Hands-on Exercise 3: Training Convolutional Neural Networks on Images

In this hands-on exercise you experiment with a browser-based implementation of a convolutional neural networks. The goal of the exercise is to acquire an understanding of the inner workings of convolutional neural networks.

- The CNN has the following configuration:

Input	32 x 32 x 3
Conv	32 x 32 x 16 filter size 5 x 5 x 3 (stride 1)
Relu	32 x 32 x 16
Pool	16 x 16 x 16 pooling size 2 x 2 (stride 2)
Conv	16 x 16 x 20 filter size 5 x 5 x 16 (stride 1)
Relu	16 x 16 x 20
Pool	8 x 8 x 20
Conv	8 x 8 x 20
Relu	8 x 8 x 20
Pool	4 x 4 x 20 pooling size 2 x 2 (stride 2)
fc	1 x 1 x 10
softmax	1 x 1 x 10

1. Train the network and examine each layer of the network by looking at the images and reading the numbers. The gradients (for the activations and weights) visualise the changes due to learning. Try to understand the transformations between each pair of layers. If training takes too long, you may want to load a pre-trained network which classifies about 80% correct.

2. The interface allows you to vary the learning rate, momentum, batch size and weight decay. While training from scratch, vary each of the parameters to see the effect on training. What is the purpose of each parameter?

Although the CNN illustrates the training of a real deep learning network, its complexity may hamper a full understanding of what happens. A simpler network and task (MNIST) can be found here.

3. Examine the MNIST CNN configuration and compare it to that of the CNN trained on the CIFAR data. What are the main differences and what motivates the different network structure?
4. Questions 1 and 2 for the MNIST CNN.