

Finn Fassbender
Mathis Welker

1

 Σ

Assignment Sheet Nr. 5

Question 1 Boat-MNIST: Compete on an image classification task on SeaDronesSee

For image classification purposes a convolutional neural network (CNN) was chosen, similar to the LeNet-5 architecture for digit recognition on the MNIST database of handwritten digits. CNNs outperformed all other models on this task (LeCun et al., 1998).

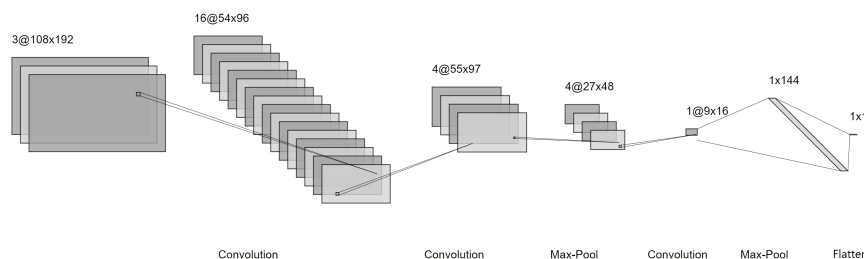


Figure 1: Visualization of the network layers.

A standard Stochastic Gradient Descent optimizer was chosen for training, using learning rate $\alpha = 0.1$ with batch size 64. Data augmentation was not performed as data exploration showed the ratio of boat/no-boat pictures to be approximately 1 to 3 (1900 out of 5271; no differences in magnitude, roughly equally distributed between training and validation set). To prevent overfitting of the model, aggressive dropout regularization was implemented following both MaxPool steps (50% and 25% respectively). The resulting neural net would be thus be trained to be as resilient as possible, whilst minimizing overutilization of specific neurons. This meant that the model could be safely trained for up to approx. 2500 epochs without needing to worry about significant overfitting.

Final validation loss was 99.93% (1505/1506 images correctly classified).

The summary of the model during training is as follows (using `torchinfo.summary()`):

```

1 model size: 2426 Parameters      0.01MB
2 =====
3 Layer (type:depth-idx)           Output Shape           Param #
4 =====
5 Net                               [64, 1]                --
6 |-Conv2d: 1-1                     [64, 16, 54, 96]       1,216
7 |-Conv2d: 1-2                     [64, 4, 55, 97]        1,028
8 |-Dropout: 1-3                    [64, 4, 27, 48]        --
9 |-Conv2d: 1-4                     [64, 1, 27, 48]        37
10 |-Dropout: 1-5                   [64, 1, 9, 16]         --
11 --Linear: 1-6                    [64, 1]                145
12 =====
13 Total params: 2,426
14 Trainable params: 2,426
15 Non-trainable params: 0
16 Total mult-adds (M): 757.52
17 =====
18 Input size (MB): 15.93
19 Forward/backward pass size (MB): 54.06
20 Params size (MB): 0.01
21 Estimated Total Size (MB): 69.99
22 =====

```