International Journal of Engineering Applied Sciences and Technology, 2023
Vol. 8, Issue 02, ISSN No. 2455-2143, Pages 42-46
Published Online June 2023 in IJEAST (http://www.ijeast.com)

MOOD PLAYER: EMOTION BASED MUSIC RECOMMENDATION SYSTEM

Kaif Ahmad, Aditi Prashar, Jitesh Aggarwal, Pratham Tripathi, and Riddhi Maniktalia
SCOPE, Vellore Institute of Technology,
Vellore, Tamil Nadu, India}

Abstract: **Music can create, change and lighten one's mood, and when technologies of artificial intelligence meet with music it can be a great combination. Selecting appropriate songs according to our emotion from such a wide range is a very tedious and time-consuming task. It takes a lot of time and effort to manually group songs on a playlist and annotate them according to the user's emotional condition at the moment. To automate this process, many algorithms have been proposed. However, the current algorithms are inefficient, use more hardware, which raises the system's overall cost, and are less accurate. This paper aims at recognition of human facial expressions through a model based on 68 landmarks approach of the training dataset, further selecting the appropriate playlist for the user. The Emotion Based Music Recommendation System has the potential to enhance the music listening experience by providing users with music that not only matches their preferences but also their emotional state, thereby increasinguser engagement and satisfaction.**

Keywords: **Emotion, Music Player, OpenCV, CLAHE, Faciallandmarks, CNN, SVM and HOG:**

I. INTRODUCTION

Facial expressions are a great way to express emotions, and music has always been agreat source to lighten the mood or make the moments more enjoyable. Capturing, recognizing and detecting the emotion that is being voiced by a person and displayingappropriate songs can increasingly calm the mind of a user and provide satisfaction.It's often that a person has a lot of music in his/her playlist. So instead of doing ahectic task of selecting a song from that huge playlist, the user will randomly select asong which might not be in accordance with his/her mood and can disappoint him/her.

Currently the music players are in accordance with the playlist given. The song we play in our playlist randomly depends on our mood. Often, we want to hear songsin accordance to our mood, for that we often have to do a long search, which is time consuming. Once we select a song it is not necessary that the next song will be in the same category we need depending on our mood. Every time we were required to select songs according to our mood from the playlist.

So, this process takes a lot of time which often becomes irritating. Due to all these reasons, we should have an application that can play songs in accordance with our mood and which can save a lot of our time. This application will extract the emotion of the user by capturing the facial feature of the user by his/her device's camera therefore detecting the emotion of the user.

Music lovers frequently find it difficult to manually create and arrange playlists when working with a sizable song collection. Also, it can be difficult to keep track of every song and make sure that each one is played because unused songs take up storage space, which results in waste. Users frequently have to choose music manually based on their tastes and emotions, which might take time. Playlists are used by existing music programmes to organise tracks, but users must manually edit each song's playlistbecause it is difficult to replace or adjust it. Also, the playlist's song order can change every time, and songs that are regularly played might not be given priority or might even be excluded. It is difficult to find specific songs because there are currently no programmes that enable users to play songs instantaneously without choosing them manually or from a playlist. Thus it would be quite helpful to have a music player that chooses songs based on the user's current mood.

Facial Expression based Music Player is an interactive, sophisticated and innovative mobile (Android) based application to be used as a music player in a different manner. A music player is designed to capture the human's emotion using the web cam interface. The main objective of our project is to develop an application that will help the user to listen to music according to his emotion, save time which users spends in searching for music in their playlists and connect the emotion recognition system with online popular streaming and media service providers like Spotify, so that it will satisfy maximum user needs.

II. RELATED WORK

To successfully identify human emotions, a number of methods have been put out and implemented. Seven fundamental emotions have received the majority of this attention because to their universality across cultures,

ages, and other characteristics.For example, Rohit[1] et al used the Viola Jones algorithm, which is used for face detection and then forwards it to the SMV. Different emotion categories with different weights has already been assigned to SMV and as soon as the emotion gets closestto the assigned weight-age and is detected, output is transfered to the audio modulewhich further plays the song associated with the mood accordingly. H. Immanuel[2] et al obtained the required frames from the recorded video of the person or subject using the camera, by preprocessing the facial expressions and, further converted into (AUs) that is a Sequence of Action units from the frames and images obtained from the camera. The Facial Action Coding System (FACS) using a combination of AUs(64) is brought into work to describe all the recorded facial expressions Further, the extracted emotions are categorized as Happy, Sad, Angry, and Surprise. Sarvesh[13] et al had a very similar approach.
Another approach was by Anukirti[3] et al, they obtained the Binary images from
the converted gray-scale and RGB images recorded. Their input was aas preprocessedimages in into the face detection block. Detection of Face is further carried out us- ing Viola-Jones algorithm. For Expression Recognition, only eyes and mouth are taken into consideration to reduce time complexity and for real time performance. Simi- lar method was adopted by Nikhil[4] et al, where Binary images obtained from RGB format, perform OpenCV and Facial Detection using Haar Cascade, using effective ex- tracting processes that feature, point detection, lip feature detection and eye feature detection.
The proposed system by Rahul[5] et al involves various modules like testing image, input image, face detection, landmark point extraction, training of SVM & Trained SVM and Music Player. Captured Image is Obtained using a webcam. JAFFE database is brought into use to train image, further detection is operated using Viola-Jones Algorithm. 68 landmark points were detected and processed to SMV for training and testing using the landmark detection feature. Sayali[6] et al performed face de- tection through Viola Jones algorithm. Feature extraction is carried out using an appearance-based method. Here in the Proposed work, two facial expression databases are to be used. FG-NET Facial Expression and Emotion Database that consists recorded base of MPEG video files with spontaneous emotions. Further Classifier is operated for a system that is (SVM) with Radial based Kernel Working.
Sulaiman[7] et al Proposed work contains combination of two Vanilla CNN models in addition to few transfer learing models. All the above methods and models are executed and compared. After analysing the results , ensemble with soft voting is per-formed to five models to improve their accuracy to 76.12%.As a result GAP model uses the least amount of parameters in comparison to other model in pace with achieving a decent percent of accuracy. The paper by Aayush[8] et al involves use of Electroen- cephalography (EEG) signals. EEG signals detects Emotions and EEG records the electrical activity within the neurons of the brain. Independent Component Analysis (ICA) and Machine Learning techniques such as Support Vector Machine (SVM) and Linear Discriminant Analysis (LDA) are used to classify EEG signals into seven differ- ent emotions. Higher Accuracy, prevention of over-fitting and working well coordinated with non-linear Date are Pros of SVM, it also recognizes Emotions with a decent Ac- curacy of 87.5% and 92.5% respectively whereas in the case of LDA the accuracy fellto 82.5% and 87.5% respectively.
A machine learning based algorithm approach was taken up by Vinay[9] et al. The proposed model is assembled using SMV Machine Learning Algorithm. Uploading of the image is performed and then using OPENCV (PYTHON) extraction of featuresof uploaded images is performed, further by the application of these features on SMV training model we predict the obtained User's Emotion. Acquired Output is convertedas Input and using it, the emotion of the user will be used to discover the song. Further, a list associated with emotion will be displayed to the user and he can choose and play accordingly. S. Deebika[10] went along with the SMV machine learning algorithm as well.
A novel methodology was taken up by Rishi[11].This work is categorised into two models, the Image Classification Model and the Music Generation Model.The Music Generation Model is essentially a Doubly Stacked LSTM architecture. This is per- formed after Categorization and identification of facial expression into these major emotional segments: Angry, Disgust, Fear, Happy, Sad, Surprise and Neutral, which is performed by (CNN) Convolutional Neural Networks. Pushkar[12] et al carry for- warded there research using the Haar Cascade classifier for face detection and SMV classifier for extracting emotions.
Chavi[14] used the Haar Cascade classifier just like other researchers, but used to the Fisherface method to detect emotions. Here the RGB to binary image helped in facial recognition. The novelties of Jayshree[15] et al was in the use of JavaCV library, and in Nanda[16] et al work was the high accuracy. Also, Hafeez[17] et al used the Frontal Cart property to convert RGB images to binary.
Very few systems have been developed to manage the creation of a music playlist using human emotions, despite the fact that numerous methods have been presented to distinguish face emotions and emotions in an audio stream. The ideas that have already been created and studies that have been done to govern playlist generation using human emotions either employ extra hardware, or human speech. The research given in this study

International Journal of Engineering Applied Sciences and Technology, 2023
Vol. 8, Issue 02, ISSN No. 2455-2143, Pages 42-46
Published Online June 2023 in IJEAST (http://www.ijeast.com)

tries to address a number of flaws in the literature. By creating an automated music recommendation system that would create a personalised playlist based on a user's facial expression without needing any additional hardware, this article tries to address the shortcomings observed in the above papers.

III. METHODOLOGY

Dataset

The FER-2013 dataset from Kaggle website [18] consists of 28,000, 48x48 pixel grayscale images. The training set consists of 28,709 images and the public test set consists of 3,589 images and 3,500 labelled images in the development set. For our model we only require happy, neutral, sad and angry images from the dataset which is 5,212 test im-ages and 21,005 training images. FER-2013 is labelled as one of seven emotions, such as happy, sad, angry, afraid, surprise, disgust, and neutral, with happy being the most prevalent emotion, providing a baseline for random guessing of 24.4%.

Facial Emotion Recognition System

There are 3 major steps in our Facial Emotion Recognition System:

- Image pre-processing
- Facial Landmark Prediction
- Expression Classification (Feature Extraction)

Image pre-processing

The dataset consists of images captured in a wide range of lighting conditions. Thus, to ensure that all images are equalized to similar lighting conditions, Contrast Limited Adaptive Histogram Equalization (CLAHE) is performed on all the images in the dataset using OpenCV built-in function.

Facial Landmarks Prediction

For the face detection in our image, we used a pre-trained detector based on Histogram of Oriented Gradients (HOG) features, and a linear Support Vector Machine (SVM) in a sliding window detection approach. Then to extract the key facial features from the input image, we used the facial landmark predictor and the shape prediction which will mark the 68 coordinates (Fig 1) on the image categorizing different facial features

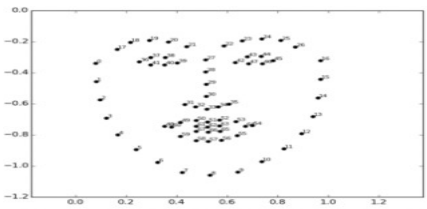

Figure 1: 68 facial landmarks

Expression Classification

We trained Convolutional Neural Network over the dataset which has an architecture that consists of Convolution layer, Activation layer, pooling layer, and Fully Connected layer (as shown in Fig 2). These are all interconnected so that CNNs can process and perceive data in order to classify the emotions of the image.

Convolution Layer: CNNs use image recognition and classification in order to detect emotions. They are made up of neurons with learnable weights and biases. Each specific neuron receives numerous inputs and then takes a weighted sum over them, where it passes it through an activation function and responds back with an output.

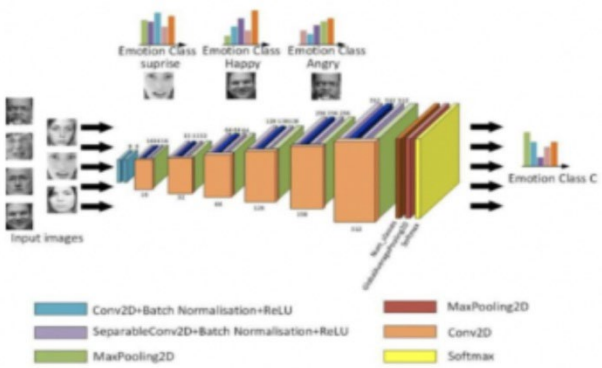

Figure 2: Expression recognition CNN Layers

Activation Layer: We have applied ReLu (Rectifier Linear Unit) in our model. This rectifier function is used to increase non-linearity in CNN.

Pooling Layer: In our model, we used max pooling layer. It uses a 2 cross 2 maxfilter with a stride of 2, this is a non-overlapping filter, which is used to return the max value in the feature within the region.

Fully Connected Layer: This involves transforming the entire pooled feature map matrix into a single column which is then fed to the neural network for processing. With the fully connected layers, we combined these features together to create a model. Finally, we apply the activation function softmax to classify the output.x

The architectural diagram of our project is shown in Fig 3.

Music recommendation System

After the facial recognition is done by the model, the song based on the user's emotionis played on Spotify. This is based on simple REST principles, the Spotify Web API endpoints return JSON metadata about music artists, albums, and tracks, directly from the Spotify Data Catalogue.

International Journal of Engineering Applied Sciences and Technology, 2023
Vol. 8, Issue 02, ISSN No. 2455-2143, Pages 42-46
Published Online June 2023 in IJEAST (http://www.ijeast.com)

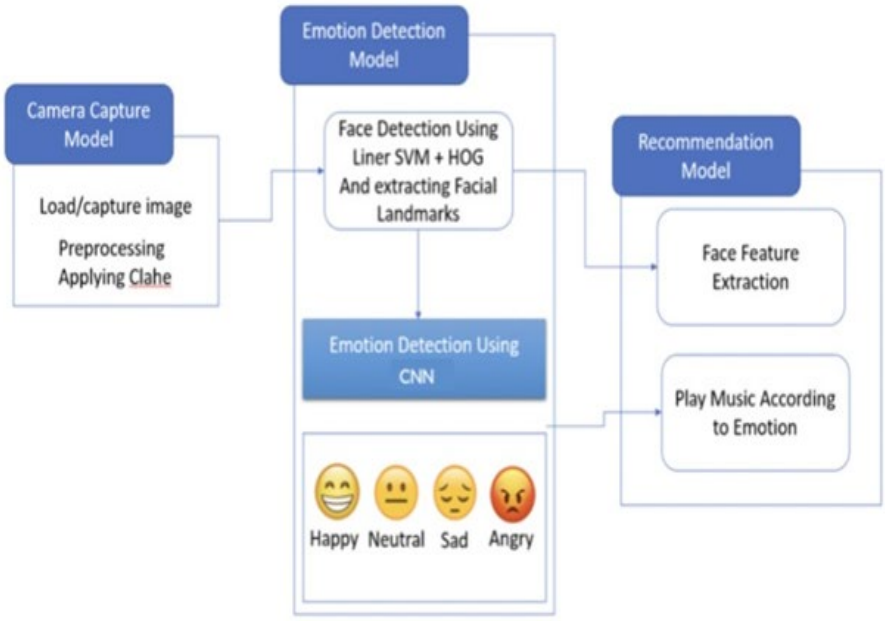

Figure 3: Architectural Diagram

Figure 4: Web API connecting Spotify with our software

Web API also provides access to user related data, like playlists and music thatthe user saves in the Your Music library. Such access is enabled through selective authorization,◊ by the user. The API (shown in Fig 3) provides a set of endpoints, each with its own unique path. To access private data through the Web API, such as user profiles and playlists, an application must get the user's permission to access thedata. Authorization is via the Spotify Accounts service. The image beneath shoes how the model is working.

IV. EXPERIMENTAL RESULT AND ANALYSIS

Implementation and experimentation in this paper is carried out on Windows8, 32 bit operating system and Intel's i5 M460(2.53 GHz) processor.

Result of CNN Model

Coming to the Trained Datasets Parts and outcome we got after facial and emotion detection from available datasets, the complete set of dataset was divided in ratio of 7:3 which then was used for training and testing purpose respectively. Results were quite satisfying as the trained data displayed accuracy of 73.43% and loss of 78% withit.

The average accuracy obtained high value around 70% [Fig 5] with average loss of 77% [Fig 6] in recognizing the emotion. Evaluation Accuracy also stood strong with 73.4362% and Evaluation Loss with 78.333%.

Human Emotion Aware Music Player

The music player runs by the following three steps

◊ To capture emotion by enabling webcam.

Enter space bar to capture the frame at that respective moment and pass the emotion label as an input to the music player(Spotify)

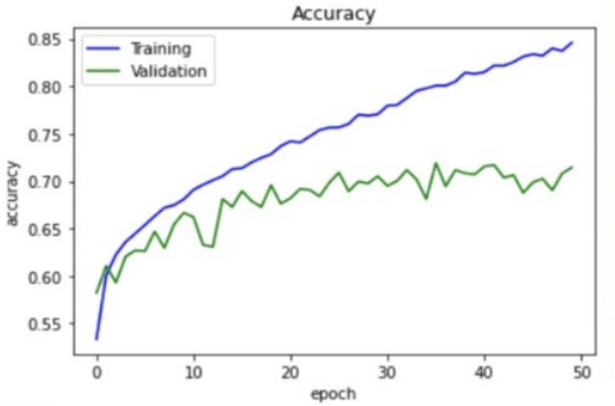

Figure 5: Model accuracy after training CNN

International Journal of Engineering Applied Sciences and Technology, 2023
Vol. 8, Issue 02, ISSN No. 2455-2143, Pages 42-46
Published Online June 2023 in IJEAST (http://www.ijeast.com)

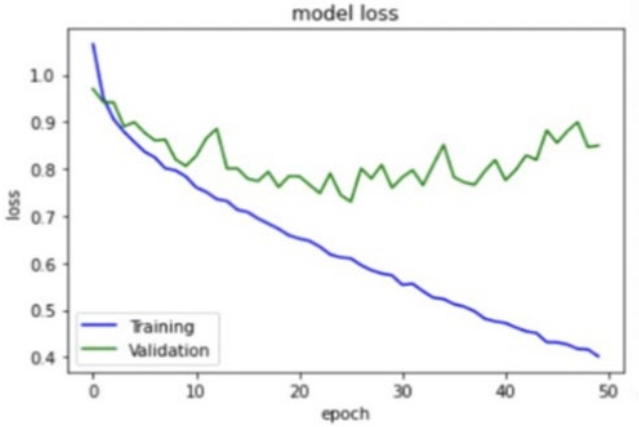

Figure 6: Model accuracy after training CNN

◊ The respective track id of the resultant emotion detected is played.

V. CONCLUSION

In the field of image analysis and computer vision facial expression recognition is a challenging problem and it has managed to receive great deal of attention over the lastfew years because of its many applications in various domains. This paper proposesa human facial expression recognition model based on 68 landmarks approach of the training dataset. Some of the drawbacks regarding this model were: High resolution Camera and good lighting is also required to get the more accurate Facial Expression. Also, dataset needs to be trained more, so it's a very time taking process. The field of research in expression recognition is an area which can be further explored and improved. The system was designed using the facial landmarks scheme and was testedunder various scenarios for the result that would be obtained. It is seen that the classifier has an accuracy of more than 55 percent for most of the test cases, which is pretty good accuracy in terms of emotion classification. It can also be seen that the classifier can accurately predict the expression of the user in a real-time scenario whentested live for a user.

VI. FUTURE WORK

The proposed algorithm was successful in crafting a mechanism for playing music based on facial emotion detection. In future, we will work on reducing the time required to train the classifier and the run time of the application.

VII. REFERENCES

[1] Rohit Kadam, Darshan Gosavi, Chaitanya Sankhe, Amruta Mhatre. Effective Music Player Integrated onUser's Mood.

[2] H. Immanuel James, J. James Anto Arnold, J. Maria Masilla Ruban, M. Tami- larasan, R. Saranya. Emotion Based Music Recommendation System.

[3] Anukriti Dureha. An Accurate Algorithm for Generating a Music Playlist based on Facial Expressions.

[4] Nikhil Zaware, Tejas Rajgure, Amey Bhadang, D. D. Sapka. Emotion BasedMusic Player.

[5] Rahul Hirve, Shrigurudev Jagdale, Rushabh Banthia, Hilesh Kalal ,K.R. Pathak. EmoPlayer: An Emotion Based Music Player.

[6] Sayali Chavan, Ekta malkan, Dipali Bhatt, Prakash Paranjape. XBeats-An emo-tionbasedmusic player.

[7] Sulaiman Muhammad, Safeer Ahmed, Dinesh Naik. Real Time Emotion Based Music Player Using CNN Architectures.

[8] Aayush Bhardwaj, Ankit Gupta, Pallav Jain, Asha Rani, Jyoti Yadav. Classifi- cation of human emotions from EEG signals using SVM and LDA Classifiers.

[9] Vinay p,Raj Prabhu,Bhargav Satish Kumar Y, Jayanth P, A. SUNEETHA. Fa-cial Expression Based Music Recommendation System.

[10] S. Deebika, K. Indira, et al., A machine learning based music player by de- tecting emotions, in: 2019 Fifth International Conference on Science Tech-nology Engineering and Mathematics (ICONSTEM), Vol. 1, IEEE, 2019,pp. 196–200

[11] R. Madhok, S. Goel, S. Garg, Sentimozart: Music generation based on emo- 45tions., in: ICAART (2), 2018, pp. 501–506.

[12] P. Mahajan, P. Khurad, P. Chaudhari, et al., Face player: Facial emotion- based music player, IJRAR-International Journal of Research and AnalyticalRe- views (IJRAR) 7 (1) (2020) 894–897

[13] S. Pal, A. Mishra, H. Mourya, S. Dicholkar, Emo-music (emotion basedmusic player), Tech. rep., EasyChair (2020)

[14] C. Ralhan, K. Mohan, K. V. Raj, P. A. Reddy, P. Saiprasad, Emotion baseds- mart music player

[15] Prof. Jayshree Jha, Akshay Mangaonkar, Deep Mistry, Nipun Jambaulikar, Prathamesh Kolhatkar. A Media Player which operates depending on Human Emo-tions.

[16] Nanda. M. B, Rithik. V Gopal, Soumyadeep Ghosh. Emotion Based MusicPlayer

[17] Hafeez Kabani, Sharik Khan, Omar Khan, Shabana Tadvi. Emotion Based Music Player.

[18] Dataset FER-2013: https://www.kaggle.com/msambare/ fer2013