Journal of Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.7699798

EMOTIONAL DETECTION OF FACE FOR SONG RECOMMENDATION USING DEEP LEARNING TECHNIQUE

[1] Dr. V. Divya, [2] Mrs. U. Hemamalini, [3] Dr. S. Arunarani
[1]Assistant Professor, Department of Information Technology, VELS Institute of Science, Technology and Advanced Studies (VISTAS), Chennai.
Email: divyavenkatraman1992@gmail.com
[2]Assistant Professor, Department of Information Technology VELS Institute of Science, Technology and Advanced Studies (VISTAS), Chennai.
Email: hemababu2501@gmail.com
[3]Assistant Professor, Department of Computer Applications, SRMIST, Kattankullathur.
Email: arunaprabhu07@gmail.com

ABSTRACT

The facial expressions of a user can be used to infer the user's state of mind or emotion. These expressions can be generated using the live feed provided by the system's camera. There is a significant amount of research being carried out in the field of computer vision and machine learning (ML), which is the process of teaching computers to recognise different human emotions and states of mind. The ability to read people's feelings is made possible by a number of different approaches provided by machine learning. The use of the mobilenet model with Keras is one method of this kind. This method generates a trained model with a small size and simplifies the process of Android-ML integration. People frequently turn to music as a technique for regulating their moods, particularly when they want to shift a poor mood, improve their level of energy, or alleviate tension. Additionally, if you listen to the appropriate genre of music at the appropriate moment, this may be beneficial to your mental health. Therefore, there is a close connection between the emotions of humans and music. In the system that we have proposed, a mood-based music player will be constructed. This player will do real-time mood detection and will make song recommendations depending on the detected mood. This adds a new function to the standard music player applications that are already pre-installed on our mobile devices. Customer satisfaction is an essential advantage that may be gained through mood detection. The purpose of this system is to analyse the user's image, make an educated guess as to the user's expression, and then make song recommendations that are relevant to the user's current state of mind.

Keywords: Machine Learning, Deep Learning, mobilenet, keras, Haar face detector, spotify.

INTRODUCTION

Every day, each and every person goes through a lot of struggles, and music is the stress reliever that may help relieve all of the tension that is experienced. If this is the case, one of the most important aspects of being able to hear the song is having a player who can adapt to the listener's state of mind so that the song is played appropriately. Recognizing human emotions is thought of as a global consistency; nonetheless, it depicts variety across persons based on the abilities they possess. The various methodologies that can be utilised in the categorization of feelings that are now available can be broken down into three categories: knowledge-based, statistical, and hybrid approaches. However, there are a number of challenges involved in getting information about music, such as categorising songs according to their genre, searching for specific singers, etc. The implementation that has the greatest potential is one in which music recommendations are generated on the basis of the content. It has been discovered that a feeling descriptor is helpful when attempting to describe a music taxonomy. When a sensation is analysed, it is first interpreted as a string of quantities, and then a mapping of the experience onto actual numbers is performed. A circumflex model has been developed in order to represent human emotions. Within this model, the impact of each emotion is represented over two bipolar dimensions. The aspects in question are alertness and sleep, as well as pleasantness and unpleasantness. As a result, each word is portrayed as combining a sense of pleasure and excitement. Valence and arousal are the two dimensions that are included in the model that was suggested by Thayer. There is a two-dimensional emotional plane in Thayer's model. This plane can be subdivided into four quadrants, and each of these quadrants can have one of 11 emotion adjectives placed on it. In the context of the network, different mental states such as happiness, sadness, anger, contempt, fear, surprise, and calmness are accounted for. The chance of moving from one state to another is factored into the computation of transitioning between states. Researchers in the field of emotion detection presented a method based on the analysis of two variables, tempo and articulation, which map to a variety of genres, including joyful, sad, and furious. If we take a step ahead and then apply the worry information, we will get another proactive music recommendation. Facial expressions are a natural method to express emotions, moods, and sensations. The usage of the boosted cascade allows for the efficient design to detect a face. It is thought that the Haar utilised in the Viola Jones face detector is quick in spotting frontal faces. However, given that the faces displayed are in an uncontrolled context, Haar's simplicity makes it challenging to identify. The user must spend a great deal of time searching for the right music to listen, which could be made easier. This chore of having to explore songs according to one's attitude is seen to be labor-intensive. Despite the fact that there are many face detection systems, they have various limitations, such as poor picture quality caused by illumination. The fact that each person's facial looks and pattern variations make identifying a face the most challenging. There are pattern variances since people's skin tones and differences in face features occur. Common items like a beard and spectacles are almost always present. Additionally, being 3-dimensional things, faces, the influence of Lighting may contribute to the variation. It is also possible to use a pre-processing step to lessen the impact of light.

EXISTING AND PROPOSED METHODOLOGY

The following is a list of characteristics that may be found in the many music players that are already installed in computers: Party Shuffle, Playlists, Music Squares, and Manual Selection of Music all require the user to manually classify the songs in line with designated moods for just four basic emotions: Passionate, Calm, Joyful, and Excitement. To identify songs that would improve his mood and emotional condition on conventional music players, a user had to actively search through his playlist. A number of music players have been created in the modern world with capabilities like fast forward, reverse, variable playback speed (seek and time compression), local playback, streaming playback with multicast streams, volume modulation, genre categorization, and other functions of a like kind. Additionally, these devices can play music. Even though these features satisfy the user's basic needs, the user must still manually browse the playlist of music and choose songs based on his or her present mood and behaviour. That is what a user would anticipate, who periodically feels the urge and want, depending on his emotions and mental state, to browse through his playlist. The multilayer perceptron convolution neural network technique was created specifically for the recognition of information about pictures with just two dimensions. It consists of an output layer, a sample layer, a convolution layer, and an input layer. These four layers make up its structure. In a deep network architecture, the sample layer and the convolution layer each could include several copies of themselves. Although CNN is less constrained than the Boltzmann machine, all connections in convolutional neural network algorithms must come before and after the layer of neurons in the adjacent layer. Each neuron simply has to sense the local area of the picture; it is not necessary for it to feel the entire image. Each neuron's parameters are also synchronised, meaning they are all set to the same value. This makes it possible for neurons to share weights and equips them all with the same convolution kernels for the deconvolution picture. Some of the foundational periods of CNN are the local receptive field, weight sharing, and subsampling by employing time or space. In order to extract features and condense the amount of the training parameters, these periods were created. The main advantage of the CNN algorithm is that it can learn automatically from the training data without needing to manually extract features. The feature mapping's surface has equal neuron weights across the board, allowing the network to learn concurrently while maintaining a modest degree of complexity. using a framework for resampling that is based on the stability of time, scale, and deformation displacement. There is a chance that the input data and the network's architecture will be a perfect match. It offers special benefits for image processing.

MODULE CLASSIFICATION

The process of song recommendation is broken down into three distinct modules: facial expression detection, music mood classification, and user interface. The Facial Emotion Detection model needs to begin by locating a face in the picture before it can learn to categorise the subject's expression as either sad, joyful, tranquil, or enthusiastic. A music mood classification model should use elements of the song like its tone or valence in order to determine the feeling that is expressed by it. The user interface was created using HTML and flask, and it allows them to upload their own picture and playlist if they so like. The user should subsequently be presented with the emotion that the app recognised combined with some randomly selected, mood-appropriate music.

Facial Expression Model

Image processing and detection are both aspects of the neural network known as a CNN, which stands for "convolutional neural network." The input layer, the convolutional layers, the dense layers, and the output layer are the components that make up a CNN. The image is broken down into its component parts using convolutional layers, which aid in identifying the particular emotion being conveyed. In order to accurately detect emotions, you must first locate a face inside an image and then put that face through a model that can determine its expression. After that, the haarcascade frontal default model was utilised in order to successfully recognise the faces on their own. A frontal default version of the haarcascade model will be found to be the most accurate and most applicable after a series of tests involving a variety of haarcascade models corresponding to various areas of the face. The haarcascade model generates a printout that consists of the four coordinates of the bounding box that encompasses the detected face. When applying the haarcascade model, we found that the face in the photograph had to be looking directly into the camera; it couldn't be angled in any way, and all of the facial features had to be discernible. The use of a customised model with tailored input parameters did not produce satisfying outcomes. In spite of the fact that the accuracy was high (in the 90% percentile), it was only around 55% accurate in the testing dataset. The initial model made predictions about false positives for fear, in which cheerful faces were misclassified as ones filled with terror. In addition, there were not many instances in which the emotion "sad" was misidentified. The VGG16 performed far better than the earlier custom model, which only achieved an accuracy of 59% in the testing dataset. Nevertheless, the model has a fault, which is illustrated by the confusion matrix that can be found below. The flaw is that quite a few photographs that were sad were classified as neutral. However, the VGG16 model did not perform as well as expected when it came to classifying more nuanced emotions, such as neutral or sad feelings. For example, it did a better job categorising more extreme and expressive emotions such as happiness and rage. The VGG16 model required a significant amount of room as well.

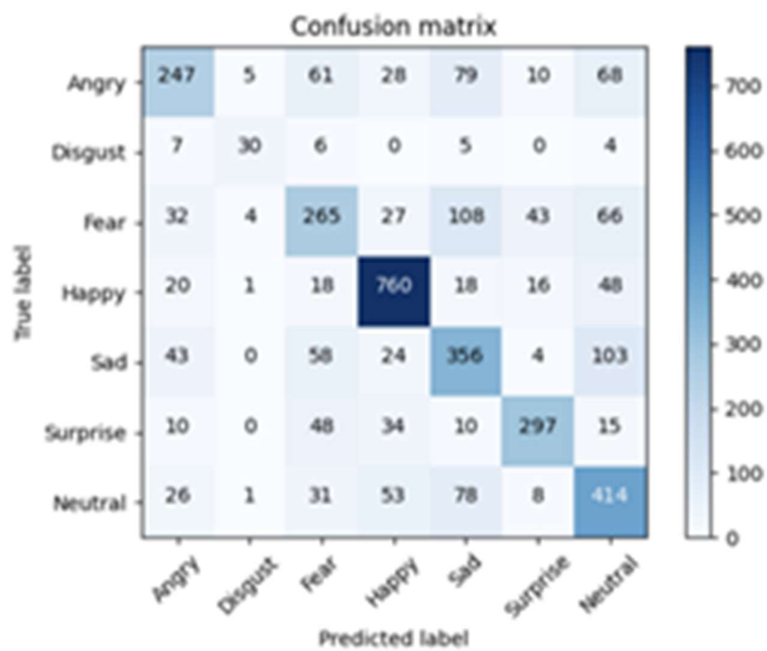

Figure 1: CONFUSION MATRIX OF CUSTOM MODEL

The ultimate solution is to make use of a VGG16 model that has specific parameters, which is represented in figure 2. This model was more accurate than the one that came before it when it came to predicting the more nuanced emotions, such as neutral and sadness. Even though the final model still causes some confusion between happy and neutral faces, production is moving forward with it because it produces the best results.

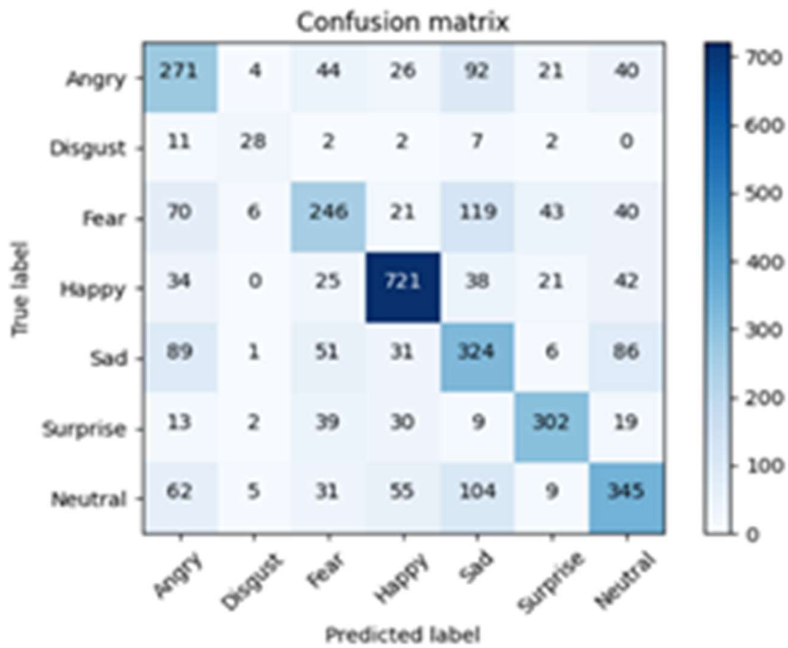

Figure 2: CONFUSION MATRIX OF VGG16 MODEL

Music Mood Classifier

The ability to categorise songs according to their effect on the listener was the overarching goal of this study. It was necessary for us to train a model that could use a variety of aspects of a song, such as its tone, danceability, acousticness, energy, pace, and so on, to identify whether or not a song was happy, sad, calm, or energy-pumping. To begin, we searched through the prepared playlists on Spotify and identified a variety of playlists that each featured a specific kind of music. The majority of the tunes that we picked for the happy music were pop songs. Pop and electronic dance music were both employed to create upbeat songs. Mellower tracks were picked for the sorrowful music we were playing. We utilised LoFi for the more tranquil songs. The Spotify Application Programming Interface (API) as well as the Spotipy Python Library are utilised in order to extract song information and then save it into the datasets. At the conclusion of the process, a dataframe was produced that included the following characteristics of each song: "danceability," "acousticness," "energy," "instrumentalness," "liveness," "valence," "loudness," "speechiness," and "tempo." The following stage consists of training the classifier, which was accomplished with the help of TensorFlow and Keras. It is necessary to convert the model into numpy models in order to make the Pandas dataframes compatible with it. Following this step, we then use Sklearn to do a 66/33 split on the dataset that contains song attributes and labels. The final model was created using Keras; it was trained using the Scikit-Learn Keras classifier function and included a single dense layer that used the ReLU activation function. At the end of the process, the model achieves an accuracy of 74% during training and 70% during testing.

	danceability	acousticness	energy	instrumentalness	liveness	valence	loudness	speechiness	tempo
0	0.667	0.01230	0.830	0.000000	0.1910	0.701	-5.715	0.0749	113.030
1	0.647	0.21900	0.822	0.000000	0.0908	0.962	-4.662	0.1830	160.01
2	0.620	0.33200	0.832	0.000000	0.1040	0.481	-4.848	0.0360	144.92
3	0.448	0.21000	0.838	0.000131	0.1310	0.254	-5.257	0.0557	139.863
4	0.673	0.18500	0.886	0.000000	0.0826	0.795	-4.440	0.0431	97.012
...	...	...	...	...	...	...	...	...	...
1195	0.760	0.04510	0.941	0.000000	0.6490	0.805	-6.356	0.0688	116.013
1196	0.284	0.00583	0.704	0.000115	0.0424	0.100	-7.276	0.1860	191.702
1197	0.625	0.02740	0.833	0.000000	0.2640	0.581	-5.389	0.0347	126.013
1198	0.794	0.03060	0.758	0.000001	0.2690	0.821	-6.289	0.0713	115.990
1199	0.810	0.01670	0.956	0.281000	0.3500	0.961	-5.753	0.0367	129.996

1200 rows × 9 columns

Figure 3: ATTRIBUTES OF DATAFRAMES

The challenge of keeping track of a large amount of music has gotten increasingly difficult. The fact that the Spotify API was unable to retrieve information for each and every one of the songs included in the dataset is one of the most significant mistakes. You will, in essence, need to go through the playlist song by song in order to figure out which song is causing an error so that you can manually remove it. If you want to simplify the process of training and analysing data, you should separate all of the code for the music classifier into two different notebooks: one for data and one for training. Pandas is used to visualise our dataframes in the data notebook. We spent a significant amount of time making sure that all of our data sizes were consistent, that all of the songs worked with the API, etc. All of the data are labelled with a single hot encoding, and the dataframes are exported as.csv files so that they can be stored and used at a later time. Importing the data, preprocessing or splitting the data, creating a model, training the model, and testing it were all covered in the second notebook, which made use of the.csv files from the previous notebook. It also outputs the finished model as a.pkl file in order to save the hyperparameters so that it can be used for future predictions.

Application Interface

The user is requested to input an image in the final design, which is used to initiate the classification and music recommendation processes. When an individual chooses a file to upload and then clicks the "Submit" button, a function is activated that recognises that action and then gives back a particular feeling. The music mood classifier has already completely categorised the default playlist that is placed into the app for your convenience. A function takes the emotion that was identified as an input and then searches for a song that corresponds to that feeling and plays it. Following this, the app will display to the user an embedding of the song's Spotify page alongside the picture that they have uploaded.

SONG DETECTION USING DEEP LEARNING

The positive effects of music on human wellbeing may be implemented in a broad range of settings; the use cases that were covered in the section that came before this one suggest the possibility for the practical application of the effects of music on people. A generalist system that considers the particular, specific qualities of each user is necessary in order to handle these and other conceivable eventualities. This part focuses on the development of a solution that aims to identify individual features of the physical and emotional impact of music-related components in a range of situations and combine those aspects with well-established generic approaches. The system's primary goal is to locate the musical tracks that are the closest (closest) to the abstract etalon one, which is defined by a particular set of musically-related characteristics, irrespective of the actual intent, such as whether it is necessary to change a user's emotional state or to retain and maintain it the same. This task is carried out whether it is necessary to alter a user's emotional state or to preserve and keep it constant. The first tactic is a simple one in which the ultimate goal, or the targeted position in a user's emotional distance, is plainly stated. This point contains a vector of several music-related, person-related, and situation-related factors attached to it within the context of the music-driven emotion therapy, each of which has a value. In this case, the system searches for the adjacent or nearby music tracks by using the related distance measurement function. In a different, more sophisticated version of the system, the changeover to the desired point is easier and less forceful, and it is done while taking into consideration the user's present position in the emotional space. In this instance, the system bridges the discrepancy between extracted features reflecting the present location and the ones that are planned. It is anticipated that employing this mode would make the user feel less anxious and obtrusive. The multidimensional (multi-featured) "delta" distance (step) computation must, thus, utilise a more complicated algorithm that takes into account the particular qualities of each specific user. The Personalized Emotion Transformation Model (PETM) for each user must thus be recorded throughout the phase of data gathering and model construction. Modify the model throughout the system's continuing operation to ensure that it consistently reflects the user's current personality. This holds true for the aforementioned two situations. Constant data collection and analysis of user input will therefore enable on-the-fly model training and raise the amount of personalisation in predictions. Using the data received from the first survey, the system creates a general user profile, sometimes referred to as a GUP, as its first action. Additionally, a user's initial music-driven emotional model (MEM) is created using the GUP system. This model may also be used with an average model that shows the closest group of users that are grouped together according to their GUPs. If we knew what the GUPs were for every user, we could do this. The emphasis should be placed on the system's individualization and customization because the goal is to build one that is more efficient. The basic music-driven user model is therefore further customised and improved in the second stage of the system. Each customised model must decide which musical genres can support different state transitions in order to conduct transitions of the personal conditions in a range of circumstances.

We worked to enhance the trial version of the system and validate the correctness of the recommendation model. The core element of the system is a web service. This service collects input on listening experiences and musical qualities before categorising music songs in order to provide further suggestions. The Spotify platform, the MuPhonic tool, and the music player are all integrated into the mobile application. Users have the choice to listen to music on Spotify or the player; this has no impact on the data collecting or suggestion process. The MuPsych tool's data will be the main source of information used in the early stages of the experimental prototype's model training process. Because the application only offers rudimentary data and personality tests, users may keep total anonymity. This enables the programme to accurately classify people so that it may provide suggestions in the future that are more accurate. The degree of activity, emotional state, and rationale of the listener are only a few examples of the variables that affect the context in which music is listened to. The smartphone application queries the user about their mood and the reasons they are listening to music when they first start playing music, after five minutes of listening, and once they have completed listening. The MuPsych tool is meanwhile capturing the musical qualities of the song being listened to. In addition to the interactive user response, the data processing engine is in charge of gathering all the data. According on user clusters, contextual information, listening preferences, and track characteristics, the music recommendation system classifies music tracks. The datasets that MuPsych has already gathered as well as data that is regularly updated from the user side provide the foundation for this categorization. The second stage of the process involves the system delivering generic playlists that are created with the activity environment and audience group's mood in mind. The system then offers these playlists. The system will continuously classify tracks and modify the customised recommendation model's parameters as more user-specific data is obtained. The mobile application's graphical user interface encourages participatory listening with feedback, which in turn affects the creation of new playlists. Trial trials showed that for the specific settings of mood and activity, created playlists' main musical characteristics, such as energy, valence, tempo, and loudness, adequately matched those of the music recordings in the MuPsych datasets. As a consequence, we may infer that the system selects songs with characteristics that are similar to those that were examined during the changes in emotional state that took place while listening. Additional elaborations, thorough experiments, and system validations are needed to determine how ideas impact mood and how well they match wants.

PERFORMANCE MEASURE FOR EMOTIONAL FACE DETECTION

In the area of accuracy, the word "estimated time of arrival," or "ETA," most commonly refers to the anticipated arrival time, while in the realm of technology, it refers to the anticipated end time of a calculation process in general. The issue is too narrowly focused on calculating the amount of time it will take for a group of lengthy scripts to finish running in parallel with one another while processing data and producing lists. The amount of time needed to complete each one is subject to change based on the historical data that was taken into consideration. The period at which all of the images are processed individually, both forward and backward to the network, is referred to as an "epoch." In most cases, we feed training data to a neural network for more than one epoch in a variety of different patterns. This allows for improved generalisation when the network is presented with data that it has not seen before. If there is a large but limited training dataset, then the neural network will have the opportunity to revisit the earlier data in order to fine-tune the model parameters. This will ensure that the model is not overly reliant on the most recent few data points while it is being trained. The term "loss" refers to nothing more than the prediction mistake made by a neural network, while the term "loss function" refers to the mechanism that is used to calculate the loss. A loss function is used in the optimization of the machine learning algorithm. The loss is estimated based on the training set and the validation set, and its significance is assessed by contrasting it with the model's efficiency in both of these sets. In the training or validation sets, it shows the overall number of mistakes made for each unique occurrence. A model's loss value after each optimization step shows whether or how how poorly the model performs. In order to measure an algorithm's efficacy in a clear and accessible way, accuracy metric is used. The correctness of a model is frequently determined after the model and is expressed as a percentage. After the model parameters have been established, the accuracy of the model is frequently measured as a percentage. It is a measurement of the model's prediction accuracy in relation to the training set, which is the real information itself.

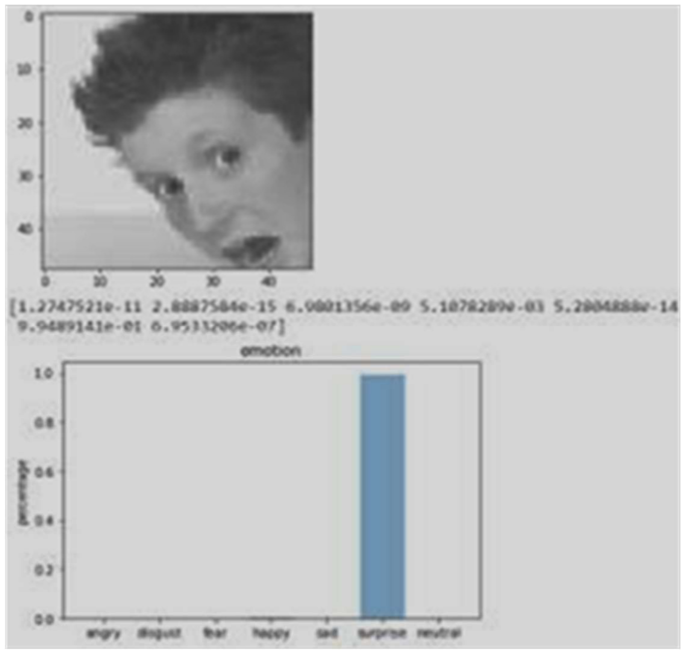

Figure 4: Emotion Detected: Surprise with 99.489 accuracy

The figure 4 depicts the emotion detection of the surprise face with the percentage of 99.489 accuracy. Similarly figure 5 depicts the emotion detection of sad face with the accuracy of 99.99 percentage.

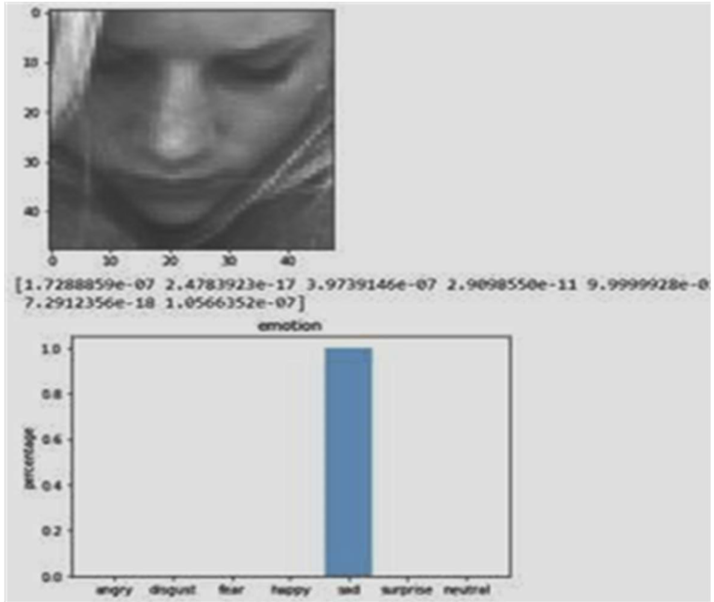

Figure 5: Emotion Detected: Sad with 99.99 percentage of accuracy

The figure 6 depicts the emotion detection of angry face with accuracy of 99.612 percentage.

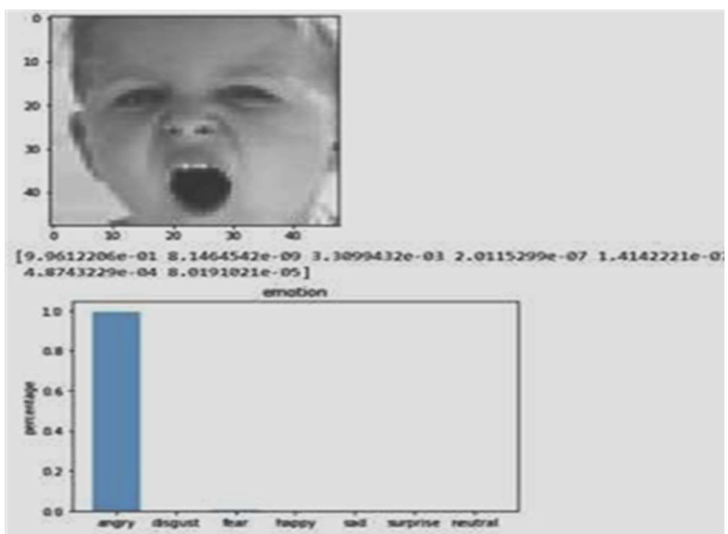

Figure 6 : Emotion Detected: Angry with 99.612 percentage of accuracy

CONCLUSION

The process of developing suggestions must take into account a number of factors, including the particular circumstances at hand, individual preferences, sentiments, and sentiments. Current music recommendation systems have the difficulty of bridging the disparity between personalisation, human sentiments, contextual preferences, and emotional factors when it comes to offering music. In this study, we suggested an emotion-driven recommendation system that takes into consideration both the preferences of the individual and the particular life and activity conditions. The research's technique attempts to provide people with the greatest number of advantages that might come from the act of listening to music. Making progress steadily in building a greater music collection requires Making the system aware of the method through which it creates suggestions is crucial. The system's mission is to learn each individual user's hearing aspirations, sentiments, and contextual preferences in order to choose the musical selections that are most appropriate for them. Data from several sources are fed into the system to achieve this. Considering the types of data needed for the recommendation system as well as potential retrieval techniques, we made several observations. The main data processing methods are explained within the purview of this study, and an experimental prototype has been created. Furthermore, a significant amount of the data must be utilised to train the machines in order for systems for machine learning to achieve the best accuracy in their predictions and to make them more or less relevant. The process of gathering data is now being performed out. In the same vein, a system of this kind need a significant amount of scientific research, as well as collaboration with psychologists, to refine and examine the model for true ideas and reduce any possible risks associated with its usage. The next phase will likely involve doing further work on the development and testing of the recommendation engine after a significant amount of the data has been collected, executing empirical experiments and assessing the project's impact. In this view, the work that is considered to represent further progress might be seen of as the production of music by artificially intelligent systems endowed with certain musical characteristics intended to alter emotional states in people.

REFERENCES

[1] G. Heamalathal, C.P. Sumathi, A Study of Techniques for Facial Detection and Expression Classification, International Journal of Computer Science & Engineering Survey(IJCSES) Vol.5, No. 2, April 2014.

[2] Bernhard Fröba Christian Küblbeck Real-Time Face Detection Using Edge-Orientation Matching AVBPA '01 Proceedings of the Third International Conference on Audio- and Video-Based Biometric Person Authentication Pages 78-83 Springer-Verlag London, UK ©2001 table of contents ISBN:3-540-42216-1

[3] Oliver Jesorsky, Klaus J. Kirchberg, and Robert W. Frischholz BioID AG, Berlin, Germany, Robust Face Detection Using the Hausdorff Distance. Third International Conference on Audio- and Video-based Biometric Person Authentication, Springer, Lecture Notes in Computer Science, LNCS-2091, pp. 90–95, Halmstad, Sweden, 6–8 June 2001.

[4] PAUL VIOLA Microsoft Research, One Microsoft Way, Redmond, WA 98052, USA And MICHAEL J. JONES Mitsubishi Electric Research Laboratory, 201 Broadway, Cambridge, MA 02139, USA Robust Real-Time Face Detection International Journal of Computer Vision 57(2), 137–154, 2004_c 2004 Kluwer Academic Publishers. Manufactured in The Netherlands.

[5] http://changingminds.org/techniques/body/parts_body_la nguage/face_body_language.ht

[6] The psychological functions of music listening Christine Städtler and David Huron US National Library of Medicine National Institutes of Health Published online 2013 Aug 13. Prepublished online 2013 May 24.

[7] Face recognition using Histograms of Oriented GradientsO. Déniz, G. Bueno, J. Salido, F. De la Torre Universidad de Castilla-La Mancha, E.T.S. Ingenieros Industriales, Avda. Camilo Jose Cela s/n, 13071 Ciudad Real, Spain Carnegie Mellon University, Robotics Institute, 211 Smith Hall, 5000 Forbes Ave., Pittsburgh, PA 15213, USA Pattern Recognition Letters 32 (2011) 1598–1603
[8] Histograms of Oriented Gradients for Human Detection Navnee Dalal And Bill Triggs INRIA Rhone-Alps,655 Avenue de l'' Europe, Montbonnot 38334, France
[9] Porawat Visutsak Emotion Classification through Lower facial Expressions using Adaptive Support Vector Machines
[10] Vahid Kazemi and Josephine Sullivan One millisecond face alignment with an Ensemble of regression Trees Computer Vision Foundation CVPR2014
[11] Melanie Dumas Department of computer science,University of california,san Diego CA 92192- 0114
[12] Lundqvist, D., Flykt, A., & Öhman, A. (1998). The Karolinska Directed Emotional Faces -KDEF, CD ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet, ISBN 91-630-7164-9.
[13] Li-Fen Chen and Yu-Shiuan Yen.(2007). Taiwanese Facial Expression Image Database. Brain Mapping Laboratory, Institute of brain Science, National YangMing University, Taipei, Taiwan.
[14] Michael J. Lyons, Shigeru Akemastu, Miyuki Kamachi, Jiro Gyoba.Coding Facial Expressions with Gabor Wavelets, 3rd IEEE International Conference on Automatic Face and Gesture Recognition, pp. 200-205 (1998).
[15] Davis E. King. Dlib-ml: A Machine Learning Toolkit. Journal of Machine Learning Research 10, pp. 1755- 1758, 2009
[16] C. Chang and C.-J. Lin. LIBSVM : a library for support vector machines. ACM Transactions on Intelligent Systems and Technology, 2:27:1--27:27, 2011.
[17]. Tao Li, Mitsunori Ogihara. Detecting Emotion in Music. ISMIR (International Conference on Music Information Retrieval),2003.
[18]. Namrata Mahajan, Harshad Mahajan. Detecting Emotion in Music. International Journal of Electrical and Electronics Research ISSN 2348- 6988 (Vol. 2), Issue 2, pp: (56-60), Month: April - June 2014.
[19]. Hafeez Kabani, Sharik Khan, Omar Khan, Shabana Tadvi. Emotion Based MusicPlayer. International Journal of Engineering Research and General Science Volume 3, Issue 1, January-February, 2015.
[20]. Abhishek R. Patel, Anusha Vollal, Pradnyesh B. Kadam, Shikha Yadav, Rahul M. Samant. MoodyPlayer: A Mood based Music Player. International Journal of Computer Applications (0975 8887) Volume 141 No.4, 2016.