Deep Representation Learning with Target Coding Supplementary Material

Shuo Yang[1]**, Ping Luo**[2,1]**, Chen Change Loy**[1,2]**, Kenneth W. Shum**[1]**,** and **Xiaoou Tang**[1,2]
[1]Department of Information Engineering, The Chinese University of Hong Kong
[2]Shenzhen Key Lab of CVPR, Shenzhen Institutes of Advanced Technology,
Chinese Academy of Sciences, Shenzhen, China
{ys014,pluo,ccloy,xtang}@ie.cuhk.edu.hk, {wkshum}@inc.cuhk.edu.hk

Abstract

This document provides the details of the CNN structure and parameters used in our paper, to facilitate future reimplementing of our work. We also describe our data processing methods for each dataset. In addition, the Hadamard Code with different lengths is attached at end.

CNN Structure and Learning Parameters

STL-10, MNIST

We used the same network structure on STL-10, MNIST following existing works (Srivastava and Salakhutdinov 2013) and removed two fully-connected layers. The deep network has three convolutional layers each followed by a max pooling layer. The last hidden layer is a locally-connected layer. The dropout rate for the first two convolutional layers is 0.25. For the last convolutional layer and locally-connected layer the dropout rate is 0.4 or 0.5. In fact, there are no significant difference between them.

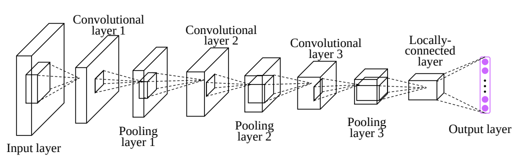

Figure 1: The network structure of STL-10, MNIST.

CIFAR-100

The CIFAR-100 deep network following existing works (Srivastava and Salakhutdinov 2013) has three convolutional layers, three max pooling layer, one local connected layer and two fully-connected layers. The dropout rate for the first two convolutional layers is 0.25. For the last convolutional layer, local connected layer and two fully-connected layers the dropout rate is 0.5.

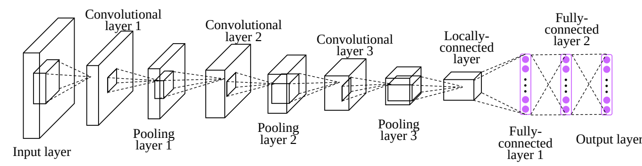

Figure 2: The network structure of CIFAR-100.

ImageNet-2012

The ImageNet-2012 deep network following existing works (Krizhevsky, Sutskever, and Hinton 2012; Zeiler and Fergus 2014) has five convolutional layers, three max pooling layers and two fully-connected layers. The dropout rate for the two fully-connected layers the dropout rate is 0.5.

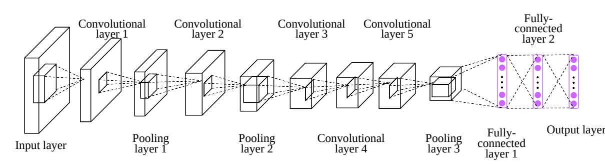

Figure 3: The network structure of ImageNet-2012.

Datasets

Overview

MNIST Following Larochelle et al (Larochelle et al. 2007), we use the MNIST extension dataset which is generated with additional factors of variation to the MNIST images. Specifically, the MNIST-rot-back-image generate with two procedures. The digits were first rotated by an angle generated uniformly between 0 and 2π radians, then a random patch from a black and white image was used as the background for the digit image. The MNIST-rot-back-image consists of 12000 training images and 50000 testing images.

STL-10 STL-10 dataset is an image recognition dataset with 96×96 natural RGB images in 10-classes. Compared with CIFAR-10 dataset STL-10 dataset has less labeled data, it has 5000 training images with 500 images per class and 8000 testing images with 800 per class. It has 100,000 unlabeled data for the unsupervised learning.

CIFAR-100 The CIFAR-100 dataset consists of 60000 natural 32×32 RGB images belonging to 100 classes. This dataset contains 50000 training images and 10000 test images. There are 500 training images and 100 testing images per class. 100 classes are divided into 20 groups of 5 classes each. For example, the superclass flowers contains orchids, poppies, roses, sunflowers, tulips, and the superclass insects contains bee, beetle, butterfly, caterpillar, cockroach.

ImageNet-2012 The ImageNet-2012 dataset consists of 1.35 million natural RGB images belonging to 1000 classes. This dataset contains 1.2 million training images and 150,000 validation test images. There are about 1000 training images and 100 testing images per class. The validation and test data are not contained in the training data.

Data preprocessing

All the processed data were used on 1-of-k and Hadamard Code based CNN model.

MNIST No preprocessing.

STL-10 Deep model training following the pre-defined STL-10 partition is infeasible due to the small number of samples in each fold. Therefore, deep learning studies (Coates and Ng 2011) tend to adopt different protocols for STL10. Note that our main goal is to compare the effectiveness of Hadamard coding vs. 1-of-K coding. We trained both coding schemes using the same data partition. Following the augmentation strategy in Hinton's (Krizhevsky, Sutskever, and Hinton 2012) work, we first resized the original images to 64×64 then randomly cropped 48×48 patches 20 times from resized images and test on the center 48×48 patch. We only used the labeled data and ignored the 100,000 items of unlabeled data.

CIFAR-100 Following the same process with STL-10, we first resized the original images to 40×40 then randomly cropped 32×32 patches 15 times from the resized images and test on the center 32×32 patch.

ImageNet-2012 Following the same process with Hinton's (Krizhevsky, Sutskever, and Hinton 2012) work, we first resized the original images to 256×256 then cropped 224×224 patches from four corners and the center of the resized images then flipped the cropped patches.

Hadamard Code(HC) Codewords

HC-63, HC-127, HC-255 for 10 classes

For HC-63 every two rows in the table represent one codeword. Similar with HC-63, a codeword in HC-127 and HC-255 has four rows and eight rows respectively.

Table 7: HC-63.

1 1 1 0 1 0 1 0 1 1 0 0 1 1 0 1 1 1 0 1 1 0 1 0 0 1 0 0 1 1 1 0 0 0 1 0
1 1 1 1 0 0 1 0 1 0 0 0 1 1 0 0 0 0 1 0 0 0 0 0 1 1 1
0 0 0 0 0 1 1 1 1 1 1 0 1 0 1 0 1 1 0 0 1 1 0 1 1 1 0 1 1 0 1 0 0 1 0 0
1 1 1 0 0 0 1 0 1 1 1 1 0 0 1 0 1 0 0 0 1 1 0 0 0 0 1
1 1 1 1 0 0 1 0 1 0 0 0 1 1 0 0 0 0 1 0 0 0 0 0 1 1 1 1 1 1 0 1 0 1 0 1
1 0 0 1 1 0 1 1 1 0 1 1 0 1 0 0 1 0 0 1 1 1 0 0 0 1 0
0 0 0 1 0 0 0 0 0 1 1 1 1 1 1 0 1 0 1 0 1 1 0 0 1 1 0 1 1 1 0 1 1 0 1 0
0 1 0 0 1 1 1 0 0 0 1 0 1 1 1 1 0 0 1 0 1 0 0 0 1 1 0
0 0 0 0 1 0 0 0 0 0 1 1 1 1 1 1 0 1 0 1 0 1 1 0 0 1 1 0 1 1 1 0 1 1 0 1
0 0 1 0 0 1 1 1 0 0 0 1 0 1 1 1 1 0 0 1 0 1 0 0 0 1 1
0 1 0 0 0 0 0 1 1 1 1 1 1 0 1 0 1 0 1 1 0 0 1 1 0 1 1 1 0 1 1 0 1 0 0 1
0 0 1 1 1 0 0 0 1 0 1 1 1 1 0 0 1 0 1 0 0 0 1 1 0 0 0
1 1 1 1 1 1 0 1 0 1 0 1 1 0 0 1 1 0 1 1 1 0 1 1 0 1 0 0 1 0 0 1 1 1 0 0
0 1 0 1 1 1 1 0 0 1 0 1 0 0 0 1 1 0 0 0 0 1 0 0 0 0 0
1 1 1 0 0 1 0 1 0 0 0 1 1 0 0 0 0 1 0 0 0 0 0 1 1 1 1 1 1 0 1 0 1 0 1 1
0 0 1 1 0 1 1 1 0 1 1 0 1 0 0 1 0 0 1 1 1 0 0 0 1 0 1
1 0 1 0 1 1 0 0 1 1 0 1 1 1 0 1 1 0 1 0 0 1 0 0 1 1 1 0 0 0 1 0 1 1 1 1
0 0 1 0 1 0 0 0 1 1 0 0 0 0 1 0 0 0 0 0 1 1 1 1 1 1 0
0 0 0 1 1 1 1 1 1 0 1 0 1 0 1 1 0 0 1 1 0 1 1 1 0 1 1 0 1 0 0 1 0 0 1 1
1 0 0 0 1 0 1 1 1 1 0 0 1 0 1 0 0 0 1 1 0 0 0 0 1 0 0

Table 8: HC-127.

1 1 0 0 1 0 1 0 1 0 1 1 1 1 1 1 1 0 0 0 0 0 0 1 0 0 0 0 0 1 1 0 0 0 0 1
0 1 0 0 0 1 1 1 1 0 0 1 0 0 0 1 0 1 1 0 0 1 1 1 0 1 0 1 0 0 1 1 1 1 1 0
1 0 0 0 0 1 1 1 0 0 0 1 0 0 1 0 0 1 1 0 1 1 0 1 0 1 1 0 1 1 1 1 0 1 1 0
0 0 1 1 0 1 0 0 1 0 1 1 1 0 1 1 1 0 0
1 1 1 1 0 0 0 0 0 0 1 0 0 0 0 0 1 1 0 0 0 0 1 0 1 0 0 0 1 1 1 1 0 0 1 0
0 0 1 0 1 1 0 0 1 1 1 0 1 0 1 0 0 1 1 1 1 1 0 1 0 0 0 0 1 1 1 0 0 0 1 0
0 1 0 0 1 1 0 1 1 0 1 0 1 1 0 1 1 1 1 0 1 1 0 0 0 1 1 0 1 0 0 1 0 1 1 1
0 1 1 1 0 0 1 1 0 0 1 0 1 0 1 0 1 1 1
1 0 1 0 0 0 1 1 1 1 0 0 1 0 0 0 1 0 1 1 0 0 1 1 1 0 1 0 1 0 0 1 1 1 1 1
0 1 0 0 0 0 1 1 1 0 0 0 1 0 0 1 0 0 1 1 0 1 1 0 1 0 1 1 0 1 1 1 1 0 1 1
0 0 0 1 1 0 1 0 0 1 0 1 1 1 0 1 1 1 0 0 1 1 0 0 1 0 1 0 1 0 1 1 1 1 1 1
1 0 0 0 0 0 0 1 0 0 0 0 0 1 1 0 0 0 0
1 0 1 0 1 1 1 1 1 1 1 0 0 0 0 0 0 1 0 0 0 0 0 1 1 0 0 0 0 1 0 1 0 0 0 1
1 1 1 0 0 1 0 0 0 1 0 1 1 0 0 1 1 1 0 1 0 1 0 0 1 1 1 1 1 0 1 0 0 0 0 1
1 1 0 0 0 1 0 0 1 0 0 1 1 0 1 1 0 1 0 1 1 0 1 1 1 1 0 1 1 0 0 0 1 1 0 1
0 0 1 0 1 1 1 0 1 1 1 0 0 1 1 0 0 1 0
0 0 0 1 1 0 1 0 0 1 0 1 1 1 0 1 1 1 0 0 1 1 0 0 1 0 1 0 1 0 1 1 1 1 1 1
1 0 0 0 0 0 0 1 0 0 0 0 0 1 1 0 0 0 0 1 0 1 0 0 0 1 1 1 1 0 0 1 0 0 0 1
0 1 1 0 0 1 1 1 0 1 0 1 0 0 1 1 1 1 1 0 1 0 0 0 0 1 1 1 0 0 0 1 0 0 1 0
0 1 1 0 1 1 0 1 0 1 1 0 1 1 1 1 0 1 1
0 1 1 1 0 1 0 1 0 0 1 1 1 1 1 0 1 0 0 0 0 1 1 1 0 0 0 1 0 0 1 0 0 1 1 0
1 1 0 1 0 1 1 0 1 1 1 1 0 1 1 0 0 0 1 1 0 1 0 0 1 0 1 1 1 0 1 1 1 0 0 1
1 0 0 1 0 1 0 1 0 1 1 1 1 1 1 1 0 0 0 0 0 0 1 0 0 0 0 0 1 1 0 0 0 0 1 0
1 0 0 0 1 1 1 1 0 0 1 0 0 0 1 0 1 1 0
1 0 1 0 0 1 0 1 1 1 0 1 1 1 0 0 1 1 0 0 1 0 1 0 1 0 1 1 1 1 1 1 1 1 0 0 0
0 0 0 1 0 0 0 0 0 1 1 0 0 0 0 1 0 1 0 0 0 1 1 1 1 0 0 1 0 0 0 1 0 1 1 0
0 1 1 1 0 1 0 1 0 0 1 1 1 1 1 0 1 0 0 0 0 1 1 1 0 0 0 1 0 0 1 0 0 1 1 0
1 1 0 1 0 1 1 0 1 1 1 1 0 1 1 0 0 0 1
0 0 1 0 0 0 0 0 1 1 0 0 0 0 1 0 1 0 0 0 1 1 1 1 0 0 1 0 0 0 1 0 1 1 0 0
1 1 1 0 1 0 1 0 0 1 1 1 1 1 0 1 0 0 0 0 1 1 1 0 0 0 1 0 0 1 0 0 1 1 0 1
1 0 1 0 1 1 0 1 1 1 1 0 1 1 0 0 0 1 1 0 1 0 0 1 0 1 1 1 0 1 1 1 0 0 1 1
0 0 1 0 1 0 1 0 1 1 1 1 1 1 1 0 0 0 0
0 1 1 1 1 1 1 1 0 0 0 0 0 0 1 0 0 0 0 0 1 1 0 0 0 0 1 0 1 0 0 0 1 1 1 1
0 0 1 0 0 0 1 0 1 1 0 0 1 1 1 0 1 0 1 0 0 1 1 1 1 1 0 1 0 0 0 0 1 1 1 0
0 0 1 0 0 1 0 0 1 1 0 1 1 0 1 0 1 1 0 1 1 1 1 0 1 1 0 0 0 1 1 0 1 0 0 1
0 1 1 1 0 1 1 1 0 0 1 1 0 0 1 0 1 0 1
0 1 1 1 0 0 1 1 0 0 1 0 1 0 1 0 1 1 1 1 1 1 1 0 0 0 0 0 0 1 0 0 0 0 0 1
1 0 0 0 0 1 0 1 0 0 0 1 1 1 1 0 0 1 0 0 0 1 0 1 1 0 0 1 1 1 0 1 0 1 0 0
1 1 1 1 1 0 1 0 0 0 0 1 1 1 0 0 0 1 0 0 1 0 0 1 1 0 1 1 0 1 0 1 0 1 1 0 1 1
1 1 0 1 1 0 0 0 1 1 0 1 0 0 1 0 1 1 1

Table 9: HC-255-1.

1 1 0 0 0 1 1 1 0 0 0 0 1 1 0 0 0 0 0 1 1 1 0 1 1 0 0 0 0 1 0 1 0 1 1 0
0 1 0 0 1 1 1 0 0 1 1 1 0 1 0 1 0 1 1 1 1 1 1 1 1 0 1 1 0 1 1 0 0 1 1 1
1 0 0 0 1 1 0 1 0 1 1 1 0 0 1 0 0 0 0 1 1 1 1 0 1 1 1 0 1 1 1 1 0 1 0 0
0 0 0 1 0 0 0 0 0 0 1 0 1 1 0 1 1 1 1 1 0 0 1 1 0 1 1 1 0 0 0 1 0 1 1 1
0 1 0 0 1 1 0 0 1 0 1 0 1 0 1 0 0 1 0 0 1 0 0 0 1 0 1 0 0 0 0 1 0 0 1 1
0 1 0 0 0 1 1 1 1 1 0 1 0 1 1 0 1 0 0 1 0 1 0 0 1 1 1 1 1 1 0 0 0 0 0 0
0 1 1 0 1 1 0 1 0 1 0 0 0 1 0 0 1 0 1 1 1 1 0 0 1 0 1 1 0 0 0 1 0 0 0 1
1 0 0
1 1 0 1 1 1 0 1 1 1 1 0 1 0 0 0 0 0 1 0 0 0 0 0 0 1 0 1 1 0 1 1 1 1 1 0
0 1 1 0 1 1 1 0 0 0 1 0 1 1 1 0 1 0 0 1 1 0 0 1 0 1 0 1 0 1 0 0 1 0 0 1
0 0 0 1 0 1 0 0 0 0 1 0 0 1 1 0 1 0 0 0 1 1 1 1 1 0 1 0 1 1 0 1 0 0 1 0
1 0 0 1 1 1 1 1 1 0 0 0 0 0 0 0 1 1 0 1 1 0 1 0 1 0 0 0 1 0 0 1 0 1 1 1
1 0 0 1 0 1 1 0 0 0 1 0 0 0 1 1 0 0 1 1 0 0 0 1 1 1 0 0 0 0 1 1 0 0 0 0
0 1 1 1 0 1 1 0 0 0 0 1 0 1 0 1 1 0 0 1 0 0 1 1 1 0 0 1 1 1 0 1 0 1 0 1
1 1 1 1 1 1 1 0 1 1 0 1 1 0 0 1 1 1 1 0 0 0 1 1 0 1 0 1 1 1 0 0 1 0 0 0
0 1 1
0 1 0 0 1 1 0 0 1 0 1 0 1 0 1 0 0 1 0 0 1 0 0 0 1 0 1 0 0 0 0 1 0 0 1 1
0 1 0 0 0 1 1 1 1 1 0 1 0 1 1 0 1 0 0 1 0 1 0 0 1 1 1 1 1 1 0 0 0 0 0 0
0 1 1 0 1 1 0 1 0 1 0 0 0 1 0 0 1 0 1 1 1 1 0 0 1 0 1 1 0 0 0 1 0 0 0 1
1 0 0 1 1 0 0 0 1 1 1 0 0 0 0 1 1 0 0 0 0 0 1 1 1 0 1 1 0 0 0 0 1 0 1 0
1 1 0 0 1 0 0 1 1 1 0 0 1 1 1 0 1 0 1 0 1 1 1 1 1 1 1 1 0 1 1 0 1 1 0 0
1 1 1 1 0 0 0 1 1 0 1 0 1 1 1 0 0 1 0 0 0 0 1 1 1 1 0 1 1 1 0 1 1 1 1 0
1 0 0 0 0 0 1 0 0 0 0 0 0 1 0 1 1 0 1 1 1 1 1 0 0 1 1 0 1 1 1 0 0 0 1 0
1 1 1
1 0 0 0 1 0 0 1 0 1 1 1 1 0 0 1 0 1 1 0 0 0 1 0 0 0 1 1 0 0 1 1 0 0 0 1
1 1 0 0 0 0 1 1 0 0 0 0 0 1 1 1 0 1 1 0 0 0 0 1 0 1 0 1 1 0 0 1 0 0 1 1
1 0 0 1 1 1 0 1 0 1 0 1 1 1 1 1 1 1 1 0 1 1 0 1 1 0 0 1 1 1 1 0 0 0 1 1
0 1 0 1 1 1 0 0 1 0 0 0 0 1 1 1 1 0 1 1 1 0 1 1 1 1 0 1 0 0 0 0 0 1 0 0
0 0 0 0 1 0 1 1 0 1 1 1 1 1 0 0 1 1 0 1 1 1 0 0 0 1 0 1 1 1 0 1 0 0 1 1
0 0 1 0 1 0 1 0 1 0 0 1 0 0 1 0 0 0 1 0 1 0 0 0 0 1 0 0 1 1 0 1 0 0 0 1
1 1 1 1 0 1 0 1 1 0 1 0 0 1 0 1 0 0 1 1 1 1 1 1 0 0 0 0 0 0 0 1 1 0 1 1
0 1 0
0 0 1 1 1 1 1 0 1 0 1 1 0 1 0 0 1 0 1 0 0 1 1 1 1 1 1 0 0 0 0 0 0 0 1 1
0 1 1 0 1 0 1 0 0 0 1 0 0 1 0 1 1 1 1 0 0 1 0 1 1 0 0 0 1 0 0 0 1 1 0 0
1 1 0 0 0 1 1 1 0 0 0 0 1 1 0 0 0 0 0 1 1 1 0 1 1 0 0 0 0 1 0 1 0 1 1 0
0 1 0 0 1 1 1 0 0 1 1 1 0 1 0 1 0 1 1 1 1 1 1 1 1 0 1 1 0 1 1 0 0 1 1 1
1 0 0 0 1 1 0 1 0 1 1 1 0 0 1 0 0 0 0 1 1 1 1 0 1 1 1 0 1 1 1 1 0 1 0 0
0 0 0 1 0 0 0 0 0 0 1 0 1 1 0 1 1 1 1 1 0 0 1 1 0 1 1 1 0 0 0 1 0 1 1 1
0 1 0 0 1 1 0 0 1 0 1 0 1 0 1 0 0 1 0 0 1 0 0 0 1 0 1 0 0 0 0 1 0 0 1 1
0 1 0

Table 10: HC-255-2.

0 1 0 1 1 1 0 1 0 0 1 1 0 0 1 0 1 0 1 0 1 0 0 1 0 0 1 0 0 0 1 0 1 0 0 0
0 1 0 0 1 1 0 1 0 0 0 1 1 1 1 1 0 1 0 1 1 0 1 0 0 1 0 1 0 0 1 1 1 1 1 1
0 0 0 0 0 0 0 1 1 0 1 1 0 1 0 1 0 0 0 1 0 0 1 0 1 1 1 1 0 0 1 0 1 1 0 0
0 1 0 0 0 1 1 0 0 1 1 0 0 0 1 1 1 0 0 0 0 1 1 0 0 0 0 0 1 1 1 0 1 1 0 0
0 0 1 0 1 0 1 1 0 0 1 0 0 1 1 1 0 0 1 1 1 0 1 0 1 0 1 1 1 1 1 1 1 1 0 1
1 0 1 1 0 0 1 1 1 1 0 0 0 1 1 0 1 0 1 1 1 0 0 1 0 0 0 0 1 1 1 1 0 1 1 1
0 1 1 1 1 0 1 0 0 0 0 0 1 0 0 0 0 0 0 1 0 1 1 0 1 1 1 1 1 0 0 1 1 0 1 1
1 0 0
1 1 1 1 1 0 1 1 0 1 1 0 0 1 1 1 1 0 0 0 1 1 0 1 0 1 1 1 0 0 1 0 0 0 0 1
1 1 1 0 1 1 1 0 1 1 1 1 0 1 0 0 0 0 0 1 0 0 0 0 0 0 1 0 1 1 0 1 1 1 1 1
0 0 1 1 0 1 1 1 0 0 0 1 0 1 1 1 0 1 0 0 1 1 0 0 1 0 1 0 1 0 1 0 0 1 0 0
1 0 0 0 1 0 1 0 0 0 0 1 0 0 1 1 0 1 0 0 0 1 1 1 1 1 0 1 0 1 1 0 1 0 0 1
0 1 0 0 1 1 1 1 1 1 0 0 0 0 0 0 1 1 0 1 1 0 1 0 1 0 0 0 1 0 0 1 0 1 1 1
1 1 0 0 1 0 1 1 0 0 0 1 0 0 0 1 1 0 0 1 1 0 0 0 1 1 1 0 0 0 0 1 1 0 0 0
0 0 1 1 1 0 1 1 0 0 0 0 1 0 1 0 1 1 0 0 1 0 0 1 1 1 0 0 1 1 1 0 1 0 1 0
1 1 1
1 1 1 0 1 0 1 0 1 1 1 1 1 1 1 1 0 1 1 0 1 1 0 0 1 1 1 1 0 0 0 1 1 0 1 0
1 1 1 0 0 1 0 0 0 0 1 1 1 1 0 1 1 1 0 1 1 1 1 0 1 0 0 0 0 0 1 0 0 0 0 0
0 1 0 1 1 0 1 1 1 1 1 0 0 1 1 0 1 1 1 0 0 0 1 0 1 1 1 0 1 0 0 1 1 0 0 1
0 1 0 1 0 1 0 0 1 0 0 1 0 0 0 1 0 1 0 0 0 0 1 0 0 1 1 0 1 0 0 0 1 1 1 1
1 0 1 0 1 1 0 1 0 0 1 0 1 0 0 1 1 1 1 1 1 0 0 0 0 0 0 0 1 1 0 1 1 0 1 0
1 0 0 0 1 0 0 1 0 1 1 1 1 0 0 1 0 1 1 0 0 0 1 0 0 0 1 1 0 0 1 1 0 0 0 1
1 1 0 0 0 0 1 1 0 0 0 0 0 1 1 1 0 1 1 0 0 0 0 1 0 1 0 1 1 0 0 1 0 0 1 1
1 0 0
0 1 1 1 0 0 0 0 1 1 0 0 0 0 0 1 1 1 0 1 1 0 0 0 0 1 0 1 0 1 1 0 0 1 0 0
1 1 1 0 0 1 1 1 0 1 0 1 0 1 1 1 1 1 1 1 1 0 1 1 0 1 1 0 0 1 1 1 1 0 0 0
1 1 0 1 0 1 1 1 0 0 1 0 0 0 0 1 1 1 1 0 1 1 1 0 1 1 1 1 0 1 0 0 0 0 0 1
0 0 0 0 0 0 1 0 1 1 0 1 1 1 1 1 0 0 1 1 0 1 1 1 0 0 0 1 0 1 1 1 0 1 0 0
1 1 0 0 1 0 1 0 1 0 1 0 0 1 0 0 1 0 0 0 1 0 1 0 0 0 0 1 0 0 1 1 0 1 0 0
0 1 1 1 1 1 0 1 0 1 1 0 1 0 0 1 0 1 0 0 1 1 1 1 1 1 0 0 0 0 0 0 0 1 1 0
1 1 0 1 0 1 0 0 0 1 0 0 1 0 1 1 1 1 0 0 1 0 1 1 0 0 0 1 0 0 0 1 1 0 0 1
1 0 0
0 1 1 0 1 0 1 0 0 0 1 0 0 1 0 1 1 1 1 0 0 1 0 1 1 0 0 0 1 0 0 0 1 1 0 0
1 1 0 0 0 1 1 1 0 0 0 0 1 1 0 0 0 0 0 1 1 1 0 1 1 0 0 0 0 1 0 1 0 1 1 0
0 1 0 0 1 1 1 0 0 1 1 1 0 1 0 1 0 1 1 1 1 1 1 1 1 0 1 1 0 1 1 0 0 1 1 1
1 0 0 0 1 1 0 1 0 1 1 1 0 0 1 0 0 0 0 1 1 1 1 0 1 1 1 0 1 1 1 1 0 1 0 0
0 0 0 1 0 0 0 0 0 0 1 0 1 1 0 1 1 1 1 1 0 0 1 1 0 1 1 1 0 0 0 1 0 1 1 1
0 1 0 0 1 1 0 0 1 0 1 0 1 0 1 0 0 1 0 0 1 0 0 0 1 0 1 0 0 0 0 1 0 0 1 1
0 1 0 0 0 1 1 1 1 1 0 1 0 1 1 0 1 0 0 1 0 1 0 0 1 1 1 1 1 1 0 0 0 0 0 0
0 1 1

References

Coates, A., and Ng, A. Y. 2011. Selecting receptive fields in deep networks. In *NIPS*.

Krizhevsky, A.; Sutskever, I.; and Hinton, G. E. 2012. ImageNet classification with deep convolutional neural networks. In *NIPS*.

Larochelle, H.; Erhan, D.; Courville, A.; Bergstra, J.; and Bengio, Y. 2007. An empirical evaluation of deep architectures on problems with many factors of variation. In *ICML*, 473–480.

Srivastava, N., and Salakhutdinov, R. 2013. Discriminative transfer learning with tree-based priors. In *NIPS*.

Zeiler, M., and Fergus, R. 2014. Visualizing and understanding convolutional neural networks. In *ECCV*.

Table 1: The details of model structure of STL-10, MNIST.
FN = Filter Number, FS = Filter Size, AF = Activation Function

Dataset	Conv 1			Conv 2			Conv 3			Local		
	FN	FS	AF	FN	FS	AF	FN	FS	AF	FN	FS	AF
STL-10	64	5	RELU	64	5	RELU	256	5	RELU	256	5	RELU
MNIST	64	5	RELU	64	5	RELU	64	3	RELU	64	3	RELU

Table 2: The details of learning parameters of STL-10, MNIST.
IW = Initial Weights, LR = Learning Rate

Dataset	Conv 1		Conv 2		Conv 3		Local	
	IW	LR	IW	LR	IW	LR	IW	LR
STL-10	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
MNIST	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001

Table 3: The details of model structure of CIFAR-100
FN = Filter Number, FS = Filter Size, AF = Activation Function
NN = Neuron Number, FC = Fully Connected

Dataset	Conv 1			Conv 2			Conv 3			Local			FC 1		FC 2	
	FN	FS	AF	FN	FS	AF	FN	FS	AF	FN	FS	AF	NN	AF	NN	AF
CIFAR-100	64	5	RELU	96	5	RELU	256	3	RELU	256	3	RELU	2048	RELU	2048	RELU

Table 4: The details of learning parameters of CIFAR-100.
IW = Initial Weights, LR = Learning Rate

Dataset	Conv 1		Conv 2		Conv 3		Local		FC 1		FC 2	
	IW	LR	IW	LR	IW	LR	IW	LR	IW	LR	IW	LR
CIFAR-100	0.01	0.001	0.01	0.001	0.01	0.001	0.01	0.001	0.01	0.001	0.01	0.001

Table 5: The details of model structure of ImageNet-2012
FN = Filter Number, FS = Filter Size, AF = Activation Function
NN = Neuron Number, FC = Fully Connected

Dataset	Conv 1			Conv 2			Conv 3			Conv 4			Conv 5			FC 1		FC 2	
	FN	FS	AF	FN	FS	AF	FN	FS	AF	FN	FS	AF	FN	FS	AF	NN	AF	NN	AF
AlexNet	96	11	RELU	256	5	RELU	384	3	RELU	384	3	RELU	384	3	RELU	4096	RELU	4096	RELU
Clarifai	96	7	RELU	256	5	RELU	384	3	RELU	384	3	RELU	384	3	RELU	4096	RELU	4096	RELU

Table 6: The details of learning parameters of ImageNet-2012.
IW = Initial Weights, LR = Learning Rate

Dataset	Conv 1		Conv 2		Conv 3		Conv 4		Conv 5		FC 1		FC 2	
	IW	LR	IW	LR	IW	LR	IW	LR	IW	LR	IW	LR	IW	LR
AlexNet	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Clarifai	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01