Improving Convolutional Neural Network

Aliaksandr Hladkou
Dr. Alvin Grissom II
CS-477A: Artificial Intelligence
Ursinus Colloge
`GitHub Link`

November 25, 2019

Abstract

At the current time, there is a class of deep neural networks called Convolutional Neural Network (CNN) that is one of the most successful neural network in completing vision tasks such as image and video recognition. The purpose of this research work is to achieve highest possible accuracy of CNN using CIFAR-10 dataset from Keras library. As an example of improving the model, I will try to redesign the model. However, I will be focusing mostly on activation functions in order to achieve best possible accuracy result. In addition to personal research, there will be used several methods and concepts from other projects.

1 Introduction

Kunihiko Fukushima introduced *neocgnition* which included convolutional and downsampling layers. In addition, there was introduced a method called *max-pooling*. Max-poolong method takes downsampling units and computes the maximum of activations in its path.
A CNN structure is like many other Neural Network structures. It is formed by using a set of layers that are represented as a stack of layers and that convert input volume to output volume. To be more specific, a CNN has *convolutional layer, pooling layer, normalization layer, fully connected layer*, and *loss layer*. In addition, neurons arranged in 3D with *width, height, depth (or the number of channels)*.
As it was mentioned above, this work will be based on the CIFAR-10 dataset which has 10 classes. As an example of architecture, I will start with the input volume. It has the size of the input volume of 32x32x3 in which last dimension represents three (R, G, B) color channels. The convolution layer computes the output of neurons. Relu layer applies activation function in which Exponential Linear Unit (Elu) provided the best result. Normalization layer normalize the activations of the previous layer at each batch. In the next layer, the model

will max pooling operation for spatial data.Finally, in fully-connected layer the model will compute the class scores. It will also give volume of size 1x1x10.
After setting this model it was reasonable to try using different activation function in order to improve the model. Also, as another benefit it was a great to discover what effect to the model each activation function would make. By the end of the project, it will be demonstrated that *elu* activation function will provide the result of this model with accuracy of close to 90 percent.

2 Related Work

2.1 Convolutional Neural Network

By taking the standard model from Keras library, there were several test made. The base model called **cifar10cnn** which can be found in Keras library along with man other models. However, by discovering several sources I noticed that making several rearrangements in this model made better accuracy as well as loss results.
After training several times the first model, it can be found that the accuracy is below expectations. The following table represents few tests on the model:

Epochs	Activation	Test Accuracy
10	Relu	0.6491
25	Relu	0.7433
50	Relu	0.7921

As can be seen on the table above, test results on accuracy does not meet expected requirements. Therefore, by changing the learning rate and activation functions I tried to get better results on this model.

2.2 Activation Functions

Researched by Forest Agostinelly, Matthew Hoffman and Peter Sadowski, the CNN model in this project will implement several abstract ideas and methods for comparing different activation functions. Additionally, learning Abhijeets Kumar's method on achieving better accuracy using the same model, I was able to significantly improve the accuracy and the speed of that model.

3 Experiments

As it was shown in the previous section, test accuracy of the model did not provide expected accuracy. Therefore, the first step of rearrange the model was changing the learning rate based on the number of epochs. Also, I added the additional layer for batch normalization in order to apply a transformation that, as described in Keras documentation, maintains the mean activation close to 0 and the activation standard deviation close to 1. After running the several test, results were already better than of initial model:

Epochs	Activation	Test Accuracy
10	Relu	0.7937
25	Relu	0.8235
50	Relu	0.8312

After several experiments, test accuracy was significantly improved. Furthermore, training time of improved model decreased providing faster training for the model. Changing *Dropout* parameter to lower could be one of the reasons of improving the speed. In addition, it helped to prevent the network from overfitting.

Then, the decision was made toward *ELU* activation function. ELU is similar to the RELU activation function, except the fact that ELU can produce negative inputs and more smooth graph. According to the definition, ELU tend to converge cost to zero faster and produce more accurate results:

Epochs	Activation	Test Accuracy
10	Elu	0.8141
25	Elu	0.8335
50	Elu	0.8678

Clearly, test accuracy much better with ELU activation function. It slightly affected the time to complete training of each epoch longer. Also, it was important to add *z-score* because the gradients are always under control.

4 Conclusion

By making several experiment with CNN model and changing the activation functions, it can be concluded that it is possible to achieve the accuracy above 80 percent. By adding batch normalization, changing learning rate and changing the dropout to prevent from overfilling it is also possible to reduce the amount of times required for training. In addition, these tests shows the improved performance in deep neural networks.

References

[1] Forest Agostinelli & Matthew Hoffman & Peter Sadowski(2015) LEARNING ACTIVATION FUNCTIONS TO IMPROVE DEEP NEURAL NETWORKS.

[2] Abhijeet Kumar (March 24, 2018) Achieving 90 percent accuracy in Object Recognition Task on CIFAR-10 Dataset with Keras: Convolutional Neural Networks

[3] Hasselmo, M.E., Schnell, E. & Barkai, E. (1995) Dynamics of learning and recall at excitatory recurrent synapses and cholinergic modulation in rat hippocampal region CA3. *Journal of Neuroscience* **15**(7):5249-5262.