IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 11, Issue 5, May 2023

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105015 |

Emotion Based Music Playlist Using CNN

Dr. K. Mariyappan[1], Dhruv Sharma[2], Jayant Gadhia[3], Anika Shrivastava[4], Kabita Gautam[5]

[1]Associate Professor, Faculty of Engineering and Technology, Jain University, Bangalore, India

UG Student, Department of Information Science and Engineering, Faculty of Engineering and Technology, Jain University, Bangalore, India[2,3,4,5]

ABSTRACT: Music can regenerate brain cells and ease emotions, when users are in a negative mood. As the world begins to corrode with mental disorders, Music Therapy has been introduced in hospitals to help the patients improve their mental health and overall, well-being. Life is not a plateau, the ups and downs we face every now and then is not well accepted by many individuals The proposed Project plays an effective role in generating a music playlist based on the user`s current mood. As we know, music can reach those corners of the heart and mind that no therapist or beloved one can. In this project, facial expressions are captured through webcam and such expressions are fed into Deep learning algorithm which gives most probable emotion. Along with CNN, methodologies for capturing and detecting emotions, we will be using Deepfake for a clean emotion detection. Further evolving it by inserting a pre-determined set of music based on the mood for generating the playlist with respect to the emotion captured. With an innovative way of utilizing and implementing music in one`s life, we can better the mental health of many.

KEYWORDS: Facial expression, Deep learning, CNN, Deep Face, Pygame

I. INTRODUCTION

Music has an ability to lighten moods and uplift one's confidence. Under normal circumstances, people have to browse through songs or manually search for songs according to their liking and mood which at times can just become an annoying task. Facial expressions are best way of expressing mood of a person. They are produced by the often-involuntary manipulation of muscles in the face in a manner and fashion that could change the position and shape-both absolute and relative of facial features such as eyes, eye lids, eye lashes, nose, lips, cheek muscles, etc.

This oftentimes creates wrinkles and bulges in the face typically conveying a user's emotional state. This project includes recognizing of the facial expressions of a user and play songs according to their emotion. The facial expressions are captured using a webcam and face detection. Which is further fed into Deep Learning Algorithm which defines the emotion of the user. Deepfake will be used to detect a clean emotion while CNN, methodology will be used to capture and detect emotions.

II. RELATED WORK

A) 2019 29th International Telecommunication Networks and Applications Conference (ITNAC)
DaiLun Chiang, JihHsiang Yang, ZiYuan Huang, FeiPei Lai
Music Response Based on Real-time Facial Expression Recognition
Information and communication technology wanted to develop an App that plays music lists based on emotion [1]. Their setup had a personalized multimedia interactive platform and an operational structure that helped them assist the elderly in "independent living" and "social participation" through technology. In this regard, an instant and interactive home companion model was built to enable care recipients to maintain an independent lifestyle.
Methodology: The camera obtains users' facial expression; the image is transmitted to the iPad's Home App automatically. They built an App to record images obtained by the home app, transfer the image to the desktop computer, and use the CNN model to compute expressions in the image. The Erica emotion App was implemented on Swift for the iOS platform and had a quite simple interface. When playing proper music as a response, playlists are divided into negative and happy ones. If the expression result is anger, disgust, scared, or sad, then a negative emotional song will be played, if happy, a positive emotional song list will be played. The dataset was from Kaggle. In the model the size of each photo has been adjusted to 48*48-pixel values in the data pre-processing stage. In the second model all the disgust and fear data are deleted in the test and training data. The third version is based on the results of the second version plus some photos downloaded from the Internet, so it improves accuracy.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105015 |

Results: The accuracy of each expression in third dataset is: 84% for happiness; 83% for surprise, 58% for angry, 55% for sadness and 56% for neutral.
Gap: It was that they were not able to detect the emotions very accurately because of the fingers in the images and side face of the person, thus made them remove few of the images as well as few emotions.

B) International Journal for Research in Applied Science & Engineering Technology, 2020
Prof. Jaychand Upadhya, Sharan Shetty, Vaibhav Murari, Jarvis Trinidade
Mood Based Music Playlist Generator Using Convolutional Neural Network
The main motive was to create a Music Playlist Generator basically by detecting the mood of the user just by their facial emotion [2]. They have solved the issue of the manual finding of songs to suit one's mood along with creating a high accuracy CNN model for Facial emotion recognition. Through the webcam, the emotional state had been deduced from facial expressions. They have also created a neural network model; the CNN classifier was used.
Methodology: Data Processing- Convolutional Neural Networks was used for creating the model to predict the mood/emotion of the user. The model was trained on the FER 2013 Dataset. Model Training and Hyperparameter Tuning- 80% of the data was set aside for training, 10% for validation, and 10% for testing. CNN architecture was with 6 Convolutional layers along with a Maxpooling layer and finally the flatten and dense layers. They used 'relu' as the activation function and for the dense layer, 'softmax' was used. Batch Normalization- It is the technique of adding extra layers to a deep neural network to make it faster and more stable.
Results: To test the overall performance, the proposed model was trained for 100 epochs on the FER2013 database. The model obtained an accuracy of 88.3% with a 0.9 loss on the validation data and 97.42% accuracy with 0.09 loss on the training data from FER2013.
Gaps: The proposed system can be enhanced by integrating audio emotion recognition to facial emotion recognition to create an even more effective model. The current system only recommends songs in English; however, the system can be enhanced by detecting the language spoken by the user and accordingly generating songs in that particular language. The current system requires manually mapping songs to their moods, audio emotion recognition can be used to automate this task and map the songs easily. One can further create an entire user account system allowing users to save the recommended songs to their account for future use.

C) 2021 International Conference on Artificial Intelligence and Machine Vision(AIMV)
Kevin Patel, Rajeev Kumar Gupta
Song Playlist Generator System Based on Facial Expression and Song Mood
This paper proposed a Deep Learning based approach for the playlist generation based on human current mood with the help of user's past history of song selection [3]. First was user mood prediction using his/her face. Second was song mood detection using neural network as function of classifier KerasClassifier, so that user can experience personalize recommendation based on his/her mood and his/her listened history.
Methodology: Face-Detection- They used camera with the help of OpenCV library to detect the face. Emotion Detection- They have used Convolutional Neural Network (CNN) for emotion detection task. They have collected users` past listened songs by using Spotify API, past 50 listened songs were collected. They have proposed song mood detection methodology with KerasClassifier and separated the whole list into different lists based on song moods. Later concat different mood specific list based on facial expression.
Results: They classified the facial expression into 4 moods and got testing accuracy of 83% with loss of 0.4532. They have achieved 83% of accuracy with 4 classes those are: Calm song, Happy song, energetic song, Sad Song classification.
Gaps: Finding best path of song mood that need to occur according to user mood by evaluating more user data and activity. There are many aspects like emotion detection using surrounding audio feature processing, evaluation of the system based on skipping ratio of songs and many other factors.

D) 2018 International Conference on Mechatronics, Electronics and Automotive Engineering (ICMEAE)
Sergio A. Navarro-Tuch, Rodrigo Solís-Torres, Rogelio Bustamante-Bello, Ariel A. López-Aguilar, Guillermo Gonzales-Archundia, Omar Hernández-Gonzales
Variation of facial expression produced by acoustic stimuli
They have used cell phone camera to detect the facial expression from the user and propose a list of personalized songs in accordance with the emotion detected [4]. We can see that there is a different response between synthetic audio and the ones generated by human beings, the standard deviation values gave out a low deviation indicating that the results obtained are reliable, thus we can see that there was a change of emotion when an audio generated by humans was played.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105015 |

Methodology: Through the data obtained using iMotions, the analyzed data was processed using MATLAB and Excel. MATLAB was used to analyze the raw data using the data analysis tools, in order to determine the impact of artificially synthesized versus human composed audios. The data was obtained from three experiments, each experiment had different conditions, each experiment had different number of participants, the length and type of audio segment changed for each experiment. The data obtained from these three experiments were analyzed using FACET of iMotions. For the next stage we analyzed Neutral Emotion, the purpose of this analysis is of outmost importance for future stages for a proper development of a recurrent neural network to develop an algorithm to create music sequences.
Result: All the data was gathered, mean and standard deviation for each subject, to graph the results of mean vs. standard deviation. At the end all the mean data was gathered for each subject with their respective stimulus, MATLAB was used to generate a matrix of the mean data, we obtained the derivative with respect to time of the matrix, then the result was normalized. The new dispersion graphs were obtained from matrix's data of the normalized means.
Gap: The synthetic audio and the audio produced by the algorithm didn't have a higher probability of showing a change in Neutral Emotion this is because there is no pattern that the subjects find interesting. This result may lead us to determine the kind of base stimuli for the audio generator to part from. The protocol presented and developed for this work allowed us to find the main problematics and improvement points for the research using audio stimuli.

E) 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT)
Krupa KS, Ambara G, Kartikey Rai, Sahil Choudhury
Emotion aware Smart Music Recommender System using Two Level CNN.
Music plays an important role in elevating one's mood [5]. 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. Emotion recognition is a domain gaining relevance in recent times. Automated decision-making systems can read a person's state of mind and thus recognize one's emotional status to be applied for a variety of applications that include recommendations. The paper work proposes a CNN based approach to recommend music by analysing 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.
Methodology: A. Emotion detection using facial expression - For detecting emotions, FER2013 dataset is used. The facial image of the user is captured using a webcam. Every image is pre-processed and rescaled to an array with 48x48 grey scale values and tested with the testing data. B. Emotion detection using semantic analysis on chatbot interaction – They have used chat box to detect emotions using conversations or voice commands. Emotion recognition in conversation (ERC) is achieved using deep learning-based technique. The chatbot implementation uses Google's Text-to Speech (gTTS) and Speech-to-Text API to make the system interactive. Pydub along with Pyaudio module is used. The chatbot is trained on IEMOCAP dataset and modelled with text using CNN. Detection of emotions from text conversations with chatbot involves a) Emotion Expression and b) Emotion Detection using Deep Learning approach to classify the emotion.CNN is directly applied on these word embeddings without any prior information on their semantic context. A vector of words from IEMOCAP dataset are mapped to the context and used to train our model to predict the words in the response text.
Result: The network is trained for 60 epochs initially and the accuracy comes to be around 63% for 60 epochs. When tested for 100 epochs the accuracy came to be around 88%. Music plays a major role in handling the stressful situations and emotions triggers of the user. it is required to recommend music that suits the current emotional needs of the user. the proposed CNN based model detects the emotion and generate the playlist accordingly.
Gap: The accuracy was seen to be 60% initially and went on to being 88%, which shows the model has a good learning rate. However, the predictions need to be closer to the real values, in order to work effectively. We will be implementing these learnings in our project to get an accuracy of 98% and more.

III. PROPOSED METHODOLOGY

Facial expression: Facial expressions are one of the more important aspects of human communication. The face is responsible for communicating not only thoughts or ideas, but also emotions. What makes the communication of emotions interesting is that it appears as if some of these expressions of emotion (e.g., anger, disgust, fear, happy, sad, surprise, and to a lesser extent contempt, embarrassment, interest, pain, and shame) may be biologically hardwired, and are expressed the same way by all peoples of all cultures. This contrasts with other views that all facial expressions are a product of social learning and culture. Darwin was the first to propose that some of these facial expressions of emotion have their origins in the evolution of the human species. These expressions helped the organism survive because it appeared important to social animals like humans or chimpanzees to express these imminent behaviours implied by the emotions (running away in fear, attack in anger) so they could avoid conflict, danger, or allow approach,

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105015 |

and so forth. However, these emotion expressions are not immune to modification by social learning; different cultures learn different display rules to manage their expressions of emotion. Although the current evidence supports Darwin's basic premise it is not without controversy. Future technological advances will allow facial expression research to expand to address many of the important issues that remain.

Deep Learning: Deep learning is a branch of machine learning which is completely based on artificial neural network as neural network is going to mimic the human brain so deep learning is also a kind of mimic of human brain. In deep learning, we don't need to explicitly program everything. The concept of deep learning is not new. It has been around for a couple of years now. It's on hype nowadays because earlier we did not have that much processing power and a lot of data. As in the last 20 years, the processing power increases exponentially, deep learning and machine learning came in the picture. A formal definition of deep learning is- neurons. Deep Learning is a subset of Machine Learning that is based on artificial neural networks (ANNs) with multiple layers, also known as deep neural networks (DNNs). These neural networks are inspired by the structure and function of the human brain, and they are designed to learn from large amounts of data in an unsupervised or semi-supervised manner.

Convolutional Neural Network: In deep learning convolutional neural network (CNN/ConvNet) is a class of deep neural networks, most commonly applied to analyze visual imagery. Now when we think of a neural network, we think about matrix multiplications but that is not the case with ConvNet. It uses a special technique called Convolution. Now in mathematics convolution is a mathematical operation on two functions that produces a third function that expresses how the shape of one is modified by the other. Convolutional neural networks are composed of multiple layers of artificial neurons. Artificial neurons, a rough imitation of their biological counterparts, are mathematical functions that calculate the weighted sum of multiple inputs and outputs an activation value. When you input an image in a ConvNet, each layer generates several activation functions that are passed on to the next layer. The first layer usually extracts basic features such as horizontal or diagonal edges. This output is passed on to the next layer which detects more complex features such as corners or combinational edges. As we move deeper into the network it can identify even more complex features such as objects, faces, etc.

Pygame: It is a Python wrapper for the SDL Library, which stands for Simple DirectMedia Layer. SDL provides cross-platform access to your system's underlying multimedia hardware components, such as sound, video, mouse, keyboard, and joystick. Pygame started life as a replacement for the stalled PYSDL Project.Pygame is a popular Python library that enables the development of 2D video games and multimedia applications. It provides a simple and intuitive interface for handling graphics, sound, and user input. Pygame is built on top of the Simple DirectMedia Layer (SDL), a powerful cross-platform multimedia library. It abstracts away the complexities of SDL and provides a higher-level API for game development, making it accessible to beginners while still offering flexibility for advanced users.

IV. PSEUDO CODE

1. Import necessary libraries like cv2, DeepFace
2. Initialize Pygame mixer
3. Define emotion playlists
4. Open webcam
5. Start video capture and emotion analysis loop:
   - Read frame from webcam
   - Detect faces in the frame
   - Analyze emotions using DeepFace library
   - Display dominant emotion on the frame
   - Check if 's' key is pressed to capture the emotion:
     - Capture the emotion and check if it has a corresponding playlist
     - Display the available songs in the playlist
     - Ask user to choose a song from the playlist
     - Play the chosen song using Pygame mixer
   - Check if 'q' key is pressed to quit the program
6. Stop and quit Pygame mixer
7. Release the webcam and close all windows

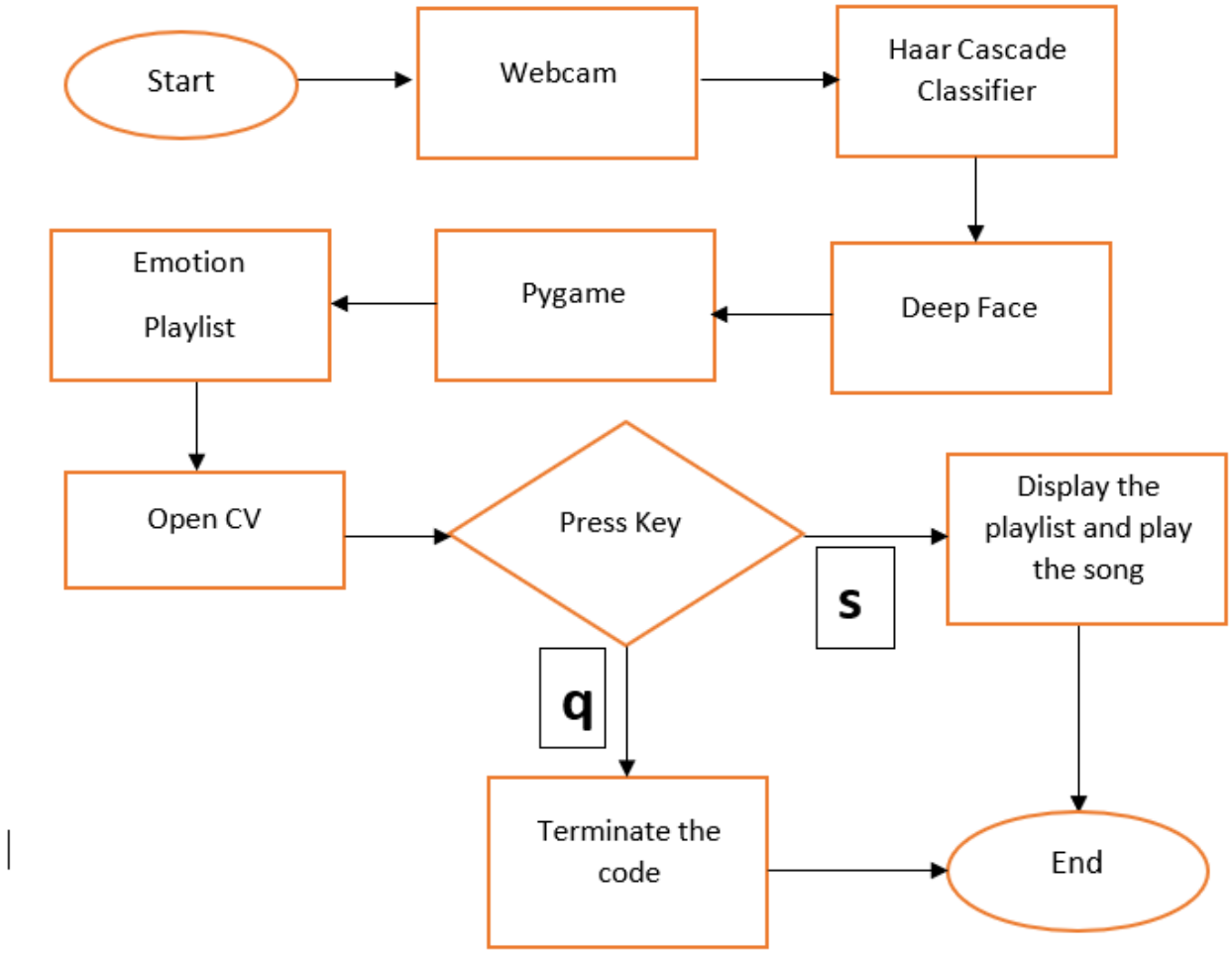
International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105015 |

V. SYSTEM ARCHITECTURE



VI. SIMULATION RESULTS

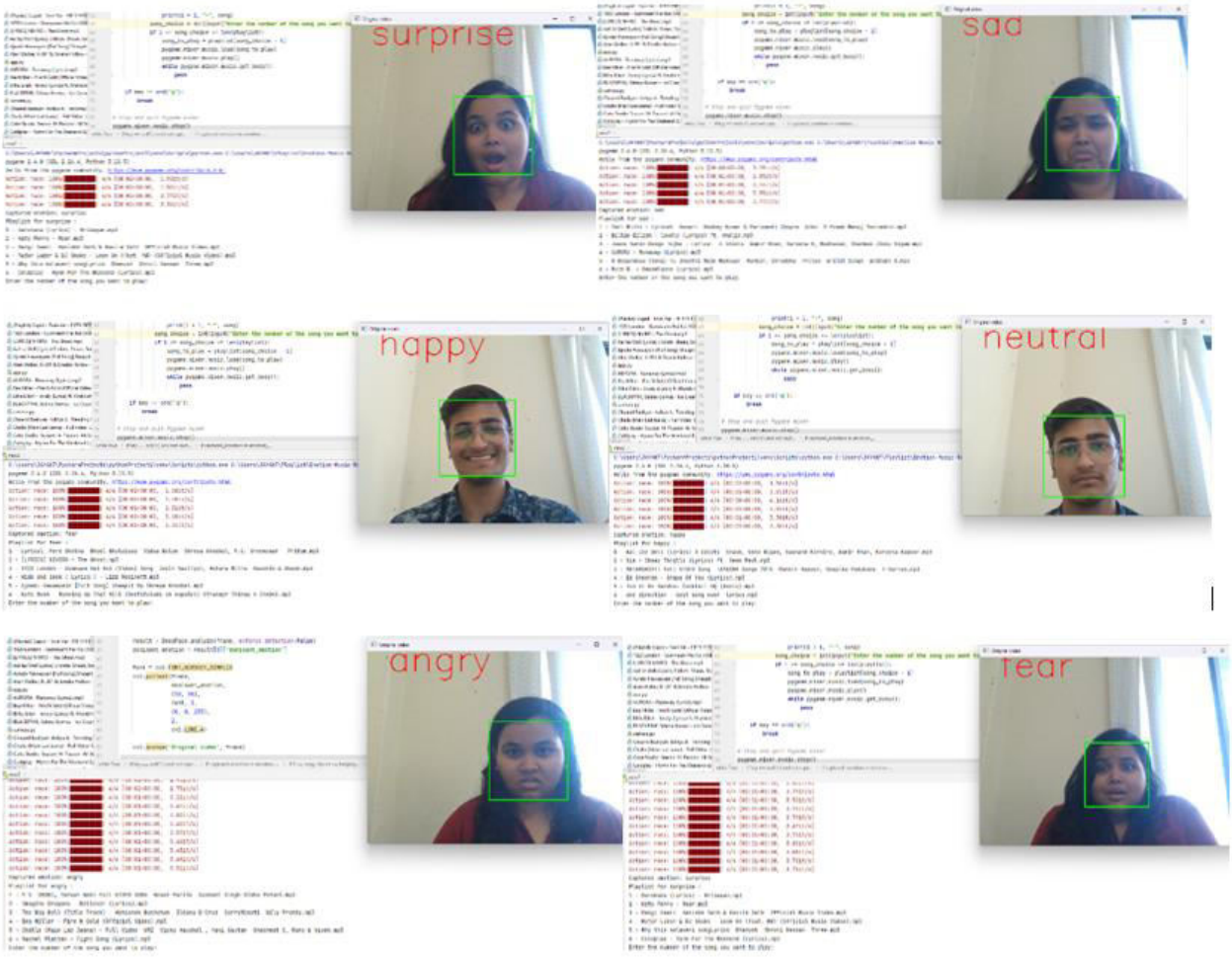
We have used OpenCV to open the camera and capture the video,Pygameto play the audio and get the result. We have trained our model using different expressionsby Deep Face.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105015 |



VII. CONCLUSION AND FUTURE WORK

In today`s world, mental health plays an important role in one`s well-being. Music tends to calm your brain cells. We read a lot of Literature Reviews that helped us get a clear insight of our project. This project that creates a playlist with respect to one`s emotions, will help in soothing one`s mind and even uplift their mood. The simulation results have been achieved by using various libraries and technologies such as OpenCV,Deep Face and Pygame to play music based on the user's emotional state as well as let the user select the song from the list. In the future, we plan on customizing the music recommendations according to the user`s music history by taking it from Spotify. Currently, this project only plays the song but in future we plan to display it onthe web and even develop an application.

REFERENCES

[1] DaiLun Chiang, JihHsiang Yang, ZiYuan Huang, FeiPei Lai, Music Response Based on Real-time Facial Expression Recognition, 2019 29th International Telecommunication Networks and Applications Conference (ITNAC), DOI: 10.1109/ITNAC46935.2019.9078004, 2020.

[2] Prof. Jaychand Upadhya, Sharan Shetty, Vaibhav Murari, Jarvis Trinidade, Mood Based Music Playlist Generator Using Convolutional Neural Network, International Journal for Research in Applied Science & Engineering Technology (IJRASET), ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538, Volume 10 Issue III Mar 2022- Available at www.ijraset.com.

[3] Kevin Patel, Rajeev Kumar Gupta, Song Playlist Generator System Based on Facial Expression and Song Mood, 2021 International Conference on Artificial Intelligence and Machine Vision (AIMV), DOI: 10.1109/AIMV53313.2021.9670976, 2021.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105015 |

[4] K S Krupa; G Ambara; Kartikey Rai; Sahil Choudhury, Emotion aware Smart Music Recommender System using Two Level CNN, 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT), 10.1109/ICSSIT48917.2020.9214164, 2020.
[5] S. Kim and H. Kim, "Deep explanation model for facial expression recognition through facial action coding unit," in 2019 IEEE International Conference on Big Data and Smart Computing (BigComp). IEEE, 2019, pp. 1–4.
[6] A. D. Rasamoelina, F. Adjailia, and P. Sinak, "Deep convolutional neural network for robust facial emotion recognition," in 2019 IEEE International Symposium on INnovations in Intelligent SysTems and Applications (INISTA). IEEE, 2019, pp. 1–6.
[7] Vijayakumar, T. "Posed Inverse Problem Rectification Using Novel Deep Convolutional Neural Network." Journal of Innovative Image Processing (JIIP) 2, no. 03 (2020): 121-127.
[8] Smys, S., Joy Iong Zong Chen, and Subarna Shakya. "Survey on Neural Network Architectures with Deep Learning." Journal of Soft Computing Paradigm (JSCP) 2, no. 03 (2020): 186-194.
[9] B. Bozorgtabar, D. Mahapatra, and J.-P. Thiran, "Exprada: Adversarial domain adaptation for facial expression analysis," Pattern Recognition, vol. 100, p. 107111, 2020.
[10] Sungheetha, Akey, and Rajesh Sharma. "Design an Early Detection and Classification for Diabetic Retinopathy by Deep Feature Extraction based Convolution Neural Network."Journal of Trends in Computer Science and Smart technology (TCSST) 3, no. 02 (2021): 81-94.
[11] Rahul Ravi, S.V Yadhukrishna, Rajalakshmi, Prithvi raj, "A Face Expression Recognition Using CNN& LBP",2020 IEEE.

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor: 8.379

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details