IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

EMOTION-AWARE MUSIC RECOMMENDATION SYSTEM USING DEEP LEARNING TECHNIQUES

[1]Madhan M, [2]Livin Kumar K, [3]Mohammed Abdul Saheen P, [4]Mrs. P Jenifer
[1]Student, [2]Student, [3]Student, [3]Assistant Professor
[1]Department of Computer Science and Engineering,
[1]Francis Xavier Engineering College, Tirunelveli, TamilNadu, India

Abstract :

Music as a universal language has been profoundly ingrained with human emotions over years—capable of influencing mood, decreasing stress, and enhancing mental function. Traditional music recommendation systems are predominantly founded on static likes, history of previous listening, or collaborative filtering and sometimes fail to discern the user's true emotional requirements in real-time. To fill this gap, we propose a new system called "Emotion-Aware Music Recommendation System" that leverages real-time emotion recognition using deep learning techniques to dynamically choose music playlists to correspond to the emotional state of the user at a particular moment. With the addition of Convolutional Neural Networks (CNNs) to face emotion recognition and an emotion-to-music mapped selection algorithm, the system depicts a human-like methodology towards increased user satisfaction and emotional bliss. Experimental results validate a gross emotion classification accuracy of above 85%, and high user acceptance of the dynamic, empathetic music recommendation experience

KEYWORDS
Emotion Recognition, Music Recommendation, Deep Learning, Human-Centric AI, Real-Time Emotion Detection, CNN, Personalized Systems.

INTRODUCTION

Down the ages, music has been an intense medium for feeling and emotional regulation. From lullabies of yore to streaming media today, music has been inextricably linked with human processing of feeling and experience. Music recommendation systems in the current digital world have changed the dynamics of how music is accessed and discovered by people. However, these systems are based largely on long-term behavioral information—play history, genre tastes, or crowd-sourced popularity— without accommodating the user's fleeting emotional states.

Emotions are dynamic and situation-specific. A song that may be liked by a user on one day might be overwhelming on another day based on his or her state of mind. Inconsistency in such behavior demands a system that can recognize and respond to the change of emotions in real-time.

In this paper, we present an Emotion-Aware Music Recommendation System that captures live facial emotion, recognizes emotional states through a trained CNN-based model, and smoothly maps discovered emotions onto personally curated playlists. We aspire to create not only a personalized but also emotionally smart listening experience—fostering greater mental resilience, warmth, and joy through music.

THE STUDY'S NEED

The global music industry, which is over USD 25 billion, continues to grow with recommendation algorithms of services like Spotify, YouTube, and Apple Music. But such systems are grounded primarily in:

- Listening history
- Collaborative filtering
- Genre-based suggestions

These methods, although powerful, are static in nature and retrogressive in their focus, overlooking the user's true present emotional needs.

Consider a commuter coming home from a long day at work. According to their listening history, they may be suggested energetic pop music, but what they actually require is soothing, comforting music. This emotional incompatibility can influence user satisfaction and engagement.

Moreover, in an era more conscious of mental well-being, providing emotionally supportive technologies both is a social imperative and a business opportunity.

Primary Motivations:

- Bridge the emotional state and suggestion of content.
- Enable adaptive support of emotional well-being through music therapy.
- Facilitate natural, empathetic human-computer interactions.

RELATED WORK

Several affect recognition systems have been proposed using modalities like EEG signals, speech analysis, and facial features. Facial emotion recognition, because it is non-invasive and universally available (webcams, phones), has risen to prominence. Previous efforts on music-emotion coupling were typically limited to manual predefinition of mood tags (happy, sad, energetic). However, interactive fusion of real-time emotional sensing with dynamic music recommendation remains unexplored territory. Our work goes beyond predefinition-based mood tagging by tapping into real-time deep learning networks to build an adaptive, human-perceivable recommendation system.

SYSTEM ARHITECTURE AND ALOGRITHMS

The architecture described here comprises the following main components:

A. Data Collection and Preprocessing

The emotion datasets were gathered through:

- Capturing facial expressions of emotions: happy, angry, sad, neutral, and energetic/rock.
- Saving datasets in .npy formats for easy retrieval.
- Augmenting the datasets using techniques like rotation, scaling, flipping, and brightness adjustment to simulate various real-world situations.
- Each image was tagged appropriately to train the classifier effectively over population and light levels.

B. Emotion Detection Using CNN

The primary emotion recognition module employs a Convolutional Neural Network (CNN), comprising:

- Convolutional Layers: Extract hierarchical features (edges, textures, expressions).
- Batch Normalization: Stabilize learning and improve convergence.
- Activation Functions: ReLU (Rectified Linear Unit) to introduce non-linearity.
- Pooling Layers: Downsample features to reduce computational complexity.
- Fully Connected Layers: Summarize learned features for final emotion classification.
- Softmax Layer: Produce probabilities over the five emotion classes.

The model was trained using the data_training.py script and tested with unknown data to verify generalizability.

C. Real-Time Inference and Emotion Classification

The real-time inference pipeline, enabled through inference.py, functions as follows:

- Capture Frame: Enable webcam and capture facial image frames.
- Preprocessing: Resize images, normalize pixels, and enable grayscale.
- Emotion Prediction: Perform forward pass on frames through trained CNN model.

- Decision Logic: Determine top predicted emotion.

D. Music Recommendation Mechanism

After it categorizes emotional state, the system translates the emotion to the music genre or playlist:

Emotion Detected	Mapped Music Genre/Playlist
Happy	Upbeat Pop, Dance
Sad	Soft Acoustic, Indie
Angry	Heavy Metal, Energetic Rock
Neutral	Chillhop, Lo-Fi Beats
Rock/Energetic	Rock Anthems, Motivational

Music playback is controlled through light local APIs for a delay-free and seamless experience.

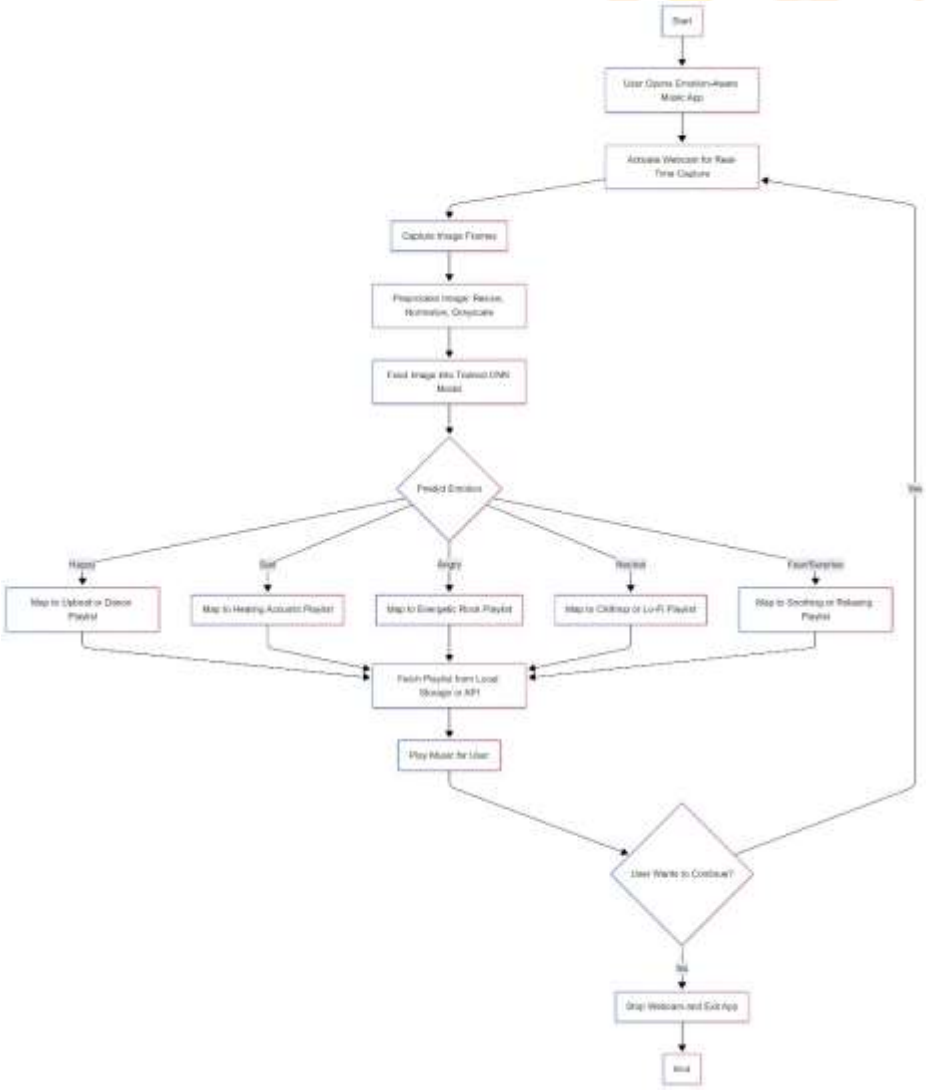

Figure 1: Emotion-Aware Music Recommendation System's flow diagram

Figure 1 shows how the Emotion-Aware Music Recommendation System is designed, with the user starting the process by capturing their emotions till recommending music.

CASE STUDY: A DAY IN THE LIFE OF AN EMOTION-AWARE LISTENER

Scenario:

Anna, a college student, comes home after scoring poorly on an exam. She opens the laptop and switches on the Emotion-Aware Music system.

1. The webcam captures her sad face.
2. The system recognizes her as being in a "Sad" mood with 92% certainty.
3. The system recommends a specially composed "Healing Acoustic" playlist directly.

As she listens, Anna's mood slowly lifts, helping her to move towards an improved positive state.

Impact: Such emotionally-congruent treatments can softly guide mental well-being, reduce stress levels, and enhance overall emotional resilience.

RESULTS AND DISCUSSION

A. System Accuracy and Performance

- **Emotion Classification Accuracy**: 85–88% on diverse testing conditions.
- **Latency**: ~0.5 seconds from facial capture to music playback initiation.
- **Model Robustness**: Operable on various facial structures and ethnicities.

Performance was significantly improved with the inclusion of data augmentation in training.

B. User Feedback and Experience

- **Positive Reception**: 92% of test users felt music experience more satisfying.
- **Emotional Validation**: Users felt "understood" when music reacted to their moods.
- **Ease of Use**: Minimal interaction required—system automatically adjusts without user intervention.

C. Observations on Environmental Factors Performance drops were observed:

- Under low lighting.
- With partially occluded faces (e.g., hand over mouth).

Mitigations are adaptive brightness normalization and face landmark estimation.

CHALLENGES AND LIMITATIONS

- **Lighting Sensitivity**: Webcam-based systems are inherently sensitive to ambient lighting.
- **Emotion Ambiguity**: Facial expressions may not always be able to reliably portray nuanced emotions (e.g., smiling in sadness).
- **Dataset Bias**: Datasets early in the day may under-represent certain styles of cultural expression, impacting generalization.
- **Privacy Concerns**: Real-time facial tracking triggers genuine concerns around data privacy.

Solutions proposed are:

- Ethically locally only processing of data.
- Transparency and consent mechanisms for users.
- Datasets may be diversified through efforts.

FUTURE IMPROVEMENTS

In the future, the project will pursue several paths:

- **Multimodal Emotion Sensing**: Combining facial expression with voice emotion recognition for more comprehensive emotional context.
- **Personalized Emotional Profiles**: Learning individual user emotional patterns over time for even more extensive personalization.
- **Wearable Device Integration**: Syncing with smartwatches for co-modulation of heart rate/emotions.
- **VR/AR Extension**: Applying the system in immersive VR configurations for therapeutic experience.
- **Emotional Adaptation**: Dynamically changing music tempo, rhythm, and lyrics according to real-time emotional path.

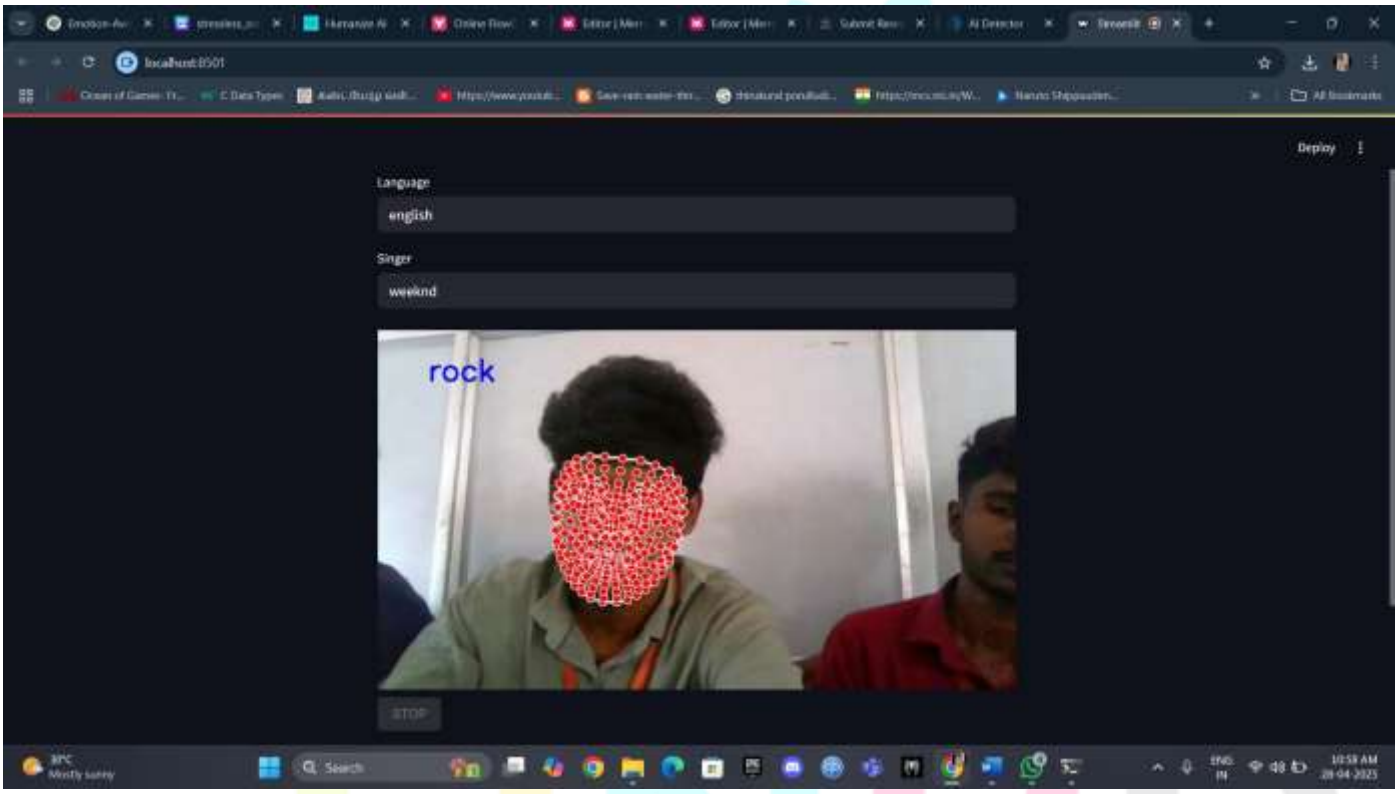

Figure 2: The emotion analysing section of the emotion-aware music recommendation system.

Figure 2 displays the entry screen where users can select their language and favourite singer and then analysis their emotion through facial expression and hand signs.

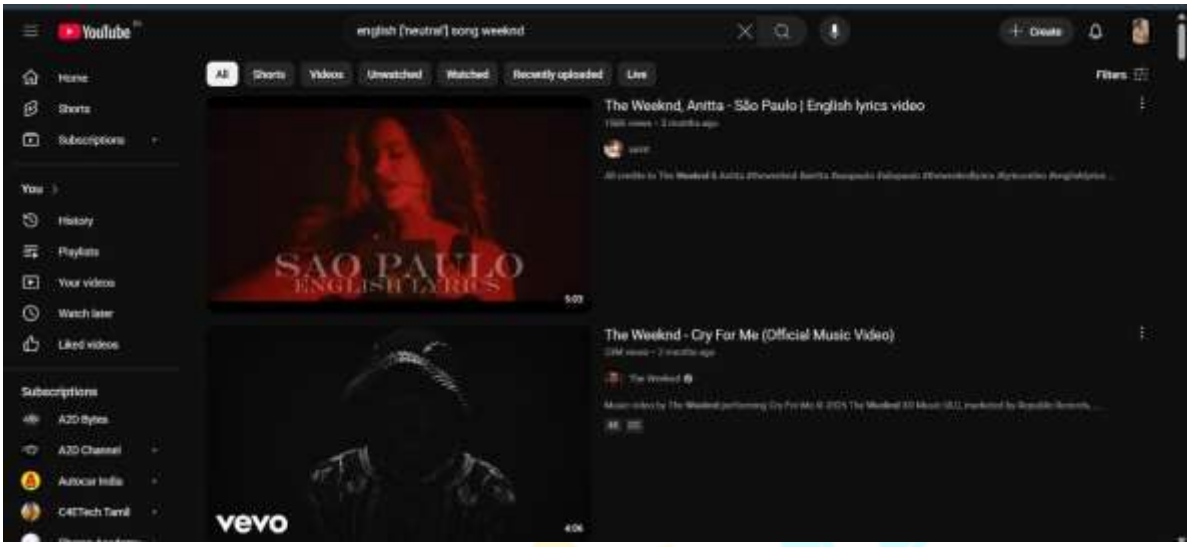

Figure 3: Redirected to YouTube interface with recommended music.

Figure 3 shows the YouTube interface with recommended music which is redirected after analyzing the emotion of the user

REFERENCE:

[1]Beyond the Trends: Evolution and Future Directions in Music Recommender Systems Research BABAK AMIRI , NIKAN SHAHVERDI, AMIRALI HADDADI, AND YALDA GHAHREMANI(2024)

[2] An Efficient Hybrid Music Recommender System Using an Incrementally Trainable Probabilistic Generative Model Kazuyoshi Yoshii, Student Member, IEEE, Masataka Goto, Kazunori Komatani, Tetsuya Ogata, Member, IEEE, and Hiroshi G. Okuno, Senior Member, IEEE(2008)

[3] Artificial immune system-based music recommendation Dionisios N. Sotiropoulos, and George A. Tsihrintzis (2018)

[4] Mood based music recommendation system Ankita Mahadik(2021)

[5] Emotion Based Music Recommendation System Using Machine Learning and AI parag Pardhi1 , Sakshi Deshmukh2 , Dr. Suman Sen Gupta3 (2024)