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 based Music Recommendation System

[1] Shravani Ashok More, [2]Payal Balkrishna Shete, [3]Dnyanraj Jaywant Chavan, [4]Tejas Tanaji Patil, [5]Miss. Jyoti N. Gurav

[1]Student, [2]Student, [3]Student,[4] Student,[5]Guide
[1]Department of Computer Engineering,
[1]D. Y. Patil Technical Campus, Talsande, Kolhapur, India

Abstract: 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 and including volume modulation, genre classification etc. The motivation of this work comes from the possibility of reducing the human effort in creating music playlists manually, thus generating them automatically based on the user's emotional state. The human face plays an important role in knowing an individual's mood. The required input is extracted from the human face directly using a camera. One of the applications of this input can be for extracting the information to deduce the mood of an individual. This data can then be used to get a list of songs that comply with the "mood" derived from the input provided earlier.

IndexTerms **– Face Emotion Detection, Emotion Classification.**

I. INTRODUCTION

People frequently use their facial expressions to convey their feelings. It has long been recognized that music may change a person's disposition. A user's mind can be gradually calmed down and an overall nice effect can be produced by capturing and recognizing the emotion being uttered by the person and playing appropriate tunes matching the one's mood. The goal of the project is to record a person's facial expressions as they exhibit emotion. The web camera interface for computer systems is used by a music player to record human emotion. The software takes a user's image and uses image segmentation and image processingtechniques to extract information from a target person's face in an effort to determine what emotion they are attempting to convey. The idea attempts to uplift the user's mood by playingmusic that fits their needs while simultaneously taking their photograph. Facial expression recognition has been the most effective method of expression analysis known to humanity since ancient times. Facial expressions are the best way that people may deduce or evaluate the emotion, sentiment, or thoughts that another person is attempting to express.

II. Motivation

Our motivation in this work is to use emotion recognition techniques with wearable computing devices to generate additional inputs for music recommended system's algorithm, and to enhance the accuracy of the resulting music recommendations. Emotion recognition from only GSR signals. In this study we are enriching signals with PPG and propose a data fusion based emotion recognition method for music recommendation engines The proposed wearable attached music recommendation framework utilizes not only the user's demographics but also his/her emotion state at the time of recommendation. Using GSR and PPG signals we have obtained promising results for emotion prediction.

III. LITERATURE REVIEW

[1] S. Metilda Florence and M. Uma proposed the music recommendation system The main goal of their music recommendation system is to provide users with suggestions based on their tastes. Analyzing a user's facial expressions / emotions can lead to an understanding of the user's current emotional or mental state. It is well known that people make use of facial expressions to express what they want to say and the meaning of words more clearly. Developing a recommender system helps users decide what type of music they should listen to, and helps them reduce their stress levels. Users don't have to waste time searching for songs or looking up songs based on their mood, as thebest track would be shown to them according to their requirements. The captured images of the user will help in identifying the songs/playlist. The system is still not able to record all the emotional states correctly due to the lack of images inthe image dataset. The image that is fed into the classifiershould be taken in a well-lit atmospherefor the classifier to give accurate results.[2020]

[3]. . Immanuel James, J. James Anto Arnold, J. Maria Masilla Ruban Sara In this paper, the focus is on detecting human

emotions for developing emotion-based music players, which approaches have been developed in the available music players to detect emotions, which method their music player follows to detect human emotions, and how it is better to use their system for emotion detection. It also briefly describes playlist generation, and emotion classification. By using a pycharm analysis tool, they have developed a software that understands user emotion based on facial expression. Theyhave integrated Python code into the web service, allowing them to play music based on facial expressions. [2019]Shiha, Mohammed, and Serkan Ayvaz .The face is an important aspect in assessing human emotions and moods. Emotions are extracted with the help of a camera. The face is given as an inputto the process of recognizing facial emotions, and the music is played automatically based on the emotions. This system develops a prototype of recommendation, a dynamic music recommendationsystem based on human emotions. Based on all human listening patterns, songs are trained for all emotions. Feature extraction and machine learning techniques have been integrated. Once the mood is derived from the input image, the user will play a song that suits the particular mood. This system provides a high level of accuracy on real facial images. The Pygame package is used to interpret the sound libraries in the python programming language.[2017]

[4]. Diah Anggraeni Pitalokaa, Ajeng Wulandaria implemented the proposed method and compared the pre-processing methods for facial expression recognition. Based on the experimental results obtained, face detection and cropping to capture the regions of interest has been declared to achieve the best improvement in CNN performance. The global contrast normalization step contributes moreaccuracy than other normalization techniques to accuracy but not as good as getting the ROI. Graph Convolutional Network (GCN) tries to reduce the distribution of data so that different contrast values are not present. The proposed CNN model works better at 32x32 and 64x64 resolutions. It seems the capacity of the model satisfies the complexity task for facial expression recognition on those resolutions. Performance of CNN can be boosted using data augmentation like combining data from step cropping and adding noises.[2017]

[5]. Xiao-Wei Wang, Dan Nie, and Bao-Liang Lu. In it, he suggested manually separating playlistsand annotating songs according to the user's current emotional state. This is laborious and time consuming. Numerous algorithms have been proposed to automate this process.However, becauseexisting algorithms were slow, the use of additional hardware (such as EEG systems and sensors) increased the cost of the entire system but had less accuracy. This paper presents an algorithm that automates the process of generating audio playlists, based on the user's facial expressions, saving the time and effort spent manually running the process. The algorithm proposed in this paper aspiresto reduce the total computational time and the cost of the designed system. It also aims to improve the accuracy of the designed system. The facial expression recognition module of the proposed algorithm is validated by testing the system against user-dependent and user-independent datasets.[2014]

[6]. Fang-Fei Kuo et al and Suh-Yin Lee et al.With the growth of digital music, the development of music recommendations is helpful for users. The existing recommendation approaches are based onthe user's preference for music. However, sometimes, recommending music according to the emotion is needed. In this paper, they have proposed emotion-based music recommendation, which was based on the association discovery from film music. They investigated the music feature extraction and modified the affinity graph for association discovery between emotions and music features. Experimental results show that the proposed approach achieves an accuracy of 85% on average.[2005]

IV. PROBLEM STATEMENT

Real time dataset allows us to capture the person's image at the particular instant based on which songs can be suggested which complies with his mood.

V. PROPOSED SYSTEM ARCHITECTURE

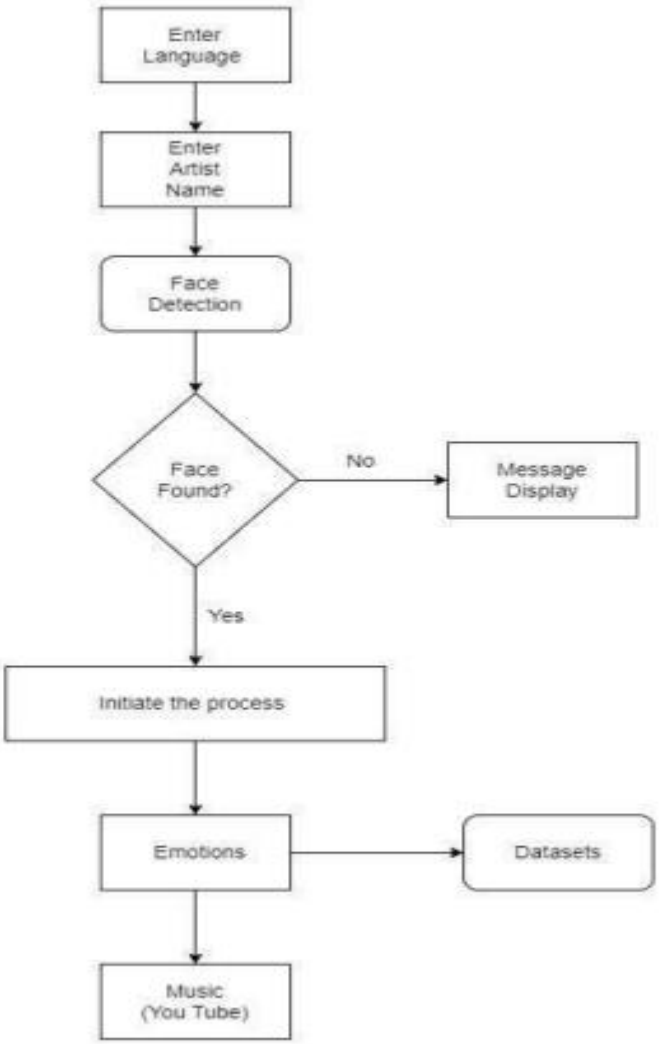

Fig: Proposed System Architecture

VI. METHODOLOGY

Convolutional Neural Network is a Deep Learning algorithm which can take in an input image, assign importance to various aspects/objects in the image and be able to differentiate one from the other. CNNs have an input layer, an output layer, and hidden layers. The hidden layers usually consist of convolutional layers, ReLU layers, pooling layers, and fully connected layers. In a convolutional layer, neurons only receive input from a subarea of the previous layer. In a fully connected layer, each neuron receives input from every element of the previous layer. Convolutionallayers apply a convolution operation to the input. This passes the information on to the next layer. Pooling combines the outputs of clusters of neurons into a single neuron in the next layer. Fully connected layers connect every neuron in one layer to every neuron in the next layer.

In this system requires mainly three modules i.e.,

Module 1:Face Detection
Module 2:Mood detection
Module 3:Song Recommendation

Module 1: Face Detection Module

Real time datasets are used for emotion classification by using a web camera or mobile camera. Images are captured after getting permission from the user. The captured images are compared withFER 2013 datasets for emotion classification. FER 2013 datasets consist 37887 grey scale images with 7 different emotions where 0 is Disgust, 1 is Fear, 2 is Happy, 3 is Sad, 4 is Surprise and 5 is Neutral.

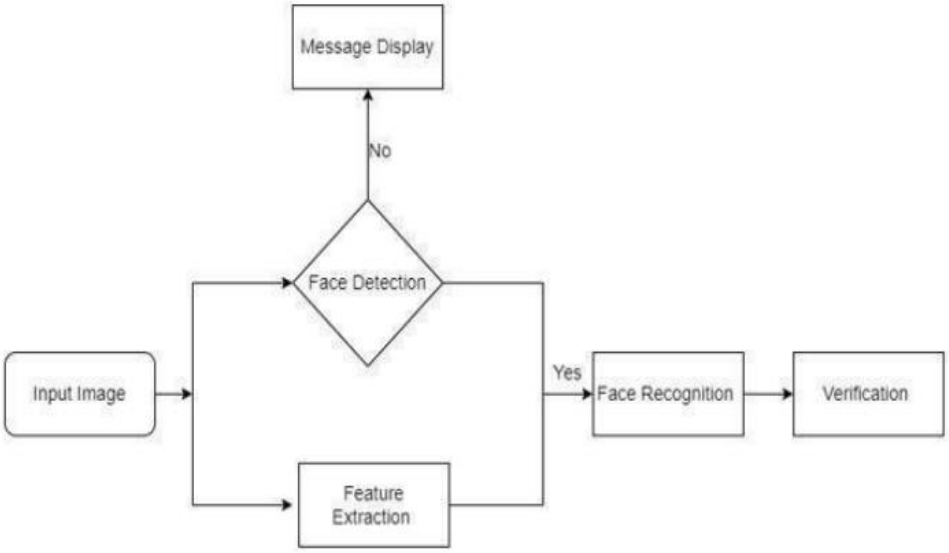

Fig: Face Detection Module

Module 2: Mood Detection Module

The user's image is captured using the camera / webcam. Once the image is captured, the capturedimage frame is converted from the webcam feed to a grayscale image, improving the performance of the classifier. It is used to identify the face present in the image. When the conversion is complete, the image is sent to the classification algorithm. The classification algorithm can use feature extraction technology to extract faces from the frames in webcam feeds. Individual features are retrieved from the extracted face and sent to the trained network torecognize the emotions expressedby the user. These images are used to train the classifier. This allows us to extract the location of facial landmarks from those images based on the knowledge we had already acquired from the training set when the classifier is presented with an entirely new and unknown image set. It returnsthe coordinates of the newly recognized face landmarks. The network is trained using the dataset. In this way, the emotions expressed by the user are identified.

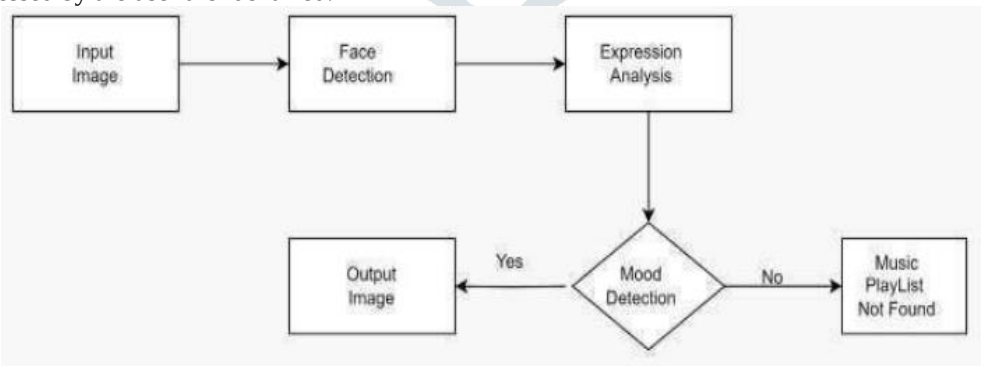

Fig: Mood Detection Module

Module 3: Song Recommendation Module

After the emotion of the user is extracted, the music/audio clip based on the voiced emotion is displayed to the user i.e. a list of songs based on the emotion is displayed. Based on the regularity that the user would listen to the songs are displayed in that order. For example, if the emotion or facial feature is categorized under happy, then songs from the happy database will be displayed to the user. The dataset obtained is made to undergo preprocessing to make it suitable for model Training. After preprocessing, the dataset is split into 70% training and 30% test dataset. Different machine learning algorithms such as Convolutional Neural Network (CNN), Convolutional LSTM (Long Short Term Memory) Network (ConvLSTM) are used to determine which algorithm shows the best accuracy.

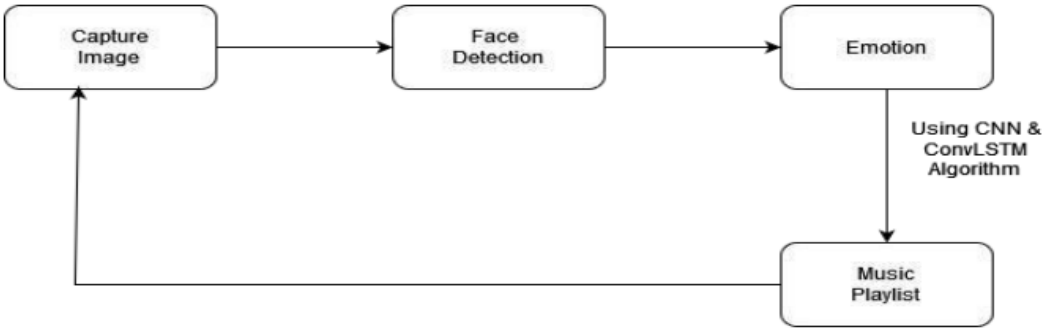

Fig: Song Recommendation Module

VII. REQUIREMENTS

- HARDWARE REQUIREMENTS
  - Processor: Intel I3 processor
  - Storage Space: 500 GB. Screen size: 15” LED
  - Devices Required: Web camera, Mouse and a Keyboard
  - Minimum Ram: 4GB and a good Internet connection
- SOFTWARE REQUIREMENTS
  - OS: Windows 7 and above /UBUNTU
  - Programming Language: Python
  - Software: JetBrains PyCharm Community Edition 2017.1.4 x64
  - Additional requirements: TensorFlow

VIII. REFERENCES

[1] Florence, S. Metilda, and M. Uma. "Emotional Detection and Music Recommendation System based on User Facial Expression." IOP Conference Series: Materials Science and Engineering. Vol.912. No. 6. IOP Publishing, 2020.

[2] James, H. Immanuel, et al. "EMOTION BASED MUSIC RECOMMENDATION SYSTEM." EMOTION 6.03(2019)

[3] Shiha, Mohammed, and Serkan Ayvaz. "The effects of emoji in sentiment analysis." Int. J. Comput. Electr. Eng.(IJCEE.) 9.1 (2017): 360-369.

[4] Pitaloka, Diah Anggraeni, et al. "Enhancing CNN with preprocessing stage in automatic emotion recognition." Procedia computer science 116 (2017): 523-529

[5] Xiao-Wei Wang, Dan Nie, and Bao-Liang Lu. "Emotional state classification from EEG data using machine learning approach." Neurocomputing, Volume 129. ISSN 0925-2312. Elsevier, 2014

[6] Kuo, Fang-Fei, et al. "Emotion-based music recommendation by association discovery from film music." Proceedings of the 13th annual ACM international conference on Multimedia.