Facial Expression Based Emotion Recognition

G. Appa Rao[1], N. Viswanadha Reddy[1], K. Sujatha [2]

[1]Assistant Professor, CSE, DIET-Anakapalle, Visakhapatnam.
{apparaog@diet.edu.in, viswanadhareddyn@diet.edu.in}
[2]Professor, CSE, DIET-Anakapalle, Visakhapatnam.
{sujathak@diet.edu.in}

Abstract. The concept of recognition of emotional expressions is mainly giving a system the ability to know the actual state of emotion of a person by using facial expressions. Understanding the human facial expressions and the study of expressions has many aspects. Machine learning systems can be trained to recognize emotional expressions from im-ages of human faces, with a high degree of accuracy in many cases. However, implementation can be a complex and difficult task; various algorithms are used in each phase of the process. The computers when used for detecting emo-tions can be used in wide range of areas such as identifying criminals and also can be used by many organizations to identify the expression of the customer regarding the company's product.

Keywords: Human Facial Expression, Machine Learning, Emotion Detection

1 Introduction

Facial Expression Recognition is an Image Classification problem located within the wider field of Computer Vision. Image Classification problems are ones in which images must be algorithmically assigned a label from a discrete set of categories. In FER systems specifically, the images are of human faces and the categories are a set of emotions. Ma-chine learning approaches to FER all require a set of training image examples, each labeled with a single emotion cate-gory.
A standard set of seven emotion classifications are often used are: Anger, Disgust, Fear, Happiness, Sadness, Surprise, and Neutral.

Classifying an image based on its depiction can be a complicated task for machines. It is straightforward for humans to look at an image of a bicycle and know that it is a bicycle, or to look at a person's face and know that they are smiling and happy.

Two different approaches are used for facial expression recognition, both of which include two different methodologies, exist [6]. Dividing the face into separate

action units or keeping it as a whole for further processing appears to be the first and the primary distinction between the main approaches. In both of these approaches, two different methodologies, namely the 'Geometric based' and the 'Appearance-based' parameterizations, can be used.Making use of the whole frontal face image and processing it in order to end up with the classifications of 6 universal facial expression prototypes: disgust, fear, joy, surprise, sad ness and anger; outlines the first approach. Here, it is assumed that each of the above-mentioned emotions have characteristic expressions on face and that's why recognition of them is necessary and sufficient. Instead of using the face images as a whole, dividing them into some sub-sections for further processing forms up the main idea of the second approach for facial expression analysis. As expression is more related with subtle changes of some discrete features such as eyes, eyebrows and lip corners; these fine-grained changes are used for analyzing automated recognition.There are two main methods that are used in both of the above explained approaches.

Geometric Based Parameterization is an old way which consists of tracking and processing the motions of some spots on image sequences, firstly presented by Suwa et al to recognize facial expressions[7]. Cohn and Kanade later on tried geometrical modeling and tracking of facial features by claiming that each AU is presented with a specific set of facial muscles [8]. The disadvantages of this method are the contours of these features and components have to be adjusted manually in this frame, the problems of robustness and difficulties come out in cases of pose and illumination changes while the tracking is applied on images, as actions & expressions tend to change both in morphological and in dynami-cal senses, it becomes hard to estimate general parameters for movement and displacement. Therefore, ending up with robust decisions for facial actions under these varying conditions becomes to be difficult. Rather than tracking spatial points and using positioning and movement parameters that vary within time, color (pixel) information of related re-gions of face are processed in Appearance Based Parameterizations; in order to obtain the parameters that are going to form the feature vectors.

Different features such as Gabor, Haar wavelet coefficients, together with feature extraction and selection methods such as PCA, LDA, and Adaboost are used within this framework. For classification problem, al-gorithms like Machine learning, Neural Network, Support Vector Machine, Deep learning, Naive Bayes are used. Raghuvanshi A. et al have built a Facial expression recognition system upon recent research to classify images of hu-man faces into discrete emotion categories using convolutional neural networks [9]. Alizadeh, Shima, and Azar Fazel have developed Facial Expression Recognition system using Convolutional Neural Networks based on Torch model [10].

2 FER System

2.1 System Architecture:

The facial expression recognition system is implemented using convolutional neural network. The block diagram of the system is shown in following figures:

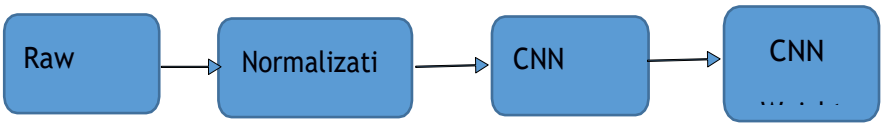


Figure 1(a): Training Phase

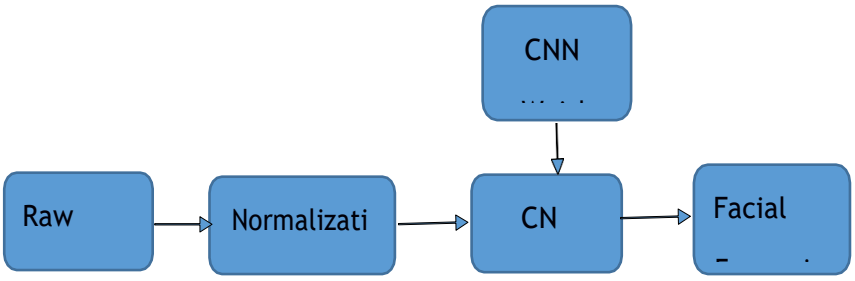


Figure 1 (b): Testing Phase

During training, the system received a training data comprising grayscale images of faces with their respective expres-sion label and learns a set of weights for the network. The training step took as input an image with a face.

Thereafter, an intensity normalization is aplied to the image. The normalized images are used to train the Convolu-tional Network. To ensure that the training performance is not affected by the order of presentation of the examples, validation dataset is used to choose the final best set of weights out of a set of trainings performed with samples pre-sented in different orders. The output of the training step is a set of weights that achieve the best result with the train-ing data. During test, the system received a grayscale image of a face from test dataset, and output the predicted ex-pression by using the final network weights learned during training.

Its output is a single number that represents one of the seven basic expressions.

2.2 Working Of CNN:

How Does A Computer Read an Image?

Consider this image of the New York skyline, upon first glance you will see a lot of buildings and colors. So how does the computer process this image? The image is broken down into 3 color-channels which is Red, Green and Blue. Each of these color channels are mapped to the image's pixel.Then the computer recognizes the value associated with each pixel and determine the size of the image.

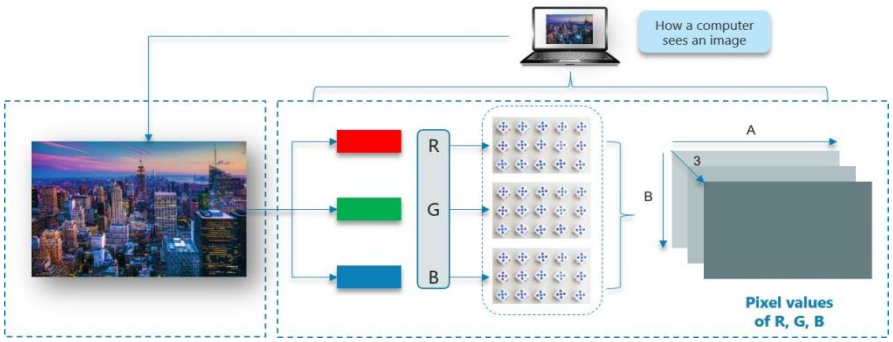
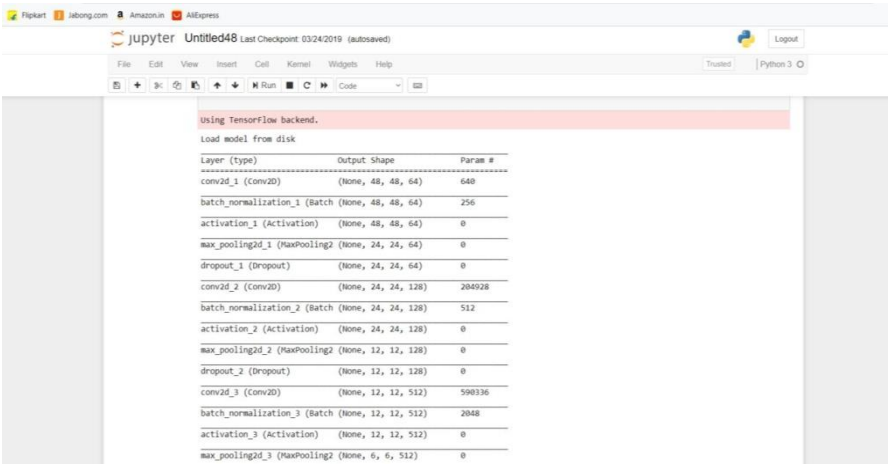


Figure.3: Reading RGB image.

3 Results and Analysis

Here, we have taken a color image of smiling woman. We have converted this RGB image to a grayscale image and then converted it pixels into an csv file format and given it as an input for our program.



CNN architecture for facial expression recognition as mentioned above was implemented in Python. Along with Python programming language and some of the packages as sqlite, Numpy, tensor flow, pillow, pip, matplotlib, keras, pandas etc, libraries were used. Training image batch size was taken as 30, while filter map is of size 20x5x5 for both convolution layer. Validation set was used to validate the training process. In last batch of every epoch in validation cost, validation error, training cost, training error are calculated. Input parameters for training are image set and corresponding output labels. The training process updated the weights of feature maps and hidden layers based on hyper-parameters such as learning rate, momentum, regularization and decay.

4 Conclusion and Future Enhancements

Convolutional Neural Networks is a popular deep learning technique for current visual recognition tasks. Like all deep learning techniques, Convolutional Neural Networks are very dependent on the size and quality of the training data. Given a well-prepared dataset, Convolutional Neural Networks are capable of surpassing humans at visual recognition tasks. However, they are still not robust to visual artifacts such as glare and noise, which humans are able to cope. The theory of Convolutional Neural Networks is still being developed and researchers are working to endow it with properties such as active attention and online memory, allowing Convolutional Neural Networks to evaluate new items that are vastly different from what they were trained on. This better emulates the mammalian

visual system, thus moving towards a smarter artificial visual recognition system. In this project, a LeNet architecture based four-layer convolution neural network is implemented to classify human facial expressions i.e. happy, sad, surprise, fear, anger, disgust, and neutral. The system has been evaluated using Accuracy, Precision, Recall and F1-score. The classifier achieved accuracy of 65.4%. In the future work, the model can be extended by allowing new pre-trained models such as AlexNet [11] and VGGNet [12], to get the more classifier accuracy.

References

[1] Shan, C., Gong, S., & McOwan, P. W. (2005, September). Robust facial expression recognition using local binary patterns. In Image Processing, 2005. ICIP 2005. IEEE International Conference on (Vol. 2, pp. II-370). IEEE.

[2] Chibelushi, C. C., & Bourel, F. (2003). Facial expression recognition: A brief tutorial overview. CVonline: On-Line Compendium of Computer Vision, 9.

[3] Convolutional Neural Networks (LeNet) – Deep Learning 0.1 documentation". Deep Learning 0.1. LISA Lab. Retrieved 31 August 2013.

[4] Matusugu, Masakazu; Katsuhiko Mori; Yusuke Mitari; Yuji Kaneda (2003). "Subject independent facial expression recognition with robust face detection using a convolutional neural network" (PDF). Neural Networks. 16 (5): 555–559. doi:10.1016/S0893- 6080(03)00115-1. Retrieved 17 November 2013.

[5] LeCun, Yann. "LeNet-5, convolutional neural networks". Retrieved 16 November

[6] C. Zor, “Facial expression recognition,” Master’s thesis, University of Surrey, Guildford, 2008.

[7] Suwa, M.; Sugie N. and Fujimora K. A Preliminary Note on Pattern Recognition Of Human Emotional Expression, Proc. International Joint Conf Pattern Recognition, pages 408-410, 1978

[8] Recognizing action units for facial expression analysis YI Tian, T Kanade, JF Cohn IEEE Transactions on pattern analysis and machine intelligence 23 (2), 97115.

[9] Raghuvanshi, Arushi, and Vivek Choksi. "Facial Expression Recognition with Convolutional Neural Networks." Stanford University, 2016

[10] Alizadeh, Shima, and Azar Fazel. "Convolutional Neural Networks for Facial Expression Recognition." Stanford University, 2016.

[11] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. ”Imagenet Classification with deep convolutional neural networks.” Advances in Neural information processing systems. 2012.

[12] Simonyan, Karen, and Andrew Zisserman. ”Very deep convolutional networks for large-scale image recognition.” arXiv preprint.arXiv:1409.1556 (2014).