Deep Learning Paper Template

Priyank Bhatia
New York University
Center for Urban Science + Progress
1 MetroTech Center, 19th Floor
Brooklyn, NY 11201
pb1672@nyu.edu

Emil Christensen
New York University
Center for Urban Science + Progress
1 MetroTech Center, 19th Floor
Brooklyn, NY 11201
erc399@nyu.edu

Peter Varshavsky
New York University
Center for Urban Science + Progress
1 MetroTech Center, 19th Floor
Brooklyn, NY 11201
pv629@nyu.edu

Abstract

The abstract paragraph goes here!

1 Introduction

The architecture of the model provided is a convolutional neural network with two stages and a classifier. The input has 3 feature maps, each 32x32 pixels.The first stage performs convolutions using a tanh squashing function with a 5x5 filter to produce 64 feature maps to which L^2 pooling with pooling size 2x2 is applied. A subtractive normalization module using a Gaussian kernel of size 7 is then applied before feeding the 64 14x14 outputs to the next stage.The second stage performs the convolutions and poolings with the same filter sizes, pool sizes, number of feature maps, and normalization function as the first stage except that the 16-dim feature maps are projected into 256-dim maps before feeding the resulting data into a 2-layer non linear classifier with 128 hidden units, which uses the tanh function for non linear transformation. A Class Negative Log-Likelihood (nll) Loss function is used along with Stochastic Gradient Descent (SGD) for optimization and learning rate of 1e-3, mini-batch size of 1, and momentum 0 number of max iterations is equal to 1. Despite our attempts to improve the model, we found that the original model (as provided by Clement Farabets tutorial) yielded the best results. Using dataset cited: [1]

2 Architecture

3 Learning Techniques

4 Training Procedure

5 Results

References

[1] Yuval Netzer, Tao Wang, Adam Coates, Alessandro Bissacco, Bo Wu, and Andrew Y. Ng. Reading digits in natural images with unsupervised feature learning. In *NIPS Workshop on*

Deep Learning and Unsupervised Feature Learning 2011, 2011.