International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-8 Issue-4, April 2019

Emotion Recognition System based on Facial Expressions

Abhilasha Dey, T.Sujithra, Barsha Anand

***Abstract**: Emotion recognition has been a challenging thing in the industry for many years. Leveraging our knowledge skills and science on the data acquired and extracting useful information is not an easy task. However, since the birth of deep learning in computer vision, emotion recognition has turned into a wildly-tackled research problem. In this paper we are implementing a convolution neural network (CNN) building framework to design real time CNNs. We are making a real-time emotion recognition system which accomplishes the task of face detection and emotion classification simultaneously in one step using our proposed CNN framework.*

***Keywords:** convolution neural network (CNN), Emotion Recognition*

I. INTRODUCTION

Emotions play a major role in human life. Our emotions reflect our moods and feelings at various kinds of situations or time. It reflects how we feel. Emotion displayed on our faces is the way we communicate without expressing ourselves using words. As humans we can produce hundreds of facial expressions, which vary in terms of relations, hidden agenda, and meaning.

Figure 1.

The emotions we express are often translated into a communication medium by the minute changed in our numerous discreet features of the face. Each facial feature has its own role, at times addition or subtraction of it, changes the entire meaning of the emotion. To extract the correct emotion from the subject we first need to carefully extract the facial features which help us identify what the subject is trying to express. The main contenders for this are- eye size (enlarge, normal or relaxed), turning of the lips (upwards, downwards or relaxed), temple of the head, the space between the brows and the arching of the eyebrows (raised, relaxed or pulled together, displaying tension or sorrow) Each of the above feature, considered singly or in combination with others, truly conveys the underlying emotion. The CNN model is first trained with hundreds of images before deploying a real time detection system.

Figure 2.

II. PROBLEM DEFINITION

Throughout the literature survey our primary goal was to find areas that would concern the fleet manager the most while managing and operating his EVs. Our research has led us to the following conclusions. First, the fleet manager needs to ensure that the vehicles in his fleet are available for deployment all the time, that is, he doesn't have cancel a trip because no vehicles are charged enough to make the trip possible. Second, all the record of the trips made, his clientele and his drivers needs to be stored in one single centralized system, that would make his work simple and efficient. And third, he needs a charging infrastructure management system that would enable him to keep a track of his chargers and allow him to manage energy efficiently.

III. METHODS

3.1 Convolutional Neural Networks

CNN stands for convolutional neural networks. Neural networks are a machine learning technique designed to model to replicate the structure of the human brain. CNNs have wide applications in image and video recognition, recommender systems and natural language processing.

Manuscript published on 30 April 2019.
***** Correspondence Author (s)
Abhilasha Dey, Dept of Computer Science and Engineering, SRM Institute of Science and Technology, Chennai, Tamil Nadu, India.
Dr. T Sujithra Dey, Assistant Professor, Dept of Computer Science and Engineering, SRM Institute of Science and Technology, Chennai, Tamil Nadu, India..
Barsha Anand, Dept of Computer Science and Engineering, SRM Institute of Science and Technology, Chennai, Tamil Nadu, India.

These neurons learn how to convert a picture of dog (input signals) into corresponding the label “dog” (output signals) , which forms the basis of automated recognition.

Step 1: Convolution

The first few layers that receive an input signal are called as convolutional filters. Here the network tries to label the input signal based on its past learning. If the input signal looks like a “dog” then the previous “dog” reference signal will be mixed or convoluted with this input signal. The resulting output is then passed on to the next layer. Signals to be propagated proficiently for better accurate identification and end results.

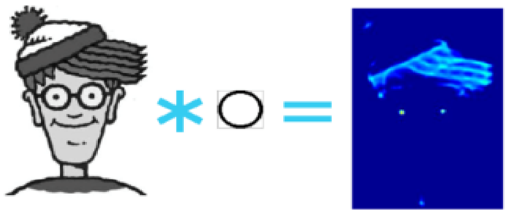

Figure 3. Convolving Bob with a circle filter – eyes.[3]

Convolution is translationally invariant. This means that each convolutional layer has a specific feature associated with it and represents it (e.g. Mouth, ear, nose) and the CNN learns which features belong to the resulting reference (i.e. dog). The output is dependent of the presence of the feature, whether is present or not, irrespective of the location of the feature. Hence the dog could be sitting, sleeping or standing, and the CNN will correctly interpret it.

Step 2: Sub-sampling

To reduce the sensitivity of the filters to noise and variations, the convolutional layers are “smoothened out” which is also known as subsampling. This is done by taking the averages or the maximum over a sample of the signal. Types of sampling for images include – reducing the size, color contrast across RGB channels.

Figure 4. Sub sampling Bob by 10 times, thereby creating a lower resolution image.

Step 3: Activation

Activation layer like the human brain which controls the transmission of neurons, controls the flow of signals from one layer to another. Neurons with past reference associations, will activate and set of more neurons, aiding CNN has many self-embedded and compatible activation signals which facilitate the signal propagation process, the most common of all is ReLU – Rectified Linear Unit, known for its fast training speed.

Step 4: Fully Connected

The completion of steps 1 to 3, results in the last layer being entirely connected, essentially meaning that all the neurons of the previous layers are linked to each neuron in the following layers. Hence it imitates logical reasoning at the complex level, wherein all probable paths from the I/P to the O/P are considered.

Step 5: Loss (Training phase)

Neural network, during the training phase has an additional layer - the loss layer. This serves as a feedback mechanism, telling the neural network whether it branded the inputs suitably or not, and how much guess work was playing a role, thereby helping the neural network to strengthen the right concepts and realign its understanding as it trains. Loss layer is the last layer of this phase.

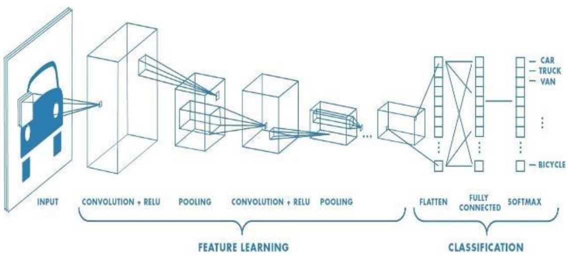

Figure 5. Flow diagram for convolutional neural network

IV. SYSTEM REQUIREMENTS

H/W Requirements

Input device: Keyboard and Mouse.
Processor: Any Processor above 500 MHz.
Hard Disk : 4 GB
Ram : 4 GB
Output device : VGA and High-Resolution Monitor.

S/W Requirements

OS : Windows 8 or higher
Programming : Python 3.6

V. ALGORITHM DESIGN

Image Data Input Parameters: number of images, image’s height, image’s width, number of channels, number of levels per pixel.
Image Preprocessing: Aspect Ratio: Square. Cropping, giving importance to the middle of the image.
Image Scaling: Use 48x48 to use an image size
Dimensionality reduction: Make a RGB into a Gray image (reduces 3 channels to 1)

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

Architecture Inputs:

- Number of Layers
- Number of Neurons per Layer
- Regularization Parameters
- Learning Rate
- Dropout Rate

- Activation Function (linear, sigmoid, tanh, ReLu)
- Algorithm for Weight updates (SGD, Adam, RMSProp…)

VI. IMPLEMENTATION

For the implantation of the idea and proof of concept, a shallow CNN is created. The network has two convolutional layers and one fully connected layer. The first convolutional layer has 32X32 filters, stride size = 1, batch normalization and dropout rate, with the absence of maxpooling. The convolutional layer has 64 3X3 filters, with stride size = 1, batch normalization, dropout rate and maxpooling with filter size 2X2. The fully connected layer has 512 hidden layers and the loss function – Softmax. The CNN essentially has Rectified Linear Unit (ReLU) as the activation layer, present in all layers.

The following briefly explains the flow of the program, one function after the other, highlighting the main functions that are essential in fulfilling the program's objective.

- Imagecreate
- Pre-process
- Train
- Split-data
- Test

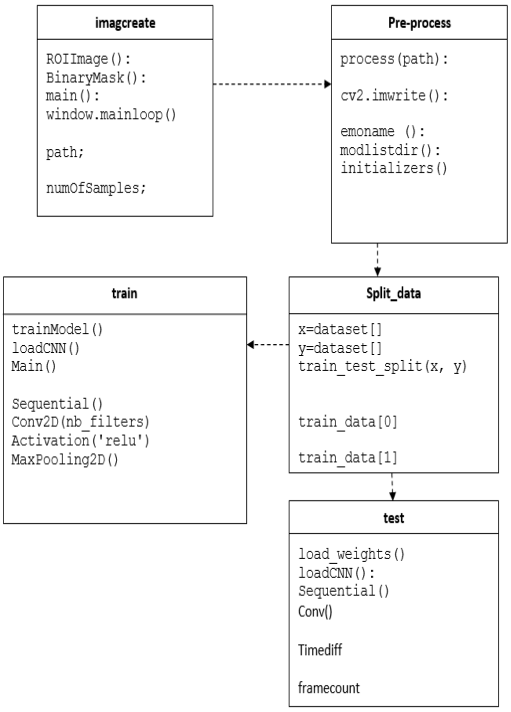

Figure 6 Class diagram expressing the properties and functions of each class

1) Face Capture

The python program face_capture.py is executed at first. Here the generate_faces() function captures the video of the face and sorts into smaller frames and takes clipping of those frames and stores them as an image. The images taken are converted into greyscale. The pool of pictures captured are essentially the "Training Set" that will be used further for the training of the CNN model. While running the program the input parameters fed are the emotion type and the number of images to be clicked.

python face_capture.py -e smile -n 100

The above command will capture 100 images while the subject poses with a smiling face. If needed the program can be interrupted in the middle and its execution can be stopped by hitting the letter 'q'.

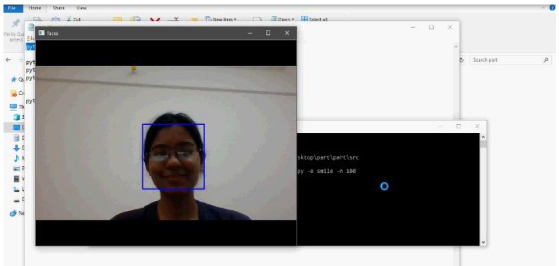

Figure 7: Running the face_capture.py function

2) Dataset Creation

After running the Face_capture program, there are images that have been clicked and needed to be stored into named folders categorized based on emotions, to be used later for emotion-wise training of the CNN.Here the hundred pictures clicked of the "smile" emotion are stored in a folder named – smile, which is created by the dataset_creator.py python program.

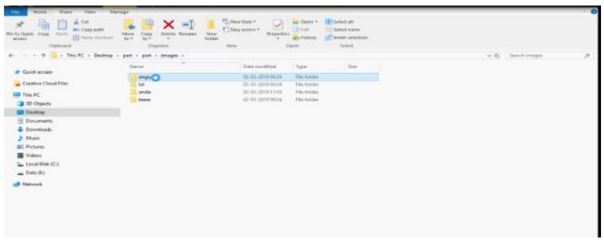

Figure 8 Running the dataset_creator.py function

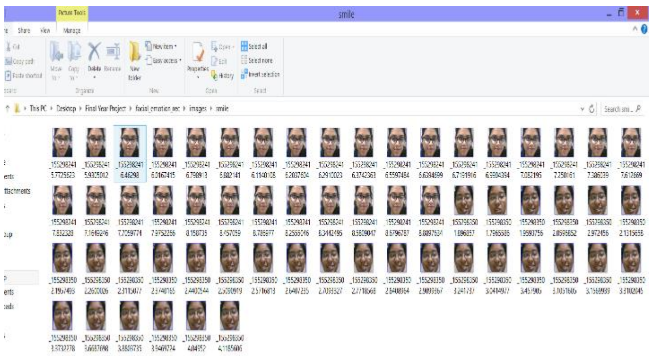

Figure 9 Pool of pictures stored in each folder after executing dataset_creator.py function

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

3) Data Manager

During the deployment of the real-time CNN system for detection of emotion, there will be an emoji or emoticon that will pop up on the screen depicting the emotion the subject's face is displaying. In the program data_manager.py the images clicked are mapped to the corresponding emoji.
Another important and crucial task carried out in this step is the image preprocessing. Here the captured imaged are

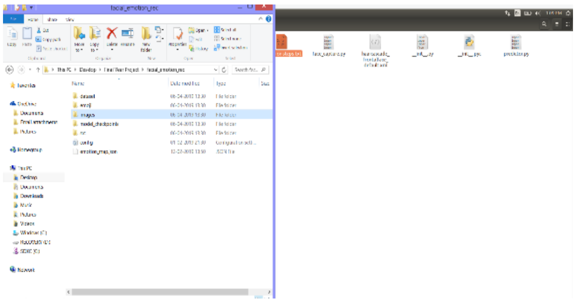

Figure 10 Running the data_manager.py function

4) Training the data

Here is where the real task begins. Up till now all the images have been preprocessed and stored in separate folders named according to the emotion – smile, sad and angry. Now the python program traioner.py opens each folder one by one and trains the CNN model with the facial features extracted.

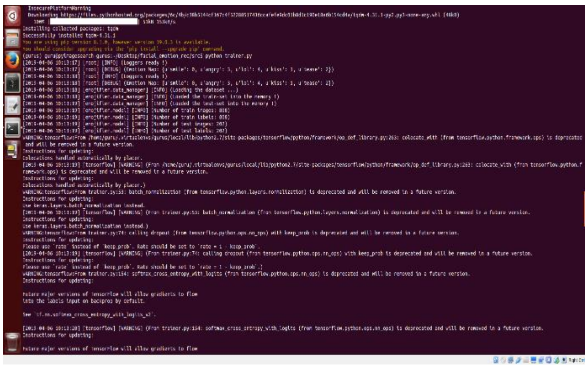

Figure 11 (a) Running the trainer.py function

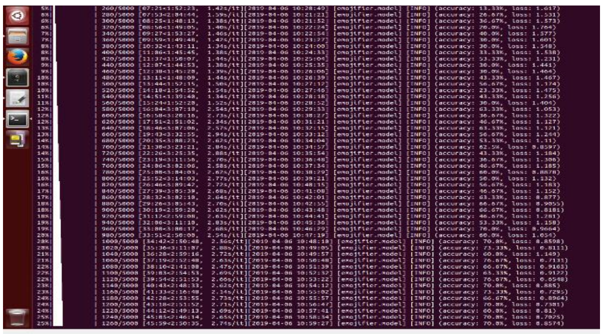

Figure 11(b) Training of the CNN model

cropped to remove all the background noise and have all the images that will be used for training to have only the face of the subject. The face spanning from the upper-limit of the face to the chin and from one ear end to the other is cropped

5) Real time emotion prediction

In the real-time emotion detection system, the face_capture.py is called once again and the subject's face is captured as a video, from which on a frame-to-frame basis pictures of the face is captured. The predictor.py program detects the facial features of the real-time face and calculates an estimated guess by matching the characteristics of the facial features extracted to the ones already stored as a result from the trainer.py.

VII. RESULTS

In general, Neural Network consists of different hidden layer. In most of Conv2D will have two hidden layers with 16 or 32 neurons and more, Hidden layers are multiplied with different random weight of image pixel data which is between 0 to 1. But in Conv2D Neural Network was design with same two hidden layers and each hidden layers consists of large set of neurons i.e. we used 512 neurons are taken and this are multiplied with random weights. In the Convolutional Layer, 5x5 filters are applied (extracting 5x5-pixel sub-regions), with ReLU activation function. Here Table 1 displays some labeling of characters.

Layer Type	Layer operation	No of neurons	Feature map size	Window size	Total parameters
C1	Conv2D	512	168 X 168	5X5	5000

Table 1: Sample data representation of labeling.

The five layers of CNN along with Adam Optimization, forms the testing phase of the neural network. There are two layers for convolution, one layer for max pooling or sub-sampling and one Flatten layer.

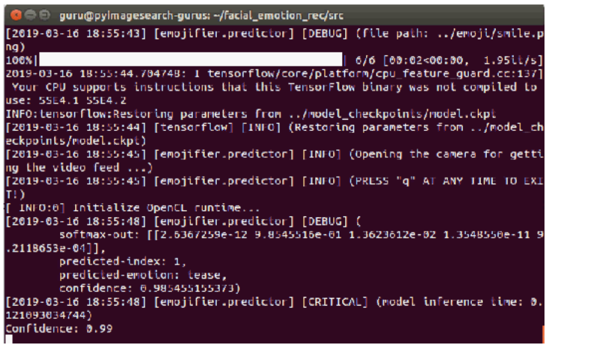

Figure 12(a) Result- displaying the predicted index, emotion label and level of confidence

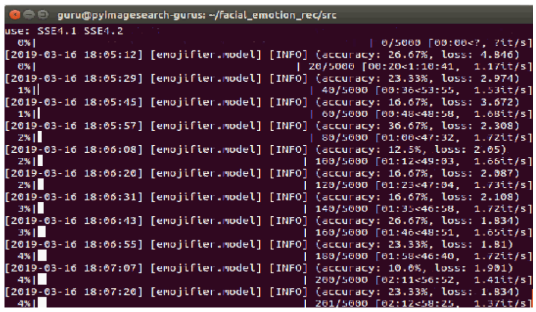

Figure 12(b) Result- displaying the accuracy% and loss%

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

VIII. CONCLUSION

The facial recognition was performed in a two-step process-human face detection, followed by facial expression extraction and recognition. The entire project was more focused and aimed towards the latter step. All the extraction features were carefully chosen and finalized to be used in the emotion characterization of the subject. A real-time data base was created based on the subject of choice and the CNN model was trained to work independently. Post the completion of the data set training, real time images of subjects are captured, and detection and recognition are performed, with the calculation of the confidence percentage and displaying of the result on screen.

REFERENCES

1. Raghav Puri, Mohit Tiwari, Shivendra Goel, Nitish Pathak, "Emotion Detection using Image Processing in Python" Conference: international Conference on Big Data, Computer Science and Information Technology (ICBDCSIT), At Chandigarh, India, January 2011
2. W.N. Widanagamaachchi & A.T. Dharmaratne, "Emotion Recognition with Image Processing and Neural Networks", University of Colombo School of Computing, Sri Lanka. October 2009.
3. http://www.ais.uni-bonn.de/deep_learning/images/ Convolutional_NN.jpg
4. Algobeans,(2016, January) [Online] Available : https://algobeans.com/2016/01/26/introduction-to-convolutional-neural-network/
5. Paweł Tarnowski, Marcin Kołodziej, Andrzej Majkowski, Remigiusz J. "Emotion recognition using facial expressions", Rak Warsaw University of Technology, Warsaw, Poland. International Conference on Computational Science, ICCS 2017, 12-14 June 2017, Zurich, Switzerland.
6. Mehmet S. Unluturk, Kaya Oguz, Coskun Atay, "Emotion Recognition Using Neural Networks", Department of Software Engineering, Izmir University of Economics, TURKEY, 10th WSEAS International Conference on NEURAL NETWORKS. January 2009
7. Kavita Kushwah, PGOI, INDORE, PROF. Vikrant Sharma, (CSE)PGOI, INDORE, Upendra Singh, "Neural Network Method Through Facial Expression Recognition", International conference of Electronics, Communication and Aerospace Technology (ICECA), 20-22 April 2017, Conference Location: Coimbatore, India, Published.
8. Harsh Pokharna, (2016, July) [Online] Available : https://medium.com/technologymadeeasy/the-best-explanation-of-convolutional-neural-networks-on-the-internet-fbb8b1ad5df8
9. Mohammed Saaidia, Dept. Génie-Electrique ,Université M.C.M Souk-Ahras, Souk-Ahras, Algeria Narima Zermi and Messaoud Ramdani, "Facial Expression Recognition Using Neural Network Trained with Zernike, Moments", Dept. Electronique, Université B.M Annaba, Annaba, Algeria.2014 4th International Conference on Artificial Intelligence with Applications in Engineering and Technology. Date of Conference: 3-5 Dec. 2014, Date Added to IEEE Xplore: 10 December 2015Conference Location: Kota Kinabalu, Malaysia, Published.
10. Ujjwal Karn (2016, August) [Online] Available: https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/
11. Adit Deshpande (2016 July) [Online] Available:
12. https://adeshpande3.github.io/A-Beginner%27s-Guide-To-Understanding-Convolutional-Neural-Networks/
13. Vibhor Nigam (2018, Aug) [Online] Available: https://towardsdatascience.com/understanding-neural-networks-from-neuron-to-rnn-cnn-and-deep-learning-cd88e90e0a90
14. Paul van Gent (2016, August) [Online] Available: http://www.paulvangent.com/2016/08/05/emotion-recognition-using-facial-landmarks/

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)