Song Recommendation System Based on Facial Emotion

Mr. Aswin Jeba Mahir A[1], Mr. Dinesh K[2], Mr. Arjun R[3], Mr. Dheenadhayalan A[4]

[1]Assistant Professor, Department of Information Technology, SRM Valliammai Engineering College, Kattankulathur, Tamil Nadu, India.

[2,3,4]Student, Department of Information Technology, SRM Valliammai Engineering College, Kattangulathur, Tamil Nadu, India.

Email Id: *jebamahir@gmail.com[1], dinesh033g@gmail.com[2], thearjun207@gmail.com[3], dheenadhayalen@gmail.com[4]*

Abstract

This research is aiming to enhance the user experience in music consumption by incorporating real-time facial emotion analysis. Emotions play a fundamental role in shaping individual preferences, and leveraging facial expressions as a means of understanding user's emotional states can significantly contribute to personalized music recommendations. Our proposed system begins by capturing real-time facial expressions using a webcam or analyzing static images. These facial expressions are then processed through a CNN-based emotion recognition model trained to classify emotions such as happiness, sadness, anger, and more. The CNN model extracts high-level features from facial images, enabling accurate emotion recognition. Using the detected emotional state as input, our system employs a recommendation algorithm tailored to the user's current emotional state to suggest relevant music or videos from YouTube.

Keywords: *OpenCV; CNN; Streamlit; TensorFlow*

1. Introduction

The facial emotion-based song recommendation system consists of three main components: the face recognition system, the emotion classification system and the song recommendation system.[1] The facial emotion recognition system captures a video of the listener's face and extracts facial features such as eye movement, eyebrow position, and mouth shape. These features are then used to determine the listener's emotional state using a trained machine learning model. The song recommendation system uses a recommendation algorithm to generate song recommendations based on the listener's emotional state. The database consists of different emotions (happy, sad, angry, neutral). By integrating facial emotion analysis with the YouTube platform, we aim to create a user-friendly and immersive experience.[2] Users can receive real-time music recommendations based on their facial expressions while navigating through videos or engaging with the content. This system goes beyond traditional recommendation methods, offering a more dynamic and emotionally resonant interaction with the music content available on YouTube.[5] By employing sophisticated deep learning architectures such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), or transformer-based models, we seek to capture the intricate interplay between audio features and emotions. Furthermore, our system will prioritize user interaction and feedback, allowing users to express their emotional preferences explicitly [3].

2. Method

2.1. Facial Emotion Recognition System Video Capture

Capturing video using OpenCV and TensorFlow involves utilizing the OpenCV library in Python to access and process frames from a video source, such as a webcam.[4] To capture the video, we need to install OpenCV using pip install OpenCV-python and import the OpenCV library using import cv2 and to initialize the video capture, create a VideoCapture object to access the video source, which can be a webcam (0 for default webcam) by

using cap = cv2. VideoCapture (0)

2.2. Data Pre-Processing

CNN (Convolutional Neural Network)

Data preprocessing for Convolutional Neural Networks (CNNs) is a crucial step to ensure that the input data is in a suitable format for the model.

Load your image dataset, ensuring that images are organized into appropriate folders (e.g., one folder for each class). Augment your dataset by applying transformations to the images, such as rotation, flipping, zooming, and changes in brightness or contrast. This helps the model become more robust and generalize better to different variations of the input data. Resize all images to a consistent size. This ensures uniformity in the input dimensions and reduces computational complexity.

2.3. Emotion Classification

Train a machine learning model (e.g., Convolutional Neural Network - CNN) using a dataset of annotated facial expressions is show in Figure 1. Integrate the trained model to classify the user's emotional state based on extracted facial features. The system identify emotion such as happy, sad, angry, surprise, neutral. And an additional feature rock which is recognized by hand movement in Figure 2.
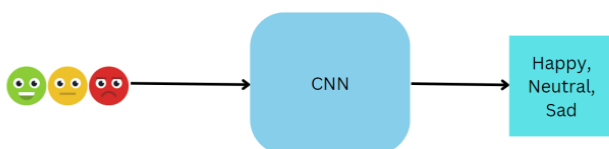

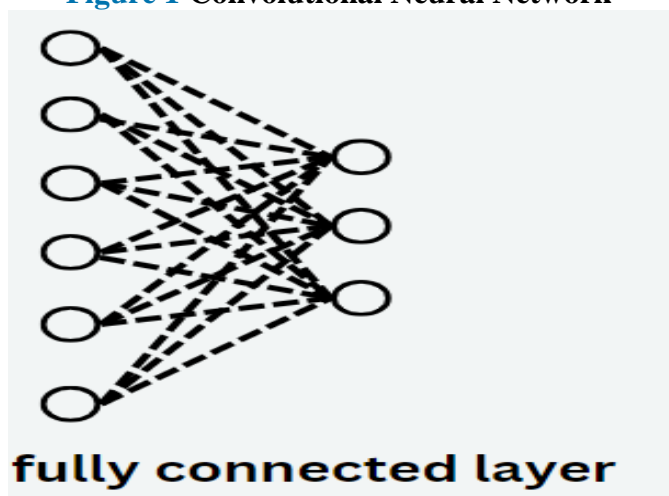**Figure 1 Convolutional Neural Network**

Figure 2 Fully Connected Layer

3. Results and Discussion

The facial emotion-based song recommendation system successfully decodes user's emotional states using their facial expressions.[6] Real-time recommendations, dynamically tailored to detected emotions, offer an engaging and personalized music experience. The YouTube API integration enhances the system's music library diversity, allowing seamless navigation through content.[7] Privacy measures and user controls prioritize ethical usage. Continuous learning mechanisms refine recommendations over time, ensuring adaptability and personalization. Challenges, including cultural diversity considerations and potential biases, highlight avenues for future improvements. Positive results indicate a promising future for emotion-aware music recommendation systems.

Conclusion

In conclusion, the facial emotion-based song recommendation system presents a robust and innovative approach to redefining user interactions with music streaming platforms. Through the effective integration of facial emotion recognition, real-time recommendation algorithms, and seamless user interfaces, the system successfully decodes users' emotional states and provides dynamic, personalized music suggestions. The utilization of technologies like OpenCV, TensorFlow, and MediaPipe, coupled with the YouTube API integration, broadens the system's capabilities, ensuring a diverse and immersive music library. Privacy measures and user controls prioritize ethical considerations, fostering trust and user engagement. The continuous learning mechanisms further elevate the system's adaptability and refinement over time. Despite challenges, such as cultural diversity and potential biases, the positive results and user-centric design underscore the system's potential to revolutionize the music discovery experience. The facial emotion-based song recommendation system stands at the forefront of AI-driven music interaction, promising a future where music recommendation systems are intricately attuned to the emotional preferences of individual users.

References

[1]. Anagha S.Dhavalikar and Dr. R. K. Kulkarni, "Face Detection and Facial Expression Recognition System" 2014 Interntional Conference on Electronics and Communication System (ICECS -2014).

[2]. Yong-Hwan Lee, Woori Han and Youngseop Kim, "Emotional Recognition from Facial Expression Analysis using Bezier Curve Fitting" 2013 16th International Conference on Network-Based Information Systems.

[3]. Arto Lehtiniemi and Jukka Holm, "Using Animated Mood Pictures in Music Recommendation", 2012 16th International Conference on Information Visualisation.

[4]. F. Abdat, C. Maaoui and A. Pruski, "Humancomputer interaction using emotion recognition from facial expression", 2011 UKSim 5th European Symposium on Computer.

[5]. T.-H. Wang and J.-J.J. Lien, "Facial Expression Recognition System Based on Rigid and Non-Rigid Motion Separation and 3D PoseEstimation," J. Pattern Recognition, vol. 42, no. 5, pp. 962-977, 2009.

[6]. Renuka R. Londhe, Dr. Vrushshen P. Pawar, "Analysis of Facial Expression and Recognition Based on Statistical Approach", International Journal of Soft Computing and Engineering (IJSCE) Volume-2, May 2012.

[7]. Anukriti Dureha "An Accurate Algorithm for Generating a Music Playlist based on Facial Expressions"