iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** V **Month of publication:** May 2026

DOI: https://doi.org/10.22214/ijraset.2026.81712

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

Emotionally Aware Music Playback: MoodSync

Diya Goyal[1], Sakshi Gupta[2], Simarjit Kaur[3] (Supervisor)
Department of Computer Science and Engineering Chitkara University Institute of Engineering and Technology

Abstract: ***In recent era stress, anxiety, and emotional exhaustion are rising because of accelerated lifestyles and constant digital exposure. Due to this mental health challenges has heightened. To help overcome these challenges, music is a proven therapy for managing emotions. A study was happened recently, which showed that 256 participants attended a music therapy session. The impact this session created was remarkable as an exponential increase in emotional resilience and overall well-being was observed.***

But the music playback system is not designed according to the user's mood. The selection of music is intervened by the user. To bridge this gap, this paper is introducing MoodSync. It is an AI-driven system that captures the mood of the user via webcam face detection and according to the language selected it plays the music.

This helps to deliver adaptive music therapy. The strategy is to capture the user's facial expression through a webcam and then analysing it with Convolution Neural Network (CNN). The research uses multiple trending technologies too support emotional health. Mood Sync provides a flexible, scalable framework that can advance emotion-responsive technologies in mental health support and human-computer interaction domains.

Index Terms: Facial Emotion Recognition, Convolutional Neural Networks, Mental Health Applications, Computer Vision, Deep Learning, Personalized Emotional Therapy.

I. INTRODUCTION

A. Background of Mood Sync

In today's era, rising levels of stress and emotional imbalance are very common. Despite multiple music streaming platforms available, most of them rely on user's intervention and they ignore real-time emotional states. Mood Sync aims to bridge this gap by introducing an intelligent system that understands human emotions through facial expressions and responds with contextually appropriate music.

B. Evolution of Emotion-Aware Systems

Advancements in technologies have enabled machines to interpret human emotions through facial recognition and other ways. Technologies such as deep learning and computer vision have significantly improved the accuracy of emotion recognition.

C. Importance in Modern Education

In this fast-paced world, mental health awareness has become so crucial among students and working professionals. To help combat this situation, mood sync comes into action. Mood Sync actually incubates music therapy which is used as a non-invasive method to regulate the emotions. It is an automated system which plays music according to the user's mood and language preference.

D. Purpose of the Study

The objective of this research is to combat the shortcomings of conventional music system, which rely on user's intervention. Such approaches generally fail to incorporate real-time emotional awareness and cannot adapt dynamically to the listener's present emotional state.

The primary scope of this study is to bridge this gap by incubating emotion recognition technique with music mechanisms, thereby creating a more personalized music playback system. This work contributes to the development of affective computing applications in entertainment and well-being by enabling systems that respond sensitively to users' immediate affective needs rather than static profiles alone. Figure 1 shows the interface our system where the mood of the user is recognized.

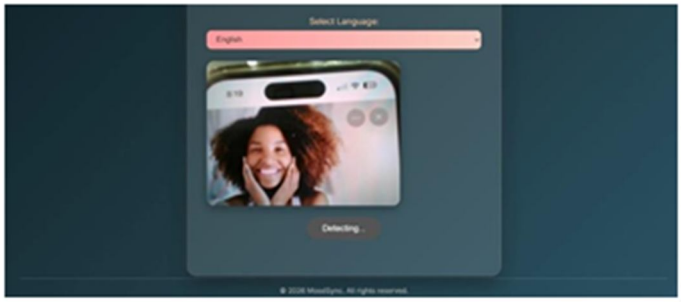


Fig. 1. Detection of mood.
The figure represents MoodSync's Interface during detection. It also showcases the language of music selected as English.

II. LITERATURE REVIEW

Some recent studies have demonstrated the progress in facial recognition using CNN where high accuracy is achieved in classifying core emotions such as sadness, anger, joy etc. Some instances include like a model was evaluated on controlled datasets where accuracies of 95-98.5% was achieved [2], while performance on more challenging, real-world datasets like FER2013 typically ranges from 64–81%, depending on preprocessing, model complexity, and techniques like transfer learning or attention mechanisms. In the era of computers, researches has been progressing consistently integrating emotion detection with personalized content delivery.

Orthodox music playback systems have been largely depending on collaborative filtering, content-based methods, historical patterns and popularity metrics. More recent works, however, incorporates contextual signals- incubating mood, activity, time of day etc. to create more adaptive suggestions [3]. Nevertheless, few systems fully combine real-time facial recognition with music playback systems. Mood Sync fills this crucial gap through the seamless integration of computer vision for live facial analysis, deep learning models for accurate emotion classification.

Mood Sync captures this idea by integrating emotion recognition through webcam using computer vision and according to the language chosen, our system plays the music [6]. If the mood detected is sad, the music played will be comforting and joyful to console the user's mood and daily activity hustle.

Fig. 2. Research Framework of the Study. The image represents the methods which led up to the development of MoodSync, the steps involved in emotion detection and music playback and the outcomes of the proposed system.

III. PROPOSED METHODOLOGY

The proposed methodology outlines the step-by-step process of Mood Sync: capturing facial input via web cam, detecting and classifying the user's emotion.

A. System Design

This study inculcates a technical architecture keeping in mind the flexibility, maintainability, and seamless integration across its core components. The system consists of four main modules [5]. The Emotion Detection Module, built in Python, uses OpenCV for real-time face detection and preprocessing. It is paired with CNN to accurately identify the user's current emotional state from facial expression. At last the User interface is a responsive web-based platform that enables we-cam access, showing real time emotion feedback.

B. Data Preparation

The system uses emotion-labeled datasets to train the CNN model. The dataset used in this project has 35,685 facial recognized emotions, where each image is resized to 48x48 pixels [1]. Later, these resized images are converted into grayscale format to reduce the computational complexity. While reducing this complexity, all key facial features were preserved. The data is divided into training and testing data. The training data is used by the model to learn the patterns whereas testing data is used to evaluate the model's performance on the unseen data. The images are grouped into emotion-specific folders, where angry emotion has nearly 958 files and similarly for other emotions as well.

C. Data Analysis

The music generation in Mood Sync uses the detected emotion and user's language preference (English, Hindi, or Punjabi) to build a dynamic mood detection. The dataset chosen for mood sync highlights a class imbalance, with certain emotions like happy and neutral having more samples of images [4]. After collecting the responses, the data is organized and prepared for analysis. Noisy or the inconsistent data is removed for the better accuracy for the study.

D. Evaluation of Key Factors

The collected data is analyzed to evaluate key factors affecting Mood Sync performance: emotion detection accuracy and reliability, input image quality, music playback effectiveness, and real-time responsiveness. This helps assess how well the system adapts to the user's emotional state, delivers personalized mood-based music, and supports emotional well-being. Fig 3 is a generalized study pattern where each research starts with an objective followed by reviewing other studies and then later on accordingly selecting the dataset for our own research.

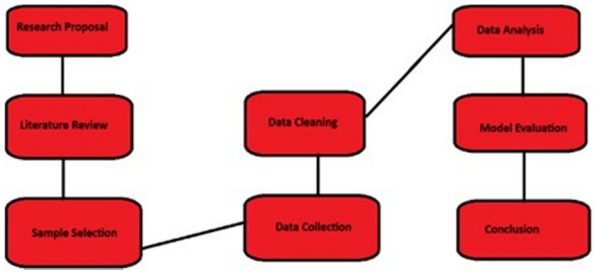


Fig. 3. Research Methodology Process. The image showcases the steps involved in the development of the research paper, inclusive of the proposal, review sample collection and cleaning and final presentation.

IV. PROPOSED FORMULATION

A. Emotion Detection Factors

The performance of mood sync is determined by various critical factors which influence the robustness of emotion detection. The first factor is facial expression recognition accuracy. The system is reliant on CNN to effectively classify emotions which is further dependent on the quality of training data and the model's ability to generalize across diverse facial features and expressions. The second factor is image quality and resolution. High quality input images allow precise detection of facial details such as eye movement, lip curvature and micro expressions. Low quality or blurred images can lead to incorrect predictions. The third factor is lightning conditions. Poor lightning, shadows and overexposure can lead to the distortion of facial features. This has been addressed through normalization and contrast adjustment. Further, the facial alignment can influence results. Misalignment or partial obstruction (e.g., face turned sideways or obstructed by an object) can negatively affect the results. Further, the robustness of the CNN model plays a critical role. A well-trained model with appropriate layers, activation functions, and regularization techniques ensures reliable real-time performance.

B. User Engagement

Engagement refers to how actively users interact with the system and respond to the recommended music. Higher user engagement indicates better user satisfaction and system relevance. The following factors have been incorporated to enhance user engagement. First, real time emotion detection, which ensures that the recommendations align closely with the user's current emotional state. Second, personalized music recommendations, which increase the likelihood of users connecting with the recommended music, resulting in prolonged usage.

Third, user friendly interface, which ensures smooth navigation, minimal response delay and clear output presentation, enhancing user retention. Fourth is multilingual support, which enhances accessibility and inclusivity for users from different backgrounds.

C. Recommendation Effectiveness Indicators

This metric uses a set of quantitative and qualitative indicators to assess system performance. First is user satisfaction, which can be measured through explicit feedback like ratings or likes, or implicit behaviour such as continued listening and reduced skipping of tracks. Second is interaction duration, which assesses the amount of time a user spends on interacting with the system. Longer session durations typically imply higher effectiveness. Fourth is the frequency of system usage, wherein repeated usage can be an indicator of user trust and satisfaction. Fifth is skip rate and playback continuity where a lower skip rate and higher playback continuity can indicate effectiveness.

D. Emotion detection

Emotion detection forms the core component of the Mood Sync system as it indicates the effectiveness of the CNN model. Any misclassification may result in irrelevant recommendations thereby impacting user experience. Further, a strong correlation exists between emotion detection accuracy and the perceived quality of recommendations. This relationship can be determined by comparing the model's prediction accuracy with user feedback metrics. Furthermore, the accuracy of emotion detection can directly impact the user's psychological state, highlighting the system's therapeutic significance. Therefore, analysing and optimizing emotion detection performance is critical for improving the overall effectiveness, reliability and user acceptance of the proposed system

Fig. 4 User Experience WorkFlow. The image showcases the steps involved in emotion detection on the client-side of the proposed system.

V. RESULTS

The proposed system was tested in real time to evaluate its performance. The observations suggest that the system was able to accurately detect emotions and play context relevant music.

During testing, the system was able to identify the following emotions – happiness, sadness, anger, calmness, surprise and neutrality. Sample observation indicates that the detection of happy emotion led to the playback of upbeat music. The detection of sad emotion led to the playback of happy or calming music. The detection of anger led to the playback of calming music. This shows the capability of the system to accurately map emotions to music.

Further, the system showed efficient time performance. This is corroborated through the fact that the system takes only 45 seconds between detection and playback of music.

Further testing results showed that out of 10 samples, the system shows accurate detection in eight of the samples.

The following table (Table 1.1) is a confusion matrix table which indicates the result of testing.

ACTUAL EMOTION	PREDICTED HAPPY	PREDICTED SAD	PREDICTED ANGRY	PREDICTED NEUTRAL	TOTAL
HAPPY	2	0	0	0	2
SAD	0	2	0	0	2
ANGRY	0	1	1	0	2
NEUTRAL	0	0	1	1	2
TOTAL	2	3	2	1	10

Table 1.1. Confusion Matrix depicting the accuracy of the proposed model.

Overall, the results indicate that with 80% accuracy MoodSync is able to effectively identify the emotions and map it to relevant music.

VI. CONCLUSION

In conclusion, the paper complements the current strides in the field of AI, by the proposal of an emotionally aware music playback system – MoodSync. MoodSync has been explored as a potential supplement in the sphere of mental health and well-being, which can utilize music therapy for patient treatment. Further, the paper traversed through various factors which lead to the effectiveness of the proposed system, including the strength of the CNN model, the variety in the input dataset and the focus on user satisfaction.

The results of the research further indicate that MoodSync was able to effectively identify emotions and play context-relevant music which highlights MoodSync's capability to positively influence the user's mood.

However, the system is still sensitive to factors such as the lighting conditions, low quality or blurred images, which can lead to unexpected results. This brings to the fore, the need for further improvements in the proposed system, including advanced model training and supplemental optimization.

Despite the challenges, MoodSync can be considered as a pragmatic and constructive solution in the discipline of mental health.

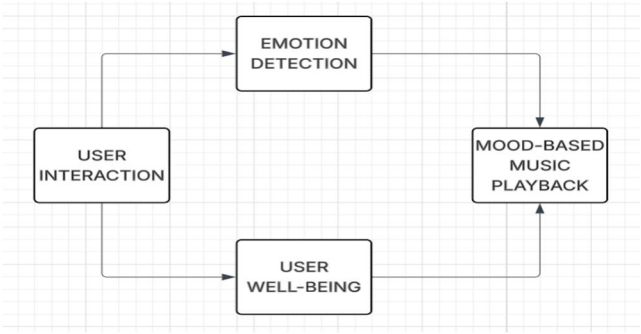
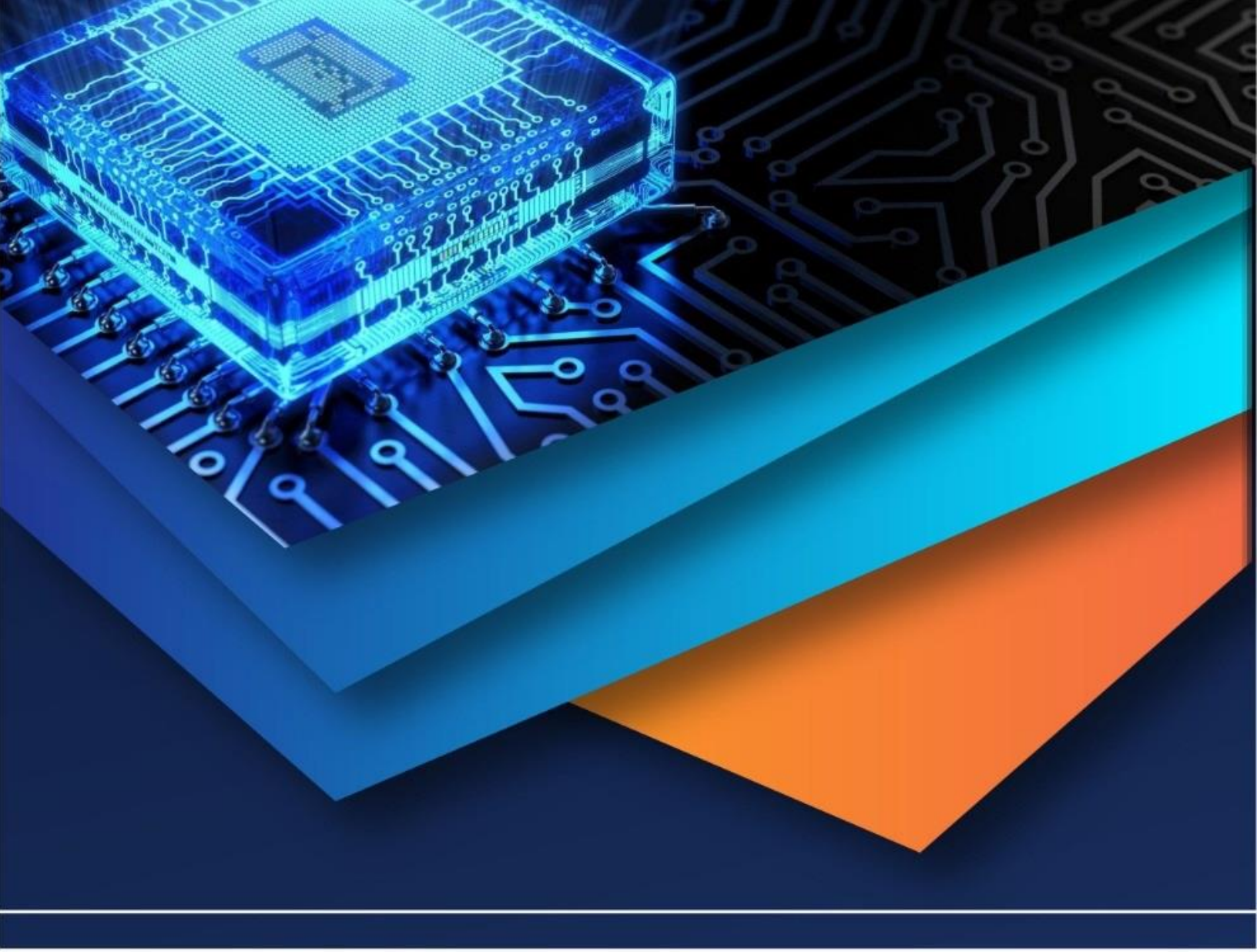


Fig 5. Overall Impact of Mood Detetction. The image represents the relation between the propsed system and benefits to the user

REFERENCES

[1] HTTPS://WWW.KAGGLE.COM/DATASETS/ANANTHU017/EMOTI ON-DETECTION-FER
[2] Jaiswal, A., Raju, A.K. and Deb, S., 2020, June. Facial emotion detection using deep learning. In 2020 international conference for emerging technology (INCET) (pp. 1-5). IEEE.
[3] Chutia, T. and Baruah, N., 2024. A review on emotion detection by using deep learning techniques. Artificial Intelligence Review, 57(8), p.203.
[4] Pereira, R., Mendes, C., Ribeiro, J., Ribeiro, R., Miragaia, R., Rodrigues, N., Costa, N. and Pereira, A., 2024. Systematic review of emotion detection with computer vision and deep learning. Sensors, 24(11), p.3484.
[5] Halkiopoulos, C., Gkintoni, E., Aroutzidis, A. and Antonopoulou, H., 2025. Advances in neuroimaging and deep learning for emotion detection: A systematic review of cognitive neuroscience and algorithmic innovations. Diagnostics, 15(4), p.456.
[6] Gaddam, D.K.R., Ansari, M.D., Vuppala, S., Gunjan, V.K. and Sati, M.M., 2021, November. Human facial emotion detection using deep learning. In ICDSMLA 2020: Proceedings of the 2nd International Conference on Data Science, Machine Learning and Applications (pp. 1417-1427). Singapore: Springer Singapore.
[7] Chen, Y. and He, J., 2022. Deep learning-based emotion detection. Journal of Computer and Communications, 10(02), pp.57-71.
[8] Liao, Y.J., Wang, W.C., Ruan, S.J., Lee, Y.H. and Chen, S.C., 2022. A music playback algorithm based on residual-inception blocks for music emotion classification and physiological information. Sensors, 22(3), p.777.
[9] Devarajan, D. and Rajkumar, K., 2024, May. EmoSync: Facial Emotion Detection for Adaptive Music Playback. In 2024 3rd International Conference on Artificial Intelligence For Internet of Things (AIIoT) (pp. 1-6). IEEE.
[10] Bakariya, B., Singh, A., Singh, H., Raju, P., Rajpoot, R. and Mohbey, K.K., 2024. Facial emotion recognition and music recommendation system using CNN-based deep learning techniques. Evolving Systems, 15(2), pp.641-658.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY