Exercise 4
Machine Learning in Graphics & Vision

Driton Goxhufi   4233242   driton.goxhufi@student.uni-tuebingen.de

Damir Ravlija   5503184   damir.ravlija@student.uni-tuebingen.de

Task 1

(c) It seems that using the window size 3 leads to better visual results because although window size 3 produces a lot of artifacts and the image is noisy, it has more details than the images produced by larger window sizes which suffer from loss of details as parts become too smooth.

(d) We think that block matching approach fails to lead to good estimations around homogeneous regions because since these regions have mostly the same color values, multiple window blocks in these regions have similar sums of absolute difference (SAD) values.

Task 2

(c) When the initial parameters with 64 filter numbers and 1000 iterations were used, loss did not decrease significantly, but a certain downward trend was visible (Figure 1). The best training loss and accuracy that we got after training 1000 iterations was:
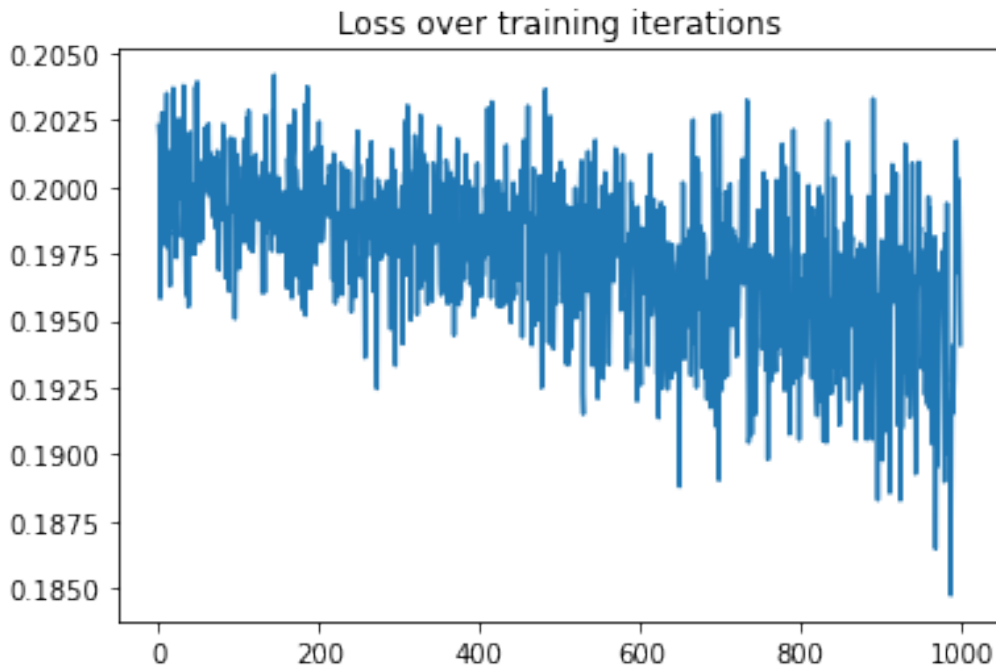
Loss:0.1965   Accuracy: 0.557

Figure 1: Task 2.c) Training loss over iterations when training with filter number 64 and 1000 iterations.

When we increased the number of iterations to 5000 the model continued improving (Figure 2) and the best loss and accuracy achieved was:
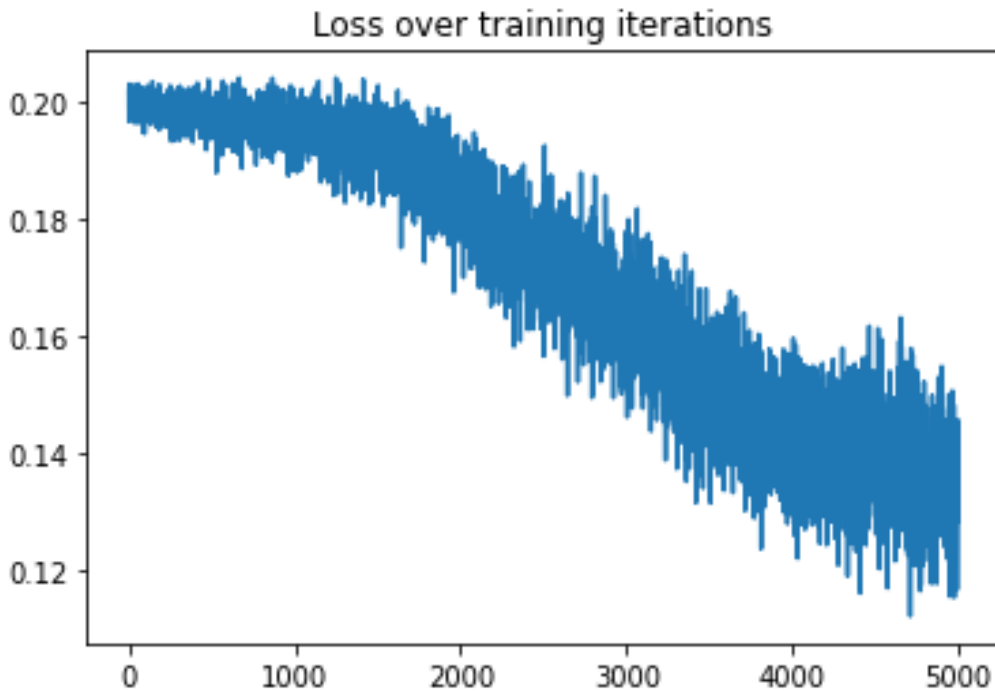
Loss:0.1238   Accuracy: 0.746

Figure 2: Task 2.c) Training loss over iterations when training with filter number 64 and 5000 iterations.

When the number of filters was decreased to 32, the loss did not significantly decrease with the increased number of iterations (Figure 3). On the other hand, increasing the filter number to 128 did not result in significantly improved results compared to using filter number 64.

From our findings it seems that the increased number of iterations is advantageous only with the increasing filter number since the smaller network with filter number of 32 could not significantly improve in 5000 iterations. Furthermore, since the networks with filter numbers 64 and 128 improve even after more than 5000 iterations, it seems that 1000 iterations is not enough to train the model, while the decreased loss and improved accuracy that are achieved with larger filter numbers become visible only if they are accompanied by increased number of iterations (Figure 3).

(d) Predictions by the Siamese Neural Network seem to be better than the ones obtained by the block matching algorithm because although the visualization is much smoother than block matching with window size 3, it doesn't suffer from as much detail loss as visualization obtained using block matching with larger window sizes. Difference in the predictions is the most dominant in the texture-less road region near the camera.

2

(a) Filter number 32, Iterations 1000

(b) Filter number 32, Iterations 5000

(c) Filter numbers 128, Iterations 1000

(d) Filter number 128, Iterations 5000

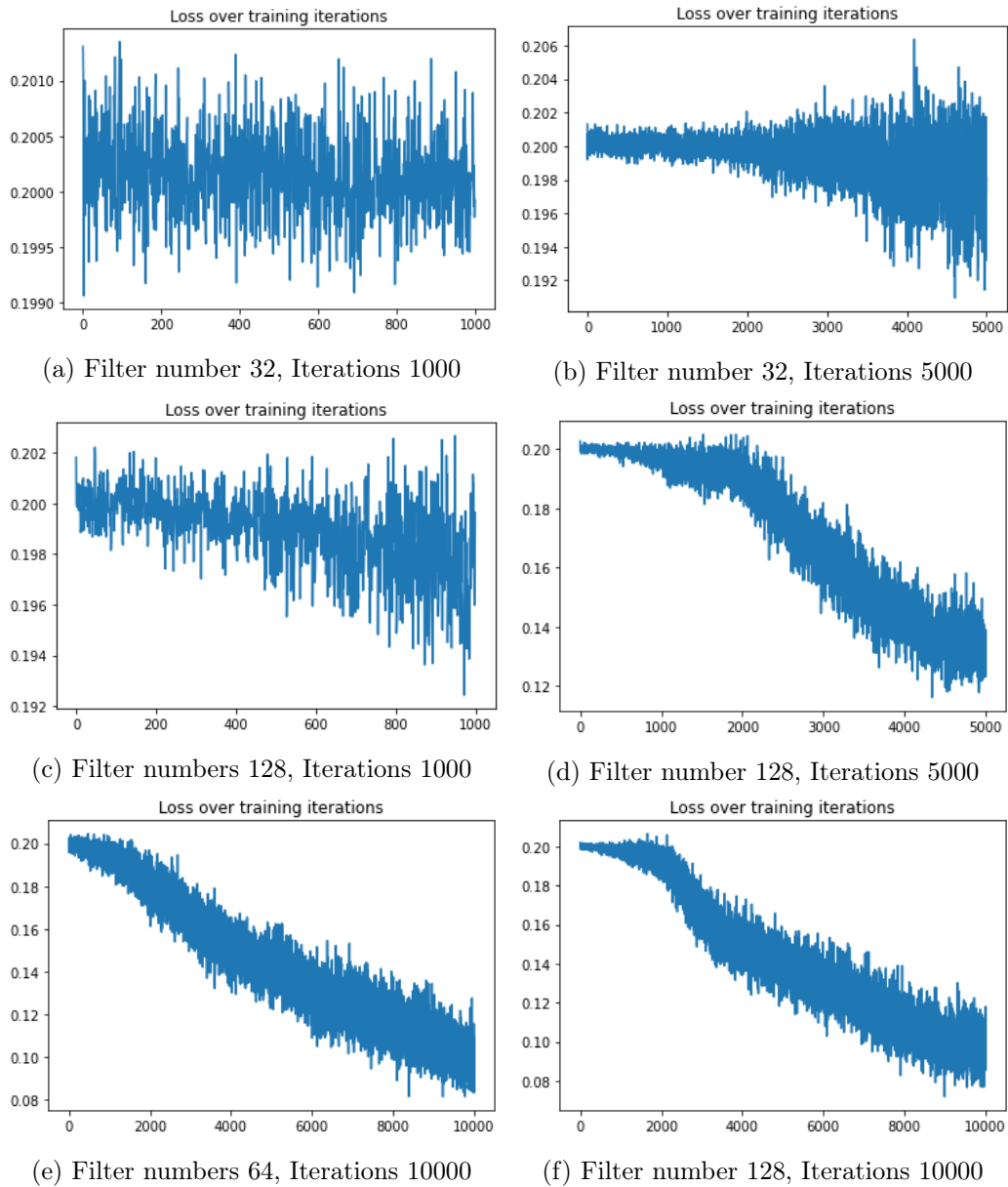(e) Filter numbers 64, Iterations 10000

(f) Filter number 128, Iterations 10000

Figure 3: Task 2.c) Training loss over iterations.