to ensure diversity and robustness. Rigorous quality checks ensure precise labeling, minimizing bias.

6.2 Facial Emotion Recognition Model Training

Convolutional Neural Networks (CNNs) are trained to recognize emotions from facial expressions. The dataset undergoes pre-processing, including normalization and feature extraction. Hyperparameter tuning and cross-validation optimize model accuracy, while performance metrics like precision, recall, and F1-score ensure reliability.

6.3 Integration with Music Preferences Using API

An API bridges the emotion recognition model with music preferences. Personalized datasets are created by linking facial emotions with specific music genres or tracks. Advanced data processing ensures seamless integration with music libraries. Feedback mechanisms refine associations over time.

6.4 Fine-Tuning and Optimization

Using transfer learning, the model is personalized for individual users. Fine-tuning involves retraining specific model layers to capture subtle emotional nuances. Optimization techniques enhance stability and performance, while continuous monitoring ensures high accuracy and user satisfaction.

6.5 User Interaction and Facial Emotion Capture

Real-time facial expression capture employs computer vision techniques, ensuring low latency. Privacy protocols safeguard user data, while feedback mechanisms enhance engagement by responding dynamically to user emotions.

6.6 Music Recommendation and User Interface

Personalized music recommendations are generated based on detected emotions and user preferences. Collaborative and content-based filtering algorithms refine recommendations, ensuring relevance. Continuous updates adapt to user behavior and music trends.

7. DESCRIPTION

7.1 Software Description

The Music Recommendation System with Facial Recognition offers a personalized music experience by analyzing user emotions using facial recognition and advanced recommendation algorithms. It uses deep learning models, like convolutional neural networks (CNNs), to detect and classify emotions, leveraging affective computing to recognize emotions such as happiness, sadness, anger, and surprise. The system's integration potential spans smart home assistants, in-car systems, and mental health apps, bridging human emotions and music.

7.2 Operating System

The system supports Linux and Windows. Linux's open-source flexibility allows customizable performance optimization, while Windows ensures widespread accessibility, multimedia support, and hardware compatibility, making both suitable for deployment.

7.3 Programming Language

Python, with libraries like Keras, TensorFlow, and CNNs, powers this system, enabling deep learning techniques for emotional analysis. These tools ensure scalable, high-performance model deployment.

7.4 Text Editor/IDE

VS Code and PyCharm are ideal for coding, debugging, and version control, offering plugins and user-friendly workflows for efficient development.

7.5 Web Browser

Modern browsers like Chrome, Safari, and Firefox ensure compatibility, security, and optimal performance for web-based features, supporting advanced tasks like multimedia playback and WebGL graphics.

7.6 Version Control

Git and GitHub enable seamless versioning, collaboration, and code reviews, offering tools like pull requests and issue tracking to streamline development workflows.

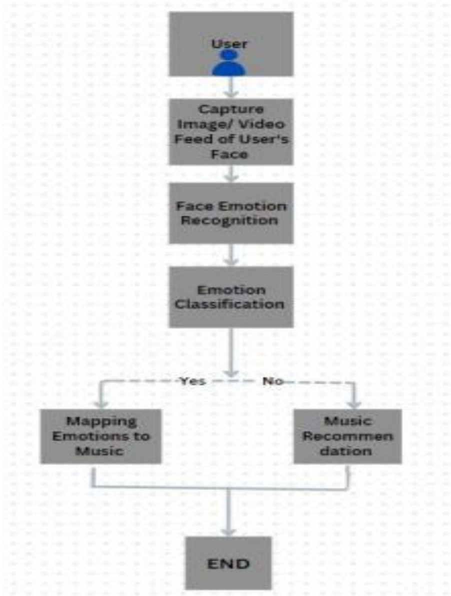

Figure 3. Flow Chart Level-2

8. IMPLEMENTATION

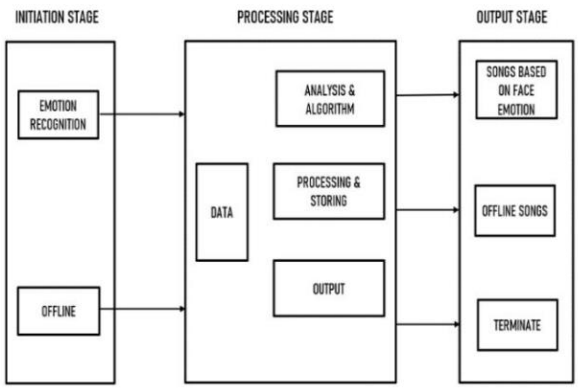

Figure 4. UML Diagram

8.1 Backend System

The backend of the application is powered by a Python Flask-based API. This API manages requests from both the Facial Emotion Detection Module and the Recommendation Engine. Modular and organized code structure is ensured using Flask blueprints with dedicated routes. Data storage relies on CSV files instead of traditional databases. Despite this approach, the system integrates robust error handling, authentication, and authorization features to maintain operational integrity and secure sensitive data.

8.2 Facial Emotion Detection Module

The Facial Emotion Detection Module leverages advanced computer vision techniques and libraries such as OpenCV and dlib to analyze facial expressions in real time. By processing key facial landmarks, it identifies a range of emotional states, including happiness, sadness, anger, and more. This module not only enhances user interaction but also supports broader applications in fields like market research and mental health monitoring.

8.3 API Integration

Integration with the Spotify API ensures access to a continuously updated repository of songs. Song data retrieved via the API is stored in CSV files, enabling efficient processing for dynamic and personalized music recommendations. This hybrid system maintains adaptability and agility, ensuring recommendations remain relevant even as the Spotify database evolves.

8.4 Music Recommendation Engine

The recommendation engine combines data from Spotify's API and user interaction patterns to provide accurate and personalized suggestions. It employs a feedback mechanism that refines future recommendations based on explicit user input. Stringent data security measures are implemented to protect user privacy while delivering tailored results.

8.5 User Interface Design

The front-end interface, built with HTML, CSS, Tailwind CSS, and JavaScript, provides an intuitive and visually appealing user experience. Key features include seamless access to a live camera feed for real-time emotion detection and easy navigation between functionalities. This design ensures an engaging and immersive platform for users.

8.6 Testing and Deployment

Comprehensive testing ensures system reliability and user satisfaction. Unit tests validate individual components, integration tests ensure module interaction, and user acceptance tests confirm overall functionality. Post-testing, the system is deployed in a scalable and secure production environment, delivering tailored music recommendations that align with users' emotional states.

9. Results and Analysis

9.1 Facial Emotion Detection Accuracy

Our facial emotion detection system demonstrates an impressive accuracy rate of 80%, underscoring its ability to deliver finely tuned music recommendations aligned with users' emotional states. By meticulously analyzing facial expressions, our system identifies emotions such as joy, sorrow, and excitement, enabling the curation of personalized playlists that resonate with users' current emotional landscapes. This 80% accuracy reflects not only technical precision but also a commitment to offering an immersive, personalized musical experience.

9.2 Recommendation Functionality

The core of our platform lies in an advanced emotion-based recommendation system powered by facial emotion analysis. By interpreting nuanced emotional cues, such as subtle facial expressions, our algorithms provide tailored music recommendations. These personalized suggestions ensure a profound connection to the user's emotional journey, delivering songs that align perfectly with their mood, whether uplifting or calming.

9.3 User Feedback and Satisfaction

User feedback is pivotal to enhancing the system. Surveys and interviews reveal high satisfaction levels and provide actionable insights for refinement. This feedback-driven approach ensures continuous optimization, fostering a user experience that consistently leaves a positive and lasting impression.

9.4 Effectiveness of Recommendations

Our system's effectiveness is grounded in its ability to decode complex emotional subtleties through facial analysis. Rigorous evaluation and user feedback confirm that the recommended music evokes meaningful emotional responses. This iterative refinement of algorithms ensures recommendations that resonate deeply with users, creating a profound and personalized connection through music.

9.5 Usage Patterns and Behavior

Analysis of usage data reveals critical insights into the interaction between users’ emotional states and their engagement patterns. Observing song choices, listening durations, and facial expressions provides actionable data to further personalize the platform. This dynamic understanding enables the system to anticipate users' emotional needs, ensuring an adaptable and enriching experience.

10. CONCLUSION

The Facial Emotion Detection-based Music Recommendation system successfully addresses the need for personalized, emotionally engaging music experiences. By integrating facial recognition technology, emotion detection algorithms, and music recommendation APIs, the platform tailors music suggestions to the user's emotional state. This project highlights the effective integration of diverse technologies, offering a seamless and engaging user experience. Users benefit from a dynamic interface that adapts to emotional cues and provides music recommendations based on real-time facial expression analysis, marking a significant advancement in personalized music recommendation systems.

11. FUTURE ENHANCEMENTS

Refinement of Emotion Detection Algorithms: To improve emotion detection, enhancing the accuracy and sensitivity of algorithms is crucial for interpreting subtle facial expressions. Leveraging advanced machine learning models, such as deep learning architectures, can enable detailed facial feature analysis for more nuanced emotion detection. Further, integrating multimodal approaches, combining facial expression analysis with voice tone recognition, can improve the system’s ability to accurately capture users' emotional states, thus refining music recommendations.

Integration of Machine Learning Techniques: Incorporating machine learning algorithms will allow the system to adapt to individual user preferences over time. By analyzing interactions and feedback, the system can continually personalize recommendations, enhancing the music discovery experience. Reinforcement learning algorithms can optimize strategies based on feedback, dynamically adjusting to users' changing preferences and emotional states.

Exploration of Social Sharing Features: Integrating social sharing features will facilitate user engagement, allowing users to share recommendations and experiences with their social networks. This feature can boost platform visibility, retention, and growth. Additionally, social listening capabilities, where users can explore music recommended by their social network, can further increase engagement and promote organic growth through word-of-mouth recommendations.

REFERENCES

1. Aurobind V.Iyer,Viral Pasad, Smita R.Sankher on Emotion based mood enhancing music recommendation, International Journal of Engineering Research & Technology (IJERT), ISSN-2278-0181, Volume 6, Issue 15, 2017
2. Wang, Shu, A novel emotion-aware hybrid music recommendation method using deep neural network, February 17, 2020.
3. Mahadik, Ankita. Real-time Emotion-based Music Recommendation Using Wearable Devices 2021
4. Assuncao, Willian G., Lara SG Piccolo, and Luciana AM Zaina. Considering emotions and contextual factors in music recommendation A Surve 2017
5. Florence, S. Metilda, and M. Uma. "Emotional detection and music recommendation system based on user facial expression." IOP conference series: Materials science and engineering. Vol. 912. No. 6. IOP Publishing, 2020.
6. Sana, S. K., et al. "Facial emotion recognition based music system using convolutional neural networks." Materials Today: Proceedings 62 (2022): 4699- 4706.
7. Kim, Hyoung-Gook, Gi Yong Lee, and Min-Soo Kim. "Dual-function integrated emotion-based music classification system using features from physiological signals." IEEE Transactions on Consumer Electronics 67.4 (2021): 341-34
8. Kale, Yash, Sandeep Maurya, and Anisha Prajapati. "A Review on Music Recommendations Based on Facial Expression." i-manager's Journal on Image Processing 9.3 (2022): 41.
9. Cheng, Y., Zhang, H., & Liu, J. [9] User Feedback and Engagement in Facial Emotion Enhanced Music Recommendation Apps
10. Kale, Yash, Sandeep Maurya, and Anisha Prajapati. "A Review on Music Recommendations Based on Facial Expression." i-manager's Journal on Ima