IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 2, May 2025
Impact Factor: 7.67

Emotion-Driven Music Recommendations with CNN

Dr. V. Kalaivazhi[1], U. Suirya[2], K. Santhosh Kumar[3]
Associate Professor (**M.Tech, Ph.D**), Department of Information Technology[1]
Students, B. Tech., Final Year, Department of Information Technology[2,3,4,5]
Anjalai Ammal Mahalingam Engineering College, Thiruvarur, India

Abstract: *With the rapid growth of digital media and artificial intelligence, personalized content delivery has become a major focus in enhancing user experience. Music, being a deeply emotional and personal medium, offers a unique opportunity for intelligent systems to understand and respond to users' moods. This project, titled Emotion-Driven Music Recommendation Using Convolutional Neural Networks (CNN), proposes an innovative approach to music recommendation that leverages deep learning techniques to detect user emotions from facial expressions. By using CNNs, the system can accurately classify real-time facial images into emotion categories such as happy, sad, angry, or neutral. Once the emotion is identified, the system dynamically recommends songs that align with the user's current mood, thereby creating a more immersive and emotionally adaptive listening experience. The proposed model is trained on a facial emotion dataset and integrated with a music recommendation engine. This approach not only enhances user satisfaction but also demonstrates the practical application of machine learning in human-centered design. The system aims to provide a seamless, emotion-aware recommendation framework that goes beyond traditional static playlists, making music listening more engaging and intuitive..*

Keywords: Emotion recognition, Music recommendation, Convolutional Neural Networks (CNN), Facial expression analysis, Deep learning, Real-time emotion detection, Personalized music, Human-centered AI, Emotion-aware system, User experience enhancement

I. INTRODUCTION

In today's digital era, the demand for personalized user experiences has significantly increased, especially in the entertainment industry. Music, being a powerful medium of emotional expression, plays a crucial role in influencing and reflecting human

emotions. Traditional music recommendation systems primarily rely on user behavior, preferences, or listening history, often failing to adapt dynamically to the user's real-time emotional state. This project aims to bridge that gap by developing an emotion-driven music recommendation system using Convolutional Neural Networks (CNN).

Emotions can be effectively detected through various physiological signals, facial expressions, or voice tone. Among these, facial expression analysis provides a non-invasive and practical approach to emotion recognition. By leveraging deep learning techniques, particularly CNNs, the system can accurately interpret facial cues and map them to specific emotional states such as happiness, sadness, anger, or surprise.

The core idea of this system is to detect the user's emotional state in real-time through facial recognition, and then recommend songs that match or improve the user's mood. This innovative approach not only enhances user satisfaction but also demonstrates how artificial intelligence can be applied to create more human-centric, responsive systems.

With the integration of emotion recognition and music streaming, the system offers a personalized, adaptive experience that could revolutionize how users interact with music platforms. This project thus stands at the intersection of affective computing, machine learning, and multimedia systems, presenting a compelling case for the next generation of intelligent applications.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 2, May 2025
Impact Factor: 7.67

II. CNN (CONVOLUTIONAL NEURAL NETWORKS)

Convolutional Neural Networks (CNNs) are a specialized type of deep learning model designed for processing data that has a grid-like topology, such as images. CNNs are highly effective for tasks involving visual recognition, making them ideal for emotion detection from facial expressions in this project.

A CNN consists of multiple layers that automatically and adaptively learn spatial hierarchies of features from input images. The main layers in a CNN include:

• Convolutional Layers: These layers apply filters to the input image to create feature maps. The filters detect patterns such as edges, textures, and shapes, which are essential for identifying facial expressions.

• Pooling Layers: Pooling (usually max pooling) reduces the spatial dimensions of the feature maps, helping in downsampling while retaining the most important features. This reduces computational complexity.

• ReLU (Rectified Linear Unit) Layers: These introduce non-linearity to the model, enabling it to learn complex patterns in the data.

• Fully Connected Layers: These layers connect every neuron in one layer to every neuron in the next and are used to classify the learned features into output categories, such as different emotions.

In the context of the emotion-driven music recommendation system, the CNN is trained on facial expression datasets. When a user's image is input to the system, the CNN analyzes it and classifies the emotion in real-time. The predicted emotion is then used to recommend music that matches or elevates the user's mood.

CNNs are preferred in this application due to their high accuracy, automatic feature extraction, and ability to handle image data efficiently. Their ability to generalize across different face types, lighting conditions, and angles makes them robust and reliable for real-world applications.

III. PROBLEM STATEMENT

• Many music recommendation systems rely on past listening history or user ratings, which may not reflect a person's current mood. This makes it difficult to get the right music at the right time. Manually choosing songs can also be time-consuming and inconvenient.

• Existing AI-based emotion detection methods require complex processing and large datasets, making them unsuitable for real-time use. Additionally, many systems do not adapt instantly to changes in mood. Privacy concerns also arise when facial data is stored for AI processing. There is a need for a simple, fast, and private way to recommend music based on real-time emotions.

• The proposed system captures a user’s facial expressions and instantly suggests music that matches their mood. This eliminates the need for manual selection and enhances the listening experience. The system is lightweight, does not require AI training, and works efficiently in real time. It provides a seamless way to enjoy music based on emotions without complexity.

IV. EXISTING SYSTEM

• CAMC Music is a novel context-aware multi-criteria music recommendation system.

• It aims to improve recommendation accuracy without relying on:

I. Explicit user ratings

II. Music content features

III. User-specific attributes

• Framework & Methodology:

I. Integrates contextual information into a multi-criteria decision-making (MCDM) framework.

II. Generates Top-N music recommendations tailored to the user's current context.

• Contextual Relevance Evaluation:

Evaluates the impact of three core elements:

I. User

II. Music genre

III. Surrounding context

DOI: 10.48175/IJARSCT-26230

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 2, May 2025
Impact Factor: 7.67

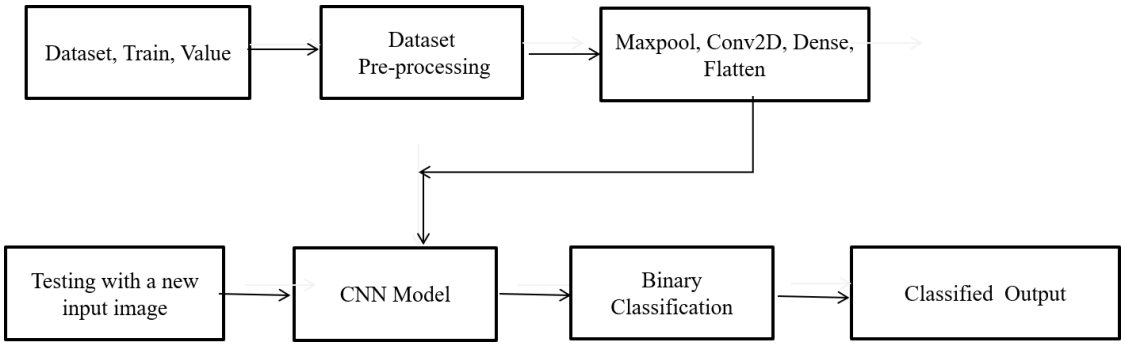
V. PROPOSED SYSTEM

- The proposed system is a real-time music recommendation platform that leverages facial recognition to analyze users' emotions and suggest suitable music.
- A camera captures the user's face, and the system interprets facial expressions to determine their emotional state. Based on the detected emotion, a predefined rule-based aligns different moods with corresponding music genres.
- Joyful expressions trigger upbeat music, while sad expressions prompt soothing melodies, ensuring an emotionally resonant experience. The system operates in real time. It enhances user engagement by seamlessly adapting to emotional changes and promoting mental relaxation. This approach offers an intuitive and personalized listening experience, making music selection effortless and immersive.
- The system's simplicity eliminates the need for complex AI models while still delivering effective results. Designed for entertainment and emotional well-being, it can be integrated into various platforms, such as mobile applications and smart home systems. By aligning music with emotions, the proposed system creates a dynamic and responsive musical experience for users.

ADVANTAGES OF THE PROPOSED SYSTEM

- Real-Time Emotion-Based Recommendations – The system instantly detects a user's emotional state through facial recognition and provides music suggestions that align with their mood, ensuring an immediate and relevant listening experience.
- No Need for Historical Data – Unlike traditional recommendation systems that rely on past listening behavior, this approach eliminates dependency on previous user data, making it effective for new users.
- Simple and Lightweight – The system follows a rule-based approach to map emotions to music genres, avoiding complex AI models and reducing computational requirements, making it faster and more efficient.
- Enhances Emotional Well-Being – By aligning music with the user's current emotional state, the system promotes relaxation, improves mood, and provides an overall better emotional experience.
- Seamless and Hands-Free Experience – Users do not need to manually select songs; the system automatically adjusts recommendations, making music selection effortless and intuitive.
- Privacy-Friendly Approach – Since it does not store facial data for AI training, the system ensures better user privacy compared to AI-based emotion recognition methods that require large datasets.
- Versatile and Adaptive – The system can be integrated into multiple platforms, including mobile applications and smart home devices, allowing for a wide range of real-world applications.
- Instant Mood Adaptation – Unlike traditional systems that take time to update preferences, this system dynamically adjusts recommendations based on real-time facial expressions, providing a more engaging and responsive user experience.

VI. SYSTEM ARCHITECTURE DIAGRAM



DOI: 10.48175/IJARSCT-26230

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 2, May 2025
Impact Factor: 7.67

VII. MODULES

List of Modules:
1. Facial Expression Detection Module
2. Emotion Classification Module
3. Music Recommendation Module
4. Real – Time Processing Module
5. User Interface Module

1. Facial Expression Detection Module:
The Facial Expression Detection Module is responsible for capturing the user's face using a camera and analyzing their facial expressions in real time. This module utilizes image processing techniques and facial feature extraction to identify key points on the face, such as the position of the eyes, eyebrows, and mouth.
By analyzing these facial landmarks, it determines subtle changes in expression that indicate the user's emotional state.
This process ensures accurate detection of facial movements, forming the foundation for the emotion classification system.

2. Emotion Classification Module:
Once facial expressions are detected, the Emotion Classification Module categorizes them into predefined emotional states such as happy, sad, neutral, angry, or surprised.
This module employs deep learning techniques, particularly convolutional neural networks (CNNs), to classify emotions with high accuracy. By training on diverse datasets of facial expressions, the system can recognize even subtle mood variations.
The classification process is designed to be efficient, ensuring minimal delay between detecting an expression and identifying the corresponding emotion.

3. Music Recommendation Module
The Music Recommendation Module maps the detected emotion to a suitable music genre using a predefined rule-based system. For example: ▪ Happy → Upbeat pop, dance, or energetic tunes
▪ Sad → Soft acoustic, classical, or slow ballads
▪ Angry → Rock, heavy metal, or intense music
▪ Neutral → Lo-fi, jazz, or ambient music
Instead of relying on complex AI models, this module simplifies the recommendation process by directly linking emotions to genres, ensuring a seamless and intuitive user experience. The music selection is optimized to enhance or balance the user's current mood.

4. Real-Time Processing Module
To maintain a smooth and immersive experience, the Real-Time Processing Module ensures instant emotion detection and music selection. This module optimizes the processing speed of facial recognition and classification to reduce lag.
It utilizes lightweight models and efficient algorithms to process user input without noticeable delays. Additionally, it continuously monitors changes in facial expressions, updating music recommendations dynamically as the user's emotions shift.

5. User Interface Module
The User Interface Module provides a simple, interactive, and user-friendly platform where users can see their detected mood and receive music recommendations effortlessly.
• The UI displays the following key elements: Real-time camera feed with emotion detection results.
• A visual indicator of the identified emotion (e.g., emoji or text label).

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 2, May 2025
Impact Factor: 7.67

- Suggested music genre and the currently playing song.
- Options to manually adjust recommendations or skip songs.

VIII. CONCLUSION

The project "Emotion-Driven Music Recommendation Using Convolutional Neural Networks (CNN)" presents a novel approach to enhance user experience by integrating artificial intelligence with emotional analysis. By utilizing CNNs to classify user emotions from facial expressions or audio cues, the system can dynamically recommend music that aligns with the user's current emotional state. This not only makes the recommendation process more personalized but also more intuitive and adaptive in real-time.

The implementation of CNNs proves highly effective in emotion recognition due to their ability to learn spatial hierarchies in data, making them ideal for processing facial images or spectrograms of voice inputs. The trained model can categorize emotions such as happiness, sadness, anger, surprise, and more, with significant accuracy. Once the emotion is identified, a recommendation engine maps this to a curated music database, offering users a playlist that matches or enhances their mood.

This system has broad applications, from entertainment platforms and mental wellness apps to smart assistants and emotion-aware gaming environments. It also highlights the potential of merging deep learning models with user-centric technologies to create adaptive, intelligent systems.

In conclusion, this project demonstrates that CNN-based emotion detection can significantly enhance the personalization of music recommendations, offering a promising step towards emotionally intelligent media systems. Future work can explore multi- modal emotion analysis (combining facial, voice, and textual data) and integration with wearable technology to further enrich the user experience.

REFERENCES

[1]. Singh, M., & Mehr, S. A. (2023). Universality, domain-specificity and development of psychological responses to music. Nature Reviews Psychology, 2(6), 333–346.

[2]. Lacker, B., & Way, S. F. (2024, May). Socially-motivated music recommendation. In Proceedings of the International AAAI Conference on Web and Social Media (Vol. 18, pp. 879–890).

[3]. Schäfer, K., & Eerola, T. (2020). How listening to music and engagement with other media provide a sense of belonging: An exploratory study of social surrogacy. Psychology of Music, 48(2), 232–251.

[4]. Nowak, R., & Bennett, A. (2022). Music sociology: Value, technology, and identity. Routledge.

[5]. Dingle, G. A., Sharman, L. S., Bauer, Z., Beckman, E., Broughton, M., Bunzli, E., Davidson, R., Draper, G., Fairley, S., Farrell, C., & Flynn, L. M. (2021). How do music activities affect health and well-being? A scoping review of studies examining psychosocial mechanisms. Frontiers in Psychology, 12, 713818.

[6]. Meyer, J., & Manfredi, V. (2024). Auxiliary speech by vocal tract modulation and musical surrogacy. In The Oxford Handbook of Language and Music.

[7]. Schedl, M., Zamani, H., Chen, C.-W., Deldjoo, Y., & Elahi, M. (2018). Current challenges and visions in music recommender systems research. International Journal of Multimedia Information Retrieval, 7, 95–116.

[8]. Dinnissen, K., & Bauer, C. (2023, June). Amplifying artists' voices: Item provider perspectives on influence and fairness of music streaming platforms. In Proceedings of the 31st ACM Conference on User Modeling, Adaptation and Personalization (pp. 238–249).

[9]. Yin, H., Qu, L., Chen, T., Yuan, W., Zheng, R., Long, J., Xia, X., Shi, Y., & Zhang, C. (2024, January 21). On-device recommender systems: A comprehensive survey. arXiv preprint arXiv:2401.11441.

[10]. Elbir, A., & Aydin, N. (2020). Music genre classification and music recommendation by using deep learning. Electronics Letters, 56(12), 627–629.