Assignment 3

Karthik Kumarasubramanian

November 11 2019

1 Introduction

The CIFAR-10 data consists of 60,000 (32 X 32) color images in 10 classes, with 6000 images per class. There are 50,000 training images and 10,000 test images in the official data. The label classes in the data-set are:

- Airplane.
- Automobile.
- Bird.
- Cat.
- Deer.
- Dog.
- Frog.
- Horse.
- Ship.
- Truck.

2 Architecture

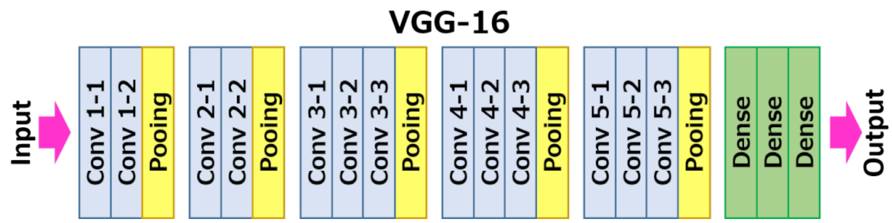

Figure 1: VGG-16 Architecture.

The architecture used for classification is loosely based of VGG-16 architecture. Convolution layer is set of 3 operations: Convolution, Activation and Batch normalization. Sometimes, Dropout layer is kept after Pooling in lieu of regularization. Also, Multiple dense layer can be kept after flattening layer before finally keeping output dense layer. The image is passed through a stack of convolutional (conv.) layers, where the filters were used with a very small receptive field: 3 X 3 . Spatial pooling is carried out by max-pooling layers, which follow some of the conv. layers. Max-pooling is performed over a 2 X 2 pixel window, with stride 2. The output layer is dense layer of 10 nodes (as there are 10 classes) with softmax activation. Regularization becomes important as number of parameters (weights) increases in order to do learning of weights from memorization of features towards generalization of features.

3 Regularization

Deep neural nets with a large number of parameters are very powerful machine learning systems. However, over-fitting is a serious problem in such networks.

3.1 Dropout

Dropout is a technique for addressing this problem. The key idea is to randomly drop units from the neural network during training. The reduction in number of parameters in each step of training has effect of regularization. In this method variant dropouts are used after Max-Pooling layer.

3.2 Kernel Regularizer

allows to apply penalties on layer parameters during optimization. These penalties are incorporated in the loss function that the network optimizes. This argument in convolutional layer is nothing but L2 regularisation of the weights. This penalizes peaky weights and makes sure that all the inputs are considered. During gradient descent parameter update, the above L2 regularization ultimately means that every weight is decayed linearly.

3.3 Batch Normalization

Normalizes the activation of the previous layer at each batch, i.e. applies a transformation that maintains the mean activation close to 0 and the activation standard deviation close to 1.

4 Optimization

The optimization method used for this classification is RMSprop. The RMSprop optimizer is similar to the gradient descent algorithm with momentum. The RMSprop optimizer restricts the oscillations in the vertical direction. The difference between RMSprop and gradient descent is on how the gradients are calculated.

5 Results

The accuracy of the model is proven to be at 80.5 % and the top-k accuracy is proven to be at 52 %. The Mean Average Precision (mAP) proven to be at 87 %.

6 References

- https://towardsdatascience.com/cifar-10-image-classification-in-tensorflow-5b501f7dc77c
- https://medium.com/@udolf15/building-a-image-classifier-using-cnn-with-cifar-10-dataset-5682afa4f51
- https://appliedmachinelearning.blog/2018/03/24/achieving-90-accuracy-in-object-recognition-task-on-cifar-10-dataset-with-keras-convolutional-neural-networks
- https://stackoverflow.com/questions/51951358/keras-how-to-get-top-k-accuracy