Journal of Engineering Sciences
Vol 11, Issue 6,June/2020
ISSN NO:0377-9254

Mood Based Music Player

Sajida Begum[1], Sakshi Manjari[2], Pranali Sawant[3], Prof. Gayatri Hegde[4]

1 - Student - Information Technology, Pillai College of Engineering, India
2 - Student - Information Technology, Pillai College of Engineering, India
3 - Student - Information Technology, Pillai College of Engineering, India
4- Faculty - Information Technology, Pillai College of Engineering, India

Abstract- Facial expressions are one of the best ways to determine how a person is feeling currently. The aim of the project is to generate a website where user's mood will be identified based on their facial expression and music will be generated according to their emotional state. The project consists of two models: 1. Emotion detection model and 2. Music generation model. In emotion detection model , facial expressions will be identified and classified into seven sentiment categories : "Happy, Sad, Angry, Neutral, fear, Sad, Surprised and Disgust" using CNN(Convolutional neural Network). In music generation model, music will be generated according to the identified emotion using LSTM architecture (Long Short- Term Memory). LSTM is a recurrent neural network (RNN) architecture that remembers values over arbitrary intervals.

I. INTRODUCTION

Music plays a very important role in everyone's life as its one of the important sources of entertainment as well as a way to help in dealing with one's present emotional state. Sometimes it can be very time consuming and hectic to find music depending upon the current mood of an individual. Hence, we have developed a mood-based music player where music will be generated according to the emotional state of the user. To determine the emotional state of a person, facial expressions are one of the best ways as it reflects how the person is feeling currently. In this paper, firstly, we capture the facial expressions of the user as it is one of the best ways to determine how the person is feeling currently using Convolutional Neural Network (CNN). The facial expressions are categorized into 7 categories - "Happy, Sad, Angry, Neutral, fear, Sad, Surprised and Disgust" Secondly, music is generated depending upon the captured expressions that indicates certain emotions using Long - Short Term Memory (LSTM).

II. LITERATURE SURVEY

It has been observed that there are majorly two methods for Emotion Detection (Image classification) : using CNN (Convolutional Neural Network) and SVM (Support Vector Machine).In Music generation, method varied depending on the type of music files used as dataset such as WAV, MIDI, ABC, etc. (i) The paper SentiMozart: Music Generation based on Emotions [2] represents the project in two parts: First, emotion of people is captured from their images and is categorized into 7 major categories: Angry, Disgust, Fear, Happy, Sad, Surprise and Neutral using Convolutional Neural Network (CNN). Second, music is generated based on these emotions using Long Short-Term Memory (LSTM). (ii) The paper MoodyPlayer: A Mood Based Music Player [1]represents following stages: in first stage, face detection is done from an image, for this various techniques are used such as model based face tracking which includes real-time face detection using edge orientation matching, Robust face detection using Hausdorff distance , weak classifier cascade which includes Viola and Jones algorithm, and Histograms of OrientedGradients (HOG) descriptors. In the next stage, features are extracted from the detected face. (iii) **A Survey: Expression Based Music Player [5]**, this paper deals with connecting emotion of the user along with music systematically. Expression based Music Player involves the image processing, facial feature detection, expression classification and audio feature extraction. In many research papers on Emotion Detection with Music Generation, the biggest

problem with the systems was manual setting of the user's emotion. After the survey it has been observed that for emotion detection, CNN performed much better than SVM. Additionally, SVM required many other steps including image processing, face detection, facial feature extraction, etc.

III. IDENTIFY, RESEARCH AND COLLECT IDEA

We reviewed some implementation techniques like SVM (Support Vector Machine), CNN (Convolutional Neural Network) for emotion detection and RNN (Recurrent Neural Network), LSTM (Long Short-Term Memory) for music generation. After thoroughly reviewing and understanding, out of these methods we chose CNN for emotion detection because of its high accuracy and LSTM for music generation as it can maintain information in memory for longer period of time. Dataset used for emotion detection was Fer2013 (the collection

of 35,887 grayscale images of 48 X 48-pixel dimensions). The data consists of 48x48 pixel grayscale images of faces. The task was to categorize each face based on the emotion shown in the facial expression into one of seven categories (0=Angry, 1=Disgust, 2=Fear, 3=Happy, 4=Sad, 5=Surprise, 6=Neutral). Following are the statistics of the number of images present in the dataset for each emotion:

0: -4593 images- *Angry*
1: -547 images- *Disgust*
2: -5121 images- *Fear*
3: -8989 images- *Happy*
4: -6077 images- *Sad*
5: -4002 images- *Surprise*
6: -6198 images- *Neutral*

MIDI files were used to train the music generation model. We collected the MIDI files and prepared a dataset of music files for each emotion.

IV. METHODOLOGY

Mood Based Music player consists of two models Emotion detection and music generation model.

Emotion Detection

Emotion detection model uses Convolutional Neural Network (CNN) for recognizing facial expressions. In our proposed system, the CNN model is trained on FER2013(Facial Expression Recognition) dataset. The CNN model is trained with images as batch which contains 64 images for 20 epochs. The model will generate output with 7 possibilities of the input image. The weights are optimized using Adam optimization algorithm i.e. the network weights are updated in an iterative approach. Adam optimization technique is chosen over Stochastic gradient descent (SGD) because in Adam optimization the learning rates are learned on a per parameter basis unlike SGD which has a single global learning rate that is applied to all the parameters. L2 regularization technique is used to avoid overfitting of the model. Loss is calculated using categorical cross entropy which is usually used to calculate loss in a model which performs classification based on labels. In the system the webcam captures the image and then the captured image is fed as an input image to the model for detecting the emotion of the user. Using Haar-Cascade face detection algorithm, the face is detected in the captured image. The model now predicts the facial expression of the detected face.

Music Generation

The dataset used for music generation consists of music files of MIDI (Musical Instrument Digital Interface) format for each emotion. Unlike mp3, midi files cannot be played, as midi files do not contain actual audio. It contains information about the notes, chords, velocity of the notes, etc. Basically, they are an instructional file that provide information like what notes are being played, for how long are they played, how loud the notes are, etc. For music generation midi files are preferred over wav or other format files because comparatively they are small in size and very informative. Before training the model, these files are preprocessed i.e. encoded in the appropriate format that is more suitable for training the data. LSTM (Long Short-Term Memory) is used for music generation. Music generation requires a neural network which would remember a sequence for a longer period of time and RNN (Recurrent Neural Network) are capable of doing this. RNN consider each and every past event and the present event for predicting the future event. But in music generation, it is not necessary for remembering all the past events which is where LSTM play an important role in music generation. The optimization technique used in music generation is the same as the one used in emotion

detection model. The loss function used during training is sparse cross entropy loss function. For generating music, the model is given an array consisting of randomly selected 50 notes. Now the model generates an array containing 300 notes by predicting the further notes. These notes are then decoded back to the original midi format. After decoding the midi file is then converted to mp3 format which is then played on the website after the emotion is detected.

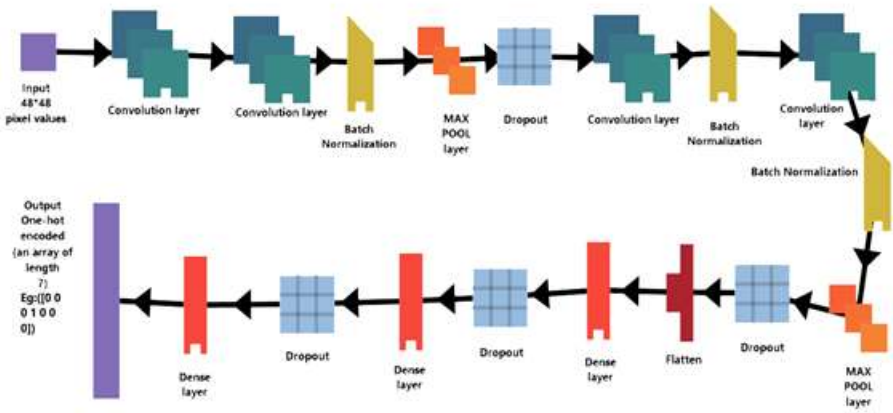

Fig 1: Emotion Detection Architecture

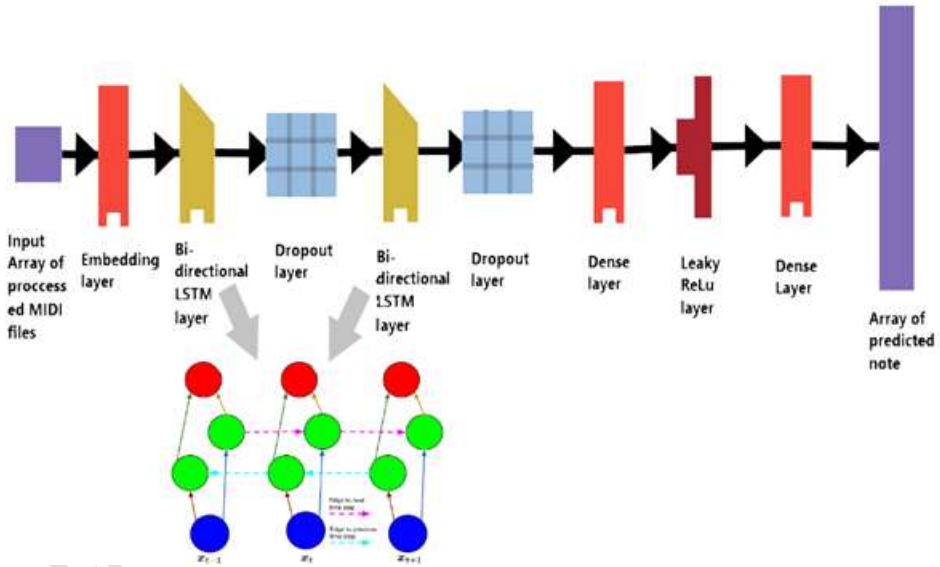

Fig 2: Music Generation Architecture

V. EXPERIMENTAL RESULT AND ANALYSIS

Tools Used

In Emotion Detection, Keras was primarily used for classification. Along with it the libraries used were Pandas, NumPy, OpenCV. For Music Generation, TensorFlow was used. PrettyMidi library was used for encoding and decoding of the midi files.

Results

The accuracy of the trained model is evaluated using the testing dataset which is split from the original dataset. 10% of the original dataset is used for testing and 10% for validation. The Emotion Detection model's accuracy is measured by directly comparing the model's predicted output to the actual output of the image. In this system, the model's accuracy is evaluated using Categorical Cross Entropy. Performance of a classification model in which output is a probability value between 0 and 1 is given by Cross-entropy. The results of the proposed application

are shown below. The image is given as an input through webcam to the website. OpenCV is used to process the images taken by the webcam. The Fig shows emotions detected such as happy, sad, angry, surprise, neutral, fear images. The predicted array will show the probabilities of all the emotions. These probabilities lie between the range 0 and 1. The emotion which has the highest probability is taken as the final emotion of the user. The total accuracy of the emotion detection model is 64.6% for testing data. The validation accuracy is 69.7%. Given below is the demo of emotion detection model-

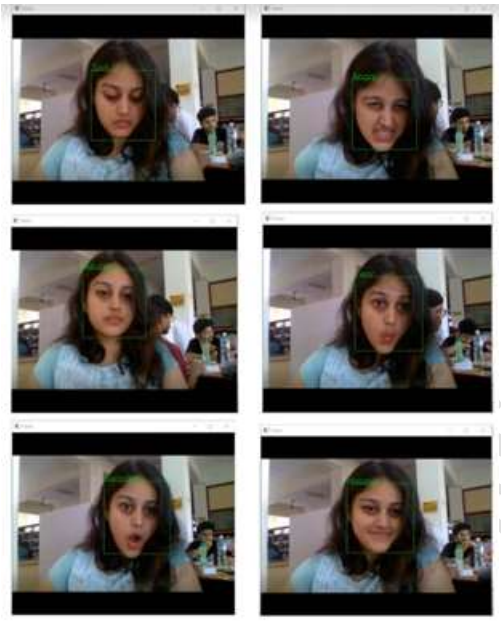

Fig 3: Output images with emotion detected

VI. FUTURE SCOPE

In future, this project can be further extended by using Voice Sentiment analysis can be applied for recognizing user's emotion along with the facial emotion recognition. In future mood-based music player can be incorporated in social media apps. It can also be used in several meditational apps for soothing user's mood.

VII. CONCLUSION

This mood-based music player detects the user's emotions and generates music according to the user's mood. In many existing systems, predefined playlists are provided to the user but in our proposed system the user is given a choice if he/she wants to listen to the generated tunes or he/she wants to listen to the predefined playlists. In this system, CNN works efficiently in recognizing the emotion of the user with an accuracy of 64.6%. The music generation model every time generates a unique tune for each emotion.

ACKNOWLEDGEMENT

We would like to express our gratitude towards our guide Prof. Gayatri Hegde who encourages us as well as helps us to solve our queries. We would like to thank her for her kind cooperation and encouragement which helped us in completion of this project. We give a special gratitude to our honorable principal Dr. Sandeep Joshi who always encourage us and motivate us to do innovative things that will increase our knowledge. We would also like to thank our H.O.D of Information Technology Department Dr. Satishkumar Varma for his guidance and also for giving us the opportunity to implement the project.

REFERENCES

[1] Abhishek R. Patel, Anusha Vollal, Pradnyesh B. Kadam, Shikha Yadav, Rahul M. Samant, "MoodyPlayer: A Mood based Music Player", International Journal of Computer Applications (0975 – 8887): Volume 141 – No.4, May 2016

[2] Rishi Madhok, Shivali Goel and Shweta Garg, "SentiMozart: Music Generation based on Emotions," Delhi Technological University, New Delhi, India: In Proceedings of the 10th International Conference on Agents and Artificial Intelligence (ICAART 2018)

[3] Maruthi Raja S K, Kumaran V, Keerthi Vasan A, Kavitha N, "Real Time Intelligent Emotional Music Player" Saranathan College of Engineering. Trichy, India: Journal for Research | Volume 03| Issue 01 | March 2017

[4] Karan Mistry, Prince Pathak , Prof. Suvarna Aranjo, ''Mood based Music Player" International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056 Volume: 04 Issue: 3 | Mar -2017

[5] Celina Jenefer C, Leena. S, Nirmala Devi M, Dr. J. SelvaKumar, ''A Survey: Expression Based Music

Player" Department of Computer Science and Engineering, Sri Ramakrishna Engineering College, Coimbatore, Tamil Nadu, India:2017 IJSRST | Volume 3 | Issue 3

[6] Aditya Gupte, Arjun Naganarayanan, Manish Krishnan, ''Emotion Based Music Player – XBeats" Department of Computer Engineering, SIES GST, Mumbai University, Navi Mumbai, India: International Journal of Advanced Engineering Research and Science (IJAERS) [Vol-3, Issue-9, Sept-2016]

[7] Hemanth P, Adarsh, Aswani C.B, Ajith P, Veena A Kumar, "EMO PLAYER: Emotion Based Music Player", Department of Computer Science, Saintgits College of Engineering, Pathamuttom, Kottayam, Kerala, International Research Journal of Engineering and Technology (IRJET) Volume:
05 Issue: 04 | Apr-2018

[8] Sri Charan Nimmagadda, "Emotion Based Music Player", Department of Computing Sciences Texas A&M University - Corpus Christi Corpus Christi, Texas

[9] Rahul Hirve, Shrigurudev Jagdale, Rushabh Banthia, Hilesh Kalal & K.R. Pathak, "EmoPlayer: An Emotion Based Music Player", Imperial Journal of Interdisciplinary Research (IJIR) Vol-2, Issue-5, 2016

[10] Harshada Sonkamble, Prof. Ujwala.V. Gaikwad, "Emotion Recognition Based on Efficient Self Organized Map", Terna Engineering College, Nerul, International Research Journal of Engineering and Technology (IRJET) Volume: 04 Issue: 04 | Apr - 2017