A DEEP LEARNING APPROACH TO MUSIC RECOMMENDATION BASED ON FACIAL EMOTION RECOGNITION

Sowmya Kannan
(RA1911026040059)
Dept of CSE, SRM Institute of Science and Technology, Vadapalani, Chennai, India.

Anju Vilashni
(RA1911026040048)
Dept of CSE, SRM Institute of Science and Technology, Vadapalani, Chennai, India.

Yegu Sanjana
(RA1911026040008)
Dept of CSE, SRM Institute of Science and Technology, Vadapalani, Chennai, India.

ABSTRACT: *A user's emotional state can be detected by their facial expressions and speech however our model is solely based on the facial expressions of the user. These expressions can be captured using a live camera. Facial recognition systems are now working throughout the world by governments and private companies. Different algorithms that are provided by machine learning can be utilised to recognise human emotions. Additionally, music has an effect on us in a way that other noises don't; it can help someone feel and think better. It has the ability to treat conditions like sleeplessness, depression, and anxiety. The system's main goal is to properly recommend a playlist by identifying facial emotion from the user's image as collected by the live camera. The suggested system will save time and money.*

KEYWORDS: Recognition, Artificial intelligence, OpenCV Application, Convolutional Neural Network.

I. INTRODUCTION

AI is a field of technology that uses computers to mimic human intelligence, allowing them to make defensible decisions. AI plays a significant role in a person's daily life and has opened up a new area of the rapidly developing technology, from self-driving automobiles to recognising faces on your mobile lock screen. A popular application of AI is facial recognition. It is one of the more sophisticated types of biometric authentication that uses the live camera to recognise and verify a person using facial features in an image or video from a database. The programme initially looks for human eyeballs before moving on to the lips, nostrils, and iris. Once all the facial traits have been recorded, further verification is carried out using pre-trained models that contain both positive and negative images to ensure that the face is indeed that of a human. Facial recognition technology has seen a rise in interest and investment over the past few years.

By utilising a music recommender system, the music service provider is able to anticipate the user's musical preferences and then present them with relevant songs based on their prior musical exposure. The general idea mentioned in the article is a system that will recommend playlists by identifying the user's mood from facial expressions. Future applications of emotion detection in robots and other areas will allow for effective sentimental analysis without the need for a human analyst.

Music Recommendation can be applied in wide range of areas such as studying, stress, music therapy and many others. Several research has been carried out in the field of music to find out the influence they have on the physiological and emotional state of a human and it has been concluded that it has a significant and remarkable impact on human feelings and thoughts. Also, music therapy is considered as an effective improvement given to humans in the treatment of depression and anxiety.

II. **RELATED WORK** (Existing Algorithms)

A. Eigenface based algorithm

The most popular face recognition algorithm use Eigenface based methods. It is renowned for simplicity, having low sensitive and doing better performances involving small databases. This approach makes advantage of the location of eyes, nose and mouth on a face and the relative distance between these objects on the face. Principal Component Analysis is mathematical technique that can be used to extract these distinctive characteristics referred to as eigenfaces. Generally, a face is identified as a face by calculating the relative distance of eigenfaces.

B. Artificial neural network - based algorithm

Face recognition software widely uses Artificial Neural Networks. Once a face has been discovered, an ANN will be used to determine the face's identity and recognition by weighing the facial data. ANN imitates the biological neuron system in the human brain. The neuron transmits signals to all the neurons in the subsequent layer after receiving a signal from the layer above. The many types of ANN include radial basis function networks, back propagation neural networks and feed forward neural networks.

C. AdaBoost Algorithm

In 2001, Viola and Paul proposed the first real-time face identification algorithm by using the AdaBoost algorithm to detect faces. The AdaBoost face detection algorithm is a combination of the AdaBoost algorithm, the cascade classifier, and the integral image. It begins by swiftly calculating the face's Haar-like features using the integral image. Then, it builds an AdaBoost classifier using the week classifiers produced by Haar-like features in accordance with weighted voting. Finally, it creates a cascade classifier using the powerful classifiers to speed up classification detection.

D. Hybrid approach of Music Recommendation System

Recommendation system generally uses either collaborative filtering or content-based filtering. The weakness with this approach is the cold start problem. Insufficient user history records or lack of access to many songs will lead the recommendation results to be inaccurate and affect the system in suggesting appropriate songs to the user. as the system has not gathered sufficient information it affects in providing a set of songs to the user.

III. **METHODOLOGY**

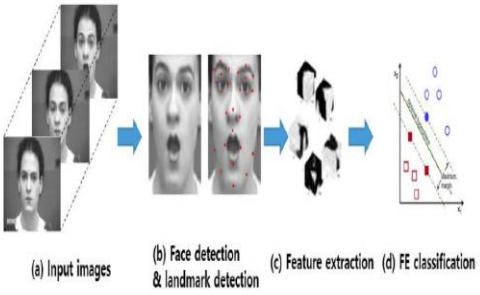
The facial emotion recognition-based music recommendation system is a programme that focuses on using a live camera to implement mood detection in real time. It is a working prototype for a brand-new item that has two key components: mood detection and music recommendation.

A. Mood Detection Module

This Module is divided into two parts:

- Face Detection – The ability to recognise a face in any image that is captured and to differentiate between a face and an object in that image. The bounding box, a fictitious rectangle that serves as a reference point, and the coordinates of the faces that were found are the outputs. The Python library OpenCV was taken into consideration for this work. OpenCV is a comprehensive and enormous open-source framework for image processing, machine learning, and computer vision. It is capable of processing live camera photos to recognise human faces and objects. Human eyes build a lot of information from what they see. Machines are given the ability to see everything; they translate the images that are acquired by cameras into numbers, store them in memory, and then translate images into numbers using the pixel value. The tiniest component of a computer picture or piece of graphics that may be displayed or represented on a digital display device is a pixel. The numbers indicate the strength of the image at a specific point. The computer extracts that value from each pixel and stores the findings in an array for interpretation and result generation. The channels are arranged by OpenCV in RGB order. An RGB is a mixture of the colours red, green, and blue that results in a new colour.
- Mood Detection – The emotions on our face can be classified as happy, angry, sad

or neutral. We are utilising convolutional neural networks for this task. This is composed of two parts: the first concentrates on removing the background from the image, and the second focuses on extracting the facial feature vector from the image. The first step in this process is to detect a image and draw a rectangle around it and the next step is to detect landmarks around it and the next step is to detect the landmarks in the detected face region. The third step is extracting the geographical and temporal features from the facial components. The final step is to use a Feature Extraction classifier and produce the recognition results using extracted features. This system solely focuses on seven different facial expressions which are happy, sad, fear, anger, disgust, neutral, surprised. The algorithm mainly aims on identifying the facial features of the given image and try to characterize them into these six categories.



FER procedure for an image

B. Music Recommendation Module

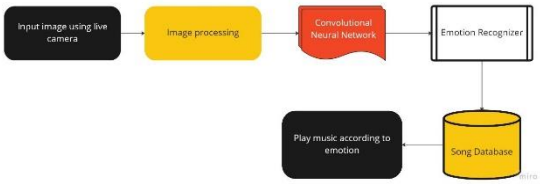
The English-language dataset of songs for the designated moods was discovered. With Spotipy, a simple Python method for the Spotify web API, you may access all of the music data the Spotify platform offers. With this tool, users can access Spotify songs and use them in any recommender system they choose. Within the Spotipy, there are two different types of authentications that take place.

IV. SYSTEM ARCHITECTURE

First, the suggested framework is set up to recognise a face in a frame taken live by the camera. The image is handled once the information picture has been identified. After that, the image is made available to OpenCV, which use machine learning methods to look for faces in images. There are already numerous classifiers for eyes, noses, and other features in OpenCV.

It is necessary to convert the BGR-captured photos to grayscale because the detection can only be performed on grayscale images. It requires the input picture, scaleFactor, and minNeighbors as its three inputs.

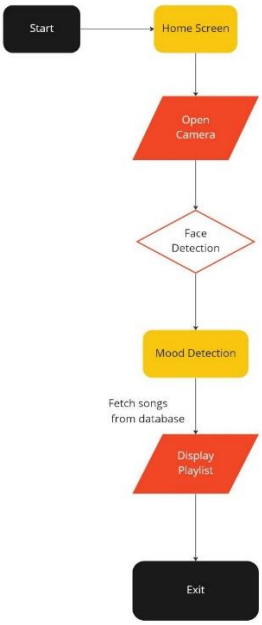
CNN analyses the detected frame to record the emotion on the recognised face.



Block Diagram

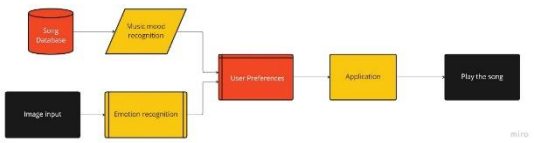
V. MODULE IDENTIFICATION

Face detection and recognition: A person's facial expressions are vivid indicators of their emotional and physical state. In this section, we'll talk about how photographs of human faces can be analysed to identify the emotions they convey. For facial recognition, a variety of algorithms are employed. In this case, the face in the image is being detected using OpenCV. For face attribute analysis, CNN is utilised.



Data Flow Diagram of the proposed system

By running the main webpage, it will trigger OpenCV which helps in capturing the images from the live camera. The next phase is where the captured images is being processed for the purpose of detecting the emotion being shown on the human's face which is done by a pre trained model to identify the accurate emotions and once the emotion is detected each emotion is stored as a key-value attribute. With the help of Spotify API app, it gets the track id and track features and then it updates into a separate csv file for each emotion which is then given a key value again and both the key values are compared to suggest the songs to the user based on the emotion been recognized.



Module Explanation

VI. SOFTWARE AND HARDWARE REQUIREMENTS

Hardware Requirements

The physical computer resources, usually referred to as hardware, the necessary hardware for this project is as follows:

- Operating System: Windows/Linux/IOS
- RAM: 8 and above
- Processor: i5 and above
- Webcam
- Minimum 16-pixel resolution camera

Software Requirements

Software requirements are concerned with outlining the pre-requirements and software resource requirements that must be installed on a computer in order for an application to work perfectly. For this project, the following software prerequisites are necessary:

- Python 3.10
- Flask 2.1.2
- Spotipy 2.19.0
- Pandas 1.4.2
- opencv_python_headless 4.5.5.64
- numpy 1.22.3
- tensorflow 2.9.0

VII. RESULTS AND DISCUSSIONS

Accurately identifying human emotions and moods is challenging because each person has distinctive face traits. However, with appropriate facial expressions, it can be recognized with some degree of accuracy. Here we propose a simple music recommendation system using facial emotion recognition. It extracts the emotions (happiness, anger, sadness, neutrality) of people's faces and proposes music.

This system detects your mood in real time with a live camera and presents playlist that match the mood. If you have a stable internet connection you can get songs from the recommended playlist.

S.No	Existing Algorithms	Accuracy (%)
1.	Eigenface based Algorithm	96%
2.	ANN based Algorithm	84.6%
3.	AdaBoost Algorithm	80%
4.	Hybrid Approach of Music Recommendation	80.7%
5.	A DL Approach to Music Recommendation System based on Facial Emotion Recognition	93.7%

Accuracy Comparison

VIII. CONCLUSION

Even though human emotions are precise and complex to be predicted by any machine with the appropriate dataset been fed to the model it is possible to predict or differentiate human emotions. The expressions on a person's face can be predicted with the help of a human's eye, nostrils etc. Once the expression is captured by the model and the expression been identified correctly the system will be able to suggest songs based on the person's mood. Because music has a significant impact on the user's emotional and physiological state, this system was created with the goal of fostering user-system interaction. Recent progress indicates that the suggestions system will have a broad scope.

REFERENCES

[1] Ankita Mahadik, Prof. Vijaya Bharathi Jagan, Shambhavi Milgir, Vaishali Kavathekar, Janvi Patel.(2021).Mood based Music Recommendation System. Don Bosco Institute of Technology Mumbai, India.

[2] Deny John Samuvel, B. Perumal, Muthukumaran Elangovan.(2020). MUSIC RECOMMENDATION SYSTEM BASED ON FACIAL EMOTION RECOGNITION.Department of Electronics and Communication Engineering, Kalasalingam Academy of Research and Education, Krishnankoil, Virudhunagar Dt., (India).

[3] NURULHUDA ISMAIL, MAS IDAYU MD. SABRI. REVIEW OF EXISTING ALGORITHMS FOR FACE DETECTION AND RECOGNITION. Department of System and Computer Technology (Multimedia), Faculty of Computer Science and Information Technology, University of Malaya, 50603 Lembah Pantai, Kuala Lumpur, MALAYSIA.

[4] H. Immanuel James, J. James Anto Arnold, J. Maria Masilla Ruban, M. Tamilarasan, R. Saranya. EMOTION BASED MUSIC RECOMMENDATION SYSTEM(2019). Student, Asst.prof Dept. of Information Technology, Valliammai Engineering College, Chennai ,Tamil Nadu.

[5] Dr.Sunil Bhutada, CH.sadhvika, Gutta.Abigna, P.Srinivas reddy.EMOTION BASED MUSIC RECOMMENDATION SYSTEM Assistant Professor, Information Technology Department,Sreenidhi Institute of Science and Technology, Yamnampet, Hyderabad. Bachelor of Technology IVth year, IT Department, Sreenidhi Institute of Science and Technology,Yamnampet, Hyderabad.

[6] Mikhail Rumiantcev, Oleksiy Khriyenko.Emotion Based Music Recommendation System University of Jyväskylä, Jyväskylä, Finland.