Pattern Recognition - Exercise 2c (CNN)

Implementation details

- We didn't use a constant learning rate. Instead, the learning rate was decreasing with 10% from one epoch to another.

- We used cross validation with 5 number of folds to determine the pair (number of epochs, learning rate) for which we were able to train the model with the highest accuracy (on the validation set that we extracted from the initial train set). Then, with these 2 values found, we trained the model once again (on the whole train dataset this time) and evaluated it using the test dataset.

- For computing the loss we used cross-entropy.

- For the optimizer, we chose Stochastic Gradient Descent.

Testing details

- Number of epochs: 8, 10, 15, 20

- Initial learning rate: 0.01, 0.05, 0.08, 0.1, 0.12, 0.13 (we tried larger values than 0.1 too because it's decreasing from one epoch to another)

- Batch size: 32

Results

- Number of epochs: 15

- Initial learning rate: 0.13

- Accuracy: 98.60%

	Accuracy		Loss	
Epoch	Train Dataset	Test Dataset	Train Dataset	Test Dataset
1	95.73%	96.37%	0.1995	0.1859
2	96.54%	96.74%	0.1388	0.1306
3	97.37%	97.30%	0.1137	0.1090
4	97.88%	97.73%	0.0990	0.0989
5	98.12%	97.94%	0.1037	0.1055
6	98.30%	98.08%	0.0963	0.0960
7	98.18%	97.76%	0.0914	0.0966
8	98.69%	98.34%	0.0819	0.0859
9	98.68%	98.33%	0.0642	0.0700
10	98.88%	98.55%	0.0642	0.0701
11	98.94%	98.44%	0.0578	0.0668
12	99.00%	98.40%	0.0574	0.0665
13	99.08%	98.50%	0.0558	0.0651
14	99.01%	98.30%	0.0499	0.0612
15	99.18%	98.60%	0.0473	0.0586