

COPY RIGHT



ELSEVIER
SSRN

2023 IJEMR. Personal use of this material is permitted. Permission from IJEMR must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. No Reprint should be done to this paper, all copy right is authenticated to Paper Authors

IJEMR Transactions, online available on 31st Mar 2023. Link

[:http://www.ijiemr.org/downloads.php?vol=Volume-12&issue=Issue 03](http://www.ijiemr.org/downloads.php?vol=Volume-12&issue=Issue 03)

10.48047/IJEMR/V12/ISSUE 03/64

Title **MUSIC RECOMMENDATION USING EMOTION RECOGNITION, GESTURES AND OBJECTS**

Volume 12, ISSUE 03, Pages: 470-475

Paper Authors

Razeena Begum.Sk, Praveen Sagar.B, Jaya Durga Prasad. P



USE THIS BARCODE TO ACCESS YOUR ONLINE PAPER

To Secure Your Paper As Per **UGC Guidelines** We Are Providing A Electronic Bar Code

Music Recommendation Using Emotion Recognition, Gestures And Objects

Razeena Begum.Sk¹, Assistant Professor, Department of Computer Science and Engineering, Andhra Loyola Institute of Engineering and Technology, Vijayawada.

Praveen Sagar.B², IV B-tech Department of Computer Science and Engineering, Andhra Loyola Institute of Engineering and Technology, Vijayawada.

Jaya Durga Prasad. P³, IV B-tech Department of Computer Science and Engineering, Andhra Loyola Institute of Engineering and Technology, Vijayawada.

Abstract

Music is an integral part of our lives, and with the advancement of technology, music players have become more sophisticated. However, finding the right song to match our current emotional state or physical gesture can be a tedious task. Current methods for creating automated playlists are often inefficient and may require additional equipment like sensors or EEG. To address this issue, we propose a novel approach that utilizes real-time face expression and gesture extraction to create personalized playlists. Unlike existing methods that rely on speech processing or require extensive computing power, our approach is both cost-effective and efficient.

By analyzing the user's facial expressions and physical gestures in real-time, our system can accurately determine their current emotional state and physical activity level. Based on this analysis, our system can generate a playlist that matches the user's mood and activity level, creating a seamless and personalized listening experience.

Deep learning has gotten a lot of attention in recent years, thanks to the development and application of big data. Convolutional neural network, as a deep learning neural network, is particularly significant in facial image identification. A model that detects facial micro-expressions and suggests music based on mood is identified using a mix of convolutional neural network micro-expression detection technology and an autonomous music recommendation algorithm

Keywords: Convolutional neural network, Deep Learning, Streamlit Application.

Introduction

Music has always been a crucial aspect of our lives, bringing joy and enhancing our overall well-being. With the advent of technology, music players have become more advanced, offering features such as fast forward, backward, variable playback speed, and volume modulation, among others. However, even with these capabilities, users still have to manually browse through the song catalogue and select songs that fit their current mood and behaviour.

A ground-breaking concept in music players is the ability to automatically play songs based on the user's emotions. By taking into account the user's feelings, this approach provides a more personalized and immersive listening experience. Under the current system, the

user still has to actively choose the music, as playing songs at random may not suit their mood. Therefore, it is up to the user to classify music into various emotional groups and then manually choose a song based on a specific feeling before the songs are played.

To implement this, the music player uses real-time face expression and gesture extraction to detect the user's emotional state and physical activity level. Depending on the emotion and gesture, music will be played from the YouTube. Each gesture has a unique name and is associated with a particular person or symbol that refers to instances from everyday life. The programmers have the ability to modify, add, or remove the emotions and gestures according to user needs.

In addition, users may occasionally want to listen to music in different languages or from various singers based on their mood. For example, if a user is feeling sad, they may choose a song that matches their mood, or they may prefer to listen to something more upbeat to improve their mood.

Overall, this approach represents a significant advancement in the field of automated playlist creation. By incorporating the user's emotions and physical activity level, the music player can provide a more personalized listening experience that meets the user's needs and preferences. As long as the software is running correctly on the system, users can choose to change the music in the YouTube according to their preferences, ensuring a seamless and immersive listening experience.

Social media: In the modern day, some users publish their current status for the day along with modified photographs and background music. Therefore, if we utilize our program in this scenario, we may instantly recommend the song without any audio editing. It enables quick and simple uploading of users' statuses.

Literature Survey

The ability to create playlists in music players is made possible by a variety of apps available on the market, but a number of procedures and approaches have been developed and enhanced using user emotions. Only a few crucial emotions are focused when advanced techniques like Voila and Jones are used. But, we also included gestures like the fist, rock, heart, etc. in this project.

The following scientific articles provide a succinct summary of the idea:

1. Both human culture and contemporary technology technologies greatly value music. In most cases, the user must actively browse through a playlist of songs in order to choose one. An efficient and accurate algorithm is used to build a list of music while taking the user's current emotion into account. Current techniques for creating automated playlists are incorrect, computationally inefficient, and sometimes require the use of extra equipment like EEG or sensors. The most natural and ancient means of

accomplishing this is through speech, which is the most natural and ancient way of expressing feelings. The suggested system allows for the recognition of user facial expressions and gestures. Also to infer facial landmarks from the expressions, hand Landmarks from the Gestures which may then be categorized to determine the user's current mood or Symbol of Gestures. When an emotion and Gesture is identified, the user will be shown name that fit the emotion or Gesture. Then user have to press the Recommend music button then automatically the music is played from the YouTube. Depending on Facial expression for music enthusiasts with smart phones and internet connection, Music Player is a terrific app. Anyone creates a profile on the system has access to the application. The requirements that can be met with this programmer are as follows:

- Create an account or sign up for one, then log in.
- Addition of Emotions and/or Gestures
- Training
- Testing

2. An intelligent agent sorts a music collection based on the feelings that each song evokes and then suggests a playlist to the user in line with their mood or their gestures. The mood of the song, or the emotion it conveys, is used to group the user's local music library at first. The music and lyrics of a song are regularly evaluated. A user takes a photo of themselves at the time they want to build a mood-based playlist. On this image, facial recognition and emotion recognition techniques are used to determine the user's sentiment and also the gestures of the user. The user is then presented with a music selection that best expresses this emotion or by hand gestures.

3. People's stress levels are rising as a result of the bad economy, mounting financial obligations, and other things. A stress-reduction activity that can be good is listening to music. But it won't work if the emotion the listeners are experiencing and the music aren't the same. Also, there isn't a music player available that can select songs based on the mood of the listener.

This study offers a solution for this problem: a facial expression-based with

the feature of extracting the hand gestures music player that can recommend songs to users depending on their feelings of melancholy, happiness, neutrality, rage or to the users ease a hand gesture that could identify the genre of the music. The appliance obtains the user's facial image or a simple hand gesture. The user's sentiment which is a face expression is determined using the classification algorithm. This essay explores two different categorization strategies: methods based on gut rate and, consequently, methods based on facial images and gestures. Afterwards the device starts playing music that matches the user's mood. Because the trial data used to develop the proposed technique can accurately define the pleasant mood. Music is important to who we are and how we live. Individuals frequently listen to music in various contexts, whether actively or passively, consciously or unconsciously, as a means of expressing their emotions.

Methodology

The recommended approach extracts facial landmarks and hand landmarks by recognizing user facial expressions or Gesture in order to categories user emotion. The user will be listen music that are appropriate for their emotion or Gesture. A fantastic resource for music lovers with a smartphone and internet connection is the Face Expression and hand Gesture Based Music Player. Anybody who creates a profile on the system has access to the application. The following user needs are those that the program is designed to meet:

- signing up for an account or creating one, then logging in
- Enter the Language and Singer Name
- Detecting emotions and Gestures
- Press the Recommend song Button
- Listen music

Architecture Diagram

The project's primary goal is to play music in accordance with a user's sentiment as indicated by their facial expression and Gesture. Due to the following two technologies:

One is the use of machine learning algorithms to read a user's emotion from their expression on their face.

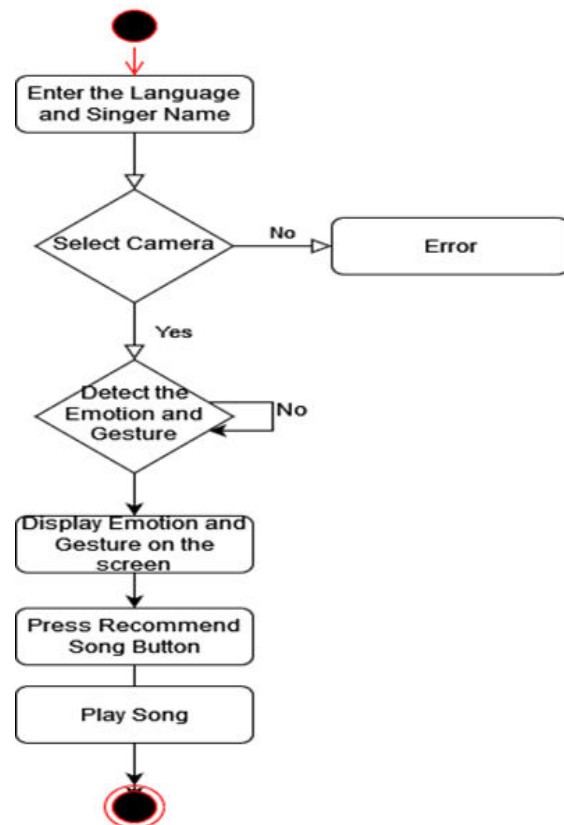
The other is a set of web development tools for creating the music player's user interface or front end.

The music player begins with an interactive user interface that gives the user a input Box with placeholder and Button.

- The first input box is for the user to enter the language of the song they wish to listen to. Because music does not have a language, we included this input box so that users may listen to any sort of language song by entering the language name.

- The second input box is for the user to enter the favorite Singer name.

- And Small Button with Recommend song.



The project's primary goal is to play music in accordance with a user's sentiment as indicated by their facial expression and Gesture. Due to the following two technologies:

One is the use of machine learning algorithms to read a user's emotion from their expression on their face.

The other is a set of web development tools for creating the music player's user interface or front end.

The music player begins with an interactive user interface that gives the user a input Box with placeholder and Button.

- The first input box is for the user to enter the language of the song they wish to listen to. Because music does not have a language, we included this input box so that users may listen to any sort of language song by entering the language name.

- The second input box is for the user to enter the favorite Singer name.

- And Small Button with Recommend song.

The project's centerpiece is the Recommend song, which plays music in response to the user's emotions and Gesture. As soon as the user enter the inputs, the algorithm uses the OpenCV library of Python to access the device's camera and collect the user's image for emotion and Gesture analysis. In order to predict the face expression and Gesture from the trained model, the acquired image is first converted to grayscale and scaled. Uses Media Pipe to extract facial and hand landmarks. It then calculates the relative position of each landmark and saves the extracted features to a NumPy array. The process is repeated until 100 samples are collected or the user stops the data collection. The saved data is in the form of a NumPy file with a user-defined name. The code can be used to collect data for machine learning models that use facial and hand gestures as input features. feedforward neural network for image classification. The specific architecture is a three-layer network with a rectified linear unit (ReLU) activation function in the hidden layers and a SoftMax activation function in the output layer. A feedforward neural network for image classification is a type of artificial neural network that uses multiple layers of processing units to classify images into different categories. The input to the network is a vector of features extracted from the image, and the output is a probability distribution over the different categories.

The network consists of an input layer, one or more hidden layers, and an output layer. Each layer is composed of a set of processing units called neurons, which are connected to neurons in adjacent layers. Each connection has an associated weight, which determines the strength of the connection.

The network is trained using a supervised learning algorithm, such as backpropagation, to adjust the weights so that the output of the network matches the desired output. During training, the network adjusts the weights to minimize the difference between the predicted output and the true output.

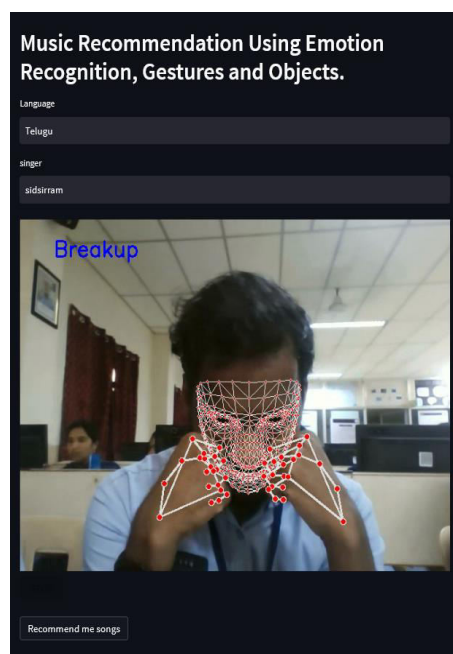
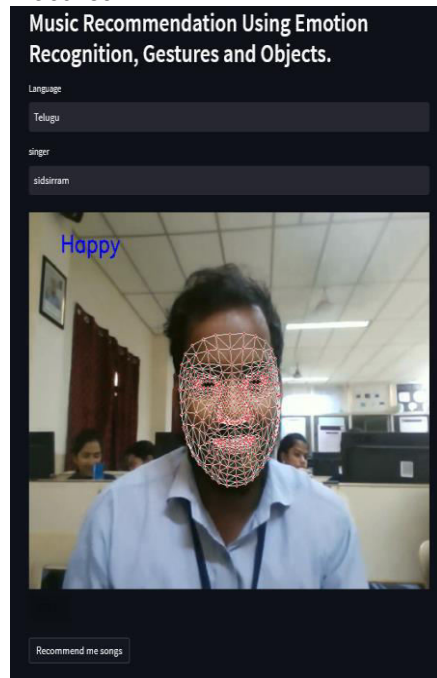
Once the network is trained, it can be used to classify new images by feeding the input vector into the network and computing the output. The category with the highest probability is chosen as the predicted class for the image.

The model is trained using categorical cross-entropy loss and the RMSprop optimizer. Following that, the algorithm predicts the emotion and Gesture in the categories of neutral, happy, sad, Rock, heart, Angry, fist . This is the way that the back end that we created operates. The music player is created using technologies like Streamlit for the user interface. Streamlit is an open-source Python library that allows developers to quickly create interactive web applications with just a few lines of code. It enables data scientists and machine learning engineers to build and share custom web apps for their models, data analyses, and visualizations, without needing extensive web development experience. The camera takes a picture of the user and sends it to the server. The script loads a pre-trained Keras model for emotion recognition from an .h5 file and a set of labels from a .npy file. It then defines a class called "Emotion Processor" that processes video frames captured by the user's webcam and extracts relevant features from the user's face and hand movements using the Media pipe library. The extracted features are fed to the emotion recognition model to predict the user's emotional state, which is then displayed on the video frame.

The web application prompts the user to enter their preferred language and singer, and then captures their emotion using the "Emotion Processor" class. Once the

user's emotion is captured, the application uses the YouTube Data API to search for videos that match the user's preferred language, singer, and emotion, and recommends the top result to the user. The recommended video is opened in the user's default web browser.

Results



Conclusion

Due to its numerous applications, the challenging topic of facial expression identification in the realm of image

analysis and computer vision has attracted a lot of attention recently. We created a music player by adding more emotions and Gestures. The main draw is "Recommend music based on facial expressions and Gestures". We created a system that is user-centric and worked to make it simpler for users to utilize. We developed an User Interface using Streamlit so that this can be operated in any system. The field of expression recognition research can be further explored and improved. Despite the small dataset size, the training is done with a feedforward neural network for image classification, which has a 96 percent accuracy rate. Future research on emotions and Gestures will broaden by including more datasets. With a collection including more photographs, these emotions and Gestures can be identified more precisely. Additionally, it would enable us to increase song selection and enhance user-based song matching. By including more emotions and Gestures like disgust, fear, dissatisfaction, confusion, etc. in the project, which may be done by expanding the dataset. Training over a larger dataset can strengthen the authorization of each registered human and eliminate the possibility of incorrect analysis. A mental health hypothesis for a certain person can be created with the analysis of each emotion in numerous contexts in psychological studies and medical diagnosis. Depending on the preferences of each user, it can also be customized for them individually.

References

- [1] 2017 Smart music player integrating facial emotion and music mood recommendation
- [2] <https://ieeexplore.ieee.org/document/8299738>
- [3] 2011 UKSim 5th European Symposium on Computer Modeling and Simulation
<https://ieeexplore.ieee.org/document/6131215>
- [4] S. Deebika, K. A. Indira and Jesline, "A Machine Learning Based Music Player

- by Detecting Emotions,\" 2019 Fifth International Conference on Science Technology Engineering and Mathematics (ICONSTEM), Chennai, India, 2019, pp. 196-200, doi: 10.1109/ICONSTEM.2019.8918890.
- [5] Arora, A. Kaul and V. Mittal, \"Mood Based Music Player,\" 2019 International Conference on Signal Processing and Communication (ICSC), NOIDA, India, 2019, pp. 333-337, doi: 10.1109/ICSC45622.2019.8938384.
- [6] Alrihaili, A. Alsaedi, K. Albalawi and L. Syed, \"Music Recommender System for Users Based on Emotion Detection through Facial Features,\" 2019 12th International Conference on Developments in eSystems Engineering (DeSE), Kazan, Russia, 2019, pp. 1014-1019, doi: 10.1109/DeSE.2019.00188.
- [7] Altieri et al., \"An Adaptive System to Manage Playlists and Lighting Scenarios Based on the User's Emotions,\" 2019 IEEE International Conference on Consumer Electronics (ICCE), Las Vegas, NV, USA, 2019, pp. 1-2, doi: 10.1109/ICCE.2019.8662061. [8] 2017 An Efficient Real-Time Emotion Detection Using Camera and Facial Landmarks.