JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

EMOTION DETECTION FOR AUDIO PLAYBACK USING AI

[1] Ameyatma Godkhindi Dixit, [2] Bharath Kumar R, [3] Darshan Gaonkar, [4] Mrs. Divya Shree H B

[1]Student, [2]Student, [3]Student, [4]Assistant Professor
[1,2,3,4] Electronics and Communication Engineering,
[1,2,3,4] Dayananda Sagar University, Bangalore, India

Abstract : *Machine learning and Artificial intelligence predict human behaviour patterns and provide an appropriate response in any given situation. As the world becomes increasingly automated, we find ourselves with a plethora of stimuli to choose from. People commonly use facial expressions to express their emotions. Music has always been known to change people's moods. In this paper, we focus on harnessing the power of AI to detect our mood by our facial expression and play a sound track associated with the emotion we are feeling. As we all know, music is a great way to acknowledge our emotions. With this paper, the expression made by us is detected and an appropriate song or music track is played. The main idea behind this paper is to automatically play audio tracks based on the user's emotions. The working is based on the detection of human emotion through the estimation of human expression. An input user image is fed to the system, the face detection and face computation are done by the image processing algorithm. By detecting the facial expression of the user, the algorithm the taps into the song data set from which it extracts a song corresponding to the given facial expression and starts playing that track. In existing systems, the user has to manually select songs, which is a cumbersome process and it may also not match the user's mood. Capturing and recognising a person's emotion and playing appropriate songs that match their mood can gently calm their mind and produce a pleasing effect.*

IndexTerms **- Face Detection, Emotions, Training, Testing, Facial Expression Recognition, Audio Recognition, Detection, Viola-Jones Algorithm, Music Player.**

I. INTRODUCTION

The concept of the emotion-based music player is to allow users to play the songs that they like based on their mood. After analysing the facial expression, the device will then play the songs that correspond to the emotion. A facial expression can be created by one or several gestures, movements, or even positions of the facial muscles. Individuals can control their facial expression and present it according to their wishes, so facial expression can be adopted as a voluntary action. However, because facial expression is so strongly linked to emotion, it is almost always an unintentional activity. An individual's expression may be shown for a few microseconds before returning to a neutral look. Facial expression analysis includes the detection and interpretation of facial motion as well as the recognition of expression. The three methodologies that enabled automatic facial expression analysis were face capture, facial data extraction and representation, and facial expression recognition. The music player will then analyze the facial expression of the user to determine the current mood of the individual. It will then play the songs that suit the individual's current mood. This paper aims to detect the user's emotion and play an appropriate audio track based on the user's estimated emotion. As the world becomes increasingly automated, we find ourselves with a plethora of stimuli to choose from. Artificial intelligence and machine learning predicts the patterns of human behaviour and delivers an appropriate result for any given scenario. In this paper, we focus on harnessing the power of AI to detect our mood by our facial expression and play a sound track associated with the emotion we are feeling. Human expression plays an important role in determining the present state and mood of a private, it helps in extracting and understanding the emotion that a private has supported various features of the face such as eyes, cheeks, forehead and mouth or even through the curve of the smile.

As we all know music is a great way to acknowledge our emotions, with this project, we can express our mood or the emotion we are feeling through an audio track. The detection of our emotion is done by processing an image of our face which is then processed by the AI to determine what mood we are in and what audio track is suitable for such a mood. The main idea behind this project is to perform audio tracks focusing on the user's emotions.
In the current system, the user must manually select the songs because arbitrarily played music may not match the user's mood. Instead, the user must categorise the songs into various emotional states and then manually select an emotional state before playing the songs. The eyes and mouth are some of the most common features that are used in facial image processing. This process involves analyzing and identifying various parts of the face. Face detection is crucial because only if a face is available will it be classified. A graphical-based classification method is also used to recognize expressions.

II. RELATED WORKS

Human emotion recognition is the first component of this system, as research has shown that facial expressions convey the majority of a message. To categorise human emotion and behaviour, numerous techniques and approaches have been proposed and implemented. Using image processing techniques and artificial intelligence, the proposed methods have only concentrated on a few of the basic emotions. Regarding the 'AI-based emotion recognition for audio playback'.

Mood Based Music Player [1], this paper will look at various classification algorithms in order to develop a clear methodology for classifying music into different mood classes and detecting the mood of the user through facial expressions, and then combining the two to create a user-tailored music playlist.

EMOSIC- A Emotion Based Music Player for Android [2], The paper proposes an accurate and efficient model for creating a playlist based on the user's current emotional state. This proposed system, which is based on real-time facial expression extraction as well as retrieving audio features from music to categorise songs into a specific emotion, will automatically generate a playlist at a low computational cost.

Emotion-Based Music Player [3], this paper presented an emotion based music player that can recommend songs based on the emotions of the user. From a smart band or a mobile camera, the application receives the subscriber's heart rate or a facial image. The pulse rate and facial image-based classification methods are both presented in this paper.

Emotion Based Music Player [4], this module uses the intensity of speech in order to detect emotion and further uses artificial neural, confusion matrix for generation of the playlist.

Smart Music Integrating and Music Mood Recommendation [5], this study basically describes the process through 3 modules, emotion module for capturing the user's details, Music classification module that uses deep learning for recognizing emotion and further Recommendation module for suggesting the playlist.

III. PROPOSED METHOD

The proposed method aims to make the process of creating a playlist more interactive for the user. The operation is based on a series of mechanisms completing their functions in a predetermined order to get the desired outcome. As seen in Fig. 1, the procedure can be described as follows.

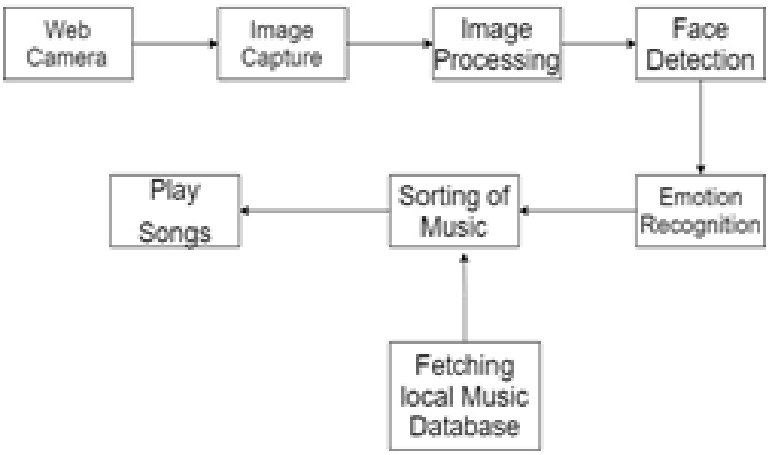

Figure 1. Block Diagram of Emotion detection for audio playback using AI.

- In the proposed system when the emotions.py file executes, Webcam will open and captures the user's photo.
- The system uses the face detection process to scan the photograph for the presence of a face, classifies the input, and generates an emotion (mood) based on the extracted expression.
- All music is retrieved from a local database by the system. After the files have been detected, they are scanned for audio features, which are then extracted.
- Following that, the extracted feature values are classified using the parameters provided.
- Based on the audio feature values to be processed, these parameters include a limited set of genre types.
- Following that, based on the feature extraction process, the songs are divided into different playlists. As a result, lists of similar songs or songs from the same genre are generated.
- The expression is then used as an input to choose a matching playlist from the previously created playlists, and the songs from the playlists are played.

IV. METHODOLOGY

Python is a high-level, general-purpose interpreted programming language. Its design philosophy emphasises code readability by using a lot of indentation. For building Emotion recognition for audio playing using AI, we choose Python as the core platform.

- Viola Jones Algorithm

The Viola Jones object detection framework was proposed by Paul Viola and Michael Jones in 2001 as the first object detection system to offer competitive object detection rates in real-time [8]. The difficulty of face detection was the motivating force behind it, despite the fact that it could be trained to recognise a variety of object types. CvHaarDetectObjects is the OpenCV implementation of this technique. The web camera will provide real-time input to the system. Face detection using the Haar cascade classifier was implemented using OpenCV libraries [7]. The Haar cascade classifier is applied in phases, with features grouped into criteria, to eliminate the undesired parts from future analysis. Face detection algorithm identifies faces and extracts necessary features from them.

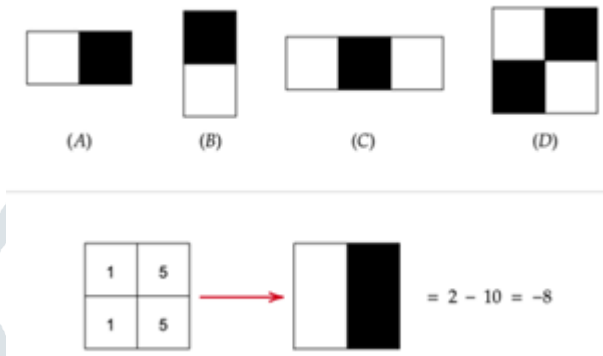

Figure 2. How to calculate Haar-like features.

Dark and light areas make up a Haar-like feature. The sum of the bright region intensities is subtracted from the sum of the dark region intensities, providing a single value. The Viola Jones Object Detection Framework solely uses the Haar-like properties displayed in Fig. 2. We may extract key information from an image, such as edges, straight lines, and diagonal lines, which we can utilize to identify an object, using the numerous Haar-like characteristics.

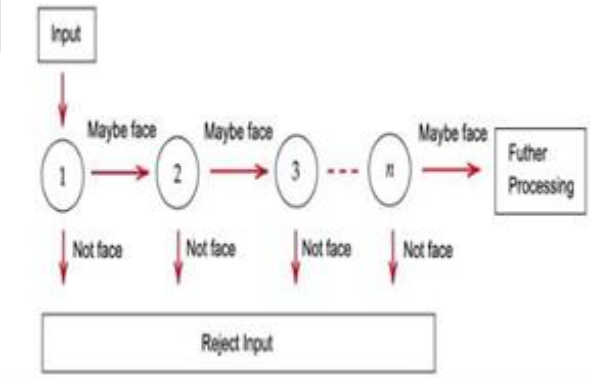

Figure 3. Cascade multi-stage classifier

A Cascade Classifier is a multi-stage classifier that can detect objects quickly and precisely. As a strong classifier passes through the phases, the number of weak classifiers increases. As indicated in Fig. 3, a sequential evaluation is done on an input. If a classifier delivers a negative result for a specific stage, the input is instantly rejected. If the output is affirmative, the input is passed on to the next stage. According to Viola & Jones, this multi-stage technique allows for the generation of simpler classifiers, which can then be used to quickly reject most negative (non-face) input while spending more time on positive (facial) data.

Tensorflow is used for defining and training neural network models. Keras is a strong and easy to use free open-source Python framework for constructing and analysing deep learning models. TensorFlow is a fast numerical calculation toolkit that allows you to design and train neural network models in just a few lines of code. Keras is a deep learning framework for estimating and constructing models.

The image detection window uses the PYTHON-VLC to detect and display the user image with emotion. Tkinter is the de facto method for creating graphical user interfaces (GUIs) in Python, and it comes standard with all Python releases. This Python framework provides an interface to the Tk toolkit by acting as a thin object-oriented overlay on top of it. Tk is a cross-platform set of 'graphical control components,' sometimes known as widgets, which can be used to build application interfaces.

As shown in Fig. 4, Emotion detection for audio playback is described, such that the working revolves around detecting emotion and playing the appropriate song; as we progress and dig deeper with each step, we examine various aspects that must be addressed, such as the dataset that can be used to detect emotion and, once emotion is detected, what type of music should be played to alleviate or maintain the mood.
Finding the proper song for one's current mood can be difficult, especially when one is in a terrible mood and needs to scour the entire playlist for that one song that calms him down. Creating a playlist is also no easy task. It takes time, patience, and a great deal of effort.

Each face is automatically registered such that it is roughly centred and fills the same amount of area. The data includes 48x48 pixel grayscale photographs of faces. The face detected from the webcam will cropped and converted to the grayscale image and processed further to detect emotion. The goal is to categorise each face into one of seven groups based on the emotion it displays (0 = Angry, 1 = Disgust, 2 = Fearful, 3 = Happy, 4 = Neutral, 5 = Sad, 6 = Surprised) [6]. The emotion of user can be identified by comparing input image with trained image dataset.
The emotion-based Audio files will be scanned and features extracted, and a playlist will be created based on the mood. Every emotion has its own subdirectory in the main directory. The songs in each subdirectory correspond to the emotion. The programmer can alter, replace, or delete songs in sub directories based on the user's requirements. In some cases, users may choose different types of tunes depending on their mood.

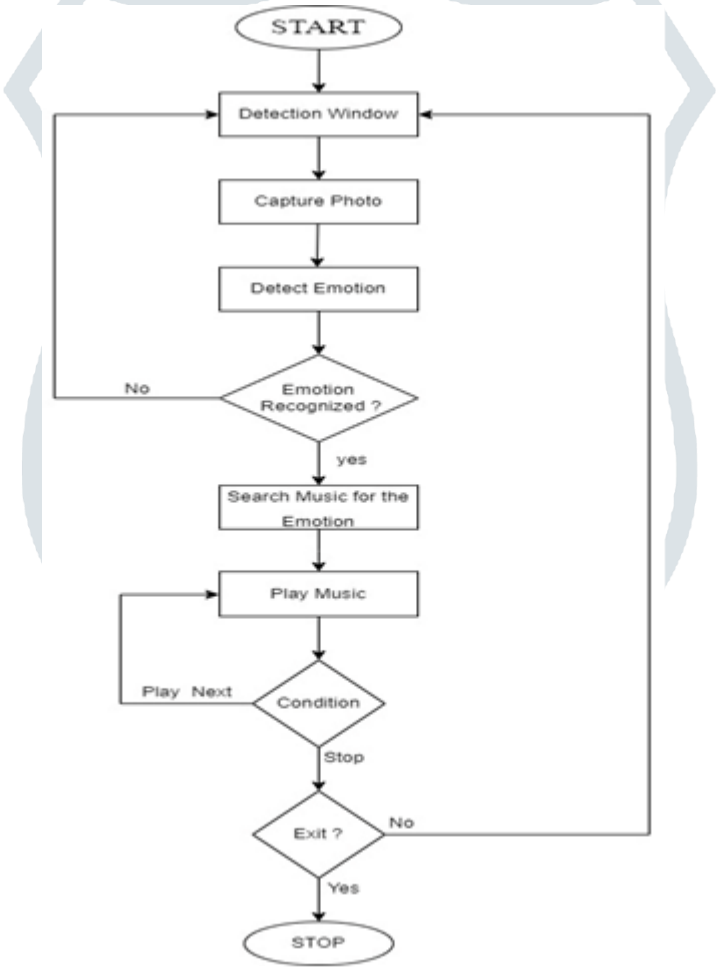

Figure 4. Flowchart of Emotion detection for audio playback.

V. EXPERIMENTAL RESULTS

The accuracy in mood recognition and music selection can be seen as a result of Emotion detection for audio playback using AI. Finding true human emotion with only one measure is difficult. However, it can be identified to some extent using facial expression. Fig. 5 described the result of mood recognition by facial expression under optimal lighting conditions. The accuracy of the model we achieved was 78%. It recognizes emotions in the way we educated it because it is a completely computer-based system. The system takes that mood and creates an accurate music playlist for it. The majority of the songs from the recommended playlist can be played by the system.

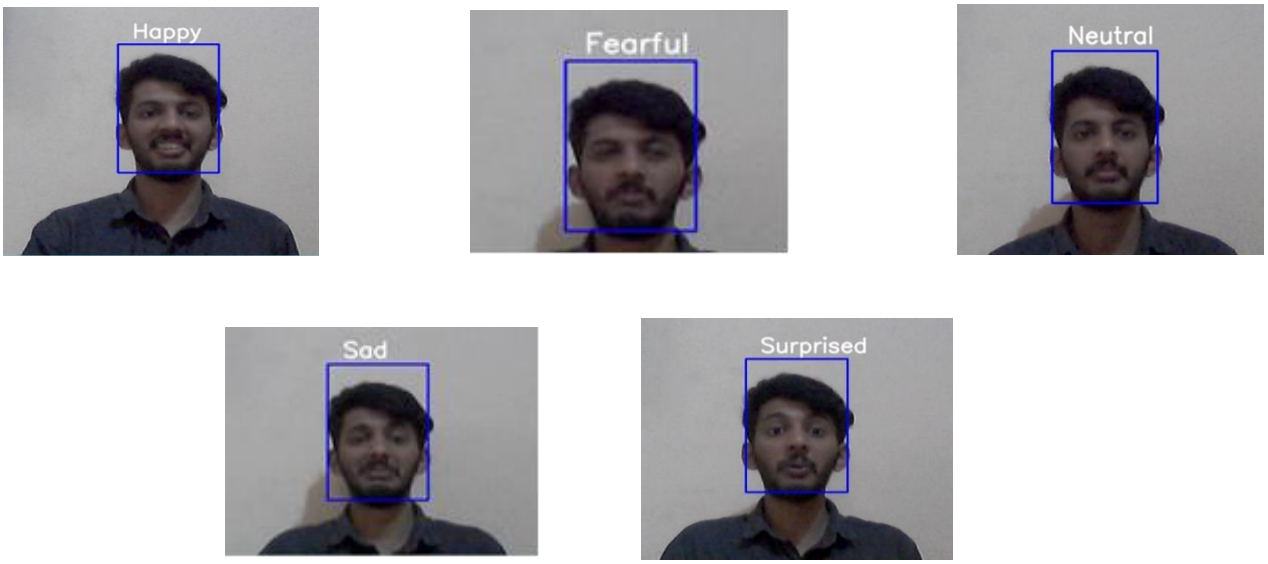

Figure 5. Screenshots of Emotion detection.

Figure 6 shows the result of music control system, the Python Tkinter module was used to create the user interface for our music player. The music player loads the recommended playlist from their modules and imports the finalised mood. And it starts playing songs from that playlist. Python's Tkinter library is used to load these sound objects and manage playing. They are controlled by this module. We added basic functionality to our music player using the Tkinter module, allowing users to pause, resume, play the previous and next song in the playlist, and stop the song on the system.

Figure 6. GUI of Music Control.

VI. CONCLUSION AND FUTURE SCOPE

In this research, we present a system for playing a music based on user's emotions. The system's major goal is to alter or sustain the user's emotional state and match music with certain characteristics. The Emotion-Based Audio Playback is designed to automate and improve the audio playback experience for physically challenged people. The application effortlessly meets the basic requirements of music listeners.

In the future, the proposed technique will aid in reducing the unpredictability of outcomes produced in extremely low light and with low camera resolution. Many additional genres and feelings can also be added. There might be a feature that maps the playlist to internet music directories like Gaana.com, Wynk.com, Jiosaavn.com, Youtube.com and Spotify.com, allowing the work of feeding music into the system to be successful.

VII. ACKNOWLEDGMENT

It gives us great pleasure to recognise the contributions and support of a vast number of people in successful completion of this project. First and foremost, we would want to express our heartfelt gratitude to Dayananda Sagar University, School of Engineering for giving us with a fantastic chance to pursue our Bachelor's Degree here. We like to express our gratitude to Dr. Vaibhav A Meshram, Department of ECE, for his consistent support and competent advise. It gives us great pleasure to convey our heartfelt gratitude to our guide, Mrs. Divya Shree H B, Assistant Professor, Dept. of ECE, for providing excellent academic guidance that enabled us to complete our task and for volunteering his time to assist us at every stage of our project work, paving the way for smooth progress and a successful conclusion. We would like to express our gratitude to our Project Coordinators, Dr. Pushpa P V and Dr. Rajashree Narandra, as well as all of the Electronics and Communication team members, for their assistance in completing the project successfully.

VIII. REFERENCES

[1] "Mood Based Music Player", Anuja Arora, Aastha Kaul, Vatsala Mittal Computer Science Department Jaypee Institute of Information Technology Noida, India, 978-1-5386-9436-7/19/2019 IEEE.

[2] "EMOSIC- An Emotion Based Music Player for Android" Karthik Subramanian Nathan, Manasi Arun, and Megala S Kannan, College of Engineering, Guindy Anna University, India, 978-1-5386-4662-5/17/2017 IEEE.

[3] "Emotion-Based Music Player" Krittrin Chankuptarat, Raphatsak Sriwatanaworachai, Supannada Chotipant, Facultys of Information Technology King Mongkut's Institute of Technology Ladkrabang Bangkok, Thailand, 978-1- 7281-0067-8/19/2019 IEEE.

[4] Hafeez Kabani, Sharik Khan, Omar Khan, Shabana Tadvi "Emotion Based Music Player" International Journal of Engineering Research and General ScienceVolume3, Issue1, January-February, 2015.

[5] Shlok Gikla, Husain Zafar, Chuntan Soni, Kshitija Waghurdekar "SMART MUSIC INTEGRATING AND MUSIC MOOD RECOMMENDATION" 2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSppNET).

[6] https://www.pyimagesearch.com/2018/09/10/keras-tutorial-how-to-get-started-with-keras-deep-learning-and-python/

[7] Haar cascade face detection in OpenCV.

[8] Anagha S.Dhavalikar, and Dr. R. K. Kulkarni, "Face Detection and Facial Expression Recognition System," Interntional Conference on Electronics and Communication System (ICECS - 2014), 2014 IEEE.