Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 4, April : 2024

Song Recommendations by Facial Expressions

(1) Mohammad Sohel Mohammed Yasin Sheikh, (2) Neeraj Sahu

(1) PG Student, Department of Computer Science & Engineering, G. H. Raisoni University, Amravati, Anjangaon Bari Road, Amravati-444701

(2) Guide, Department of Computer Science & Engineering, G. H. Raisoni University, Amravati, Anjangaon Bari Road, Amravati-444701

ABSTRACT

Facial expressions can reveal a user's emotion or mood. These expressions can be obtained from the system's camera's live feed. There is a lot of research being done in computer vision and machine learning (ML), which educates machines to distinguish distinct human emotions or moods. Machine learning provides several strategies for recognizing human emotions. One such technique is to merge a CNN model with Kera's, resulting in a small-scale trained model and simplifying ML integration. Music is commonly used to manage mood, specifically to lift a low mood, increase energy, or relieve stress. Furthermore, listening to the appropriate music at the proper time might improve mental health. The vast majority of people of all ages enjoy music, and we believe that music players should be capable of far more than just playing songs and allowing users to make playlists. As a result, listening to music (premium, ad-free, or with commercials) to unwind after work can benefit one's health. A music player should be intelligent and responsive to the user's preferences. A music player should let users organize and play tracks automatically, with minimal effort required for song selection and reorganization. An Emotion-Based Music Player improves the platform for all music lovers and automates song selection.

Keywords: facial recognition, image processing, computer vision, emotion detection, music, mood detection.

INTRODUCTION

The research seeks to capture emotions communicated through facial expressions using a music player designed to capture human emotion using a web camera interface on computing platforms. The software takes the user's image and utilizes image segmentation and image processing techniques to extract information from a target human's face to determine the

emotion being exhibited. The project's goal is to lighten the user's mood by playing tunes that match their needs while collecting the user's photograph. Facebook expression recognition is the best kind of expression analysis known to man, and all functions in the human body are intricately linked. Music can affect many parts of the human body, including heart rate, blood pressure, and other psychophysical states. Different musical characteristics, such as tempo or beat level, can elicit emotional, psychophysiological, and behavioural responses. Music-evoked emotions have a tremendous impact on people of all ages, from children to the elderly. According to Shahram Heshmet, music can provide young people with a feeling of identity. According to Ryan Hill, music has been recognized as one of the most powerful parts of society, with research indicating that youths aged 14 to 18 who listened to rap music were 2.5 times more likely to be arrested and 1.5 times more likely to engage in illegal activities. According to researchers, 20% of older persons suffer from a mental condition, such as depression or cognitive impairment. They describe depression as harmful to older individuals since it has a direct impact on their emotional, cognitive, and physical health. This discovery is crucial since music is one of the most subtle and powerful methods to influence our emotions.

Overview of the proposed work

Music has a strong emotional connection and has the potential to drastically boost a person's mood. However, the majority of music recommendations are based on user tastes that have developed over time. This research proposes a neural network-based song selection method that uses facial expressions to determine a person's mood. This method is more efficient than previous ones and streamlines the process of searching for and generating a specific playlist. A webcam or camera is used to capture a face, and data is taken from the image to evaluate an individual's mood. Machine Learning techniques are utilized to extract and categorize these motions, which are critical for mood detection. The project's goal is to create a personalized emotion-driven music recommendation system that solves the choice dilemma, discovers new musical pieces, promotes mental and physical well-being, and enhances working processes.

The suggested system combines artificial intelligence techniques with broad music recommendation and therapeutic approaches. The Convolutional Neural Network predicts the emotion based on the acquired pictures and then creates a playlist of music based on the emotion. The primary purpose is to automatically create a music playlist that alters the user's mood, which might be joyful, sad, natural, or surprised. If the topic is about negative emotion,

a playlist with the best music to lift the person's spirits will be provided. This study describes a personalized emotion-driven music recommendation system that detects emotions and generates tailored playlists. This novel approach to music suggestion has the potential to greatly improve consumers' mood and overall experience.

LITERATURE REVIEW

P. J. Rajashree (2018) According to the authors of this research, music is extremely vital in human life and modern advanced technologies. Typically, the user must actively go through the playlist of songs to select one. Here, we propose an efficient and accurate approach that can produce a playlist based on the user's present spirit and conduct. Existing methods for automated playlist building are computationally slow, less accurate, and sometimes require additional hardware such as EEG or sensors. Speech is the most ancient and natural way to express feelings, emotions, and moods, and its processing is computationally, time-consuming, and costly. This system supports real-time extraction of facial expressions as well as collecting audio elements from songs to classify into a given emotion, resulting in an automatically generated playlist with a cheap computational cost.

Cherry, K. (2018) This study offers an intelligent agent that sorts a music collection based on the emotions conveyed by each song before recommending a suitable playlist to the user based on his or her current mood. The user's local music collection is initially clustered and supported by the emotion conveyed by the song, often known as its mood. This is often computed using the song's lyrics as well as the music. When a person wants to acquire a mood-based playlist, they take an image of themselves at that time. This photograph is submitted to facial detection and emotion recognition methods, which identify the user's emotions. The user is then suggested a playlist of music that best matches their emotion.

H. P. Adarsh (2018) The study offers an intelligent agent that sorts a music collection based on the emotions conveyed by each song before recommending a suitable playlist to the user based on his or her current mood. The user's local music collection is initially clustered and supported by the emotion conveyed by the song, often known as its mood. This is often computed using the song's lyrics as well as the music. When a person wants to acquire a mood-based playlist, they take an image of themselves at that time. This photograph is submitted to facial detection

and emotion recognition methods, which identify the user's emotions. The user is then suggested a playlist of music that best matches their emotion.

As a result, the system may adjust to the user's current music preferences. Furthermore, the more the user utilizes the system, the more personalized music is frequently tailored to him or her.

Manas Sambare (2013) I researched numerous advancements in human affect recognition. He concentrated on a variety of ways to deal with audio and/or visual recordings of affective emotions. The paper gives a thorough overview of audio/visual computing approaches. The impact is described as a prototype for several emotion categories such as happiness, sorrow, fear, anger, disgust, and surprise.

Dr. Shaik Asif Husain (2020) This research focused on the issues in computing methods for developing automatic, spontaneous affect recognizers that aid in emotion detection. It also revealed several issues that had been overlooked or avoided in unit-modal posed emotion identification.

P. Lucey (2010) Proposed a proposal for automating interactions between users and music players, which learned all of a user's preferences, moods, and activities and then selected songs as a result. The device recorded the users' varied facial expressions to assess their emotions and anticipate the genre of music. This demonstrated how numerous algorithms and strategies proposed by different authors in their research may be used to link the music player to human emotions. It has therefore helped to reduce users' efforts in generating and managing playlists while also giving an exceptional experience to music listeners by bringing them the most appropriate song based on the user's present expression.

PROPOSED METHODOLOGY

The human face is a significant factor in determining an individual's mood. The camera is utilized to extract the necessary information from the human face. One application of this input could be to extract information to determine an individual's mood. The "emotion" produced from the previously provided data is utilized to generate a list of music. This time-consuming job of manually categorizing or dividing music into distinct lists is minimized, allowing for the

creation of a suitable playlist based on an individual's emotional characteristics. Facial Expression Music Player tries to scan and understand data before constructing a playlist based on the parameters entered. Thus, our proposed system focuses on detecting human emotions to develop emotion-based music players. We discuss the approaches used by existing music players to detect emotions, the approach our music player takes to detect human emotions, and how our system is superior for emotion detection. A brief overview of how our technologies work, playlist construction and emotion classification are also provided.

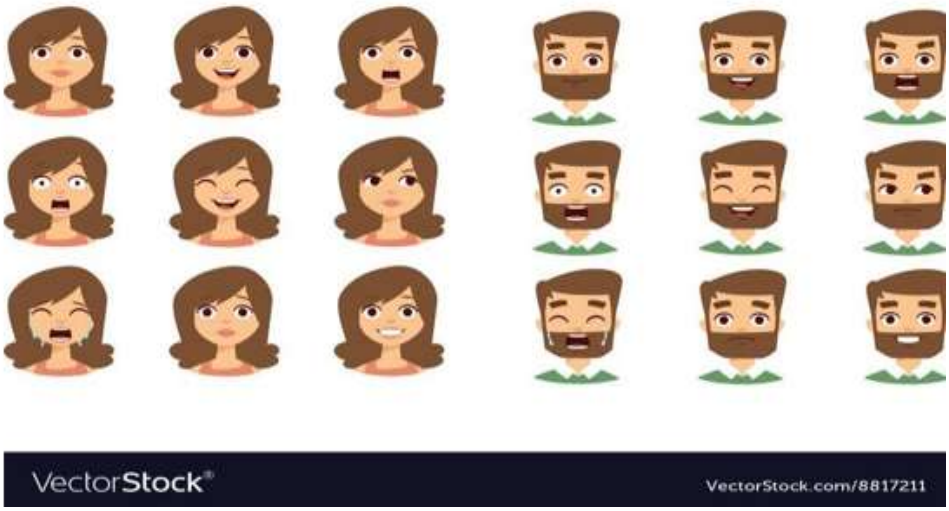


Figure 1. Different Human Emotion

1. System architecture

Humans unintentionally express their emotions, because their faces reflect them. The proposed solution allows us to facilitate interaction between the user and the music system. This project focuses mostly on the user's favoured music, which is recommended based on emotional awareness. In the initial stage of the suggested system, we provided two possibilities, each with its functionality. We have provided a collection of songs and emotions based on spatial recognition. When the application runs, it captures photos using the webcam or any other physiological device. Our main goal in this system is to create a sophisticated music player that may improve the user's mood, and music is one of the best tools for transformation. These photos collected by the system are compared to data sets, and primarily emotions are taken because humans have numerous emotions that are difficult to anticipate because they vary from person to person, therefore seven typical and easily recognized moods of the individual are chosen. Here is another alternate way that can be applied to the core premise, i.e. random

selection of songs that could assist us in brightening our spirits. The other mode is queue mode, which allows us to create our playlists. In all modes, we use individual user data rather than past user data.

The system architecture diagram displays the overall structure of the software system, as well as the relationships, restrictions, and boundaries between its components. When the user launches the web application, the main screen appears, which has two buttons for snapping and using emoji. When the user presses the take snapshot button, the camera opens and the user takes a picture. This image is used as input for the face detection software. If no face is recognized or many faces are detected, the user receives an appropriate error message. After successful single-face detection, the picture is fed into the mood detection module. The detected mood is displayed to the user, following which the play tunes button is enabled. The playlist screen displays the playlist that best matches the observed mood. where the user can choose and play music. If the user presses the utilize emoji button, a screen with six emojis will appear. Users can access the playlist by clicking on any of the emojis. To exit the app, simply touch the back button.

The suggested system can detect the user's facial expressions and extract facial landmarks that are then classed to determine the user's emotion. Once the emotion is identified, the user will be presented with songs that match their emotions. The system structure layout depicts the overall structure of the software system, as well as the relationships, constraints, and boundaries between its components. When you enter the website, you will see a large screen with two buttons: take a snap and utilize an emoji. When the user presses the "take snap" button, the camera opens and the user selects an image. This image is offered for a facial recognition algorithm. If no face is found, the song will not play. When one face is correctly identified, the image is supplied as input to the status detection module. The received mood is displayed to the user, and the "play songs" button is enabled. The playlist screen displays the proper playlist for the status discovered, allowing the user to select and play the song. When a user presses the "use emoji" button, a screen with numerous emoji appears. To find the appropriate playlist, the user can just click on any emoji. To leave the website, simply use the back button.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 4, April : 2024

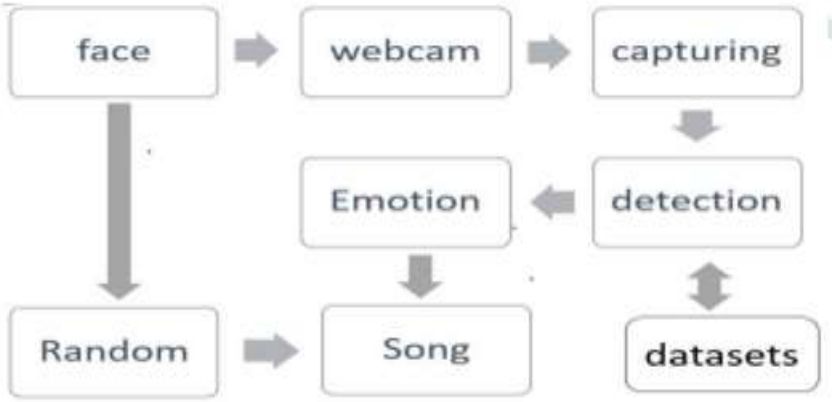


Figure 2. System Architecture

2. Face capturing

The main goal of this session is to take photographs, thus we're using a typical device, a camera, but we could use any other physiological equipment. We use the computer vision library to accomplish this. This makes it easy to interface with other libraries that also use Keras, and it is mostly used for real-time computer vision. When the initial procedure begins, it accesses the camera stream and grabs approximately ten photos for subsequent processing and emotion recognition. So, in the first phase of this project, we'll gather photographs and detect faces. We utilize an algorithm that can classify authentic photographs, and we need a large number of positive images with only faces, as well as negative images with no faces, to train the classifier. The categorized photos are collected as part of the model.

3. Face detection

Face recognition is regarded as one of the most effective ways to identify a person's mood. This image processing system uses the principal component analysis (PCA) method to reduce the face space dimensions before applying Fisher's linear discriminant (FDL) or the LDA method to obtain the feature of the image characteristics; we use this method specifically because it maximizes the training process classification between classes. It aids in the process of picture identification in Fisher's face. When we apply the matching faces method, we utilize the minimal Euclidean, which allows us to classify the user's expression that indicates their mood. Fisher faces with open CV mainly emphasises the class-specific transformation matric so, they don't take illustrative images as the subject and emotion is mainly concluded by the model that

the value evaluated from the process can help us to deduce the mood of the user by comparing the data sets that each emotion is compared with tens of stored images and scale gives the exact emotion so that it can play the music based on the recommendation made by the system by using the following steps. It also does not require any other personal information, unlike other existing apps. Linear classification is a step in the face detection process. It simplifies linear classification rather than SVM. Its purpose is to reduce computational time so that the classification process may be completed more quickly and with more accuracy.

4. Emotion classification

When the face is correctly detected, a box appears and overlays the image, extracting the face for further study. In the following stage, the previously retrieved photos will be processed using the function. The code will extract facial spatial positions from the face image using the pixel intensity values indexed at each point. It compares the supplied data to the stored data to guess which class holds the emotion. If it contains one of the four emotions (angry, sad, neutral, or joyful), and the emotion detection appears to be reducing speed, a command will be executed to reduce the speed of the wheelchair to protect the user.

5. Music recommendation

The input photographs are obtained from the web camera and used to collect real-time images. We've chosen four basic emotions because it's difficult to specify all of them, and utilizing fewer alternatives reduces compilation time and improves the outcome. It compares the values defined as a threshold in the code. The values will be transferred to carry out the online service. The music will be played based on the detected sentiment. Every song is associated with a specific emotion. When the emotion is moved to the appropriate music, the emotions are numbered, organized, and assigned to each song. We can, however, recommend a variety of models due to their accuracy. We are utilizing the winning sound for the sound mechanism, and it is the most generally used Python library for playing basic sounds for the mechanism. The obtained results are compared to the values set as a threshold. There are also other alternatives to the emotion-based method, such as the queue and random modes. In queue mode, we may create a playlist similar to other music software, and the final mode is random mode, which allows us to pick songs at random rather than in sequence, it is also one of the therapies that can help us feel better. When the song is played in response to the emotions of the users, it also conveys those emotions in the form of emoticons in four different emotions.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 4, April : 2024

ANALYSIS AND DESIGN

The mood-based music recommendation system is an application that uses real-time mood recognition. It is a prototype for a new product with two primary modules: facial expression recognition and emoji identification.

1: Facial Expression Recognition Module

This module includes two parts: mood detection and song recommendation.

Mood Detection

a: Face Detection

Capability to identify the location of a face in any supplied image or frame. The result is the discovered faces' bounding box coordinates. For this assignment, the Python library OpenCV was first explored. However, integrating it proved difficult, thus the Face Detector class in Java was suggested. This library identifies people's faces in Bitmap visual objects and returns the number of faces in a given image.

b: Mood Detection

The emotions on the face are classified as pleased, angry, sad, neutral, astonished, afraid, or disgusted. Thus, the CNN architectural model for Image Classification and Vision was utilized. There are various models, but what distinguishes Resnet Net-50 is that it requires extremely minimal computational power to execute or apply transfer learning. CNN was used with Keras to train and evaluate our model for seven categories: happy, angry, neutral, sad, surprised, fearful, and disgusted. We trained it for 25 epochs and obtained an accuracy of around 95-96%.

Music Recommendation The dataset of songs sorted by mood was found on Kaggle in two languages: Hindi and English. A search for a good cloud storage platform to store, retrieve, and query this song data based on the user's request was done. Options such as AWS and Google Cloud were discovered, but they were rejected due to their high cost and limited free storage. The mp3 versions of the songs were manually uploaded to Firebase storage and then connected to the Real Time database based on mood and language.

2: Emoji Detection Module

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 4, April : 2024

We have given our users the option of utilizing emojis to create the playlist. When the user does not want to or is unable to take a snapshot of their mood due to various reasons such as extremely high or low lighting, their camera not working properly, having a lower resolution camera that is unable to take a clear picture of their face, which in turn is unable to detect the proper mood or any other reason, the user can click on the "Use Emoji" button and select the emoji that represents the mood that they are in, or the mood that they want. Figure 1 shows a screenshot of the screen that appears to the user when they click the "Use Emoji" button. The first is the emoji for the mood "happy", the second for the mood "sad", the third for the mood "neutral", the fourth for the mood "angry", and the fifth for the mood "fearful."

4.2 Life Cycle of the System

First, we will launch our web application, after which the home screen will show. Two categories will appear on the home screen: first, with the camera turned on, and second, with emoji. If users select an open camera, they will be able to observe faces with the help of take snap, and based on face detection playlists will be displayed. If the user's mood changes, users can exit or restart again. If the user selects to use an emoji, several weather emoji will be detected based on the user's click on the mood emoji and displayed in a playlist. Enjoy our cheerful motions.

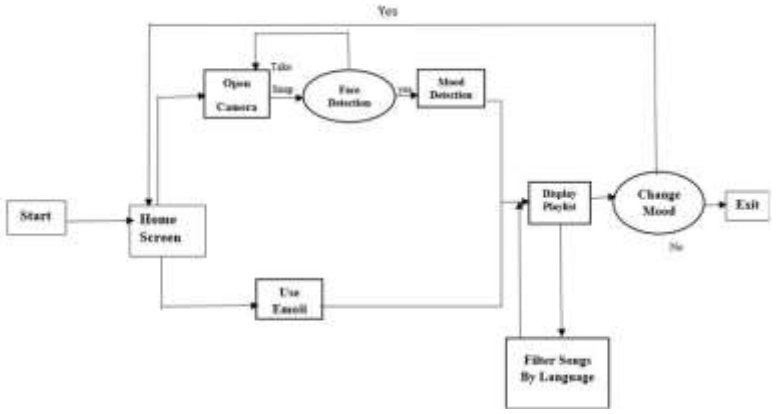


Figure 3. Flow Diagram of Emotion-Based Music Recommendation System

System Requirements

a. Hardware

The most common set of requirements defined by any operating system is the physical computer resources, also known as hardware. The hardware requirements required for this project are:

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 4, April : 2024

RAM	:	Minimum4GB
Hard Disk	:	Minimum 500+

b. Software

Software requirements define the software resources and prerequisites that must be installed on a computer for a program to work properly. These requirements or prerequisites are typically not included in the software installation package and must be installed separately before the software is installed. The software requirements for this project are as follows:

Operating System	:	Windows,
Front End	:	HTML, CSS, JavaScript
Backend	:	Python
IDE	:	Vs code
Used Software	:	Jupiter notebook, Anaconda

Diagrams

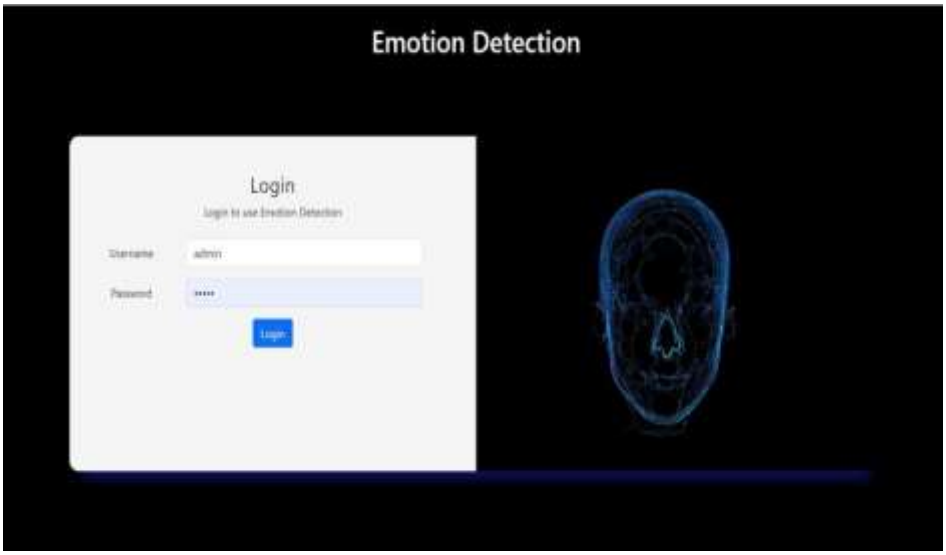


Figure 4. Home Screen

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 4, April : 2024

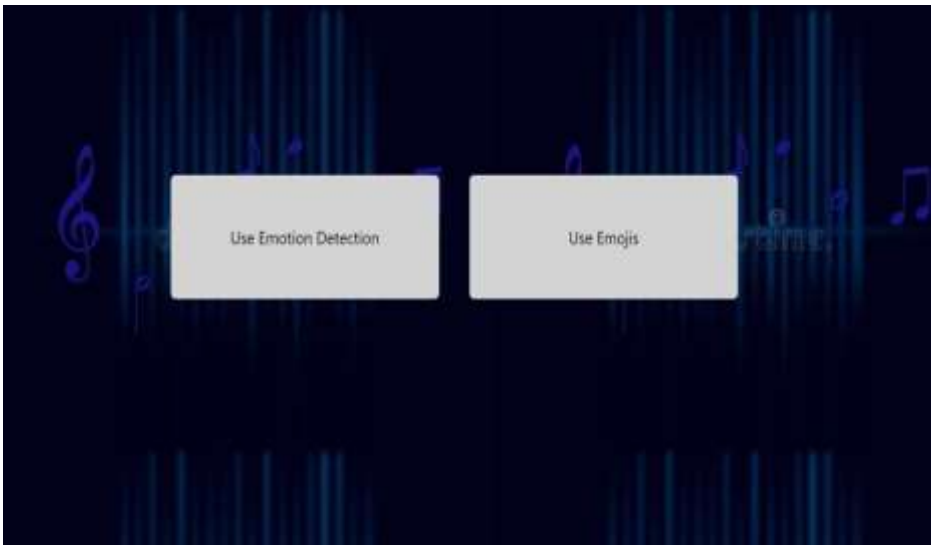
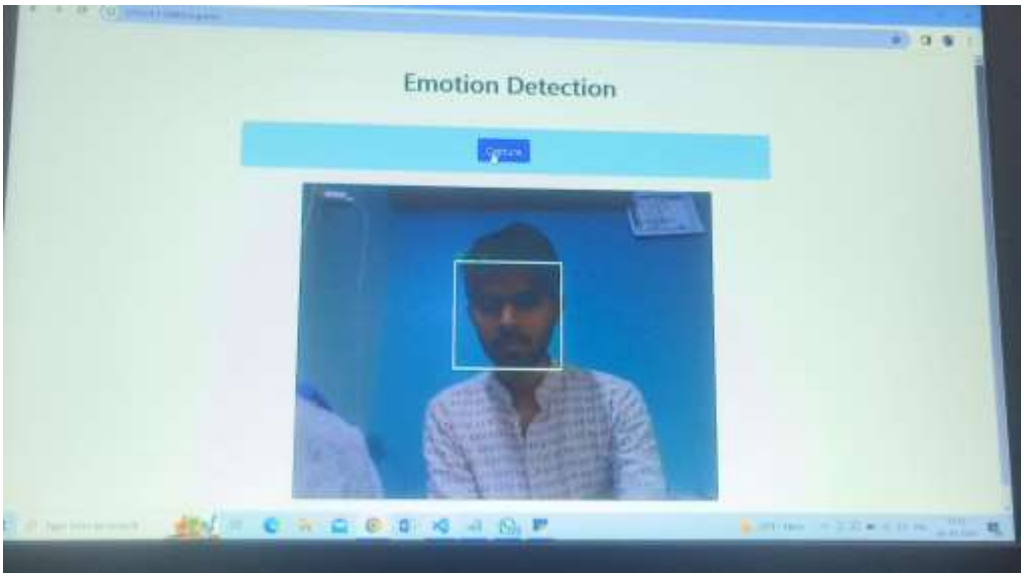
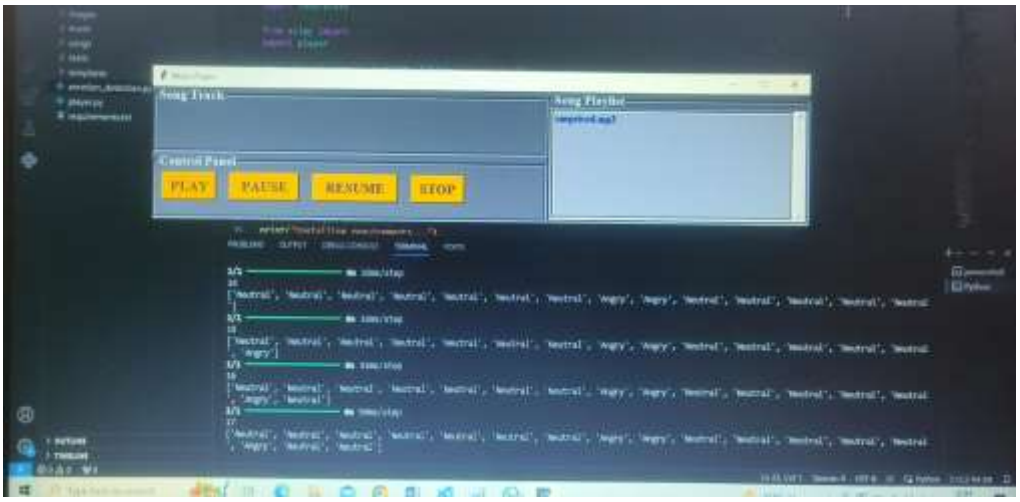


Figure 5. Two Categories

RESULT





It's tough to identify correct human emotion or mood because everyone's face is different. However, it can be identified to some extent with the right facial expressions. The device's camera should have a greater resolution. The web application that we created is up and running, and the screenshot below can be identified to some extent with what was taken while using it. Shows how it correctly detects "neutral" moods, as well as "angry," "sad," "glad," and "fearful" moods.

ADVANTAGES & LIMITATIONS OF THE PROPOSED SYSTEM

Advantages

1. The Emotion-Based Music Player technology will be extremely beneficial to people looking for music based on their moods and emotional behaviour.
2. It will help to shorten the search time. Users can play music with ease.
3. Songs will be played based on the user's mood. Users can also download music via this app.

Limitations

The drawbacks include the proposed system's inability to accurately record all emotions due to the limited availability of images in the image dataset employed. Users may have difficulty capturing their mood or facial expressions for a variety of reasons, including a

malfunctioning camera, improper lighting conditions, and so on.

Need internet access.

CONCLUSION

The purpose of this study is to aid users by proposing songs depending on their emotions using an emotion-based music player. The program employs technology to enhance the system's interaction with users in a variety of ways, making it easier for users to snap photos, measure their moods, and provide personalized playlists. The model has an accuracy of around 96% and can detect seven different moods: anger, contempt, fear, happiness, sadness, surprise, and neutral. The Android app can play music appropriate to the observed mood. Additional measures, like heart rate or body temperature, are required for the accurate detection of fear and disgust states. The project faced problems in modelling and analysis, including unfamiliarity with Python and the use All programming work was completed using free services, with the only cost being model deployment on Render. The performance of a pre-trained model in detecting human emotion was assessed, and the developed model was implemented as a web application. Validation findings revealed that the proposed method can reach an accuracy of 75%.

FUTURE WORK

This project tries to automate technology by training a system to recognize emotions through facial expressions. The concept will centre on making suggestions for films and television shows based on emotional discovery. Emotion recognition based on facial expressions is an important research area, with image processing algorithms playing a critical role in medical and human sciences fields. Researchers are always finding new ways to collect user emotions and treat them. The system has effectively collected user emotions and been tested in real-time scenarios; nevertheless, it must be evaluated under various lighting circumstances to verify its resilience. The system was also able to update its classifier and training dataset with new user photos. The system was designed using the face landmark scheme and tested in a variety of circumstances. The classifier has an accuracy of more than 80% for the majority of test cases, demonstrating good emotion classification. It can accurately predict user expressions in real-time circumstances when tested live. This project is a step toward automation and attempts to create a strong algorithm for emotion recognition in a variety of disciplines.

REFERENCES

1. S, P. J., Rajashree, S. M., S. H., & J, S. S. (2018). Update on Facial Expression-Based Music Player. International Journal of Research and Technology (JERT), 6 (15), 1-3.

2. Cherry, K. (2018). 6 Types of Basic Emotions and Their Effect on Human Behavior. Brought to Mind You: https://www.verywellmind.com/about-us

3. Engineer, A. (n.d.). Create an Android project. Retrieved from developers:https://developer.android.com/training/basics/firstapp/creating-project

4. ED, B. (2015). Hello, Android. In S. D. Pfalzer (Ed.), Hello, Android (Vol. 4, p. 319). Pragmatic bookshelf.

5. H. P., Adarsh, B., A. C., & Kumar, V. A. (2018, April). Emotional Based Music Player. International Research Journal of Engineering and Technology (IRJET), 05 (04), 1-6

6. Manas Sambare, FER2013 Dataset, Kaggle, July 19, 2020. Accessed: September 9, 2020. [Online], Available at: https://www.kaggle.com/msambare/fer2013

7. Mahmudi MA, MMA Facial Expression Dataset, Kaggle, June 6, 2020. Accessed: September 15, 2020. [Online], Available at: https://www.kaggle.com/mahmoudima/mma -facial-expression

8. Drs. Shaik Asif Hussain and Ahlam Salim Abdallah Al Balushi, “Real-time emotional separation and recognition using an in-depth learning model”, 2020 Journal. of Phys .: Conf. Ser. 1432 012087

9. P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar and I. Matthews, "Extended Cohn-Kanade Dataset (CK +): Complete database of action and emotion-specific unit," 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, San Francisco, CA, USA, 2010, pp. 94-101, doi: 101109 / CVPRW.2010.5543262.