Mozart

1st Bhagyashree Avhadkar
Computer Science Engineering
G.H Raisoni Institute of Engineering & Technology
Nagpur, India

2nd Parnavi Mandaokar
Computer Science Engineering
G.H Raisoni Institute of Engineering & Technology
Nagpur, India

3rd Shyam Zod
Computer Science Engineering
G.H Raisoni Institute of Engineering & Technology
Nagpur, India

4th Antara Bhattacharya
Computer Science Engineering
Faculty G.H Raisoni Institute of Engineering & Technology
Nagpur, India

Abstract— *Individuals will in general tune in to music dependent on their temperament and interests. This undertaking centers on making a framework to recommend tunes for client dependent on their temperament by catching outward appearances. Outward appearance is a type of nonverbal correspondence. Computer vision is an interdisciplinary field that passes on an undeniable level comprehension of advanced pictures or recordings to computers. In this framework, Computer vision parts are utilized to decide the user's feeling through outward appearances. When the feeling is perceived, the framework recommends a playlist for that feeling, saving a great deal of time for a client over choosing and playing tunes physically. Mozart additionally monitors user's subtleties like number of plays for every melody, sort tunes dependent on class and interest level, and revamps the play-list without fail.*

Keywords—SVM, Emotion Recognition, Machine Learning

I. INTRODUCTION

Music is a fundamental part of our day by day life. We tune in to tunes according to our state of mind. Music is one of the media of amusement and even confers a restorative methodology. It is imperative to play a suitable tune on the specific passionate state. Existing music player fulfills the user's fundamental prerequisites, yet the client needs to confront the assignment of physically perusing the playlist of melodies and select tunes dependent on his present state of mind and conduct. Mozart depends on the guideline of identification of human feelings to play proper melodies for current enthusiastic state. The current enthusiastic condition of individual can be effortlessly seen through their outward appearances. It tends to be accomplished with assistance of picture preparing and AI strategies. AI can be utilized for arranging music into set of specific feelings. Once, this information is available, client can be learned about his/her inclinations and propensities for tuning in by time, temperament, action, and so on Preparing on this information can produce better playlist for future tuning in. Objective behind this work is to allow every day elements to improve music proposal.

II. RELATED WORK

Different strategies and approaches have been proposed and created to characterize human passionate condition of conduct. The proposed approaches have zeroed in just on the portion of the essential feelings. With the end goal of highlight acknowledgment, facial highlights have been ordered into two significant classes, for example, Appearance-based component extraction and Mathematical based element extraction by zheng et. al [9].
A precise and effective measurable based methodology for examining extricated outward appearance highlights was proposed by Renuka R. Londhe et al. [7]. The paper was significantly centered around the investigation of the progressions in bends on the face and forces of relating pixels of pictures. Counterfeit Neural Organizations (ANN) was utilized in the arrangement extricated highlights into 6 significant all inclusive feelings like displeasure, appall, dread, glad, miserable, and shock. A Downsized Form Slope spread calculation in relationship with two-layered feed forward neural organization was utilized and was effective in getting a 92.2 % acknowledgment rate. To decrease the human exertion and time required for manual isolation of melodies from a playlist, in connection with various classes of feelings and temperaments, different methodologies have been proposed.

The point of this paper [4], is to create music framework which thinks about human feelings into account. The passionate state can be deciphered from outward appearances through the webcam. We have used the CNN classifier to construct a neural organization model. This model is prepared and exposed to recognize temperament from outward appearances utilizing OpenCV. Melodies having a place with specific opinions are ordered based on rhythm highlight in beats each moment. A framework creates music playlist dependent on that identified mind-set. Our music player will play that produced music playlist to improve client's state of mind [4].

In paper [2], they have utilized Microsoft's psychological services and Google's own movement acknowledgment

Programming interface in our methodology for quicker execution. AI can be utilized for arranging music into set of specific feelings. Once, this information is available, client can be learned about his/her inclinations and propensities for tuning in by time, temperament, movement, and so on Preparing on this information can create better playlist for future tuning in. Objective behind this work is to allow every day components to improve music suggestion.

In paper [6], the proposed framework dependent on outward appearance separated will create a playlist naturally consequently decreasing the exertion and time engaged with delivering the interaction physically. In this way the proposed framework will in general diminish the computational time engaged with getting the outcomes and the general expense of the planned framework, accordingly expanding the general exactness of the framework by utilizing Counterfeit Neural Organizations, Disarray Lattice, Viola and Jones Face Discovery.

Various methodologies have been intended to extricate facial highlights and sound highlights from a sound sign and not many of the frameworks planned have the ability to create a feeling based music playlist utilizing human feelings and the current plans of the frameworks are skilled to produce a robotized playlist utilizing an extra equipment like Sensors or EEG frameworks accordingly expanding the expense of the plan proposed. A portion of the downsides of the current framework are as per the following:

I. Existing frameworks are exceptionally mind boggling as far as time and memory necessities for separating facial highlights progressively.
ii. In view of the current enthusiastic state and conduct of a client, existing frameworks have a lesser exactness in age of a playlist.
iii. Some current frameworks will in general utilize the utilization of human discourse or now and again even the utilization of extra equipment for age of a robotized playlist, consequently expanding the all out cost caused.
iv. It requires the client to physically choose the melodies.
v. Haphazardly played tunes may not match to the state of mind of the client.
vi. Client needs to order the tunes into different feelings and afterward for playing the melodies client needs to physically choose a specific feeling.

III . METHODOLOGY USED

Mozart is computer-based software that focuses on implementing mood detection. It is a prototype of a new product that merges some separate interfaces: face detection, facial expression recognition, Playlist generation, play music.

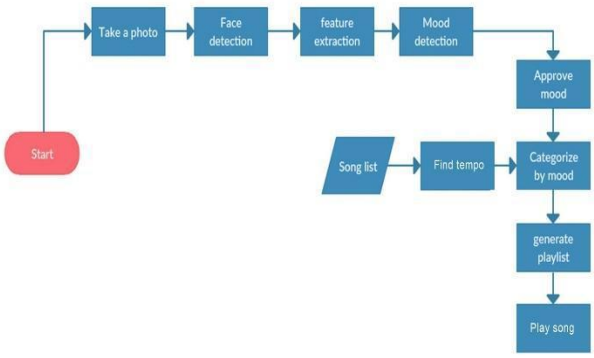


Fig 1: Flow Diagram of Mozart [4]

Mozart will first take the photo of the person, then after the image is being captured the face is detected, then the expression is being extracted from the captured image, the expression is being detected and according to the extracted emotion the song is been played from the system playlist. Also, in Mozart there is a provision of playlist generation according to the previously played songs.

Algorithm of Mozart:
begin
1. Fetch the dataset.
2. Remove duplicate entries if present.
3.If same set of independent variable has multiple dependent class then categorize them in one class.
4. Divide the dataset into test set and training set
5. Apply Random forest classifier to train the dataset.
6. Take the inputs from data frame
7. Predict the result as song by giving input to 5.
end[2]

Hardware Requirements & Software Requirements

The hardware requirements required for this project are: Intel i3,4GB RAM ,Webcam , Speaker. The software requirements that are required for this project are: Python 3.6.5 , Open CV 3.4.3, FisherFace, EEL 0.9.10.
The system has two modules and they are:

Emotion Extraction Module
a). Face Detection :
Mozart will take real time input from web camera. We have used OpenCV libraries to implement face detection through Haar cascade classifier [5]. Haar cascade classifier implemented in stages by application of features grouped into criteria such that it will discard the unwanted part from further analysis. Face detection algorithm detects face and further it will extract required features from it.
b). Loading and saving trained model:
Facial expression detection in Fisherface works with the help of trained models. Reason behind this is to allow user to take dataset according to their use. Suppose if we take a huge amount of dataset of around 25-30k it will give nice accuracy no doubt but if the situation is like that the user

of the devices are a few people. Now in such condition if we take some precise dataset with around 400-450 images as input related to the user then it will also give good accuracy with the benefit of less amount of dataset and less storage on memory to operate. As well as small memory of data give output fast which result in quick response time. Here we first tried with Cohn-Kanade dataset then we made some classification in the as our need make it to train our model.

c) Haarcascade Model:

Haarcascade model is precise face detection trained model which is provided by Open-cv. It return the co-ordinates in terms of (x, y) at (left, bottom) of face frame and it's width and height from those co-ordinates. As here in the .detectMultiScale() method it is capable of detect multiple faces and it return an array of all the faces(co-ordinates) as an element. The arguments have set according to the threshold what we need for our checking purpose. We have set it such like it doesn't affect our model accuracy.

Audio Feature Extraction Module

a). Detected Emotions :We have implemented the linking of python with javascript through eel library. Which provide us the privilege to access python methods from js as well as vice versa. Here the striating flow will be in python code as the library is implemented in python then it transfer the control to html, JS. And according to the result we show emoticons.Sad, happy,angry,neutral. According to which we can classify emotion directory for playing song we have chosen this 4 emotions.

b). Methods for playing songs: In JavaScript file we have implemented too much methods for the switching of song.

- Queue
- Based on Emotion
- Random

In the first one as queue works it has been implemented. In second one we call python code to get emotion from user's facial expression and according to that chosen next song which is also randomly and played it. In third one we directly used random function and all the methods are dynamic it can handle as change in number of songs accordingly.

c).HTML, CSS and JS concepts for online music player:

As we know the CSS give a great look to communicate and through JS we can interact with user and not look like complicated program run at console and it also give user privilege to choose any song to play.

IV. EXPEIMENTAL RESULTS

Testing and implementation is complete, and surface expressions are drawn for the user's independent and dependent data sets. We selected a dataset containing 25 face images for user-independent experiments and 10 computer packages for user-dependent experiments. We experimented with static and dynamic datasets using 350X350 images. We trained a dataset consisting of about 1000 items in the training set and 100 items in the test set. Weka tools are used to determine the accuracy of various machine learning algorithms used to train models. The accuracy of the random forest algorithm can be improved by increasing the number of trees in the forest.

Parameters	J48 Algorithm	Random Forest
Correctly Classified instances	50.86%	92.33%
Incorrectly Classified instances	49.14%	7.67%

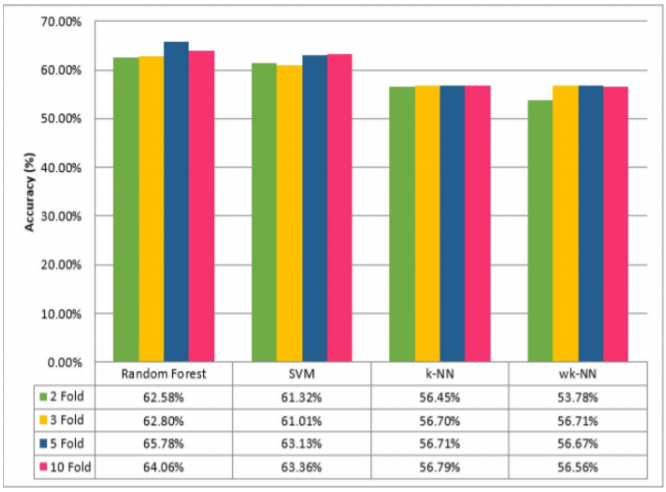
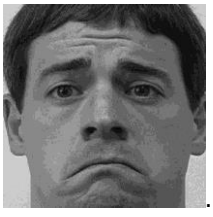


Fig 2: Comparison of Random Forest Algorithm with SVM, k-NN, wk-NN[10]

The effectiveness of Mozart can be seen in the precision of mood detection and song recommendations. Human exact emotions are difficult to find only in one dimension, but can be judged to some extent by facial expressions. The model we used reached 80% accuracy. Since this is a fully computerized system, we understand emotions the way we trained them. The system takes this mood into account and creates exactly one music playlist that fits that mood. The system can play most songs in the recommended playlist.

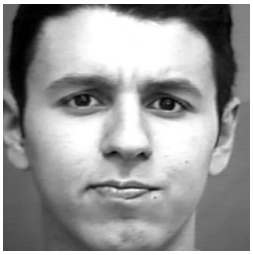
Images used:



Sad

Neutral

Happy



Angry

Mozart has songs about a variety of emotions and has proven its correctness by suggesting songs about specific emotions.

Estimation of accuracy for various categories of sound features:

Emotion	Accuracy
Happy	92.25%
Sad	89%
Angry	94%
Neutral	100%

V. CONCLUSION

Mozart will benefit consumers looking for music based on mood and emotional behavior. This improves the overall accuracy and efficiency of the system by reducing the time it takes to find music, reducing unnecessary computing time. This system not only reduces physical stress, but can also increase music therapy systems and help music therapists prescribe treatment to patients. With the additional features mentioned above, it will be a perfect system for music lovers and listeners. Mozart is used to automate and improve the quality of music players for end users. The program meets the basic requirements of traditional applications as well as undisturbed music listeners. We use technology to improve the user experience of the system in a number of ways. It simplifies the end-user experience by taking pictures with the camera, grasping emotions, and providing personal playlists through a more advanced interactive system. Users can also see unplayed songs to free up storage space.

VI. FUTURE WORK

Mozart is a new way to customize song playlists with machine learning technology. Our solution works and gives you the best playlists. In the future, the system should develop useful mechanisms for music therapy and provide music therapists with the support they need to treat patients with disorders such as mental stress, anxiety, acute depression and trauma. The proposed system also tends to avoid unpredictable future consequences that occur in very poor lighting conditions and very low camera resolutions. The system offers a variety of features such as voice assistant, voice search, lyrics assistance, and history, favorite songs based on previous searches, self-made playlists, and dark/light mode settings.

VII. REFERENCES

[1]. Sarvesh Pal, Ankit Mishra, Hridaypratap Mourya and Supriya Dicholkar "**EMO-MUSIC (Emotion based Music player)**" ; EasyChair ; **January 26, 2020**

[2]. Pranav Sarda, Sushmita Halasawade, Anuja Padmawar, Jagannath Aghav "**Emousic: Emotion and Activity-Based Music Player Using Machine Learning**" ; DOI: 10.1007/978-981-13-6861-5_16; In book: Advances in Computer Communication and Computational Sciences (pp.179-188); Project: Machine Learning ; **May 2019**

[3]. Ahmed Hamdy AlDeeb, Ghada Hassan , "**Emotion-Based Music Player Emotion Detection from Live Camera**";DOI: 10.13140/RG.2.2.25443.40482 ;Thesis for: Computer Science Bachelor Degree; Advisor: Dr. Ghada Hassan; Project: Emotion-Based Music Player ; **June 2019**

[4]. T D Kusumaningrum1, A Faqih1 and B Kusumoputro1; "**Emotion Recognition Based on DEAP Database using EEG Time-Frequency Features and Machine Learning Methods**"; International Conference on Science & Technology**(ICoST 2019)**

[5]. A. S. Mali1, A. A. Kenjale2, P.M. Ghatge3, A.G. Deshpande4 1 Dept. of CSE, Rajarambapu Institute of Technology, Shivaji University, Sakharale, India 2, 3, 4 Computer Science and Engineering, Rajarambapu Institute of Technology, Islampur, India; "**Mood based Music System**"; Institute of Electrical and Electronics Engineers(IEEE) in Computer Science and Engineering Vol.6, Issue.3, pp.27-30, **June (2018)**

[6]. Sri Charan Nimmagadda ; "**Emotion Based Music Player**"; Submitted to the Faculty of the Department of Computing Sciences Texas A&M University - Corpus Christi Corpus Christi, Texas ; **summer 2017**

[7]. Hafeez Kabani1, Sharik Khan2, Omar Khan3, Shabana Tadvi4 M.H Saboo Siddik College of Engineering, University of Mumbai, India "**Emotion Based Music Player**" International Journal of Engineering Research and General Science Volume 3, Issue 1, **January- February, 2015** ISSN 2091-2730

[8]. 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**.

[9]. Byeong-jun Han, Seungmin Rho, Roger B. Dannenberg and Eenjun Hwang "**SMERS: music emotion recognition using support vector regression**", 10thISMIR , **2009**.

[10]. Z. Zeng;"**A survey of affect recognition methods: Audio, visual, and spontaneous expressions",**IEEE. Transaction Pattern Analysis, vol 31, **January 2009**.