International Journal for Modern Trends in Science and Technology
Volume 12, Issue 04, pages 75-81.
ISSN: 2455-3778 online
Available online at: http://www.ijmtst.com/vol12issue04.html
DOI: https://doi.org/10.5281/zenodo.19324488

Edge AI - Based Emotion Recognition System with Personalized Music Recommendation

M. Vijaya Bhaskar, D. Sivalekya, B. Mokshagna, Ch. Akhila, B. Jala Venkata Janeesha, B. Omserma

Department of CSE-AI, PBR Visvodaya Institute of Technology and Science, Kavali, A.P, India

To Cite this Article
M. Vijaya Bhaskar, D. Sivalekya, B. Mokshagna, Ch. Akhila, B. Jala Venkata Janeesha & B. Omserma (2026). Edge AI - Based Emotion Recognition System with Personalized Music Recommendation. International Journal for Modern Trends in Science and Technology, 12(04), 75-81. https://doi.org/10.5281/zenodo.19324488

Article Info
Received: 28 February 2026; Revised: 18 March 2026; Accepted: 22 March 2026.

KEYWORDS

Convolutional Neural Network, Edge AI, Emotion-Based Music Recommendation, FER Python Library.

ABSTRACT

Human emotions play a vital role in influencing mental health, decision-making, and user interaction with intelligent systems. Music has a strong psychological connection with emotions and can effectively regulate mood. Traditional music recommendation systems rely on user history and cloud-based processing, which leads to latency, privacy concerns, and lack of real-time emotional awareness.

This project proposes an Edge AI-Based Emotion Recognition System with Personalized Music Recommendation using the FER (Facial Emotion Recognition) Library for real-time emotion detection. Instead of training on a separate dataset, the system directly utilizes the FER Python library, which provides pre-trained deep learning models for facial expression analysis. The system performs on- device inference, ensuring low latency, enhanced privacy, and faster response time. Based on detected emotions such as happy, sad, angry, fear, surprise, disgust, and neutral, the system automatically recommends mood-based music playlists.

INTRODUCTION

INTRODUCTION CONVOLUTIONAL NEURAL NETWORK

A Convolutional Neural Networks (CNNs)

are a class of deep learning models specifically designed for processing structured grid data such as images. CNNs are widely used in computer vision tasks including image classification, object detection, face recognition, and facial emotion recognition. They automatically learn and extract important features from images without manual feature engineering. CNNs analyze facial patterns such as eye shape, eyebrow movement, mouth curvature, and muscle variations to classify different emotional states.

- This layer applies filters (kernels) to the input image to extract important features like edges, textures, and patterns.

• Multiple filters generate feature maps that capture various visual characteristics.
• The emotion with the highest probability is selected as the final prediction.

LITERATURE SURVEY

1. Music Recommendation Based on Face Emotion Recognition

https://www.researchgate.net/publication/354855186

ABSTRACT:

This study presents a real-time music recommendation system based on facial emotion recognition. The system uses a Convolutional Neural Network (CNN) trained on the FER-2013 dataset to classify facial expressions into seven emotions. Based on the detected emotion, the system recommends music that aligns with the user's mood. The approach integrates computer vision techniques with a recommendation engine to enhance user experience. Experimental results demonstrate that emotion-based systems provide more personalized and relevant music suggestions compared to traditional methods.

TECHNIQUES

• Convolutional Neural Networks (CNN)
• FER-2013 dataset
• Real-time face detection (OpenCV)

FINDINGS

• Emotion-based recommendation improves user satisfaction
• Real-time emotion detection is feasible with CNN models

CONCLUSION

The integration of facial emotion recognition with music recommendation systems enhances personalization and user engagement.

2. Emotion-Based Music Recommendation System Integrating Facial Expression Recognition and Lyrics Sentiment Analysis

https://www.researchgate.net/publication/39174654

ABSTRACT:

This study proposes an advanced emotion-based music recommendation system that combines facial expression recognition with lyrics sentiment analysis. A CNN model is used to detect emotions from facial images, while natural language processing techniques analyse song lyrics to determine emotional tone. The system integrates both sources of information to generate more accurate and the context aware recommendations. Experimental results show improved accuracy and relevance compared to systems relying solely on facial expressions.

TECHNIQUES

• CNN-based emotion detection
• Natural Language Processing (NLP) for lyrics analysis
• Hybrid recommendation system

FINDINGS

• Combining multiple data sources improves recommendation accuracy
• Hybrid systems outperform single-input approaches

CONCLUSION

Integrating facial emotion detection with additional contextual data such as lyrics enhances the effectiveness of music recommendation systems.

3. Emotion-Aware Music Recommendation Systems

https://www.researchgate.net/publication/3 04358493

ABSTRACT:

This study explores the concept of emotion-aware music recommendation systems that utilize user mood and contextual information. The system considers emotional states along with environmental factors to recommend music. It demonstrates that emotion- driven systems significantly improve user engagement compared to traditional recommendation techniques that rely only on historical data.

TECHNIQUES

• Content-based filtering
• Context-aware recommendation
• Emotion modelling

FINDINGS

• Emotion-aware systems improve personalization
• Contextual information enhances recommendation relevance

CONCLUSION

Emotion-based recommendation systems provide a more adaptive and user-centric approach compared to conventional methods.

4. Facial Emotion Detection and Music Recommendation using Deep Learning

https://www.ijraset.com/research-paper/facial-emotion-detection-and-music-recommendation-using-deep-learning

ABSTRACT:
This research focuses on detecting human emotions using deep learning techniques and recommending music accordingly. The system uses facial image input and processes it through CNN models to classify emotions. Based on the detected emotion, appropriate music is suggested to the user. The system demonstrates the effectiveness of combining computer vision and recommendation systems for improving user experience.

TECHNIQUES
- Deep Learning (CNN)
- Facial image processing
- Emotion classification

FINDINGS
- Deep learning improves emotion detection accuracy
- Emotion-based systems enhance user interaction

CONCLUSION
Combining deep learning with recommendation systems provides an effective solution for personalized user experiences.

5. Music Recommendation Based on Facial Expression using CNN
https://joiv.org/index.php/joiv/article/view/3794
ABSTRACT:
This study presents a CNN-based system for detecting facial expressions and recommending music accordingly. The system uses image preprocessing techniques and feature extraction methods to improve classification accuracy. Experimental results show that the model achieves good performance in emotion recognition and enhances the relevance of music recommendations.

TECHNIQUES
- CNN-based classification
- Image preprocessing
- Feature extraction

FINDINGS
- CNN models achieve reliable emotion detection accuracy
- Emotion-driven recommendations improve user satisfaction

CONCLUSION
Facial expression-based personalised music recommendation systems provide a practical approach for personalized content delivery.

OBJECTIVE
The main objective of this project is to develop a real-time facial emotion recognition system using the FER Python Library to detect emotions from facial expressions. The system aims to perform on-device processing to reduce latency and improve privacy. It also focuses on automatically recommending music based on the user's detected emotional state, creating a personalized and adaptive listening experience.
- Develop a real-time facial emotion recognition system.
- Use the FER library for emotion detection.
- Reduce latency, increase security and remove cloud dependency.
- To design an automatic emotion-based music recommendation system
- Improve user experience through mood- aware interaction

NEED FOR STUDY
Traditional music recommendation systems rely mainly on listening history and cloud-based processing, which results in high latency, internet dependency, and privacy concerns. These systems do not consider the user's real-time emotional state, leading to less personalized and context-aware recommendations. Therefore, there is a need for an edge-based emotion recognition system that can detect emotions instantly and provide automatic mood-based music recommendations while ensuring faster response and improved data privacy.
- Lack of real-time emotion awareness
- Dependence on cloud processing
- Privacy concerns with data transmission
- Need for personalized and context-aware system
- Requirement for low-latency edge processing

To overcome these limitations, an edge-based facial emotion recognition system integrated with music recommendation is necessary.

By detecting emotions such as happy, sad, angry, fear, surprise, disgust, and neutral in real time, the system can provide instant and context-aware playlist suggestions

EXISTING SYSTEM

In most current music recommendation platforms, such as streaming services and smart playlists, recommendations are generated based on user listening history, genre preferences, popularity trends, and collaborative filtering algorithms.

These systems typically analyze past behavior like frequently played songs, skipped tracks, and user ratings to predict what music the user might like next. While this approach improves relevance over time, it does not adapt to the user's current emotional state.

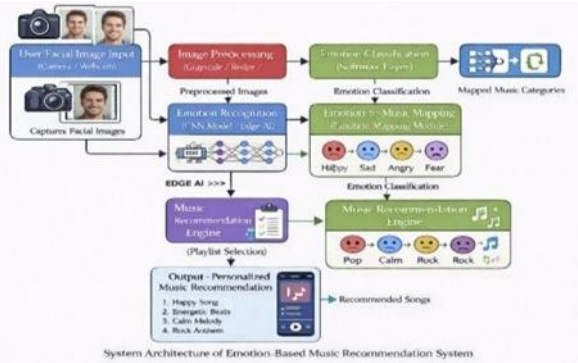
Many existing systems also rely heavily on cloud-based processing, where user data is sent to servers for analysis and recommendation generation. This framework introduces latency, requires continuous internet connectivity, and raises privacy concerns since personal data and preferences are transmitted and stored externally.

Disadvantages

- Latency: Cloud processing can take several seconds to generate recommendations, especially under poor internet conditions.
- Privacy Risks: Personal listening data and behavioral patterns are stored on remote servers.

- No Real-Time Emotion Awareness: Systems do not adjust recommendations based on how the user feels at that moment.
- Generic Recommendations: Suggestions are primarily based on historical data and popularity trends, which may not accurately reflect the user's current mood or context.
- Cloud Dependency: Existing systems rely on cloud-based servers for data processing and recommendation generation, which requires continuous internet connectivity and may cause delays.

SYSTEM ARCHITECTURE



System Architecture of Emotion-Based Music Recommendation System

MODULES

- Image Capture and Preprocessing Module
- Emotion Detection Module
- Emotion–Music Mapping Module
- Music Recommendation Module
- Edge Processing and User Interface Module

1. Image Capture and Preprocessing Module

This module captures live video frames using a webcam and prepares the images for emotion analysis. OpenCV is used for face detection and frame extraction. The detected face is resized and normalized to match the input size required by the FER library model. Proper preprocessing improves accuracy and ensures consistent input quality.

- Real-time video frames
- Detected face region
- Resized and normalized facial images

2. Emotion Detection Module

This module uses the FER (Facial Emotion Recognition) Python library to analyze facial expressions. The FER library contains pre-trained CNN models that classify emotions and provide probability scores for each emotion category. The system identifies the dominant emotion among happy, sad, angry, fear, surprise, disgust, and neutral.

- Preprocessed facial image
- Emotion probability scores
- Dominant detected emotion.

3. Emotion–Music Mapping Module

After detecting the emotion, this module maps the identified emotion to a predefined music category. Each emotional state corresponds to a specific playlist type designed to either match or regulates-music

- Happy → Energetic / Upbeat songs
- Sad → Calm / Soothing songs
- Angry → Relaxing music
- Neutral → Soft background music

4. Music Recommendation Module

This module selects and plays songs based on the mapped emotion. The playlists can be stored locally or accessed using lightweight music APIs. The system dynamically recommends songs according to the user's real-time emotional state.

- Song metadata (title, genre, mood category)
- Audio playback control
- Playlist database

5. Edge Processing And User Interface Module

This module ensures that all operations, including emotion detection and music recommendation, are performed locally on the device. It also provides a user interface that displays the detected emotion and recommended music. Since processing is done on-device, latency is minimized and privacy is enhanced.

- Real-time emotion display
- User interaction controls
- Local processing outputs

PROPOSED SYSTEM

The is an Edge AI-Based Emotion Recognition and Music Recommendation System that uses the FER Python Library to detect human emotions in real time through facial expressions. Unlike traditional cloud-based recommendation systems, this system performs all processing locally on the user's device, ensuring faster response time and improved privacy. The webcam captures facial images, and the FER library analyzes the image using a pre-trained Convolutional Neural Network (CNN) model to classify emotions such as happy, sad, angry, fear, surprise, disgust, and neutral. Once the dominant emotion is detected, the system maps it to a predefined mood-based playlist. The music recommendation module then selects appropriate songs from a local playlist database or integrated API. Because the system operates on edge devices, it minimizes latency, removes internet dependency for emotion detection, and protects sensitive facial data from being transmitted to external servers.

The proposed system creates a real-time, emotion-aware, and context-sensitive multimedia experience that enhances user engagement and supports mood regulation.

Advantages

- Real-time emotion recognition
- On-device (Edge AI) processing
- Reduced latency and faster performance
- Enhanced data privacy and security
- Automatic mood-based music recommendation
- Better human–computer interaction
- Suitable for mental wellness applications
- Improved personalization
- No cloud dependency
- Works efficiently on standard hardware

Hardware Requirements

- Processor : Intel Core i5 or higher
- RAM : 4 GB minimum
- GPU : Provides support for faster CNN training
- Storage : Minimum 20 GB free disk space
- Input Device : Camera/Image dataset
- Output Device : Monitor / Display

Software Requirements:

- Operating System : Windows/Linux
- Language : Python 3.8
- IDE : Jupyter Notebook, VS Code
- Deep Learning

Frameworks : TensorFlow, Keras

- Libraries: NumPy, Pandas, OpenCV
- Models Used : Convolutional Neural Network
- Music Data : Emotion based music database

TECHNIQUES USED IN THE PROJECT

A. CNN-Based Emotion Detection

The facial emotion recognition module uses a pretrained convolutional neural network provided by the FER library. Although pretrained, the underlying CNN operates according to the following mathematical framework.

- Input Representation:

$$X \in R^{H*W*C}$$

- Activation (ReLU):

$$f(x)=\max(0,x)$$

- Predicted emotion:

$$\hat{y} = arg \underbrace{max}_{c} P(y = c \mid X)$$

B. Performance Evaluation Metrics

The performance of the emotion detection model was evaluated using standard multi-class classification metrics.

- Overall Accuracy
- Class-wise Performance
- Precision and Recall Trends
- F1-Score Analysis

C. Mathematical Model for Personalized Music Recommendation

The personalized music recommendation module maps the detected emotional state to corresponding musical attributes using audio feature analysis. The system utilizes Spotify Web API audio descriptors such as valence, energy, tempo, and danceability to

- Each track t_j is represented by audio features:

$$t_j = (v, e, d, t^{tempo})$$

- Emotion-to-Feature Mapping

$$\phi: E \rightarrow R^4$$

D. Edge AI Mathematical Representation

The proposed system follows an Edge AI deployment model in which facial emotion inference is performed locally on the device instead of being transmitted to a remote cloud server. This design minimizes latency, enhances privacy, and reduces network dependency.

CONCLUSION

The proposed Edge AI-Based Emotion Recognition System with Personalized Music Recommendation demonstrates an effective integration of facial emotion detection and intelligent multimedia interaction. By utilizing the FER Python Library, the system is capable of detecting human emotions such as happy, sad, angry, fear, surprise, disgust, and neutral in real time through facial expressions captured via a webcam. Unlike traditional music recommendation systems that depend solely on listening history, genre preference, or popularity trends, this system adapts dynamically to the user's present emotional state. This real-time emotional awareness allows the system to generate context-sensitive and mood-based music recommendations, thereby enhancing personalization and overall user satisfaction.

Unlike traditional music recommendation systems that rely only on listening history, the proposed system adapts dynamically to the user's current mood, creating a more personalized and context-aware experience. The integration of real-time emotion detection and mood-based music selection improves human–computer interaction and supports applications in mental wellness and intelligent entertainment systems. Overall, the system demonstrates an efficient, secure, and user-friendly approach to emotion-aware multimedia recommendation.

FUTURE ENHANCEMENT

In the future, the system can be enhanced by integrating multimodal emotion detection using voice and speech analysis along with facial expressions to improve accuracy Adaptive learning can also be added so the system learns from user feedback and listening behavior over time. Integration with popular music streaming platforms through APIs can expand the music database and improve personalization. Additionally, developing the system as a mobile or web application and adding emotion history tracking can make it more user-friendly and suitable for real-world applications.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

REFERENCES

[1] V.Vijayalakshmi, P. Shrivastav, and G. Thiyagarajan, Facial Expression-based AI System for Personalized Music Recommendations, 2025.

[2] M. Athavle, D. Mudale, U. Shrivastav, and M. Gupta, Music Recommendation Based on Face Emotion Recognition, Journal of Informatics Electrical and Electronics Engineering, vol. 02, no. 02, pp. 1-11, 2021.

[3] Justin Shenk, FER: Facial Expression Recognition Python Library, GitHub repository.
[4] GeeksforGeeks, Facial Expression Recognizer Using FER Library – Deep Neural Networks.
[5] Facial Emotion Recognition In Real Time with OpenCV and FER, Cristian Velasquez, Medium article, Feb. 9, 2024.
[6] D. Bhagat, Facial Emotion Recognition Using Convolutional Neural Networks, 2024. Available:
[7] M. Athavle, D. Mudale, U. Shrivastav, and M. Gupta, Music Recommendation Based on Face Emotion Recognition, Journal of Informatics Electrical and Electronics Engineering, 2021.
[8] C. Velasquez, Real-Time Emotion Recognition in Python using OpenCV and FER, 2024.