IOSR Journal of Engineering (IOSRJEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
Special Issue || Dec. 2019 || PP 19-20
International Conference on Management Practices, Innovations & Research 2019

Convolutional Neural Network for Facial Expression Recognition

Vaishali Ingale, Rajesh Kumar, Neeraj Singh
Department of Information Technology Army Institute of Technology,Pune
Department of Information Technology Army Institute of Technology, Pune
Department of Information Technology Army Institute of Technology, Pune

Abstract: *In this project, we developed Convolutional Neural Networks (CNN) for a facial expression recognition task. Our goal is to classify each facial image into one of the seven facial emotion categories. Different depth structures of CNN models are trained using grayscale images from the Kaggle facial expression challenge [1]. To reduce overfitting, we utilized different techniques including dropout and L2 regularization (loss function).*

I. Introduction

In recent years, due to the fast-growing social networks, photos and videos that include faces constitute a great proportion of visual data over the Internet, using machine to do the facial recognition is becoming commonly used. Facial expression, as the most expressive and direct way to communicate emotion in humans, draws a lot of attractions. However, although facial expressions can be easily recognized by human beings, reliable facial expression recognition by machine is still a great challenge. Moreover, it is an interesting and challenging problem due to its wide range of applications such as human–computer interaction In this paper, we present an approach based on Convolutional Neural Networks (CNN) for facial expression recognition. The input into our system is an image , thenwe use CNN to predict the facial expression label which should be one these labels: anger, happiness, fear, sadness, disgust and neutral.

II. Related Work

In recent years, researchers have made considerable progress in developing automatic expression classifiers.Some expression recognition systems classify the face into a set of emotions such as happiness, sadness and anger.There have been several developments in the techniques used for facial expression recognition: Bayesian Networks, Neural Networks and the multilevel Hidden Markov Model (HMM). Some of them contain drawbacks of recognition rate or timing. Usually, to achieve accurate recognition two or more techniques can be combined; then, features are extracted as needed. The success of each technique is dependent on pre-processing of the images because of feature extraction.

III. Dataset

The dataset for this project is from Kaggle facial expression challenge [1], which is comprised of 48 × 48 pixel gray-scale images of human faces. The training set consists of 28,709 examples. Each image is categorized into one of the seven classes that express different facial emotions: "anger", "disgust", "fear", "happiness", "sadness", "surprise", and "neutral".

IV. Experiment

We developed CNNs with variable depths to evaluate the performance of these models for facial expression recognition. A typical CNN architecture contains all or some of the following layer types:
[Conv(ReLU) → Max-pooling] with Dropout× M → [Fully-connected(ReLU)]× N → Output

The first part of the network refers to the first kind of layers, of which usually contains Convolutional layer with ReLU activation function and Max-pooling layer. We can include spatial max-pooling, dropout and even batch normalization and ReLU nonlinearity in the first step. After using M(Training dataset) times of the first layers, the network is led to Fully-Connected layers that always have affine operationand ReLU nonlinearity, and can also include batch normalization and dropout. Finally, the network is followed by the affine layer connects to the class nodes, in which the scores are computed and then softmax (Keras loss function) is used to calculate the probability. This CNN model gives us a lot of freedom to customize the number of Convolutional and Fully-Connected layers, as well as the existence of dropout and batch normalization method. Furthermore, the number of filters, size of filters, strides, and zero-padding can be specified by users. Along with dropout techniques, we included L2 regularization in our implementation. We implemented the models from scratch, and then using TensorFlow and Keras to build deeper CNN.The model was trained on a NVIDIA 940MX GPU.

To observe the effect of adding convolutional layers and FC (fully connected) layers to the network, we trained a deeper CNNwith 4 convolutional layers and two FC layers. The first convolutional layer had 64 3×3 filters, the second one had 128 5×5 filters, the third one had 512 3×3 filters and the last one had 512 3×3 filters. In all the convolutional layers, we have a stride of size 1, batch normalization, dropout, max-pooling and ReLU as the activation function. The hidden layer in the first FC layers had 256 neurons and the second FC layer had 512 neurons. In both FC layers, same as in the convolutional layers, we used batch normalization, dropout and ReLU. Also we used Softmax as our loss function

To explore the deeper CNNs, we also trained networks with 5 and 6 convolutional layers, but they did not increase the classification accuracy. Therefor we considered the model with 4 convolutional layers and 2 FC layers as the best network for our dataset.

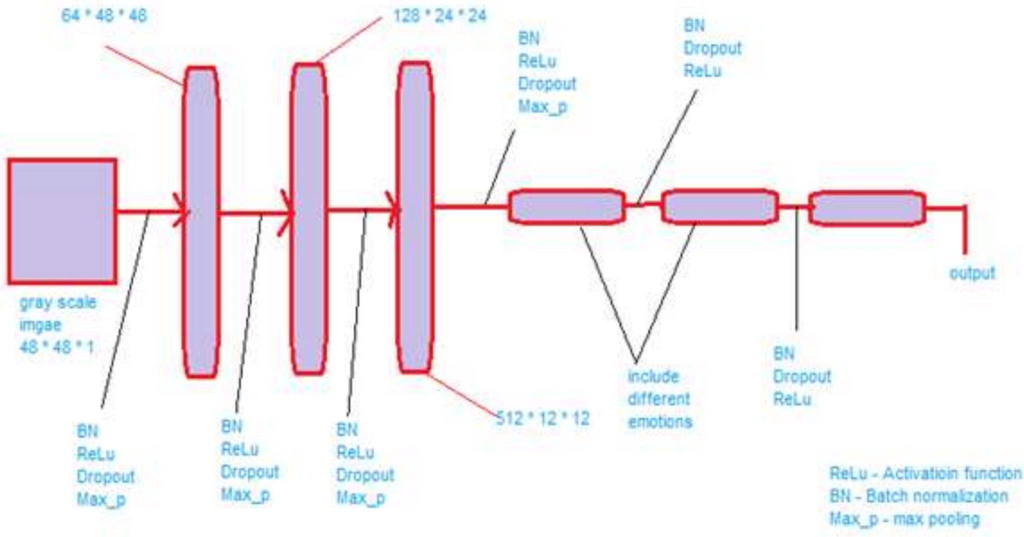

Figure 1: The architecture of the deep network: 4 convolutional layers and 2 fully connected layers

V. Conclusion

In this project, we built several Convolutional Neural Network to recognize facial expression from gray-scale pictures inspired by Kaggle facial expression recognition challenge [1]. We used different depth of Conv(ReLU) - Max-Pooling architecture and Softmax loss. For regularization, we experimented with L2 regularization and dropout in our final deeper CNN.Some of the difficulties with improving this is that the images are very small and in some cases.We have only explore certain architecture but we believe that adding more layers and more filters would further improve the network.

References

[1]. https://www.kaggle.com/c/challenges-inrepresentation-learning-facial-expressionrecognition-challenge/data

[2]. Dalal, Navneet, and Bill Triggs (2005). Histograms of oriented gradients for human detection. Computer Vision and Pattern Recognition (CVPR), IEEE Computer Society Conference on.

[3]. Bettadapura, Vinay (2012). Face expression recognition and analysis: the state of the art. arXiv preprint arXiv:1203.6722

[4]. Y. Tian, T. Kanade, and J. Cohn. Recognizing action units for facial expression analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(2), 2001.

[5]. Anbang Yao, Junchao Shao, Ningning Ma, and Yurong Chen. Capturing au-aware facial features and their latent relations for emotion recognition in the wild. In Proceedings of the 2015 ACM on International Conference on Multimodal Interaction, pages 451–458. ACM, 2015.

[6]. Samira Ebrahimi Kahou, Vincent Michalski, Kishore Konda, Roland Memisevic, and Christopher Pal. Recurrent neural networks for emotion recognition in video. In Proceedings of the 2015 ACM on International Conference on Multimodal Interaction, pages 467–474. ACM, 2015.

[7]. Deepali Aneja, Alex Colburn, Gary Faigin, Linda Shapiro, and Barbara Mones. Modeling stylized character expressions via deep learning. In Asian Conference on Computer Vision, pages 136–153. Springer, 2016.