194.077 Applied Deep Learning - 2021WS

Assignment 3 – Final Report

Felix Reuß (1431542)

Despeckling C-band SAR data using a Deep Residual U-Net

Beat the classics

1. Introduction – Problem definition

Microwave remote sensing and in particular synthetic aperture radar (SAR) data is an important input for a wide range of applications. These include environmental monitoring, soil moisture retrieval, flood mapping, yield prediction, and many more. However, it comes with a downside that limits the interpretability of the data. The interaction of the emitted wave with the surface leads to destructive and constructive interference which results in an inherent noise in the measurement, the so-called speckle effect.

2. Deep Learning for SAR Image despeckling

Current state-of-the-art approaches to remove this noise are based on basic image processing algorithms which average pixel values in a sliding window to reduce the noise. Yet, this averaging come with the costs of a loss of information and a lower resolution. First approaches to use machine learning or in particular deep learning for this application were hindered by the lack of noise-free reference data. To overcome this, several studies propose to use optical images converted to grey-scale instead [1]. Due to differences in the acquisition geometry and the simplified noise model, these datasets often fail to represent the characteristics of real SAR acquisitions.

3. Goal of the project

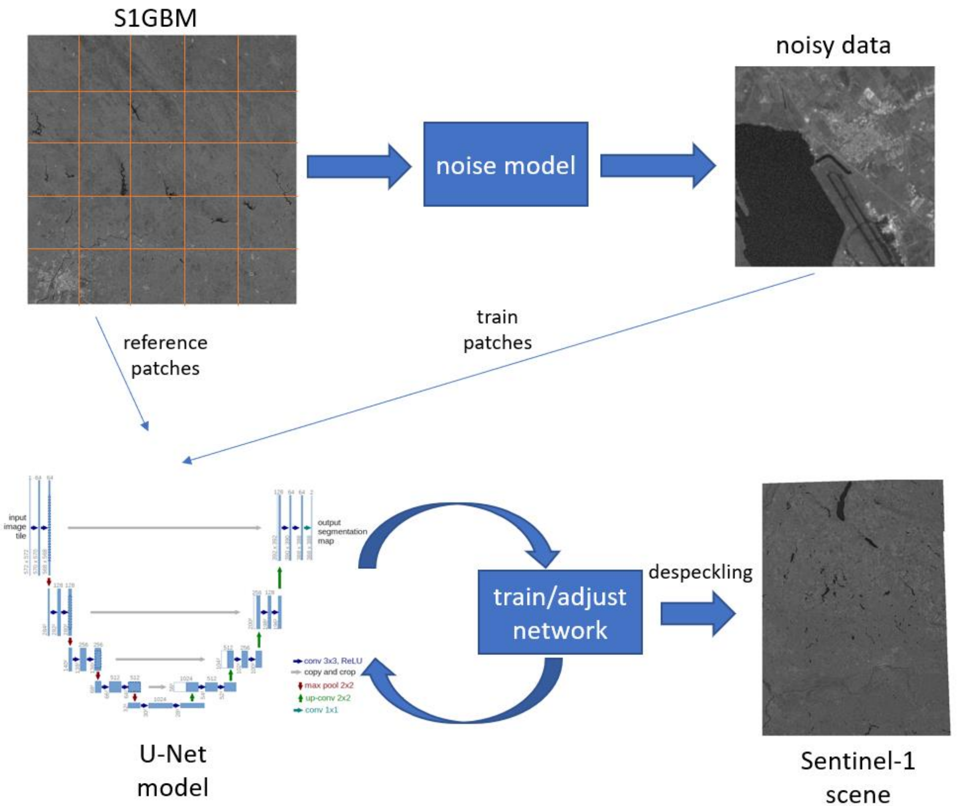

Figure 1: Illustration of the workflow, including the data preparation, model training, and the final application.

This project will overcome this obstacle by using the global Sentinel-1 backscatter model (S1GBM) as a (de facto) noise-free reference [2]. This data is built by averaging measurement of the Sentinel-1 C band SAR mission over two years on a global scale, leading to an extreme noise reduction. Figure 1 illustrates the workflow of the project. Artificial noise was added to the patches of the S1GBM to obtain a large-scale training and reference dataset. This dataset was then used to train and evaluate a U-Net for noise reduction. In a final step, the trained network, in the following called Despeckle Net, was applied on Sentinel-1 acquisitions to test the speckle removal on “real data”. Given the complexity of the task and the limited amount of time, the study is designed as a proof-of-concept.

4. Data

Figure 2: Illustration of a noise-free SAR image of Vienna on the left and an (original) noisy scene on the right.

The S1GBM is provided in the Equi7 Tiling Grid [3]. Eleven tiles over Europe covering different land cover types were chosen. For simplification, only VV polarized data was used. To generate pairs of noisy and noise-free scenes, the speckle model proposed by Singh, P. & Pandey R. S. [4] was applied to the S1GBM data. The here proposed formula was extended by a noise level factor, to introduce the noise level dependency of the landcover and can be written as follows:

$$(1)\ \ N = I + w \times \left(\frac{Std}{3200}\right) \times (I \times G)$$

Where N is the output noise array, I is the input image array, w is a noise weight (empirically set to 0.7). Std is the temporal pixel standard deviation over two years and G is the Gaussian noise array.

To feed this data to the Despeckle Net, a tiling workflow was created, splitting each tile into patches of 400x400 pixels. The patches of ten of these tiles were randomly split into 80% training and 20% validation data. The entire eleventh tile was used for testing to get a large-scale impression of the model output. After the successful training of the network, the model was tested on three Sentinel-1 acquisitions to test the transferability of the model from the artificial noise data to data with real speckle noise.

5. Methodology & Training

The U-Net architecture was originally proposed by Ronneberger et al. [5] as a semantic image segmentation network. In recent years, several studies also used the architecture for image denoising with the extension of residual blocks [6,7]. A recent study indicates the good performances of this architecture to remove Gaussian noise from various image types [8]. Here, Resize Convolutions were used in the upsampling part of the network as proposed by Aitken et al. [9]. Previous studies have shown that Deconvolutions tend to introduce checkerboard-like artifacts in regression applications. To avoid this, the Resizing Convolution blocks use a Bilinear or Bicubic Resizing Layer followed by one or several Convolutional Layers.

Within the training process, several (hyper-) parameters were tuned. Especially the number of filters, kernel size, learning rate, loss function, and activations functions had a significant impact on the results. In the end, a network with a kernel size of five, five downsampling, and five upsampling blocks, resulting in approximately 800,000 parameters was used. Adam optimizer with a learning rate of 0.0002 was used with the Leaky Relu function after the convolution blocks and a linear activation at the output layer.

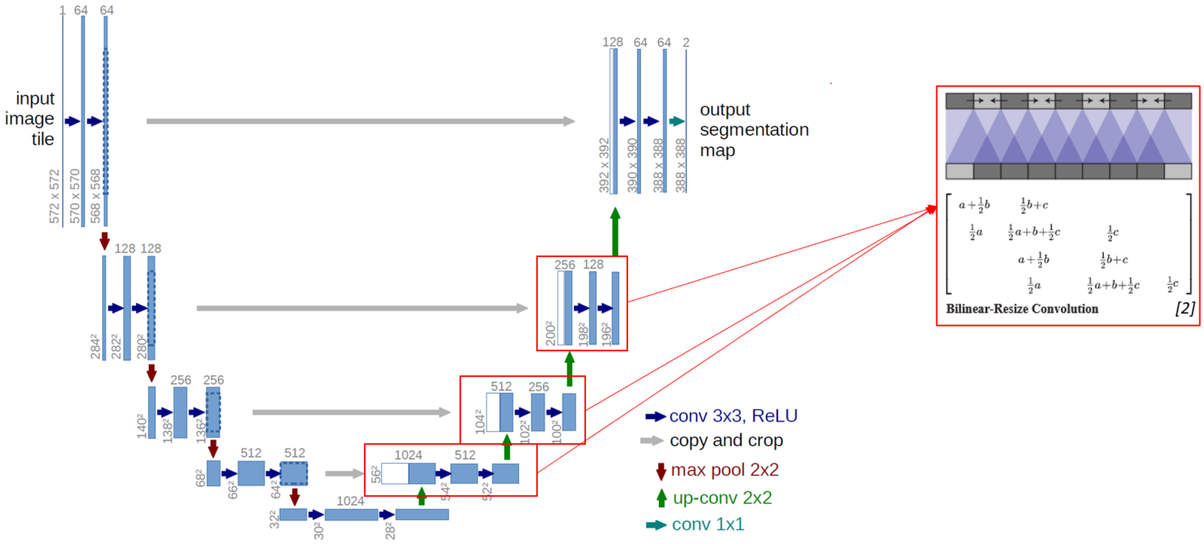

Figure 3: Illustration of the residual U-Net with resize convolutions

6. Results

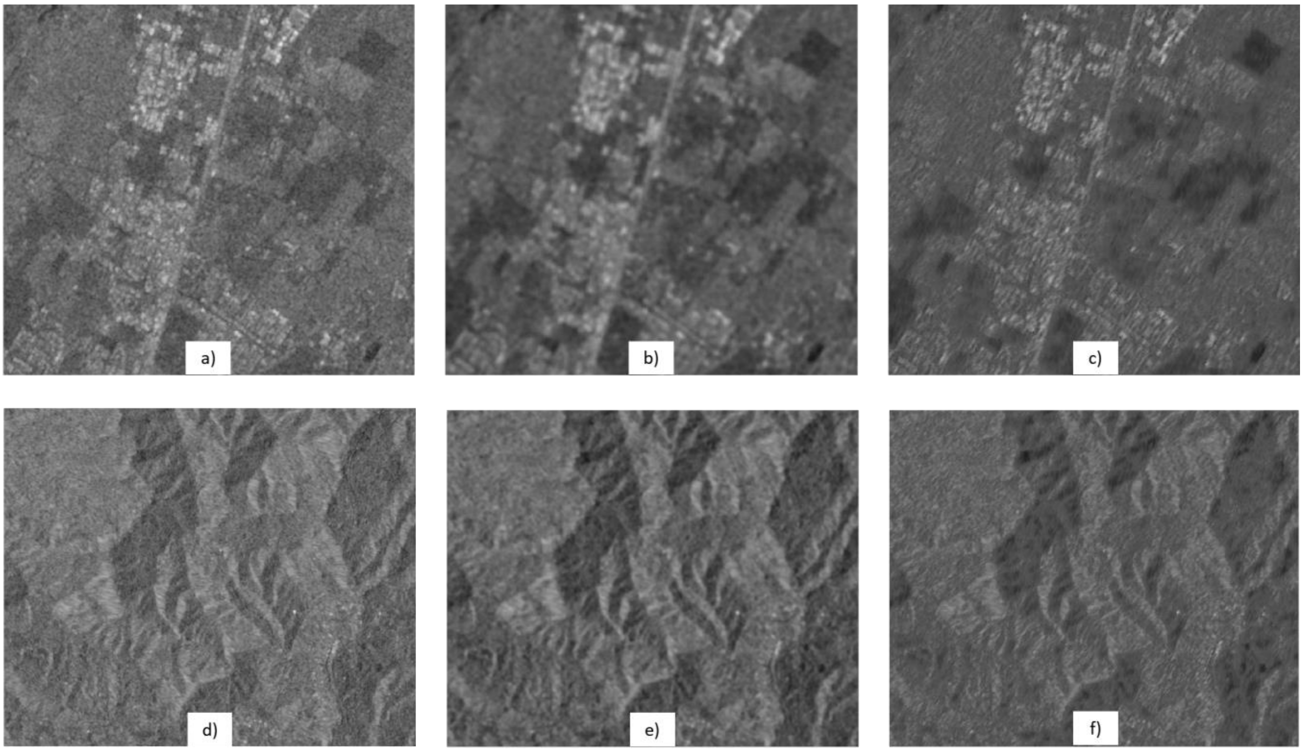

Figure 4: Original image (left) results from Lee filter (middle) and the Despeckel Net (right).

To validate the model and allow a conclusion if the previously defined goal was achieved, the mean squared error (MSE) was calculated on the validation data. One the same validation data, the Lee filter was applied for speckle removal. Subsequently, the MSE was calculated again. Table 1 shows the achieved results for both speckle models. For comparison, also the difference between the noisy data used as an input and the reference was calculated. As the table shows, the Lee Filter achieved a significantly better MSE. However, quantitative error metrics are inadequate to properly measure some aspects like edge-preserving or over-smoothing. Therefore, in addition, a visual inspection of both models was carried out on the real SAR data. Figure 4 shows a comparison of the obtained results.

*Table 1 shows the achieved accuracies for the Lee Filter and the Despeckle Net compared to the original train data. * Data was scaled by the factor 100.*

	Noisy data (baseline)	**Lee Filter**	**Despeckle Net**
MSE	**102***	**21***	**46***

7. Application

Speckle removal is typically done as an intermediate step in the processing chain that prepares ground range detected data to be analyse ready for the user. For the pre-processing chain of the Sentinel-1 datacube hosted at EODC [10] this workflow is implemented in Python. To allow the integration of the here developed model in the workflow, a Python function was written, that can take an image or array of any size and resolution as an input. Using a user-defined pixel overlap, it applies the despeckling model on it and afterward georeferences it and writes it as GeoTIFF to disk.

8. Insights

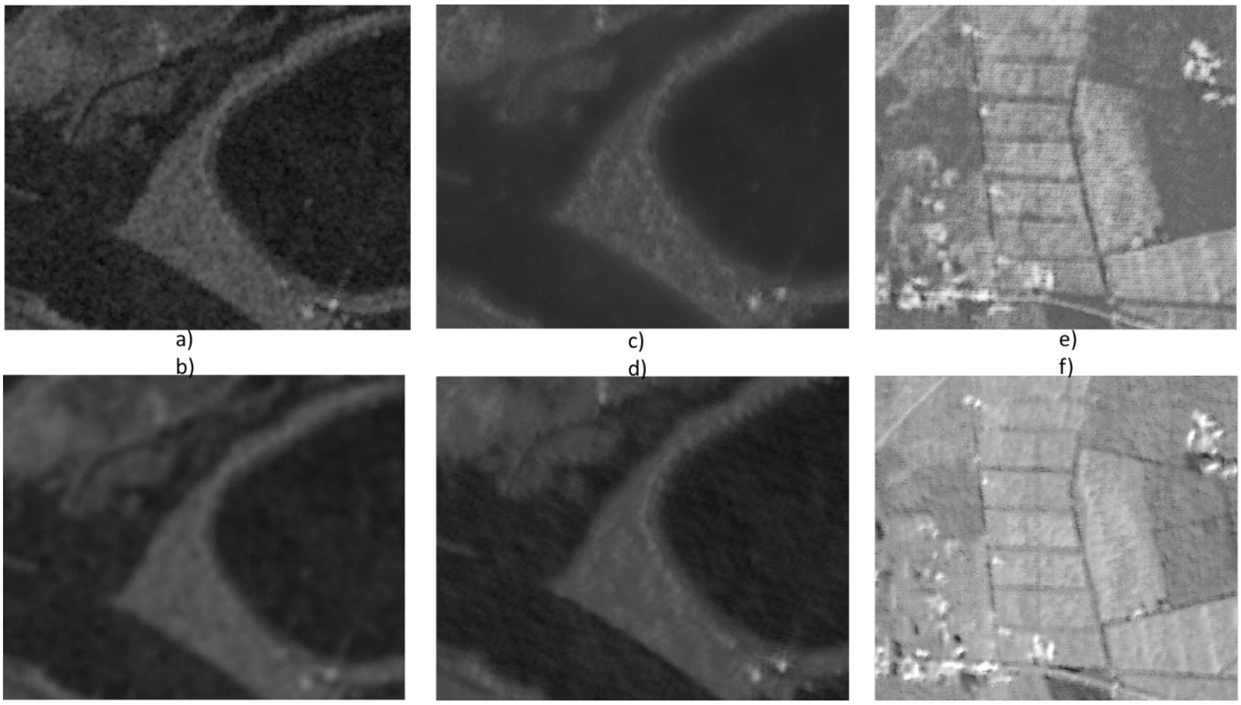

Figure 5: Comparison of achieved with the Lee filter and the Despeckle Net. Subfigure a) shows an original image patch, b) the Lee filter result, c) Despeckle Net with a kernel size of 5 and d) with a kernel size of 3. Figure e) shows results using deconvolution blocks and f) Resize Convolutions.

8.1. Impact of the kernel size

Figure 5b) and c) show a comparison of a water area with a 5x5 respectively 3x3 kernel. As illustrated, especially the homogenous water bodies show a better smoothing effect and less noise with larger kernels. However, this comes at the cost of edge-preserving and vanished small objects like roads.

8.2. Transpose Convolution vs Resize Convolutions

As outlined in section 5, Transpose Convolutions tend to produce checkerboard artifacts. This effect could be observed especially in agricultural areas and is illustrated in Figure 5e). Figure 5f) in contrast shows the same patch using the Despeckle Net with Resize Convolutions. Here, the checkerboard artifacts do not appear.

8.3. Impact of the landcover

Figures 4) and 5a-c) show that the performance of the Despeckle Net highly depends on the land cover. The network outperforms the Lee filter on water surfaces as it removes much more noise. For cities and other small objects it also shows clear advantages as it is not based on spatial averages and preserves more details here. In mountainous areas, the network barely removed any noise and the Lee filter showed better results. The same applies for agricultural areas with present plants and thus high respectively heterogenous backscatter values. Here the Despeckle Net tends to introduce patterns with cloud-like textures.

8.4. Project Evaluation

The final work plan shows the original estimated hours were underestimated especially for the network training. Here technical issues and missing experience respectively literature for this specific application led to higher time consumption. However, those issues are hard to prevent and common in Deep Learning projects. The project delivered valuable insights and promising results. Given that, there is no major point that I would change when doing the project again.

9. Conclusion & Outlook

This project demonstrated that a deep Residual U-Net trained on noisy SAR data based on the S1GBM shows potential to remove speckle noise from real SAR acquisitions. The desired goal of a better performance compared to the Lee filter was not achieved. However, visual interpretation showed good results for cities and water bodies. With more time to adapt the noise model and the network architecture, the performance can likely be improved also for other landcover classes and current limitations can be overcome.

Bibliography

1. Vitale, S.; Ferraioli, G.; Pascazio, V. Multi-Objective CNN-Based Algorithm for SAR Despeckling. *IEEE Trans. Geosci. Remote Sens.* **2021**, 1–14, doi:10.1109/TGRS.2020.3034852.
2. Bauer-Marschallinger, B.; Cao, S.; Navacchi, C.; Freeman, V.; Reuß, F.; Geudtner, D.; Rommen, B.; Vega, F.C.; Snoeij, P.; Attema, E.; et al. The Sentinel-1 Global Backscatter Model (S1GBM) - Mapping Earth’s Land Surface with C-Band Microwaves 2021.
3. Bauer-Marschallinger, B.; Sabel, D.; Wagner, W. Optimisation of Global Grids for High-Resolution Remote Sensing Data. *Comput. Geosci.* **2014**, *72*, 84–93, doi:10.1016/j.cageo.2014.07.005.
4. Singh, P.; Pandey, R.S. Speckle Noise: Modelling and Implementation. *Int. Sci. Press* **2016**, *9*, 8717–8727.
5. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *ArXiv150504597 Cs* **2015**.
6. Komatsu, R.; Gonsalves, T. Comparing U-Net Based Models for Denoising Color Images. *AI* **2020**, *1*, 465–487, doi:10.3390/ai1040029.
7. Fracastoro, G.; Magli, E.; Poggi, G.; Scarpa, G.; Valsesia, D.; Verdoliva, L. Deep Learning Methods For Synthetic Aperture Radar Image Despeckling: An Overview Of Trends And Perspectives. *IEEE Geosci. Remote Sens. Mag.* **2021**, *9*, 29–51, doi:10.1109/MGRS.2021.3070956.
8. Saharia, C.; Ho, J.; Chan, W.; Salimans, T.; Fleet, D.J.; Norouzi, M. Image Super-Resolution via Iterative Refinement. *ArXiv210407636 Cs Eess* **2021**.
9. Aitken, A.; Ledig, C.; Theis, L.; Caballero, J.; Wang, Z.; Shi, W. Checkerboard Artifact Free Sub-Pixel Convolution: A Note on Sub-Pixel Convolution, Resize Convolution and Convolution Resize. *ArXiv170702937 Cs* **2017**.
10. Wagner, W.; Bauer-Marschallinger, B.; Navacchi, C.; Reuß, F.; Cao, S.; Reimer, C.; Schramm, M.; Briese, C. A Sentinel-1 Backscatter Datacube for Global Land Monitoring Applications. *Remote Sens.* **2021**, *13*, 4622, doi:10.3390/rs13224622.