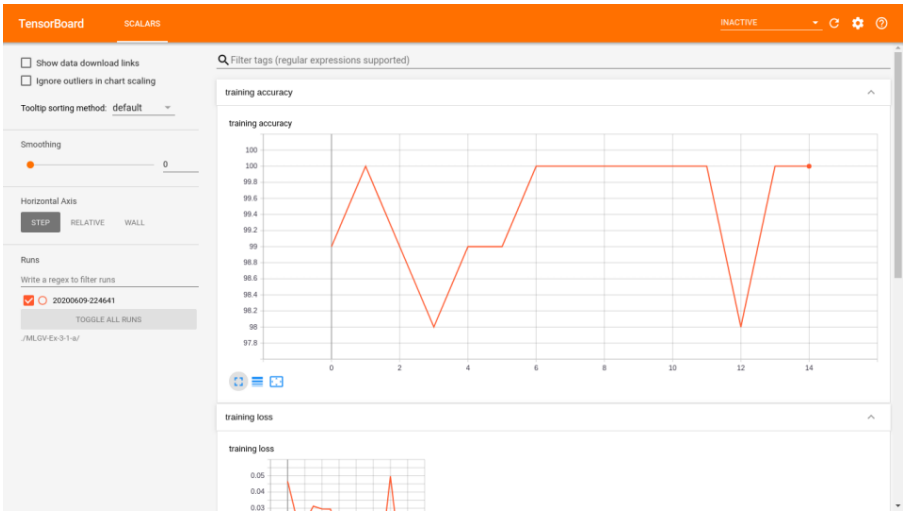
Exercise 3

Machine Learning in Graphics & Vision

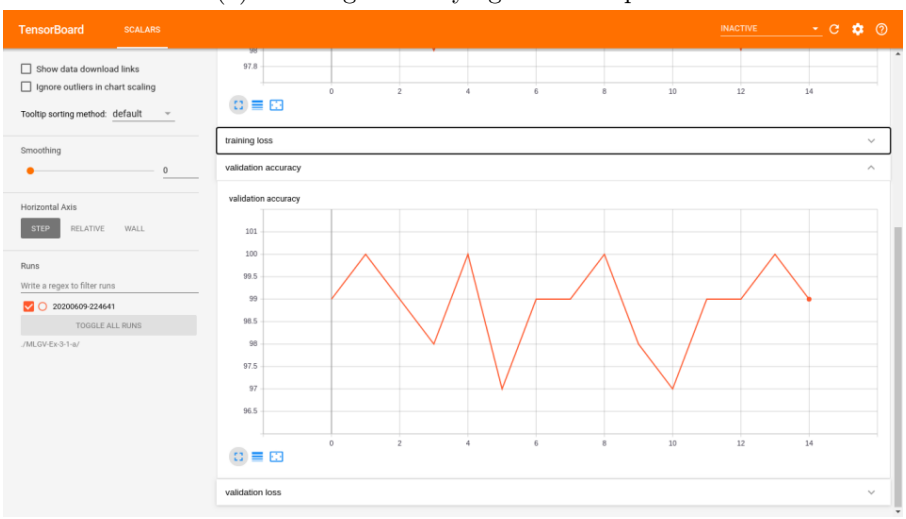
Driton Goxhufi 4233242 driton.goxhufi@student.uni-tuebingen.de
Damir Ravlija 5503184 damir.ravlija@student.uni-tuebingen.de

1 Task 1

(a) After 15 epochs our model achieves 98.92 validation accuracy 1. State-of-the-art accuracy on classifying MNIST digits reported in the literature is 99.84 [1].



(a) Training accuracy against the epochs



(b) Validation accuracy against the epochs

Figure 1: Plot of results from task 3.1.(a)

(b) Model summary is shown in Figure 2. Output shape of `torch.nn.conv2d` with zero padding for an input of shape $[B, T, T, C]$ and stride 2 kernel-size 3 can be computed in the following way

[1] https://paperswithcode.com/sota/image-classification-on-mnist

Driton Goxhufi 4233242
Damir Ravlija 5503184

(under the assumption that B is batch size, first T is the number of channels, second T is the input height and C is the input width):

- Batch size B stays the same.
- As the kernel-size is larger than 2 and there are zero padding, both height and width have to decrease and since the stride 2 is also used, the resulting height is ceil($\frac{T}{2}$) and the resulting width is floor($\frac{C}{2}$).
- The resulting output shape: $[B, T, \text{ceil}(\frac{T}{2}), \text{ceil}(\frac{T}{2})]$ if we assume that the number of output channels is the same as the number of input channels T.

```
----------------------------------------------------------------
        Layer (type)               Output Shape         Param #
================================================================
            Conv2d-1          [-1, 32, 28, 28]             320
              ReLU-2          [-1, 32, 28, 28]               0
            Conv2d-3          [-1, 32, 28, 28]           9,248
              ReLU-4          [-1, 32, 28, 28]               0
            Conv2d-5          [-1, 32, 28, 28]           9,248
              ReLU-6          [-1, 32, 28, 28]               0
            Linear-7                [-1, 256]       6,422,784
              ReLU-8                [-1, 256]               0
            Linear-9                 [-1, 10]           2,570
================================================================
Total params: 6,444,170
Trainable params: 6,444,170
Non-trainable params: 0
----------------------------------------------------------------
Input size (MB): 0.00
Forward/backward pass size (MB): 1.15
Params size (MB): 24.58
Estimated Total Size (MB): 25.74
```

Figure 2: Model summary obtained using `torchsummary` package task 3.1.(b)

(c) After training 15 epochs we get 92.07 validation accuracy.