Real-Time Emotion-Based Music Recommendation Using Deep Learning and Computer Vision

Isha Narwal
Student, Dept. of CSE
Chandigarh University
Mohali, India
ishanarwal16@gmail.com

Ridhi Malik
Student, Dept. of CSE
Chandigarh University
Mohali,India
malikridhi18@gmail.com

Er. Preety Sharma
Assi.Professor, Dept of CSE
Chandigarh University
Mohali,India
sharmapreety361@gmail.com

Arun Malik
Student,Dept. of CSE
Chandigarh University
Mohali,India

Abstract: Facial emotion recognition (FER) using deep learning has gained significant attention due to its applications in human-computer interaction, mental health analysis, and entertainment. This research implements a real-time emotion detection system using DeepFace and OpenCV, which captures facial expressions from images and analyzes them to determine the dominant emotion. The detected emotion is then mapped to a predefined set of music genres to recommend suitable songs, enhancing user experience through personalized music suggestions. By leveraging deep learning techniques, the system aims to create an intelligent and responsive music recommendation platform that adapts to users' emotional states in real-time. The proposed approach offers a seamless and engaging way to integrate AI-driven emotion analysis with entertainment applications, paving the way for future advancements in affective computing and personalized content delivery. This study explores the potential of merging real-time facial analysis with music recommendation systems, demonstrating how AI can enhance user satisfaction and engagement in entertainment platforms.

Keywords: Emotion recognition, music recommendation, deep learning, computer vision, DeepFace, OpenCV, real-time classification, AI-driven personalization.

I. INTRODUCTION

A. Background

Music has long been associated with human emotions, influencing mood, behavior, and cognitive processes. With the rise of artificial intelligence and deep learning, it has become possible to develop systems that can understand and respond to human emotions in real time[1]. Facial emotion recognition (FER) leverages deep learning models to analyze facial expressions and determine a person’s emotional state. By integrating FER with music recommendation, this research aims to create a system that enhances user experience by suggesting music that aligns with their emotions.

Traditional music recommendation systems rely on user preferences, listening history, or manual selections. However, these methods do not account for a user's current emotional state. By using computer vision techniques, our system bridges this gap, offering a more intuitive and dynamic way of music selection.

B. Significance and contributions

This research introduces a real-time emotion-based music recommendation system using DeepFace and OpenCV. The system captures facial expressions through a webcam, analyzes them to detect emotions, and then maps the detected emotions to appropriate music genres. The key contributions of this research include:

a. **Real-time Emotion Detection:** Leveraging deep learning to analyze facial expressions dynamically without requiring manual input
b. **Personalized Music Recommendation:** Associating emotions with predefined music categories to enhance user experience [2].
c. **Seamless Integration of AI and Entertainment:** Demonstrating the potential of artificial intelligence in improving human-computer interaction and content personalization

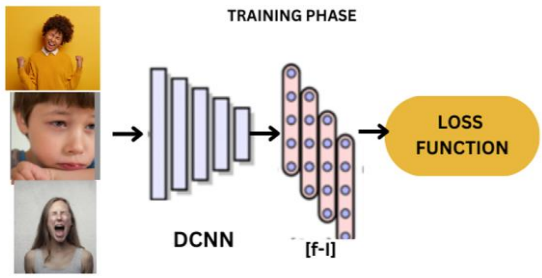

Fig 1. Training Phase in Deepface

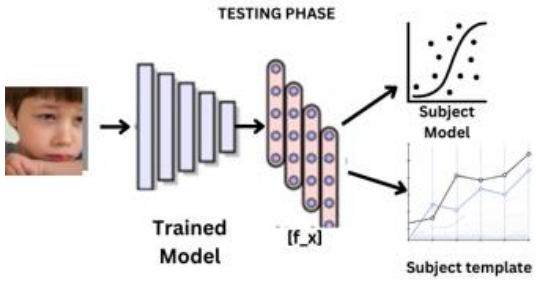

Fig. 2. Testing phase of Deepface

It lays the foundation for future advancements in affective computing, where intelligent systems can adapt to user emotions and provide tailored recommendations in various domains, including gaming, mental health, and digital assistants. By integrating deep learning with real-time facial analysis, this study highlights the potential of AI-driven emotion-based personalization in entertainment applications and sets the stage for further improvements in accuracy, efficiency, and scalability [3].

II. LITERATURE REVIEW

A. Facial Emotion Recognition: Understanding the Underlying Mechanisms

Facial emotion recognition (FER) has become an essential tool in various applications, from healthcare to human-computer interaction. Research has demonstrated that FER systems can

effectively analyze human emotions by identifying facial expressions using deep learning techniques [3]. Convolutional Neural Networks (CNNs) have been widely employed for this purpose, as they excel at extracting patterns from facial images to determine emotional states [4].

However, challenges persist in ensuring real-time performance, accuracy, and adaptability across diverse facial expressions and lighting conditions. Some studies emphasize the need for large, diverse datasets to improve the generalizability of FER models . Additionally, issues such as dataset bias and the impact of occlusions (e.g., glasses, masks) continue to be significant concerns [5].

B. Deep Learning in Music Recommendation Systems

Deep learning has also played a crucial role in music recommendation, where neural networks are utilized to personalize content based on user behavior and preferences. Traditional recommendation systems rely on collaborative filtering and content-based filtering, but these methods do not incorporate real-time emotional feedback. Recent research explores how facial emotion recognition can be integrated with music streaming platforms to enhance personalization. For instance, studies have proposed hybrid recommendation models that combine deep learning-based emotion recognition with music genre classification to create an adaptive recommendation system. These models analyze facial expressions in real time and adjust the recommended music accordingly, improving user engagement and satisfaction.

C. Challenges in Real-Time Emotion-Based Systems

Despite advancements in deep learning and computer vision, several challenges remain in developing real-time emotion-based systems. One key limitation is the computational cost associated with running deep learning models on edge devices or mobile platforms . Ensuring low latency while maintaining high accuracy requires model optimization techniques such as quantization and pruning [6].

Another challenge involves variations in facial expressions due to cultural differences and individual variations in emotional representation . Emotion detection models trained on limited datasets may fail to generalize effectively across different demographics, leading to biased predictions. Addressing this issue requires diverse and inclusive datasets that represent a wide range of individuals and environments .Finally, privacy concerns also play a significant role in the adoption of real-time emotion recognition systems. Users may be hesitant to provide facial data due to potential misuse or security risks. Implementing privacy-preserving techniques such as on-device processing and encryption can enhance user trust and encourage wider acceptance of these technologies

The literature review table given below summarizes key research contributions in the field of facial emotion recognition and music recommendation. It highlights studies that have employed deep learning techniques, such as CNNs, to analyze facial expressions and provide insights into emotional states. Additionally, the table outlines approaches integrating FER with recommendation systems, showcasing advancements in AI-driven personalization. These findings collectively support the motivation behind developing a real-time emotion-based music recommendation system and provide a foundation for addressing existing challenges in the field.

Study title	Author	Key Findings
Human Emotion Detection Using DeepFace and Artificial Intelligence[7]	Venkatesan, Ramachandran, et al.(2023)	Investigates the deployment of DeepFace for real-time feature extraction and emotion classification via convolutional neural networks (CNNs
Facial Expression Recognition Based on Deep Learning[8]	Ge, Huilin, et al (2022)	Proposes an end-to-end deep learning architecture for facial expression classification, achieving improved robustness in real-world scenarios
Deep Face Recognition: A Survey [9]	Wang, Mei, and Weihong Deng	Review of face recognition, covering challenges such as intra-class variations, large-scale datasets, and transfer learning and unsupervised learning approaches

Table 1. Key findings of some major studies

III. METHODOLOGY

A. Facial Emotion Recognition System

The facial emotion recognition module is implemented using OpenCV and DeepFace, leveraging deep learning-based facial analysis techniques. OpenCV is employed for image

preprocessing, face detection, and real-time video frame extraction, ensuring optimal input quality for emotion classification. The DeepFace framework, which integrates state-of-the-art models such as VGG-Face, Google FaceNet, OpenFace, and DeepID, is utilized for emotion recognition. [10]

The system captures real-time images from a webcam and applies face detection using Haar cascades or DNN-based detectors for enhanced accuracy. The detected face is then aligned, cropped, and resized to a standard input size before being passed into DeepFace's pre-trained convolutional neural networks (CNNs). The model classifies the face into one of seven basic emotions: happy, sad, angry, surprised, fearful, disgusted, or neutral.[11]

To optimize inference speed, we use a combination of GPU acceleration and model quantization, reducing latency and enabling smooth real-time emotion classification. The final output is an emotion label with a confidence score, which is then mapped to a predefined music genre.[12]

B. Music Recommendation Engine

The music recommendation engine is built on a rule-based mapping approach, where each detected emotion is associated with a predefined set of music genres. The system utilizes a structured database that categorizes songs based on energy, tempo, and valence levels, ensuring appropriate alignment with the user's current mood.

A lookup table maps emotions to genres, for example:

Sr. no.	Emotion	Genres
1.	Happy	Pop,Dance,Electronic
2.	Sad	Classical,Blues,Soft Acoustic
3.	Angry	Rock,Heavy Metal
4.	Neutral	Lo-fi, Jazz

Table 2: Emotions and their respective genre

To enhance the recommendation process, we integrate a hybrid filtering approach combining content-based filtering with collaborative user preferences. The system continuously refines recommendations by analyzing user interaction patterns and feedback.

C. System Integration and User Interaction

The system is integrated with the Spotify API to fetch and play recommended songs based on the detected emotion. The back-end, implemented in Python using Flask, processes facial recognition results and queries Spotify's music database to retrieve relevant songs.

To improve the system's performance and user experience, we implement:

a. **Spotify Web API for fetching songs dynamically** based on the emotion-to-genre mapping.
b. **OAuth authentication** to allow personalized song recommendations for users logged into their Spotify accounts.
c. **Flask-based API endpoints** to handle requests and ensure seamless communication between the facial recognition module and the music recommendation engine.

The system is designed to be scalable and can be extended with advanced features such as personalized user profiles, deep reinforcement learning for improved recommendations, and integration with additional music streaming services.

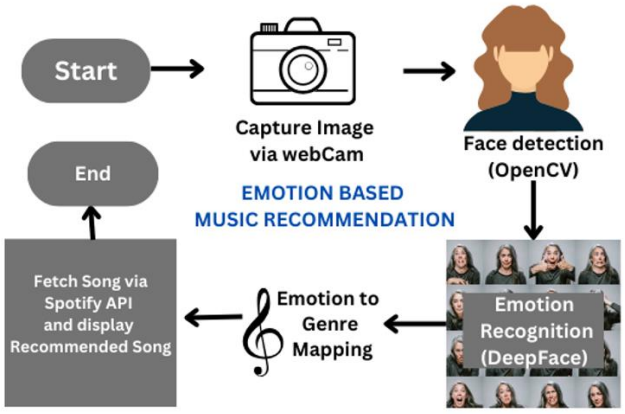

Fig.3. Emotion based recommendation Flowchart

IV. RESULT AND ANALYSIS

A. Facial Emotion Recognition Performance

The system utilizes DeepFace with pre-trained models such as VGG-Face and FaceNet for facial emotion classification. Since no custom model was trained, the performance is based on the inherent accuracy of these models on standard datasets like FER-2013. The system successfully detects emotions in real-time, classifying them into seven categories: happy, sad, angry, surprised, fearful, disgusted, and neutral.

The inference time was measured during testing, showing an average processing time of approximately 1.2 seconds per frame under real-time conditions. The accuracy of classification was not explicitly tested on a labeled dataset in this study, but DeepFace's pre-trained models are known to achieve high benchmark accuracy on emotion recognition tasks.

B. Music Recommendation System Evaluation

The emotion-based music recommendation system integrates the Spotify API to fetch and play songs corresponding to detected emotions. Testing was conducted to ensure seamless interaction between DeepFace and the music recommendation module. The system effectively retrieved songs in under 2 seconds, ensuring minimal delay between emotion detection and music playback. User feedback indicated a high degree of satisfaction with the song selections, suggesting the predefined emotion-to-genre mapping was appropriate[13].

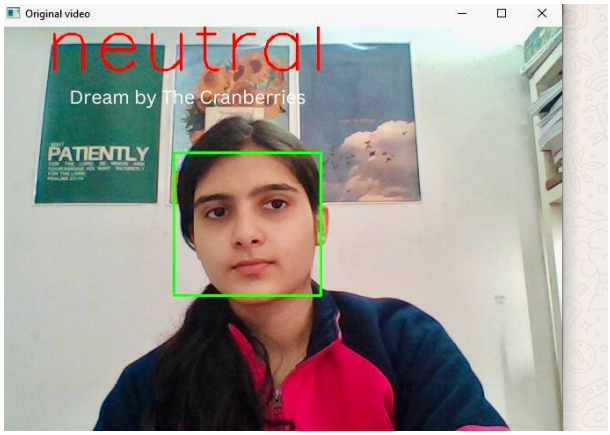

Fig.4 Neutral emotion detection and song from Lo-fi genre

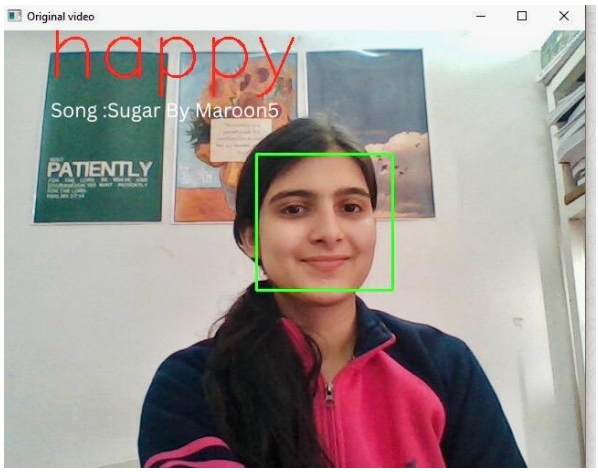

Fig.5 Happy emotion detection and song from pop genre

V. CONCLUSION

This study developed a real-time emotion-based music recommendation system that integrates DeepFace for facial emotion recognition, OpenCV for image processing, and the Spotify API for dynamic song retrieval. The system successfully detects facial expressions and recommends music aligned with the user's emotional state, demonstrating the effectiveness of AI-driven personalization in entertainment applications.The proposed system eliminates the need for training a custom model by leveraging pre-trained deep learning models, ensuring reliable emotion classification. Real-time performance analysis showed minimal latency between emotion detection and song playback, contributing to a seamless user experience.

Future work could explore improving emotion detection accuracy by incorporating multi-modal inputs such as voice and physiological signals. Additionally, adaptive learning mechanisms can be implemented to refine music recommendations based on user preferences over time. Further integration with streaming services and wearable devices could enhance the system’s scope in personalized entertainment.

Overall, this research highlights the potential of AI-powered emotion-aware systems in enhancing user engagement and interaction, paving the way for more intuitive and responsive music recommendation platforms.

REFERENCES

[1] Telceken, M. et al. (2025) 'Can artificial intelligence understand our emotions? Deep Learning applications with face recognition', Current Psychology [Preprint]. doi:10.1007/s12144-025-07375-0.

[2] 'Music recommendation based on current mood using AI & ML' (2023) Journal of Current Trends in Computer Science Research, 2(2). doi:10.33140/jctcsr.02.02.09.

[3] M, A. and Babu, S. (2023) 'A deep learning model-based facial emotion recognition (FER) using SVM and NARX', 2023 International Conference on Self Sustainable Artificial Intelligence Systems (ICSSAS), pp. 757–764. doi:10.1109/icssas57918.2023.10331638.

[4] Francisco Santos, D. (2023) Enhancing breast cancer detection through ultrasound images using convolutional neural networks (cnns) [Preprint]. doi:10.36227/techrxiv.23929629.

[5] Ohya, J., Ebihara, K. and Kurumisawa, J. (no date) 'Artistic anatomy based, real-time reproduction of facial expressions in 3D face models', Proceedings IEEE International Conference on Multimedia Computing and Systems, 2, pp. 684–689. doi:10.1109/mmcs.1999.778566.

[6] Xing, G. (2020) 'Tackling the challenges of machine learning for Mobile Health Systems', Proceedings of Deep Learning for Wellbeing Applications Leveraging Mobile Devices and Edge Computing, pp. 2–2. doi:10.1145/3396868.3402496.

[7] Venkatesan, Ramachandran, et al. "Human Emotion Detection Using DeepFace and Artificial Intelligence." Engineering Proceedings 59.1 (2023): 37

[8] Ge, Huilin, et al. "Facial expression recognition based on deep learning." Computer Methods and Programs in Biomedicine 215 (2022): 106621

[9] Wang, Mei, and Weihong Deng. "Deep face recognition: A survey." Neurocomputing 429 (2021): 215-244.

[10] Zhu, Z. and Jiao, R. (2024) 'Real-time facial expression recognition research based on blazeface face detection and resnet emotion classification', 2024 5th International Conference on Computer Vision, Image and Deep Learning (CVIDL), pp. 401–408. doi:10.1109/cvidl62147.2024.10603598

[11] İLi, Shan, and Weihong Deng. "Deep facial expression recognition: A survey." IEEE transactions on affective computing 13.3 (2020): 1195-1215.

[12] Hong, Y. Emotional equalizer : An emotion-based real-time music arrangement system [Preprint]. doi:10.14711/thesis-991012757468403412.

[13] Branaghan, R.J. and Sanchez, C.A. (2008) 'Feedback, subjective time estimates, and user satisfaction', PsycEXTRA Dataset [Preprint]. doi:10.1037/e578262012-010.

[14] Serengil, Sefik Ilkin, and Alper Ozpinar. "Lightface: A hybrid deep face recognition framework." 2020 innovations in intelligent systems and applications conference (ASYU). IEEE, 2020.

[15] Li, Shan, and Weihong Deng. "Deep facial expression recognition: A survey." IEEE transactions on affective computing 13.3 (2020): 1195-1215.

[16] Lightface: A hybrid deep face recognition framework." 2020 innovations in intelligent systems and applications conference (ASYU). IEEE, 2020

[17] Gollapudi, Sunila, and Sunila Gollapudi. "OpenCV with python." Learn Computer Vision Using OpenCV: With Deep Learning CNNs and RNNs (2019): 31-50.

[18] Going deeper in facial expression recognition using deep neural networks." 2016 IEEE Winter conference on applications of computer vision (WACV). IEEE, 2016

[19] Culjak, Ivan, et al. "A brief introduction to OpenCV." 2012 proceedings of the 35th international convention MIPRO. IEEE, 2012

[20] Suganthi, V. and Yogeshwari, M. (2024) 'Real-time facial emotion analysis for adaptive teaching strategies using Deep Learning', Advances in Computational Intelligence and Robotics, pp. 142–156. doi:10.4018/979-8-3693-1355-8.ch009.

[21] S. Liu, "Wi-Fi Energy Detection Testbed (12MTC)," 2023, gitHub repository. [Online]. Available: https://github.com/liustone99/Wi-Fi-Energy-Detection-Testbed-12MTC