Assignment 2: Convolutional Networks

Due date: 01.04.2021 - 23:59.
Total: 2 points.

For this homework, you will design and train a convolutional network capable of counting the digits in an image. You will get a dataset of 5000 training images and 1000 testing images, each of size 100×100. These images are formed by placing N (1 to 5) digits from the MNIST dataset (28×28 digits) onto a noisy background. The goal is to predict the number of digits in the image.

The dataset is found in the course Drive folder in mnist_count_test.pickle and mnist_count_train.pickle. This assignment is very similar to the tasks solved in the third laboratory about convolutions, and you could use the code given in the lab as a starting point for the homework.

You must implement two methods for this task.

1 Method 1: Without training on the counting dataset.

Just use a model trained to classify images on MNIST dataset. Use this model in a sliding window approach, by splitting the large image into smaller 28×28 patches or convert the model into a fully convolutional network to predict maps representing the probability of digits being present at that location. Without additional training, use these maps to predict the number of digits in the image.

You can use the training set of the dataset to choose different hyper-parameters of the method (such as different thresholds) by observing the effect on the accuracy. You must compute and report the accuracy on the validation set.

2 Method 2: Train a new convolutional network to predict the number of digits.

Design a new convolutional network to directly predict the number of digits. Your task is to investigate the best network architecture and successfully train a model.

You must investigate the following:

- Follow the training and testing loss and accuracy after each epoch. Plot the loss / accuracy for train and test set in the same figure.
- Investigate the effect of pretraining on the accuracy, depending on the number of samples used in training. Train two models, where one has some layers pre-trained on MNIST classification and the other one has all the layers randomly initialized. Compare their accuracy when using $20\%, 50\%, 100\%$ of the training data. Plot the final accuracy of all models in the same figure, where the Ox axis represents the fraction of data used and the Oy axis represents the final accuracy.
- investigate other aspects of a network architecture such as loss function, number of layers, regularisation.