

KECE470: Pattern Recognition
School of Electrical Engineering, KOREA UNIVERSITY
(Homework #4) Deep Neural Networks
Report containing the code, results, discussions

In this homework, we will use MNIST dataset which can be downloaded from <http://yann.lecun.com/exdb/mnist/>. You will write a program to construct and train a Convolutional Neural Network, then use it to predict class label of the test data. **Please submit a report file and codes, respectively.**

Please answer the following questions.

1. **(Download MNIST dataset)** The data has been divided into several sets for training and test. You will randomly take 10% of the training set as validation set.
2. **(Build the CNN Model)** Next, you will build 2 kinds of Convolutional Neural Network
A Network: Conv-Pool-Conv-Pool
B Network: Conv-Conv-Pool-Conv-Conv-Pool,
the architectures are as follows and you will take ReLU as the activation function in these two networks,

A network		
Layer	Number of Filters	Padding
Input Image	-	-
Conv2d (f=3, s=1)	32	Valid
MaxPool(f=2, s=2)	-	Valid
Conv2d (f=3, s=1)	64	Valid
MaxPool (f=2, s=2)	-	Valid
Conv2d (f=3, s=1)	128	Valid
MaxPool (f=2, s=2)	-	Valid
Flatten	-	-
Dense	-	-
Softmax	-	-

B network		
Layer	Number of Filters	Padding
Input Image	-	-
Conv2d (f=3, s=1)	16	same
Conv2d (f=3, s=1)	16	same
MaxPool(f=2, s=2)	-	Valid
Conv2d (f=3, s=1)	32	Valid
Conv2d (f=3, s=1)	32	Valid
MaxPool (f=2, s=2)	-	Valid
Flatten	-	-
Dense	-	-
Softmax	-	-

3. **(Training and Evaluation)**
 - a) Describe which network is better? (Analyzing from the accuracy on the validation set)
 - b) For the network you choose, plot the loss curves on training set and validation set, observe and explain their trend.
 - c) Evaluate the trained network which you choose on the test set, print the accuracy.
 - d) Explain how to prevent overfitting problem when training the deep network.