IARJSET

ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Vol. 8, Issue 7, July 2021

DOI: 10.17148/IARJSET.2021.8756

Facial Emotion Recognition and Detection in Python Using Deep Learning

Nikhil Rai[1], Dhirender Gulair[2], Jawad Shiningwala[3], Mayuri H. Molawade[4]

Students, Department of Computer Engineering, Bharati Vidyapeeth (Deemed to be University) College of Engineering, Pune, India-411043[1,2,3]

Assistant Professor, Department of Computer Engineering, Bharati Vidyapeeth (Deemed to be University) College of Engineering, Pune, India-411043[4]

Abstract: Human facial emotion recognition (FER) has attracted the eye of the research network for its promising applications. Mapping one of a kind facial expressions to the respective emotional states are the primary task in FER. The classical FER consists of two most important steps: feature extraction and emotion recognition. presently, the Deep Neural Networks, particularly the Convolutional Neural network (CNN), is extensively used in FER with the aid of distinctive feature of its inherent feature extraction mechanism from pictures. numerous works were reported on CNN with only some layers to clear up FER issues. but, wellknown shallow CNNs with straightforward getting to know schemes have restricted characteristic extraction capability to seize emotion data from high-resolution pictures. A notable disadvantage of the most current techniques is that they consider only the frontal pictures (i.e., ignore profile perspectives for convenience), despite the fact that the profile perspectives taken from different angles are essential for a practical FER system. For growing a highly correct FER system, this study proposes a completely Deep CNN (DCNN) modeling thru transfer learning (TL) technique wherein a pre-skilled DCNN model is followed through changing its dense top layer(s) well suited with FER, and the model is great-tuned with facial emotion data. a novel pipeline strategy is brought, wherein the training of the dense layer(s) is accompanied via tuning each of the pre-skilled DCNN blocks successively that has brought about gradual improvement of the accuracy of FER to a better level.

Keywords: convolutional neural network (CNN); deep CNN; emotion recognition; transfer learning

I. INTRODUCTION

Human beings regularly have different moods and facial expressions changes consequently. Human emotion recognition performs a completely crucial role in social relations. the automated recognition of emotions has been an active analysis subject matter from early eras. in this deep learning system user's emotions using its facial expression can be detected. real-time detection of the face and deciphering different facial expressions like happy, sad, angry, afraid, surprise, disgust, and neutral. and many others. This system can locate six different human emotions. The trained model is capable to hit upon all of the noted emotions in real-time. an automated facial expression recognition system has to carry out detection and site of faces throughout a cluttered scene, facial feature extraction, and facial expression classification. The facial expression recognition system is enforced victimization of Convolution Neural network (CNN). A CNN model is trained on FER2013 dataset. FER2013 Kaggle faces expression dataset with six facial features labels as happy, sad, surprise, fear, anger, disgust, and neutral is used during this project. as compared to the alternative datasets, FER has extra variant within the pictures, which includes face occlusion, partial faces, low-contrast pictures, and eyeglasses. This system has capability to monitor human beings emotions, to discriminate among emotions and label them accurately and use that emotion information to guide thinking and behavior of specific individual.

II. LITERATURE REVIEW

Facial expression is the common signal for all humans to convey the mood. There are many attempts to make an automatic facial expression analysis tools as it has application in many fields such as robotics, medicine, driving assist systems, and lie detector. Since the twentieth century, Ekman et al. defined seven basic emotions, irrespective of culture in which a human grows with the seven expressions (anger, feared, happy, sad, contempt, disgust, and surprise). In a recent study on the facial recognition technology (FERET) dataset, Sajid et al. found out the impact of facial asymmetry as a marker of age estimation. Their finding states that right face asymmetry is better compared to the left face asymmetry. Face pose appearance is still a big issue with face detection. Ratyal et al. provided the solution for variability in facial pose appearance. They have used three-dimensional pose invariant approach using subject-specific descriptors. There are many issues like excessive makeup pose and expression which are solved using convolutional

IARJSET
ISSN (O) 2393-8021, ISSN (P) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
Vol. 8, Issue 7, July 2021
DOI: 10.17148/IARJSET.2021.8756

networks. Recently, researchers have made extraordinary accomplishment in facial expression detection, which led to improvements in neuroscience and cognitive science that drive the advancement of research, in the field of facial expression. Also, the development in computer vision and machine learning makes emotion identification much more accurate and accessible to the general population. As a result, facial expression recognition is growing rapidly as a sub-field of image processing. Some of the possible applications are human–computer interaction, psychiatric observations, drunk driver recognition, and the most important is lie detector.

III. EXISTING SYSTEM

In Emotion recognition using mind activity the developer Robert Horling's has used mind activities that is toughest challenge to do because it turn out to be expensive, complex and additionally time consuming when we attempt to measure human mind with Electroencephalography. they've used existing data and the result of their analysis were 31 to 81 percent accurate and from which even through the usage of Fuzzy logic 72 to 81 percentage only for two classes of emotions. seemingly, the division among different emotions isn't always (only) based totally on EEG signals it depends on some another.

IV. PROPOSED SYSTEM

Considering the anomalies within the existing system computerization of the entire activity is being recommended after initial analysis. This application of feature extraction of facial expressions with combination of neural network for the recognition of various facial emotions (happy, sad, angry, fear, surprised, neutral and so on..). human beings are capable of generating heaps of facial moves at some point of verbal exchange that fluctuate in complexity, intensity, and meaning. This paper analyses the limitations with existing system Emotion recognition using brain activity. In this paper by using an existing simulator i've achieved ninety seven percent accurate outcomes and it is simple and simplest way than Emotion recognition using brain activity system. Purposed system relies upon human face as we understand face also reflects the human brain activities or emotions. In this paper neural network has been used for higher outcomes. in the end of paper comparisons of existing Human Emotion recognition system has been made with new one.

A. CNN Architecture

The networks are program on pinnacle of keras, operating on Python, using the keras learn library. This surroundings reduces the code's complexity, because only the neuron layers want to be formed, instead of any neuron. The software additionally gives real-time remarks on training progress and overall performance, and makes the model after training easy to save and reuse. In CNN architecture initially we need to extract input picture of 48*48*1 from dataset FERC-2013. The network starts with an input layer of 48 via 48 which suits the input data size parallelly processed via comparable models that is functionality in deep learning, after which concatenated for better accuracy and getting features of pictures flawlessly as shown in Fig.1 which is our proposed model,**Model-A.** There are sub models for the extraction of CNN features which share this input and both have same kernel size. The outputs from these feature extraction sub-models are flattened into vectors and concatenated into one lengthy vector matrix and transmitted to a completely connected layer for evaluation before a final output layer permits for classification.

This models consists of convolutional layer with 64 filters each with size of [3*3], accompanied by using a local contrast normalization layer, maxpooling layer, followed by using one more convolutional layer, max pooling, flatten respectively. After that we concatenate two comparable models and linked to a softmax output layer which can classify seven emotions. We use dropout of 0.2 for decreasing over-fitting. it's been applied to the completely connected layer and all layers include units of rectified linear units (ReLu) activation function.

First we are passing our input picture to convolutional layer which includes 64 filters each of size 3 by 3, after that it passes via local contrast normalization can remove average from neighbourhood pixels leads to get quality of feature maps, accompanied by ReLu activation function. most pooling is used to reduce spatial size reduction so processing pace will increase. we're using concatenation for getting functions of pictures (eyes, eyebrows, lips, mouth and so on) flawlessly so that prediction accuracy progressed as compared to previous model. moreover, it is followed by completely connected layer and softmax for classifying seven emotions. A 2nd layer of maxpooling is introduced to reduce the quantity of dimensionality. right here, we use batch normalization, dropout, ReLu activation function, categorical cross entropy loss, adam optimizer, softmax activation function in ouput layer for seven emotion classification.

IARJSET

ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Vol. 8, Issue 7, July 2021

DOI: 10.17148/IARJSET.2021.8756

In **Model-B**, previously proposed by Correa et al , the network begins with a 48 by 48 input layer, which fits the dimensions of the input data. this layer is preceded by way of one convolutional layer, a local contrast normalization layer, and one layer of maxpooling, respectively. more convolutional layers and one completely linked layer, linked to a softmax output layer, complete the network. Dropout has been implemented to the completely connected layer and all layers contain units of ReLu.

V. EXPERIMENT DETAILS

To evaluate the 2 models (Model-A and Model-B) stated above on their emotion detection capability. This phase describes the data used for training and testing, explains the information of the used data sets and evaluates the outcomes received using two different datasets with two models.

A. Datasets

Neural networks, and specifically deep networks, needs massive amounts of training data. further, the selection of pictures used for the training is liable for a huge part of the eventual model's overall performance. It means the need for a data set that is both high quality and quantitative. numerous datasets are to be had for research to understand emotions, starting from a few hundred high resolution pictures to tens of heaps of smaller pictures. the 2, we will be debating in this work, are the japanese female Face Expression (JAFFE), facial expression recognition challenge (FERC-2013) which includes seven emotions like anger, surprise, happy, sad, disgust, fear, neutral.

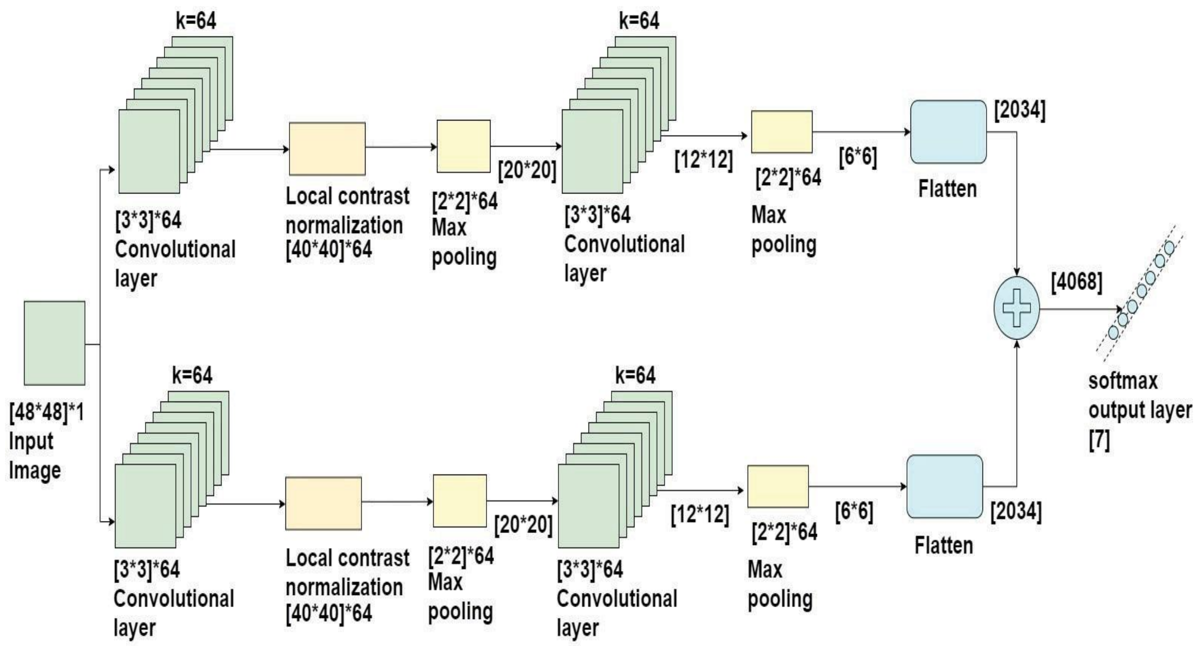

Fig.1:- Network Architecture

B. Training Details

We train the network using GPU for 100 epochs to make sure that the precision converges to the ultimate. The network might be trained on a bigger set than the one formerly described in an try to enhance the model even more. training will take place with 20,000 pictures from the FERC-2013 dataset rather than 9,000 pictures. The FERC-2013 database also makes use of newly designed verification (2000 pictures) and sample sets (1000 pictures). It indicates number of emotions within the final testing and validation set after training and testing our model. The accuracy might be better on all validation and test sets than in previous runs, emphasizing that emotion detection using deep convolutional neural networks can enhance the overall performance of a network with greater information.

C. Results Using Proposed Model

In emotion detection we're using 3 steps, i.e., face detection, features extraction and emotion classification using deep learning with our proposed model which offers higher end result than preceding model. in the proposed method,

IARJSET

ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Vol. 8, Issue 7, July 2021

DOI: 10.17148/IARJSET.2021.8756

computation time reduces, validation accuracy will increase and loss also decreases, and in addition performance evaluation achieved which compares our model with preceding existing model. We tested our neural network architectures on FERC-2013 and JAFFE database which incorporates seven primary emotions like sad, fear, happiness, angry, neutral, surprised, disgust.

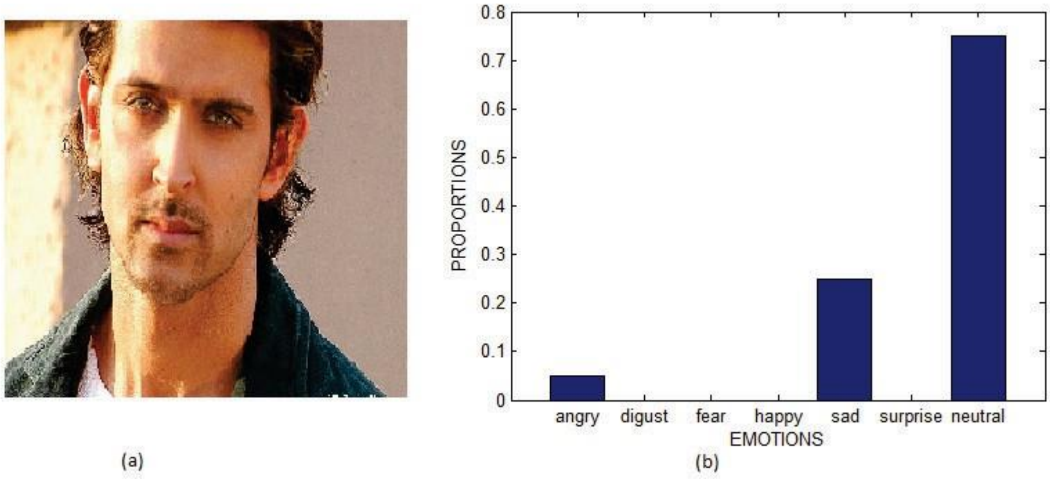

Fig. 2: (a) Image, (b) Proportion of emotions

Fig.2 suggests the proportions of detected emotions in a single picture of FER dataset. Fig.2(a) indicates the picture, while the detected emotion proportions are shown in Fig.2(b). it's far clearly observable that neutral has better proportion than other emotions. Which means, the emotion detected for this picture (in Fig.2(a)) is neutral.

Similarly, performance is evaluated for all of the test pictures of the dataset. We've achieved 75 percent for neutral, 69 percent for sad, 68 percent for surprise, 63 percent for disgust, 65 percent for fear and 56 percent for angry. On a mean we are getting average accuracy of 70.14 percent using our proposed model.

The confusion matrix of classification accuracy is proven in table I. We get a mean validation accuracy of 70.14 percent using our proposed model in facial emotion detection using the FER dataset.

Fig. 4: Prediction of emotion.

Fig.4 indicates the result of test sample associated with surprise emotion from JAFFE dataset, and our proposed model additionally predicted the identical emotion with decreased computation time in comparison to preceding current model B. further, performance is evaluated for all of the test samples of JAFFE dataset. when we're using JAFFE dataset we are getting validation accuracy of 98.65 percent that is higher than preceding result and it takes much less computational time per step.

DOI: 10.17148/IARJSET.2021.8756

VI. FUTURE SCOPE

As human facial expression recognition is a very fundamental technique, it is beneficial to assess the mood or emotional state of a topic underneath observation. As such, super ability lies untapped in this area. The simple concept of a machine being able to recognize the human emotive state may be put to use in innumerable eventualities.

VII. CONCLUSION

This was our project of system design about “Facial Emotion recognition and Detection” developed in Django based on Python programming language. The development of this system takes numerous efforts from us. We assume this system gave numerous delight to everyone. although every task is never said to be ideal in this development field even more improvement can be viable in this application. We discovered so many things and received plenty of information about development field. we are hoping this could prove fruitful to us.

VIII. REFERENCES

1) http://ieeexplore.ieee.org/document/7013165/
2) https://www.analyticsvidhya.com/blog/2018/08/a-simple-introductionto-facial-recognition-with-python-codes
3) https://realpython.com/face-recognition-with-python/
4) http://web.stanford.edu/class/ee368/Project_Autumn_1617/Reports/report_pao.pdf
5) https://sci-hub.se/10.1109/INCET49848.2020.9154121
6) https://link.springer.com/article/10.1007/s42452-020-2234-1