Journal of Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.777518

AN APPROACH TO RECOMMEND MUSIC USING FACIAL EMOTION GESTURES

Ongole Gandhi[1], P V Rama Gopala Rao[2], Venkata Krishnarao L[3], Yakobu Dasari[4], Dr. M. Anusha[5], R Prathap Kumar[6]

[1,4,6] Assistant Professor, VFSTR Deemed to be University, Vadlamudi, A.P., INDIA

[2]Assistant Professor,Teegala Krishna Reddy Engineering College, Hyderabad, Telangana, INDIA

[3]Assistant Professor,PVP Siddhartha Institute of Technology, Vijayawada, A.P., INDIA

[5] Associate Professor, Koneru Lakshmaiah Education Foundation, Vaddeswaram, A.P., INDIA

Abstract- Music plays a vital role in our everyday life. Life without music cannot be imagined. Music changes our mood; Whatever our mood might be, the only thing we do in all of our moods is to listen to music. We also listen to music when working, driving, travelling and even when reading a comic or a story. Music can induce a clear emotional response in its listeners. The pitch and rhythm of the music are managed in the areas of the brain that deal with emotions and mood. Thus, music plays an important role in enhancing our mood. As elders have said "Face is the Index of the Mind", the mood of a person can be known by looking at the face of the person. The abstract of this system/ project is to build an automated system that builds playlists and plays the songs according to the mood of the user by directly discerning the facial emotions of the user. This model requires a camera to capture the face of the user and then the mood of the user is recognized by CNNs. Then the playlist is recommended to the user based on the discerned "Mood" of the user. This disposes of the tedious and monotonous task of physically gathering tunes into various records and helps in creating a suitable playlist dependent on a person's passionate highlights. Hence, the proposed system can be used to build a music recommendation system based on the facial emotion gestures of the user.

Keywords – Mood, Music, CNN, Facial emotion gestures

I.INTRODUCTION

Music prompts a reasonable passionate reaction in itsaudience. Melodic inclinations have been exhibited to be exceptionally associated with character qualities and mind-sets. Facial emotions are the most common and natural methods of passing on feelings, temperaments and sentiments. Convolutional Neural network, as a Deep Learning Neural Network, assumes a critical part in face image recognition. Cognitiontechnology of CNN and Music Recommendation System based on Facial Emotion Gesturesis created to distinguish a model that perceives facial articulations and prescribes music as indicated by comparing mind-set of the user or client.

Human beings have the innate capacity to see somebody's face and conjecture their mind-set. This capacity if learnt by an electronic gadget - computer, humanoid robot or a mobile gadget - can have important applications in reality. Music, an instrument for stirring emotions and feelings, is undeniably more remarkable than language. Music is something which takes

advantage of our emotional centre as human beings [1]. Accordingly, paying attention to good music can assist us with lifting our mind-set from a negative sense to a positive sense. For example, focusing on lively tunes when the individual is feeling grim can assist him with arising his difficulty and start feeling better. This framework proposes one such application, emotion-based music recommendation. Emotion of the client can be effortlessly speculated by taking a gander at his/her face. For this reason, face detection and emotion recognition, examining the fiducial highlights from his/her face is essential.

Fig 1.1 Facial expression with different emotions[2]

The issues related with face detection incorporate foundation components, lighting conditions, posture and facial demeanour. This space of face detection and emotion detection is as of now a functioning space of examination because of advancement of Virtual Reality and Augmented Reality. Constant face detection and recognition frameworks have restricted usefulness because of the fluctuating nature of pictures as a result of the issues related like foundation, enlightenment, and so on Thus, innovative work for arrangements identified with these issues is a continuous work.

Using regular music players, a client expected to actually mastermind his playlist and select tunes that would diminish his/her attitude and energetic experience. This task was work genuine and an individual every now and again went up against the trouble of showing up at an appropriate once-over of songs. Different frameworks which recognize the disposition of the client by utilizing facial appearance have their time and memory intricacy generally high and subsequently flop in accomplishing an ongoing presentation. Regardless of whether they perceive the temperament of the client then their choice of melodies for making a playlist is with the end goal that it will simply pick tunes mirroring the current mind-set of the client and won't attempt to improve his mind-set in any capacity. In this way, if the client is dismal, In the current frameworks, client is furnished with a rundown of melodies with pitiful emotion which can corrupt his/her mind-set further and can prompt misery.

Along these lines, the framework proposed will distinguish the emotion of the client from his facial articulations. It will then, at that point furnish the client with a playlist of melodies, paying attention to which the client will feel good.

Fig 1.2 Facial expression with 7 different emotions [3]

A. Convolutional Neural Network

A CNN gets a picture as a contribution to the type of a 3D Matrix. The underlying two measurements contrast with the width and height of the image in pixels while the third one identifies with the RGB potential gains of each pixel. CNNs comprises of the accompanying successive modules (every one may contain more than one layer)Convolution

- ReLu activation function
- Pooling
- Fully connected layers
- Output layer

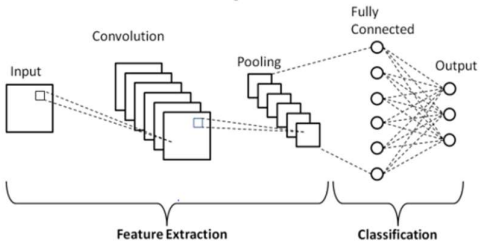

Fig 1.3 Schematic representation of Basic Convolutional Neural Network[4,5]

B. Convolution Layer :

he part connected with doing the convolution movement in the underlying portion of a Convolutional Layer is known as the Kernel/Channel. Convolution activity is a component savvy network increase activity. Convolutional layers take the three-dimensional information framework and they pass a channel (otherwise called convolutional channel) over the image, applying that to a little window of pixels at the same time (for instance , 3x3 pixels) and this window, being moved until the entire picture has been separated. The convolutional action registers the dab consequence of the pixel regards in the current channel window close by the heaps described in the channel. The yield of this movement is the last tangled picture. The focal point of picture request CNN's is that as the model trains what it really does is that it learns the characteristics for the channel matrices that enable it to remove huge features (shapes, surfaces, concealed districts, etc) in the image. Each convolutional layer applies one new channel to the tangled image of the past layer that can eliminate one more part. Accordingly, as more channels are stacked, the more features the CNN can remove from an image.

The three components that go into the convolution activity are:

- Input image
- Feature detector
- Feature map

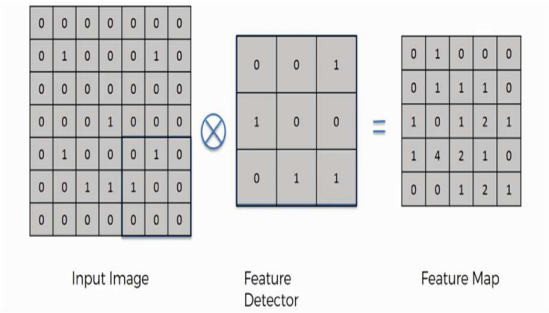

Fig 1.4 Feature Map generation through convolution operation[5]

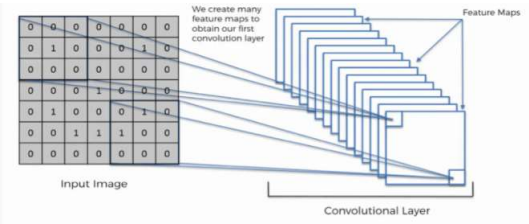

Fig 1.5 Creation of Convolution Layer[6,7]

C. **ReLu Layer:**

After every convolution activity, CNN applies a Rectified Linear Unit (ReLu) function to the yield of the convolved picture. If the convolved image has negative values, it replaces them with '0'.It also introduces nonlinearity into the model.

D. Pooling Layer:

Pooling is the interaction where measurement of the convolved picture is decreased It does as such to diminish handling time and the registering power required. During this cycle, it ensures the fundamental component information. There are a couple of procedures that can be used for pooling. The most generally perceived ones are Max pooling and Typical pooling. In our application, we will use max pooling as it is the best an enormous segment of the events. Max pooling is fundamentally equivalent to the convolution cycle. A window slides over the component guide and thinks tiles of a predefined size. For each tile, max pooling picks the greatest worth and adds it to another component map. In this manner, the face highlights are separated utilizing convolution and pooling layers.

Fig 1.6 Maximum Pooled Feature Map[6]

E. Fully connected layer:

Fully connected layers are a fundamental segment of Convolutional Neural Networks (CNNs), which have been demonstrated fruitful in perceiving and ordering pictures for computer vision. The CNN cycle starts with convolution and pooling, separating the picture into highlights, and investigating them freely. The consequence of this interaction takes care of into a fully connected neural organization structure that drives the last arrangement choice.

In the Fully Connected Layer, all neurons of one layer are connected to all neurons in the following layer.

F. Output Layer:

The last fully connected layer is the yield layer which applies a SoftMax capacity to the yield of the past fully connected layer and returns a likelihood for each class.

II. LITERATURE SURVEY

R.L.Rosa et al[8]. classified the music through sentences posted in various social media networks they did not classify through sentiments of users they classified through informal text which are posted in social media network where the text describes the emotions of a person whether he is happy, sad, or neutral through these sentence sentiment intensity analysis is done and using a new lexicon-based sentiment measure named it as "Enhanced Sentiment Metric(ESM)".Sentiment analysis is a technique of natural language processing and text analysis these techniques are applicable to various fields like e-commerce, multimedia learning, etc, this analysis research is based on semantics, tag-based extractions, SVM. These techniques are used to describe emotions based on words and differentiating the emotions related to words. Here the sentiment analysis is divided into three. 1.lexicon-based 2. corpus-based 3. hybrid based. Lexicon based approach uses a word dictionary by using these words positive scale words and negative scale words are divided and these are represented using +, - symbols. The corpus-based approach uses a machine learning method where a large amount of data is collected

Xuan Zhu et al[9]. came up with an integrated music recommendation system technique that consists of automatic music genre classification, music emotion classification, and music similarity query with the help of AdaBoost algorithm, they proposed a new tempo feature called (LMFC) and it comes with timbre features which improve the efficiency of music classification here in this process where an interface is implemented where the user needs to press two buttons and the music will be played according to his genre selection of music there is also another feature where there will be user-selected songs where user can play a song on his selection the main goal of this process is to divide the large music databases into separate selections according to the specified genre

Jukka Holm et al[10]. proposed a technique where music will be differentiated according to pictures here where the user interacts with pictures where they collect pictures according to emotions like happy, sad, etc where by using these pictures music will be recommended where 40 Finnish people used this and it turns out to be 85% accuracy is showed on music selection there are few drawbacks where sometimes pictures selection will be failed and music recommendation to few pictures were also failed. Where they developed a hi-fi prototype application for a touch screen pc was developed and with the help of Nokia Research Centre's Super Music service music will be streamed.

Sanghoon jun et al[11]. focused on perceiving music emotion dependent on abstract human emotions they acquainted three stages with accomplish this. Various sorts of music highlights are removed from the music signal. Those music highlights are noticed and planned into certain qualities knows as (AV esteems) by a fluffy surmising motor. At last, emotion will be shown dependent on AV esteems. Later introduced a feedback method where user can give feedback

where they can rectify the flaws in the interface. There are few drawbacks where the user interface for expressing emotion is hard.

Ziyang Yu et al[12]. proposed that As there is the rapid growth of big data, deep learning with the help of these technologies where they combined micro-expression recognition technology of convolutional neural network and automatic music recommendation algorithm by using this technology where music will be suggested automatically by recognizing the emotion of a person there are few basic steps to achieve these firstly pictures need to be collected of all expressions later pre-processing of the images should be done and finally, classification should be done after implementing this there was 62.1% of recognition rate here the recognition will not be done by past listening history it will be on the user's present emotions after achieving this there are few problems to be solved sometimes the accuracy of the image processing is slower and music recommendation algorithm should be improved for more accuracy.

Kyoungro Yoon et al[13]. suggested a self-sufficient and versatile recommendation framework that depends on the client's disposition and verifiable input to suggest tunes with no earlier information about the client inclinations. This strategy assembles self-ruling a dormant factor model from the accessible online information of numerous clients dependent on the affiliations extricated between the client, tune, client disposition, and tune emotion. It utilizes a blend of the Support Learning structure and Page-Hinkley test to customize the overall melody map for every mind-set by client understood prize.

Aurobind V.Iyer et al[14]. From the Emotion Based Mood Enhancing Music Recommendation system, stated the use of a different approach rather than asking the users mood from the application, the use of the photograph of the user which is sent through Fishers Face algorithm which recognizes the emotion of the user and then selects playlists from the servers.

Qing-Qiang Liu et al[15]. stated that the centre of the recommendation structure is the development of the item versus client emotion model by two-dimensional cover spaces, which assumes a significant part in passing on emotions in items. This technique explores the item include extraction and proposes some connected coordinating with calculations for the development of the item versus client emotion model.

Yu-Hao Chin et al[16]. pitched the idea of an emotion profile-based music recommendation framework. Inside the arranged algorithmic guideline, two inclination profiles region unit made misuse call worth in help vector machine (SVM) and upheld present moment and future alternatives severally. The perceived emotion, feeling profile, and private chronicled question results region unit took care of into the prescribed module to think of the outcome music list.

Chaima Dhahri et al[17]. drafted a self-ruling and versatile recommendation framework that depends on the client's mind-set and implied criticism to suggest tunes with no earlier information about the client inclinations. Our technique assembles independently an idle factor model from the accessible online information of numerous clients dependent on the affiliations removed between the client, melody, client state of mind, and tune emotion. It utilizes a blend of the Support Learning structure and Page-Hinkley test to customize the overall tune map for every disposition as indicated by client implied reward.

III. PROPOSED METHOD

The proposed system contains three modules namely Data Augmentation, Model Training&Testing, Face Detection & Emotion Recognition and Music Recommendation.

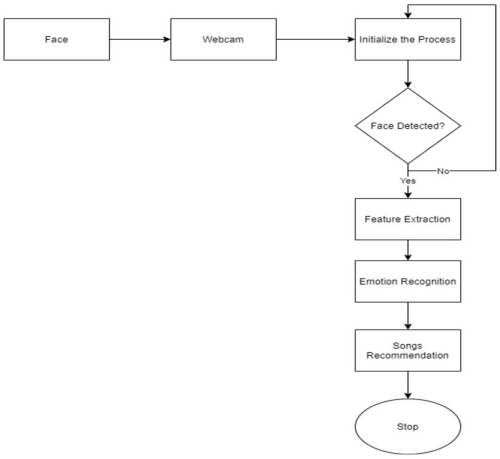

Fig. 3.1. Flow Diagram of the Proposed System

Modules:

A. Data Augmentation
B. Model Training & Testing
C. Face Detection &Emotion Recognition
D. Music Recommendation

A. Data Augmentation:

Data augmentation in data analysis are strategies used to expand the measure of data by adding somewhat altered duplicates of previously existing data or recently made engineered data from existing data. The strategies, for example, cropping, padding, and level flipping are ordinarily used to prepare enormous neural networks. It goes about as a regularizer and decreases overfitting when training a machine learning model.

B. Model Training & Testing:

DeepFace is the facial recognition framework utilized by Facebook for labelling pictures. It was created by scientists at Facebook AI Research (FAIR) at the 2014 IEEE Computer Vision and Pattern Recognition Conference (CVPR)[18].

In current DeepFace recognition there are 4 stages:

- Detect,
- Alignment,
- Represent,
- Classify

Lately, an enormous number of photographs have been slithered via web search tools, and transferred to informal organization by using this enormous data

The DeepFace framework manages a face check estimation, coordinated by man-made intellectual competence (AI) systems using neural association models. A neural organization is a collection of arranged neurons that perform various activities and gives potential results dependent on the emotion of a human being

As Facebook utilizes an undeniable level interpretation of this strategy, the methods are a touch more created and clarified than these. Adding the 3D change and piece-wise relative change in the framework, the estimation is locked in for passing on more exact results

The in-depth level DeepFace recognition fills in as underWorking of DeepFace Recognition

i. Storing of the images to process

At the point at the point when we transfer pictures to the structure, the DeepFace apparatus checks the image of the human face. Whether or not you are transferring two photos of comparative individuals, taken from different focuses, those are a comparable thing for individuals yet not for the PC. It would be treated as two individuals by figuring estimations.

ii. Matching

Then, it starts the 'Matching' measure. This face recognizing confirmation is finished by making and sending a sign to the Neural organization, a nine-layer neuron network having a wide assortment of human faces approx. 120 million.

iii. Using of Neural Network for Feature Detection

By tolerating the sign, fake neurons develop association among neurotransmitters All of the relationship between two neurons can be considered as a course. These courses pass the sign ahead with one small step at a time to recognize the transferred picture . These neurons check with all current pictures to pass the sign the wired way (picture is as of now existing or coordinates with any face) or awful lighting and demonstrates as that image isn't coordinating for certain different faces. In the chief case, the sign continues with the course of the current face and stores the image in the past collection by means of milestone testing, however in the ensuing case, The sign diverts its course on different results like if the picture comprises of huge eyes, it checks whether it had enormous eyes or little eyes. Expect the image is having enormous eyes, the sign starts to go towards all gigantic eyes' photos. Then further it can perceive whether the eyes are dull or gritty shaded. Accept it is dull, and a while later it can make further capabilities like the distance between eyes, shades of eyebrows, sorts of lips, nose, etc

iv. 68 Fiducial point testing

By following rehashing ID courses utilizing stimulated consciousness (AI), the DeepFace goes for 68 fiducial focuses testing, every human face is having 68 explicit facial focuses that can make a match

v. Encoding & Mapping

The apparatus checks design by planning with the testing picture with various photos of an equivalent individual and another person. Then it encodes the right picture. Later by encoding, it glances through the information like name and address of that encoding and saves the picture in that assortment.

vi. Working of each stage in DeepFace Recognition

Initial stage is detect where it detects the image or face of the image to put in simple words where it captures the image and later alignment is done the main process happens in the alignment stage, Here the face I.e., detected will be transformed into 3D image, the main objective of this alignment is

The target of this game plan part is to create frontal face from the info picture that may contain faces from different stance and bearings. The strategy proposed in this paper used 3D

frontalization of faces subject to the fiducial (face include focuses) separate the frontal face[16]. The whole cycle is appeared beneath.

Given an info picture, we at first recognize the face using six fiducial focuses. These six fiducial focuses are 2 eyes, tip of the nose and 3 focuses on the lips[18].These component focuses are used to recognize faces in the image. In this movement we make the 2D-face picture managed from the main picture using 6 fiducial core interests.

In the third step, we apply the 67 fiducial point map with their relating Delaunay Triangulation on the 2D-adjusted altered picture. This movement is done to change the out of plane turns. In this movement, we besides make a 3D-model utilizing a traditional 2D to 3D model generator and plot 67 fiducial spotlights on that really.

The last stage is frontalization of arrangement. we add the lingering part to x-y bearings of 3D twist since it lessens debasement in 3D-twist. Finally, frontalization is refined by doing piece-wise relative on Delaunay triangulation that we made on 67-fiducial focuses[18].

Any simulated intelligence or deep learning system needs adequate training information so it can 'learn'. With a gigantic customer base, Facebook has enough pictures to investigate various roads in regards to. The gathering used more than 4 million facial pictures of more than 4000 people therefore. This computation plays out a lot of exercises for perceiving faces with human accuracy levels.

The final result is a face portrayal which is created from a 9-layer profound neural net. This neural net has more than 120 million factors, which are planned to various privately connected layers. Furthermore, the end-product is Facebook can distinguish if the two pictures address a comparative individual. The site can do it, autonomous of ecological factors light, camera point, and shadings wearing on face for instance facial make-up. Amazingly, this computation works with 97.47 percent precision, which is basically comparable to regular eye accuracy 97.65 percent.

C. Face Detection& Emotion Recognition:

In this module, we use load model library form keras.models. We first capture the user face through web camera using cv2 library and read a frame from that video. Now we use Cascade Classifier to detect the face in the image(frame) detected and we crop the face and store it another object. We convert this image to grey color and resize this image to 48*48-pixel image as the model is trained on the same size of images which is 48*48.Now,we convert this image to array using img_to_array function imported from keras.preprocessing library. Now we use the load_model function to load the model that is already saved. Since we use 'SoftMax' function, it gives the probability for each class, so among them we take the maximum probability and we mark it as the emotion of the user. If the face is not found in the window, then "Face Not Detected' message appears. Thus, emotion of the user is identified.

D. Music Recommendation:

First, we create a folder named music system under this folder we create a folder called 'Music', in this folder again we create 5 more folders which corresponds to each of the 5 emotions. The folder names are Angry, Happy, Sad, Fear, Neutral.In each of these 5 folders, we play corresponding songs that match the user mood. In the code, first we need to import the necessary libraries tkinter,matplotlib,pygame.We recommend music to the users using tkinter.Tkinter is the standard GUI library for Python. Python when joined with Tkinter gives a quick and simple approach to make GUI applications [19]. We first create an object for

tkinter.Then,we define the size of the GUI window. We then store the necessary images in different objects by specifying their paths for later use.We create a class called 'player' in which we define 5 methods. First method is create_frames() which creates 3 frames of different sizes as needed. Another function tract_widgets() is to display an image and for song tracking which displays the song name of the song that is currently being played.

Another function retrieve_songs() is to fetch the songs from the system according to the emotion detected and to display them on the frame right side. The function control_widgets() is to control the music player which can be play the music, pause the song, go to previous song, go to next song, increase or decrease volume. This is how we created a music player using tkinter.

ALGORITHM:

Step 1:Create a sequential model.
Step 2:Initialize n with 5.
Step 3:WHILE (n<9):
    Add Convolutional layer to this model with no of filters of size 2^n.
    Increment value of n.
    Add Convolution layer with no of filters of size 2^n.
    Add Activation function ReLu.
    Apply Max Pooling layer.
    Increment value of n.
  END WHILE
Step 4:Add a Flatten Layer.
Step 5:Add a Dense Layer.
Step 7:Add another Dense Layer.
Step 8:Add SoftMax Activation Function to get the probability for each class.

IV. EXPERIMENTAL RESULTS

The results, i.e., Output Screenshots of the proposed system under different testcases is as follows.

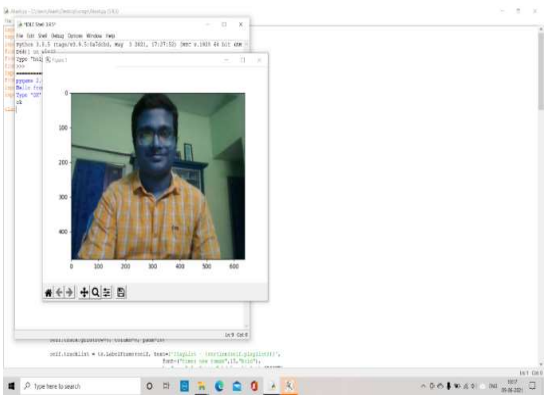

Fig. 4.1. Captured a Happy Face

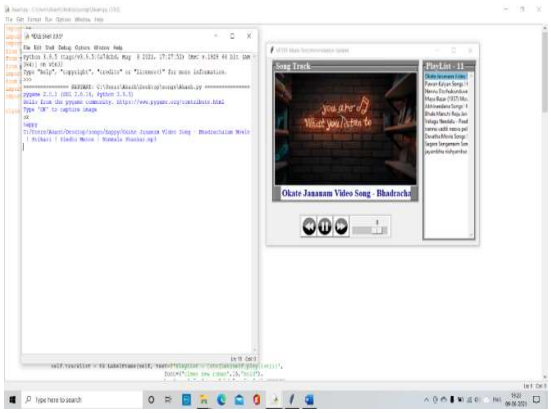

Fig. 4.2. Output for a Happy Face

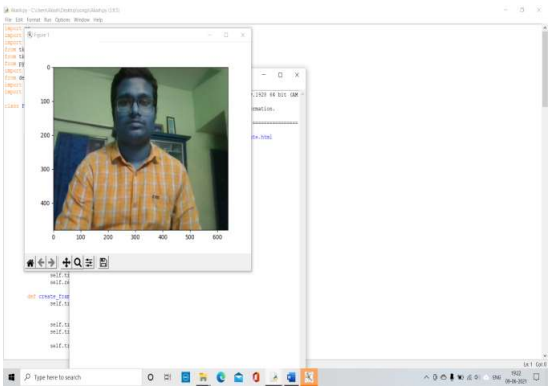

Fig. 4.3. Capturing a neutral Face

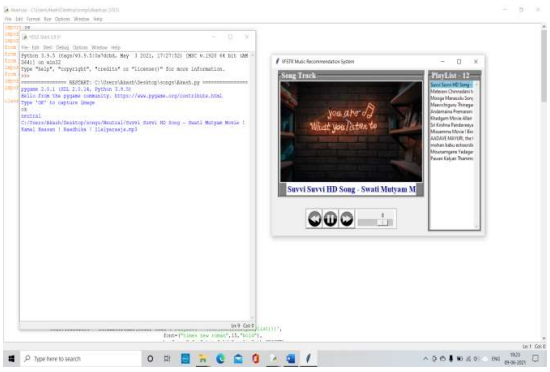

Fig. 4.4. Output for Neutral Emotion

V. CONCLUSION & FUTURE WORK

We would like to do much accurate model in the future and also a modem which detects all the basic 7 emotions which are happy, sad, anger, neutral, fear, disgust, surprise accurately. In the future we want to develop an android app for emotion-based music recommendation system which can be installed in our Android phones easily. We also plan to develop some special features where quotation of some great personalities will be recommended based on the user emotions if the user is detected as sad a song was suggested according to that song a quote will be displayed so that the user can be motivated. People tend to listen lyrics a lot at the same time

if quotation were showed they tend to feel motivated or feel active and where they can share it in social media this help at least few people to motivate. And other feature is that if user uses the application most then some points will be added and these points can be used to purchase singer albums. One more thing we want to include is we wish to recommend songs of different languages like Hindi, English, Tamil etc.

REFERENCES

[1] "Music & How It Impacts Your Brain, Emotions." https://psychcentral.com/lib/music-how-it-impacts-your-brain-emotions#1 (accessed Jun. 09, 2021).

[2] "RAVDESS Emotional speech audio | Kaggle." https://www.kaggle.com/uwrfkaggler/ravdess-emotional-speech-audio (accessed Jun. 08, 2021).

[3] "Smart Doll - Development Eyes of Things." http://eyesofthings.eu/?page_id=1884 (accessed Jun. 08, 2021).

[4] "Schematic diagram of a basic convolutional neural network (CNN)... | Download Scientific Diagram." https://www.researchgate.net/figure/Schematic-diagram-of-a-basic-convolutional-neural-network-CNN-architecture-26_fig1_336805909 (accessed Jun. 08, 2021).

[5] V Ramakrishna Sajja, Hemantha Kumar Kalluri, "Classification of Brain Tumors Using Convolutional Neural Network over Various SVM Methods", Ingénierie des Systèmes d'Information, Vol. 25, No. 4, August, 2020, pp. 489-495.

[6] "The Ultimate Guide to Convolutional Neural Networks (CNN) - Blogs SuperDataScience - Machine Learning | AI | Data Science Career | Analytics | Success." https://www.superdatascience.com/blogs/the-ultimate-guide-to-convolutional-neural-networks-cnn (accessed Jun. 08, 2021).

[7] V Ramakrishna Sajja, Hemantha Kumar Kalluri, "Brain Tumor Segmentation Using Fuzzy C-Means and classification using SVM", Lecture notes in Networks and Systems, Proceedings of Smart Technologies in Data Science and Communication, 2019, Vol.105, PP.197-204.

[8] R. L. Rosa, D. Z. Rodríguez, and G. Bressan, "Music recommendation system based on user's sentiments extracted from social networks," *IEEE Trans. Consum. Electron.*, vol. 61, no. 3, pp. 359–367, 2015, doi: 10.1109/TCE.2015.7298296.

[9] X. Zhu, Y. Y. Shi, H. G. Kim, and K. W. Eom, "An integrated music recommendation system," *IEEE Trans. Consum. Electron.*, vol. 52, no. 3, pp. 917–925, 2006, doi: 10.1109/TCE.2006.1706489.

[10] A. Lehtiniemi and J. Holm, "Using animated mood pictures in music recommendation," *Proc. Int. Conf. Inf. Vis.*, pp. 143–150, 2012, doi: 10.1109/IV.2012.34.

[11] S. Jun, S. Rho, B. J. Han, and E. Hwang, "A fuzzy inference-based music emotion recognition system," *IET Conf. Publ.*, no. 543 CP, pp. 673–677, 2008, doi: 10.1049/cp:20080398.

[12] Z. Yu, M. Zhao, Y. Wu, P. Liu, and H. Chen, "Research on Automatic Music Recommendation Algorithm Based on Facial Micro-expression Recognition," *Chinese Control Conf. CCC*, vol. 2020-July, pp. 7257–7263, 2020, doi: 10.23919/CCC50068.2020.9189600.

[13] K. Yoon, J. Lee, and M. U. Kim, "Music recommendation system using emotion triggering low-level features," *IEEE Trans. Consum. Electron.*, vol. 58, no. 2, pp. 612–618, 2012, doi: 10.1109/TCE.2012.6227467.

[14] A. V. Iyer, V. Pasad, S. R. Sankhe, and K. Prajapati, "Emotion based mood enhancing music recommendation," *RTEICT 2017 - 2nd IEEE Int. Conf. Recent Trends Electron. Inf. Commun. Technol. Proc.*, vol. 2018-Janua, pp. 1573–1577, 2017, doi: 10.1109/RTEICT.2017.8256863.

[15] Q. Q. Liu, J. Zhu, and T. Kong, "RSED: A novel recommendation based on emotion recognition methods," *Proc. - 2009 Int. Conf. Inf. Eng. Comput. Sci. ICIECS 2009*, no. 2, pp. 1–4, 2009, doi: 10.1109/ICIECS.2009.5364111.

[16] Y. H. Chin, S. H. Lin, C. H. Lin, E. Siahaan, A. Frisky, and J. C. Wang, "Emotion profile-based music recommendation," *Proc. - 2014 7th Int. Conf. Ubi-Media Comput. Work. U-MEDIA 2014*, pp. 111–114, 2014, doi: 10.1109/U-MEDIA.2014.32.

[17] C. Dhahri, K. Matsumoto, and K. Hoashi, "Mood-aware music recommendation via adaptive song embedding," *Proc. 2018 IEEE/ACM Int. Conf. Adv. Soc. Networks Anal. Mining, ASONAM 2018*, no. 2, pp. 135–138, 2018, doi: 10.1109/ASONAM.2018.8508569.

[18] "Deep Face Recognition - GeeksforGeeks." https://www.geeksforgeeks.org/deep-face-recognition/ (accessed Jun. 09, 2021).

[19] "Python - GUI Programming (Tkinter) - Tutorialspoint." https://www.tutorialspoint.com/python/python_gui_programming.htm (accessed Jun. 09, 2021).