Proceedings of the 7th International Conference on Intelligent Computing and Control Systems (ICICCS-2023)
IEEE Xplore Part Number: CFP23K74-ART; ISBN: 979-8-3503-9725-3

A Novel Emotion based Music Recommendation System using CNN

Tejaswini Priyanka V
Department of computer science and engineering
Gokaraju Rangaraju Institute of Engineering and Technology
Hyderabad, India
priyankavedula20@gmail.com

Y Reshma Reddy
Department of computer science and engineering
Gokaraju Rangaraju Institute of Engineering and Technology
Hyderabad, India
reshureshmareddy02@gmail.com

Dharani Vajja
Department of computer science and engineering
Gokaraju Rangaraju Institute of Engineering and Technology
Hyderabad, India
dhaanivajja@gmail.com

G Ramesh
Department of computer science and engineering
Gokaraju Rangaraju institute of engineering and technology
Hyderabad, India
ramesh680@gmail.com

S Gomathy
Department of computer science and engineering
Gokaraju Rangaraju Institute of Engineering and Technology
Hyderabad, India
gomathy1420@gmail.com

2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS) | 979-8-3503-9725-3/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICICCS56967.2023.10142330

Abstract— **Music has a unique emotional connection with humans. It is a means of connecting individuals from all over the world. On the other hand, it is a highly difficult task to generalize music and claim that everyone would prefer and enjoy the same type. Emotion-based music selection is important because it can assist humans in reducing stress. Its major purpose is to accurately predict the user's emotions, and play songs depending on the user's preferences. Using Human Computer Interaction (HCI), the proposed bot recognizes human emotions from facial emotions. Another significant challenge is the extraction of facial features from the user's face. The proposed CNN Algorithm is utilized in the proposed model to properly capture and recognize the user's face from the live webcam stream and to detect emotions based on facial factors such as lips and eyes. Also, an additional option will be provided for people to make a good choice manually.**

Keywords— Face Recognition, Image Processing, Emotion Detection, Music, and Mood detection.

I. INTRODUCTION

Music and emotion have a strong connection; it can be influenced by each other. A common way for people to express their emotions is through facial expressions. At the same time, certain music can change a person's emotional state. Emotion-based music recommendation is much needed as it helps people to relieve stress and listen to relaxing music that suits their current emotions. The primary objective of this research work is to capture users' emotions through facial expressions. Here, the music player intends to capture human emotions using the computer's webcam feature. Proposed application takes a picture of the user and then the image processing techniques extracts the features of the face and attempts to recognize the emotion that the person is attempting to express. [1]

The goal of researchers is to equip computers with remarkable perceptual skills, enabling them to collaborate with humans as close companions. This involves expanding the range of abilities of computers to facilitate human-like interaction, such as identifying human traits, speaking, listening, and even inferring human emotions. By using sophisticated video cameras and microphones, researchers are implementing a non-invasive method of detecting user behavior through enhanced sensory abilities. This method enables computers to comprehend a user's intentions, gaze direction, and even perceive their physical or emotional state. The ultimate objective of the emotion recognition system is to accurately recognize human emotions. [2]

When users have hundreds of songs, it becomes challenging to manually create and organize playlists. Additionally, it can be hard to keep track of all the songs, and unused ones may unnecessarily occupy device memory, necessitating the user to remove them manually. Users also have to choose songs manually based on their preferences and mood, and rearranging and playing music can be difficult if the play style varies. To address these issues, the project has incorporated machine learning techniques, involving facial scans and feature tracking [11], to determine the user's emotion and offer a personalized playlist based on it. The emotion recognition module plays a vital role in identifying the user's emotions and provide entertainment in the form of emotion-based music. The application comprises three primary modules: questionnaire, mood detection, and music recommendation

II. LITERATURE SURVEY

Charles Darwin, an eminent scientist, recognized facial expressions as an indicator of human emotions, intentions, and ideas. Rosalind Picard presented the importance of emotions to the computing community in 1997. Affective computing has two components: one to enable computers to recognize emotions and allow them to categorize the emotions. Facial emotions are important in the decision-making process and emotion recognition is an important step towards developing an adaptable computer system. Efforts have been directed towards developing a smart,

adaptable computer system capable of detecting a user's emotional state. Incorporating emotions into computers can also increase computer users' productivity. Dryer and Horowitz (1997) revealed that people with similar or complementary personalities collaborate more effectively. Furthermore, Dryer (1999) demonstrated that people perceive their computers with personalities. Therefore, it is vital to develop computers that can work well with their users.

In 2010, Renu Nagpal, Pooja Nagpal, and Sumeet Kaur introduced a novel method for emotion detection in heavily corrupted noisy environments by combining Mutation Bacteria Foraging Optimization (MBFO) and Adaptive Median Filter (AMF) in a cascading manner. The approach involved eliminating noise from the image using MBFO and AMF, followed by identifying local, global, and statistical features from the image. The researchers discovered that the proposed technique was suitable for identifying emotions in the presence of salt and pepper noise levels as high as 90%. Future research may entail utilizing the same technique to detect emotions in the presence of other forms of noise.

III. PROPOSED MODEL

The proposed MRS (Music Recommendation System) application incorporates an emotion detection module as its primary feature. The emotion detection module is crucial in identifying the user's emotions, which in turn provides entertainment in the form of music according to their mood. The application comprises three main modules: Questionnaire, Mood detection, and Music recommendation. Upon opening the application, the user is presented with two options: selecting a mood from a questionnaire and clicking on the recommend button to send the message, or clicking on the "detect emotion" button, which triggers the chatbot application to start the emotion detection process.

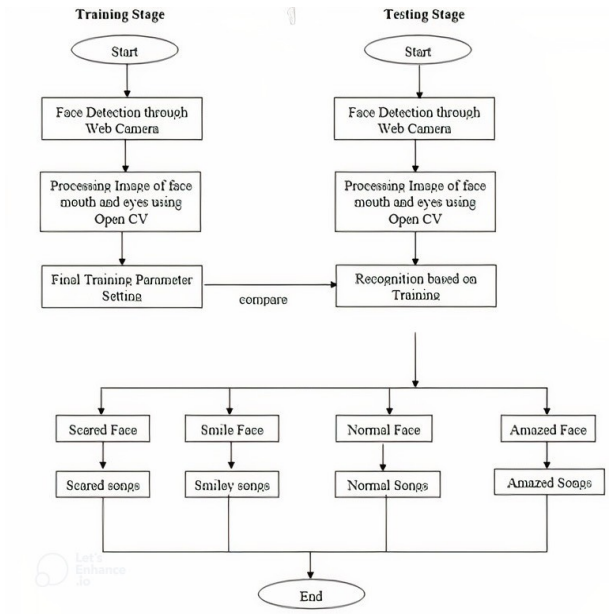

Fig.1. Flowchart of proposed system [16]

IV. IMPLEMENTATION

Implementation steps:

1. Open anaconda
2. Select project environment
3. Launch spyder
4. Run app.py
5. Go to Microsoft edge and type localhost:5000
6. Click sign up and register yourself
7. Using the registered email, user can sign and fill out the form which exists on the next page
8. After navigating to the button (sit straight and put your face in the middle else it will show Please look steadily)
9. After getting the maximum emotion which is captured will be shown on the next webpage if it's wrong, there is a questionnaire also to know the emotional state.

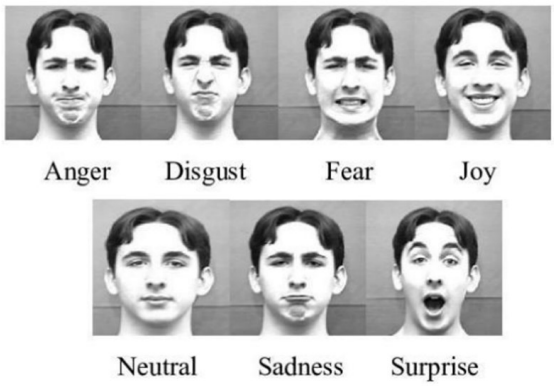

Fig. 1. The seven basic emotions

Fig.2 The seven basic emotions

METHODOLOGY

A. OPENCV [12]

OpenCV is a large open-source image processing data library. In OpenCV, the CV is a shortened form of Computer Vision, which is a computational concept first surfaced in the 1950s, when neural networks were used to locate the edges of items, subsequently progressing to written-by-hand text, dialogue, and languages. OpenCV currently includes the datasets for face recognition. OpenCV is defined as a field of study that assists computer frameworks with photographs, and video recordings. OpenCV supports Python, C++, Java, and other programming languages. [7]

B. CNN (CONVOLUTIONAL NEURAL NETWORK) [13,14,15]

Convolutional Neural Networks (CNNs), the advanced version of Artificial Neural Networks (ANNs) can be trained to produce the results. The Convolutional Neural Network (CNN) model used here recognizes the entire face of the user and taken it as an input and eliminates the noise/error from more profound face designs. Based on the estimated value, the exact projections will be made.

The next phase includes fine-tuning analysis of face locations. At each of the three layers, different convolutional layers are combined to improve accuracy and efficiency. Before collecting the information and develop predictions for images it has never seen, the model is trained with a huge number of photos. [5] CNN is divided into two sections:

First is hidden layers and feature extraction- In this segment, the network will perform a series of convolutions and pooling to recognize the facial features.

Second is classification to highlight and classify the obtained features.

C. MAX POOLING

The technique of selecting the most extreme component from the element map is referred to as max pooling. As a result, the maximum pooling layer generates a component map that includes the most highlights from the prior element map. The input image should be 28 × 28 pixels in size. While zoom into an open field of 5*5, a few highlights in this image will be recorded. While eliminating few highlights, MaxPooling is recommended. Since the images will have edges and inclinations, Max pooling cannot be implemented in the early stages of convolutional neural network [6]

D. CONVOLUTION

Convolution is a numerical methodology that permits two arrangements of data to be consolidated. Convolution is applied to the information on CNN to channel the data and assemble an element map. To perform convolution, the portion iteratively crosses the info picture, performing grid increase component by component. Assuming that the CNN has layers upon layers, the preparation stage will consume more processing time. To deal with and train the neural networks, a Convolution requires a huge dataset.

The figure below shows the structure of the CNN.

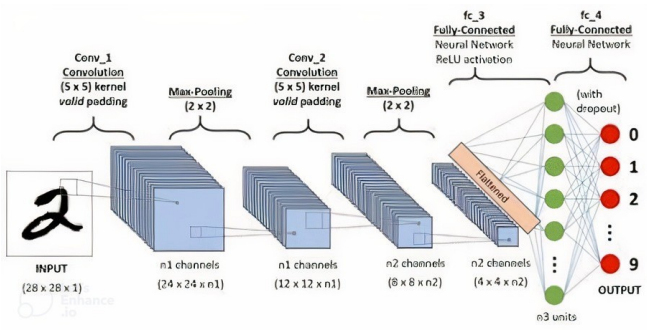

Fig.3. CNN Structure

In the proposed model, CNN Algorithm has been used for accurately detecting the user's face from the live webcam feed and to detect the emotion being expressed by the user from the facial features.

Traditionally, the Haar Cascade algorithm is used to detect emotions from facial features extracted from the user but as research suggests CNN can perform both detection and extraction of facial features using concepts like convolution and max pooling. Henceforth CNN has been used.

E. EMOTION RECOGNITION

Emotional acknowledgement refers to the relationship that includes recognizing human emotions. Individuals' ability to recognize the emotions of others varies widely. Emotion recognition is one of the several facial features. Face recognition software is being employed to allow modification in order to explore and deal with a human's face. AI recognizes various appearances in order to connect them with more information. This may be used for a variety of reasons and allows professionals to recognize an individual's sentiments. [9]

F. FACIAL EXTRACTION

Data extraction is the most important stage in design identification and data mining. Using specific rules, the significant element subset is extracted from the initial information at this step. Since all removed highlights may not significantly add to the structure, it is preferable to eliminate suitable element space [10]. It also determined that only a few pieces are generally required and selected. As a result, a massive amount of data can be reduced to work computationally faster and relatively smaller. As a result, trained element selection is a critical step towards achieving efficient face recognition and biometric verification. The primary aim of component extraction is to reduce machine preparation and computational complexity in order to achieve aspect reduction [4].

G. ARTIFICIAL INTELLIGENCE

Artificial Intelligence (AI) mimics human understanding in computers, allowing them to think like people and replicate their behaviours when programmed to do so. The sample features of AI mimic the features of human brain like reasoning and problem-solving. Despite the fact that multidisciplinary research covers a variety of opinions, advances in AI are prompting a paradigm shift in almost all the fields.

V. RESULT

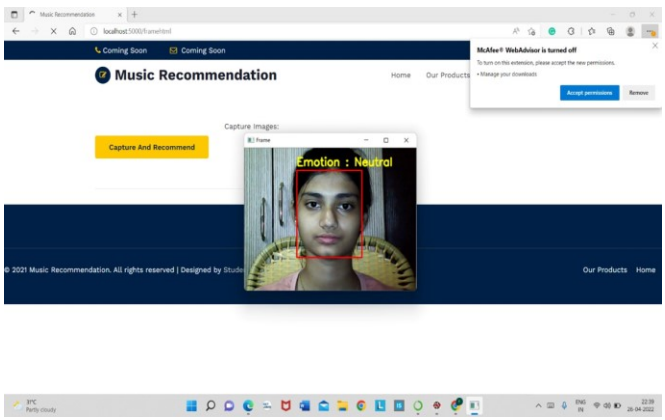

Fig.4. "Neutral" mood

Fig 4 shows the detection of emotion through webcam

when clicked on capture and recommend. Users can confirm the emotion to be correctly detected else may recapture it. Neutral emotion is detected when there is no change in facial features.

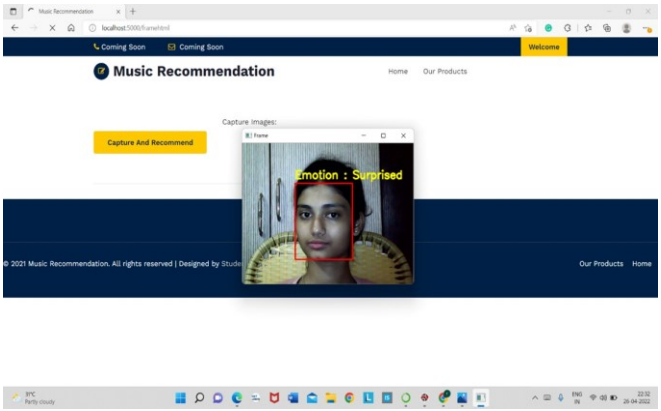

Fig.5. "Surprised" mood

In Fig 5, "Surprised" emotion is detected based on few changes in facial features like eye brows, mouth, nose etc.

For the respective moods detected, a list of songs from Spotify playlists is played in the Spotify app.

A. Metrics:

- Accuracy

  Accuracy = TN + TP/ TP+FP+TN+FN

  When the "detect emotion" option is selected, songs appropriate for the user's emotion are played with an accuracy of approximately 96%~.

  Accuracy for the proposed CNN algorithm is 97.21%.

- Loss

  Loss = 0.0215

- F1

  It is the harmonic mean of precision and recall. Overall to avoid Type 1 errors more than Type 2. To do so, there's an F1 Score. F1 Score is the harmonic mean of precision and recall.

  Precision = TP/ TP+FP = 100

  Recall = TP/ TP+FN = 97.18

  F1 Score = 98.57

VI. CONCLUSION AND FUTURE SCOPE

This study has successfully developed automatic facial expression recognition system for the purpose of creating an emotion-based music player. Facial emotion analysis has been extensively researched and applied, beginning with psychological research. Manual face analysis previously employed by psychologists has now been replaced by suitable computer software. Various image processing algorithms have been developed to meet the demands of the facial emotion recognition system. This project not only covers the theoretical foundation, but also provides a framework for designing and implementing an Emotion-based music player. The proposed system processes videos, extracts facial expressions, recognizes basic emotions, and displays a list of songs based on the emotions. Simple methods were used to develop the proposed music emotion recognition system. It extracts a person's facial expressions, such as happiness, anger, surprise, and neutrality to offer music suitable for the individual's emotion. Although the system cannot handle major head rotations and obstacles, it allows head movements. The future work will be dedicated to improve the recognition rate of the system.

ACKNOWLEDGMENT

We are very grateful to our Project guide Mr. G Ramesh for his extensive patience and guidance throughout our project work. We express our immense gratitude to all who have directly or indirectly contributed to our needs at the right time. This has been for the development and success of our project work.

REFERENCES

[1]. Dr. Vishwanath Kharad: "Research Paper on the Chatbot Development for Educational Institute" Social Science Research Network (SSRN), 2020.

[2]. Abhigya Verma, Chandana Kuntala, Pragya Khatri, Sristi, Sukhmani Kaur, A K Mohapatra, Shweta Singhal: "University Chatbot System using NLP" Social Science Research Network(SSRN), 2022.

[3]. Payal Jain: "College Enquiry Chat Bot using Iterative Model" International Journal of Scientific Engineering and Research (IJSER), 2019.

[4]. Harshala Gaude, Vedika Patil, Prachi Vishe, Sonali Kolpe: "College Enquiry Chat-Bot System" International Journal of Engineering Research & Technology(IJERT), 2020

[5]. Saksham Saraswat, Siddhartha Mishra, Vikas Mani, Shristi Priya: "GALGOBOT – The College Companion Chatbot" International Conference on Intelligent Computing and Control Systems(ICICCS), 2021

[6]. Neelkumar P. Patel, Devangi R. Parikh, Darshan A. Patel, Ronak R. Patel: "AI and Web-Based Human-Like Interactive University Chatbot (UNIBOT)" International Conference on Electronics, Communication, and Aerospace Technology (ICECA), 2019

[7]. Udhayakumar Shanmugam, Sowjanya Mani, Sneha Sivakumar, Rajeswari: "Human-Computer text conversation through NLP in Tamil using Intent Recognition" International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN), 2019

[8]. Kumar, S.K., Reddy, P.D.K., Ramesh, G., Maddumala, V.R. (2019). Image transformation technique using steganography methods using LWT technique. Traitement du Signal, Vol. 36, No. 3, pp. 233-237. https://doi.org/10.18280/ts.360305

[9]. Y. Sara, J. Dumne, A. Reddy Musku, D. Devarapaga and R. Gajula, "A Deep Learning Facial Expression Recognition based Scoring System for Restaurants," 2022 International Conference on Applied Artificial Intelligence and Computing (ICAAIC), Salem, India, 2022, pp. 630-634, doi: 10.1109/ICAAIC53929.2022.9793219

[10]. Somasekar, J Ramesh, G "Beneficial Image Preprocessing by Contrast Enhancement Technique for SEM Images", IJEMS Vol.29(6) [December 2022], NISCPR-CSIR, India.

[11]. G. Ramesh et al., "Feature Selection Based Supervised Learning Method for Network Intrusion Detection", International Journal of Recent Technology and Engineering (IJRTE), ISSN: 2277- 3878, Volume-8, Issue-1, May 2019.

[12]. Ramesh, G., Anugu, A., Madhavi, K., Surekha, P. (2021). Automated Identification and Classification of Blur Images, Duplicate Images Using Open CV. In: Luhach, A.K., Jat, D.S., Bin Ghazali, K.H., Gao, XZ., Lingras, P. (eds) Advanced Informatics for Computing Research. ICAICR 2020. Communications in Computer and Information Science, vol 1393. Springer, Singapore. https://doi.org/10.1007/978-981-16- 3660-8_52

[13]. Authors: Rahul Chauhan, Kamal Kumar Ghanshala, R.C Joshi

"Convolutional Neural Network (CNN) for Image Detection and Recognition" (ICSCCC 2018).

[14] Sivakumar Depuru, Anjana Nandam, S. Sivanantham, K Amala, V. Akshaya, M. Saktivel, "Convolutional Neural Network based Human Emotion Recognition System: A Deep Learning Approach", 2022 Smart Technologies, Communication and Robotics (STCR), pp.1-4, 2022.

[15] Natarajan, V.A., Kumar, M.S., Patan, R., Kallam, S. and Mohamed, M.Y.N., 2020, September. Segmentation of Nuclei in Histopathology images using Fully Convolutional Deep Neural Architecture. In 2020 International Conference on Computing and Information Technology (ICCIT1441) (pp. 1-7). IEEE.

[16] Roy, Dharmendra, Anjali CH, G. Kavya Sri, B. Tharun, and K. Venu Gopal. "Music Recommendation Based on Current Mood Using Ai & Ml." (2023).