International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 05 Issue: 01 | Jan -2021 ISSN: 2582-3930

A ML based Song Recommendation according to Facial Expression and Controlling Music Player Based on Hand Gestures

Ayansh Namdeo[1] , Devyani Rathore[2] , Krishna Raghuvanshi[3] , Monika Barfa[4] ,Nikhil Dubey[5], Prof. Akshita Mishra[6]

[1,2,3,4,5] *Student, Department of Information Technology*

[6]*Assitant Professor, Department of Information Technology*

[1,2,3,4,5,6]*Acropolis Institute of Technology and Research, RGPV, Indore, India*

Abstract —The human face is an essential organ of an individual‘s physique and it particularly performs an necessary position in extraction of an individual‘s conduct and emotional state. Manually segregating the listing of songs and producing an terrific playlist primarily based on an individual‘s emotional aspects is a very tedious, time consuming, labor intensive andupheld task. Various algorithms have been proposed and developed for automating the playlist era process.

Facial expression and hand gesture are structure of nonverbal communication. Computer imaginative and prescient is an interdisciplinary discipline that helps bring a high-level perception of digital photos or movies to computers. In this system, laptop imaginative and prescient factors are used to decide the user’s emotion thru facial expressions and hand gestures. Once the emotion is recognized, the gadget suggests a play-list for that emotion, saving a lot of time for a person over deciding on and enjoying songs manually.

Emotion-Based Music Player additionally maintains music of user’s important points like variety of performs for every song, kinds songs primarily based on class and activity level, and reorganizes the play-list each and every time. The gadget additionally notifies person about the songs that are by no means performed so that they can be deleted or modified. In this system, we uses webcam to detect hand gestures made by the users to perform basic operations such as volume up and volume down.

Keywords—Emotion recognition, Hand gesture detection, Computer vision, Camera, MusicPlayer, Categorization , Recommendations.

1.INTRODUCTION

Music plays a very important role in enhancing an individual‘s life as it is an important medium of entertainment for music lovers and listeners and sometimes even imparts a therapeutic approach. In today‘s world, with ever increasing advancements in the field of multimedia and technology, various music players have been developed with features like fast forward, reverse, variable playback speed (seek & time compression),local playback, streaming playback with multicast streams. Although these features satisfy the user‘s basic requirements, yet the user has to face the task of manually browsing through the playlist of songs and select songs based on his current mood and behavior. The introduction of Emotion recognition, Hand gesture detection and Music player in the traditional music players provided automatically parsing the playlist based on various classes of emotions and moods.

Emotion recognition, Hand gesture detection and Music player is a technique which deals with classifying a received audio signal, by considering its various audio features into various classes of emotions and moods, whereas Music player is a field that extracts some critical information from an audio signal by exploring some audio .Though both and Music player included the capabilities of avoiding manual segregation of songs and generation of playlist, yet it is unable to incorporate fully a human emotion controlled music player. Although, human speech and gesture are a common way of expressing emotions, but facial expression is the most ancient and natural way of expressing feelings, emotions and mood.

Gesture recognition involves a human-computer interaction wherein human gestures are processed and recognized by the machine. It is now possible to control the music based on hand gestures.

The main objective of this paper is to design an efficient and accurate algorithm that would generate a playlist based on current emotional state and behavior of the user and control volume of the music based on hand gestures. The facial expression would categorize into 5 different types of facial expressions likeanger, happy, surprise, sad, and neutral. The hand gesture would categorize into 2 different types like thumbs up and thumbs down. Using the laptop device’s camera, we can capture the user’s facial expression and hand gesture. There are many emotion recognition systems which take captured images as input and determine the emotion. For this application, we are using pretained model MobileNet for recognition of emotion and Convolutional Neural Network (CNN) model for recognition of hand gesture. The system includes a music player that organizes songs based on the user’s emotions and preferences and the volume of music can be controlled by the user’s hand gesture. This player also suggests user’s songs to play based on their emotion.

2.LITERATURE SURVEY

1).Different methods and procedures have been proposed and developed to classify human emotional country of behavior.

The proposed methods have centered solely on the some of the fundamental emotions. For the motive of function recognition, facial elements have been classified into two primary classes such as Appearance-based characteristic extraction and Geometric primarily based function extraction via zheng et. al. Geometric based totally characteristic extraction approach regarded solely the structure or important outstanding factors of some necessary facial facets such as mouth and eyes. In the gadget proposed with the aid of Changbo et. al [3], round a whole of fifty eight foremost landmark factors used to be viewed in crafting an ASM. The look based totally extraction function like texture, have additionally been regarded in unique areas of work and development. An environment friendly technique for coding and imposing extracted facial elements collectively with multi-orientation and multi-resolution set of Gabor filters was once proposed by using Michael Lyons et. [2] al.

2)In correct and environment friendly statistical based totally method for inspecting extracted facial expression facets used to be proposed via Renuka R. Londhe et al.[5] . The paper was once majorly targeted on the find out about of the modifications in curvatures on the face and intensities of corresponding pixels of images. Artificial Neural Networks (ANN) was once used in the classification extracted elements into 6 essential commonplace thoughts like anger, disgust, fear, happy, sad, and surprise. A Scaled Conjugate Gradient back-propagation algorithm in correlation with two-layered feed ahead neural community was once used and was once profitable in acquiring a 92.2 percent focus rate. In order to decrease the human effort and time wished for guide segregation of songs from a playlist, in correlation with special lessons of thoughts and moods, more than a few strategies have been proposed.

3)Thayer[8] proposed a very beneficial 2-dimenesional (Stress v/s energy) mannequin plotted on two axes with thoughts depicted via a 2- dimensional co-ordinate system, mendacity on both two axes or the four quadrants shaped through the 2-dimensional plot. The song temper tags and A-V values from a complete 20 topics have been examined and analyzed in Jung Hyun Kim‘s work, and primarily based on the consequences bought from the analysis, the A-V aircraft used to be labeled into eight regions(clusters), depicting temper with the aid of records mining environment friendly k-means clustering algorithm.

4)Various tactics have been designed to extract facial points and audio facets from an audio sign and very few of the structures designed have the functionality to generate an emotion based totally song playlist the use of human feelings and the present designs of the structures are successful to generate an computerized playlist the use of an extra hardware like Sensors or EEG structures thereby growing the price of the plan proposed. Some of the drawbacks of the current gadget are as follows :

i. Existing structures are very complicated in phrases of time and reminiscence necessities for extracting facial elements in real time.
ii. Based on the modern-day emotional nation and conduct of a user, present structures possess a lesser accuracy in era of a playlist.
iii. Some present structures have a tendency to rent the use of human speech or from time to time even the use of extra hardware for era of an automatic playlist, thereby growing the complete value incurred.

This paper mainly pursuits and focuses on resolving the drawbacks concerned in the present machine via designing an computerized emotion primarily based song participant for the technology of personalized playlist based totally on consumer extracted facial aspects and for this reason heading off the employment of any extra hardware. It additionally consists of a temper randomized and appetizer characteristic that shifts the temper generated playlist to any other identical degree of randomized temper generated playlist after some duration.

5)Hand Gesture cognizance machine affords us an innovative, natural, consumer pleasant way of interplay with the laptop which is greater acquainted to the human beings. Hand gesture has the herbal potential to represents thoughts and movements very easily, consequently the use of these distinctive hand shapes, being recognized by means of gesture awareness gadget and interpreted to generate corresponding event, has the attainable to grant a greater herbal interface to the pc system. By maintaining in thinking the similarities of human hand structure with 4 fingers and one thumb, this paper objectives to existing a actual time device for hand gesture consciousness on the foundation of detection of some significant form primarily based points like orientation, centre of mass (centroid), fame of fingers, thumb in phrases of raised or folded fingers of hand and their respective place in image.

6)Gesture cognizance does no longer reflect onconsideration on any pores and skin colour and/or structure of fingers. Once fed in to the system, the facts is saved as per the given practise and hence can be given the quintessential instructions to act in accordance to the gestures. The algorithm "Weighted Averaging Analysis" can be used in our challenge which offers extra environment friendly result. About 96% of the signs and symptoms are efficiently classified. Such outcomes are so environment friendly that it does now not want any education like in neural networks. The processing is pretty quick given that no state-of-the-art calculus is required whilst jogging the program. The reminiscence requirement is additionally very much less as we are doing statistical evaluation and now not the database matching like in ordinary picture processing.

3.METHODOLOGY

The proposed machine entails an emotion track suggestion gadget that affords the technology of a playlist in accordance to the user's emotional nation and manage the extent of song based totally on hand gesture.

The working can be referred to as follows:

1) The proposed machine works through first presenting an easy adequate interface which prompts the consumer to create an account or login.
2) In the subsequent step, the consumer digital camera is invoked with ideal permissions and a actual time graphical enter (image) is supplied to the system.
3) The gadget first exams for the presence of a face in theenter the use of the face detection process, then classifies the enter and generates an output which is an Emotion (mood) primarily based on the expression extracted from the actual time graphical input.
4) After this the categorized expression acts as an enter and is used to choose a terrific playlist from the in the beginning generated playlists and the songs from the playlists are delivered in the queue.
5) In the subsequent step, person can play songs from the queue.
6) There is a webcam button on the interface, by means of clicking on it the person digital camera is invoked with appropriate permissions and an actual time graphical enter (image) is furnished to the system.
7) After this the labeled gesture acts as an enter and is used to control the volume of the music.

A.Facial Expression Recognition:

In this undertaking we are offering the actual time facial expression consciousness of 5 most simple human expressions: **ANGER, HAPPY, NEUTRAL, SAD, SURPRISE.**This mannequin can be used for prediction of expressions in actual time video. The gadget robotically detects the face the usage of HAAR cascade then its vegetation it and resize the photograph to a unique dimension and supply it to the mannequin for prediction. The mannequin will generate 5 chance values corresponding to 5 expressions. The best chance cost to the corresponding expression will be the anticipated expression for that image. The dataset is taken from kaggle dataset. For facial expression cognizance we are the usage of pertainedMobilenet model.

Some of the examples of facial expression images are:

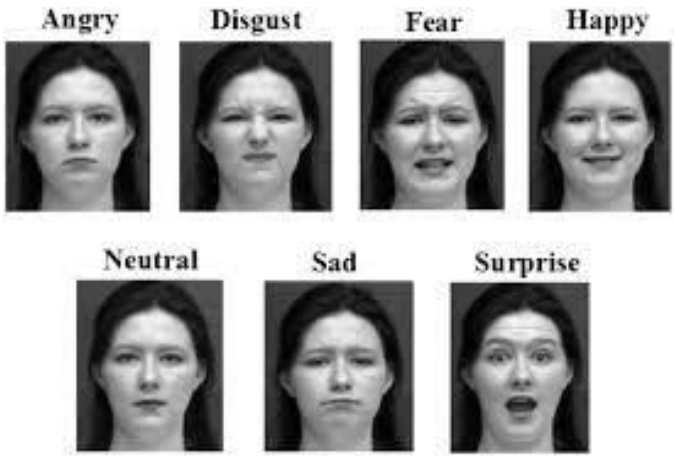

Fig-1: Facial Expression

Table -1 Estimated Accuracy for different categories of emotion

Emotions	Accuracy
Sad	0.0653416
Happy	0.8810852
Surprise	0.8119784
Neutral	0.08923215
Angry	0.6773573

B. Hand Gesture Recognition

In this system, we are providing the actual time hand gesture attention of two human handgestures: Thumbsup and Thumsdown. This mannequin can be used for prediction of expressions in actual time video. The machine robotically captures the picture of hand gesture then its vegetation it and resize the photograph to a particular measurement and supply it to the mannequin for prediction. The mannequin will predict the photograph and supply favored output. The pics for dataset are taken from google. For hand gesture focus we are the usage of CNN model. This system can increase or decrease the volume of the music based on the user's hand gesture.

Some of the examples of hand gesture are:

Fig -2:Hand Gestures

Table -2 Estimated Accuracy for different categories of hand Gesture

Hand Gestures	Accuracy
Thumbs up	0.721135
Thumbs down	0.743421

Thereby lowering the pointless computational time and thereby growing the standard accuracy and effectivity of the system. The device will no longer solely decrease bodily stress however will additionally act as a boon for the tune remedy structures and might also additionally aid the tune therapist to therapize a patient. Also with its extra elements stated above, it will be a whole gadget for track fans and listeners.

The future scope in the application would to format a mechanism that would be beneficial in song remedy cure and furnish the track therapist the assist wanted to deal with the sufferers struggling from problems like intellectual stress, anxiety, acute despair and trauma. The proposed gadget additionally tends to keep away from in future the unpredictable outcomes produced in severe awful mild prerequisites and very terrible digital camera resolution.

4. SNAPSHOTS OF MODULES

A.Login Module

In this user have to login for accessing the application after that grant permission to access camera to capture image.

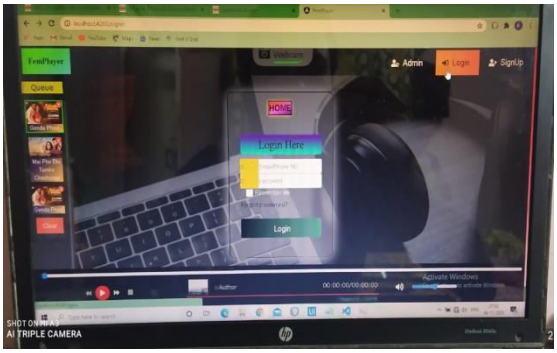

Fig-3: Login Module

B.Face Recognition Module

In this after capturing the image the application proceesed it ang generate the song list suitable for his mood.

Fig-4: Face Recognition Module(Happy)

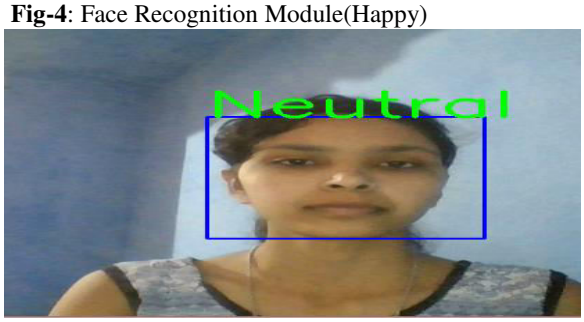

Fig-4.1: Face Recognition Module (Neutral)

C.Music Player Module

In this generated song list comes in queue so user can select a song his choice and playing in player.

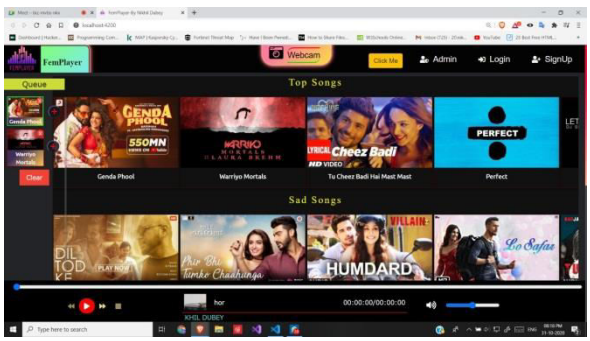

Fig .5: Music Player Module

In this we also add hand gestures detection to change volume like thumbs up- volume up and thumbs down - volume up.

ACKNOWLEDGMENT

It gives us great pleasure in presenting this project report titled:"A ML based Song Recommendation According to Facial Expression and Controlling Music Player based on Hand Gestures". On this momentous occasion, we wish to express our immense gratitude to the range of people who provided invaluable support in the completion of this project. Their guidance and encouragement has helped in making this project a great success.

We express our gratitude to our project guide Prof Akshita Mishra, who provided us with all the guidance and encouragement and making the lab available to us at any

time. We also would like to deeply express our sincere gratitude to Project coordinators.

We are eager and glad to express our gratitude to the Head of the Information Technology Dept. Dr. Kamal K. Sethi, for his approval of this project. We are also thankful to him for providing us the needed assistance, detailed suggestions and also encouragement to do the project.

REFERENCES

1. Samuel Strupp, Norbert Schmitz, and KarstenBerns, ―Visual-Based Emotion Detection for Natural Man-Machine Interaction‖.

2. Michael lyon and Shigeru Akamatsu, ―Coding Facial expression with Gabor wavelets., IEEE conf. on Automatic face and gesture recognition, March 2000

3. Kuan-Chieh Huang, Yau-Hwang Kuo, Mong-Fong Horng, Changbo,Emotion Recognition by a novel triangular facial feature extraction method‖.

4. S. Dornbush, K. Fisher, K. McKay, A. Prikhodko and Z. Segall ,Xpod- A Human Activity and Emotion Aware Mobile Music Player‖, UMBC Ebiquity, November 2005.

5. Renuka R. Londhe, Dr. Vrushshen P. Pawar, ―Analysis of Facial Expression and Recognition Based On Statistical Approach‖, International Journal of Soft Computing and Engineering (IJSCE) Volume-2, May 2012.

6.Russell, ―A circumplex model of affect‖, Journal of Personality and Social Psychology Vol-39(6), pp1161–1178 , 1980.

7. Sanghoon Jun, Seungmin Rho, Byeong-jun Han and Eenjun Hwang, ―A fuzzy inference-based music emotion recognition system‖,VIE,2008.

8. Thayer ‖ The biopsychology of mood & arousal‖, Oxford University Press ,1989

9. Z. Zeng ―A survey of affect recognition methods: Audio, visual, and spontaneous expressions,‖IEEE. Transaction Pattern Analysis, vol 31, January 2009

10. www.ijergs.org

11.www.wikipedia.com

12.www.ijsrd.com