iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** III **Month of publication:** March 2022

DOI: https://doi.org/10.22214/ijraset.2022.40668

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

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

Mood Based Music Playlist Generator Using Convolutional Neural Network

Prof. Jaychand Upadhyay[1], Sharan Shetty[2], Vaibhav Murari[3], Jarvis Trinidade[4]
[1, 2, 3, 4]*Department of Information Technology, Xavier Institute of Engineering, Mumbai University, Maharashtra*

Abstract: ***Music is used in one's everyday life to modulate, enhance, and plummet undesirable emotional states like stress, fatigue, or anxiety. Even though in today's modern world, with the ever-increasing advancement in the area of multimedia content and usage, various intuitive music players with various options have been developed, the user still has to manually browse through the list of songs and select songs that complement and suit his or her current mood. This paper aims to solve this 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 can be deduced from facial expressions. To create a neural network model, the CNN classifier was used. This model is trained and tested using OpenCV to detect mood from facial expressions. A playlist will be generated through the system according to the mood predicted by the model. The model used for predicting the mood obtained an accuracy of 97.42% with 0.09 loss on the training data from the FER2013 Dataset.***
Keywords: Convolutional neural networks, Facial emotion recognition, FER2013, Music, OpenCV.

I. INTRODUCTION

Facial expressions are important because they reveal important information about a person's emotions [14]. The next generation of computer vision systems could be heavily influenced by present computer systems that incorporate dynamic user interactions. Face emotions are a major prospect in security, entertainment, and human-machine interaction (HMI). A person's emotions can be expressed through their lips and eyes. In today's environment, having a vast music playlist is commonplace. Music is a crucial source of entertainment for music lovers, and it can even provide a therapeutic method in some cases. Music can lift one's mood, while also boosting one's happiness and helping to reduce anxiety. From ancient times to today's modern world of technology that encompasses unlimited streaming services, music is an integral part of the human experience [13].

The link between music and mood is thus so evident that its effects are often known to improve one's emotional and psychological state. The proposed system detects one's mood by facial emotion recognition using convolutional neural networks and generates a music playlist based on the mood detected. Unlike human beings detecting facial expressions can be a formidable task for machines. However, with the advancement in technology, the task is made somewhat possible with Convolutional Neural Network proving to be an efficient solution for facial emotion recognition.

The model will be trained on the FER 2013 Dataset [15]. The data consists of grayscale facial images of faces. The images are 48 pixels in height and width. The training set comprises 28,709 examples and the public test set comprises 3,589 examples. Keras tuner will be used to optimize the model and choose the accurate number of convolutional layers, flattened layers, and dense layers required to gain the highest accuracy. Once the face has been detected with the help of a haar-cascade classifier in the image captured by the user through the webcam, the image is passed on to the trained and optimized CNN model to predict the Mood of the user. A playlist will be generated through the music Database according to the mood predicted by the model. Songs will be chosen randomly from the respective mood category. The system will generate those selected songs through external storage and display them to the user to listen to the songs to enhance their moods. This paper aims to provide an alternative for suggesting songs to a user which is via the user's facial emotion. The paper also aims to create a high-accuracy CNN model for detecting facial emotions.

II. RELATED WORKS

The preceding music players had a variety of characteristics. The following are the characteristics that are present in the previously created Music players: i) Song chose manually. ii) Shuffle Songs. iii) Create a playlist

There are different ways of emotional analysis by using facial expression, gestures, body movement, speech, etc [3]. Many research has been conducted with different approaches for detecting and classifying the physiological behavioral and emotional status expressed on the face by the users. The facial digital image is pre-processed and subjected to different feature extraction and classification algorithms.

©IJRASET: All Rights are Reserved

Convolutional neural networks (CNN) are a good option for facial emotion recognition. The author of this paper [1] wants to improve the performance of a convolutional neural network by adjusting its hyperparameters and architecture and therefore converting human facial photos into different sorts of emotions. The Random Search algorithm was used to create and train models with many hyperparameter combinations and architectures from a large range of possible solutions. Thus, when this method based on convolutional neural networks was used for the FER2013 dataset, the results revealed an accuracy of 72.16 percent, which was deemed acceptable given the optimization limitations.

Facial emotional processing is one of the maximum vital sports in powerful calculations, engagement with humans and computers, system vision, online game testing, and client studies. Facial expressions are a shape of nonverbal verbal exchange, as they screen a person's internal emotions and feelings. As the quickest verbal exchange medium of any type of information. Facial expression popularity offers higher information of a person's mind or perspectives and analyzes them with the presently trending deep getting to know methods. This paper [2] presents a quick assessment of the one-of-a-kind FER fields of software and publicly available databases utilized in FER and researches the state-of-the-art and cutting-edge critiques in FER of the use of convolutional neural network algorithms. Finally, it's miles located that everybody reached top results, particularly in phrases of accuracy, with one-of-a-kind rates, and the use of one-of-a-kind information sets, which affects the results. They have supplied diverse databases of random pictures received from the real global and different laboratories to stumble on human feelings accurately. Researchers have discovered and checked their styles in lots of datasets to confirm the neural community architecture's feasibility. The developing extent of information, processing time, and computational sources wanted for the studies methods will affect the algorithms utilized in classification, good enough memory, and version performance.

The work presented in this paper [3] focuses on a system that proposes songs to users based on their mood. It provides a CNN-based approach that evaluates multimodal emotional information recorded by facial movements as well as semantic analysis of the user's speech or text interactions, hence enhancing the system's real-time judgment on detected emotions. The model of the proposed system includes modules for identifying facially displayed emotions and sentiments communicated through a chatbot conversation, resulting in a robust music recommender system. When an emotion is recognized, the system proposes a song to match that mood, saving the user time from having to manually select music.

This study [4] presents a music recommendation system that extracts the user's image and predicts the mood based on the extracted image.

The user's sentiment is detected via a network. The HELEN dataset is used to learn visual features from the user's face as well. The user's picture is taken and then depending on the user's mood/emotion an appropriate song from the playlist will be played, corresponding to the needs of the user. Once this has been detected, an appropriate song is selected by the music player that would best match the mood of the user. Users don't have to spend time searching or searching for songs and tracks that best suit the user's mood will be detected and played automatically by the music player. The system can also capture new images of the user and appropriately update its training and classification datasets.

The different datasets that are widely used for facial emotion identification, as well as the various deep learning modules that are employed for face emotion detection, are examined in this study [5]. It shows the several CNN models and their accuracies, as well as the datasets that CNN was applied to. This research also provides an overview of current advances in facial expression identification for sensing emotions using various deep learning architectures.

Databases	Descriptions	Emotions
MultiPie [10]	More than 750,000 images captured by 15 view and 19 illumination conditions	Anger, Disgust, Neutral, Happy, Squint, Scream, Surprise
MMI [11]	2900 videos, indicate the neutral, onset, apex and offset	Six basic emotions and neutral
GEMEP FERA [12]	289 images sequences	Anger, Fear, Sadness, Relief, Happy
SFEW [13]	700 images with different ages, occlusion, illumination and head pose.	Six basic emotions and neutral
CK+ [14]	593 videos for posed and non-posed expressions	Six basic emotions, contempt and neutral
FER2013 [15]	35,887 grayscale images collect from google image search	Six basic emotions and neutral
JAFFE [16]	213 grayscale images posed by 10 Japanese females	Six basic emotions and neural
BU-3DFE [17]	2500 3D facial images captured on two view -45°, +45°	Six basic emotions and neutral
CASME II [18]	247 micro-expressions sequences	Happy, Disgust, Surprise, Regression and others
Oulu-CASIA [19]	2880 videos captured in three different illumination conditions	Six basic emotions
AffectNet [20]	More than 440.000 images collected from the internet	Six basic emotions and neutral
RAFD-DB [21]	30000 images from real world	Six basic emotions and neutral

Fig. 1 Summary of some FER Databases [5]

The goal of this assignment [6] is to extract features from the human face and discover emotion and play songs according to the emotion detected. Facial expressions have been captured by a picture tool or an in-built camera. Existing strategies use preceding records to signify tune and different algorithms used are normally slow. Emotion-Based-tune-participant has the functionality to stumble on feelings that is the face of the consumer with the assistance of system studying a set of rules the use of python. Based on the detected person's mood song listing could be displayed/propose to the person. It makes use of the python Eel library to choose a random tune without any order. The emotion of a consumer is extracted via way of means of taking pictures the photograph of the consumer via a webcam. The captured picture is improved via way of means of the method of dimensional reduction. This includes taking the preliminary or the number one records taken from the human face this is reduced to many different classes.

The goal of this work [7] was to identify emotions in human-robot interactions. Its goal is to extract deep and abstract features from facial emotional photos, as well as to mitigate the impact of deep learning's complicated structure and sluggish network updates on facial emotion recognition. Convolution and maximum pooling are employed to extract features, and then the ridge regression algorithm is utilized to recognize emotions. When the network needs to grow, the incremental learning method dynamically updates the network. K-fold cross-validation was used to verify the experimental results. The suggested paper's recognition accuracy on the JAFFE database is higher than the state of the art, such as Local Binary patterns with softmax and deep attractive multi-path convolutional neural networks, according to the recognition findings.

The proposed paper's main contribution is the application of deep learning and broad learning to facial emotion recognition. The use of CNN for extracting deep and abstract features from facial emotional images is the key contribution of this research. Simultaneously, BL is used to achieve quick network weight updating, eliminating the problems of complicated network topology and sluggish updating speed. To cut down on retraining time, incremental learning was implemented, which can be quickly reconstituted in large-scale expansion. The paper had an accuracy of 92.85 percent.

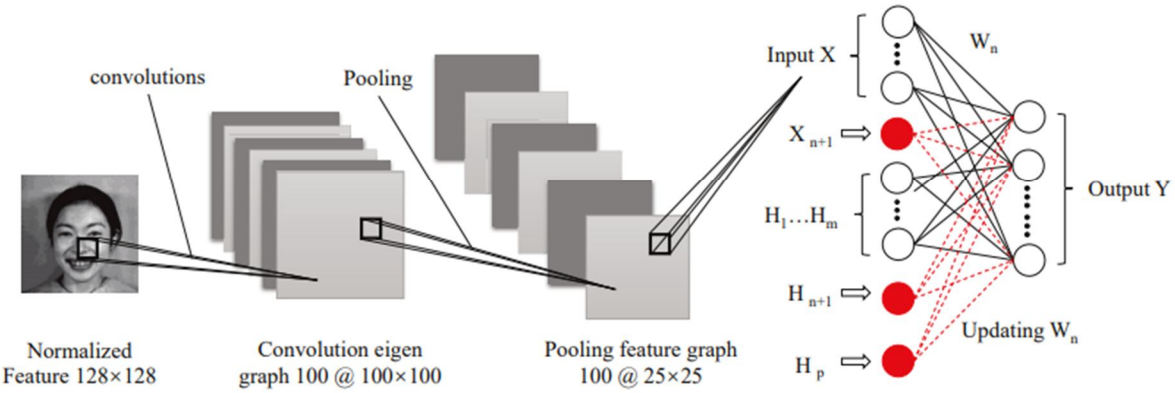


Fig. 2 The structure of CNNBL with efficient incremental reconstruction [7]

This study [9] focuses on developing an effective music recommendation system that employs Facial Recognition algorithms to assess the user's emotion. The constructed algorithm would prove to be more effective than previous systems. The system's overall goal is to recognize facial emotion and quickly recommend tunes. The proposed system saves both time and money. A simple system for music recommendation based on face expression identification is proposed here. It suggests music by analyzing a person's facial expressions: happy, angry, surprised, and neutral. To recognize a face from a static image, the recommended framework must first be prepared. The image is processed once the information picture has been perceived. SVM classifiers are used to analyze the image for subtleties to detect the emotion portrayed by the face. The feeling classifier uses the details recovered from the image to determine to feel.

The proposed method in the study [10] extracts the current mood of an individual user by scanning and analyzing the user's facial expressions. After the emotion has been retrieved, the user will be presented with a playlist of music that matches the mood. Image processing and facial detection technologies are included in the system. Images are still the most common input, and they are further analyzed to determine the user's sentiment. At the start of the application, images are taken. The photographs are captured using Web-Cam. The image that was previously captured will be preserved and sent to the rendering process. The suggested system in the research aims to create a music playlist by first performing facial detection using HAAR traits, then feature point detection, and finally PCA-based emotion prediction. The goal of the Mood Depending Music Player with Real-Time Facial Expression Extraction is to give the user a pleasant music listening experience by presenting appropriate music content based on his or her current mood.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue III Mar 2022- Available at www.ijraset.com

III. METHODOLOGY

A. Data Processing

Convolutional Neural Networks are used for creating the model to predict the mood/emotion of the user. The model is trained on the FER 2013 Dataset [15]. The data consists of grayscale facial images of faces. The images are 48 pixels in height and width. The training set comprises 28,709 examples and the public test set comprises 3,589 examples. The faces have been automatically registered such that they are more or less centered in each image and take up roughly the same amount of area.



Fig. 3 Few images of the FER 2013 Dataset [15]

B. Model Training and Hyperparameter Tuning

After getting the data and performing resizing and rescaling we went ahead and separated the data into training, testing, and validation sets. The dataset consisted of around 35 thousand images split into 6 different moods, Happy, Sad, Fear, Anger, Neutral, and Surprise. 80% of the data was set aside for training, 10% for validation, and 10% for testing. The first CNN architecture was with 6 Convolutional layers along with a Maxpooling layer and finally the flatten and dense layers. We used 'relu' as the activation function and for the dense layer, 'softmax' was used. The model was trained for 100 epochs and achieved an accuracy of 51.83%
Keras Tuner [19] is a library that helps you pick the optimal set of hyperparameters for your TensorFlow program. Upon using the Keras tuner we trained the model with the best model as given by the Keras tuner. An accuracy of 56.19 % was achieved, around 5% more than the first model.

C. Batch Normalization

It is the technique of adding extra layers to a deep neural network to make it faster and more stable [21]. The standardizing and normalizing procedures are performed by the new layer on the input of a previous layer.
Batch Normalization is a two-step process. First, the input is normalized, and later rescaling and offsetting are performed.
Normalization of the Input:
The process of altering data to have a mean of zero and a standard deviation of one is known as normalization. We have our batch input from layer h in this phase, and we must first calculate the mean of this hidden activation.

$$\mu = \frac{1}{m}\Sigma h_i \tag{1}$$

Here, m is the number of neurons at layer h.
Once we have meant at our end, the next step is to calculate the standard deviation of the hidden activations.

$$\sigma = \left[\frac{1}{m}\Sigma(h_i - \mu)^2\right]^{\frac{1}{2}} \tag{2}$$

Furthermore, we already have the mean and standard deviation. These values will be used to normalize the hidden activations. We'll do this by subtracting the mean from each input and dividing the result by the total of standard deviation and the smoothing term (ε).

$$h_{i(\text{norm})} = \frac{(h_i - \mu)}{\sigma + \varepsilon} \tag{3}$$

Rescaling of Offsetting
The final process consists of rescaling and offsetting the input γ (gamma) and β(beta) are two BN algorithm components that come into play here. These parameters are used to rescale (γ) and shift (β) the vector containing the results of preceding operations.

$$h_i = \gamma h_{i(\text{norm})} + \beta \tag{4}$$

These two parameters are learnable, and the neural network ensures that the best values of γ and β are chosen throughout training. This will allow each batch to be accurately normalized.

©IJRASET: All Rights are Reserved

D. Best Model Parameters

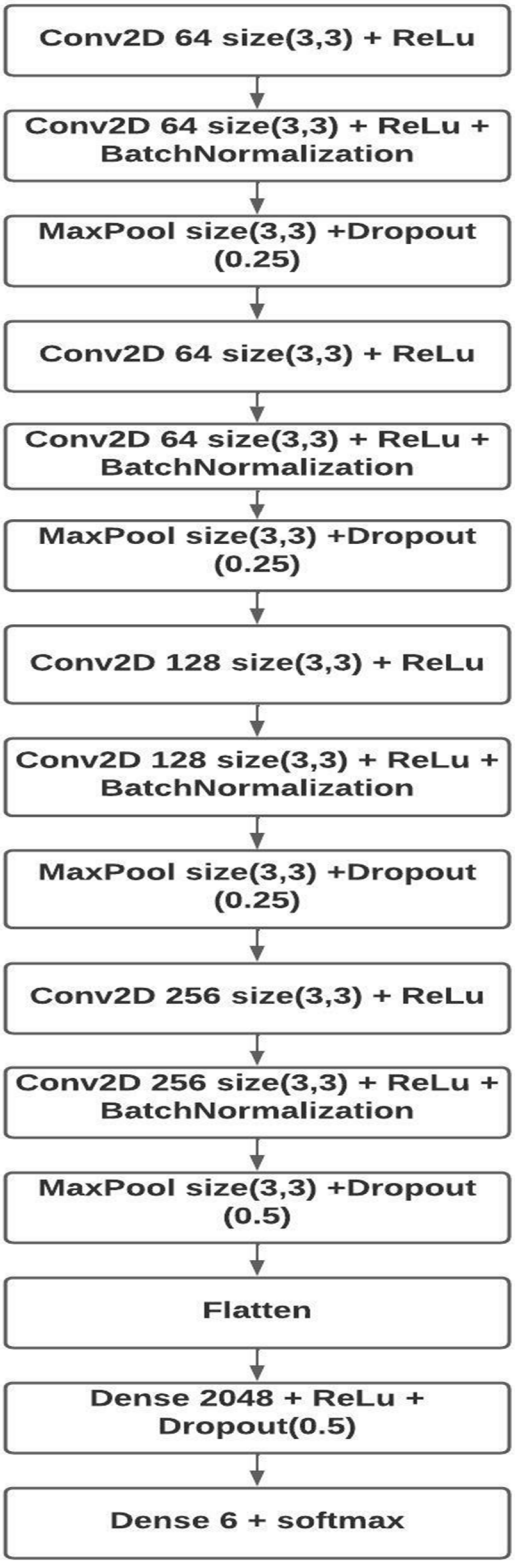


Fig. 4 CNN Architecture of Proposed Model

This model performed the best giving an accuracy of 88.3%

E. Flow of the Proposed System

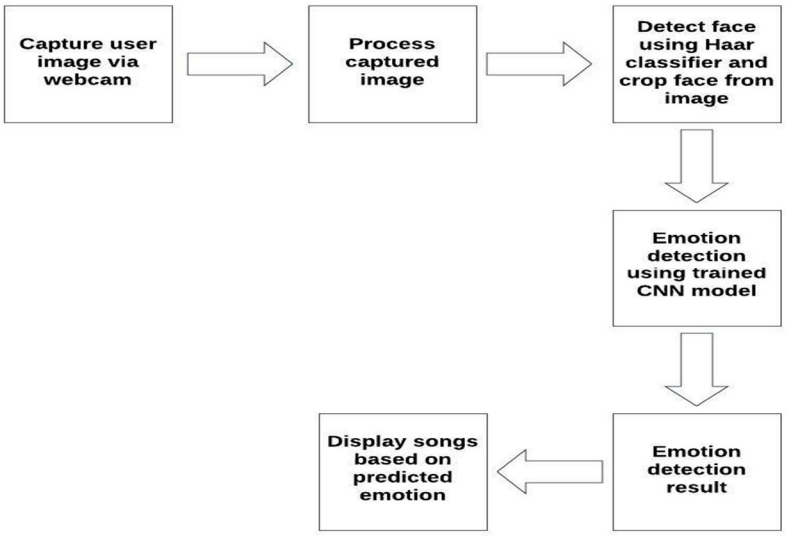


Fig. 5 Flow of the proposed system

A user will be made to capture his image using the webcam, with the help of Web RTC the user's webcam will be accessed to show his face, the user can decide when to click the capture button to capture the image, the image will be captured at that very instance and will be displayed on a canvas on the website. After the image has been acquired and the user is satisfied with it, the user can click the upload button, which will transmit the captured image to the server using AJAX. The image will be delivered as a base 64 string, which will be converted back to an image later. The image will then be sent to a program to detect if there is a face in the image if so, then just the face will be cropped from the captured image and then sent to the Trained CNN model which will then predict the emotion through the image. Haar cascade frontal face detection [22] is used to detect and adapt the face from the original image to remove unnecessary parts for each pipeline. Image resizing is applied to change the size of the input images to 48 × 48 for the CNN classifier. Once the emotion is predicted a playlist will be generated based on the emotion.

IV. RESULTS

To test the overall performance, the proposed model was trained for 100 epochs on the FER2013 database. For this final training the database was augmented in form of horizontal mirroring, ±10° rotations, ±10% image zooms, and ±10% horizontal and vertical shifting. In this stage, the learning rate was set the optimizer chosen was Adam with a learning rate of 0.001, and batch size was set to 64. Following the training, 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.

The reported measurement has been obtained in Kaggle on an NVIDIA TESLA P100 GPU.

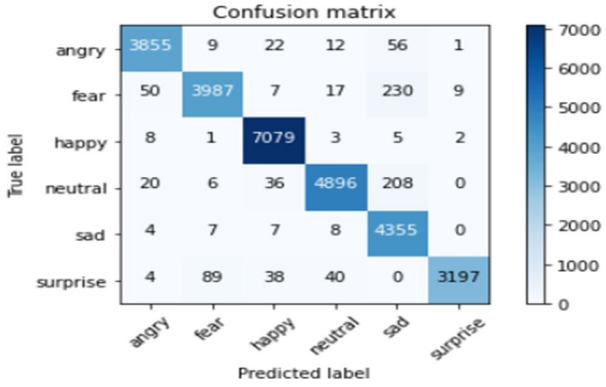


Fig. 6 Confusion Matrix for Training Data

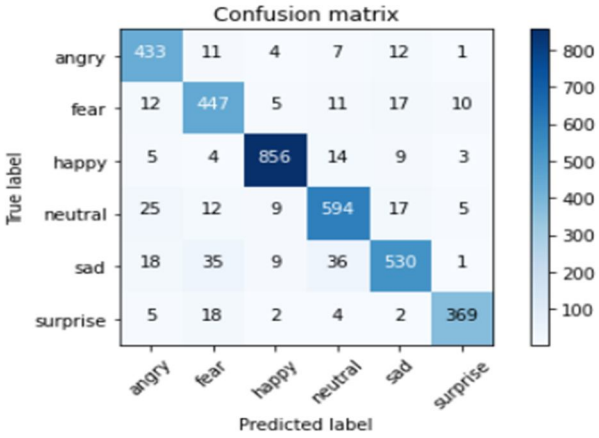


Fig. 8 Confusion Matrix for Testing Data

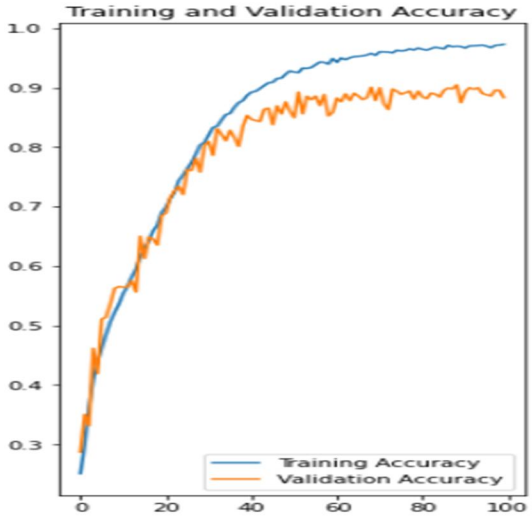


Fig. 8 Training and Validation Accuracy curve

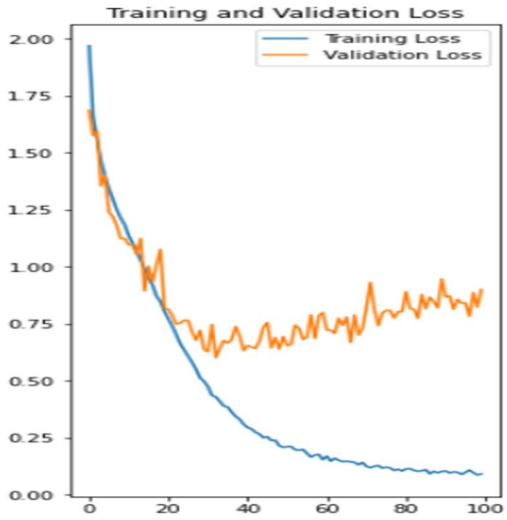


Fig. 9 Training and Validation Loss curve

TABLE I
RESULTS AND RANKING OF THE PROPOSED MODEL

MODEL	ACCURACY
N. Zhou et al. "A Lightweight Convolutional Neural Network for Real-Time Facial Expression Detection" [24]	87.71
Khanzada, Amil et al. "Facial Expression Recognition with Deep Learning." [16]	75.8
M. Georgescu et al. "Local Learning With Deep and Handcrafted Features for Facial Expression Recognition" [17]	75.42
S. Li and W. Deng, "Deep Facial Expression Recognition: A Survey" [18]	75.1
Adrian Vulpe-Grigorasi, Ovidiu Grigore "Convolutional Neural Network Hyperparameters Optimization for Facial Emotion Recognition" [1]	72.16
Proposed Model	88.3

From the table it can be seen that the difference in accuracy between the proposed model and few other models, the proposed model has only 3332806 parameters.



Fig. 10 Taking user's image as Input



Fig. 11 Mood Detected and Songs generated accordingly

V. CONCLUSION

The project proposes a system that some of the top music providers like Amazon, Spotify, etc do not provide as of yet. The main motive was to create a Music Playlist Generator basically by detecting the mood of the user just by their facial emotion. The old or the current system requires the user to manually go and search for a playlist based on his mood, our proposed system makes this task very simple by just capturing an image of the user and detecting the user's mood and generating a playlist according to the mood detected to enhance the user's mood. The Proposed CNN model with the help of batch normalization achieved an accuracy of 97.42% on the training data. However, 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.

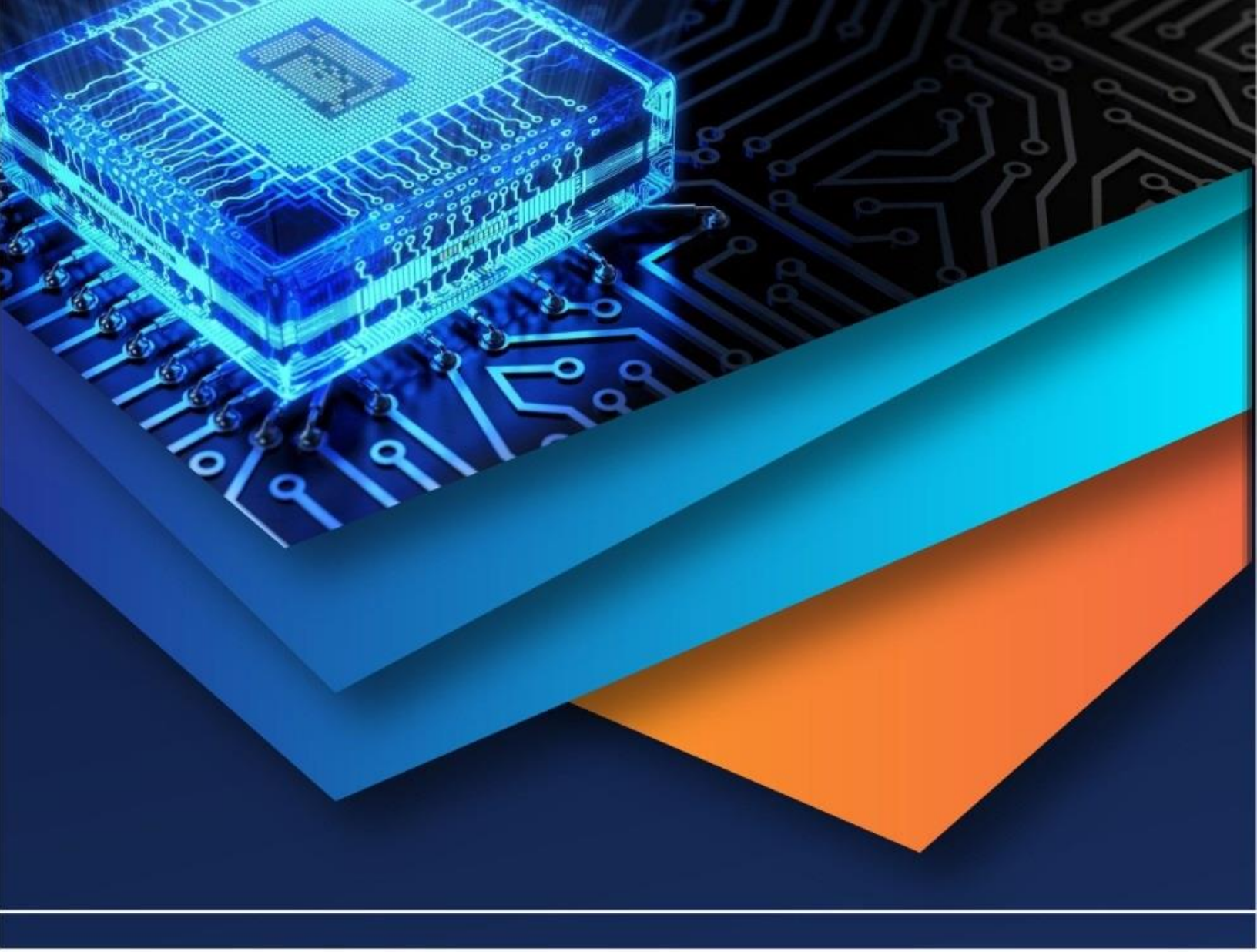
REFERENCES

[1] Adrian Vulpe-Grigorasi, Ovidiu Grigore "Convolutional Neural Network Hyperparameters Optimization for Facial Emotion Recognition", The 12th INTERNATIONAL SYMPOSIUM ON ADVANCED TOPICS IN ELECTRICAL ENGINEERING, 2021.
[2] Sharmeen M. Saleem Abdullah, Adnan Mohsin Abdulazeez "Facial Expression Recognition Based on Deep Learning Convolution Neural Network: A Review", Journal of Soft Computing and Data Mining, 2021.
[3] Krupa K S, Ambara G, Kartikey Rai, Sahil Choudhury "Emotion aware Smart Music Recommender System using Two Level CNN", Proceedings of the Third International Conference on Smart Systems and Inventive Technology, 2020.
[4] S Metilda Florence,M Uma "Emotional Detection and Music Recommendation System based on User Facial Expression", 3rd International Conference on Advances in Mechanical Engineering, 2020.
[5] Wafa Mellouk, Wahida Handouzi "Facial emotion recognition using deep learning: review and insights", The 2nd International Workshop on the Future of Internet of Everything (FIoE) 2020.
[6] CH.sadhvika, Gutta.Abigna, P.Srinivas reddy, Dr.Sunil Bhutada "EMOTION BASED MUSIC RECOMMENDATION SYSTEM", Journal of Emerging Technologies and Innovative Research (JETIR),2020.
[7] Luefeng Chen, Min Li, Xuzhi Lai, Kaoru Hirota, Witold Pedrycz "CNN-based Broad Learning with Efficient Incremental Reconstruction Model for Facial Emotion Recognition", IFAC PapersOnLine 53-2, 2020.
[8] MinSeop Lee, Yun Kyu Lee, Myo-Taeg Lim and Tae-Koo Kang "Emotion Recognition Using Convolutional Neural Network with Selected Statistical Photoplethysmogram Features" Applied Sciences MDPI, 2020.
[9] Deny John Samuvel, B. Perumal, Muthukumaran Elangovan "Music recommendation system based on facial emotion recognition", 3C Tecnolog´ıa. Glosas de innovaci´on aplicadas a la pyme, 2020.
[10] Prof.Sumeet Pate, Miss.Shreya Zunjarrao, Miss.Poonam Harane, Mr.Akshay Choudhary "Mood Based Music Player Using Real Time Facial Expression Extraction", International Journal for Research in Engineering Application & Management (IJREAM), 2019.
[11] Ahlam Alrihaili, Alaa Alsaedi, Kholood Albalawi, Liyakathunisa Syed "Music Recommender System for users based on Emotion Detection through Facial Features", Developments in eSystems Engineering (DeSE), 2019.
[12] Swapnil V Dhatrak, Ashwini G Bhange, Abhishek K Sourabh, Prof. Yogesh A Handge "A Survey Emotion Based Music Player through Face Recognition System" International Journal of Innovative Research in Science, Engineering and Technology, 2019.
[13] https://en.wikipedia.org/wiki/Music
[14] https://en.wikipedia.org/wiki/Facial expression
[15] https://www.kaggle.com/msambare/fer2013
[16] Khanzada, Amil et al. "Facial Expression Recognition with Deep Learning." ArXiv abs/2004.11823, 2020.
[17] Georgescu, R. T. Ionescu and M. Popescu, "Local Learning With Deep and Handcrafted Features for Facial Expression Recognition," in IEEE Access, vol. 7, 2019.
[18] S. Li and W. Deng, "Deep Facial Expression Recognition: A Survey," IEEE Transactions on Affective Computing, 2020.
[19] https://www.tensorflow.org/tutorials/keras/keras_tuner
[20] https://www.jeremyjordan.me/hyperparameter-tuning/
[21] https://www.analyticsvidhya.com/blog/2021/03/introduction-to-batch-normalization/
[22] https://docs.opencv.org/3.4/db/d28/tutorial_cascade_classifier.html
[23] V. Christlein, L. Spranger, M. Seuret, A. Nicolaou, P. Král, and A. Maier, "Deep Generalized Max Pooling," arXiv Prepr., vol. 1908.05040, pp. 1–7, 2019.
[24] N. Zhou, Renyu Liang, and Wenqian Shi, "A Lightweight Convolutional Neural Network for Real-Time Facial Expression Detection," 2020.
[25] C. Zhang, P. Wang, K. Chen, and J. K. Kämäräinen, "Identity-Aware Convolutional Neural Network for Facial Expression Recognition," J. Syst. Eng. Electron., vol. 28, no. 4, pp. 784–792, 2017
[26] M. Sun, Z. Song, X. Jiang, J. Pan, and Y. Pang, "Learning Pooling for Convolutional Neural Network," Neurocomputing, vol. 224, pp. 96–104, 2017
[27] A. Ruiz-Garcia, M. Elshaw, A. Altahhan, and V. Palade, "Stacked deep convolutional auto-encoders for emotion recognition from facial expressions," Proceedings of the 2017 International Joint Conference on Neural Networks (IJCNN 2017), 2017
[28] A. T. Lopes, E. de Aguiar, A. F. De Souza, and T. OliveiraSantos, "Facial expression recognition with Convolutional Neural Networks: Coping with few data and the training sample order," Pattern Recognit., vol. 61, pp. 610–628, 2017

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

[29] A. Majumder, L. Behera, and V. K. Subramanian, "Automatic Facial Expression Recognition System Using Deep Network-Based Data Fusion," IEEE Trans. Cybern., vol. 48, no. 1, pp. 103–114, 2018
[30] S. Alizadeh and A. Fazel, "Convolutional Neural Networks for Facial Expression Recognition," arXiv Prepr., vol. 1704.06756, pp. 1–8, 2017
[31] T. Chang, G. Wen, Y. Hu, and J. Ma, ''Facial expression recognition based on complexity perception classification algorithm,'' 2018
[32] M. Z. Uddin, W. Khaksar, and J. Torresen, ''Facial expression recognition using salient features and convolutional neural network,'' IEEE Access, vol. 5, pp. 26146–26161, 2017
[33] D. Hazarika, S. Gorantla, S. Poria, and R. Zimmermann, ''Self-attentive feature-level fusion for multimodal emotion detection,'' in Proc. IEEE Conf. Multimedia Inf. Process. Retr. (MIPR), 2018
[34] K.-Y. Huang, C.-H. Wu, Q.-B. Hong, M.-H. Su, and Y.-H. Chen, ''Speech emotion recognition using deep neural network considering verbal and nonverbal speech sounds,'' in Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), 2019
[35] D. Kollias, P. Tzirakis, M. A. Nicolaou, A. Papaioannou, G. Zhao, B. Schuller, I. Kotsia, and S. Zafeiriou, ''Deep affect prediction inthe-wild: Aff-wild database and challenge, deep architectures, and beyond,'' Int. J. Comput. Vis., vol. 127, pp. 1–23, 2019
[36] E. Marinoiu, M. Zanfir, V. Olaru, and C. Sminchisescu, ''3D human sensing, action and emotion recognition in robot assisted therapy of children with autism,'' in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2018
[37] Z. Du, S. Wu, D. Huang, W. Li, and Y. Wang, ''Spatio-temporal encoderdecoder fully convolutional network for video-based dimensional emotion recognition,'' IEEE Trans. Affect. Comput., early access, 2019
[38] P. Barros, N. Churamani, and A. Sciutti, ''The FaceChannel: A light-weight deep neural network for facial expression recognition,'' in Proc. 15th IEEE Int. Conf. Autom. Face Gesture Recognit. (FG), 2020
[39] M. Koujan, L. Alharbawee, G. Giannakakis, N. Pugeault, and A. Roussos, ''Real-time facial expression recognition 'in the wild' by disentangling 3D expression from identity,'' in Proc. 15th IEEE Int. Conf. Autom. Face Gesture Recognit. (FG), 2020
[40] G. Zeng, J. Zhou, X. Jia, W. Xie, and L. Shen, ''Hand-crafted feature guided deep learning for facial expression recognition,'' in Proc. 13th IEEE Int. Conf. Autom. Face Gesture Recognit. (FG), 2018

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)