Deep Learning
Practical 1

Conaire O Codlatin Lachtna (s3524302) [a]   Kaustav Datta (s3573141) [b]

November 1, 2019

Abstract

We study the performance of different deep architectures like the AlexNet and ResNet50 on the CIFAR-10 dataset. We examine the performance of these neural networks with different settings such as dropout and activation functions.

1 Introduction

Image classification is the task of taking an image as an input and giving as an output the class (or probability of the class) the image belongs to. Convolutional Neural Networks (CNNs) are a class of Deep Neural Networks which have been proven to perform very well in image classification tasks. A CNN usually consists of the following types of layers - convolutional layers, ReLU layers, pooling layers and fully connected layers. In this paper we will explore the performance of a couple of CNNs on an image classification task.

2 The Data Set

The CIFAR-10 dataset consists of 60000 32x32 colour images in 10 mutually exclusive classes, such as automobile, truck and bird with 6000 images per class. There are 50000 training images and 10000 test images [1].

3 CNN Architectures

3.1 AlexNet Inspired Network

AlexNet [5] is a famous CNN designed by Alex Krizhevsky. This network was the winner of the 2012 ImageNet Large Scale Visual Recognition Challenge with a top-5 error of 15.3%, 10.8% better than the next best score.

The input size of the original AlexNet was 256x256. Here we use a modified version [3] to accept CIFAR-10 images which are of size 32x32. To allow for the smaller input size the kernel sizes were reduced for the convolution layers and the pool sizes used for max pooling. Batch normalization is also used instead of local response normalization as it is said to be more effective.

3.2 Residual Neural Network (ResNet)

Residual Neural Networks were introduced at the ILSVRC 2015 [2] by Kaiming He et al [4]. It was a novel architecture featuring skip connections and heavy batch normalization. The biggest advantage of this neural network was that it made the training of substantially deep networks less complex than previous networks like VGGNet [6].

For our study, we used a particular variant of the ResNet called ResNet50 which is 50 layer Residual Network.

4 Comparison Techniques

The performance of the CNNs were compared by using different combinations of settings such as batch normalization, dropout, activation functions and also using weights pre-trained on ImageNet images.

Activation functions are an important part of neural networks as they can be used to introduce the non-linearity required to estimate non-linear functions. In this study we compare the effectiveness of four non-linear activation functions.

first activation function we will use is sigmoid. This function was one of the first to be widely used. A problem this function has is that as you move further from the origin the rate of change of the output becomes very small with respect to the input which leads to ”vanishing gradients”. This means that at a certain point the network can't ”learn” anymore.

Second, Rectified Linear Units (ReLU) is one of the most popular activation functions today. Part of this reason for this is that it is resistant to the vanishing gradients problem

Third, Exponential Linear Unit (ELU) is a variation on RelU that tends to converge faster and produce more accurate results.

The last function we will look at is tanh. This is also said to be more effective than sigmoid as the gradients is stronger due to steeper derivatives.

Batch normalization increases the stability of the network by normalizing the output of the previous activation layer. For example, if a network is trained only on red cars batch normalization, will make the network more accurate if it is tasked with recognizing a car that is white. It works by subtracting the batch mean from the output and dividing by the batch output.

Dropout is a technique that reduces overfitting by setting the output of each hidden neuron to 0 with a probability set by the dropout rate hyper-parameter, in this case we used 0.2. This reduces ”complex co-adaptations of neurons, since a neuron cannot rely on the presence of particular other neurons” which in turn makes the network learn more robust features that are useful with more types of random input.

The networks were trained on 50000 samples and validated on 10000 samples. They were trained across 20 epochs.

5 Results

Weights pre-trained on ImageNet				
Model	Activation	Normalization	Dropout	Accuracy(%)
ResNet50	ReLU	✓	✓	96.11
ResNet50	sigmoid	✓	✓	93.67
ResNet50	ELU	✓	✓	96.45
ResNet50	tanh	✓	✓	94.08
ResNet50	ReLU	✓		95.99
ResNet50	sigmoid	✓		94.82
ResNet50	ELU	✓		96.41
ResNet50	tanh	✓		93.87

Randomly initialized weights				
Model	Activation	Normalization	Dropout	Accuracy(%)
ResNet50	ReLU	✓	✓	92.43
ResNet50	sigmoid	✓	✓	92.8
AlexNet	ReLU	✓	✓	64.62
AlexNet	sigmoid	✓	✓	35.79
AlexNet	ELU	✓	✓	66.73
AlexNet	tanh	✓	✓	28.64
AlexNet	ReLU			51.75
AlexNet	ReLU		✓	55.84
AlexNet	ReLU	✓		64.08
AlexNet	ReLU	✓	✓	64.62

6 Discussion

We can see that ResNet50 beats the AlexNet implementation. This is because of the advantages which comes with the design of Residual Networks like less layers which solves the problem of vanishing gradients and the degradation problem. We also see better results with dropout regularization because it helps the network generalize better. ELU was the most effective activation function, followed by ReLU. There was no clear winner between tanh and sigmoid. Finally, we also see that models pre-trained on ImageNet weights out-perform models which are trained from scratch with random weights.

References

[1] The cifar-10 dataset. `https://www.cs.toronto.edu/~kriz/cifar.html`.

[2] Large scale visual recognition challenge 2015. `http://image-net.org/challenges/LSVRC/2015/`.

[3] Robert Crowe. google-alexnet. `https://github.com/robertcrowe/google-alexnet/`.

[4] Kaiming He et al. Deep residual learning for image recognition. 2015.

[5] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 25*, pages 1097–1105. Curran Associates, Inc., 2012.

[6] Simonyan and Zisserman. Very deep convolutional networks for large-scale image recognition. 2015.