STM JOURNALS

Journal of
Artificial Intelligence
Research & Advances

ISSN: 2395-6720
Volume 11, Issue 2, 2024
May–August
DOI Journal: 10.37591/JoAIRA

https://journals.stmjournals.com/joaira

Review JoAIRA

Real-time Facial Recognition with Convolutional Neural Networks for Personalized Music Therapy

Shivam Kumar[1], Rameshber Goswami[1], Yash Tripathi[1], Srishti Malu[1], Anjuli Dubey[2,*]

Abstract

In this project, a web-based application has been developed that integrates computer vision-based facial recognition, multiple algorithms, and machine learning approaches. The given system obtains a user's emotions in the real-time frame by analyzing facial expressions such as eyes, mouth, the forehead, and so on. It detects emotions like happiness, sadness, that is neutrality, or rock. For a given detected emotion, language, and a user's chosen artist, the system recommends a song or music that matches a user's mood. The deep learning model used in the given system is trained on the FER-2013 dataset, which is nothing more than an annotated dataset of facial images. The system is based on real-time video feeds, which recommend an emotional level based on a user's emotional state. Thus, the presented system is a widely innovative tool that has the potential to revolutionize music consumption and generally improve the user's state through accurate mood-based music selection. A performance assessment involving a labeled image dataset resulted in a detection accuracy rate of around 81%. Subsequent user studies ensured the system's ability to recommend music tunes, which accurately reflect the user's mood and personality.

Keywords: Deep learning, computer vision, deep CNN, facial recognition technology, emotion recognition, music recommendation system, vision, machine learning, convolutional neural network, mood detection

INTRODUCTION

Facial recognition technology has become increasingly popular across various sectors such as marketing, healthcare, and security. Despite its widespread use, its application in the music industry is still relatively unexplored. The ability of this technology to analyze human emotions through facial expressions offers a unique opportunity to improve the music listening experience by providing personalized song recommendations tailored to the user's mood. This study aims to delve into the potential of facial recognition to identify emotions accurately and to suggest music accordingly, which could significantly alter how we interact with digital platforms and contribute to emotional well-being [1–5].

Music profoundly affects our emotions, yet navigating a vast library to find songs that resonate with our current mood can be daunting. Here, facial recognition can be pivotal. It can evaluate facial expressions in real-time to pinpoint the user's current emotional state. Using machine learning algorithms, this system can identify a range of

***Author for Correspondence**
Anjuli Dubey
E-mail: anjuli.dubey@poornima.org

[1]Student, Department of Computer Engineering, Poornima College of Engineering, Jaipur, Rajasthan, India
[2]Assistant Professor, Department of Computer Engineering, Poornima College of Engineering, Jaipur, Rajasthan, India

Received Date: May 06, 2024
Accepted Date: June 30, 2024
Published Date: July 10, 2024

Citation: Shivam Kumar, Rameshber Goswami, Yash Tripathi, Srishti Malu, Anjuli Dubey. Real-time Facial Recognition with Convolutional Neural Networks for Personalized Music Therapy. Journal of Artificial Intelligence Research & Advances. 2024; 11(2): 67–76p.

emotional indicators, including happiness, sadness, anger, surprise, and neutrality. Integrating these technologies allows for a sophisticated music recommendation system that adapts song selections to the user's detected emotions, offering a tailored and engaging listening experience [6].

This pioneering method could establish a model for similar technologies across various industries, broadening the applications of facial recognition. This research explores the implementation of facial recognition in the music industry and examines its broader implications. The ultimate objective is to develop a cost-effective and efficient facial recognition system that not only identifies emotions but also enhances the music listening experience by selecting songs that align with the user's emotional state [7, 8].

LITERATURE REVIEW

In their paper, Dharsini *et al.* explored the potential applications of facial recognition technology for emotion detection across multiple sectors [9]. They emphasized the deep link between music and emotions, proposing an advanced music recommendation system that leverages facial recognition to ascertain a user's emotional state. The authors claim that their algorithm outperforms existing methods in efficiency, significantly cutting down the time and effort needed for manual operations. Their system aims to identify emotional cues from facial expressions and suggest music in a cost-effective and time-efficient manner.

Naik *et al.* explored the development of music player systems capable of interpreting facial expressions to gauge emotions [10]. Their research advocates the use of Convolutional Neural Networks (CNN) for both recommending music and recognizing emotions. Their strategy centers on utilizing computer vision and machine learning to interpret human emotions, aiming to create a music recommendation system based on these emotional insights.

In their paper, Samuvel *et al*. explored the use of face recognition technologies in fields such as security and digital video processing [11]. They emphasized music's powerful emotional influence and suggested creating an effective music recommendation system that leverages facial recognition to determine the user's mood. This system aims to enhance the user experience by aligning music choices with emotional states, leveraging the emotive power of music.

The proposed algorithm is designed to surpass the efficiency of existing music recommendation systems by identifying facial expressions to suggest songs, focusing on both time and cost efficiency. Athavle *et al.* introduced a novel approach for automatic music playing based on facial emotion detection, which contrasts with traditional methods involving manual sorting or wearable devices [12]. Their system utilizes a Convolutional Neural Network for detecting emotions and employs Tkinter & Pygame for generating music recommendations. It aims to enhance accuracy while minimizing computational time and costs, with tests conducted on the FER2013 dataset for emotion detection from facial features.

Mahadik *et al.* explored using computer vision and machine learning to assess a user's mood from live camera feeds [13]. Their system, integrating the compact MobileNet model with Keras, facilitates Android-ML integration, offering a music recommendation feature that adds value to conventional music player applications. Music, known for its universal appeal and mood-regulating properties, is at the core of this system designed to boost user satisfaction.

The paper by Joshi *et al.* reviewed challenges in music recommendation, particularly in matching songs to emotions [14]. They proposed that integrating Natural Language Processing and Deep Learning could allow systems to interpret emotions from texts and facial cues more effectively. By comparing various deep learning models, the optimal one was chosen for a music recommendation application that accepts inputs as text or facial expressions, enhancing accuracy through a CNN for emotion detection.

Chidambaram *et al*. discussed the potential of facial recognition to analyze distinct facial features and determine mood [15]. Their system creates a personalized or default playlist based on this mood assessment, obviating the need for manual song selection. They provided insights into the system's design for an emotion-based music player, including details on playlist creation and emotion classification, highlighting the practical application and efficiency of the proposed system.

Sashank *et al*. highlighted the pervasive stress in today's fast-paced society in their research paper, and discussed how music can mitigate stress when appropriately aligned with the user's current mood [16]. They noted a gap in the market: music applications typically do not suggest songs based on the listener's emotional state. To address this, they proposed a music player app that uses facial recognition to detect emotions such as anger, happiness, and sadness. The app allows users to upload a current or past photo, or to enter their mood via text, to receive tailored music recommendations that can enhance their emotional well-being.

Supriya *et al*. took this concept further in their work on an emotion-based music player [17]. Their project, rooted in affective computing, employs facial recognition to assess the user's emotions and automatically selects music that corresponds to these emotions, thereby simplifying the search process. The system employs a webcam to capture the user's image and analyze facial features to recognize emotions. A Support Vector Machine (SVM) model is used to classify these emotions, integrating both facial and auditory data to create a cost-effective playlist tailored to the user's current emotional state. This innovative approach streamlines the process of music selection, making it more responsive and personalized to the user's emotional needs.

PROPOSED SYSTEM

The objective of this project is to create a real-time emotion detection system employing computer vision and machine learning methods. The system will assess facial features like eyes, mouth, and forehead contours to recognize emotions like happiness, sadness, anger, surprise, neutrality, and rock. Using the identified emotion, in addition to user-chosen language and artist preferences, the system will recommend a suitable genre of music.

To enable this, a substantial dataset of facial expressions will be employed to train a facial recognition algorithm, using machine learning methods to ensure precise emotion interpretation. This training will help build a deep learning model that can recognize emotional states from facial images.

The deep learning model will be specifically calibrated to identify emotions from real-time video feeds effectively. After training, this model will be integrated into the facial recognition system, which will then be capable of real-time emotion detection. Utilizing the user's facial expressions, the system will accurately assess emotional states and recommend music that aligns with both the detected mood and the user's past music preferences, thereby enhancing the personalized listening experience.

The proposed project centers on several key objectives designed to enhance how users interact with music based on their emotional state.

Data Collection

Assemble a comprehensive dataset of facial images, each tagged with specific emotions such as happiness, sadness, and neutrality. This dataset will form the core training material for the deep learning model.

Model Training

Train a deep learning model using the gathered dataset to recognize different emotions depicted in facial images. This training process will involve sophisticated machine learning techniques to ensure the model can accurately interpret emotional expressions.

System Implementation

Integrate the trained model into a facial recognition system that can detect emotions from real-time video or still images. This system will be engineered to process input promptly and accurately, ensuring that emotional assessments are both timely and reliable.

Emotion-Song Matching

Develop a mechanism within the system to link detected emotions to specific categories of music stored in a database. This involves creating a logical mapping from each recognized emotion to a corresponding set of songs that reflect or complement the mood represented by that emotion.

Music Recommendation

Based on the emotion identified and the user's historical preferences and mood, recommend songs from the matched category. This recommendation engine will tailor selections to enhance the user's listening experience, making it more personalized and emotionally resonant.

Overall, this project aims to leverage cutting-edge technology in computer vision and machine learning to create a user-centric music recommendation system that not only understands but also responds dynamically to the emotional cues of its users as shown in Figure 1.

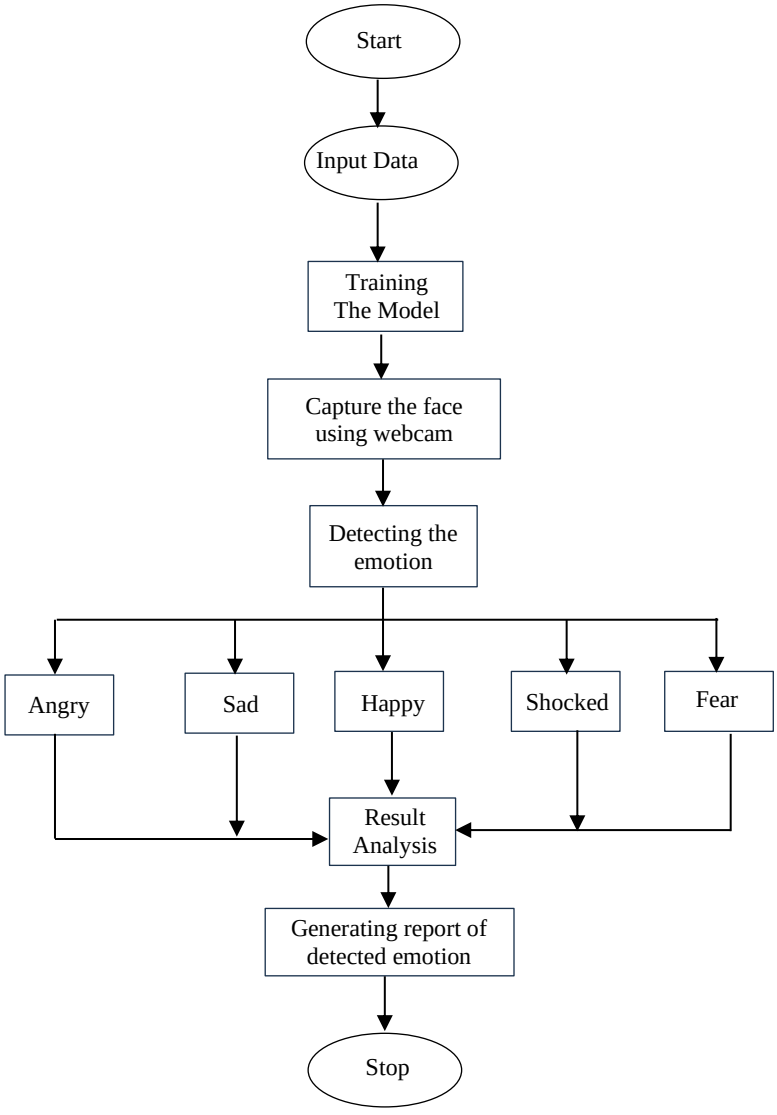

Figure 1. Proposed approach.

OBJECTIVE

The primary aim of this endeavor is to create an advanced facial recognition system capable of precisely determining a user's mood through the analysis of their facial expressions and physiological cues. To support this, the project will also create a comprehensive song database, categorized by mood, genre, and artist, which will be used to tailor music recommendations to the user's current emotional state.

Key Objectives Include

Facial Recognition System Development

Design and implement a system that uses advanced algorithms to interpret facial expressions and physiological cues effectively, providing an accurate gauge of the user's emotional condition.

Song Database Creation

Compile a diverse collection of music that encompasses various moods, genres, and artists. This database will serve as the foundation for the recommendation engine, allowing it to suggest music that aligns with the user's detected mood.

User Interface Design

Create a user-friendly interface that makes it simple for users to submit their mood data and receive music recommendations. This interface will be accessible to users of all technological proficiencies and will include features for saving favorite tracks and creating personalized playlists based on mood fluctuations.

System Testing and Evaluation

Rigorously test the system to validate its accuracy and effectiveness in real-world scenarios. This testing will determine how well the system can detect mood variations and recommend appropriate music to a variety of users.

By accomplishing these objectives, the project aims to revolutionize the music listening experience, making it a dynamic and responsive interaction that adjusts to the user's emotional state in real time. This could be particularly beneficial for individuals who find it challenging to articulate their emotions or select music that complements their current mood. The user-friendly interface will ensure that all users, regardless of their technical skills, can enjoy a seamless and personalized music experience tailored to their emotional needs.

METHODOLOGY

This research work outlines a detailed approach for developing a comprehensive music recommendation system that integrates facial recognition technology. The project is structured in several key phases, each crucial to ensuring the system's overall effectiveness and user satisfaction [18–21, 9].

Development of the Facial Recognition Algorithm

- *Data Collection:* Gather a dataset of facial images from participants, annotated with diverse emotions, captured under controlled lighting conditions to maintain consistency.
- *Data Preprocessing:* Enhance the quality of the collected facial images by eliminating noise and other extraneous details that could affect the accuracy of the emotion detection process.
- *Feature Extraction:* Apply computer vision and machine learning techniques to the preprocessed images to extract relevant features that accurately identify various facial expressions and associated emotions such as happiness, sadness, rock, and neutrality, as shown in Figure 2.

Creation of the Song Recommendation System

This system will use facial recognition algorithms to detect emotional states and recommend songs based on them. It will also incorporate user preferences for specific singers and languages to further tailor the music selections to individual tastes.

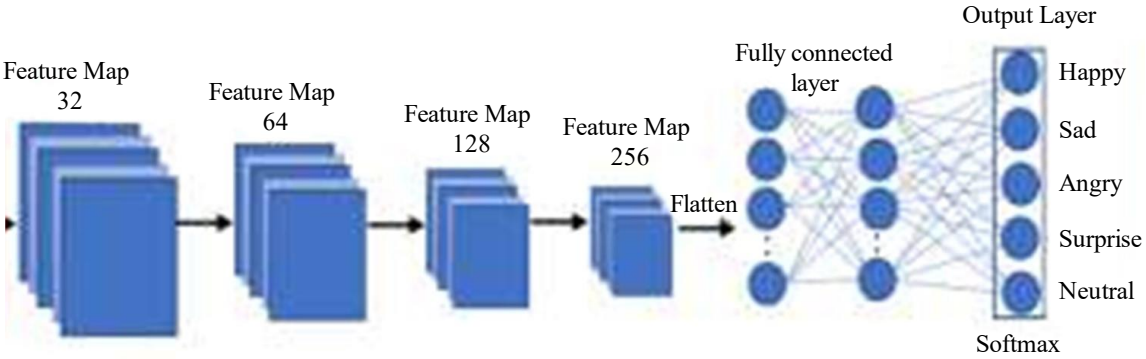

Figure 2. Convolution neural network Architecture.

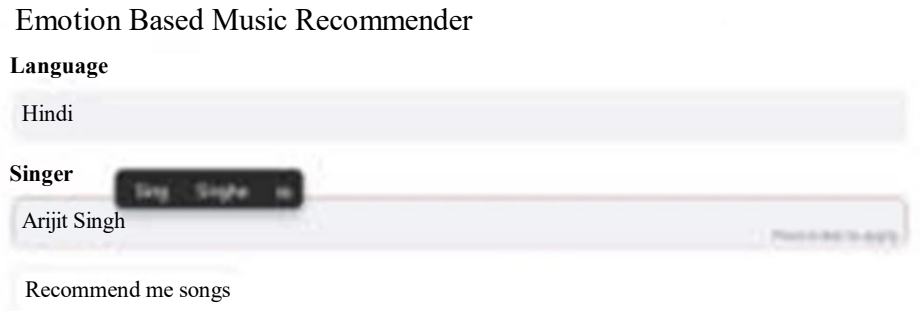

Figure 3. Hosted Application.

Validation through User Study

A user study will be undertaken to evaluate the efficiency of the facial recognition and song recommendation system. This study will evaluate the accuracy of the mood detection and the appropriateness of the song recommendations, providing crucial feedback that can be used to refine the system.

By following these steps, the project aims to create a dynamic and responsive music recommendation platform that not only understands but also adapts to the emotional needs of its users, thereby personalizing the music listening experience to an unprecedented degree.

EXPERIMENTAL SETUP

We conducted an extensive study with a limited number of participants to evaluate the effectiveness of our music recommendation system, which utilizes facial recognition technology. This study was designed to test the system's ability to accurately detect participant moods from facial expressions and to recommend songs that align with these moods as shown in Figure 3.

Data Collection

Participants were subjected to various stimuli to elicit six predefined emotions: happiness, sadness, anger, surprise, neutral, and rock. These emotions were induced through different methods such as watching videos or listening to specific types of music. During these sessions, their facial expressions were captured using a high-definition camera in a controlled setting to ensure consistency.

Experimental Design

The study utilized a within-subjects design, meaning each participant experienced all emotional stimuli in a random sequence. This approach helped mitigate any bias that could arise from the order of exposure. The controlled environment with consistent lighting was crucial for maintaining the integrity of the facial data captured.

Data Analysis

Advanced computer vision techniques and machine learning algorithms were utilized to process the gathered facial images. The primary tool was a deep learning model that had been pre-trained on a large dataset to recognize facial expressions associated with specific emotions. We tested this model on the new images collected from participants to measure its effectiveness in real-time mood detection.

The model's output was used to trigger recommendations from a song database, which was organized by mood, language, and artist preferences. This structured approach allowed us to assess not only the accuracy of the emotion detection but also the relevance of the song recommendations made by the system as shown in Figure 4.

RESULTS

The correlation between personality traits, emotions, and musical preferences is well-documented, reflecting the profound interaction between the brain's emotional processing centers and the musical elements of metre, timber, rhythm, and pitch. This connection formed the theoretical basis for our facial recognition-based music recommendation system, which was tested extensively in our study.

Performance Metrics

In our evaluation, the system demonstrated high efficacy in mood detection and music recommendation, as sown in Table 1:

- *F1 Score:* Achieved 82.35%, indicating a strong balance between precision and recall in our emotion classification.
- *Total Classification Accuracy:* Recorded at 81.50%, showing that the system correctly identified the mood of the participants in most instances.
- *Precision:* At 87.50%, this metric underscores the system's ability to limit false positives in detecting emotional states.

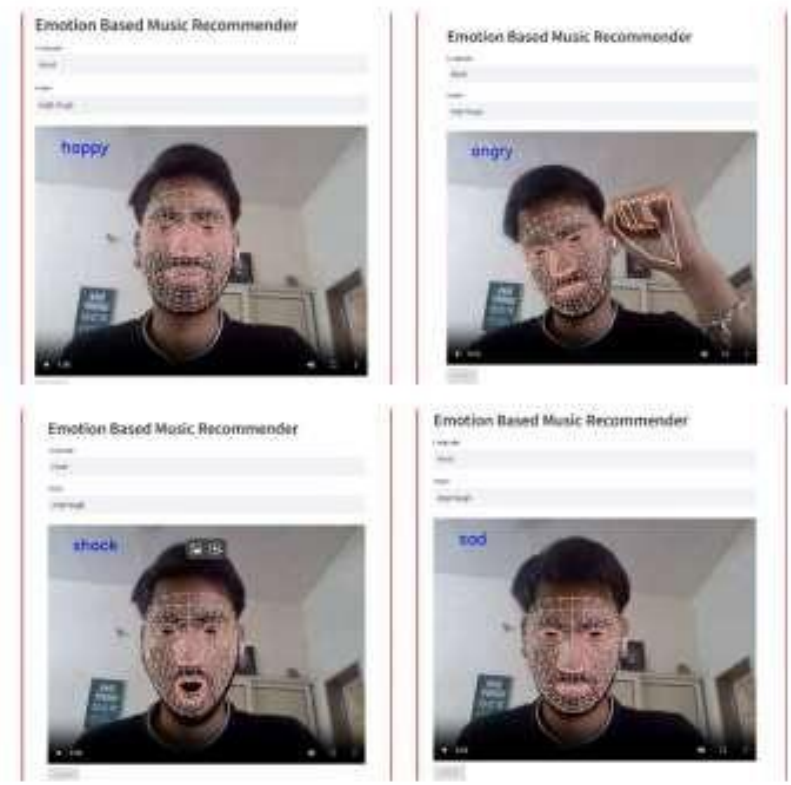

Figure 4. Emotion Detection.

Table 1. Validation and Testing accuracy.

Metric	Value	Description
F1 Score	82.35%	Measures the test's accuracy and considers both the precision and the recall to compute the score; a higher score indicates better accuracy and balance.
Total Classification Accuracy	81.50%	Indicates the overall rate of the system correctly identified and classified the participants emotions.
Precision	87.50%	Reflects the accuracy of positive predictions, highlighting the system's ability to identify true positives among all positive results.
Participant Satisfaction	High	Participants expressed a strong satisfaction with both the accuracy of mood detection and the suitability of the music suggestions.

Music Classification and Recommendation

The music classification module, integral to tailoring song recommendations, performed exceptionally well. It successfully matched music choices to the mood categories derived from the participants' facial expressions, ensuring that the music resonated with their current emotional states.

Study Findings and Participant Feedback

The research results validated the efficacy of the suggested system, with participants expressing considerable satisfaction towards the music recommendations. They particularly valued the relevance and suitability of the selections tailored to their detected emotions. The positive feedback highlighted not only the system's technical accuracy but also its practical utility in enhancing the music listening experience.

This outcome emphasizes the promise of combining advanced recognition technologies with personal entertainment systems. By accurately interpreting users' emotions through facial expressions and recommending music that aligns with their mood, the system offers a more personalized and emotionally responsive listening experience. This capability could transform user interactions with music platforms, making them more engaging and supportive of emotional well-being.

CONCLUSION AND FUTURE SCOPE

The primary aim of this project is to enhance the user's emotional wellbeing by accurately detecting their mood and delivering music that complements their emotional state. Facial expression recognition, a technique with a rich history, serves as a powerful tool for understanding and interpreting human emotions. Employing this technology in the music industry to suggest songs based on mood detection is a burgeoning field with significant promise.

Potential of Facial Recognition in Music Recommendation

The application of facial recognition technology for mood-based music suggestion represents an innovative approach to personalize the music listening experience. By aligning music with the listener’s current emotions, the system promises to make the experience more personal and immersive, potentially transforming how users interact with music platforms.

Future Directions and Applications

The proposed system offers several avenues for further development:

- *Integration with Wearable Devices:* Linking the system with wearable technology could allow for continuous mood monitoring and music adjustment, providing a seamless way to enhance user mood throughout the day.
- *Expanding Emotional Range:* Current capabilities could be expanded to detect a broader spectrum of emotions, providing even more nuanced music recommendations.
- *Language Adaptation:* Making the system multilingual would broaden its applicability, catering to a global audience with varied linguistic needs.

- *Application in Other Fields:* Beyond music, this technology could be adapted for movie recommendations, where viewer mood could dictate the genre of movies suggested, enhancing the viewing experience.
- *Mental Health Therapy and Marketing:* The technology could also find applications in mental health therapy, aiding in emotion regulation strategies, or in marketing, to better gauge consumer responses and tailor advertisements based on emotional reactions.

Enhancing Wellbeing and User Experience

Continued research and development of this technology could substantially benefit human wellbeing and enrich musical experiences, making them more responsive and tailored to individual emotional states. This could lead to broader implications for user interaction across various digital platforms, offering more adaptive and empathetic user technology interactions.

REFERENCES

1. Baddur RP, Shedole S. A novel approach for sentiment analysis using deep recurrent networks and sequence modeling. Recent Pat Eng. 2020 Dec 1; 14(3): 403–11.
2. Ni R, Yang B, Zhou X, Cangelosi A, Liu X. Facial expression recognition through cross-modality attention fusion. IEEE Trans Cogn Develop Syst. 2022 Feb 9; 15(1): 175–85.
3. Narimisaei J, Naeim M, Imannezhad S, Samian P, Sobhani M. Exploring emotional intelligence in artificial intelligence systems: A comprehensive analysis of emotion recognition and response mechanisms. Ann Med Surg. 2024 Jun 21. 86(8):4657–63.
4. De Witte M, Pinho AD, Stams GJ, Moonen X, Bos AE, Van Hooren S. Music therapy for stress reduction: a systematic review and meta-analysis. Health Psychol Rev. 2022 Jan 2; 16(1): 134–59.
5. Yang X, Dong Y, Li J. Review of data features-based music emotion recognition methods. Multimedia systems. 2018 Jul; 24: 365–89.
6. Sharma VP, Gaded AS, Chaudhary D, Kumar S, Sharma S. Emotion-based music recommendation system. In 2021 IEEE 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO). 2021 Sep 3; 1–5.
7. Sana SK, Sruthi G, Suresh D, Rajesh G, Reddy GS. Facial emotion recognition based music system using convolutional neural networks. Materials Today: Proceedings. 2022 Jan 1; 62: 4699–706.
8. Vaishnavi G, Sumathi R, Anvitha K, Bathineed D, Nikhitha B, Vanaja K. Music Recommendation Based on Facial Expressions and Mood Detection using CNN. In 2023 IEEE International Conference on Computer Communication and Informatics (ICCCI). 2023 Jan 23; 1–4.
9. Visnu Dharsini S, Balaji B, Kirubha Hari KS. Music recommendation system based on facial emotion recognition. J Comput Theor Nanosci. 2020 Apr 1; 17(4): 1662–5.
10. Kullayappa KC, NaikCh.Hima Bindu M.Hari Babu J.Sriram Pavan. Emotion Based Music Recommendation System using CNN. Int J Appl Eng Res. 2020; 5(2): 2666–2795.
11. Samuvel DJ, Perumal B, Elangovan M. Music recommendation system based on facial emotion recognition. 3C Tecnologia. 2020 Mar 1: 261–71.
12. Athavle M, Mudale D, Shrivastav U, Gupta M. Music recommendation based on face emotion recognition. J Infor Electr Electron Eng. 2021 Jun 9; 2(2): 1–11.
13. Mahadik A, Milgir S, Patel J, Jagan VB, Kavathekar V. Mood based music recommendation system. Int J Eng Res Technol. 2021 Jun; 10(06): 553–559.
14. 7Joshi S, Jain T, Nair N. Emotion based music recommendation system using LSTM-CNN architecture. In 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT). 2021 Jul 6; 01–06.
15. Plichoski GF, Chidambaram C, Parpinelli RS. A face recognition framework based on a pool of techniques and differential evolution. Inf Sci. 2021 Jan 8; 543: 219–41.
16. Sashank MS, Maddila VS, Krishnasai P, Boddu V, Karthika G. Mood-Based Music Recommendation System Using Facial Expression Recognition and Text Sentiment Analysis. J Theor Appl Inf Technol. 2022 Oct 15; 100(19): 5667–5674.

17. Supriya LP, Khilar R. Affective music player for multiple emotion recognition using facial expressions with SVM. In 2021 IEEE 5th international conference on I-SMAC (IoT in social, Mobile, analytics and cloud) (I-SMAC). 2021 Nov 11; 622–626.
18. Eliyajer G, Natarajan B, Bhuvaneswari R, Elakkiya R. A Novel Approach for Song Recommendation System Using Deep Neural Networks. In 2023 IEEE World Conference on Applied Intelligence and Computing (AIC). 2023 Jul 29; 382–387.
19. Shabu SJ, Janaardhan C, Bhaskar K, Mary AV, Refonaa J, Dhamodaran S. Music Recommendation System based on Facial Expression. In 2023 IEEE 4th International Conference on Electronics and Sustainable Communication Systems (ICESC). 2023 Jul 6; 908–912.
20. Gupta M, Venkatesh SN, Suraskar R, Praveena K, Jegadesan S, Suneetha S. Enhancing Music Recommendations with Emotional Insight: A Facial Expression Approach in AI. In 2023 7th IEEE International Conference on Electronics, Communication and Aerospace Technology (ICECA). 2023 Nov 22; 450–457.
21. Chen X, Tang TY. Combining content and sentiment analysis on lyrics for a lightweight emotion-aware Chinese song recommendation system. In Proceedings of the 2018 10th International Conference on Machine Learning and Computing. 2018 Feb 26; 85–89.

