Deep Learning
Assignment 3 Report

Tsinghua University

Albert Millan – 2019280366

Contents

1 Introduction

This report explains the setup and methodology involved to deploy convolutional neural network with cross-entropy loss and dropout.

2 ConvNet

The convolutional layer was implemented using im2col and col2im methods. Rather than performing convolution operations on an element-per-element basis, the aforementioned methods group all elements involved in the operation into a single matrix. Successively, it performs a dot product by the stretched filters. The speed-up increase can be up to $\tilde{2}30$ times faster than a naive implementation using for-loops. Upon successful implementation of the loss function and gradient computation, an accuracy on the test set of 93.33% was obtained with the provided hyperparameters. The model was further tuned, achieving a final accuracy of 96.15%. Table 1 outlines the values of the hypermarameters employed to obtain such result. Considering the model was far from converging with the initial parameters, the number of training iterations where doubled to 20 and learning rate set to 0.003. This model was then re-trained with an additional dropout layer after the first fully-connected layer, with a dropout ratio of 0.5.

Type	**Initial Value**	**Tuned Values**
Batch size	100	100
Learning rate	0.001	0.003
Weight decay	0.005	0.005
Max epoch	10	20

Table 1: Hyperparameters employed to train the model.

3 ConvNet vs MLP

The time per epoch required to train the ConvNet is significantly larger than that of MLP, despite using speed-up algorithms. Notwithstanding, the convergence rate for ConvNet is steeper at early stages of the training process. ConvNet also converges at a smaller loss distance compared to its counterpart MLP under cross-entropy loss, while also achieving a slightly superior accuracy when fine-tuned.

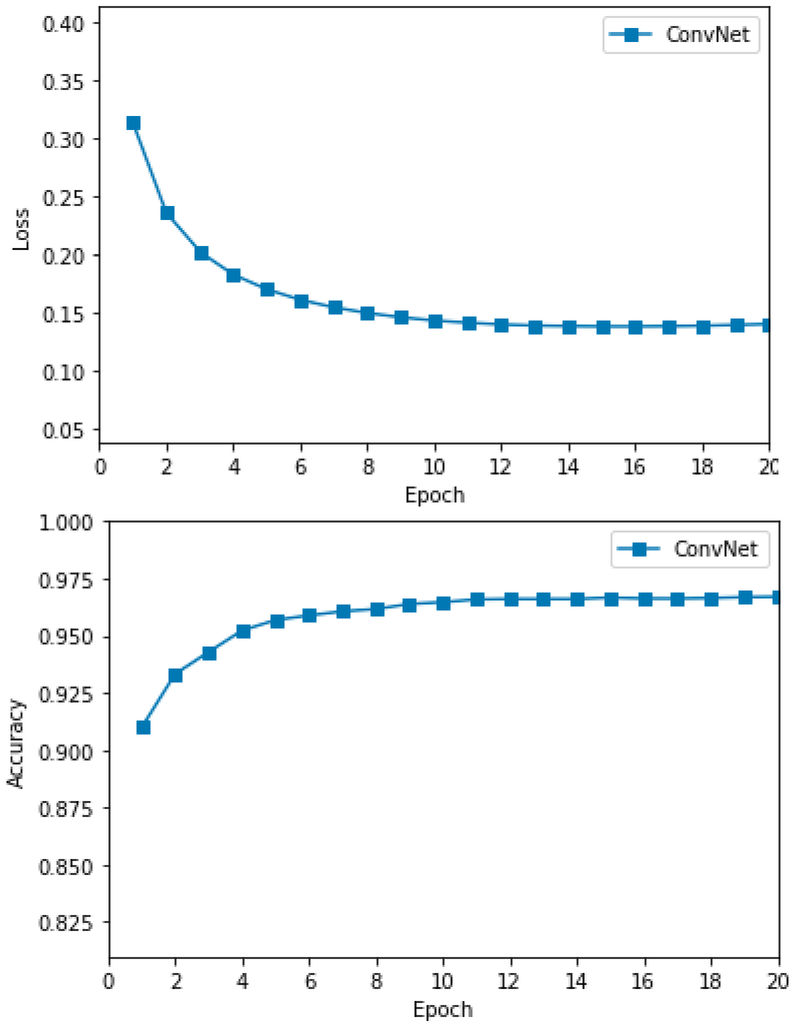

Figure 1: CNN loss and accuracy over epochs.