International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 5, pp 273-277, May-2023

https://doi.org/10.47001/IRJIET/2023.705036

Music Recommendation System

[1]Aman Tamboli, [2]Harshwardhan Almnur, [3]Vivek Patil, [4]Uday Patil, [5]Sheetal Nirve

[1,2,3,4]Student, Computer Engineering, KJ College of Engineering and Management Research, Pune, Maharashtra, India

[5]Professor, Computer Engineering, KJ College of Engineering and Management Research, Pune, Maharashtra, India

***Abstract* - Music plays a significant role in improving and elevating one's mood as it is one of the important source of entertainment and inspiration to move forward. Recent studies have shown that humans respond as well as react to music in a very positive manner and that music has a high impact on human's brain activity. Now-a-days, people often prefer to listen to music based on their moods and interests. This work focuses on a system that suggests songs to the users, based on their state of mind. In this system, computer vision components are used to determine the user's emotion through facial expressions. Once the emotion is recognized, the system suggests a song for that emotion, saving a lot of time for a user over selecting and playing songs manually. Conventional method of playing music depending upon the mood of a person requires human interaction. Migrating to the computer vision technology will enable automation of such system. To achieve this goal, an algorithm is used to classify the human expressions and play a music track as according to the present emotion detected. It reduces the effort and time required in manually searching a song from the list based on the present state of mind of a person. The expressions of a person are detected by extracting the facial features using the Haar Cascade algorithm and CNN Algorithm. An inbuilt camera is used to capture the facial expressions of a person which reduces the designing cost of the system as compared to other methods.**

Keywords: Recommender System, Convolution Neural Network (CNN), Deep Learning, image processing, Artificial intelligence, Classification.

I. INTRODUCTION

Facial expressions are one of the natural means to communicate the emotions and these emotions can be used in entertainment and Human Machine Interface (HMI) fields. In today's world, with the advancements in the areas of technology various music players are deployed with features like reversing the media, fast forwarding it, streaming playback with multicast streams. Although these features satisfy the basic requirements of the user, yet one has to manually surf for the song from a large set of songs, according to the current circumstance and mood. This is a time consuming task that needs some effort and patience. The main objective of this work is to develop an intelligent system that can easily recognize the facial expression and accordingly play a music track based on that particular expression/emotion recognized. The seven universally classified emotions are Happy, Sad, Anger, Disgust, Fear, Surprise and Neutral. The main objective of this work is to develop an intelligent system that can easily recognize the facial expression and accordingly play a music track based on that particular expression/emotion recognized. The seven universally classified emotions are Happy, Sad, Anger, Disgust, Fear, Surprise and Neutral. The algorithm that is used in developing the present system is Haar Cascade algorithm which utilizes eigen faces to extract the facial features. The designed algorithm is very efficient due to less computational time taken hereby increasing the performance of the system.

This work finds its applications in various domains like Human Computer Interaction (HCI), therapeutic approach in health care etc. Most of the time the digital music is sorted and put together based on attributes such as artist, genre, albums, language, and popularity and so on. Many of the available online music streaming services recommend music based on user's preferences and his previous music listening history that employ content based and collaborative filtering recommendations. But these recommendations may not suite the current mood of the user. The manual classification of songs by learning user's preference of emotion is a time consuming task. So, recommendations can also be achieved using the physiological and emotional status of the user which are mainly captured from the user's facial expression, gestures, pulse rate, movement, speech/text interactions etc. Several work is carried to detect emotions using facial landmarks to extract the features. Nguyen et al. detected three kinds of emotions namely positive, negative and blank using 68 facial landmarks. This system work proposes a CNN based approach to recommend music by analyzing the multimodal emotional information captured by facial movements and semantic analysis of the speech/text interactions of the user, thus, intensifying the decision of the system on recognized emotions in real-time. Machine learning has become very popular in recent years. Depending on the type of application and the dataset available, certain types of machine learning techniques are more appropriate than others for different applications. The main types of learning algorithms include supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning. A neural network (NN)

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 5, pp 273-277, May-2023

https://doi.org/10.47001/IRJIET/2023.705036

is a technique of machine learning that is generally effective at extracting critical features from complex datasets and deriving a function or model that expresses those features. The NN utilizes a training dataset to first train a model. After the model is trained, the NN can then be applied to new or previously unseen data-points and classify the data based on the previously trained model.

II. LITERATURE SURVEY

It was observed in a cross-database experiment [1] that raw features worked best with Logistic Regression for testing RaFD (Radboud Faces Database) database and Mobile images dataset. The accuracy achieved was 66% and 36% respectively for both using CK+ dataset as a training set. The additional features (distance and area) reduced the accuracy of the experiment for SVM (Support Vector Machine) from 89%. The algorithm that had been implemented generalized the results from the training set to the testing set better than SVM and several other algorithms. An average accuracy of 86% was seen for RaFD database and 87% for CK+ database for cross-validation=5.The main focus was feature extraction and analysis of the machine algorithm on the dataset. But accurate face-detection algorithms become very important if there are multiple people in the image.

One of the works [10] was tested by deriving expression from the live feed via the system's camera or any pre-existing image available in the memory. It has been implemented using Python 2.7, OpenCV and NumPy. The objective was to develop a system that can analyse the image and predict the expression of the person. The study proved that this procedure is workable and produces valid results.

There has also been research done on the Music Recommendation System. According to one such research [11], a preliminary approach to Hindi music mood classification has been described, that exploits simple features extracted from the audio. MIREX (Music Information Retrieval Evaluation eXchange) mood taxonomy gave an average accuracy of 51.56% using the 10-fold cross validation. In addition to this, there is an article [10] that states that the current music recommendation research results from the perspective of music resources description. It is suggested that there is a lack of systematic research on user behaviour and needs, low level of feature extraction, and a single evaluation index in current research. Situation was identified to be an important factor in the music personalized recommendation system. Finally, it was concluded that when the weights given to all the contextual factors were the same, greatly reduced the accuracy of the recommendation results.

In a particular system [8], Anaconda and Python 3.5 softwares were used to test the functionality and Viola-Jones and haar cascade algorithms were used for face detection. Similarly, KDEF (Karolinska Directed Emotional Faces) dataset and VGG (Visual Geometry Group) 16 were used with CNN (Convolution Neural Network) model which was designed with an accuracy of 88%, for face recognition and classification that validated the performance measures. However, the results proved that the network architecture designed had better advancements than existing algorithms.

Another system [9] used Python 2.7, OpenSource Computer Vision Library (OpenCV) & CK (CohnKanade) and CK+ (Extended Cohn-Kanade) database which gave approximately 83% accuracy. Certain researchers have described the Extended Cohn-Kanade (CK+) database for those wanting to prototype and benchmark systems for automatic facial expression detection. The popularity and ease of access for the original Cohn-Kanade dataset this is seen as a very valuable addition to the already existing corpora. It was also stated that for a fully automatic system to be robust for all expressions in a myriad of realistic scenarios, more data is required. For this to occur very large reliably coded datasets across a wide array of visual variabilities are required (at least 5 to 10k examples for each action) which would require a collaborative research effort from various institutions

III. METHODOLOGY

Emotion Based Music Recommendation System is Web Application that focuses on emotion detection. It is a prototype of a new product that comprises following modules:

A) Face Detection

Face detection is one of the applications which is considered under computer vision technology. This is the process in which algorithms are developed and trained to properly locate faces or objects in object detection or related system inimages. This detection can be real-time from a video frame or images. Face detection uses such classifiers, which are algorithms that detect what's either a face (1) or not a face (0) in an image. Classifiers are trained to detect faces using numbers of images to get more accuracy. OpenCV uses Haar Cascades. A Haar classifier is used for face detection where the classifier is trained with predefined varying face data which enables it to detect different faces accurately. The main aim of face detection is to spot the face within the frame by reducing external noises and other factors. It is a machine learning-based approach where the cascade function is trained with a group of input files.

B) Feature Extraction

While performing feature extraction, we treat the pre-trained network that is a sequential model as an arbitrary

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 5, pp 273-277, May-2023

https://doi.org/10.47001/IRJIET/2023.705036

feature extractor. Allowing the input image to pass on it forward, stopping at the pre-specified layer, and taking the outputs of that layer as our features. Starting layers of a convolutional network extract high-level features from the taken image, so use only a few filters. As we make further deeper layers, we increase the number of the filters to twice or thrice the dimension of the filter of the previous layer. Filters of the deeper layers gain more features but are computationally very intensive. Doing this we utilized the robust, discriminative features learned by the Convolution neural network. The outputs of the model are going to be feature maps, which are an intermediate representation for all layers after the very first layer. Load the input image for which we want to view the Feature map to know which features were prominent to classify the image. Feature maps are obtained by applying Filters or Feature detectors to the input image or the feature map output of the prior layers. Feature map visualization will provide insight into the interior representations for specific input for each of the Convolutional layers within the model.

C) Emotion Detection

Convolution neural network Architecture. Convolution neural network architecture applies filters or feature detectors to the input image to get the feature maps or activation maps. Feature detectors or filters help in identifying various features present in the image such as edges, vertical lines, horizontal lines, bends, etc. After that pooling is applied over the feature maps for invariance to translation. Pooling is predicted on the concept that once we change the input by a touch amount, the pooled outputs don't change. We can use any of the pooling from min, average, or max. But max-pooling provides better performance than min or average pooling. Flatten all the input and giving these flattened inputs to a deep neural network which are outputs to the class of the object. The class of the image will be binary, or it will be a multi-class classification for identifying digits or separating various apparel items. Neural networks are as a black box, and learned features in a Neural Network are not interpretable. So basically, we give an input image then the CNN model returns the results. Emotion detection is performed by loading the model which is trained by weights using CNN. When we take the real-time image by a user then that image was sent to the pre-trained CNN model, then predict the emotion and adds the label to the image.

D) Music Recommendation

Songs Database We created a database for Bollywood Hindi songs. It consists of 100 to 150 songs per emotion. As we all know music is undoubtedly involved in enhancing our mood. So, suppose a user is sad then the system will recommend such a music playlist which motivates him or her and by this automatic mood will be delighted. Music Playlist Recommendation. By using the emotion module, real-time emotion of the user is detected. This will give the labels like Happy, Sad, Angry, Surprise, and Neutral. Using the os.listdir() method in python we connected these labels with the folders of the songs database which we have created. This will result in the recommended playlist for the user in the GUI of the music player by showing captions according to detected emotions. We have used a library called Pygame for playing the audio as this library supports playing various multimedia formats like audio, video, etc.

IV. RESULTS AND DISCUSSIONS

The system detects the mood in real time and a playlist is displayed for that emotion accurately. It is able to fetch and play the songs from the recommended playlist in the presence of a stable internet connection. The playlist displays the songs to the user in list view, and when the user selects a song to be played, the song is played in a media player, that comes with all the necessary buttons such as play, pause, shuffle, next and previous, along with a seek bar. For "angry", "Fear", "disgust" and "surprise" moods, devotional, motivational and patriotic songs are suggested to the user. Hence, the user is also provided with mood improvement.

A) Main Home Page

Figure 1: Home Page

B) Registration Page

Figure 2: Registration Page

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 5, pp 273-277, May-2023

https://doi.org/10.47001/IRJIET/2023.705036

C) Login Page

Figure 3: Login Page

D) User Home Page

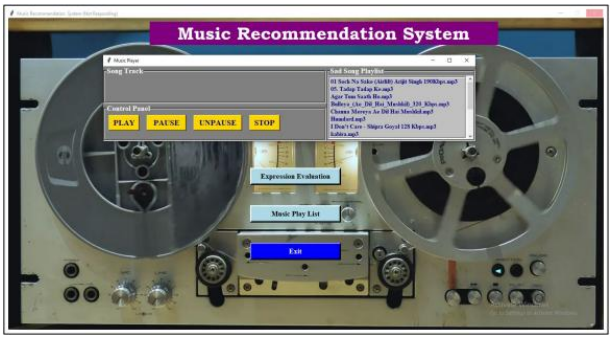

Figure 4: Result 1

E) Result

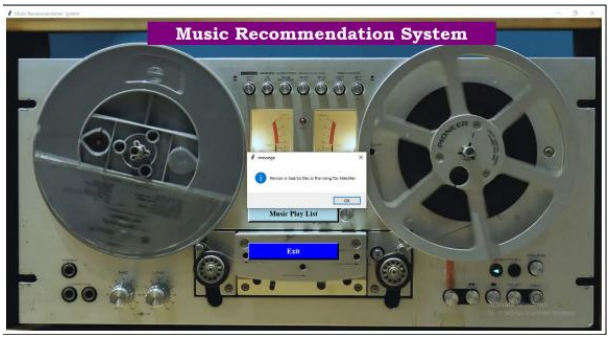

Figure 5: Result 2

V. CONCLUSION

The proposed work presents facial expression recognition system to play a song according to the expression detected and also classify music Type. It uses Haar Classifier for Face detection and CNN approach to extract features, and emotion detection. In this work, real images i.e. user dependent images are captured utilizing the in-built camera.

REFERENCES

[1] Raut, Nitisha, "Facial Emotion Recognition Using Machine Learning" (2018). Master's Projects. 632. https://doi.org/10.31979/etd.w5fs-s8wd

[2] Hemanth P,Adarsh ,Aswani C.B, Ajith P, Veena A Kumar, "EMO PLAYER: Emotion Based Music Player", International Research Journal of Engineering and Technology (IRJET), vol. 5, no. 4, April 2018, pp. 4822-87.

[3] Music Recommendation System: "Sound Tree", Dcengo Unchained: Sıla KAYA, BSc.; Duygu KABAKCI, BSc.; Işınsu KATIRCIOĞLU, BSc. and Koray KOCAKAYA BSc. Assistant : Dilek Önal Supervisors: Prof. Dr. İsmail Hakkı Toroslu, Prof. Dr. Veysi İşler Sponsor Company: ARGEDOR.

[4] Tim Spittle, lucyd, GitHub, , April 16, 2020. Accessed on: [Online], Available at: https://github.com/timspit/lucyd

[5] A.Abdul, J. Chen, H.-Y. Liao, and S.-H. Chang, "An Emotion-Aware Personalized Music Recommendation System Using a Convolutional Neural Networks Approach," Applied Sciences, vol. 8, no. 7, p. 1103, Jul. 2018.

[6] Manas Sambare, FER2013 Dataset, Kaggle, July 19, 2020. Accessed on: September 9, 2020. [Online], Avalilable at: https://www.kaggle.com/msambare/fer2013

[7] Mahmoudi MA, MMA Facial Expression Dataset, Kaggle, June 6, 2020. Accessed on: September 15, 2020. [Online], Available at: https://www.kaggle.com/mahmoudima/mma-facial-expression

[8] Dr. Shaik Asif Hussain and Ahlam Salim Abdallah Al Balushi, "A real time face emotion classification and recognition using deep learning model", 2020 Journal. of Phys.: Conf. Ser. 1432 012087.

[9] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar and I. Matthews, "The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression," 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, San Francisco, CA, USA, 2010, pp. 94-101, doi: 10.1109/CVPRW.2010.5543262.

[10] Puri, Raghav & Gupta, Archit & Sikri, Manas & Tiwari, Mohit & Pathak, Nitish & Goel, Shivendra. (2020). Emotion Detection using Image Processing in Python.

[11] Patra, Braja & Das, Dipankar & Bandyopadhyay, Sivaji. (2013). Automatic Music Mood Classification of Hindi Songs.

[12] Lee, J., Yoon, K., Jang, D., Jang, S., Shin, S., & Kim, J. (2018). MUSIC RECOMMENDATION SYSTEM BASED ON GENRE DISTANCE AND USER PREFERENCE CLASSIFICATION.

ISSN (online): 2581-3048
Volume 7, Issue 5, pp 273-277, May-2023
https://doi.org/10.47001/IRJIET/2023.705036

[13] Kaufman Jaime C., University of North Florida, "A Hybrid Approach to Music Recommendation: Exploiting Collaborative Music Tags and Acoustic Features", UNF Digital Commons, 2014.

[14] D Priya, Face Detection, Recognition and Emotion Detection in 8 lines of code! Towards data science, April 3, 2019. Accessed on: July 12, 2020 [Online], Available at: https://towardsdatascience.com/facedetection-recognition-and-emotion-detection-in-8-lines-of-codeb2ce32d4d5de

[15] bluepi, "\Classifying Different Types of Recommender Systems, November 14, 2015. Accessed on: July 7, 2020. [Online], Available on: https://www.bluepiit.com/blog/classifying-recommendersystems/#:~:text=There%20are%20majorly%20six%20types,system%20and%20Hybrid%20recommender%20system

AUTHORS BIOGRAPHY

Aman Tamboli, Student, Computer Engineering, KJ College of Engineering and Management Research, Pune, Maharashtra, India.

Harshwardhan Almnur, Student, Computer Engineering, KJ College of Engineering and Management Research, Pune, Maharashtra, India.

Vivek Patil, Student, Computer Engineering, KJ College of Engineering and Management Research, Pune, Maharashtra, India.

Uday Patil, Student, Computer Engineering, KJ College of Engineering and Management Research, Pune, Maharashtra, India.

Citation of this Article:

Aman Tamboli, Harshwardhan Almnur, Vivek Patil, Uday Patil, Sheetal Nirve, "Music Recommendation System" Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 7, Issue 5, pp 273-277, May 2023. https://doi.org/10.47001/IRJIET/2023.705036

****

