Assignment 2: Classifying STL-10 with a deep feed-forward convolutional neural net

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 Data

Designed to study the use of unlabeled data for image classification, the STL-10 dataset [1] consists of three sets of 96x96 pixel RGB color images: 5000 labeled training images, 8000 labeled test images, and 100000 unlabeled images. Each image belongs to one of 10 categories. In this submission we only utilize the labeled portion of STL-10 and attempt to improve classification quality by image augmentation and depth of network.

2 Architecture

Two versions of a feed-forward convolutional neural net architectures were compared. The first architecture was an implementation of the baseline model by Christian Puhrsch with one convolutional layer of 23 7x7 pixel learned filters with a stepwise of 2, ReLU nonlinearity, 2 sq pixel pooling with step size 2, 50% dropout, a 50-node fully connected layer, pointwise ReLU and LogSoftMax and, and a negative log likelihood criterion. The second model used the same basic architecture but consisted of two convolutional layers of 200 5x5 pixel filters each, and a 400-neuron fully connected linear layer. Spatial pooling, rectified linear unit nonlinearities and dropout were applied as in the first network.

3 Preprocessing and augmentation

The original 5000 training images were split into training and validation sets of sizes 4500 and 500. To ameliorate the small training size and improve feature invariance the 4500 training images were cloned twice. The first cloned set was flipped horizontally, and the second was rotated counter-clockwise by 0.35 radians. This yielded an augmented training set of 13500 images. We further attempted to augment the training set using contrast HSV color space adjustments similar to contrast2 in [2], small random translations and rotations, but ran into training convergence issues, possibly due to coding errors. Augmented data were converted to YUV color space. Training images were globally normalized. Validation and test images were globally normalized using training mean and standard deviation. All images were further locally normalized and given a 2-pixel zero pad.

4 Learning Techniques

Forward and back prop. Learning rate 0.1. Rate annealing 0.001 (check), metaparameters determined with validation.

5 Training Procedure

6 Results

Model 1 - One convolutional layer			
	train error	validation error	test error
train with validation	...	...	...
train without validation	...	...	...

Model 2 - Two convolutional layers			
	train error	validation error	test error
train with validation	87.82%	60.0%	61.69%
train without validation	...	N/A	...

References

[1] A. Coates, H. Lee, and A.Y. Ng. An analysis of single-layer networks in unsupervised feature learning. In Geoffrey Gordon, David Dunson, and Miroslav Dudk, editors, *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, volume 15 of *JMLR Workshop and Conference Proceedings*, pages 215–223. JMLR W&CP, 2011.

[2] Alexey Dosovitskiy, Jost Tobias Springenberg, Martin Riedmiller, and Thomas Brox. Discriminative unsupervised feature learning with convolutional neural networks. *CoRR*, abs/1406.6909, 2014.