HT2020: IL2230 HADL Lab 3C
Handwritten Digits Recognition from MLP to CNN

Zhonghai Lu

October 25, 2020

1 Introduction

Hand-written digits recognition has been a long standing problem since the beginning of machine learning and pattern recognition. There are certainly many different approaches. In the lab, we try to use neural networks to address this problem. In particular, we use a CNN called LeNet-5 [1] to do this task. LeNet-5 is perhaps the first successful application of convolutional neural networks (CNNs) in real-life problems. CNNs can automatically extract features in a way much more efficient and effective than manually-made features. After feature extraction by convolutional layers, the neural network can perform classification using fully connected layers.

2 Project objectives

The purpose of the project is to use and contrast different neural networks to solve an image recognition problem. It has a few intended learning outcomes after completing the project:

- be able to design and train MLPs.
- be able to design and train CNNs, which is a special kind of MLPs.
- be able to compare the performance of different neural networks, in particular, understand the benefits of convolutional layer in automatic feature extraction.

The project shall be carried out using PyTorch. It is suitable as a small-size team of 3-4 students (typically 4 students) in one group.

3 Project tasks

The project intends to solve an image recognition problem, specifically, handwritten digits. The data set is the well-known MNIST, which consists of only 10 classes from 0, 1, 2, ... to 9.

The MNIST database of handwritten digits has a training set of 60,000 examples, and a test set of 10,000 examples. The digits have been size-normalized and centered in a fixed-size image. It is a good database for testing learning techniques and pattern recognition methods on real-world data while spending minimal efforts on data pre-processing and formatting.

Your tasks are as follows:

1. Design and train an MLP.

   You can consider two variants: (1) MLP1: MLP with 1 hidden layer. (2) MLP2: MLP with 2 hidden layers.

   For MLP1, you can try different number of neurons for the hidden layer, say from 30 to 300 neurons with a step size of 30. Certainly you can try own options.

   For MLP2, you can try different numbers of neurons for the two hidden layers. One possible choice is 300 neurons for the first hidden layer and 100 neurons for the second hidden layer.

[1] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. "Gradient-based learning applied to document recognition." Proceedings of the IEEE, 86(11):2278-2324, November 1998.

2. Write and train a neural network implementing the LeNet-5 model.

   LeNet-5 is perhaps the first CNN which successfully deals with hand-written digits recognition problem.

   Note that the images were centered in a 28x28 image. The standard LeNet-5 assumes the image size 32x32.

   After adaption, record the classification accuracy of your adapted CNN.

3. Compare the inference performance, namely, classification accuracy, run-time, memory consumption, of the two approaches.

   Hint: Use PyTorch Profiler to characterize execution time and memory consumption.

3.1 Documentation

There are two deliverables as documentation:

1. Write a technical report (names of team members on the first page), describing your procedures for implementing the tasks, reporting results you obtained, and discussing your results.

2. Upload your source code and the report to the Canvas course page. The source code should accompany a simple Readme.txt for how to run the code and obtain the graphs/numbers in your report.

   After your lab results have been approved by the lab assistant, submit your report via the Canvas course website:

   https://kth.instructure.com/courses/20640

3.2 Aids

Links to online materials.

- PyTorch recipes and tutorials.

  Tutorials: https://pytorch.org/tutorials/

  Recipes: https://pytorch.org/tutorials/recipes/recipes_index.html

  PyTorch profiler: https://pytorch.org/tutorials/recipes/recipes/profiler.html

- The MNIST dataset.

  http://yann.lecun.com/exdb/mnist/

  This website gives detailed description of the MNIST dataset. Read this as the first step of understanding the dataset and its data structure.