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 6, Issue 5, April 2026
Impact Factor: 8.2

An AI-Based Facial Expression-Driven Mood Detection and Music Recommendation System

A Research Paper on Emotion-Aware Personalized Music Recommendation Using Facial Expression Recognition and Web Technologies

Pritam Rangari, Sharvari Shalgar, Nirja Sartale
Department of Information Technology
AISSMS Institute of Information Technology, Pune, India
pritamrangari125@gmail.com, shalgarsharvari@gmail.com, nirjas66@gmail.com

Abstract: *With the explosive emergence of the Artificial Intelligence (AI), it became possible to develop emotion-sensing computer applications that could interact with users intelligently. This paper provides information on how I developed an intelligent Mood Detection and Music Recommendation System based on AI. In the proposed system, the user's current emotional state is detected from their facial expressions captured via a webcam and recommended music based on the emotion detected in real-time. Facial expressions of the user are recognized using a machine learning model trained for emotion detection and classified into one of the four main emotional categories: Happy, Sad, Angry, and Neutral, where each mood is associated with a suitable type of music. External music APIs used to fetch music for each emotion include Spotify Web API, which are dynamically displayed on an interactive dashboard. The front-end of the system comprises HTML, CSS, Tailwind CSS, and JavaScript code while the back-end consists of emotion detection services, music recommendation services, user profile management, and chat services using a REST API framework. Other functionalities of the system include user authentication, user profile, recommendation history, and live chatting services.It can be observed from the system that the concept of emotion-aware computing has immense scope for revolutionizing the domain of digital entertainment by supplying the relevant information based on the psychological state of the users rather than depending upon their past choices to generate digital entertainment. Experimental analysis demonstrates that real-time detection of emotions becomes effective in a bright environment, mood-based song selection, and an interactive interface.*

Keywords: *Artificial Intelligence*

I. INTRODUCTION

The combination of Artificial Intelligence and Human Computer Interaction has brought about a fresh perspective in the domain of personal digital experiences. One such crucial area is affective computing, wherein the primary responsibility of the field is to develop systems capable of detecting, analyzing, and to the feelings expressed by other individuals. Being an exceptionally effective universal tool, music exerts an immense influence on controlling one's mood and emotional state. Despite the increasing popularity of music streaming platforms such as Spotify, Apple Music, and YouTube Music, there has not yet been a breakthrough in the development of more precise recommendation algorithms based on the user's actual emotional state, which, perhaps, is one of the most relevant pieces of information regarding his/her musical preferences at a particular moment. The use of Facial Expression Recognition (FER) techniques appears to be among the most promising directions in recognizing users' feelings. With the latest advances in the field of deep learning and, in particular, Convolutional Neural Networks (CNNs), it is possible to achieve high accuracy of recognizing human emotions based on the analysis of a webcam feed. This ability to detect the emotion along with the music recommendation engines will allow one to develop an intelligent software

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 6, Issue 5, April 2026 Impact Factor: 8.2

solution that will recognize the state of mind of the person and provide the user with the proper music according to their mood without any human intervention whatsoever. The main problem with conventional music recommendation systems is that there is no emotional context at the moment when the recommendation takes place and this causes significant gap between the actual recommendation and user's needs in the certain emotional context [1]. There were attempts to combine FER with recommendation systems, yet a majority of the existing solutions was not aimed at implementing an online interactive web-based platform with real-time recommendations [4][5]. Clearly, there is a need for a fully functional, online and emotion-driven system and this research aims to fill this gap by proposing an AI-based Mood Detection and Music Recommendation system. The algorithm is based on detecting facial expressions via normal webcam, processing the image via emotion detection models, identifying emotions and selecting the appropriate category of music from Spotify using its API. The web-based interface delivers a smooth experience, including authentication, an animated dashboard, mood control, live chat, and profile management. This is a perfect example of how affective computing can transform digital entertainment and personalized content distribution.

II. LITERATURE REVIEW

As can be seen from the previous review, there is extensive research on FER in recommender systems. The following table gives an overview of the important papers on the subject and their contributions related to the problem under discussion and directly relevant to the design of our proposed project. The main gap in research identified through the literature review is that while individual aspects of the technology, like facial emotion detectors, deep learning models, and music recommendation API's, have been made independently in plenty, not enough research has been conducted to develop a system which combines all these parts in an interactive web-based solution. Highly accurate models (¿ 90) have proved to be very slow [5][9] or need high computational power, while lighter solutions often sacrifice accuracy for speed [8]. Literature on using social media sentiment analysis poses major privacy issues, and most of the reviewed projects fail to include important features such as user authentication and real-time chat. Our solution tries to fill this gap by combining a convolutional neural network based facial emotion detector and the Spotify API into a fully functional web application.

III. PROBLEM STATEMENT

In the existing modern music streaming platforms, recommendation engines rely on historic user data such as past consumption trends, user ratings and collaboratively filtered interests. While these approaches bring some element of personalization to their operations, they are inherently reactive rather than adaptive. It means that the recommendations are based on what the user did in the past rather than how he/she feels now. This limitation is extremely crucial for the creation of personalized digital music experiences. Some specific problems in the current systems that the suggested project aims to overcome include:

- The current solutions are unable to capture and account for the actual emotional state of the user and therefore cannot make recommendations according to it.
- The existing music streaming platforms have no capability to use a webcam camera to track facial emotions and provide corresponding recommendations.
- Solutions that base on historic behavioral patterns cannot be adaptable in cases of the change of mood, thus making the users less satisfied.
- Lack of sophisticated web applications based on artificial intelligence mood detection technologies with built-in interactive features, such as authentication, profile, chat and dynamic dashboards.
- The existing recommendation models based on emotions do not provide a sufficiently high-quality user interface to be released into production and integrated with popular music websites and platforms.
- These problems lead to the development of an algorithm based on computer vision and artificial intelligence, which analyzes emotions in real time and makes suggestions for music according to the particular context, providing a real-time update and a convenient user interface.

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 6, Issue 5, April 2026
Impact Factor: 8.2

IV. PROPOSED SYSTEM

The proposed Mood Detection and Music Recommendation System is a full-stack, AI-enabled web app that will be used to customize the delivery of music content based on the emotional state of the user being detected. In contrast to traditional music recommendation systems, the system does not rely on past records to make any recommendations but instead takes into consideration facial expressions taken in real-time using a web camera. This guarantees that the suggestions will always be in line with the mood of the user as opposed to the previous behavior. The main functionality of the system is structured around five key features: user verification, real-time facial emotion recognition, mood-to-music mapping, personalized song recommendation, and interactive dashboard. When logging in, the users receive access to a dynamic dashboard where the mood detection feature can be turned on. The webcam captures a frame of an image that is live and it is sent to the backend server in a base64-coded message. The picture is sent to the emotion recognition model from the backend using the /predict endpoint, which will return one of the following four categories of emotion: Happy, Sad, Angry, or Neutral. The program, after detecting the emotion, uses a predefined mood mapping algorithm for correlating the detected category of emotion to an appropriate genre of music. For instance, Happy emotions are mapped to vibrant and lively music, Sad to relaxed and calming music, while Angry to relaxing and stress-free music, while Neutral to a balanced type of music. The /songs endpoint requests the Spotify Web API to provide a list of songs/albums based on the mapped music category. These results will be dynamically displayed on the user dashboard depending on the mood label detected. Additionally, the program is capable of tracking the facial expression in real-time and updating the music recommendations when there is a change in emotional state. Apart from the ability to detect moods and recommend music based on the same, the application has a chat module that allows users to interact with the system and a profile management component where users can keep track of their preferences and recommendations.

V. SYSTEM ARCHITECTURE

The system is designed as a six-layer framework to ensure modularity, scalability, and efficient communication between all components. The architecture can be represented as follows:

User → Frontend Layer → Backend Layer → Database Layer → External API Layer → Output Layer

It appears in conjunction with the registration form or the login screen. The dashboard is reached after authentication, and all the functions of the system are accessible. The entire process of user interactions and the visual representation takes place through the Frontend Layer. The Frontend Layer, which is constructed using HTML, CSS, Tailwind CSS, and JavaScript, displays the login and registration screens, the dashboard, the mood detection screen, the music recommendations screen, the user's profile screen, and the live chat screen.

Backend Layer acts as the processor and controller for the entire system. It is developed using RESTful API architecture. The back end receives images from the front end, invokes the emotion detecting model, and returns the music recommendation and performs all database transactions. The most common API endpoints are /predict, to identify emotions; /songs, to return recommended music; /chat/send and /chat/fetch, to utilize the messaging module; and conventional authentication endpoints, to register and log in as users. Persistent application data is stored in the Database Layer through SQLite (mood app.db). User data, user profiles, history of emotions identified, history of suggested songs, and chat logs are kept in the database layer. This enables the system to offer personalized experience even after logging out. The External API Layer is the combination of Spotify Web API with the emotion detection model using artificial intelligence. Emotion detection model receives a

Sr.	Title / Author / Year	Focus Area	Methodology	Dataset	Key Results	Challenges
1	Smart Music Player Integrating	Intelligent music player	ANN	Facial images	Automated mood	Limited classes; light sensitivity

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 6, Issue 5, April 2026 Impact Factor: 8.2

	FER – Gilda et al. (2017)				rec	
2	Emotion Based Music Rec. – Iyer et al. (2017)	Mood enhancement	Viola–Jones, OpenCV	Facial images	Effective mood rec	Lower accuracy than deep learning
3	FER and Music Rec. System – Gadagkar et al. (2024)	Real-time rec	MobileNetV2	Live streams	High efficiency	Dependency on cam-era quality
4	FER Music System Using DCNN – Ashwini et al. (2024)	Deep learning detection	Deep CNN	Image datasets	Improved robustness	High compute requirements
5	Rec. System Using FER – Parmar (2026)	Emotion-aware rec	Feature extraction	Live images	Personalization	Privacy; real-time overhead
6	Movie & Song Recommender System (2023)	Multimedia rec	classifiers	Image dataset	Better engagement	Lighting sensitivity
7	Smart Sound: AI Face Detection (2025)	Real-time rec	Deep Learning	Live feed	Adaptive rec	High computational cost
8	FER Music Rec. Social Integration (2024)	Multimodal rec	FER + Sentiment	Images + Social text	Better context	Social data dependency
9	FER Music Rec. Using DL (2025)	detection	CNN	Public datasets	High accuracy	Training complexity
10	CNN-based Emotion & Music System	rec	CNN + Data Aug	FER-2013	>90% accuracy	High power requirements
11	Emotion-based Music Rec.	Trad. ML rec	HOG/LBP + SVM	JAFFE / CK+	75–85% accuracy	Less robust than DL
12	Deep Learning + Rec. Framework	Hybrid framework	VGG, ResNet	FER-2013	~90% accuracy	Complex architecture
13	Lightweight Real-time System	Mobile support	Light CNN+MTCNN	FER-2013	Fast inference	Reduced accuracy
14	CNN-based Music Rec. System	rec	Deep CNN	FER-2013	95.9% training acc	possible overfitting
15	Hybrid Facial Emotion Model	Integrated rec	CNN + Engine	Mixed / Live	User satisfaction	Difficulty with mixed emotions

Table 1

Emotion detection model receives a facial image encoded in base64 and returns the emotion state that was detected. The API, in turn, retrieves a list of recommended songs and playlist and the songs are then provided according to the category of music matched from the mood map, thus providing recommendations for appropriate music within an appropriate context. Finally, the Result Layer provides the output from the system to the user through the front-end

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 6, Issue 5, April 2026
Impact Factor: 8.2

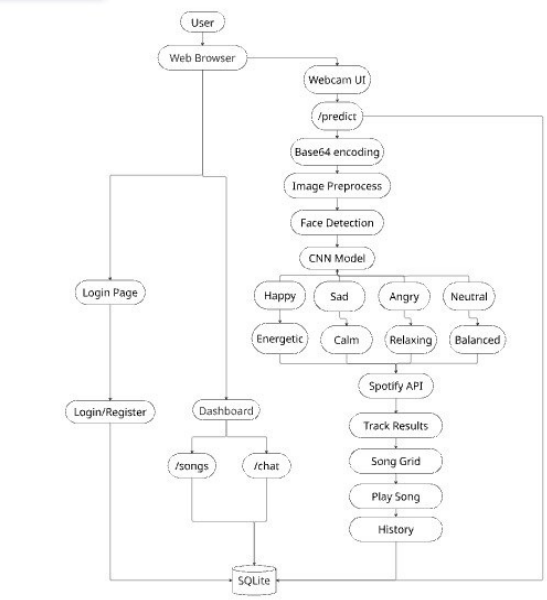
dashboard. Both the emotion detected and the recommended songs are displayed through SQLite (mood app.db). User data, user profiles, history of emotions in a nice and animated graphical interfacein a nice and animated graphical interface. In real-time, as soon as a change in emotion is detected, the system recommends the appropriate songs instantly.

VI. METHODOLOGY

The architecture entails a stepwise procedure involving user authentication, image acquisition, preprocessing, emotion detection, moods classification, music recommendation and visualization of results. These processes are described in details below.

- Authentication Process: To use the system, the user logs in through a login or registration portal. The user's login credentials are verified in the process and upon successful verification, the user is redirected to the dashboard.
- Starting the Process: In the dashboard, the user clicks on the 'Start Mood Detection' button to activate the mood detection service. The device camera is activated using WebRTCgetUserMedia and the system begins capturing images from the feed.
- Frame Capture & Image Encoding: Using a front-end JavaScript code, a single frame is captured from the feed provided by the camera. After that, this frame is encoded into a base64 string and sent in the form of JSON payload through an HTTP POST request to the predict API endpoint.
- Preprocessing Stage: Before an image passes to the neural network for detection of emotions, it undergoes preprocessing which entails scaling down to a specific size required by the model, greyscaling or normalization, noise filtering and irrelevant background elimination.
- Emotion Recognition: In the process of identifying emotions, the AI-based facial emotion recognition system evaluates the image. This model evaluates key elements in the face, such as eyebrow placement, eye aperture size, nasal labial fold, and lip movement.
- Mood Assignment: After identification of the mood, suitable genre of music is selected by using pre-defined strategies: Happy mood -> Music that is energetic and cheerful; Sad mood -> Music that is calm and soothing; Angry mood -> Music that is relaxing and stress relieving; Neutral mood -> Music that is balanced and light.

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 6, Issue 5, April 2026
Impact Factor: 8.2



System Architecture

VII. IMPLEMENTATION

7.1 Frontend Implementation

The front-end part of the application is designed as a one-page web application using technologies such as HTML5, CSS3, Tailwind CSS, and vanilla JavaScript. The UI contains 5 key views namely, Login Page, Registration Page, Main Dashboard, Profile Management Page, and Chat Interface. These views are all responsive as well as visually appealing. After successfully logging in, users are directed to the dashboard where the central functionality of the application lies. The webcam video is accessed via WebRTCgetUserMedia API, and frames are decoded into images using HTML5 Canvas and encoded into base64 format.

7.2 Backend Implementation

The backend is an API server that utilizes a REST API architecture and enables request routing, coordination of emotion detection, recommendation of music, and management of users' information. The crucial API endpoints include the following: The /predict endpoint will receive the image encoded in base64 format and will deliver the label of the detected emotion; the /songs endpoint will make a request to the Spotify API and provide the list of appropriate songs.

7.3 Implementation of the Database

The database management system used by the application involves storing the database in a local file called mood app.db. This schema includes information on users (including usernames, passwords, email IDs, and other user-related details), mood detection information along with their time stamps and user emotions, songs recommended by the application, and messages from the chat feature of the application.

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 6, Issue 5, April 2026
Impact Factor: 8.2

7.4 Mood Detection Implementation

Mood detection begins with a process that takes place on the client-side. To be more precise, a video frame will be captured using the webcam, then displayed on the HTML5 canvas element and converted into a Base64 string of a JPEG image. The latter will act as an input for fetch() API request.

7.5 Music Recommendation Implementation

Once the emotion name has been received, then the recommendation system kicks in by sending a request at the /songs API endpoint with the mood property. Once that request has been received at the backend, it is converted to suitable filters for the Spotify API call, whereupon the API is called using OAuth 2.0 access tokens.

7.6 Chat System Implementation

The chat application runs in real-time using a polling technique. The JavaScript on the client's side sends repeated calls to the server at regular intervals to check whether there is any new message available. To send messages, users can use a POST method call to the server at the /chat/send address.

VIII. RESULTS

Development and validation of the Mood Detection and Music Recommendation Application proved to be successful, especially concerning the functionality of the application. Results of empirical studies revealed that the emotion recognition function was capable of distinguishing between four predetermined emotions: Happy, Sad, Angry, and Neutral. Prompt results were obtained after the image acquisition with no signs of latency.

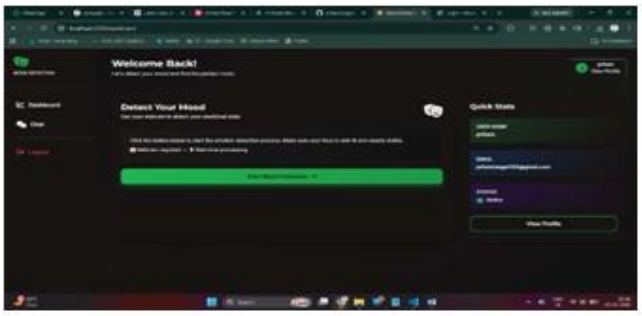
TABLE II: SYSTEM PERFORMANCE AND COMPARISON

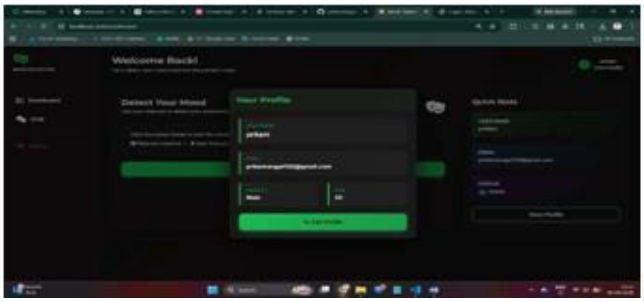
Parameter	Existing Systems	Proposed System	Improvement
Recommendation Ba-sis	Listening history, trends	Real-time facial emotion	Contextual & adaptive
Personalization	Generic playlists	Mood-matched songs	Highly personalized
User Interaction	Search/browse	Webcam + dashboard	Automatic
Real-time Adaptation	Not supported	Continuous monitoring	Dynamic updates
Emotion Categories	None	Happy, Sad, Angry, Neutral	4 moods mapped

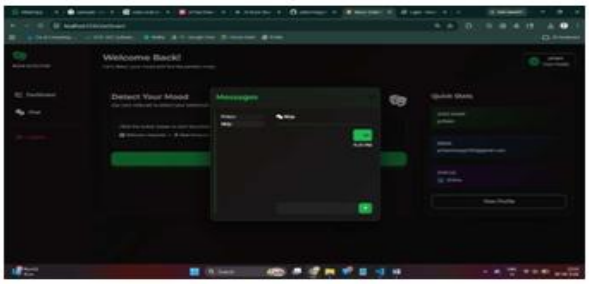
Output:

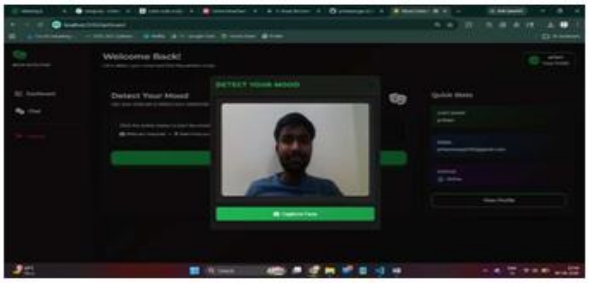
DOI: 10.48175/IJARSCT-32848

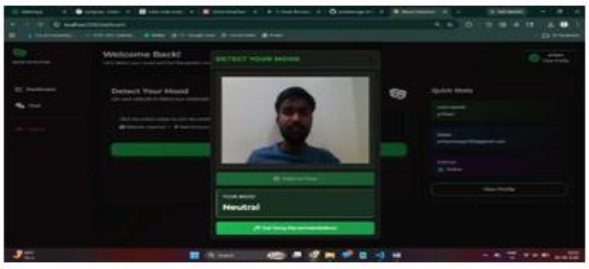
IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 5, April 2026
Impact Factor: 8.2





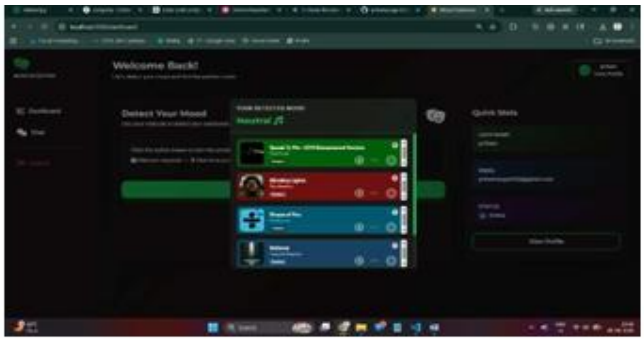






DOI: 10.48175/IJARSCT-32848

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 6, Issue 5, April 2026
Impact Factor: 8.2



IX. FUTURE SCOPE

Future Implementations: In the future, the application could leverage more sophisticated architectures like Vision Transformers (ViTs) or CNNs using an attention mechanism trained on bigger and diverse datasets for more precise classifications of emotions.

Multi-modal Emotion Detection: The system could also be developed further by incorporating different emotional data types such as speech emotions detection, typing behavior detection, and detection of physiological states.

Mobile Applications: Creating native applications for mobile devices will allow the application to be available to more users and enable offline emotion detection by conducting model inference on the device itself.

Support for More Music Services: The inclusion of YouTube Music, Apple Music, and Deezer could reduce dependence on any third-party services.

Monitoring of Mental Well-being: The system could detect prolonged instances of negative emotion reactions and offer appropriate wellness advice in order to align with future digital health applications.

Personalized Emotion Mappings: Machine learning approaches could be used for learning individual emotional preferences and adapting the mapping algorithm accordingly.

Deployment in Cloud: Horizontally scaling in cloud-based architectures such as AWS or GCP will enable simultaneous access by multiple users and ensure high availability.

FER that Protects Privacy: The privacy-preserving aspect should be explored further for future work, including performing emotion recognition on-device or utilizing federated learning to minimize facial image uploads.

X. CONCLUSION

Mood Detection and Music Recommendation System has been designed, structured, and implemented in the current paper. In particular, the proposed software helps address the main limitation of the classical methods of music recommendations, which is a lack of emotion context in real-time. Through a combination of computer vision, machine learning/deep learning approaches, and the Spotify Web API, a full-stack web application has been developed. The conducted experiment shows that the system yields accurate results in normal circumstances.

REFERENCES

[1]. P. Ekman and W. V. Friesen, Facial Action Coding System (FACS), 1978.
I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning, 2016.
A. Mollahosseini et al., ”Facial Expression Recognition using DNNs,” WACV, 2016.
[2]. S. Li and W. Deng, ”Deep Facial Expression Recognition,” IEEE Trans., 2020.
[3]. B. Bakariya et al ”FER and Music Rec. System using CNN,” Springer, 2023.
[4]. ”Emotion-basedMusic Recommendation System,” IEEE ID: 8299738, 2018.
[5]. ”A Deep Learning Approach for FER,” IEEE ID:10134389, 2023.
[6]. ”Real-time FER and Sentiment Music Suggestion,” IEEE ID: 10690441, 2024.

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 6, Issue 5, April 2026
Impact Factor: 8.2

[7]. ”FER Based Music System using CNN,” ScienceDirect, 2022.
[8]. ”Movie & Song Recommender System,” ICAC3N,IEEE,2023
[9]. https://ieeexplore.ieee.org/document/10541832
[10]. ”Smart Sound: AI Face Detection,” ICIRCA, IEEE, 2025. https://ieeexplore.ieee.org/document/11089793
[11]. ”FER Music Rec. with Social Integration,” ICCIRT, IEEE, 2024. https://ieeexplore.ieee.org/document/10922070
[12]. ”Emotion-Based Music Rec. using DL,” ICIDeA, IEEE, 2025
[13]. https://ieeexplore.ieee.org/document/10963221
[14]. ”Improved FER using YOLO and DeepFace,” ICESC,
[15]. IEEE,2022. https://ieeexplore.ieee.org/document/9885456
[16]. Spotify Developers, ”Spotify API,” https://developer.spot ify.com/documentation/web-api/
[17]. Mozilla, ”WebRTCAPI,” https://developer.mozilla.org/ [17] OpenCV, ”CV Library,” https://opencv.org/
[18]. Tailwind CSS, ”CSS Framework,” https://tailwindcss.co m/docs