iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** V **Month of publication:** May 2023

DOI: https://doi.org/10.22214/ijraset.2023.52568

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue V May 2023- Available at www.ijraset.com

Machine Learning Based Music Recommendation System Using Facial Expression

Pooja Kedari[1], Pavan Rengade[2], Samiksha Deshmukh[3], Mr. Sharad Adsure[4], Mrs. Deepika Jaiswal[5]
[1, 2, 3, 4]*Computer Engineering, Dr. D. Y. Patil Institute of Technology, Pimpri, Pune, India*
[5]*Asst. Professor, Dr. D. Y. Patil Institute of Technology, Pimpri, Pune, India*

Abstract: ***Music plays a key role in improving your well-being, as it is one of the key sources of entertainment and inspiration to keep you going. Recent studies have shown that people respond very positively to music and that music has a great effect on a person's brain activity. It is often better to listen your favorite music these days. This job focuses on systems that suggest songs to users based on their state of mind. This computer vision system uses components to determine a user's emotion from facial expressions. When an emotion is detected, the system will suggest songs with that emotion, saving users a lot of time manually selecting and playing songs. It enables the migration of computer vision techniques for such systems to automation. To achieve this goal, we use the algorithm to classify human expressions, game play, and music tracks according to their currently detected emotions. Reduces the effort and time required to manually search for songs in the list. Recognize human facial expressions by extracting facial features using Haar Cascade and CNN algorithms.***

Keywords: Recommender System, Convolution Neural Network (CNN), Image Processing, Machine Learning, Classification etc.

I.INTRODUCTION

Facial expressions are one of the verbal means of conveying emotions, and those feelings can be exploited through verbal exchange and Human-Machine Interface (HMI). In contemporary technologically advanced world, exceptional music players are utilized in capabilities like flip Media, speedy ahead, Flow Multicast Streams and greater. While these features meet basic person needs, you still should manually look for the current temper amongst a massive variety of songs. That is a time-ingesting undertaking that calls for a few attempt and endurance. The main purpose is to develop an smart system that can without problems apprehend facial expressions and play music tracks based on a particular expression/emotion. Emotions generally categorized into three- glad, sad, and neutral. The principle intention of this work is to expand a clever system which can easily understand facial expressions and play tracks primarily based on them. The three feelings which might be usually categorised are glad, sad and neutral. The algorithm, Haar Cascade used for the face extraction characteristic. The proposed algorithm could be very beneficial due to the fast calculation time, which improves the system overall performance. This system has been carried out in diverse fields. Human-Computer Interaction (HCI), healing methods in healthcare, and extra. Virtual tune is generally categorized as created based on attributes consisting of artist, genre, album, language, and reputation[7]. A number of the available on line tune streaming offerings use collaborative filter out-based pointers to suggest music primarily based in your choices and listening history. but, these suggestions might not fit your current temper. This approach proposes a CNN-primarily based tune recommendation technique that uses data from multimodal sentiment analysis captured by using facial actions to improve the machine's selection making on emotions detected in actual time. Machine getting to know has to end up very famous in recent years. Some styles of machine studying strategies are extra appropriate than others, depending on the character of the application and the datasets available. For a ramification of the principle forms of studying algorithms encompass supervised, unsupervised, semi-supervised and reinforcement getting to know. Neural networks (NNs) are machine getting to know and generally green techniques for extracting critical capabilities from complicated datasets and inferring functions or models that represent the ones functions[4]. The NN makes use of the schooling facts set to teach the model first. After the model is skilled, the NN may be implemented to new or previously unseen records factors to classify the records primarily based on the previously educated fashions.

II.LITERATURE REVIEW

1) An Emotional Recommender System for music [1]**:** Recommender systems have become essential for users to find "what they need" in large collections of items. Meanwhile, recent studies have shown that user personality can effectively provide more valuable information to significantly improve the performance of recommenders, especially considering behavioral data captured from social network logs. In this work, they describe a new music recommendation technique based on the

identificationof personality traits, moods and emotions of a single user, based on solid psychological observations recognized by the analysis of user behavior in a social environment. In particular, user's personalities and moods have been incorporated into the content filtering approach to achieve more accurate and dynamic results.

2) Music Recommender System for users based on Emotion Detection through Facial Features [2]: Facial emotion detection has received tremendous attention due to its applications in computer vision and human-computer interaction. In this research, they propose an emotion recognition recommendation system that is able to detect the user's emotions and suggest a list of suitable songs that can improve his mood. The proposed system detects emotions, if the subject has a negative emotion, then he will be presented with a specific playlist containing the most suitable types of music that will improve his mood. On the other hand, if the detected emotion is positive, an appropriate playlist will be provided that will include different types of music that will enhance the positive emotion. The proposed recommendation system is implemented using the Viola- Jonze algorithm and PCA (Principal Component Analysis) techniques.

3) Emotional Detection and Music Recommendation System based on User FacialExpression [3]: Music plays a significant role in improving and uplifting the mood. It is often confusing for a person to decide what music to listen to from the vast collection of existing options. Analyzing the user's facial expression/emotions can lead to an understanding of the user's current emotional or mental state. This work focuses on a system that suggests songs to users based on their state of mind. The user's image is captured using a web camera. A snapshot of the user is taken and then according to the user's mood/emotion, a suitable song from the user's playlist is displayed to match the user's request.

4) Facial Expression Based Music Player [4]: The conventional way of playing music depending on a person's mood requires human interaction. The transition to computer vision technology will enable the automationof such a system. To achieve this goal, an algorithm is used to classify human expressions and play a music track according to the currently detected emotion. It reduces the effort and time required to manually search for a song from a list based on a person's current state of mind. A person's expressions are detected by extracting facial features using the PCA algorithm and the Euclidean Distance classifier. In this paper, they use an embedded camera that is used to capture a person's facial expressions, which reduces the system design cost compared to other methods.

5) Music Recommendation System Using Facial Expression Recognition Using Machine Learning [5]: The study of human emotional responses to visual stimuli such as photographs and movies, known as visual sentiment analysis, has proven to be a fascinating and difficult problem. The development of powerful algorithms from computer vision is responsible for the success of current models. Most existing models attempt to overcome the problem by recommending either robust features or more sophisticated models. Key suggested inputs are mainly visual elements from the entire image or video. Local areas have received less attention, which we believe is important for people's emotional response to the whole picture. Image recognition is used to find people in photos, analyze their emotions, and play emotion-related tunes based on their feelings. This repository achieves this goal by leveraging Google's Vision services. Given an image, it would search forfaces, identify them, draw a rectangle around them, and describe the emotions it found.

III. METHODOLOGY

Steps of system design, training dataset, and test images are considered and the following applicable steps are used to achieve the desired result. The training set is the data with a large amount of data stored and the test set is the input given for detection purposes.

1) *Image Acquisition:* In all image processing techniques, the first task is to acquire an image from the source. These images can be acquired via cameras. The image displayed here is different for each user. i.e., Dynamic image



Figure 1. Image Acquisition

2) *Pre-processing:* Pre-processing is mainly performed to remove unnecessary information from the acquired image and correct some values so that the values are the same throughout. In the preprocessing phase, the image is converted from RGB to grayscale. The eyes, nose, and mouth during pretreatment are considered the regions of interest. Detected by a cascade object detector using the Haar cascade function.



Figure 2. RGB Image

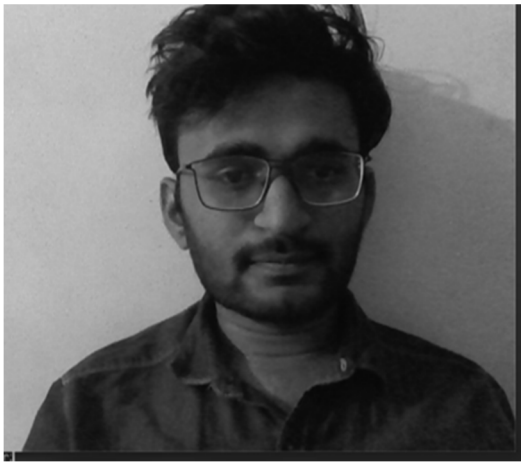


Figure 3. Grayscale Image

3) *Facial Feature Extraction:* After preprocessing, the next step is feature extraction. The extracted facial features are saved as useful information during training and testing phases. You can consider the following facial features - mouth, forehead, eyes, dimples on the face and chin, eyebrows, nose, facial wrinkles. In this work, the eyes, nose, mouth, and forehead are targeted for feature extraction as the reasons for the most attractive facial expressions. It's easy to tell if you're surprised or scared when you have wrinkles on your forehead or an open mouth. But you can never show off with a human complexion. Haar feature technology is used to extract facial features.

4) *Facial Recognition:* For Recognition and Classification of People's Facial Expressions Convolutional Neural Network classifier is used. Gets the best match to the test data from the training data set, so it better matches the currently known formula. Face recognition is an important computer vision problem for recognition and localization images. Face recognition can be performed with a traditional feature-based cascade classifier using the OpenCV library. Face detection can be achieved with a multitasking cascaded CNN.

5) *Play Music:* The last and most important part of this system is to play music based on the recognized person's current emotions. Once the user's facial expressions are classified, the corresponding user's emotional state is identified, and the collected songs of different genres by the number of emotions and put them in a list. Each emotion category has a certain number of traces. If the user expression is classified using the CNN algorithm, songs belonging to that category will be played.

Block Diagram

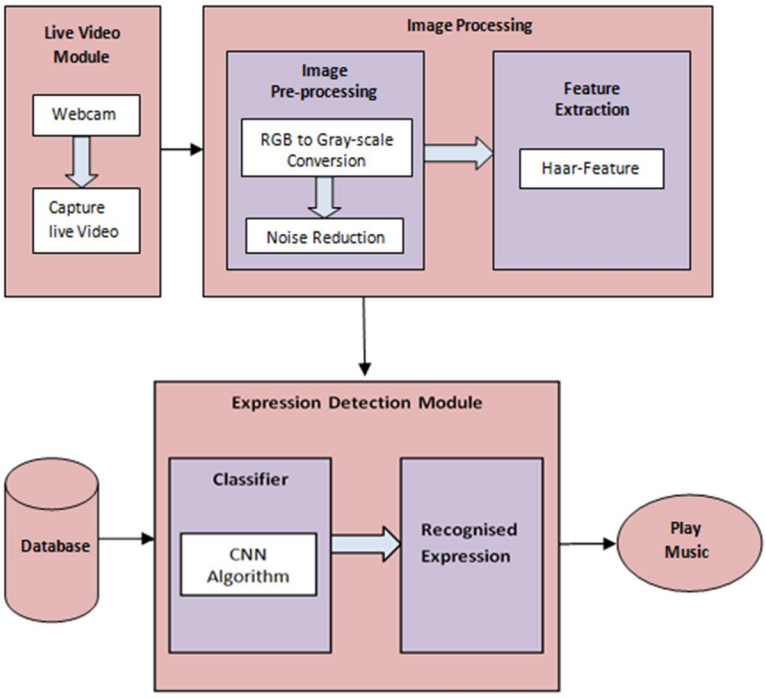


Figure 4. Block Diagram

A. Input

1) In the input phase, live video is recorded by the webcam. Detected faces are used for feature extraction in a separate process. After completing the input, the next step is preprocessing.

2) Pre-processing: During the pre-processing phase, the blurry parts of the image are denoised. It also uses the matplotlib library to convert the image from RGB to grayscale.

3) Feature Extraction: After preprocessing, extract features from the captured image using Haar features. Haar cascades are OpenCV, XML files. The Haar function considers a rectangular part of the image and divides it further into several parts. Therefore, the eyes, nose, mouth, and forehead are considered for feature extraction to present the most attractive facial expressions.

4) Classification: Use CNN algorithm for classification. Classification is the process of classifying and labeling groups of pixels in an image based on certain rules. After completing the training phase, the machine creates a model, a trained model. You need to give an image as input for testing. The model can then enter the testing phase and also provide output to the user.

B. Algorithm

Convolutional Neural Networks (CNN): Convolutional Neural Networks, a type of deep learning algorithm, are very good at reading pictures. Great algorithm for automatic processing of pictures. The picture contains RGB combination information. You can use matplotlib to import an image from a file into memory. A convolutional neural network is a special type of neural community that helps machines learn and classify images.

A convolutional neural network has three types of layers

1) Convolutional Layer: Each input neuron in a ordinary neural community is attached to the following hidden layer. A small fraction of the CNN's input layer neurons are connected to the hidden layer of neurons.

2) Pooling Layer: The dimensionality of the feature map is reduced using a pooling layer. Within the hidden layers of a CNN, there are many activation and pooling layers.

3) Fully-Connected layer:The last few layers of the network are known as fully connected layers. The output of the last pooling or convolutional layer is fed to the fully connected layer and flattened earlier being applied.

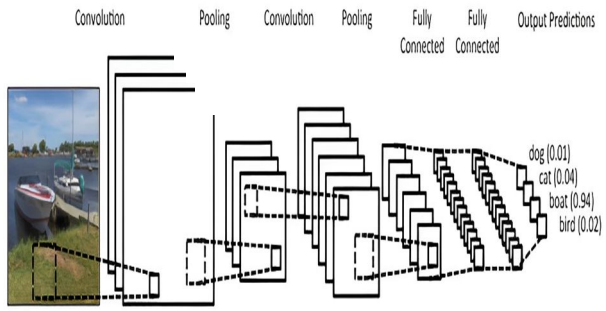


Figure 5. CNN Algorithm

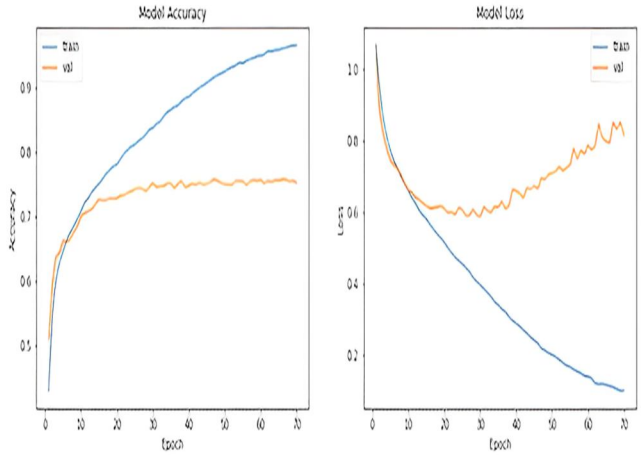


Figure 6. Model Accuracy and loss graph

C. Haar Cascade Feature

This is an object detection algorithm used to identify faces in real-time images or videos. This algorithm uses the edge or line detection feature proposed by Viola and Jones in year 2001 research paper, “Rapid Object Detection using a Boosted Cascade of Simple Features”. This algorithm contains many positive image planes and many negative images that have not been added.

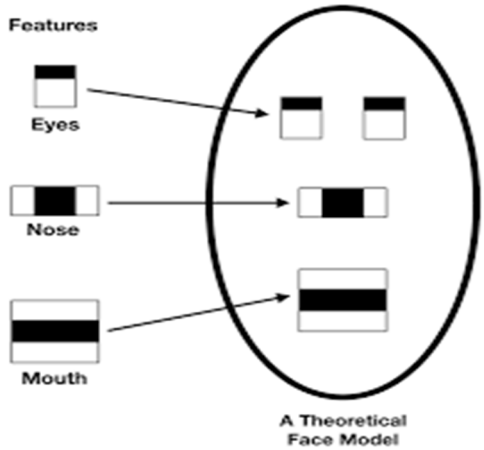


Figure 7. Haar Cascade Feature

Flowchart

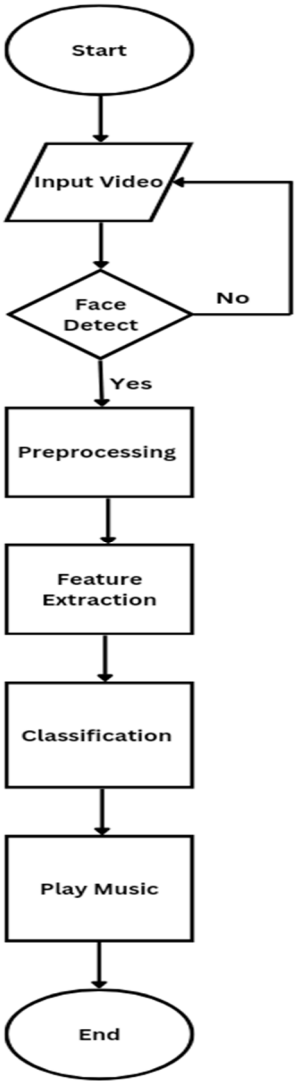


Figure 8. Flowchart

IV. SYSTEM SPECIFICATIONS

A. Hardware Specifications

1) System Processors : Intel i5
2) Speed : 2.80 GHz
3) Hard Disk :64 GB
4) RAM : 8 GB
5) Key Board: Standard Windows Keyboard

B. Software Specifications

1) Operating System: Windows 10(64 Bit) and Above.
2) IDE: Spyder
3) Programming Language: Python version 3.7, 3.8.

V. RESULT

Once the code runs, the user lands on the entry page of the system, where there will be options for registration and login be on the console. After clicking on the registration option the registration page will appear in the window/screen.



Figure 9. Registration Window

This is the registration page that will appear after clicking on the registration option, the user will be taken to the registration window where you need to enter details such as full name, address, email, etc.

After registering, login information is asked where you have to enter correct username and password for the next login process.

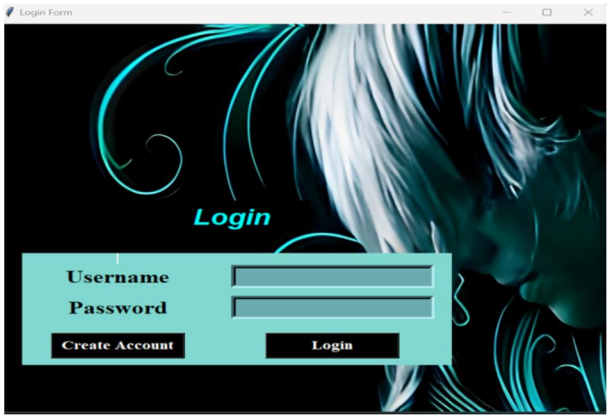


Figure 10. Login Page

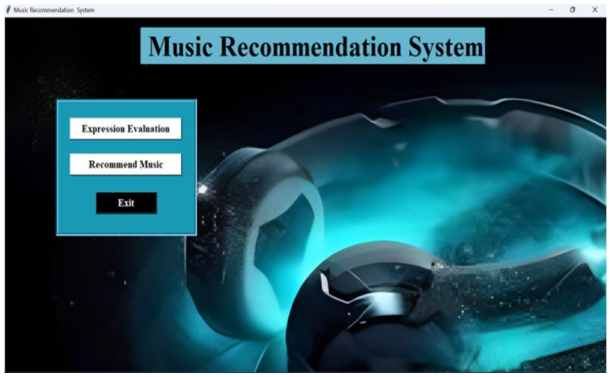


Figure 11. Home Page

Once the user logs in, home page will get displayed. There are two buttons consist of expression evaluation and recommended music. Once Expression evaluation button get clicked then the live camera will turn on, and detects the face. After detecting it will be in the pre processing phase. In that phase, the image will be converted to gray scale and further will be used for noise reduction.



Figure 12. Image Preprocessing

Once the preprocessing phase is done, feature extraction will be done using the Haar feature and then the classification will be implemented with the help of CNN Algorithm and emotion get detected.

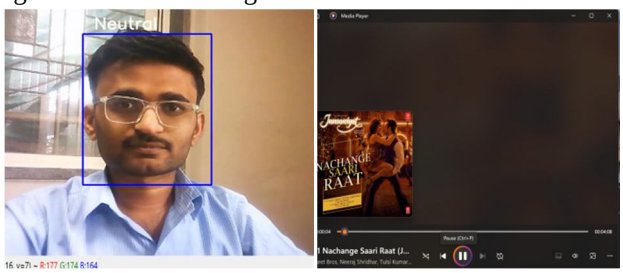


Figure 13. Song Play as per Expressions Detection (Neutral)

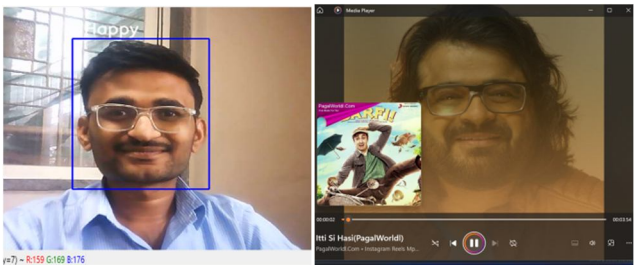


Figure 14. Song Play as per Expressions Detection (Happy)

After clicking Recommend Music button, there will be music played according to the detected emotions of the users and list get displayed.



Figure 15. Recommended Happy Song List

Once the happy face get detected it will display the list of songs for neutral mood.



Figure 16. Recommended Neutral Song List

Once the neutral face get detected it will display the list of songs for neutral mood.



Figure 17. Recommended Sad Song List

Once the sad face get detected it will display the list of songs for neutral mood.

.

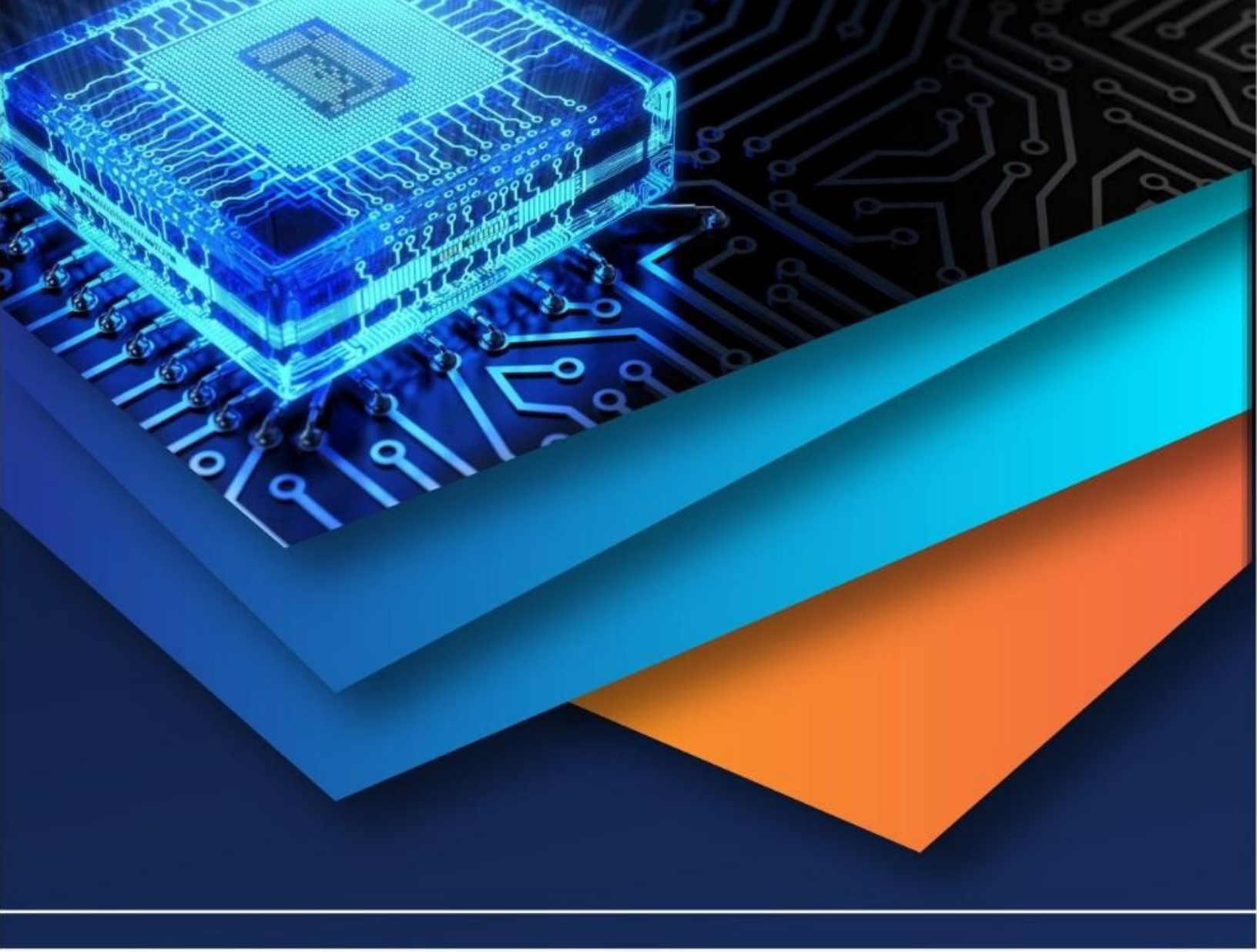
VI. CONCLUSION

The proposed work represents a facial expression recognition system to play the song according to expression detected. It uses a CNN approach for feature extraction. Our system is used to recognize the user's emotions based on the facial expression using Haar Feature. We integrate python code in to the web service and play music on the basis facial expressions like happy, sad or neutral. Recognizing emotions using facial expressions is one of the important research topics which has collected much attention in the past. It can be seen that emotion recognition with the help of image processing algorithms is increasing every day. Researchers are constantly working on the ways to solve it using different kinds image processing properties and methods. Henceforth, it is possible to extend the system on a mobile device.

REFERENCES

[1] Vincenzo Moscato, Antonio Picariello And Giancarlo Sperl, "An Emotional Recommender System For Music", Ieee Intelligent System, October 01,2020.
[2] Ahlam Alrihaili, Alaa Alsaedi, Khlood Albalawi,Liyakathunisa Syed, "Music Recommender System For Users Based On Emotion Detection Through Facial Features", Ieee,2020.
[3] S Metilda Florence1 And M Uma2, "Emotional Detection And Music Recommendation System Based On User Facial Expression",Iop Conference Series: Materials Science And Engineering,2020.
[4] Sushmita G. Kamble And Asso. Prof. A. H. Kulkarni, "Facial Expression Based Music Player", Intl. Conference On Advances In Computing, Communications And Informatics (Icacci), Sept. 21 24, 2016.
[5] B. Nareen Sai1, D. Sai. Vamshi2, Piyush Pogakwar3, V. Seetharama Rao2, Y. Srinivasulu3 1, 2, 3students, 4asst. Professor, 5assc., "Music Recommendation System Using Facial Expression Recognition Using Machine Learning",International Journal For Research In Applied Science & Engineering Technology (Ijraset), 2022.
[6] S Metilda Florence1 And M Uma2, "Emotional Detection And Music Recommendation System Based On User Facial Expression",Iop Conference Series: Materials Science And Engineering,2020.
[7] Ziyang Yu1, Mengda Zhao1, Yilin Wu1, Peizhuo Liu1, Hexu Chen,Research On Automatic Music Recommendation Algorithm Based On "Facial Micro-Expression Recognition",Proceedings Of The 39th Chinese Control Conference July 27-29, 2020.
[8] Dr. Sunil Bhutada, Ch. Sadhvika, Gutta.Abigna, P.Srinivas Reddy, "Emotion Based Music Recommendation System",Jetir April 2020.

[9] Mikhail Rumiantcev, Oleksiy Khriyenko, "Emotion Based Music Recommendation System",_Proceeding Of The 26th Conference Of Fruct Association.
[10] Krupa K S,Kartikey Rai,Ambara G,Sahil Choudhury, "Emotion Aware Smart Music Recommender System Using Two Level Cnn",Third International Conference On Smart Systems And Inventive Technology (Icssit 2020).
[11] Karthik Subramanian Nathan∗, Manasi Arun†And Megala S Kannan,Emosic- "An Emotion Based Music Player For Android",2017 Ieee International Symposium On Signal Processing And Information Technology (Isspit)



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)