Proceedings of the Fifth International Conference on Intelligent Computing and Control Systems (ICICCS 2021)
IEEE Xplore Part Number: CFP21K74-ART; ISBN: 978-0-7381-1327-2

# Design and Evaluation of a Deep Learning Algorithm for Emotion Recognition

R Raja Subramanian
*Dept of Computer Science and Engineering*
*Kalasalingam Academy Of Research And Education*
Krishnankoil, TamilNadu, India.
rajasubramanian.r@klu.ac.in

Chunduri Sandya Niharika
*Dept of Computer Science and Engineering*
*Kalasalingam Academy Of Research And Education*
Krishnankoil, TamilNadu, India.
sandyachunduri05@gmail.com

Dondapati Usha Rani
*Dept of Computer Science and Engineering*
*Kalasalingam Academy Of Research And Education*
Krishnankoil, TamilNadu, India.
urani4044@gmail.com

Parvathareddy Pavani
*Dept of Computer Science and Engineering*
*Kalasalingam Academy Of Research And Education*
Krishnankoil, TamilNadu, India.
parvathareddypavani@gmail.com

Ketepalli Poojita Lakshmi Syamala
*Dept of Computer Science and Engineering*
*Kalasalingam Academy Of Research And Education*
Krishnankoil, TamilNadu, India.
poojitakkr@gmail.com

***Abstract*—Facial emotion recognition is one of the most interesting research areas where many researchers are actively participating over the past few decades. This paper attempts to discuss about the application of emotion recognition where seven different emotions such as happy, sad, neutral, angry, surprise, fear and disgust are obtained. Humans can produce thousands of emotions in different situations which have different meanings, intensities and complexities. By using convolutional neural network (CNN) algorithm, an accuracy of about 89% has been achieved. It is the simplest way of all. For better results deep learning and neutral networks have been used. Our proposed deep learning model helps us in focusing important features in humans face to detect emotion using multiple datasets such as FER-2013 and image dataset.***

**Keywords:** ***Emotion Recognition, Neural Networks, Dataset, Types of Emotions.***

## I. INTRODUCTION

Emotions are an important part of any interpersonal communication. In every human's life emotions plays an important role. Humans are efficient in producing lot of emotions during their daily life at different situations that hold disparate meanings and intensities. At several situations humans plays different emotions based on their mood. They can be shown in disparate forms which may or may not be noticed visually. Therefore, with the help of accurate tools, the emotions are detected and recognized. For detecting emotions in face, facial emotion recognition process is used. Humans vary widely in their accuracy at finding others emotion. So, here are using technology to recognize the emotion. Emotion can be expressed through face, speech, electroencephalogram (EEG) and via text. Out of these, facial expressions are in favour as they are visible to human eye.

Emotion Recognition supports us in sensing emotions in humans at diverse situations. It is classified into 3 stages they are Face Detection, Feature Extraction, Emotion Classification. Face detection is a preprocessing stage which helps in detection emotions, During Feature extraction various features of the face is extracted, In the final stage it produces labels and model will be trained.

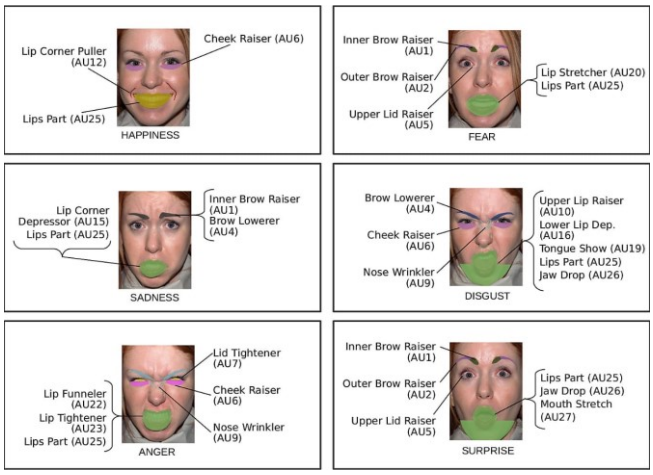


Fig 1. Emotion Recognition Through Facial Expressions.

However, with growing connections of humans with machines, researchers now aim at constructing deep learning models that can perfectly analyze affective content in conversations. This can help in creating empathetic facial emotion recognition systems, thus improving the overall human-machine interaction experience.

This paper presents a deep learning-based algorithm i.e., Convolutional Neural Networks (CNN) to detect emotions in face. As in Fig 1, it has been shown by using convolutional neural network, even a network with few layers (less than 10 layers) we can achieve a good accuracy rate. To develop a facial expression recognition system. To experiment machine learning algorithm in computer vision fields. To detect emotion thus facilitating Intelligent Human-Computer Interaction.

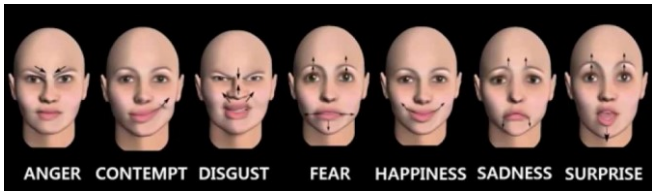


Fig 2. Different Types of Facial Expressions

2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS) | 978-1-6654-1272-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICICCS51141.2021.9432336

## II. PROBLEM STATEMENT

In the set of thousands of emotions, 7 basic human emotions are happy, sad, anger, disgust, neutral, fear, and surprise. As shown in Fig 2 These emotions contain lot of information about their state of mind. When humans are communicating each other, they produce various emotions based on their situations [2].

This application can be applied in health cares before the treatment of the patient. By this, the additional information of the patient can be obtained to provide better service during the treatment. Criminals and terrorists are always found tensed and angry. So, the criminals and terrorists can be easily found in heavy crowds based on their facial expressions.

Facial emotion recognition is important for the interpretation of facial expressions. The facial expressions are useful for efficient interaction. It is found that it is insufficient to describe all facial expressions and these expressions are categorized based on facial actions. Detecting face and recognizing the facial expression is one of the complex tasks when it is an essential to show attention to primary elements like: face configuration, orientation, location where the face is set.

## III. DATABASE

In this paper, we provide the work of experimental analysis of the proposed model on various popular facial expression datasets, such as FER2013 DATASET and IMAGE DATASET. Before moving into the results, we are going to give a brief overview of our datasets.

FER2013:

The Facial Expression Recognition 2013 (FER2013) database was first introduced in the ICML 2013 Challenges in Representation Learning. Our dataset consists of different 48x48 pixels of images which shows various emotions. This dataset contains 3 columns such as emotion, pixels and usage. In emotion column there are some numbers which indicates different emotions i.e., 0 refers to Angry, 1 refers to Disgust, 2 refers to Fear, 3 refers to Happy, 4 refers to Sad, 5 refers to Surprise, 6 refers to Neutral. In pixel column contains each image which is having a string values surrounded in quotes. Test dataset contains only pixels, but we need to detect emotion column values. This dataset carries more than 35,000 examples with 48x48 resolution, out of them most of those are taken in wild settings. Originally the training set contains 28,709 examples, and the testing set there are two categories such as public test and private test each consists of 3,589 examples.

| | emotion | pixels | Usage |
|---|---|---|---|
| 0 | 0 | 70 80 82 72 58 58 60 63 54 58 60 48 89 115 121... | Training |
| 1 | 0 | 151 150 147 155 148 133 111 140 170 174 182 15... | Training |
| 2 | 2 | 231 212 156 164 174 138 161 173 182 200 106 38... | Training |
| 3 | 4 | 24 32 36 30 32 23 19 20 30 41 21 22 32 34 21 1... | Training |
| 4 | 6 | 4 0 0 0 0 0 0 0 0 0 0 0 0 3 15 23 28 48 50 58 84... | Training |

Fig 3. Sample Dataset

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 35887 entries, 0 to 35886
Data columns (total 3 columns):
 #   Column   Non-Null Count  Dtype
---  ------   --------------  -----
 0   emotion  35887 non-null  int64
 1   pixels   35887 non-null  object
 2   Usage    35887 non-null  object
dtypes: int64(1), object(2)
memory usage: 841.2+ KB
None
```

Fig 4. Information of the Fer2013 Dataset

Fig 3 and Fig 4 depicts the detail information of dataset.

In FER dataset the images have more variations, including face occlusion (mostly with hand), partial faces, low-contrast images, and images with spectacles.

IMAGE DATASET:

As shown in Fig 5 It contains 28,000 images which Includes Seven Different Emotions Such as Angry, Fear, Disgust Happy, Neutral, Sad and Surprise. Some of the sample images of the date set is exposed in Fig 5.


Fig 5. Some Sample pictures of the Image Dataset

## IV. PROPOSED METHODOLOGY

In our model we use CNN and activation layer Softmax as output layer because with the help of Softmax layer we get more accuracy. CNNs are neural networks which are sensitive to spatial information. It is capable of recognition complex shapes and patterns in a given image. They are used for a wide variety of image classification algorithms. The CNNs for detecting facial emotions was built using Keras. The final layer of the CNN is actually changed and replaced with Softmax layer as seven output layers in our dataset i.e., FER2013 dataset. The Seven is actually a variable parameter changes with respect to dataset [3].

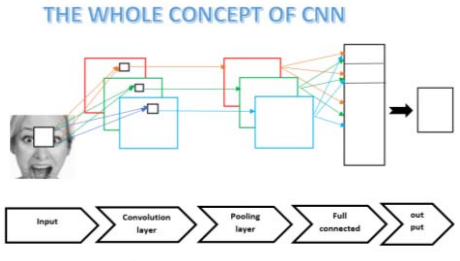


Fig 6. The concept Of Convolutional Neural Network (CNN)

### *A. Implementation*

Import the required libraries for building the model. Import NumPy for numerical calculations, Pandas for reading csv file and Keras for developing and evolution of deep learning models and the dataset will be imported. Here, this research work also import some layers, some Keras library like dense, dropout, Conv2D, Maxpooling2D.Fig 9 shows the steps to follow for building CNN model.

After importing the dataset, the first step is preprocessing. In preprocessing we are going to split the pixels i.e., in a row pixel. In try block usage column is training then we can append the pixel values into array and emotion values into an array. The same process is carrying for testing also. For building CNN with Keras we are going to convert the values into np.array form.

Second step is normalizing the data. Generally, normalization is done to scale down the value 0 to 1. Normalization helps us to train the data fast, after normalization each pixel value will be in between 0 to 1. Here we are going to subtract all the values from the mean of the data, so that average mean value becomes zero to the entire dataset from the axis zero and find the standard deviation of the dataset. The same process repeats for test dataset also.

Third step is reshaping the data. Now our dataset is in the form of array. We have to reshape our data into a format such that every image of dataset will be of same shape so that, Keras can accept the data. Here we are reshaping our images into 48 x 48 pixels. Now we are going to design neural network model using CNN algorithm with the help of Keras. First step is initializing the sequential model.

In this model we are having 3 convolutional layers, in first convolutional layer we have 2 Conv2D layers with 64 neurons, kernel size 3,3, Activation function is Relu. The process starts with the initialization of model accompanied by batch normalization layer followed by different convolutional layers with activation function ReLu. The derivative of ReLu function will be either 0 or 1 so, there will be some difference in the weight updation so, our gradient descent converges.

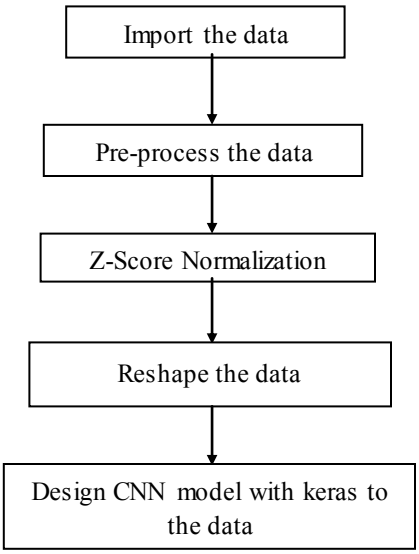


Fig 9. Steps to follow for Building CNN to data

The following are the layers of CNN:

**Convolutional Layer:** This layer is used to get features from images. Here features mean edges, intersect points which provides rich information. The number of convolution layers matter.

By using different activation functions with different number of features, we can change the architecture.

This network will have the following components:

**Activation functions:** Here in our model, we use two Activation functions i.e., Relu and Softmax which are applied to all the layers of output.

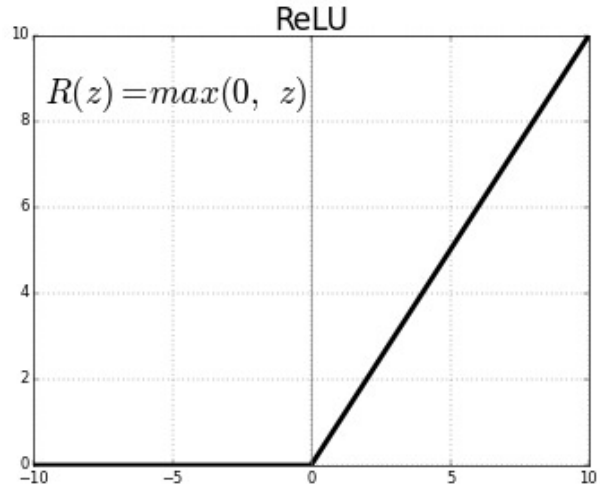


Fig 7. ReLu Activation Function

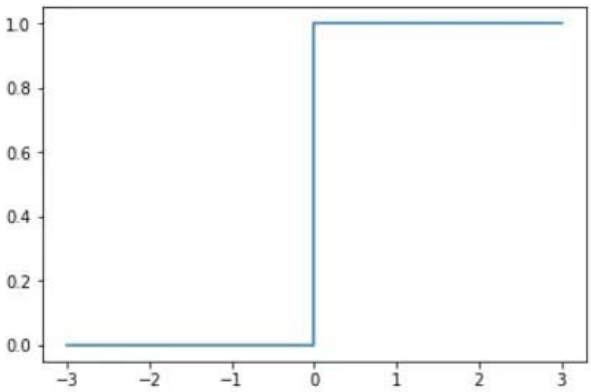


Fig 8. Derivative of ReLu Activation Function

**Pooling Layers:** It is added after the convolution layer which performs continues dimensionality reduction i.e., reduced number of parameters and computations thereby shortening training time and controls overfitting.

**Dense layers:** It is the bottom most layer of CNN. In this layer all feature data produced by convolution layers will take and do the decision making.

**Dropout Layers:** We have a deep neural network we will have many weight parameters and bias parameters due to this model may tends to overfit the dataset. For preventing overfitting of irregular turns of neurons in the network.

In dropout layer we select subset of features from input layer and subset of neurons from hidden layer, based

on the given p value some neurons and features get deactivated and some are activated.

$$\text{Dropout ratio: } 0 \leq p \leq 1$$

**Batch Normalization:** It is the method for training neural network models as the batch normalization increases the no of epochs required to train the deep neural network model decreases. It allows every layer of model to learn independently. This helps to speed up the training process.

After this, we compile the model using Adam as an optimizer, loss as categorical cross-entropy, and metrics as accuracy. After building and compiling the model, we split the data into training data and validation data. In our model we are taking the batch size as 64 with 100 epochs.

Once the training has been done, we can evaluate the model and compute loss and accuracy. We now save the model to fer.json and save the model weights in an fer.h5 file. For making predictions instead of training the model again we can make use of this file. Now by using OpenCV and webcam we test the model in Real-time manner. For building our model we use jupyter notebook. You can use many python platforms and IDEs to build the model and detect the emotion in real-time with the help of system webcam. We need to import the required libraries, then model weights which we saved earlier after training are loaded. once the model weights are loaded then we need to import haarcascade file which was build using OpenCv to detect facial emotion. Once we import the haarcascade file we had written a code to detect the faces and classify into certain emotions. The certain emotions are labelled as angry, happy, sad, surprise, neutral, disgust and fear. If we run the code it will connect to webcam and through webcam, we will detect the emotion in the face with a blue colour square box around the face [4]. To stop the code, you need to press 'q'. Anything we want that can place here to stop the execution.

## V. Experiment Results

The accuracy of the trained model for respective epochs is given in the table ACCURACY RESULTS. The best modern deep learning model outperforms the best simplistic model by almost 2% under undistinguishable conditions All measured architectures outperform the best simplistic model, including Inauguration, which has only half as many parameters. These accuracy results tell us that modern deep learning architectures model has the potential to significantly improve FER performance. Our distinct CNNs already perform competitively to previous works that exploit collaborative voting. By forming a group of 8 such CNNs, we achieve a FER2013 test accuracy of 89%. Our group of deep learning models attains the art performance of the state without operating further training data or structures, wide-ranging data expansion or demanding face registration. Deprived of any face registration, our Facial emotion recognition process is theoretically modest than earlier techniques and not caused by registration blunders. By exploiting secondary training data and wide ranging, FER definite data expansion would progress the performance additional.

In this paper we use FER2013 dataset which is the most communal and widely obtainable image dataset for CNN based Facial Emotion Recognition field. Due to various dataset the FER performance is indicative of real-world.

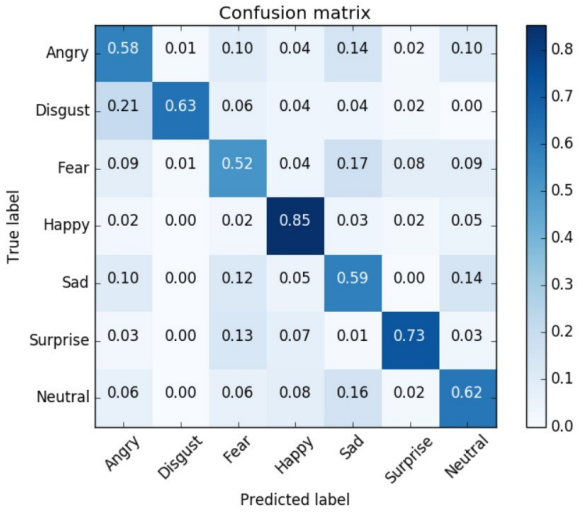


Fig 9. The confusion matrix for test set of FER Dataset model

**Parameters:** 3 convolutional layers, ReLu and softmax Activation functions, dense layer as output layer,7 outputs

ACCURACY RESULTS

| S.no | Epochs | Accuracy |
|---|---|---|
| 1 | 30 | 66.92% |
| 2 | 70 | 79.62% |
| 3 | 100 | 84.29% |

For image dataset we got an accuracy of 53 percent for 150 epochs. For increase in more no of epochs we get more accuracy.

## VI. Conclusion

The primary theme of our research paper is to build an CNN model to detect different facial emotions in real-time manner. The work done by us made the model to work in a better way such that it can recognize various emotions in different illumination and pose. In our model we use Keras which is a python library helps in deep learning model, tensor flow and OpenCV to detect emotions in face at different situations. We initially build a model and trained it such that it is capable of detecting and classifying facial emotion in real-time. The model gave good accuracy but takes much time to train the data. Here by, I conclude this paper by hoping that you got a fair knowledge, idea and understood the whole concept of designing the facial emotion recognition using CNN in real-time. You can make use of pre-trained models like VGG16, ResNet, MobileNet and EfficientNet for better performance of models.

## REFERENCES

[1] Mingli Song, Jiajun Bu, Chun Chen, and Nan Li. 2004.Audio-visual based emotion recognition-a new approach. In CVPR 2004, volume 2, pages II–II. IEEE

[2] S.Nithya Roopa' Research on Face Expression Recognition' International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8, Issue- 9S2, July 2019

[3] Christian Mancini ,Luca Falciati , Claudio Maioli and Giovanni Mirabella' "Threatening Facial Expressions Impact Goal-DirectedActions Only if Task-Relevant ,published october 2020

[4] Rohit pathar,Abhishek Adivarekar,Arti Mishra and Anusree Deshmukh "Human Emotion Recognition using Convolutional Neural Network in Real Time".

[5] A.D. Deshmukh, U.B. Shinde, D.L. Bhuyar, M.G. Nakrani, Facerecognition using Computer vision based on Internet of things for Intelligent door look system, 2019;

[6] Subramanian R.R., Seshadri K. (2019) Design and Evaluation of a Hybrid Hierarchical Feature Tree Based Authorship Inference Technique. In: Kolhe M., Trivedi M., Tiwari S., Singh V. (eds) Advances in Data and Information Sciences.Lecture Notes in Networks and Systems, vol 39. Springer, Singapore

[7] Joshva Devadas T., Raja Subramanian R. (2020) Paradigms for Intelligent IOT Architecture. In: Peng SL., Pal S., Huang L. (eds) Principles of Internet of Things (IoT) Ecosystem: Insight Paradigm. Intelligent Systems Reference Library, vol 174. Springer, Cham

[8] R. R. Subramanian, B. R. Babu, K. Mamta and K. Manogna, "Design and Evaluation of a Hybrid Feature Descriptor based Handwritten Character Inference Technique," 2019 IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS), Tamilnadu, India, 2019, pp. 1-5.

[9] R. R. Subramanian, R. Ramar, "Design of Offline and Online Writer Inference Technique", International Journal of Innovative Technology and Exploring Engineering, vol. 9, no. 2S2, Dec. 2019, ISSN: 2278-3075.

[10] Haoxiang, Wang. "Emotional Analysis of Bogus Statistics in Social Media." Journal of Ubiquitous Computing and Communication Technologies (UCCT) 2, no. 03 (2020): 178-186.

[11] Manoharan, Samuel. "Geospatial and Social Media Analytics for Emotion Analysis of Theme Park Visitors using Text Mining and GIS." Journal of Information Technology 2, no. 02 (2020): 100-107.

[12] Facial Expression Recognition usingConvolutional Neural Networks: State of the Art Christopher Pramerdorfer, Martin Kampel Computer Vision Lab, TU WienVienna, Austria.