

*Original Article*

# Emotion-Based Music Recommendation System

Pratibha Sajwan, Kunal Chopkar, Rajveer Baheti, Harshil Choudhary, and Harshita Mehta

Thakur College of Engineering & Technology,  
Mumbai, India

Corresponding Author Email: pratibha.sajwan@tctmumbai.in

Received: 17 November 2024 / Revised: 03 December 2024 / Accepted: 11 January 2025 / Published online: 30 January 2025

© The Author(s) 2025

**Abstract:** The rapid evolution of the internet and mobile technologies has revolutionized our access to a vast array of musical resources. However, a significant challenge persists in accurately discerning users' desired music based on their prevailing mood. Music possesses a profound capacity to mirror and resonate with emotions; we naturally incline toward cheerful melodies in times of elation, find solace in serene tunes during stress, and connect with melancholic songs when feeling low. To tackle this hurdle, we propose an inventive system that utilizes real-time user interactions, like conversations, to assess their emotional state. This emotional input forms the foundation for suggesting songs that match their sentiments. Our method entails categorizing songs into emotional groups such as Happy, Sad, and Neutral. Consequently, the system recommends songs aligned with the user's emotional state, empowering them to tailor their music experience to their mood. In a multimedia and technology landscape that is ever evolving, we're motivated to simplify playlist creation by replacing manual efforts with an automated system driven by the user's emotional cues. Central to our approach is real-time facial analysis through a camera, leveraging the expressive capacity of the human face. This data serves as the basis for crafting a personalized playlist in sync with the inferred emotional state. This innovative technique eradicates the laborious task of manually sorting songs, resulting in a well-suited playlist that resonates with the user's emotional disposition.

**Keywords:** Evolution of technology, Music and emotions, Real-time interactions, Automated playlist.

## 1. Introduction

The swift evolution of internet and mobile technology has transformed our access to an expansive reservoir of music resources. Notably, the music industry's inclination toward specific genres creates a challenge: accurately ascertaining user preferences based on their prevailing mood or emotional state. Music serves as a powerful medium to express and mirror emotions. Whether in times of elation, where upbeat melodies resonate, or when seeking solace amid stress, calming tunes offer respite, and during periods of melancholy, somber melodies provide comfort [1].

In response, we propose an innovative system that captures real-time user interactions, like conversations, to discern their emotional state. This emotional input serves as the cornerstone for suggesting songs that align with their feelings. Our approach involves categorizing songs into distinct emotional clusters, encompassing Happy, Sad, Confused, and Neutral sentiments. Consequently, the system recommends songs that correspond to the user's current emotional state, enabling a personalized music experience reflective of their mood [2].

The undeniable impact of music on enriching individual lives encompasses entertainment and therapeutic benefits. Amid the dynamic multimedia and technology landscape, sophisticated music players featuring variable

playback speeds and genre classifications have emerged. Our drive emanates from the aspiration to streamline the manual playlist creation process, replacing it with an automated mechanism hinged on the user's emotional state.

Central to our strategy is leveraging facial analysis through a camera to harness the human face's innate capacity for emotional expression. By directly extracting data from user facial expressions, we glean insights into their emotional disposition. This data forms the bedrock for crafting a bespoke playlist tailored to the deduced emotional state. This innovative approach eradicates the labor-intensive task of manually sorting songs, culminating in a thoughtfully curated playlist harmonizing with the individual's emotional orientation [3].

The computerized analysis and comprehension of music represents pioneering frontier music information retrieval, drawing from diverse domains such as computer science, digital signal processing, mathematics, statistics, and musicology [4]. Recent progress encompasses various areas including automatic audio genre/mood classification, music similarity computation, audio artist identification, audio-to-score alignment, and query-by-singing/humming. One prominent application is content-based music recommendation, supported by intelligent context-based suggestions derived from multidisciplinary efforts in emotion description, detection, feature-based classification, and inference-driven recommendations. Human emotion depiction often relies on models like the circumflex model, which positions affects along two bipolar dimensions: pleasant-unpleasant- and arousal- sleep. Subsequent models, like Thayer's, introduced arousal and valence dimensions to depict emotions from calm to energetic and negative to positive. Nevertheless, some emotions such as nervousness and excitement have been underexplored. Advancements in digital signal processing and feature extraction are driving rapid progress in automatic emotions detection and recognition in music with implications for music entertainment and human-computer interaction system. These developments signify an exciting fusion of technological innovation and artistic expression in the realm of music analysis and understanding [5].

In this endeavor, we introduce an innovative emotion-based and user-interactive music system. The objective is to deliver user-preferred music enriched with emotional awareness. Beginning with expert recommendations, the system offers user-tailored suggestions. Should a user decline a recommendation, they retain the autonomy to select desired music independently.

## 2. Materials and Methods

To ensure representation across all emotional states, it is advised to gather a broad dataset of images tagged with emotions (such as happiness, sorrow, rage, and surprise) from different sources. The collection ought to contain human faces with a variety of expressions [6-7]. The music dataset's emotion labels ought to correspond to the same emotion categories as those in the image dataset. Then a pre-trained deep learning model, VGGFace has been utilized for efficient emotion recognition [8-9]. For developing the emotion recognition model, the image dataset should be divided into training and validation sets. Utilize the training data to train the model, then tweak it for maximum emotion detection precision. The trained emotion recognition model's performance was analyzed on accuracy, precision, recall and F1-score performance metrics.

The algorithm that takes the emotions in the photographs into account and matches them with songs or music genres that reflect those emotions. For context-aware music suggestions, collaborative filtering and content-based filtering methods were employed. For the development, training, and application of the emotion recognition model, we have utilized the TensorFlow. For image preprocessing tasks like resizing and normalization, we used

the OpenCV library. Utilized the scikit-learn and TensorFlow Recommenders packages to implement the recommendation algorithm. To enable quick emotion detection from user-uploaded or shot photographs, integrated the trained emotion recognition model into a real-time user interface. Created the system's front-end by utilizing HTML, CSS, and JavaScript to provide a fluid and aesthetically pleasing user experience. The architecture of a CNN model is shown in Figure 1 and the visualization of feature map is shown in Figure 2.

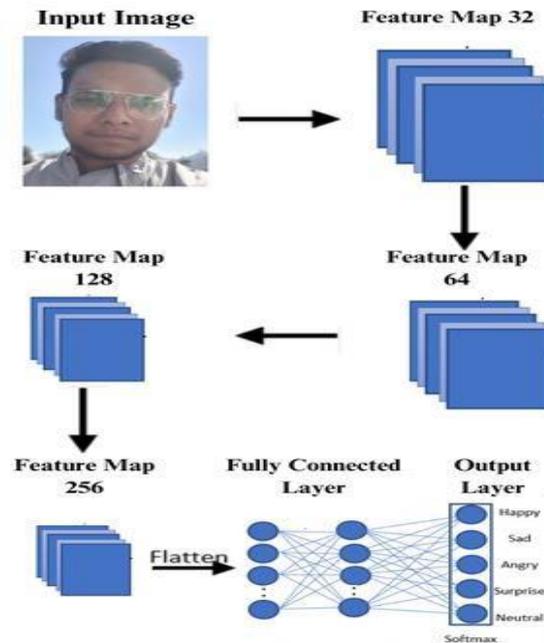


Fig.1. The architecture of a CNN model

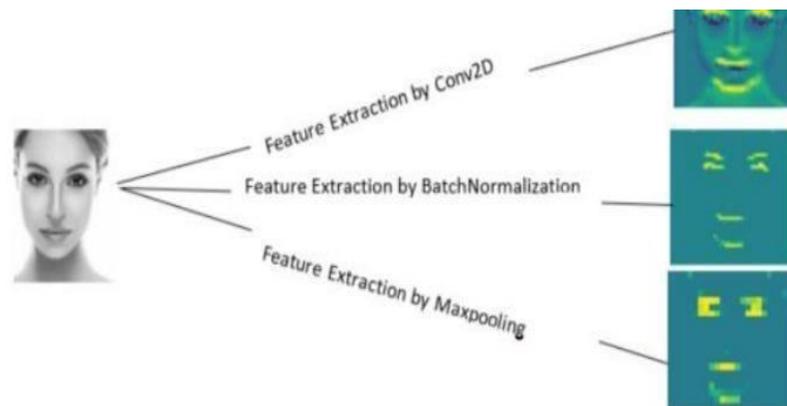


Fig 2. Visualization of feature map

The goal of the proposed image-based song recommendation system with emotion identification is to improve users' overall musical listening experiences by offering personalized, context-sensitive, and easily accessible music choices based on emotions extracted from images. A flow chart of emotion-based music recommendation system is shown in Fig.3.

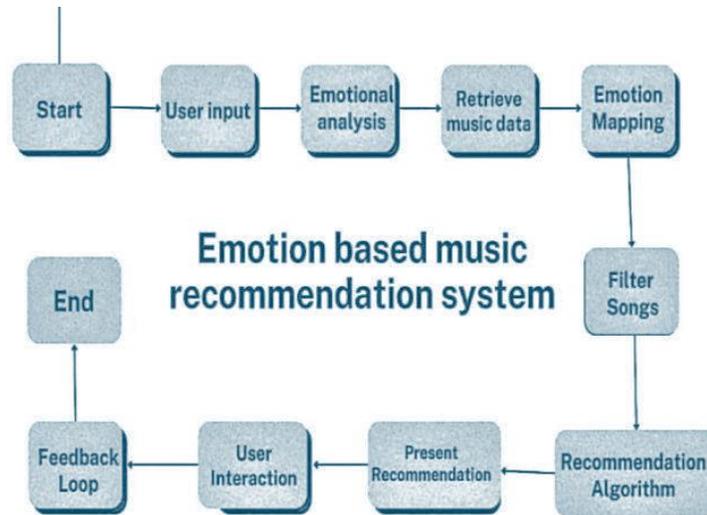


Fig.3. Emotion-based music recommendation system

### 3. Results

We examined several studies utilizing support vector machines (SVM), extreme learning machine (ELM), and convolutional neural networks. Table 1 presents a comparison of relevant algorithms, including corresponding accuracy values for each study. The incorporation of a convolutional neural network enhances the accuracy of emotion detection.

Table 1 presents a comparison of relevant algorithms

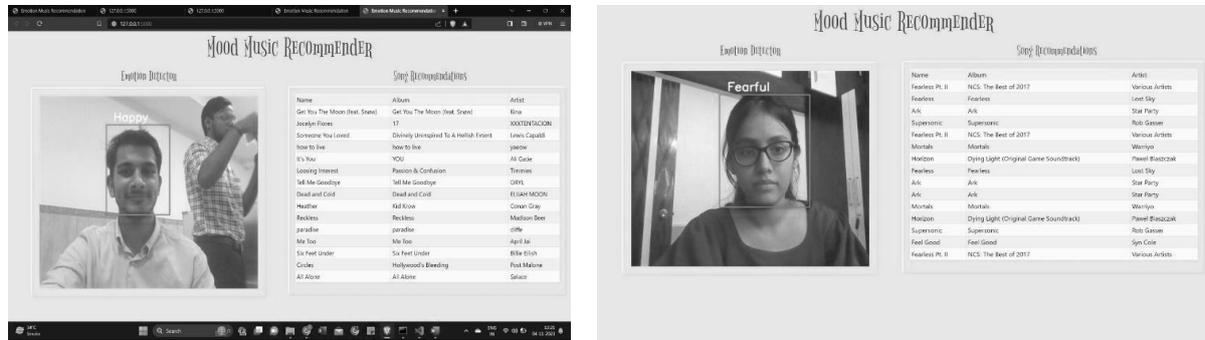
Algorithm	CNN
Testing Accuracy	0.64
Validation Accuracy	0.65

Table 2 presents the hyperparameters utilized in the trained CNN network. The learning regulates weight updates after each batch, while multiple epochs denote iterations over the entire training dataset. Batch size indicates the number of patterns presented to the network before weight adjustments. Activation functions assist the model in learning nonlinear boundaries.

Table 2. Hyperparameters utilized to train CNN model

Hyperparameters	Values
Batch size	64
No. of Classes	7
Optimizer	Adam
Learning Rate	0.0001
Epoch	75
No. of Layers	9
Activation Function	ReLU SoftMax
Loss Function	Categorical Cross-Entropy

Adam serves as an alternative optimization algorithm to stochastic gradient descent for training deep learning models. The categorical cross-entropy loss function is employed to gauge errors in deep learning models, especially in single label, multiclass classification tasks as shown in Figure 4.



(a) (b)  
 Fig.4. (a) Detection of emotions-Happy (b) Detection of emotions-Fearful

**4. Discussion**

Emotion classification has emerged as a research domain with extensive potential applications. Prior to engaging in emotion classification, the preliminary step of facial recognition holds immense significance. A noteworthy instance of this technology can be found in Apple's recent implementation of facial recognition for unlocking their devices. However, the scope of facial authentication extends to more advanced levels, such as its utilization in heightened security contexts like ATMs or office access, thereby contributing to enhanced resource integrity and safety.

Facial recognition can be considered a highly effective biometric sensor with substantial untapped potential. This technology, when elevated to more sophisticated applications, has the potential to revolutionize various domains. For example, it could be harnessed in market research to infuse brands with added value. By gathering client data and analyzing their emotions, businesses can tailor their offerings to customers with a personalized touch, bolstering customer loyalty. Moreover, envisioning a future where emotion recognition systems automatically respond to customer needs rather than relying on queries is within reach, positioning emotion recognition as a potent and valid business tool. In prominent retail environments, the prevalent practice of soliciting feedback through button presses could be transformed. Major supermarkets could leverage facial recognition technology to capture and interpret customer emotions as they queue, generating insightful data for enhanced customer service provisioning. An additional area of impact is the mitigation of accidents caused by impaired driving, notably due to alcohol consumption. Alcohol-induced impairment results in reduced driver focus, often leading to hazardous situations. Our technology has the potential to function as an alert system, notifying drivers of their unfit state to operate a vehicle safely. Furthermore, it could seamlessly integrate into autonomous vehicles, providing assistance to drivers on their journeys.

Collectively, both facial recognition and emotion detection exhibit substantial latent potential. The onus lies on us to harness this potential and derive optimal utility from these advancements.

## 5. Conclusions

Given the capacity of music to evoke emotions within users, a universal model has been deployed to suggest music selections based on the emotional states of users. Human emotions hold a crucial role in conveying an individual's inner thoughts. The principal objective of this system is to identify shifts in the user's emotional disposition and subsequently play music aligned with the user's inclinations, encompassing a wide array of musical tracks. To accomplish this, a Convolutional Neural Network (CNN) algorithm is employed for emotion classification, relying on the alterations in the shape, size, and movement of facial features such as eyebrows, eyes, and mouth. These cues categorize emotions into six fundamental types: sadness, happiness, anger, fear, disgust, and surprise, forming the basis for generating personalized playlists. The preference for the CNN algorithm over Support Vector Machines (SVM) lies in its ability to autonomously discern pivotal features in an image, eliminating the need for human intervention. Notably, CNN exhibits superior predictive accuracy compared to SVM, contributing to the system's efficacy.

Various strategies exist for employing the Music Recommender System, as illuminated through an extensive survey of the literature. We thoroughly reviewed the methodologies put forth by previous researchers and developers. During our investigation, we primarily encountered two distinct, autonomous approaches. The initial approach involved pinpointing precise emotions from facial expressions, while the second revolved around categorizing songs into primary emotional categories based on their acoustic attributes. Our system's objectives were formulated based on the outcomes of these investigations. Consequently, we made the decision to amalgamate these two methodologies, presenting a comprehensive solution for the existing issue. Contemporary technology permits the discernment of user emotions, effectively identifying emotions such as happiness, sadness, anger, neutrality, and surprise. Following emotion recognition, our proposed approach furnishes users with a playlist of music selections tailored to their emotional state. The processing of vast datasets leads to heightened memory and CPU consumption. Despite this challenge, the development process becomes more intricate and captivating. Our aim is to ensure the affordability and accessibility of this application by deploying it on a universally compatible platform. Through our facial emotion-driven music recommendation system, users will experience enhanced ease in crafting and organizing playlists.

Empirical testing of the proposed system, utilizing a webcam setup, has yielded substantial accuracy in its outcomes. Given the inherent variability of human emotions, influenced by both internal and external stimuli, attaining absolute precision of 100% remains challenging. However, with the integration of enhanced algorithms and diligent research, the potential exists to develop an emotion-based music recommendation system that approaches perfection. The proposed system has been rigorously evaluated against a web camera framework, and its implementation costs are minimal. The average estimated time required for the various modules of the proposed system is also reasonable.

## References

- [1] Gandolkar, V. K., & G, B. (2022). Using face expression detection, provide music recommendations. *International Journal for Research in Applied Science and Engineering Technology*, 10(7), 1529–1533. <https://doi.org/10.22214/ijraset.2022.45499>
- [2] R. Ramanathan, R. Kumaran, R. Ram Rohan, R. Gupta and V. Prabhu, "An Intelligent Music Player Based on Emotion Recognition," 2017 2nd International Conference on Computational Systems and

- Information Technology for Sustainable Solution (CSITSS), Bengaluru, India, 2017, pp. 1-5, doi: 10.1109/CSITSS.2017.8447743.
- [3] Bhutada, S., Sadvika, Gutta, A., & Reddy, P. S. (2020). EMOTION BASED MUSIC RECOMMENDATION SYSTEM. *Journal of Emerging Technologies and Innovative Research*, 7(4). <http://www.jetir.org/papers/JETIR2004499.pdf>
- [4] Ellis, C. L., & Salmoni, A. W. (2019). Descriptive analysis of university-student music preferences during different forms of physical activity. *Psychology of Music*, 49(2), 177–192. <https://doi.org/10.1177/0305735619847759>
- [5] Bonnette, R., Smith, M. C., Spaniol, F., Melrose, D., & Ocker, L. (2010). The effect of music listening on running performance and rating of perceived exertion of college students. *The Journal of Strength and Conditioning Research*, 24, 1. <https://doi.org/10.1097/01.jsc.0000367073.45565.4b>
- [6] Ballmann CG, Maynard DJ, Lafoon ZN, Marshall MR, Williams TD, Rogers RR. Effects of Listening to Preferred versus Non-Preferred Music on Repeated Wingate Anaerobic Test Performance. *Sports (Basel)*. 2019 Jul 29;7(8):185. doi: 10.3390/sports7080185. PMID: 31362419; PMCID: PMC6723041.
- [7] N. Sebe, I. Cohen, T. S. Huang et al., “Multimodal emotion recognition,” *Handbook of Pattern Recognition and Computer Vision*, vol. 4, pp. 387– 419, 2005.
- [8] Athavle, M. (2021). Music recommendation based on face emotion recognition. *Journal of Informatics Electrical and Electronics Engineering (JIEEE)*, 2(2), 1–11. <https://doi.org/10.54060/jieee/002.02.018>
- [9] Kathavate, S. (2021). Music Recommendation System using Content and Collaborative Filtering Methods. *IJERT*. <https://doi.org/10.17577/IJERTV10IS020071>