Introduction to Deep Learning for Computer Vision Assignment 6: Debug a deep neural network

Parm Johal
V00787710

March 28, 2020

Abstract

This report focuses on debugging a convolutional neural network during run time. The first part deals with a bug amongst the hyperparameters where a value is set too high. The second bug is found in the forward function, where the normalization step is incorrect. The third bug deals with an issue with what mode the model should be in before entering the test loop.

2.1 $Hyperparameters$

In this bug the learning rate and repetitions interval are too close together.

In particular, the learning rate is too high. This will make the model learn at a much faster rate but will introduce divergent behaviour (or in this case a warning of having NaN or inf in the input tensor). Even if there aren't any warnings, a high learning rate can still potentially give us suboptimal results.

2.2 $model.py/MyNetwork/forward$

In this bug the data wasn't properly normalized.

By attempting to normalize x with

$$x = \frac{(x - 127.5)}{127.5}$$

some of Bob's tensor values were outside the range of [-1, 1]. This could be fixed by picking a larger value than 127.5.

2.3 $solution.py/test$

For this bug, Bob forgot to set the model to evaluation mode before entering the test loop by not writing 'model.eval()'.

Since Bob is using batch normalization, He needs to set the model to evaluation mode so that the running mean/variance is used for normalization instead of the batch statistics from training mode. This is why Bob gets a lower accuracy with a smaller batch size.