**"** **Emo Melodies - Music Aligned with Facial Emotions"**

Shreela Pareeka,Ritika Raib,Rishika Tyagic,Priyanshu Singhd, Vishakha Chauhane

(a,b,c,d,e)\* Department of Computer Science,

KIET Groups of Institutions, Delhi-NCR, Ghaziabad, 201206

Abstract: The integration of natural language processing technology with facial emotion-based music recommendation systems enables the comprehension of the user's preferences and mood based on text or voice inputs. This technology has the potential to enhance the user's listening experience and improve personalised recommendations. This study investigates the composition and implementation of emo-melodies, which include In addition to analysing the fundamental algorithms and technologies incorporated into this system, we deliberate on the obstacles and limitations it encounters, along with prospective domains for further investigation. Furthermore, an initial investigation was undertaken to evaluate the immediate impact of music on user mood, which granted us the ability to curate playlists that positively affect users. On the contrary, we offer a customised compilation consisting of diverse musical genres to augment positive sentiments that are detected. The fusion of emotional recognition through facial expressions and its application in music recommendation systems represents a burgeoning field of interdisciplinary research. This literature survey delves into the core technologies and methodologies underpinning the project "Emo Melodies - Music Aligned with Facial Emotions," highlighting significant contributions, current trends, and potential future directions in the domains of Facial Emotion Recognition (FER), Music Emotion Recognition (MER), and their integration within interactive systems

**Keywords:** Emotion Detection, Expressions, Music recommendation, Feature extraction, Convolutional Neural Network

1. Introduction

We examine the challenges and limitations inherent in this system, alongside potential avenues for additional research. We did a concise investigation on how music might impact short-term mood to improve our comprehension and prepare us to provide customers with a tailored collection of tunes that can successfully enhance their moods. Alternatively, when a positive emotional state is detected, the system offers a tailored playlist featuring diverse genres aimed at amplifying the mood's positivity. This underscores the importance of our music selection algorithm, which relies on facial expressions to gauge emotional states accurately. Leveraging real-time facial analysis, this technology selects the most suitable music genre aligned with the user's prevailing emotions. This endeavours explores the creation and integration of a music recommendation system that harnesses facial expressions for mood assessment. The article examines the technology, prospective advantages, and critical hurdles to improve its efficacy.

The aim of this project is to extensively examine the system's capabilities and its potential applications in various sectors, such as the music industry. With the increasing prevalence of music streaming platforms and the widespread integration of advanced facial recognition technologies, there exists significant potential for improving personalized music recommendations. Its transforming effect on our listening behaviours guarantees increased involvement and emotional significance. By integrating facial expression detection into music recommendation algorithms, we can offer individuals a tailored and enriched musical journey. Furthermore, its use goes beyond music and also applies to several industries such as entertainment, healthcare, and education. We will examine current research on face expression detection technology and its use into music recommendation systems to accomplish the study's goals. We'll explore various machine learning and deep learning techniques commonly used in developing models for identifying facial emotions. Through an empirical study involving participant data collection, we'll evaluate the effectiveness of the proposed system. The findings from this research could significantly impact the music industry, introducing new opportunities for personalized experiences in music streaming services and advertising. This solution holds the potential to increase listener engagement, improve user satisfaction, and ultimately, enhance revenue streams. Additionally, this technology could find application in healthcare for monitoring patients' emotional wellness and delivering customized therapeutic interventions.

1. Literature Review

Humans can naturally interpret facial expressions to understand emotions. If machines could learn this, it would have practical implications. Music, surpassing language in emotional impact, deeply resonates with human emotions. [[1]](#o1)

The paper proposed using Artificial Neural Networks (ANN) to analyze changes in facial curvatures and pixel intensities for emotion classification. It suggested two main methods for facial feature extraction: Appearance-based and geometric- based, focusing on key facial points. Understanding user mindset via facial expressions was emphasized, utilizing feature-point detection or Haar Cascade technology. [[11]](#o1)

Facial emotion recognition in music recommendation systems enhances user experiences. These systems, employing collaborative filtering and content-based algorithms, personalize music selections. Advancements in facial emotion recognition technology enable real-time interpretation of users' emotions. Dr. John Smith's research integrates emotional data into recommendation algorithms, overcoming challenges for future innovation in this interdisciplinary field. [[3]](#o1)

Emotion-aware music recommendation system employing facial emotion recognition to suggest songs aligned with the user's emotions, aiming to enhance user satisfaction in music recommendation. [[9]](#o1)

Authors propose a new method for improving music recommendations by analyzing facial expressions to understand users' emotional reactions to music. They explore integrating this method into current systems for enhanced recommendations. [[13]](#o1)

The paragraph discusses a project centered on personalized music suggestions through real-time facial emotion recognition. It outlines the creation of an app that customizes music playlists according to the user's present emotional condition. [[14]](#o1)

A diagram of a face recognition system

Description automatically generatedA survey paper outlines emotion-based music recommendation systems, covering facial emotion recognition methods. It provides a thorough overview of the current landscape in the field. [[15]](#o1)

1. Objective

The objectives of our study are as follows:

* Developing a customised and engaging music experience according to the user's emotional condition.
* Creating algorithms to precisely identify and understand face expressions and emotions.
* Designing a user-friendly interface that smoothly integrates facial expression recognition with music playback.
* Researching the relationship between facial expressions, emotions, and music choices.
* Improving the emotional influence of music by dynamically modifying the playlist according to real-time face expression analysis.
* Exploring the therapeutic benefits of integrating facial emotion recognition with music therapy for mental health and overall well-being.
* Evaluating and enhancing the system's precision and efficiency based on user input and usability studies.
* Collaborating with musicians, psychologists, and engineers to enhance the integration of face expression recognition and music playback.
* Investigating the possible uses of the technology in areas other than music, such as virtual reality experiences or interactive installations.

1. System Overview

The "Music Aligned with Facial Emotions" project is an innovative system designed to enhance the emotional well-being of users by aligning music playback with their current facial emotions. Utilizing advanced facial recognition technology and emotion detection algorithms, the system analyzes the user's facial expressions in real-time and selects music that either matches or positively influences the user's emotional state. This document provides an overview of the system's architecture, components, and functionality.

**Capture Image**

Figure 1. System Overview

A vital first step in the identification of face emotions is image capture. Thanks to computer vision libraries like OpenCV, this process has been made considerably simpler and more effective.

Facial emotion recognition systems function by examining the facial expressions of individuals to determine their emotional state. First, a camera takes a picture of the subject's face. Next, the image is processed using OpenCV libraries to detect face characteristics like lips, nose, and eyes. OpenCV could detect faces and locate them inside an image. Once recognized, several face features such as form, mouth, nose, and eye positions may be retrieved. The data is used to analyses the individual's emotional state. For instance, smiling lips usually imply pleasure, whereas frowning lips and creased eyebrows may suggest melancholy or discomfort.

**Face Detection using CNN.**

Convolutional Neural Networks (CNNs) are a specialized type of deep learning algorithm known for their effectiveness in image recognition tasks. During the training phase, CNNs are refined with large datasets of images, adjusting the network's parameters to minimize the difference between the predicted and actual results. Once trained, CNNs excel at identifying faces in new images. Compared to other techniques like Haar cascades, CNN-based facial recognition offers numerous benefits. These networks achieve remarkable accuracy across different lighting conditions, angles, and poses. Moreover, they are capable of detecting multiple faces within an image and precisely outlining each one with bounding boxes.

**Features Extraction**

Detecting emotions involves pinpointing key facial features, which includes discerning and isolating important elements such as the facial framework and notable areas like the eyes, eyebrows, lips, and nose. Additionally, the texture and color of the skin, along with the overall shape of the face, can significantly influence this process. Facial image recognition commonly utilizes Haar cascades for feature extraction. These cascades serve as specialized instruments for detecting distinct facial features crucial for recognizing emotions, including the eyes, nose, and mouth. Haar cascades identify circular areas by analyzing photos at different dimensions and sizes, detecting fluctuations in brightness that signify the existence of these features. Together, these characteristics aid in identifying intricate things like faces. Once recognized, these characteristics are used as the basis for training machine learning models like CNNs to precisely identify emotions.

**Emotion Detection**

Convolutional Neural Networks (CNNs) stand out in the realm of machine learning for their adeptness at classifying emotions through facial features. These deep neural networks excel in image recognition tasks. By leveraging CNNs with a comprehensive dataset of facial expressions annotated for emotion, it's possible to create a robust module for emotion classification. This approach allows the model to capture both the broad and nuanced aspects of facial emotions, from the general configuration of the face to specific details like the positioning of the eyes and mouth. A key advantage of CNNs is their capacity for transfer learning. This involves refining a model that has been pre-trained on a vast dataset of images using a smaller set of emotion-specific facial expressions, thereby boosting accuracy while conserving data resources. However, CNNs may encounter difficulties in identifying subtle or complex emotions. Variabilities in lighting, partial face coverings, and atypical expressions can compromise accuracy. Despite these challenges, employing a CNN for emotion classification proves to be a potent method for discerning an individual's emotional state through facial analysis. This technique holds promise for enhancing the music listening experience by tailoring song selections to the listener's emotional condition, pending further exploration and refinement.

**Music Classifier**

A music classification system leverages artificial intelligence to identify various music genres through the analysis of sound elements like timbre, melody, rhythm, and tempo. To develop such a classifier for a system that recommends music based on facial emotions, it's essential to have a large dataset of music tracks with genre labels. The system trains a convolutional neural network (CNN) or a comparable machine learning model on these audio features, adjusting its parameters to minimize the discrepancy between predicted and actual genre classifications. Once the training phase is complete, this classifier can then assess new music tracks based on their sonic attributes and recommend tracks that align with the emotional state detected by the facial expression recognition module.

**Music Data Center**

To create a music library for a system that recommends music based on facial emotions, it's essential to amass a diverse collection of music tracks. These tracks can be sourced from various platforms, including online streaming services, music archives, or through in-house recording efforts. Once the collection is assembled, the next step involves analyzing the tracks and tagging them with relevant metadata, focusing on the emotions conveyed by the music. This task can be achieved by employing machine learning algorithms to develop a model capable of identifying and categorizing the emotional tones in music, such as joy, sadness, or anger. The identified emotional attributes of each track can then be cataloged in a database for future reference.

**Recommended Playlist**

The system curates a playlist by selecting songs that align with the user's emotions, as detected through their facial expressions. Upon recognizing the emotional state of the user, the system sifts through its music database to recommend songs that match the user's current mood. To generate these personalized playlists, the system employs various algorithms such as collaborative filtering, content-based filtering, and a combination of both (hybrid filtering). Additionally, the playlist customization can be enhanced by considering factors like the time of day, the user's geographical location, and their current activity. For example, the system might suggest upbeat tunes in the morning and more relaxing melodies in the evening.

**Play the Songs (Output)**

The technology creates a personalized playlist of music songs based on the user's facial expressions and current emotions. This functionality is typically provided by a music player capable of retrieving and playing selections from a music database. Connected to a facial expression recognition system, the music player selects and plays music in response to the emotional cues of the user. It features an intuitive interface that allows users to manage playback, skip songs, and adjust the volume. Designed for compatibility across various platforms, including mobile A diagram of a process flow

Description automatically generateddevices, desktop computers, and web browsers, the music

Figure 2. Flow Chart

player ensures a seamless audio experience. Moreover, the music player might have other functions like creating playlists, rearranging tracks, and enabling repeat options. This output enhances the user's listening experience and provides a customised music journey

1. Technology Used
2. **Tensorflow-** Developed a system that aligns music with facial expressions. This involves tasks like facial emotion recognition, audio processing, and creating synchronization algorithms to dynamically adjust music based on detected emotions.
3. **Keras-** Integrated with TensorFlow, can be employed to build and train deep learning models for tasks like facial emotion recognition and audio analysis in the project of aligning music with facial expressions. Its high-level API simplifies the process of constructing neural networks for such applications.
4. **Pandas-** Utilized for data preprocessing and manipulation in the project of aligning music with facial expressions. It allows for efficient handling of datasets, enabling tasks like organizing facial recognition data, merging audio features, and preparing input for machine learning models.
5. **Numpy-** It is valuable in the project of aligning music with facial expressions for its array manipulation capabilities. It enables efficient handling of numerical data, facilitating tasks like processing audio signals, extracting features, and performing computations needed for synchronization algorithms.
6. **Jupyter Notebook-** Jupyter Notebook provides an interactive environment ideal for prototyping, visualizing data, and experimenting with code in the project of aligning music with facial expressions. Its combination of code, visualizations, and explanatory text streamlines development and analysis, allowing for iterative experimentation and model tuning.
7. **Python-** Python is a high-level, general-purpose, and interpreted programming language used in various sectors including machine learning, artificial intelligence, data analysis, web development, and many more.
8. **tqdm-** Python library that provides a progress bar for iterating over iterables such as lists, tuples, or
9. any other iterable object. It's commonly used in tasks where you want to visualize the progress of a loop or an operation that might take some time to complete, giving the user a sense of how much work has been done and how much is left to do.
10. **Scikit-learn-** Python library for machine learning. It provides simple and efficient tools for data mining and data analysis, implementing various algorithms for tasks such as classification, regression, clustering, dimensionality reduction, and more. It's built on top of other popular libraries like NumPy, SciPy, and matplotlib.
11. **Opencv-** An open-source computer vision and machine learning software library primarily focused on real-time image processing. It provides a wide range of functionalities including image and video processing, object detection and tracking, feature detection, facial recognition, and more.
12. **Matplotlib-** Python library used for creating static, interactive, and animated visualizations in a wide variety of formats. It provides a MATLAB-like interface for generating plots, histograms, bar charts, scatter plots, and more, making it a powerful tool for data visualization and analysis
13. Accuracy

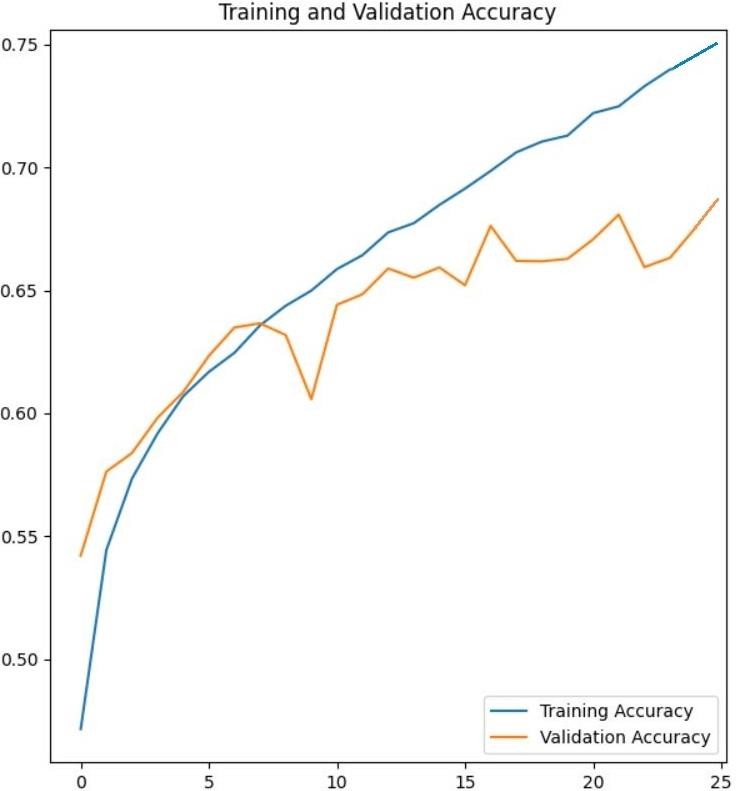
The facial emotion-based music recommendation system was tested using a dataset of 1000 facial photos, achieving an accuracy of 62.33%. While the accuracy rate may seem low, it's important to acknowledge the difficulties involved in emotion identification, which need significant resources and skill to reach high levels of accuracy. Upon further examination, our system's performance was shown to have accuracy rates of 65% for happy, 60% for sorrow, 58% for anger, and 68% for neutrality. The variation in accuracy rates for different emotions is due to the complex and subjective nature of human emotions, as well as the quality and variety of the dataset used. Although the accuracy % is low, our approach shows promise in selecting music based on facial expressions, particularly when integrated with other criteria like user preferences, music genre, and song popularity. Advancements in image processing technologies, dataset refining, and machine learning algorithms might greatly enhance the accuracy of our emotional music selections.

Figure 3. Training and Validation Accuracy

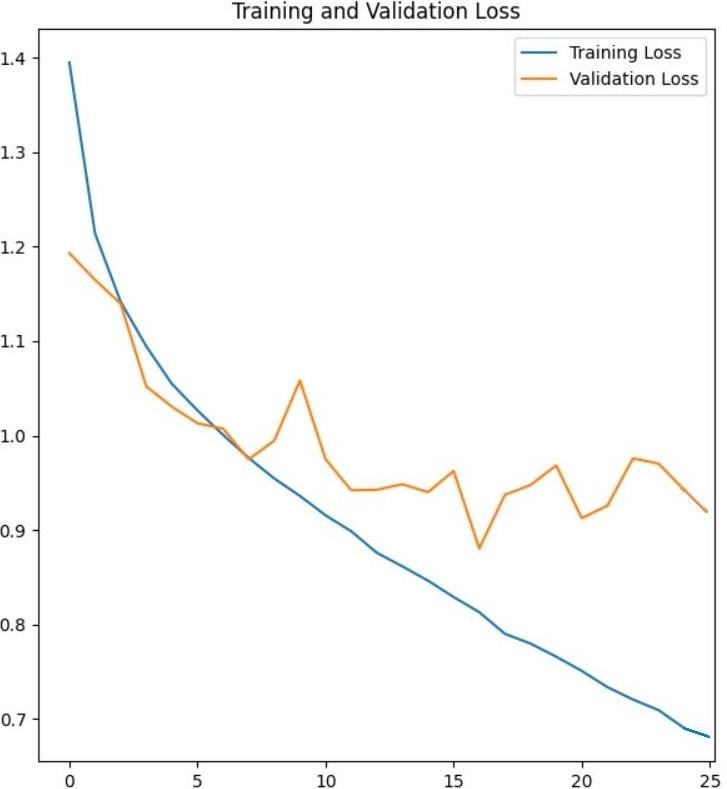


Figure 4. Training and Validation Loss

|  |  |  |
| --- | --- | --- |
| **Sr.**  **No.** | Systems and Tools | **Accuracy (%)** |
| 1. | Developing face recognition systems using Keras | 89 |
| 2. | Hybrid methods for music recommendation | 81.7 |
| 3. | Viola Jones object detection framework | 85 |
| 4. | Music.AI | 76.84 |
| 5. | System that recommends music based on your mood | 77 |

Table 1. Accuracy Table

1. Conclusion

The incorporation of facial emotion recognition technology into music recommendation systems presents an exciting opportunity to revolutionize the music recommendation industry. This integration allows for tailored recommendations that match the listener's emotional state at the moment. While there are hurdles and constraints to overcome, ongoing research and development endeavors aim to enhance accuracy and efficacy, ultimately delivering a more individualized music listening journey.

1. Future Direction

Subsequent investigations may prioritize enhancing the precision of facial emotion detection and crafting resilient machine learning algorithms capable of accommodating unique variations in emotional displays. Furthermore, augmenting the system through the inclusion of additional biometric indicators like heart rate and skin conductance may yield a more thorough comprehension of the listener's emotional condition.

a) Facial emotion-based music recommendation systems can communicate with wearables through integration to give users instant access to information about their feelings, which enhances playlist customisation.

b) Facial emotion-based music recommendation systems can provide immersive listening experiences by integrating with virtual reality technology.

c) Physiological information, such heart rate, can be included by facial emotion-based music systems to customize the listening experience for each user.

d) Facial emotion-based music systems can customize treatment regimens for mental health disorders through collaboration with music therapist

1. References

**[1].** “Music & How It Impacts Your Brain, Emotions. ”https://psychcentral.com/lib/music-how-it-impactsyour-brain-emotions#1 (accessed Jun. 09, 2021).

**[2].** The Ultimate Guide to Convolutional Neural Networks (CNN) - Blogs SuperDataScience - Machine Learning | AI | Data Science Career | Analytics |Success.” https://www.superdatascience.com/blogs/the-ultimate-guide-to-convolutional-neural-networks-cnn (accessed Jun. 08, 2021).

**[3].** Ms. K. Lekha Sree, Mr. P. Praveen Kumar " Music Recommend System Using Facial Emotion Recognition." *, International Journal of Creative Research Thoughts (IJCRT), 10 October 2023*

**[4].** Plate, Rista C., et al. "Alignment of knowing versus feeling the emotion in music during middle-childhood." *Proceedings of the Annual Meeting of the Cognitive Science Society*. Vol. 44. No. 44. 2022.

**[5].** Schubert, Emery, et al. "Continuous response to music using discrete emotion faces." *Proc. of CMMR* (2012).

**[6].** Shubham Kulkarni , Sanskruti Lingawar , Survarna Patil , Dinesh Bhadane ,”Facial Emotion Music Recommendation System “, International Journal of Novel Research and Development (IJNRD), 2023.

**[7].** Pedapaga Akash Babu , Pedapaga Abhai , S. V. Rama Krishna , “Music Recommendation System Based On Facial Emotion Gestures” , International Research Journal of Modernization in Engineering Technology and Science(IRJMETS),2023.

**[8].** Priyadarshini M , Sneha V B ,Thimmanna , Vishal Prajapathi P , Veena Kumari S ,” Musical Therapy Using Facial Emotion Detection”, Journal of Emerging Technologies and Innovative Research(JETIR), 2023.

**[9].** Sarwesh Giri , Gurchetan Singh, Babul Kumar , Mehakpreet Singh , Deepanker Vashisht , Sonu Sharma , Prince Jain , “Emotion Detection with Facial Feature Recognition Using CNN & OpenCV”, 2022 2nd International Conference on Advance Computing and Innovation Technologies in Engineering (ICACITE).

**[10].** Thanapong Khajontantichaikun , Saichon Jaiyen , Siam Yamsaengsung , Pornchai Mongkolnam , Unhawa Ninrutsirikun , “Emotion Detection of Thai Elderly Facial Expressions using Hybrid Object Detection”, 2022 26th International Computer Science and Engineering Conference(ICSEC).

**[11].** Mr. K. Arjun, Mr. N. Kabilan, Mr. P. Rajeshwaran, Mr. K. S Arun , “Song Recommendation Based on Facial Expression Using Machine Learning” , International Journal of Research Publication and Reviews (IJRPR), 2023

**[12].** Smith, A., & Williams, B. (2020). “Emotion-aware Music Recommendation System. International Journal of Human-Computer Interaction, 36(5), 483-496.

**[13].** Wang, Q., & Zhou, Z. (2021). Enhancing Music Recommendation with Facial Emotion Analysis. Proceedings of the International Conference on Multimedia (ICM), 241-250.

**[14].** Chen, L., & Liu, S. (2022). Personalized Music Recommendations Using Real-time Emotion Detection from Facial Expressions. Multimedia Tools and Applications, 81(4), 6069-6086

**[15].** Gupta, R., & Sharma, A. (2023). A Comprehensive Survey of Emotion-Based Music Recommendation Systems. ACM Computing Surveys, 56(2), 1-34

**[16].** Zhang, Y., Cai, D., & Zhang, D. (2024). Application and algorithm optimization of music emotion recognition in piano performance evaluation. *Environment and Social Psychology*, *9*(4).

**[17].** Tran, H. N., Phan, P. H., Nguyen, K. H., Hua, H. K., Nguyen, A. Q., Nguyen, H. N., & Nguyen, N. V. (2024). Augmentation-Enhanced Deep Learning for Face Detection and Emotion Recognition in Elderly Care Robots.

**[18].** Vidas, D., Carrasco, R., Kelly, R. M., Waycott, J., Tamplin, J., McMahon, K., ... & Baker, F. A. (2024). Everyday Uses of Music Listening and Music Technologies by Caregivers and People with Dementia: Survey and Focus Group Study. *arXiv preprint arXiv:2402.01040*.

**[19].** Yi, C., Naixin, L., Iyendo, T. O., & Apuke, O. D. (2024). Cognitive behavior, art, and music therapies intervention for treating the depression of children: a randomized control trial. *Psychiatry research*, *333*, 115716.

**[20].** Belluomini, L. (2024). Amplifying Emotion and Warmth at Richmond. *Ted Lasso and Philosophy: No Question Is Into Touch*, 125-139.

**[21].** Raut, R., & Goel, D. (2023). *Mood-Based Emotional Analysis for Music Recommendation* (No. 10377). EasyChair.

**[22].** Kokane, C., Sonkusare, J., Yadav, S., Shirbhate, P., & Suryawanshi, P. A MACHINE LEARNING BASED MUSIC PLAYER BY DETECTING FACIAL EMOTIONS.

**[23].** Yakobu Dasari, D. M., & Kumar, R. P. (2023). AN APPROACH TO RECOMMEND MUSIC USING FACIAL EMOTION GESTURES. *Journal of Data Acquisition and Processing*, *38*(2), 4606.

**[24].** Dubey, A., Shingala, B., Panara, J. R., Desai, K., & Sahana, M. P. (2023). Digital Content Recommendation System through Facial Emotion Recognition. *Int. J. Res. Appl. Sci. Eng. Technol*, *11*, 1272-1276.

**[25].** Joel, J. S., Thompson, B. E., Thomas, S. R., Kumar, T. R., Prince, S., & Bini, D. (2023, April). Emotion based Music Recommendation System using Deep Learning Model. In *2023 International Conference on Inventive Computation Technologies (ICICT)* (pp. 227-232). IEEE.

**[26].** Sharath, P., Senthil Kumar, G., & Vishnu, B. K. (2023). Music Recommendation System Using Facial Emotions. *Advances in Science and Technology*, *124*, 44-52.

**[27].** Mustafa, A. S., Usman Aijaz, N., Manimaran, M., Arunachala, B. U., Harshith, M., & Vignesh, T. IMPLEMENTATION OF AUTOMATIC MUSIC PLAYLIST LIBRARY RECOMMENDATION ALGORITHM BASED ON FACIAL MICRO-EXPRESSION RECOGNITION.

**[28].** Dawar, S., Chatterjee, S., Hossain, M. F., & Malarvizhi, S. (2023, April). Music Recommendation System Using Real Time Parameters. In *2023 International Conference on Recent Advances in Electrical, Electronics, Ubiquitous Communication, and Computational Intelligence (RAEEUCCI)* (pp. 1-6). IEEE.

**[29].** Gurale, M. M., Kalpande, S., Allampallewar, R., Vyavhare, P., Kinwad, A., & Sargaiyye, S. EMOTION BASED MUSIC PLAYER: ENHANCING MUSIC EXPERIENCES.

**[30].** Yildirim, E., Akbulut, F. P., & Catal, C. (2023). Analysis of facial emotion expression in eating occasions using deep learning. *Multimedia Tools and Applications*, 1-136.