Facial Expression Recognition Using Machine Learning

Manthri Bala Kiran
Data Science Master Program, Machine Learning
Högskolan Dalarna
Borlänge, Sweden
Email: h19baman@du.se

Abstract: The main aim of this project is to find the Facial Expression of humans using the Machine Learning Algorithms. Classifying an image based on its depiction can be complicated task for machines. It is easy for humans to look an image of bicycle and know it is a bicycle. To complete this task, I have used Convolution Neural Network which would be helping to classify the images and find emotions of the human face in the given images.

Keywords- Machine Learning; Deep Learning, Perdition; CNN; Hybrid-CNN; Image Classification; Data Mining;

I. INTRODUCTION

Facial Expression is the most common thing used by the humans to convey various types of feelings or meaning at various contexts. The range of the meaning is basic possibly innate of socio-emotional context such as “Surprise” or “Happy” and can be tend to be the most complex meaning such as “Carelessly”. Humans will be using Facial Expression at different context based on the environment they are been in [1].

The aim of this paper is to use different pixel image data and find the facial expression of humans. So, for this I would be using image classification technique with CNN to identify the different classes of the images (as different emotion of the images). Image classification is a process where computer will be using an image and identifying the class of the image or the probability of the image with respective to its class. A class is a label to classify the type of the image example: “Car”, “Animal”. Convolution Neural Networks are considered as the back bone for the image classification where it takes an image and assign a class and label to that image which makes that image unique [2] [3].

A Convolution Neural Network is a Deep Learning algorithm which takes the image as input, and it assigns an importance such as learnable weights and biases to the various aspects or the objects present in that image and which makes that image different from one to another [4].

II. LITERATURE REVIEW

Considering Facial Images based on the shape feature using the optimization algorithms. A classification based similarity finding is been proposed for the classification of the images based on the shape of the image like round or oval [5].

Image classification considered as the primary domain in the neural network which plays a major important role for medical image analysis. The image classification will be accepting the given input image and produces output classification for identifying whether the disease is present or not. CNN are being considered as the state-of-the-art methods for the image classification. They are been use to predict the development of the disease in the brain [6].

The method of combing Convolutional Neural Network (CNN) and eXtreme Gradient Boosting (XGBoost) will be giving a better performance. The two methods are considered as the best classifiers. CNN_XGBoost provides more precise output by integrating CNN as a trainable feature extractor to automatically obtain features from input and XGBoost as a recognizer in the top level of the network to produce results [7].

Image Classification is the core problems in computer vision field with a large variety of practical applications. Examples like: object recognition for robotic manipulation, pedestrian detention. There is a lot of attention associated with Machine Learning, specifically neural networks such as the Convolution Neural Network [8].

III. DATA SET

In this paper, I have taken the dataset of Facial Expression Recognition which has the pixel data of different Facial expression of the humans. I have taken this dataset from Kaggle [9]. The attributes of the dataset are listed as follows:

Emotion, Usage, Pixel are the attributes present in the dataset. We have also a test dataset where we have only the pixel data to set. The description of the training dataset attributes is:

- Emotion: This column has the emotion of the images in the data set. The emotions are being classified into seven categories they are- 0=Angry, 1=Disgust, 2=Fear, 3=Happy, 4=Sad, 5=Surprise, 6=Neutral.
- Usage: This column has three categories Training Set which consists of 28,709 examples, Public test which consists of 3,589 examples and Private test which consist of 3,589 examples.
- Pixels: This column has the pixel data of the image. Which is a series of numbers explaining the grayscale images of the faces. We can consider them as 48*48.

And the Test data file has only the 48*48 grayscale pixel data were we can use the dataset to predict the new set of images using our model.

IV. METHODOLOGY

a. Data Handling

Firstly, I have loaded the data into the data frame using Python. Then I have performed the data handling by taking two different steps. For the first step, I have taken the data frame divided the training frame into training, validation and test. And then performed the CNN model obtained the training accuracy and taken the test data set and predicted the emotion for the images. And also tried downloading an image from internet and checked the emotions for that image.

For the second time, divided my data into train, validation and test based on the usage column and then drop the usage column from the data sets. And later, tried to split my pixel's data to different columns and performed the CNN model and obtaining the training accuracy.

b. Exploratory Data Analysis

I tried to find the count of each emotion from the training data set. Also tried to have a look of the images present in the data set. And had a plot based on the training, validation and test set from the training data set.

c. Training and Testing data sets:

The data set has two files where one file has the emotions and another don't. The file with the emotions is been considered as the training file and it has three different categories. So I have divided them based on their categories which is good so we can see the performance of the model based on the test accuracy. The test data set is completely used for predicting the emotions based on our model.

d. Techniques Performed

I have used keras package where I have imported different layers, models, optimizers from it such as: Convolution2D, Activation, Dense, Sequential, SGD, Adam.

Using the above package, I have performed my Convolution neural network with varying my kernel sizes and epochs and tried to find my best test accuracy and training accuracy.

V. DATA EXPLORATION

Let's, first the count of emotions from our training dataset. This is the complete data set before we divide so we would be having the combinations of three categories from our usage column.

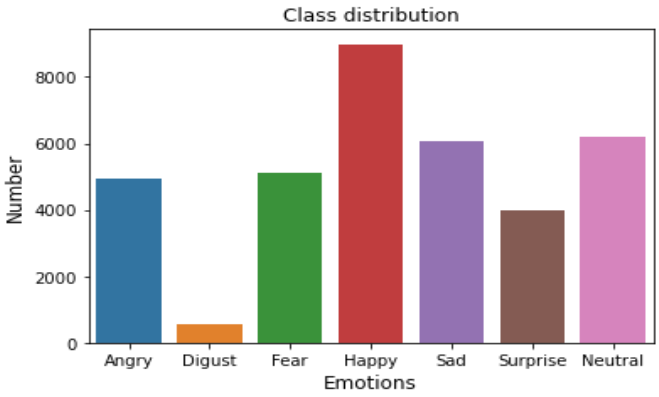

Fig 1: Emotions from the Training data set

Now let's try to a sample image from each category of the emotion from our training data set.

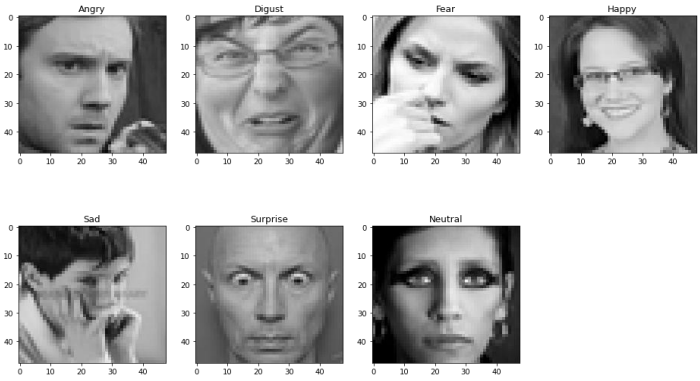

Fig 2: Sample image for each emotion

Now, let's see the emotion counts from the data set where the data set has been divided into training, validation and testing.

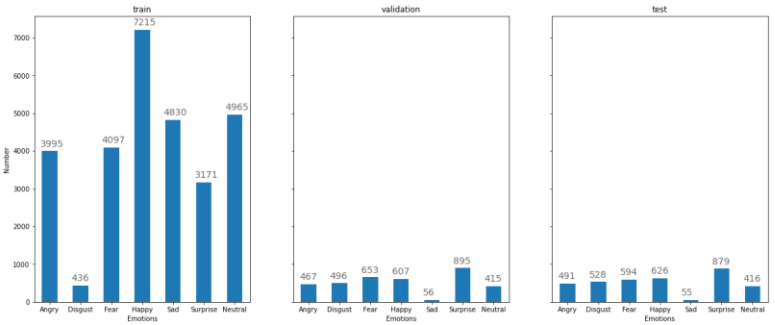

Fig 3: Emotion count of the divided training set

VI. RESULTS

Now, Let's train our model for finding the facial expression of the images from our data set. From the first time of the model prediction, I have got a training accuracy of 81% and the validation accuracy of 65%. I have used the same model for the test data from the training data set and got an accuracy of 67%.

Now, I have tried a different method for preparing my data for the model and got a training accuracy of 90% and validation accuracy of 89%. Now, I have implement the same model for the test set from the training data set and achieved an accuracy of 61%.

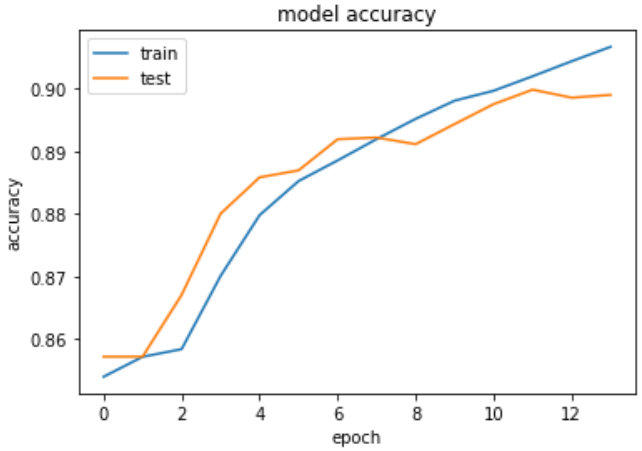

Fig 4: Training and Validation Accuracy

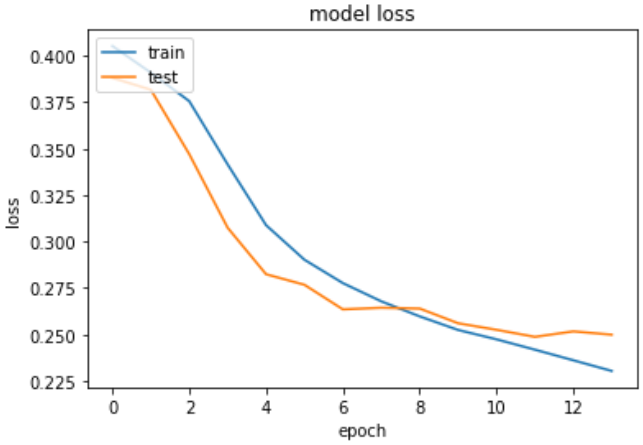

Fig 5: Training and Validation Loss

From the above figures 4 & 5 shows the training accuracy & loss, validation accuracy & loss of the training data set. The label test says the validation and label train says the training.

Using the above model which is been developed, I tried to predict the emotions from the test data set were the emotions of the images are not given. After predicting the prediction values of emotions for the test data set, I have combined the prediction values to the test data set. Now, let's see the emotion counts from the test data set.

	emotion	number
0	Angry	1009
1	Sad	1153
2	Digust	54
3	Surprise	762
4	Fear	853
5	Neutral	1462
6	Happy	1885

Fig 6: Emotion count from the Test Data Set.

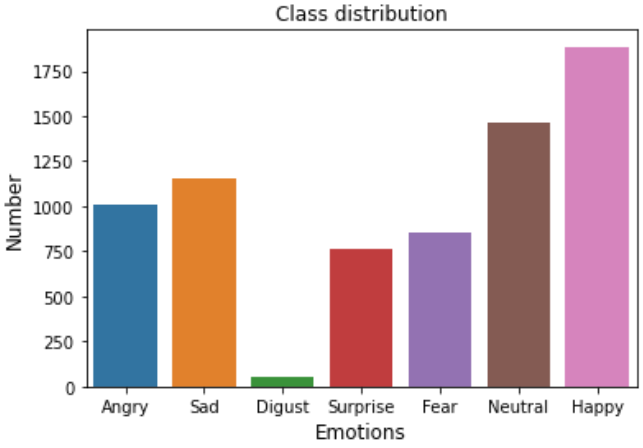

Fig 7: Plot of the emotions from the Test Data Set

Now, Let's see the sample of images from each emotion from the test data set which are been predicted from our model.

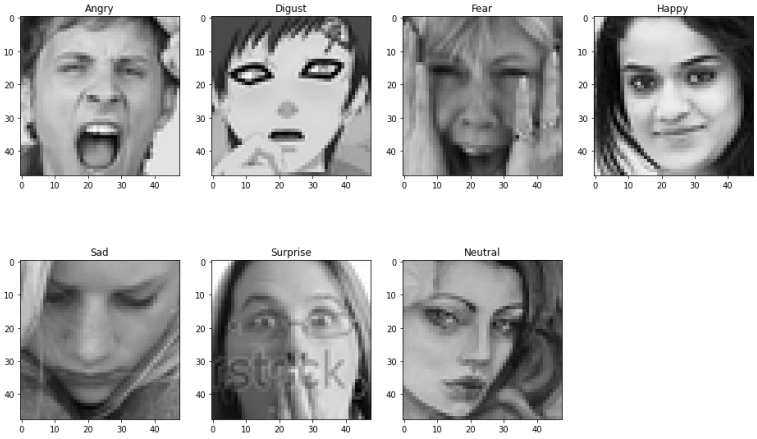

Fig 8: Images with emotions from Test Data Set

For a try, I have downloaded an image from the internet and tried predicting the emotion of that image using my model.

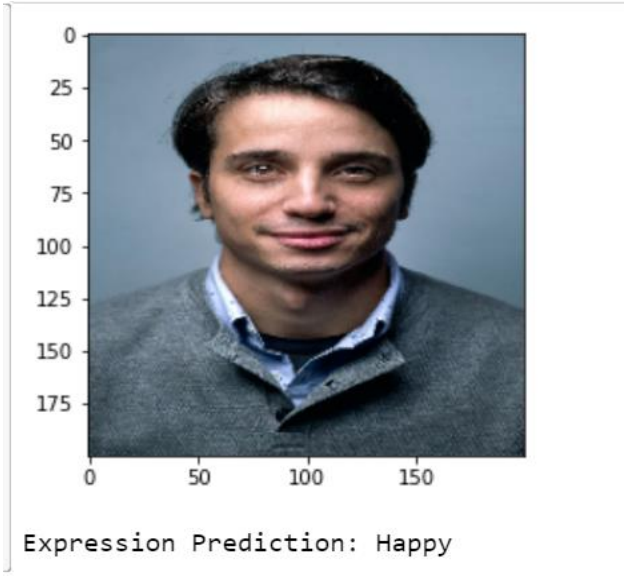

Fig 9: Predicting Emotion of the image from Internet

Now, I have divided my training, validation and testing based on the usage column and drop the usage column from my sets. And I have spited my pixel data into multiple columns and performed Hybrid-CNN and opted a training accuracy of 67% and validation accuracy of 66%.

VII. DISCUSSION

As humans, we are being trained to identify the classes of the images, objects at different stages of the life. In the same way, the machines can classify the images and their class with proper training. In the same way they can predict the emotions of the humans by images not as perfect as we humans can do but machines can predict the basic emotions of humans from the images.

VIII. LIMITATIONS

With respective to my project proposal and suggestions given from the project proposal, I have tried to perform RNN for the image classification but was not able to compare their performance for the image classification. And understood that the RNN performs better for time series data and sentence prediction and CNN works better for the image classification. But tried comparing the results with Hybrid CNN and CNN.

IX. REFERENCES

[1] F. Psychol, "Frontiers," [Online]. Available: https://www.frontiersin.org/articles/10.3389/fpsyg.2013.00115/full. [Accessed 01 11 2020].

[2] ThinkAutomation, "ELI5: what is image classification in deep learning?," ThinkAutomation, [Online]. Available: https://www.thinkautomation.com/eli5/eli5-what-is-image-classification-in-deep-learning/. [Accessed 01 11 2020].

[3] P. Vadapalli, "Using Convolutional Neural Network for Image Classification," upgrad, [Online]. Available: https://www.upgrad.com/blog/using-convolutional-neural-network-for-image-classification/#:~:text=Convolutional%20Neural%20Networks%20(CNNs)%20are,part%20of%20machine%20learning%20experiments.. [Accessed 01 11 2020].

[4] S. Saha, "A Comprehensive Guide to Convolutional Neural Networks — the ELI5 way," towards data science, [Online]. Available: https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53. [Accessed 01 11 2020].

[5] Dr.S.Vijayarani1, "An Efficient Algorithm for Facial Image Classification," semantic scholar, [Online]. Available: https://pdfs.semanticscholar.org/be8f/5dc361aa247703b073a8951d80be8b571afa.pdf. [Accessed 01 11 2020].

[6] K.Balaji, "Image Classification," Science direct, [Online]. Available: https://www.sciencedirect.com/topics/computer-science/image-classification. [Accessed 01 11 2020].

[7] X. Ren, "A Novel Image Classification Method with CNN-XGBoost Model," link.springer, [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-319-64185-0_28. [Accessed 01 11 2020].

[8] M. Hussain, "A Study on CNN Transfer Learning for Image Classification," link.springer, [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-319-97982-3_16. [Accessed 01 11 2020].

[9] R. P. Competition, "Kaggle," [Online]. Available: https://www.kaggle.com/c/challenges-in-representation-learning-facial-expression-recognition-challenge/data. [Accessed 01 11 2020].