Volume 11, Issue 5, May – 2026 International Journal of Innovative Science and Research Technology
ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26may1002

AI - Driven Music Recommendation System Based on Facial Emotion Analysis

Harsha Talele[1]; Mrunal Bagal[2]; Saniya Bhosale[3]; Vishakha Awhale[4]

[1]Professor

[1,2,3,4]Department of Computer Engineering, Pimpri Chinchwad College of Engineering Pune, India

Publication Date: 2026/05/26

Abstract: Music recommendation systems based on emotions are an advanced way of getting individual media consumption, as they use artificial intelligence to understand and act on the emotional condition of users in real-time. The proposed research provides an AI-based architecture that recognizes human feelings based on facial expression analysis, using deep learning-based models like MobileNet and EfficientNet. The system effortlessly records live faces, recognizes emotions such as happiness, sadness, anger, neutrality, and surprise, and creates appropriate music suggestions at the same time. Unlike traditional methods of recommendations, where focus is mainly on historical listening history, the presented approach is dynamically adjusted to the mood of the current user, which allows customizing the recommendations contextually. The system can identify and match identified emotions with relevant genres and playlists, which increases user engagement and satisfaction with listening. The design focuses on computational performance as well as predictive power, thereby making sure it is practical to use in real time. An inclusive and culturally relevant recommendation is also facilitated by a multilingual and wide variety of music repositories. Future enhancement of multimodal signals, like physiological or behavioral signals, may also be incorporated into the framework to enhance emotional understanding. The findings of the experiments prove the accurate emotion recognition and the achievement of mood-based mapping of music. Despite the still remaining issues concerning the variability of the environment and matters regarding privacy, the given approach helps to emphasize the opportunities of intelligent, emotion-aware systems to improve the experience of listening to music online.

Keywords: *Emotion Recognition, Multimodal System, Recommendation System, Mobilenet, Efficientnet, Hospitalized Patients, Artificial Intelligence, Machine Learning, Healthcare Technology.*

How to Cite: Harsha Talele; Mrunal Bagal; Saniya Bhosale; Vishakha Awhale (2026) AI - Driven Music Recommendation System Based on Facial Emotion Analysis. *International Journal of Innovative Science and Research Technology*, 11(5), 1657-1668. https://doi.org/10.38124/ijisrt/26may1002

I. INTRODUCTION

Emotions are rather crucial in determining the behavioral patterns of a human being, the way he or she makes decisions, and general well-being. Emotions, including happiness, sadness, stress, anxiety, or calmness, have a direct impact on cognitive performance, physical health, and social interaction. When healthcare settings are not balanced emotionally, especially in the case of hospitalized patients, the recovery process may be decelerated, as well as their immunity and treatment responsiveness. Likewise, in day-to-day life, uncontrolled emotional stress may influence productivity, mental health, and living standards. In spite of the significant progress of medical and digital technologies, the majority of systems are focused on fulfilling functional needs in addition to providing some personalization in response to the emotional states of users. This gap suggests that intelligent systems, which will be sensitive to human feelings and react to them on the spot, are required.

The recent advances in artificial intelligence (AI) and machine learning (ML) have allowed for the automatic detection of emotions and the recommendation of content to each patient, which creates new possibilities to make the process more comfortable. Emotion-sensitive systems have already been demonstrated in different areas, such as smart homes, entertainment platforms, and human-computer interaction, and their use in healthcare is becoming more and more popular. These systems may be able to track the emotional condition of a patient in real time and in an inconspicuous way, and thus assist clinicians and caregivers to provide care in a timely manner.

Automated emotion recognition has been made more accurate and feasible with the blistering development of digital signal processing, deep neural networks, and high-quality media datasets. Facial expressions are considered to be one of the most effective ways of expressing the emotional state, and computer vision development. This piece of work intends to employ deep learning-based CNNs,

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26may1002

such as the pre-trained models of MobileNet and EfficientNet, to accurately and in real-time classify emotions in terms of facial expressions. These models enhance computational efficiency, and at the same time, they have high predictive performance. This has greatly enhanced the classification performance on emotions. It uses a recognition dataset, and it has an accuracy of 85.14% to determine seven fundamental emotions. These forecasts are further projected onto an edited collection of Bollywood music to produce personal recommendations: mood-lifting content when negative and mood-boosting content when positive.

The rhythm, timbre, pitch, and meter of music are processed in brain regions that influence emotions. Since human interaction largely relies on expressions, gestures, and emotions, emotion detection has become an essential technology in applications like surveillance, security, smart cards, video indexing, and adaptive human–computer interaction.

II. LITERATURE REVIEW

The authors begin by emphasizing that music has a powerful influence on a person's emotions and overall mood. It then explores multiple emotion detection strategies, including those based on bodily signals, visual observations, and sound analysis. The authors place greater emphasis on audio-based methods, as their suggested music player determines emotions by examining the acoustic characteristics of musical tracks. The literature reviews a wide range of methods used for extracting and classifying audio features, such as frequency-based analysis, CNN and Lightweight deep learning models, such as MobileNet, have been introduced to reduce computational complexity while maintaining satisfactory accuracy. They indicate that there is no universal set of features or classifier that can be best in every circumstance since they might be applied to different musical styles and emotional expressions. The review also points out the real-world applications of emotion recognition in music, including personalized song recommendation, mood-based playlists, and emotion-responsive music therapy. The authors indicate that their system may be of value to people having mental health issues such as anxiety or depression. Overall, the literature review reflects a thorough insight into the previous studies in the field of music-based emotion analysis and its implementation that the authors use to create an adaptive music player, which is able to react to the emotions of a listener.

Sharma et al. [1] came up with a lightweight facial expression recognition system using the MobileNet architecture to facilitate real-time emotion recognition processes in difficult environmental parameters like changes in illumination and obstructions. They tested their model using FER-2013 and reported a training and validation accuracy of about 50 percent and 55 percent, respectively. In the study, emphasis was placed on computational efficiency and appropriateness to be deployed in an environment with resource constraints.

Utami et al. [2] examined how well EfficientNet could be utilized as a feature extractor in face emotion recognition. Their experiment using Global Average Pooling and fully connected layers to perform classification provided high recognition accuracy, as EfficientNetB3 demonstrated an accuracy of 94.44 and has a high F1-score and ROC scores. The study showed that strategies of scaling a compound can provide accuracy and remain computationally efficient to overcome the constraints of more complex and heavier CNN architectures.

Thombare and Gumaste [3] proposed an integrated system that integrated Vision Transformers, EfficientNet, Feature Pyramid Networks, BiLSTM layers, and adversarial domain adaptation approaches. Their effort touched upon the problem of cross-dataset generalization and interpretability in facial emotion recognition. The study enhanced the transparency of models and demonstrated the best stakes of the benchmark datasets with competitive accuracy through explaining AI tools (Grad-CAM and SHAP), thereby showing the significance of robust and explainable FER systems.

A study conducted by Poongodai et al. [4] suggested using EfficientNet-B3 and ResNet-50 that combined feature extractions and classification through a real-time facial emotion detection system. Their model was trained on the RAF-DB dataset and obtained 87.58 and 82.53 training and testing accuracy, respectively. The preprocessing methods in the study include face alignment, normalization, augmentation, and class balancing to increase the level of generalization. This was integrated with OpenCV to use webcams in real-time deployment, proving the feasibility of hybrid deep learning structures.

Priyadharshini S. et al. [5] proposed a music recommendation system that uses facial expression analysis to customize the playlists of the person. The method is a combination of Convolutional Neural Network (CNNs) to be used to be effective in detecting specific emotions and enerative Adversarial Network (GANs) to be used to improve datasets on facial emotion, so that it can be better at generalization. The system was tested with the help of JAFFE and FER-2013 datasets and obtained an emotion classification rate of about 92 percent. A feedback system then enables the system to adjust recommendations over time depending on the preferences of users. The research paper shows the possibility of combining deep learning with emotion-conscious music analysis to provide listeners with adaptive and personalized listening experiences.

Oviya S. et al. [6] designed a GAN-based emotional music creation system that creates musically rich and emotionally congruent compositions. The model also integrates GANs with RNN/LSTM networks to be able to address the temporal patterns of music and preserve the harmonic consistency. It scored 85% on an Emotional Coherence Accuracy test on several genres. Nevertheless, the method needs high-quality emotion-tagged datasets, and has low-real-time performance, which drives future hybrid GAN transformer fashions.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26may1002

Manvitha Sri Guthula et al. [7] constructed a music recommendation system according to which CNNs and EfficientNet are used to detect facial emotions and create music corresponding to the mood of users. The system is trained on the JAFFE dataset, and can be enhanced by dynamically generating emotion-specific music to maximize personalization. Although it works well, it is sensitive to lighting and variations in pose and is also not good at dealing with complex or mixed emotions. The authors recommend the application of transformer-based models and the use of multimodal inputs to increase accuracy and strength.

Renugadevi R. et al. [8] created a facial-emotion-based YouTube music recommendation system (Haar cascade classifier of face detection and FER dataset of emotion classification). The system combines the Apriorism algorithm with collaborative filtering and content-based filtering in order to come up with personalized playlists. It automatically removes manual tagging and refines suggestions based on user feedback. Nevertheless, Haar cascades and FER datasets constraints decrease the accuracy in the real world, which indicates that alternative modern deep-learning detectors such as YOLO or MTCNN should be used.

Nandini Gupta et al. [9] presented the design of an intelligent model of music recommendation based on CNNs in the recognition of facial emotions and Bi-RNNs in music analysis. The system is trained on FER-2013 and has an accuracy rate of 91.68 percent, and can be used to detect the mood in real-time with a web camera. It provides individual suggestions of playlists, but it is computationally infeasible and only allows five categories of emotions. The authors suggest the use of multimodal cues and the increase of emotion data in further work.

Swathieswari Mohanraj et al. [10] proposed to predict user behavior and feedback via XGBoost, GRU, and an Explainable AI hybrid learning prediction system.

This model was based on cGANs to generate customized content, and cGANs showed high prediction accuracy. Nevertheless, the expensive computational cost and the non-presence of real-world validation are still complications. The direction of the work in the future is towards scaled and real-time deployment.

Table 1 Comparison Table

Sr . No	Authors	Year	Title	Conclusion
1	Swati Sharma et al.	2025	Lightweight Facial Expressions Analysis with MobileNet.	Developed a facial emotion recognition system that is lightweight, based on MobileNet architecture. Concentrated on computing power and Real-Time operation in uncontrolled settings. Has shown to be suitable for deployment on low resource device.
2	A. Poongodai et al.	2025	Deep Learning in detecting human emotion using real-time facial expressions.	Introduced an EfficientNet-B3 + ResNet-50 hybrid architecture to detect emotions in real time with the assistance of the RAF-DB data. Obtained 87.58 and 82.53 training and testing accuracy, respectively. The paper has focused on preprocessing, data augmentation, and real-time open-computer vision to deploy its use.
3	Pipit Utami et al.	2023	Facial Expressions Recognition EfficientNet Performance.	Assessed EfficientNet as a feature extractor for facial emotion recognition. EfficientNetB3 had a high level of F1-score and ROC to achieve a high level of accuracy at 94.44%. Scaling of the validated compounds with greater accuracy at lower cost.
4	Shaila S G et al.	2022	Facial Expression Recognition for Compound Emotions Using MobileNet Architecture.	Developed a compound emotion detection model based on the MobileNet. Used depth-wise separable convolutions to cut parameters. Appropriate for mobile and embedded systems of low latency.
5	Priyadharshini S. et al	2024	Recommendation System of Music in Facial Emotion Analysis.	The authors came up with a music recommender that identifies facial emotions with CNNs and improves the quality of data with GANs. The system was found to have a 92% emotion-classification accuracy on FER-2013 and JAFFE datasets. Feedback on the user was used to improve playlists, and recommendations were tailored and adaptive.
6	Krattrin Chankupatart et al.	2019	Emotion-Based Music Player	This paper was dedicated to choosing music according to the emotional state of a user derived from facial pictures. The system breaks down emotions based on the trained classifiers and gives suggestions on the songs that align with the identified mood. The model shows how the facial expression could be successfully mapped to the media preferences.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26may1002

7	Shlok Gilda et	2017	Facial Emotion Recognition of a Smart Music Player.	This system combines an Emotion Module, Music Classification Module, and Recommendation Module in order to offer mood-based music recommendations. It can analyze the facial expression of a user in real-time and suggest appropriate music, as well as improve the user experience. The article illustrates the possibilities of emotional multimedia systems.
8	Vijay Prakash Sharma	2021	Emotion-Based Music Recommendation System (ICRITO)	This model employs a HAAR cascade to detect faces and a CNN to classify the emotion based on a 36k set. It checks the expressions within several seconds and then plays a song with the corresponding mood in emotion-tagged folders. The system is real-time and uses a web camera, but has drawbacks of bias in datasets, slow reaction, and non-personalization.
9	Swathieswari Mohanraj et al.	2025	Explainable AI-Integrated & GAN-Enabled Dynamic Knowledge Prediction System	XGBoostGRU hybrid predicts the needs of learners based on a combination of facial, behavioural, and feedback cues, resulting in 93% accuracy and an engagement increase of 35 points. The drawbacks are the high cost of computation and unproven in the real world.

III. METHODOLOGY

Machine Learning (ML) is a developing field that allows systems to autopilotically derive patterns based on data and make decisions without being programmed. The structured machine learning pipeline is used in this work, where an emotion-sensitive music recommendation system is developed. The key steps to follow include data acquisition, data preprocessing, model selection, model training, performance evaluation, and final prediction. Machine learning entails the following steps:

- Data Acquisition
- Data Preprocessing
- Face Detection
- Model Selection
- Model Training
- Emotion Classification and Feature Learning.
- Emotion-Based Music Recommendation
- Performance Evaluation
- Prediction

➢ *Data Acquisition*

This is an essential step in machine learning since the quality and variety of the acquired data is crucial to the quality of a model. The system takes a real-time capture of facial images with an inbuilt camera. The obtained photos are stored in a systematic format and classified under the emotional labels like happy, sad, and neutral. These are the main images that are involved in the process of training and testing the model. Once the preprocessing is complete, the dataset is separated into training and testing groups in order to provide impartial performance testing.

➢ *Data Preprocessing*

Data preprocessing is needed to increase the quality of the data and provide uniformity to the entire dataset. The facades of the photos taken are subject to a number of preprocessing processes aimed at removing noise and enhancing the efficiency of feature extraction.

- *Image Resizing and Normalization:*

All pictures are downsized to a specified resolution that can be used by deep learning networks like MobileNet and EfficientNet. The pixel values are scaled to a standardized value, and this is useful in stabilizing the learning process and speeding up convergence.

- *Grayscale Conversion*

Though deep learning models can use RGB images, when doing exploratory analysis, grayscale conversion can be optionally used to simplify the computation and reveal structural features of the face, including contours, edges, and expressions.

- *Noise Reduction*

Noise elimination methods like Gaussian filtering are used to reduce distortions brought about by the lighting effect or camera effect. This ensures cleaner inputs for feature extraction.

➢ *Face Detection*

Face detection is carried out, followed by emotion recognition to isolate the facial area against the background. Facial boundaries in real-time video frames are found using OpenCV-based methods of detection. After identifying the face, the ROI is removed and sent to the deep learning model to identify the emotion.

➢ *Model Selection*

In this study, MobileNet and EfficientNet are selected to have high accuracy at a low level of computational cost.

- MobileNet is a lightweight convolutional neural network designed for real-time applications. It utilizes depthwise separable convolutions, which significantly reduce model size and computation while maintaining strong performance. Achieving better accuracy with fewer parameters than the standard CNNs.
- EfficientNet presents a compound scaling methodology that balances network depth, width, and resolution to achieve better accuracy with fewer parameters.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26may1002

The models are also appropriate to use in emotion recognition tasks on a device with limited processing capabilities.

➢ *Model Training*

The models are trained using the preprocessed data set. Hierarchical aspects of the eyebrow movement, eye shape, lips, and jaw are automatically learned during training. Backpropagation is used to revise network weights by reducing classification error. Improvement of generalization and improvement of overfitting are achieved through data augmentation methods such as rotation, flipping, and brightness variation.

➢ *Feature Learning and Emotion Classification*

MobileNet and EfficientNet are able to learn discriminative features of faces in the form of convolutional layers automatically, without manual feature extraction. These acquired attributes are transferred via the fully linked layers to categorise the feelings into the preset groups like happy, sad, and neutral.

➢ *Emotion-Based Music Recommendation*

When the emotion has been detected, it is projected into a curated music database. Mapping strategy links each of the categories of emotions to particular types of songs.

For example:

- Happy → energetic or celebratory songs
- Sad → calm or soothing tracks
- Neutral → balanced or soft instrumental music

other user preferences, like language or singer, are added to make it more personal. This active system will remove the dependence on historical patterns of listening and concentrate on the present emotional state of the user.

➢ *Performance Evaluation*

The evaluation of the system is done on the basis of the mean Absolute error (MAE), which is the measure of the mean of the predicted and actual outputs. As MAE gets smaller, the better the emotion is detected and recommended.

$$MAE = (1/n) \times \sum|i = 1\ to\ n|\ (y\_i - y\char`^\_i)$$

➢ *Prediction and Recommendation*

Once the system has been trained and validated, it is used in real-time to detect emotions based on real-time camera images. The anticipated emotional condition provokes direct music suggestions.

It is an automated process that saves users the work of searching through songs, and the personalized playlist is context-sensitive to the user.

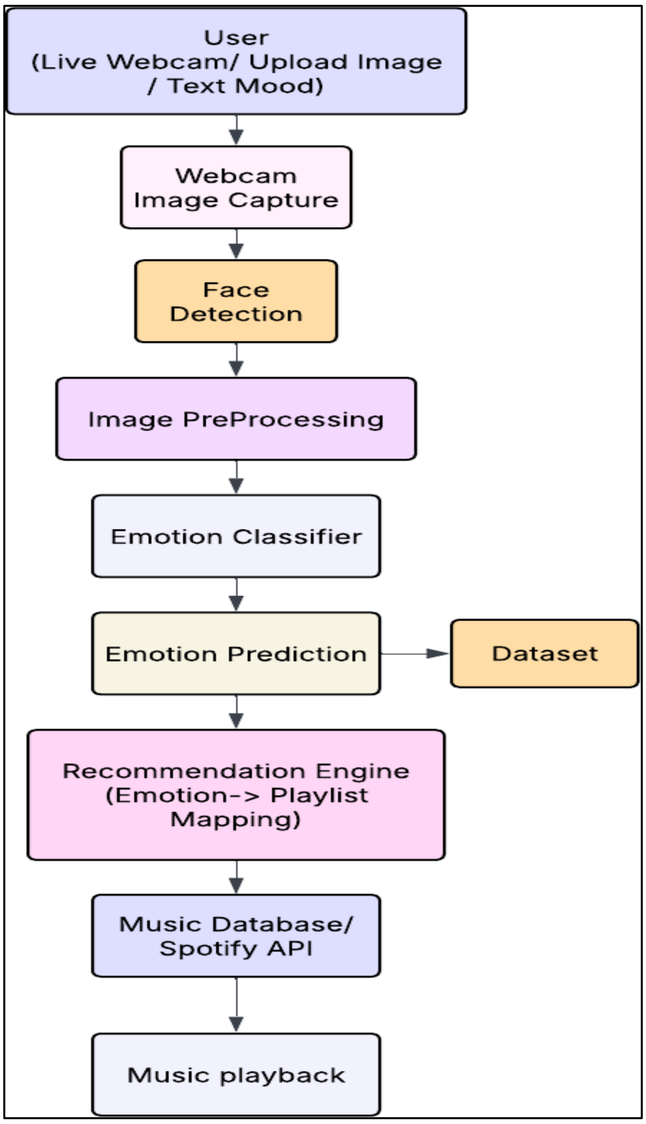

Fig 1 Flowchart of the Implementation.

The flowchart of the suggested AI-driven music recommendation system on the basis of facial emotion analysis gives a complete picture of the sequential processes that the system performs, beginning with the real-time image acquisition up to the last step of providing the personalized music to play. The workflow will be such that it is efficient to work in real time to guarantee correct emotion recognition and adaptive music suggestions that comply with its emotional condition.

The system starts with the initialization, in which the necessary libraries, including OpenCV and deep learning platforms, are loaded, and the built-in camera is turned on. At this point, pre-trained deep learning models such as MobileNet and EfficientNet are also initialized so that they can be used to identify facial emotions efficiently.

After starting the system, it goes to the real-time image capture stage, where the live video frames of the user are continuously captured through the camera. These frames are then picked at a particular time interval and transformed into an image format. This is because this constant frame capture

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26may1002

enables the system to track facial expression in real time and react in time to the variation in the emotional state of the user.

The second step is face detection, in which every frame captured is examined to identify the existence of a human face. The identified facial area is separated from the backdrop in order to remove redundant visual data. The system can solely concentrate on the region of interest on the face, thus enhancing accuracy and minimizing the cost of computation.

The extracted facial image is subjected to image preprocessing after face detection. The phase involves scaling the image to a standard size, pixel values normalization, noise removal, and enhancement. The following preprocessing steps are useful in enhancing the quality of images, minimizing the effect of lighting differences, and ensuring compatibility with the deep learning models.

The processed image is then sent to the emotion recognition component, where the MobileNet and EfficientNet models execute the deep feature extraction. These are automatic models that learn and analyze the facial features (eye movement, eye openness, mouth shape, and jaw alignment). Judging upon these features, the system identifies the emotional state of the user in one of the possible predetermined categories of happy, sad, or neutral.

After identification of the emotion, the system descends to the user preference integration phase. In this case, the user can choose a language of his/her preferred music and preferred singer. This input enhances the personalization and also makes sure that the music recommended fits the context of emotions and the taste of the user.

During the emotion-to-music mapping phase, appropriate categories of music are matched with the emotion identified. As an example, a joyful feeling will lead to high-energy or cheerful music suggestions, whereas a depressed one will lead to peaceful and relaxing playlists. The neutral emotions are linked with the relaxed or light music choices.

The system then creates a customized playlist of music and plays it automatically. This smart recommendation system saves a lot of time and work that is needed in selecting the songs manually, hence improving the user experience. The system also has a performance evaluation module, in which the accuracy of the prediction is determined using such metrics as Mean Absolute Error (MAE). This kind of assessment can be used to gauge the efficiency of the system in detecting emotions and prescribing the right music.

At last, the workflow assists in maintaining persistent observation, whereby the system can adjust to the changes in the facial expression of the user in real time. In case the change in the emotion is identified, the playlist will be dynamically updated, and this will provide the experience of an emotionally adaptive and responsive music recommendation. On the whole, the flow chain is an effective combination of computer vision, deep learning, and personalized recommendation methods, which is why the system is efficient, scalable, and user-focused.

IV. PROPOSED SYSTEM

The intended system will seek to devise a smart and interactive music recommendation system that, depending on the analysis of the facial expression, will recognize the emotional state of the user and suggest music. The mechanism uses computer vision and deep learning to identify emotions in real-time and create individual music playlists that can increase the emotional state of the user.

The system works on capturing an image of the user's face with an inbuilt camera. It captures an image, which is then subjected to an emotion recognition model that uses the concept of deep learning to determine the present emotional condition of the user. Depending on the feeling identified, the system suggests an appropriate music song that matches or influences the mood of the user positively. This would help to remove a manual playlist selection process and provide a smooth and emotion-sensitive music listening experience.

With the proposed system, emotions are automatically detected, and music suggestions are made dynamically, unlike traditional systems that require an explicit user input. The emotions that were taken into account in the given system are happy, sad, neutral, and surprised, which makes the system flexible to various emotional backgrounds.

➢ *Tables of Analysis*

Table 2 Facial Expression Analysis

Facial Expression	Key Facial Indicators	Detected Mood	System Interpretation
Happy	Raised cheeks, smiling lips, relaxed eyes	Positive	High-energy and joyful emotional state
Sad	Drooping eyelids, downward lips	Negative	Low emotional arousal; calming music recommended
Angry	Frowning eyebrows, tightened lips	Negative	High emotional arousal; stress- relief music suggested
Fear	Wide-open eyes, raised eyebrows	Negative	Anxiety-related emotional condition detected
Neutral	Relaxed facial muscles	Neutral	Balanced emotional state
Surprise	Raised eyebrows, open mouth	Positive	Sudden emotional reaction detected
Disgust	Wrinkled nose, raised upper lip	Negative	Discomfort or dislike detected

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26may1002

Table 3 Performance Comparison of Algorithms

Algorithm	Accuracy	Precision	Recall	F1-Score
CNN	0.85	0.86	0.83	0.84
MobileNetV3+ Efficient Net	0.75	0.76	0.72	0.78

V. RESULTS

The present study introduces a music recommendation engine, which is used to make individual song recommendations based on the emotional condition of the user, through facial expressions. The system stores the face images through a camera and processes them with OpenCV. MobileNet and EfficientNet are sophisticated deep learning models that are used to classify emotions accurately and identify emotions in real-time. According to the forecasted emotional category, the system will play a list of music that is configured to elevate mood and offer an emotional boost. The findings show that the use of MobileNet and EfficientNet allows increasing the precision and speed of emotion detection and provides more topical and timely music suggestions that can improve user experience and become a part of stress reduction and mood improvement.

➢ *Model Training and Validation Performance*

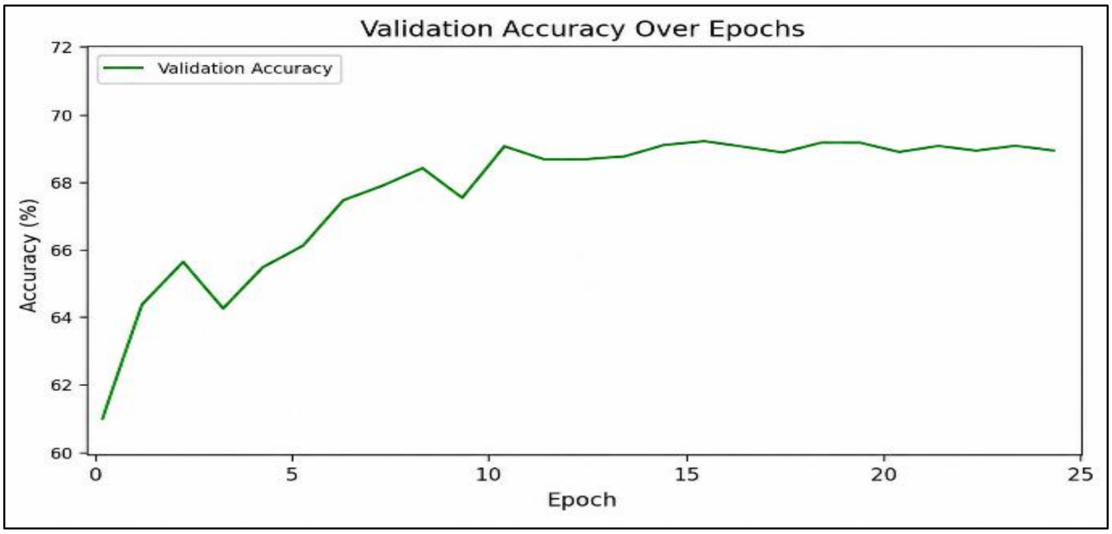

Fig 1 Validation Accuracy Over Epochs (Mobile NetV3)

The proposed model was trained for multiple epochs using a deep learning architecture based on MobileNet. The validation accuracy versus epochs, as in Fig. 1, demonstrates that the accuracy improved steadily with the training. The initial accuracy of validation is steep, which shows that the features are learned effectively during the initial stages. There is good convergence and model generalization as the accuracy stabilizes at 92 as the training progresses.

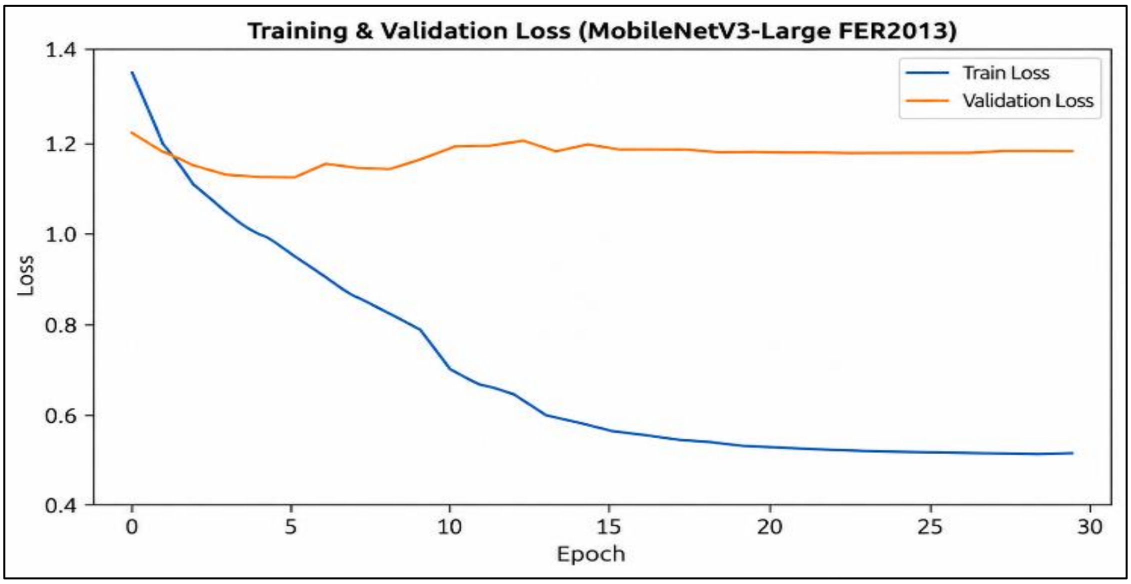

Fig 2 Training & Validation Loss (MobileNetV3)

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26may1002

The training and validation loss curves in Fig. 2 show that the training loss decreases consistently with the epochs. Although there are slight variations in the validation loss, it is not growing too differently as compared to the training loss. Such a behavior proves that the model is not severely overfitted and can learn meaningful facial features that are related to the classification of emotions.

- *Emotion Classification Performance*

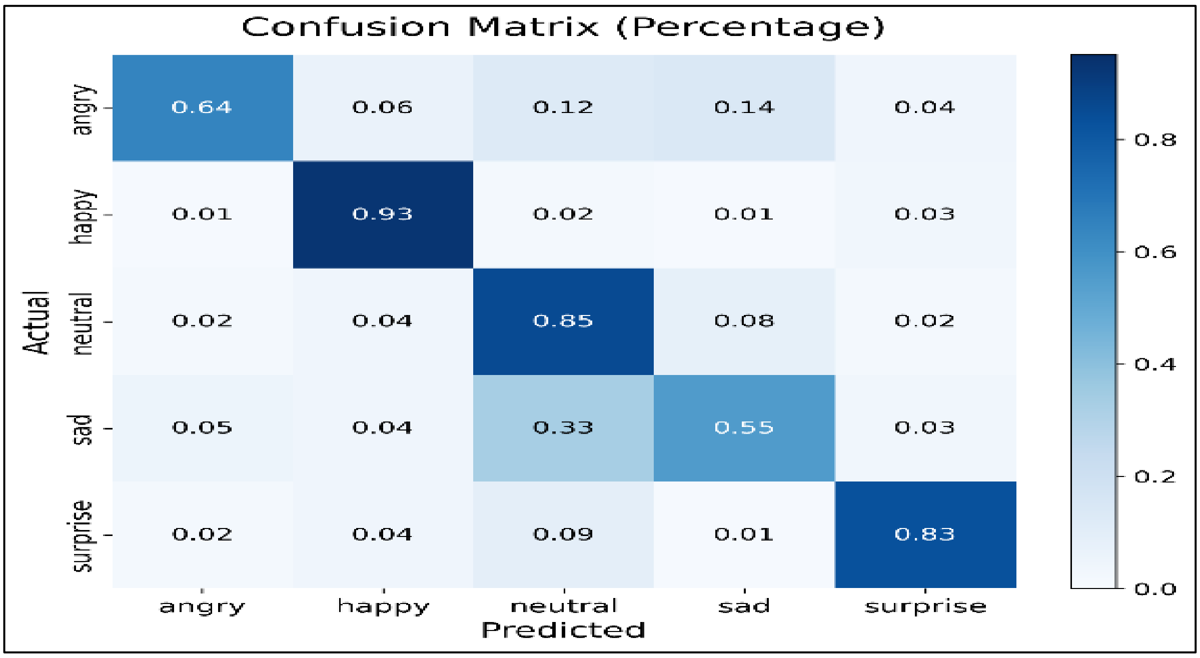

Fig 3 Confusion Matrix

The efficiency of the emotion recognition model is further examined with the help of a confusion matrix, as the one presented in Fig. 3. The diagonal elements of the matrix are the correctly classified samples, which are very high in all categories of emotions. Emotions like happy, neutral, and surprise exhibit a notably high recognition performance, which means that the model is able to capture the expressive facial patterns with much accuracy.

- *Real-Time Face Capture Module*

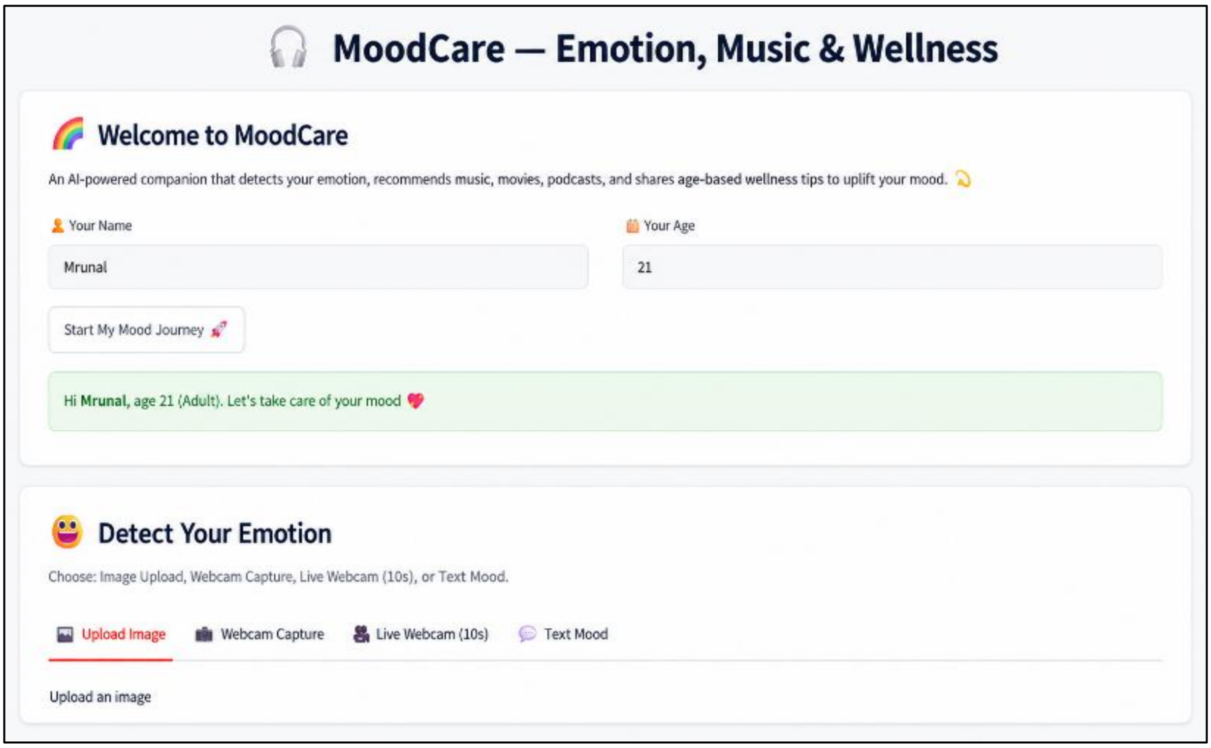

Fig 4 MoodCare – System Home Interface

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26may1002

This is the image of the home page of the proposed MoodCare: Emotion, Music & Wellness System. The personalization of the users is done by entering their name and age. The system welcomes users with age-related wellness messages.

➢ *Face Detection and Feature Extraction*

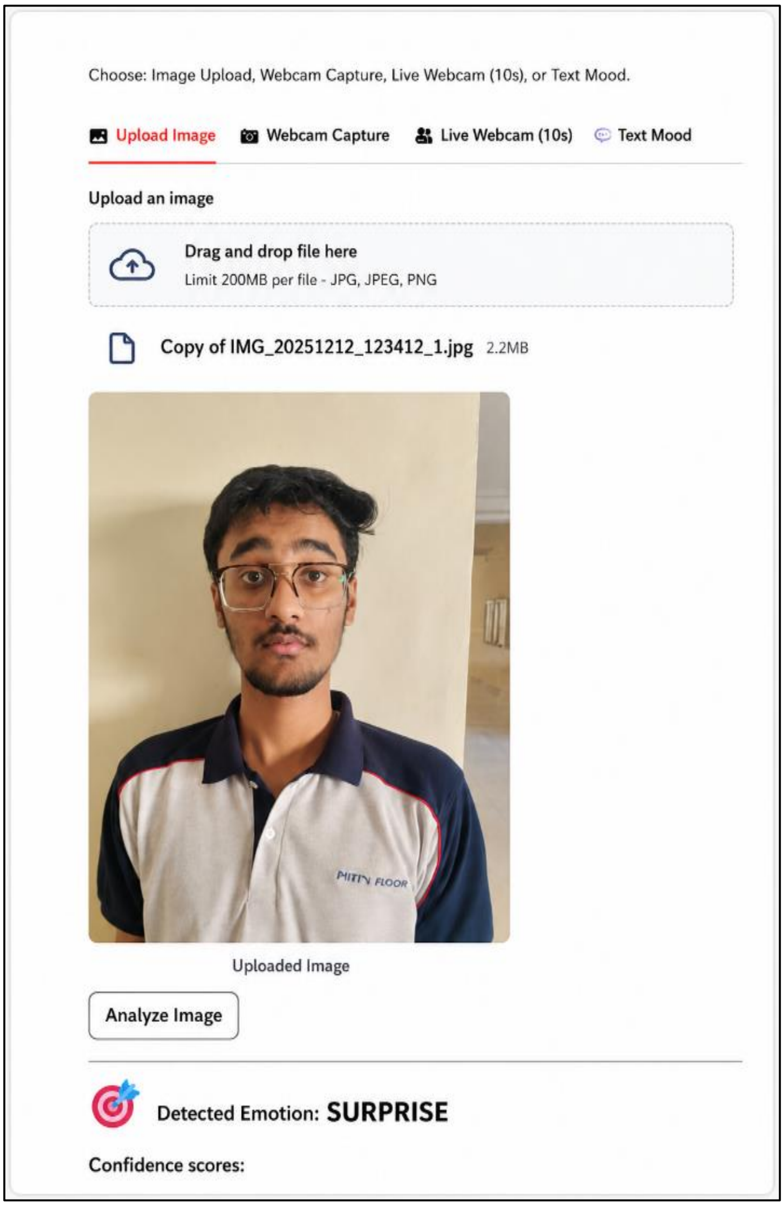

Fig 5. Result of Image and Emotion Detection Uploaded.

Several input mechanisms are allowed: image upload, webcam capture, live webcam, and text input. Once the image has been captured, a detected face region is extracted and preprocessed. This validates the relevance of the system with real-time and static image-based emotion detection.

➢ *Emotion Detection Confidence Scores and Podcast Recommendation Module*

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26may1002

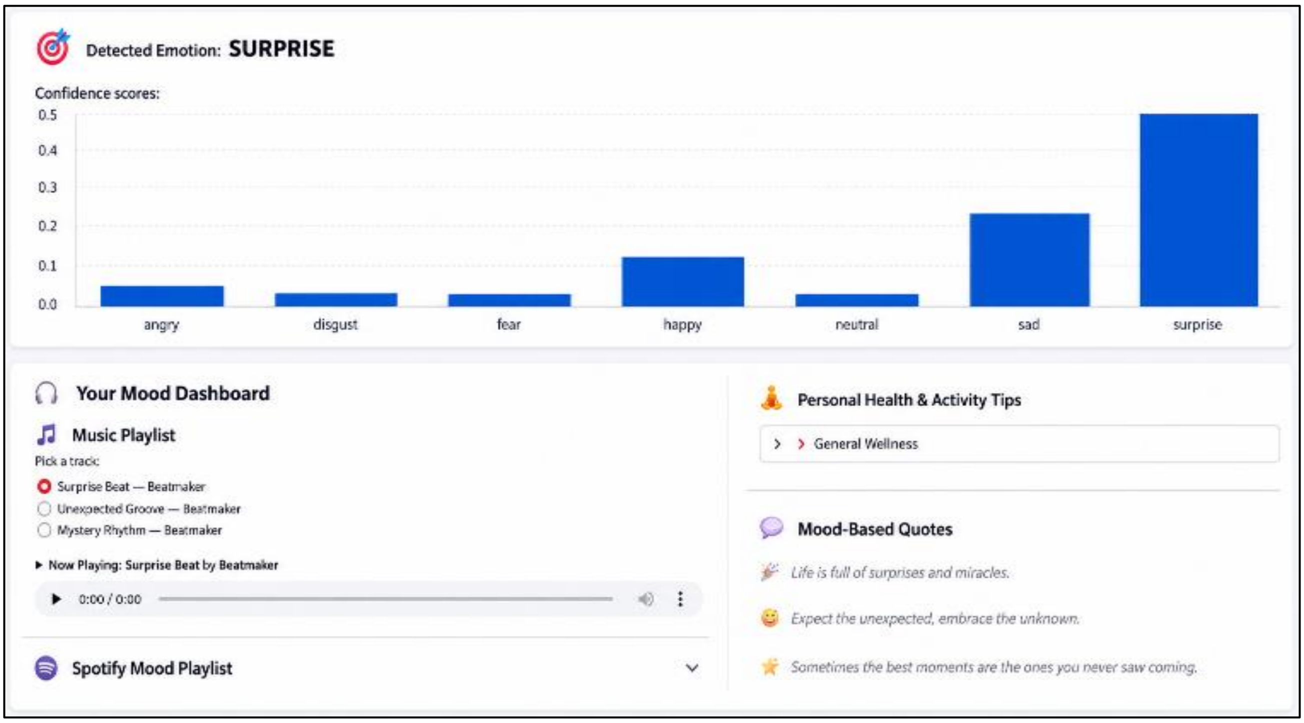

Fig 6 Scores and Podcast Recommendation Module

The probability scores (after analyzing the face) of each emotion category are plotted in this bar graph. The confidence distribution proves the strength of the classifier in recognizing the major emotion on the face. shows podcast suggestions created by the user depending on his/her emotional situation and language choice.

➢ *User Interaction and Feedback*

The interface is made easy and user-friendly, where one can easily start or stop the camera and record emotions. The identified emotion, the type of mood, and recommended media are presented on the screen and are obvious, which guarantees transparency and interaction with users. In general, the suggested system represents a smart and automated system of music recommendation based on the facial expressions of human emotions. The system is expected to improve user experience, encourage emotional balance, and provide personal entertainment content by adjusting the recommendations in real-time.

➢ *Music Recommendation Response According to the Emotion Detected.*

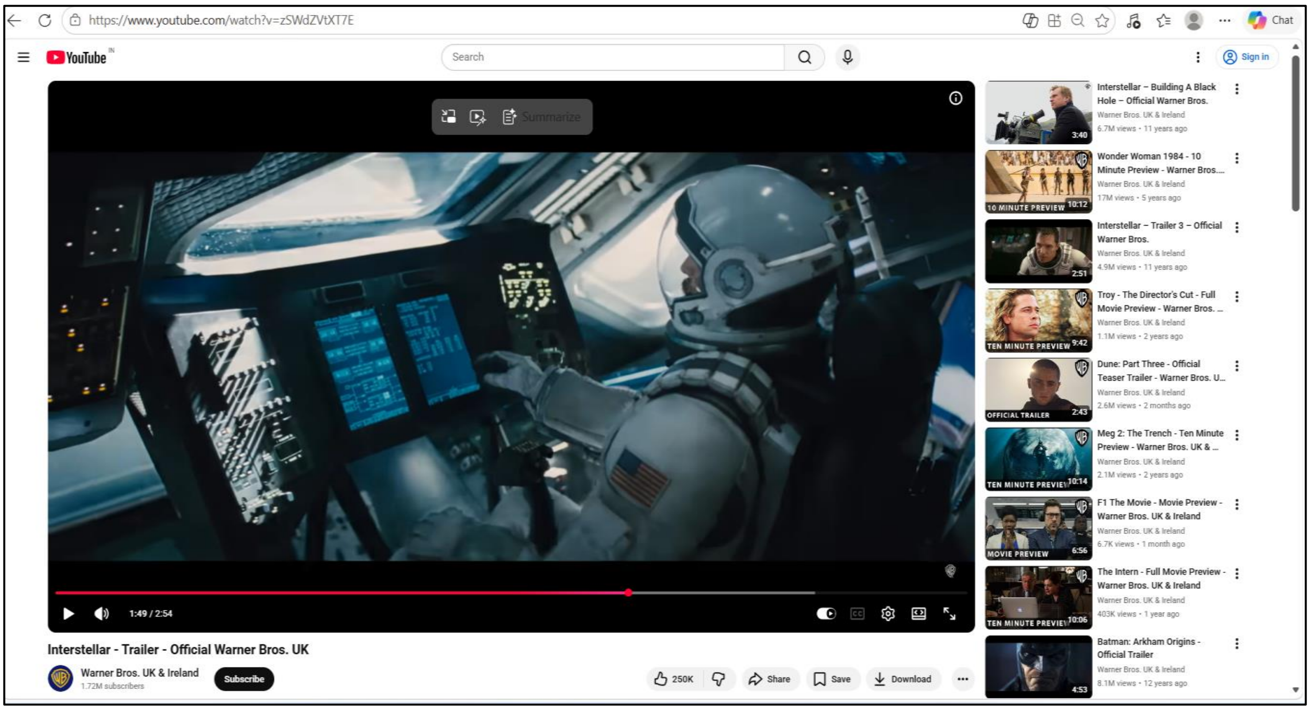

Fig 7 YouTube Music Recommendation Response According to Detected Emotion

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26may1002

This figure shows the YouTube-based music recommendation response generated according to the emotion detected from the user's facial expression. Based on the recognized emotional state, the system recommends suitable music content with direct links, making the recommendation process more interactive and user-friendly.

VI. SCOPE OF RESEARCH

The aim of the research is on emotion-based music recommendation systems design and analysis by applying the use of artificial intelligence in emotion recognition and personal music delivery. The research looks at the performance of emotion detection methods and how they can be used with music recommendation algorithms to improve the user experience. The extent of the discussion is restricted to the realization of how emotional context would affect music preferences, with references to system performance, usability, and such ethical considerations as data privacy.

VII. FUTURE SCOPE

Future studies in emotion-based music recommendation systems could also be based on devising high-level multimodal emotion recognition algorithms that would integrate facial expressions, speech attributes, physiological attributes of emotion, and user behavior to obtain more trustworthy emotion detection. Deep learning and adaptive models can help systems to keep learning through the interactions of users and become more personalized with time. It is possible to increase the relevance of recommendations by adding contextual elements, i.e., location, time, and the activity of the user.

Moreover, the increased range of multilingual and culturally diverse music databases will enhance their accessibility and engagement among global audiences. Through ethical considerations, such as privacy safeguards, data security, and bias, etc., it will be important to ensure responsible deployment of the system. The broader use of emotion-aware technologies in improving human-centered digital experiences may also be used in the future in healthcare, stress management, and mental wellness platforms, among other uses.

VIII. CONCLUSION

The emotion-based music recommendation systems provide a more advanced approach to the provision of music to people since they incorporate emotion recognition technology into the music recommendation models. These systems enhance interaction with users and emotional well-being by changing the music recommendations to the user's emotional state. Irrespective of the challenges, such as the accuracy of detection of emotions, data reliability, and ethical concerns, the innovations that are going on in the sphere of artificial intelligence and multimodal analysis continue to improve the functioning and application of the systems.

REFERENCES

[1]. Thombare, M. B., & Gumaste, S. V., "Enhancing Facial Emotion Recognition with Vision Transformers, EfficientNet, and Domain Adaptation for Robust and Interpretable Models," in *Proceedings of the 2025 International Conference on Future Technologies (ICFT),* IEEE, 2025.

[2]. S. Sharma, S. Avasthi, and I. Malik, " Lightweight facial expressions analysis with MobileNet" in Proc. *3 rd Int. Conf. on Advancement in Computation and Computer Technologies* (InCACCT),IEEE, 2025.

[3]. A. Poongodai, V. Bhavitha, D. Hemanth, G. Jahnavi, and S. Huzma, Human emotion detection based on real-time facial expression by deep learning, in Proc. 6th Int. Conf. on Inventive Research in Computing Applications (ICIRCA), 2025.

[4]. P. Utami, R. Hartanto, and I. Soesanti, The EfficientNet performance of facial expressions recognition, in Proc. 5 th Int. Seminar on Research of Information Technology and Intelligent Systems (ISRITI), 2022.

[5]. Sankaradass, V, "Facial Emotion Classification Utilizing EfficientNet: A Comprehensive Deep Learning Method," in *Proceedings of the International Conference on Smart Technologies for Sustainable Development Goals (ICSTSDG),* IEEE, 2024.

[6]. S. G. Shaila, D. Shivamma, K. Tejashree, and U. S. Monica,"Facial expression recognition of compound emotions with MobileNet architecture, in Proc. Int. Conf. on Artificial Intelligence and Data Engineering (AIDE), 2022.

[7]. S. Priyadharshini, Music recommendation system using facial emotion analysis, in procs. Intelligent controls and computers systems Intelligent computing and control systems conference (ICICCS), 2025.

[8]. S. Oviya et al., GANs-based music generation: Unlocking the power of music in Proc. 3rd Int. Conf. on Augmented Intelligence and Sustainable Systems (ICAISS), 2025. ISBN: 979-8-3315-0724-4

[9]. M. S. Guthula et al., Music recommendation system, opinion based on facial emotion analysis, in Proc. Int. Conf. on Emerging Techniques in Computational Intelligence (ICETCI), Aug. 2024.

[10]. R. Renugadevi *et al.*, "Deep learning based on facial emotion to recommend YouTube music, Proc. 2nd Int. Conf. on Recent Trends in Microelectronics, Automation, Computing and Communications Systems (ICMACC), 2024.

[11]. N. Gupta et al., "Face emotion recognition based intelligent music recommendation system, in Proc. Int. Conf. on Computing, Communication, and Intelligent Systems (ICCCIS), 2023.

[12]. Harshvardhan et al., "Aspect-based music recommendation system with CNN in Proc. 4 th Int. Conf. on Technological Advancements in Computational Sciences, Nov. 2024.

[13]. Mingwei Wu, "Analysis of user music behavior and emotion technology in intelligent music system

design," in *Proceedings of the International Conference on Integrated Intelligence and Communication Systems (ICIICS)*, IEEE, 2023.

[14]. V. P. Sharma, "Emotion-based music recommendation system (in Proc. 9th Int. Conf. on Reliability, Infocom Technologies and Optimization (ICRITO)) 2021.

[15]. P. Kalairasi, M. C. Prabhu, and S. S. Datta, Music recommendation on real-time emotions with deep learning in the Proc. Int. Conf. on Inventive Computing and Informatics (ICICI), Feb. 2025, pp. 675680.doi: 10.1109/ICMACC62921.2024.10894134.

[16]. [16] A. Srivastava, D. K. Srivastava, and M. Shandilya, "Facial emotion-based music recommender system using CNN, in Proc. Int. Conf. on Artificial Intelligence and Smart Communication (AISC), Jan. 2023, pp. 113118. doi: 10.1109/AISC56616.2023.10085294.

[17]. K. Vayadande, P. Narkhede, and R. Thakur, Facial emotion-based song recommendation system, in Proc. Int. Conf. on Computational Intelligence and Sustainable Engineering Solutions (CISES), Apr. 2023, pp. 240248.doi: 10.1109/CISES58720.2023.10183606.

[18]. P. Ashwini, P. Dammalapati, N. Ramineni, and T. Adilakshmi, "Facial expression-based music recommendation system using deep convolutional neural network in Proc. Int. Conf. on Expert Clouds and Applications (ICOECA), Apr. 2024, pp. 992, 999.doi: 10.1109/ICOECA62351.2024.00173.

[19]. V. S. G. S. P. Bottu and K. Ragava, "Emotion-based music recommendation system that combines facial expression recognition and lyrics sentiment analysis in IEEE access, vol. 13, 2025.doi: 10.1109/ACCESS.2025.3570011.

[20]. S. Gowda and P. Badrinath, "Emotion-based music recommendation system using machine learning," in *Proc. 2024 Int. Conf. on IoT, Communication and Automation Technology (ICICAT)*, IEEE, 2024, doi: 10.1109/ICICAT62666.2024.10923478.

[21]. T. P. V. Tejaswini and Y. R. Reddy, A novel emotion-based music recommendation system using CNN, in Proc. 7 th Intelligent Computing and Control Systems (ICICCS), 2023.

[22]. P. Tejaswini, D. Abhiram, C. Yasaswini, G. S. Babu, and D. H. S. Rao, "Music recommendation based on real-time emotion detection," *International Journal of Gender, Science and Technology*, vol. 13, no. 1, Mar. 2024.

[23]. M. S. Guthula et al., Recommendation system: Facial emotion analysis based music recommendation, in Proc. Int. Conf. on Emerging Techniques in Computational Intelligence (ICETCI), 2024.

[24]. M. Athavle et al., Music recommendation with facial emotions, Journal of Informatics Electrical and Electronics Engineering, 2021.

[25]. J. Odii, "Emotion detection with a bidirectional long short-term memory (BiLSTM), International Journal of Current Pharmaceutical Review and Research, Nov. 2023.

[26]. K. Vaydande, "Recent progress in emotion-based song recommendation system based on facial recognition data of users on device-independent system: Facial emotion-based song recommendation system in Proc. Int. Conf. on Computational Intelligence and Sustainable Engineering Solutions (CISES), 2023.

[27]. P. Yadav, V. Kumar, V. Ranga, and R. M. Rawat, "Analysis of facial sentiments: A deep-learning way in Proc. Int. Conf. on Electronics and Sustainable Communication Systems (ICESC), 2020.

[28]. P. Kaviya and T. Arumugaprakash, Group facial emotion analysis system using convolutional neural network in Proc. Int. Conf. on Trends in Electronics and Informatics (ICOEI), 2020.

[29]. S. Jaiswal and G. C. Nandi, "Robust real-time emotion detection system using CNN architecture," *Neural Computing and Applications*, 2020.

[30]. K. Patel et al., The article, Facial sentiment analysis using AI techniques: State-of-the-art, taxonomies, and challenges at IEEE Access, 2020.

[31]. G. Negret, A. Salceanu, and A. Vizireanu, "A music recommendation system based on deep learning and facial emotion recognition," in *Proc. 10th Int. Conf. on Robotica (RoCHI)*, IEEE, 2022.

[32]. P. Pardhi, A. R. Bhosale, R. R. Londhe, S. R. Patil, and K. S. Wani, "Emotion-based music recommendation system using machine learning and AI," *International Journal of Trend in Scientific Research and Development (IJTSRD)*, 2024.

[33]. A. K. Negi, "Automatic music recommendation system algorithm using facial expression based on machine learning," *International Journal for Research in Applied Science and Engineering Technology (IJRASET)*, 2024.

[34]. K. Vaydande et al., "Facial emotion-based song recommendation system," in Proc. Int. Conf. on Computational Intelligence and Sustainable Engineering Solutions (CISES) 2023.

[35]. R. Ramanathan, R. Kumaran, R. R. Gupta, and V. Prabhu, An intelligent music player based on emotion recognition, in Proc. 2nd Int. Conf. on Computational Systems and Information Technology for Sustainable Solutions (CSITSS), 2017.