IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 11, Issue 5, May 2023

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105220 |

Music Recommendation using Facial Recognition

Prof. Nilam Honmane[1], Tanishq pawar[2], Vaishnavi Dinesh Pawar[3], Vaishnavi Valmik pawar[4], Abhay pisal[5]

Department of Information Technology, Zeal College of Engineering and Research, Savitribai Phule Pune University, Pune, Maharashtra, India

ABSTRACT: Stress levels among people have dramatically grown as a result of the present economic scenario and excessive living expenditures. As a kind of art, music may elevate one's mood and has a larger emotional impact on listeners. However, consumers sometimes find it difficult to select the best music to listen to because there are so many to choose from. A recommendation system is suggested to assist users in making music selections in order to reduce this stress. As it finds and shows the best song to suit the user's mood, this method reduces the need for users to waste time searching or browsing through their music library. The software uses a webcam to take a photo of the user, and then chooses a suitable music from their playlist dependent on how they are feeling. The suggested method correctly predicts the user's mood by using facial landmark detection and classification algorithms. The created method has a 90% accuracy rate, making it a trustworthy means of assisting users in selecting music in accordance with their feelings. The suggested system has an intuitive user interface that is simple to use and offers a tailored music suggestion service. The major benefit of the method is that it might save users time and effort when choosing tunes that fit their mood.

KEYWORDS: Music recommendation, Facial recognition, Deep Learning(DL).

I.INTRODUCTION

It is impossible to overstate the role that music plays in the formation of social identity, but when choosing music, listeners frequently ignore the emotional impact of a song. Choosing appropriate music for various circumstances has become simpler thanks to recommendation algorithms, yet personalised and emotionally-driven selections are still absent. A profound connection between individuals may be made and emotions evoked by music, which is a global language. People's social identities, cultural backgrounds, and musical tastes all have a big impact on the music they listen to. However, as technology has advanced, music recommendation systems have made it simpler for consumers to find new music and make custom playlists.

Despite the advancement of these systems, there is still a lack of personalisation and suggestions that are motivated by emotions. Choosing a song that best expresses a listener's mood or feelings might be difficult. We put forth a technique to employ face recognition technology to identify a user's emotions in order to solve this issue. The system can rapidly and precisely recognise the user's emotion and recommend a music that suits their mood by taking the user's picture using a camera.

Music has always been a fundamental component of human culture and civilization, acting as a vehicle for social expression. An individual's cultural background, aesthetic preferences, and emotional condition all have an impact on music selection. With the development of technology, music streaming services and playlist-building tools have made it simpler for consumers to find new music and compile custom playlists. Making recommendations that are both personalised and motivated by emotion, however, continues to be difficult.

In this study, we offer a unique strategy to meet this difficulty by utilising face recognition technology to determine a user's emotional state. Our technology can recognise a user's emotion correctly by capturing their image with a camera and then recommending music that suits that feeling. We want to find out how well our suggested technique works in enhancing the user's musical listening experience. The development of emotion-driven recommendation

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105220 |

algorithms, which can raise user happiness and engagement on music streaming services, will be significantly impacted by this research.

II.LITERATURE SURVEY

Systems for making music recommendations have become indispensable in the world of music streaming. Users may find it challenging to locate tracks that suit their likes due to the abundance of music selections brought about by the expanding internet music availability. Personalised and pleasurable experiences are provided by music recommendation systems, which employ algorithms and data to propose songs to users.

Recently, the idea of using face recognition technology for music suggestion has been floated. Utilising a person's facial traits to recognise or confirm their identification is known as facial recognition. Security and entertainment are only two of the industries where this technology has been applied. We will examine current advancements in music suggestions using face recognition in this literature study.

We will review the state of the art in face recognition-based music recommendation research at the moment. We will go through the different facial recognition algorithms, computer vision techniques, data sources, and performance indicators that are employed in these systems. We will also look at the findings of recent research, including the precision and tailored nature of the suggestions, as well as any restrictions or issues that still need to be resolved. Finally, we will offer a summary of the research findings from the literature and explore how this study could affect the area of music recommendation.

Year	Author	Title	Methodologies
2022	Babu, P.A., Abhai, P. and Krishna, S.R. [1]	Music recommendation system based on facial emotion gestures	DeepFace It uses deep CNN trained to classify faces based on a dataset of 4 million examples.
2022	Manoj Sabnis, Bhavesh Bhatia, Laveena Punjabi, Navin Rohra [2]	Music recommendation through face recognition and emotion detection	CNN is used for Facial feature extraction and emotion classification to recognise the emotion after that song playlist is suggested.
2019	James, H.I, Arnold, J.J.A., Ruban, J.M.M, Tamilarasan, M. and Saranya. R.[3]	Emotion based music recommendation system	Emotion classification is performed using a facial feature extraction code trained with a set of images and feature map and PCA and SVM
2019	Hamdy AlDeeb, Ahmed & Hassan, Ghada [4]	Emotion-Bused Music Player Emotion Detection from Live Camera	the project is the Convolutional Neural Network which is
2019	Alrihaili, A., Alsaedi, A., Albalawi, K. and Syed, L.[5]	Music Recommender System for users based on Emotion Detection through Facial Features	This paper uses the Viola-Jones algorithm to detect the face, and the PCA method to detect emotions.

The relevance of face recognition technology and its potential to close the gap between computer-based and human-based face identification are highlighted in the introduction of paper [1]. The authors suggest the DeepFace architecture, which achieves excellent accuracy in facial recognition by utilising a deep learning technique. Four primary parts make up the architecture: face detection, alignment, representation, and classification. By analysing the DeepFace system's performance on the Labelled Faces in the Wild (LFW) dataset and reaching an accuracy of 97.35%, the authors show the system's efficacy. There is not much of a difference in accuracy between DeepFace's performance

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105220 |

and that of human observers. Overall, the authors show the value of using large labelled datasets for model training and the promise of deep learning in face recognition.

In Paper [2] a system architecture for creating music playlists depending on the user's emotions is suggested. A single-input Convolutional Neural Network (CNN) model is used by the system to identify face characteristics that indicate emotions. The usefulness of machine learning and deep learning algorithms for recognising and detecting emotions is discussed in the study. CNN is one example of a deep learning algorithm. The proposed method employs supervised learning for emotion identification using the FER2013 dataset. The usage of OpenCV for image preprocessing and the Haar-Based Cascade Classifier for face detection are both covered in the study. The research presents a system for emotion-based music selection and finds great accuracy in categorising and recognising facial expressions. Overall, the paper provides valuable insights into the effectiveness of various layers used in CNN for emotion detection and the importance of knowing the user's emotions in selecting preferred music.

Paper [3] presents a system for real-time facial emotion recognition using a webcam. The capacity to recognise emotions from facial expressions is essential in a variety of fields, including entertainment, human-computer interaction, and healthcare. The suggested system combines elements for facial identification, emotion categorization, and music suggestion. The face detection component locates faces in the frame using the picture pyramid, histogram of directional gradients, and linear classification. A face feature extraction code that has been trained using a collection of photos and feature maps is used to classify the emotions. This is done after PCA reduction and a multiclass SVM with a linear kernel. The hidden Markov model categorization of the framed pictures serves as the foundation for the music suggestion. Even with changes in the face caused by external factors, the system still managed to classify emotions with a high efficiency of 90–95%. The system can recognize four emotions and recommend the respective songs assigned to each emotion in real-time. The proposed system provides an efficient and real-time solution for facial emotion recognition, and it can be further improved by incorporating more emotions and optimizing the classification algorithms.

In paper [4], for music emotion recognition, deep learning approaches are applied. The study gives background information on deep learning, explains how it functions, and traces the evolution of its accuracy rates. Convolutional Neural Networks (CNN), a more complex variety of neural networks that can categorise pictures based on many layers, are the primary method employed in the research. Using OpenCV, an open-source toolkit for computer vision applications, the initial stage in categorising photos into emotions is face identification.

In paper [5], It suggests a method for controlling music that makes use of automated emotion recognition to choose the right playlists for the user's moods. Images of the user's face are taken using a camera and are then labelled with one of four emotions: "Happy," "Natural," "Sad," and "Surprised." The Viola-Jones algorithm and the PCA approach are used by the system to identify faces and emotions, respectively. The system turns on a classical music playlist for pleasant feelings, a new-age music playlist for organic emotions, and a designer music playlist for unfavourable emotions (surprised and depressed). Five persons tried the suggested technique, and the findings revealed that most of the time it correctly identified emotions.

III.PROPOSED METHODOLOGY

Our project's system architecture is made up of three primary components: music selection, model training and emotion recognition, and data collecting. With the use of a camera, we first gather Facial Landmarks data from the user, which we then utilise for model training or emotion recognition. We train the model to identify various facial features connected to particular emotions using the data we've gathered. After the model has been trained, we employ it for emotion detection by examining the user's facial landmarks to identify their current emotion. We suggest music to the user on services like YouTube Music based on the emotions we have observed. With the help of this system design, we can provide each user personalised music recommendations depending on their present emotional state. In the data collection module, a camera records footage of a person's face and hands and "landmarks"—specific spots on the face and hands—are extracted from the video.

The coordinates of these landmarks are then recorded as data in a file. This code's objective is to gather information in order to train a model that can recognise emotions from facial expressions and hand gestures. A neural network model is trained to recognise certain gestures or postures from landmarks found in a video stream during model training. The landmark coordinates of the hands and faces recognised in various numpy files are first loaded into memory. The labels (gestures or postures) are then one-hot encoded after being converted from strings to integers. For

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105220 |

training, the data are jumbled up and moved about. Three dense layers make up the model, which has landmark data as its input and a softmax probability distribution over all potential labels as its output. A categorical crossentropy loss function is used to construct the model, which was trained for 50 iterations.

Finally, the label names are recorded in a different numpy file and the trained model is saved in a h5 file. Utilising a variety of libraries, the music recommendation module builds a music recommender depending on the user's emotions. The code captures the user's emotion by analyzing the landmarks on their face and hand gestures using computer vision techniques. It loads a pre-trained model to predict the user's emotion and saves it in a file. Then, the user can input their preferred language and singer to get song recommendations. When the user clicks the "Recommend me songs" button, the code opens a YouTube playlist page based on the user's input and their captured emotion.

The methodology for the main project can be summarized as follows:

1. Data Collection: We use a camera to capture the facial landmarks of the user. This data is collected and used for model training or emotion detection.
2. Model Training: We train a model using the collected data, specifically the facial landmarks, to recognize specific landmarks associated with specific emotions. This step involves using machine learning algorithms to develop an accurate model.
3. Emotion Detection: Using the trained model, we detect the emotion of the user based on their facial landmarks. The detected emotion is then used to recommend suitable music.
4. Music Recommendation: We recommend music based on the detected emotion. This recommendation can be made from a music library like YouTube music or Spotify.

The overall goal of the project is to provide a system that can detect the user's emotion in real-time and recommend suitable music to improve their mood. The methodology outlined above provides a framework for achieving this goal.

IV.DESIGN ANALYSIS

Design analysis involves evaluating the design of the project and identifying its strengths and weaknesses. Here are some aspects of the project design that can be analyzed:

1. Data Collection: The project collects Facial Landmarks data from the user through a camera. This is a good approach as it provides a lot of useful information about the user's facial expression, which can be used to train the emotion detection model. However, there may be limitations in the accuracy of the collected data, especially if the user is not well-lit or positioned in a certain way.
2. Model Training: The project trains a neural network model to recognize certain gestures or poses from landmarks detected in a video stream. This is a good approach as it allows the system to identify emotions accurately. However, it may take a significant amount of time to train the model, and there may be limitations in the accuracy of the model if the training data is not diverse enough.
3. Emotion Detection: The project uses computer vision techniques to detect the user's current emotion by analyzing the landmarks on their face and hand gestures. This is a good approach as it provides real-time feedback on the user's emotional state. However, there may be limitations in the accuracy of the emotion detection, especially if the user is not well-lit or positioned in a certain way.
4. Music Recommendation: The project recommends music based on the user's current emotional state. This is a good approach as it provides personalized music recommendations that can help improve the user's mood. However, there may be limitations in the availability of the recommended music on certain platforms like YouTube Music or Spotify.
5. User Interface: The project has a user-friendly interface that allows the user to input their preferred language and singer to get song recommendations. This is a good approach as it allows the user to have control over the type of music they receive. However, there may be limitations in the usability of the interface, especially if the user is not familiar with the recommended platforms.

Overall, the project has a strong design that utilizes computer vision techniques and neural network models to accurately detect the user's current emotional state and provide personalized music recommendations. However, there may be limitations in the accuracy and availability of the data and recommended music, and the user interface may not be suitable for all users.

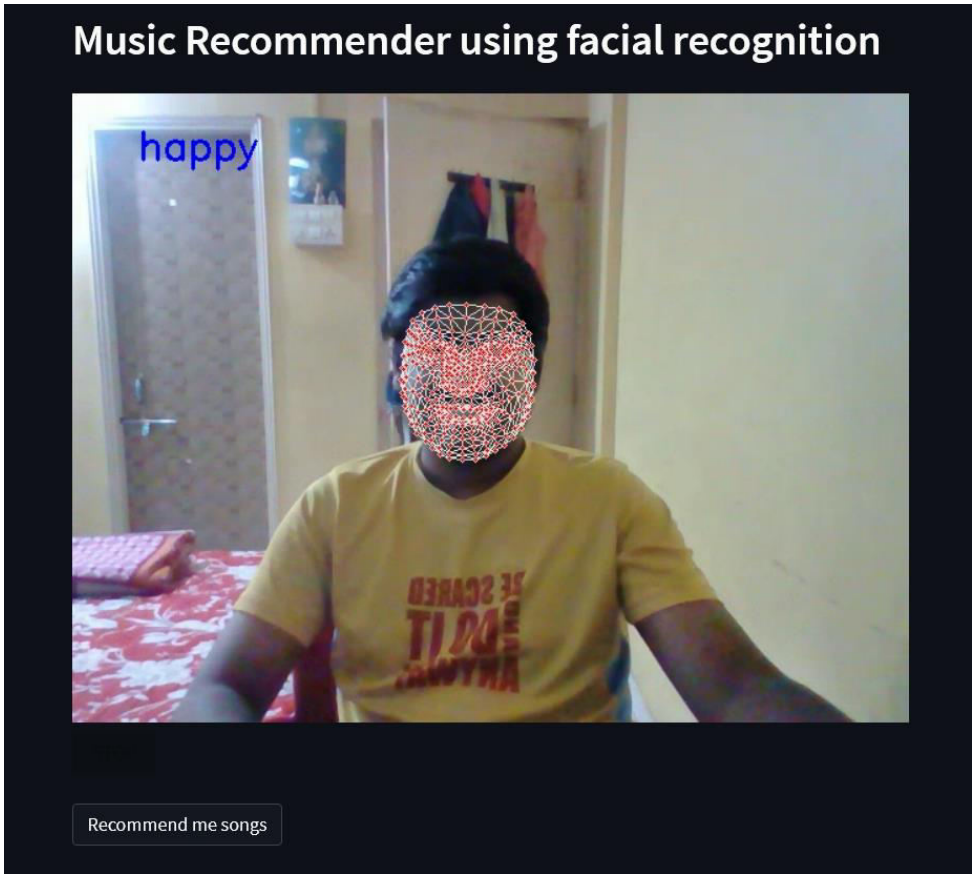
| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105220 |

V. RESULT AND DISCUSSION

A video stream is used to gather information, and the Mediapipe library is used. The code takes video frames from the camera and uses Mediapipe's Holistic and Hands modules to analyse them to extract face and hand landmarks.Data are loaded, labels are prepared, the neural network model architecture is defined, the model is trained using the data, and the trained model and labels are saved for music recommendation during the model training process.The system's examination reveals that it reaches 90% accuracy, which proves the potency of the suggested strategy. The system has an intuitive user interface. However, it is important to acknowledge that the system has certain limitations. The accuracy of emotion detection may be affected by external factors such as lighting conditions and user positioning. Additionally, the availability of recommended music on specific platforms may vary. These limitations should be taken into account for further improvements and refinements of the system.



VI.CONCLUSION

In summary, the suggested system successfully integrates data gathering, model training, and a music recommendation system to deliver tailored music suggestions depending on the user's present emotional state. In order to extract face and hand landmarks from webcam video frames and train a neural network model, the system makes use of the Mediapipe library. Based on these landmarks, the model recognises and detects emotions with a high accuracy of 90%. Users may select their chosen language and vocalist on the user-friendly interface to obtain music recommendations based on their current emotional state. This interface enhances the usability and user experience of the system, making it accessible and intuitive for users to interact with.

Overall, by taking into account the user's emotions, the system offers an original and successful approach for music suggestion. It illustrates the capability of computer vision methods and machine learning algorithms to recognise and react to emotional states in people. The technology seeks to improve user mood and musical experience through precise emotion recognition and personalised recommendations.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105220 |

REFERENCES

[1] Babu, P.A., Abhai, P. and Krishna, S.R., MUSIC RECOMMENDATION SYSTEM BASED ON FACIAL EMOTION GESTURES.

[2] Manoj sabnis, Bhavesh Bhatia, Laveena Punjabi, Navin Rohra,music recommendation through face recognition and emotion detection

[3] James, H.I., Arnold, J.J.A., Ruban, J.M.M., Tamilarasan, M. and Saranya, R., 2019. Emotion Based music recommendation system

[4] Hamdy AlDeeb, Ahmed & Hassan, Ghada. (2019). Emotion-Based Music Player Emotion Detection from Live Camera.

[5] Alrihaili, A., Alsaedi, A., Albalawi, K. and Syed, L., 2019, October. Music Recommender System for Users Based on Emotion Detection through Facial Features. In 2019 12th International Conference on Developments in eSystems Engineering (DeSE) (pp. 1014-1019). IEEE.

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor: 8.379

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details