International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 04 | Apr 2021 www.irjet.net p-ISSN: 2395-0072

FACIAL EXPRESSION RECOGNITION USING CNN WITH DATA AUGMENTATION

Swapnil Narendra Ingale[1], Ashish Anant Kadam[2], Manoj Kashinath Kadam[3], Prof. Vidya Bharde[4]

[1,2,3]*Student, Dept. of Computer Engineering, M.G.M College of Engineering and Technology, Kamothe, Maharashtra, India*
[4]*HOD. Dept. of Computer Engineering, M.G.M College of Engineering and Technology, Kamothe, Maharashtra, India*

Abstract – *In real life, people express their emotion on their face to show their psychological activities and attitudes in the interaction with other people. Humans possess natural ability in understanding facial expressions. But today is the time of intelligent machines and new discoveries in the field of human-computer interactions. Detecting emotion from facial expression has become an urgent need because of its immense applications in artificial intelligence such as human-computer collaboration, data-driven animation, human-robot communication etc. Since it is a demanding and interesting problem in computer vision, several works had been conducted regarding this topic. The objective of this research is to develop a facial expression recognition system based on convolutional neural network (CNN) with data augmentation. This approach enables to classify seven basic emotions consist of angry, disgust, fear, happy, neutral, sad and surprise from image data from FER 2013 database. The facial expression recognition is realized through building and training CNN network by keras.*

Key Words: CNN, Data augmentation, expression recognition, keras.

1. INTRODUCTION

According to reports from various web sources, the most widely used mode of communication used by humans is Facial expression. Facial expression recognition is an evolving technology in the field of human-computer interaction. Facial expression recognition has its branches spread across various applications such as virtual reality, webinar technologies, online surveys and many other fields. Even though high advancements have been witnessed in this field, there are several diplomacies that exist. The traditional feature extraction methods have slower response and lack in performance. The traditional methods have high latency or delay in their response. Through these traditional methods, it is extremely difficult to extract the required features effectively and hence, is too hazardous to utilize for real time applications. [1]

2. PROBLEM STATEMENT

In real life, people express their emotion on their face to show their psychological activities and attitudes in the interaction with other people. The primary focus of this project is to determine which emotion an input image that contains one facial emotion belongs to. Because human face is complex to interpret, emotion recognition can be specifically divided into classification of basic emotion and classification of compound emotion.

3. EXISTING SYSTEM

Like every other classification problems, the emotion recognition problem requires an algorithm to complete feature extraction and categorical classification. In order to classify an emotion, we need to extract certain feature from data and build an model that can classify the input based on the feature. The existing system consists of steps like data pre-processing, feature extraction, model construction and result generation.

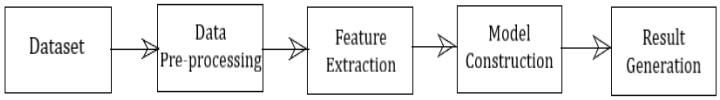

Fig-1: BLOCK DIAGRAM

4. PROPOSED SYSTEM

Convolutional Neural Network is considered as the methodology that is used with data augmentation in this research. Dataset that is used in this research has variation as data was collected from different datasets. As a result, the proposed model is not biased to any particular dataset. The event flow chart of this system is illustrated in fig. 4. At first, the model takes an image from the dataset and detects face from the image by Cascade Classifier. If face is found, then it is sent for preprocessing. Data have been augmented by ImageDataGenerator() function offered by the Keras API. At last, the augmented dataset is fed into CNN in order to predict the class.

4.1 EXPLORING DATASET

In this project, the dataset used to train the models is FER-2013. The FER-2013 dataset consists of 35887 images, of which 28709 labelled images belong to the training set and the remaining 7178 images belong to the test set. The images in FER-2013 dataset are labeled as one of the seven universal emotions: Happy, Sad, Angry, Fear, Surprise, Disgust, and Neutral. The images in FER-2013 dataset are in grayscale and is of dimensions 48x48 pixels.

4.2 TRAINING & VALIDATION

During training, to minimize the losses of Neural Network, an algorithm called Mini-Batch Gradient Descent has been used. Mini-Batch Descent is a type of Gradient Descent algorithm used for finding the weights or co-efficient of artificial neural networks by splitting the training dataset into small batches i.e., a training batch & a validation batch using Data Augmentation techniques. The ImageDataGenerator() class of Keras library in python is used to accept values or images to train.

4.3 CNN MODEL

The CNN designed is based on sequential model and is designed to have six activation layers, of which 4 are convolutional layers and the remaining 2 are fully controlled layers. The 4 convolutional layers consists of similar training techniques such as Batch Normalization, ReLu Activation function, Maxpooling and Dropout.

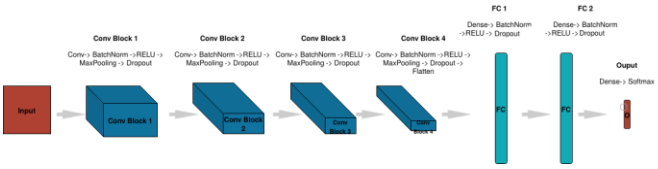

Fig-2: CNN Model

4.4 REPRESENT MODEL AS JSON STRING

The output of the CNN model is made to store as JSON string. A JSON or JavaScript Object Notation is used as it stores the data and allows faster exchange of data. The model.to_json() function of python is used to write the output of the trained model into JSON.

5. IMPLEMENTATION AND TESTING

5.1 FLASK APP & HTML TEMPLATE

An HTML file is created and designed, such that the body is coded to return the operations of the flask app. Therefore, an HTML template is created for the Flask app with a window of defined height and width that reads and runs the actions intended by the flask app.

5.2 REAL-TIME CLASSIFICATION

In this project, OpenCV's Haar cascade is used to for the detection and extraction of the region containing the face from the video feed of webcam through the flask app. The video is converted to grayscale and the detected face is contoured or enclosed within a region defined to surround the face.

6. RESULT

The CNN model designed is set to undergo 15 epochs. When trained gives an accuracy of 63% after the 15th epoch and the maximum efficiency achieved is also 64.6%. The plot of model accuracy over each epoch (from epoch_1 to epoch_15) is shown in the below

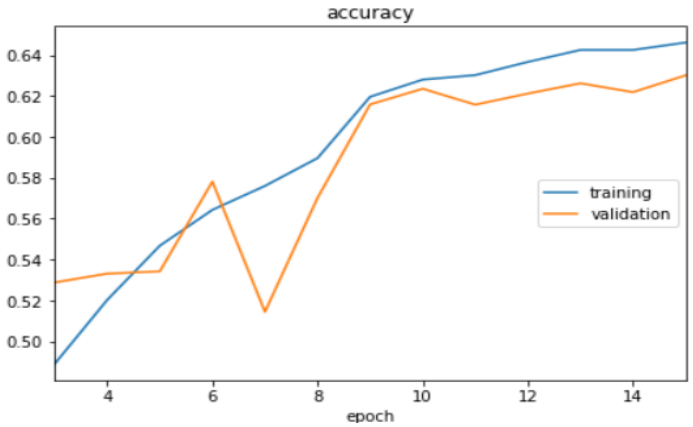

Fig-3: Accuracy vs Epoch

```
accuracy:
training   (min:    0.313, max:    0.646, cur:
validation (min:    0.346, max:    0.630, cur:
```

Fig-4: Accuracy

It has been observed that the CNN model developed can easily or flawlessly detect the following emotion types: Happy, Neutral, Sad, Surprise, Angry, Disgust and Fear.

7. CONCLUSION

Using the FER-2013 dataset, a test accuracy of 64.6% is attained with this designed CNN model. The achieved results are satisfactory as the average accuracies on the FER-2013 dataset is 65% +/- 5% and therefore, this CNN model is nearly accurate.

REFERENCES

[1] Facial Expression Recognition with Faster R-CNN (Jiaxing Lia, Dexiang Zhanga, Jingjing Zhanga, Jun Zhanga, Teng Lia, Yi Xiaa, Qing Yana, and Lina Xuna a. The School of Electrical Engineering and Automation Anhui University, Hefei 230601, China).Dr. Muhammad

Omer, Dr. Mamoona Asghar, "NUMBER PLATE RECOGNITION SYSTEM", Project Thesis.

[2] Real Time Facial Expression Recognition in Video using Support Vector Machines (Philipp Michel Computer Laboratory University of Cambridge CB3 0FD, United Kingdom), (Rana El Kaliouby Computer Laboratory University of Cambridge, Cambridge CB3 0FD, United Kingdom).

[3] Real-time Emotion Recognition From Facial Expressions (Minh-An Quinn, Grant Sivesind, Guilherme Reis CS 229 Stanford University)

[4] Baseline CNN structure analysis for facial expression recognition (Minchul Shin, Munsang Kim and Dong-Soo Kwon)