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

EXPRESSION BASED MUSIC PLAYER USING DEEPLEARNING TECHNIQUE

PURNA SUDHA SRIHASA ALLURI
Department of Master of Computer Science
Miracle Educational Society Group of Institutions
Vizianagram– 535216 (AP) India

SARAGADAM SRIDHAR
Department of Master of Computer Science
Miracle Educational Society Group of Institutions
Vizianagram– 535216 (AP) India

Abstract

Human expression plays a vital role in determining the current state and mood of an individual, it helps in extracting and understanding the emotion that an individual has based on various features of the face such as eyes, cheeks, forehead or even through the curve of the smile. Music is basically an art form that soothes and calms human brain and body. Taking these two aspects and blending them together our project deals with detecting emotion of an individual through facial expression and playing music according to the mood detected that will alleviate the mood or simply calm the individual and can also get quicker song according to the mood, saving time from looking up different songs and parallel developing a software that can be used anywhere with the help of providing the functionality of playing music according to the emotion detected. By developing a recommendation system, it could assist a user to make a decision regarding which music one should listen to helping the user to reduce his/her stress levels. The user would not have to waste any time in searching or to look up for songs and the best track matching the user's mood is detected, and songs would be shown to the user according to his/her mood. The image of the user is captured with the help of a webcam. The user's picture is taken and then as per the mood/emotion of the user an appropriate song from the playlist of the user is shown matching the user's requirement..

Keywords— Convolutional Neural Networks (CNNs) **,** Support Vector Machines (SVMs), Recurrent Neural Networks (RNNs) , K-Nearest Neighbor (KNN), Long Short-Term Memory(LSTM). Connectionist Temporal Classification(CTC)

Introduction

Music plays an important role in our daily life. Users have to face the task of manually browsing the music.Computer vision is a field of study which encompasses on how computer see and understand digital images and videos. Computer vision involves seeing or sensing a visual stimulus, make sense of what it has seen and also extract complex information that could be used for other machine learning activities. We will implement our use case using the Haar Cascade classifier.Haar Cascade classifier is an effective object detection approach which was proposed by Paul Viola and Michael Jones in their paper, "Rapid Object Detection using Boosted Cascade of Simple Features" in 2001.This project recognizes the facial expressions of user and play songs according to emotion. Facial expressions are best way of expressing mood of a person. The facial expressions are captured using a webcam and face detection is done by using Haar cascade classifier. The captured image is input to CNN which learn features and these features are analysed to determine the current emotion of user then the music will be played according to the emotion. In this project, five emotions are considered for classification which includes happy, sad, anger, surprise, neutral. This project consists of 4 modules-face detection, feature extraction, emotion detection, songs classification. Face detection is done by Haar cascade classifier, feature extraction and emotion detection are done by CNN. Finally, the songs are played according to the emotion recognized. Convolutional Neural Networks (CNN) is a specific type of Artificial Neural Network which are widely used for image classification.CNN is a type of deep learning model for processing data that has a grid pattern, such as images,which is inspired by the organization of animal visual cortex and designed to automatically and adaptively learn spatial hierarchies of features, from low- to high-level patterns. CNN is a mathematical construct that is typically composed of three types of layers (or building blocks):convolution, pooling, and fully connected layers. The first two, convolution and pooling layers,perform feature extraction, whereas the third, a fully connected layer, maps the extracted features into final output, such as classification.

A convolution layer plays a key role in CNN, which is composed of a stack of mathematical operations, such as convolution, a specialized type of linear operation.In digital images, pixel values are stored in a two-dimensional (2D) grid, i.e., an array of numbers and a small grid of parameters called kernel, an optimizable feature extractor, is applied at each image position, which makes CNNs highly efficient for image processing.

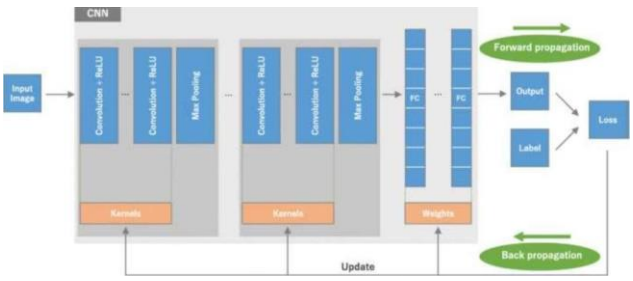


Fig-1: Basic CNN

Proposed System

Convolution neural network algorithm is a multilayer perceptron that is the special design for the identification of two-dimensional image information. It has four layers: an input layer, a convolution layer, a sample layer, and an output layer. In a deep network architecture, the convolution layer and sample layer may have multiple. CNN is not as restricted as the Boltzmann machine, it needs to be before and after the layer of neurons in the adjacent layerfor all connections, convolution neural network algorithms, each neuron doesn't need to experience the global image, just feel the local region of the image. In addition, each neuron parameter is set to the same, namely, the sharing of weights, namely each neuron with the same convolution kernels to the deconvolution image. The key era of CNN is the local receptive field, sharing of weights, subsampling by using time or space, with a purpose to extract featuresand reduce the size of the training parameters. The advantage of CNN algorithm is to avoid the explicit feature extraction, and implicitly to learn from the training data.
The proposed human emotion recognition system is of five components: input speech signal, pre- processing, feature extraction and selection, classification and finally emotions recognition. system architecture diagram is a visual representation the shown in the below

Emotion Extraction Module

The image of the user is captured with the help of a camera/webcam. Once the picture captured, the frame of the captured image from webcam feed is converted to a grayscale image to improve the performance of the classifier, which is used to identify the face present in the picture. Once the conversion is complete, the image is sent to the classifier algorithm which, with the help of feature extraction techniques can extract the face from the frame of the web camera feed. From the extracted face, individual features are obtained and are sent to the trained network to detect the emotion expressed by the user. These images will be used to train the classifier so that when a completely new and unknown set of images is presented to the classifier, it is able to extract the position of facial landmarks from those images based on the knowledge that it had already acquired from the training set and return the coordinates of the new facial landmarks that it detected. The network is trained with the help of CK extensive data set.

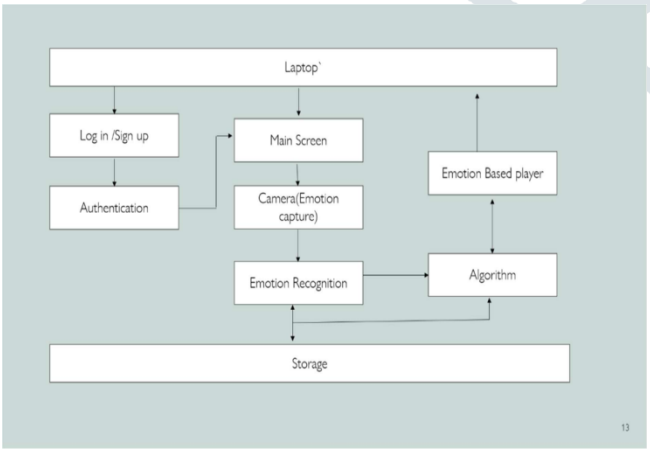
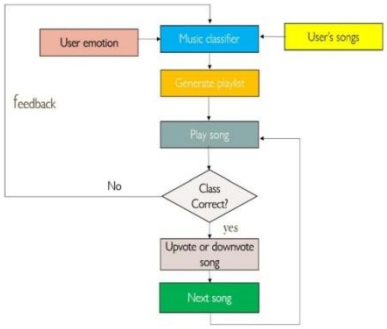
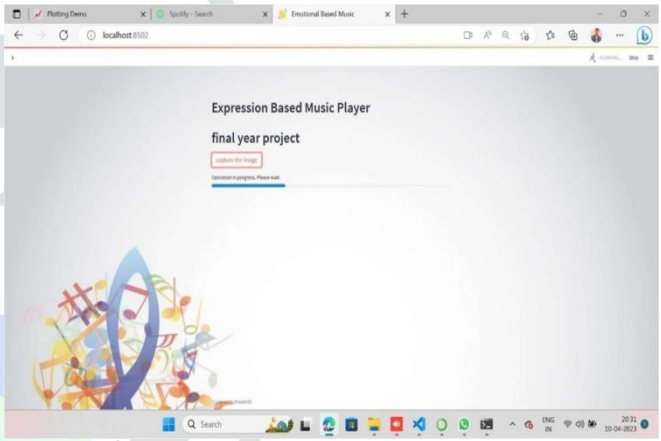


Fig 4.2 System Architecture of Emotional Based Music Player

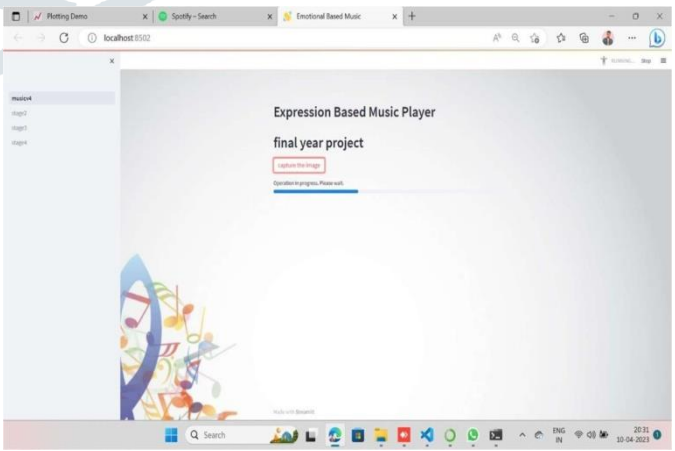


4.2 Audio Extraction Module

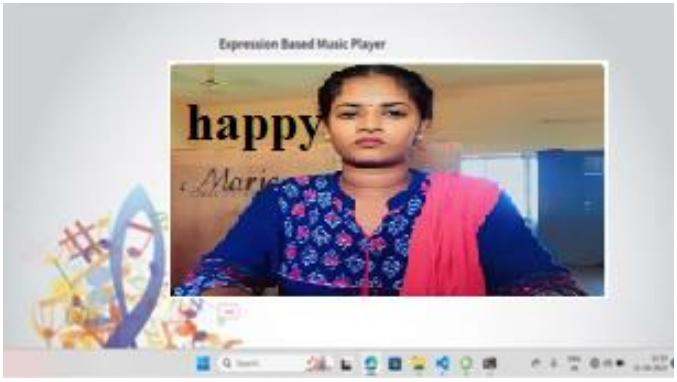
After the emotion of the user is extracted the music/audio based on the emotion voiced by the user is displayed to the user, a list of songs based on the emotion is displayed, and the user can listen to any song he/she would like to. Based on the regularity that the user would listen to the songs are displayed in that order.



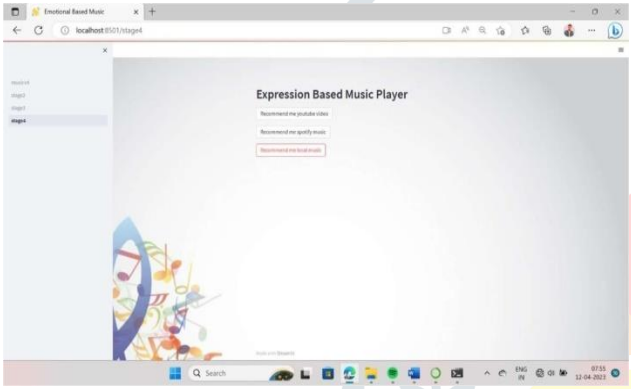
1) **Home Page of Expression Based Music Player**



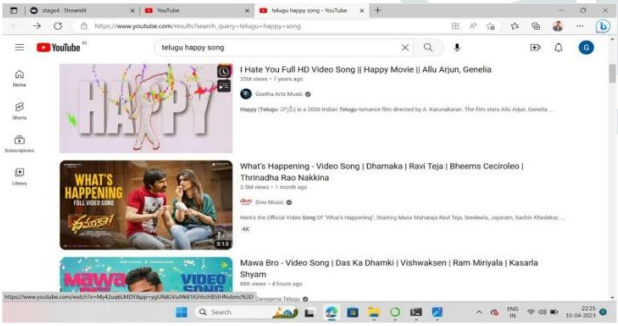
2) **Loading page for capturing the image**



3) Capturing Expression



4) Buttons Option for Music Selection



5) Selecting and Playing Music from you tube

Conclusion and Future Work

This proposed system is developed with the help of a Machine learning module in Python Script. We have used OpenCV, TensorFlow, Keras, Web Browser and of course Python Scripting to develop this automated system. This system has actually been prepared for the easeof end users and simplifying one's life. This will eventually calm the person and appreciatetheir wise choice that will help him to improve their mood. The songs will be played on YouTube along with a particular emotion playlist and using its advanced feature, one can listen to songs happily after we are generating the playlist according the emotion of the user, we developed an application for predicting the emotion of the user using Convolution neural networks and for generating the playlist we have used Spotify API. We have applied it on various imagesand achieved an accuracy of above 70%. We want to extend our work by creating a real time music player which generates playlist and play songs according to the mood of the user.In future Music Player can be enhanced with Google play music, so songs which are not present in local storage can also be played and to access the whole application in speech based. The Emotion Based Music System will be of great advantage to users looking for music based on their mood and emotional behaviour. It will help reduce the searching time for music thereby reducing unnecessary time and hence increasing the overall accuracy and efficiency of the system. The system will not only reduce physical stress but will also act as a boon for the music therapy systems and may also assist the music therapist to treat the patient. In future it can also be used to detect the sleepy mood of the driver, driving the car and many more uses. Also with its additional features mentioned above, it will be a complete system for music lovers and listeners.

[1] S.L.Happy and A. Routray, "Automatic facial expression recognition using features of salient facial patches," in IEEE Transactions on Affective Computing, vol. 6, no. 1, pp. 1-12, 1 Jan.-March 2015.

[2] Rahul ravi, S.V Yadhukrishna, Rajalakshmi, Prithviraj, "A Face Expression Recognition Using CNN & LBP",2020 IEEE.

[3] Krittrin Chankuptarat, Raphatsak Sriwatanaworachai, Supannada Chotipant Emotion-Based Music Player,978- 1-7281-0067-8/19/$31.00 ©2019 IEEE.

[4] Karthik Subramanian Nathan, manasi arun, Megala S kannan," EMOSIC — An emotion-based music player for Android,2017 IEEE.

[5] S.Deebika, K.A.Indira, Jesline, "A Machine Learning Based Music Player by Detecting Emotions",2019 IEEE.

[6]Preema J.S, Rajashree, Sahana M, Savitri H, Review on facial expression-based music player, International Journal of Engineering Re-search & Technology (IJERT), ISSN-2278-0181, Volume 6, Issue 15, 2018.

[7] AYUSH Guidel, Birat Sapkota, Krishna Sapkota, Music recommendation by facial analysis, February 17, 2020.

[8] CH. sadhvika, Gutta.Abigna, P. Srinivas reddy, Emotion-based music recommendation system, Sreenidhi Institute of Science andTechnology, Yamnampet, Hyderabad; International Journal of Emerging Technologies and Innovative Research (JETIR) Volume 7,Is-sue 4, April 2020.

[9] Vincent Tabora, Face detection using OpenCV with Haar Cascade Classifiers, Becominghuman.ai,2019.

[10] Zhuwei Qin, Fuxun Yu, Chenchen Liu, Xiang Chen. How convolutional neural networks see the world - A survey of convolutional neuralnetwork visualization methods. Mathematical Foundations of Computing, May 2018.

[11] Ahmed Hamdy AlDeeb, Emotion- Based Music Player Emotion Detection from Live Camera, ResearchGate, June 2019.

[12] Frans Norden and Filip von Reis Marlevi, A Comparative Analysis of Machine Learning Algorithms in Binary Facial Expression Recogni-tion, TRITA-EECS-EX-2019:143.

[13] P. Singhal, P. Singh and A. Vidyarthi (2020) Interpretation and localization of Thorax diseases using DCNN in Chest X-

Ray. Journal ofInformatics Electrical and Elecrtonics Engineering,1(1), 1, 1-7

[14] M. Vinny, P. Singh (2020) Review on the Artificial Brain Technology: BlueBrain. Journal of Informatics Electrical and Electronics Engi-neering,1(1), 3, 1-11.

[15] A. Singh and P. Singh (2021) Object Detection. Journal of Management and Service Science, 1(2), 3, pp. 1-20.

[16] A. Singh, P. Singh (2020) Image Classification: A Survey. Journal of Informatics Electrical and Electronics Engineering,1(2), 2, 1-9.

[17] A. Singh and P. Singh (2021) License Plate Recognition. Journal of Management and Service Science, 1(2), 1, pp. 1-14.

Online Resources:
1.https://creately.com/diagram/example/ig8yznsm1/emotion-based-music-player-usecase-
2.https://prezi.com/oes7flzuiqwp/emotion-based-music-player/

Conclusion and Future Work

This proposed system is developed with the help of a Machine learning module in Python Script. We have used OpenCV, TensorFlow, Keras, Web Browser and of course Python Scripting to develop this automated system. This system has actually been prepared for the easeof end users and simplifying one's life. This will eventually calm the person and appreciatetheir wise choice that will help him to improve their mood. The songs will be played on YouTube along with a particular emotion playlist and using its advanced feature, one can listen to songs happily after we are generating the playlist according the emotion of the user, we developed an application for predicting the emotion of the user using Convolution neural networks and for generating the playlist we have used Spotify API. We have applied it on various imagesand achieved an accuracy of above 70%. We want to extend our work by creating a real time music player which generates playlist and play songs according to the mood of the user.In future Music Player can be enhanced with Google play music, so songs which are not present in local storage can also be played and to access the whole application in speech based. The Emotion Based Music System will be of great advantage to users looking for music based on their mood and emotional behaviour. It will help reduce the searching time for music thereby reducing unnecessary time and hence increasing the overall accuracy and efficiency of the system. The system will not only reduce physical stress but will also act as a boon for the music therapy systems and may also assist the music therapist to treat the patient. In future it can also be used to detect the sleepy mood of the driver, driving the car and many more uses. Also with its additional features mentioned above, it will be a complete system for music lovers and listeners.