

Deep learning for semantic segmentation of SAR images

Sanja Šćepanović¹, Vladimir Ignatenko¹, Pekka Laurila¹

¹ ICEYE
name.surname@iceye.fi

Recent advances in deep learning (DL) techniques for computer vision have lead to the application of DL in a number of fields that rely on computer vision. Example fields are autonomous driving, computational photography, image search engines, medical diagnostics, augmented reality, and remote sensing data analytics. In this short paper, we present our work-in-progress on the application of deep learning for synthetic aperture radar (SAR) image analytics. In particular, we apply the semantic segmentation DL algorithms towards the land cover segmentation.

The tree classes of DL algorithms in computer vision are based on their main tasks. In the order of increasing complexity, they are:

- **classification**: assigning an image to a class based on what is (mainly) represented in it, for example a ship, oil tank, sea or land;
- **object detection**: detecting (and, in more advanced versions of the algorithms, localizing) presence of particular objects in an image. These algorithms can detect several objects in the given image. A notable example task in this category is ship detection in SAR images;
- **semantic segmentation**: assigning a class to each *pixel* in an image based on which class (image segment) it belongs to. These algorithms not only detect and localize objects in the image, but also output their exact areas and boundaries.

Training		Label	
Source	Sentinel-1 GRD	Source	Corine mask
Polarization	VV & VH	Res.	20 × 20m
Res. SM	23 × 23m	Fine classes	48
Res. IW	20 × 22m	Top classes	6
Count	60 × 10,000	Count	600,000

Table 1: Datasets info



Figure 1: Corine mask for Finland

DL algorithms require large datasets (on the scales of millions of images) and long training times. **Transfer learning** is an effective approach to tackle this issue by reusing DL networks pre-trained on available large datasets for a relatively similar task and then fine-tuning them at the smaller data at task.

Related work As discussed in the review of DL in remote sensing [11], the specificity of remote sensing imagery (compared to ordinary RGB or grayscale images) results in specific challenges in this area. For example, remote sensing data are georeferenced, often multi-modal, with particular imaging geometries, there are interpretation difficulties and the ground-truth or labeled data needed for DL are often still lacking. Nevertheless, a number of research papers tackling remote sensing imagery with DL techniques is already published. Object detection (and localization) is usually termed *automatic target recognition* (ATR) in the context of remote sensing, and SAR-ATR when it comes to SAR images in particular [2]. Penatti et al. [7] evaluated the transfer learning from recognition of everyday objects to remote sensing objects and showed it is successful. For instance, transfer learning is employed to tackle the lack of labeled data in ATR [4]. In general, there is already significant success in applying DL techniques for SAR-ATR and scene understanding. However, for the *semantic image segmentation* and image preprocessing, advances are yet to be achieved [9].

Semantic segmentation of SAR images for creating the land cover mask

Task Automatically segment pixels of a SAR image based on the type of land cover they represent.

Datasets We analyze a set of Sentinel-1 SAR images (Level-1 Ground Range Detected – GRD products in Stripmap (SM) or Interferometric Wide (IW) swath modes). For the labels, we use the existing Corine mask¹, providing mix of automatic and human experts labeled information for the type of land cover divided in over 40 classes. Corine is a program by the EU and stands for *coordination of information on the environment*. While a Corine mask of 100 × 100m spatial resolution is available for most of the EU territory, the Finnish Environment Institute created a more precise, 20 × 20m spatial resolution mask for Finland. We choose to work with the more precise, Finnish, mask, because its resolution corresponds better with the resolution of Sentinel-1 images, in particular, when using SM (strip map) and IW (interferometric wide swath) modes GRD products.

¹<https://land.copernicus.eu/pan-european/corine-land-cover>

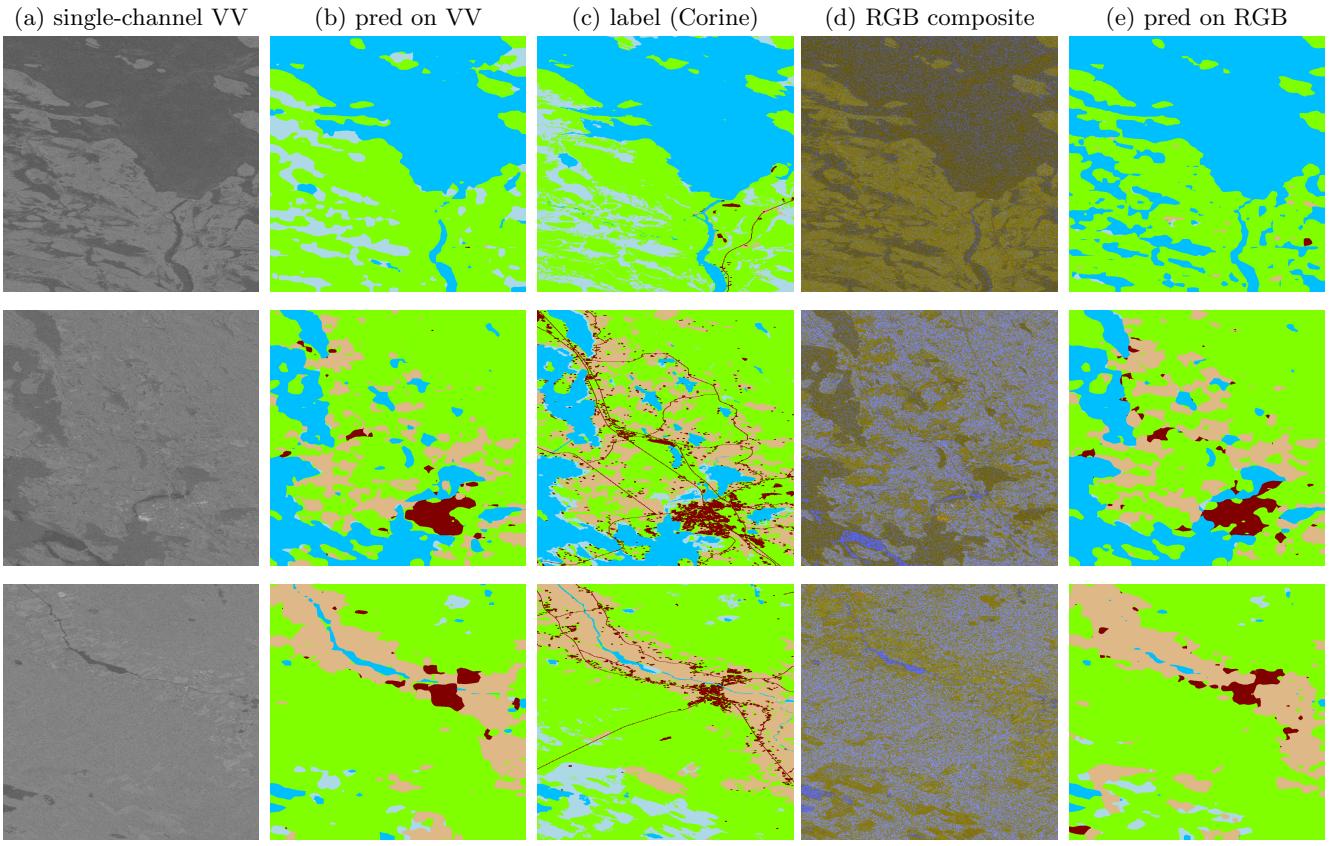


Figure 2: **DeepLab-V3 results:** (a) single-channel VV images, (d) RGB composite (VV, VH, VV-VH), (c) Corine mask labels, and (b) and (e) corresponding segmentation outputs on (a) and (d), respectively.

Hardware setup and preprocessing We develop and test DL models on a single GPU-machine (NVIDIA GeForce GTX 1080). Using this setup, the maximum image size we can work with is around 1000×1000 pixels (depending on the model) and so we split the images in a number of pieces accordingly. Moreover, the Sentinel-1 images come in two polarizations and so we try two setups for the training and testing data. The first setup we term **single-channel data** and both, VV and VH are used as separate images. The second setup we term **RGB composite data** and combined images are created by assigning VV to the red, VH to the green and the difference VV-VH to the blue channel. Note that in the second setup we end up with half the size of the single-channel dataset.

Model	DeepLab-V3		FC-DenseNet	
Dataset	single	RGB	single	RGB
Precision	.81	.87	.88	.80
F1 score	.84	.81	.82	.83
Mean IoU	.33	.32	.32	.31
Forest	.95	.94	.93	.94
Water	.58	.58	.49	.58
Swamp	.25	.11	.24	.15
Built. env.	.09	.18	.22	.10
Agriculture	.15	.29	.36	.60

Table 2: Results on single-channel and RGB composite datasets. We point in bold when the results on one of datasets are significantly better than on the other.

in better classification of the agriculture class using both algorithms. For the built environment class, the results are inconsistent.

Summary and future work Our work indicates the potential for applying existing DL models to SAR image semantic segmentation. However, there is still the need for improved precision compared to traditional established methods in the SAR image analytics. Future lines of work towards that goal include better data preprocessing, adapting existing DL models to SAR data and using SAR data with a better resolution, such as from upcoming ICEYE X-band satellite constellation.

Preliminary results We use the open-source Semantic Segmentation Suite² to test several state-of-the-art *pre-trained* semantic segmentation DL models [3] including DeepLab [1], FC-DenseNet [5], Global Convolutional Network [8], RefineNet [6], and ICNet [10]. Hence, we apply *transfer learning*. Preliminary results for two of the models: DeepLab-V3-Res152 and FC-DenseNet-Res152 are presented in Table 2. The training with the other models is currently ongoing. Example images with segmentation output of the DeepLab-V3 model are also shown in Fig. 2.

There are a couple of points to note from the preliminary results. First, we can see how the algorithms perform better on single-polarization dataset when it comes to detecting swamps compared to RGB composite dataset on which the swamp class is incorrectly identified as the water class. On the other hand, RGB composite dataset results

²<https://github.com/GeorgeSeif/Semantic-Segmentation-Suite>

References

- [1] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L Yuille. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE transactions on pattern analysis and machine intelligence*, 40(4):834–848, 2018.
- [2] S. Chen, H. Wang, F. Xu, and Y. Q. Jin. Target classification using the deep convolutional networks for sar images. *IEEE Transactions on Geoscience and Remote Sensing*, 54(8):4806–4817, Aug 2016. ISSN 0196-2892. doi: 10.1109/TGRS.2016.2551720.
- [3] Alberto Garcia-Garcia, Sergio Orts-Escalano, Sergiu Oprea, Victor Villena-Martinez, and Jose Garcia-Rodriguez. A review on deep learning techniques applied to semantic segmentation. *arXiv preprint arXiv:1704.06857*, 2017.
- [4] Zhongling Huang, Zongxu Pan, and Bin Lei. Transfer learning with deep convolutional neural network for sar target classification with limited labeled data. *Remote Sensing*, 9(9):907, 2017.
- [5] Simon Jégou, Michal Drozdzal, David Vazquez, Adriana Romero, and Yoshua Bengio. The one hundred layers tiramisu: Fully convolutional densenets for semantic segmentation. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2017 IEEE Conference on*, pages 1175–1183. IEEE, 2017.
- [6] Guosheng Lin, Anton Milan, Chunhua Shen, and Ian Reid. Refinenet: Multi-path refinement networks for high-resolution semantic segmentation. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.
- [7] O. A. B. Penatti, K. Nogueira, and J. A. dos Santos. Do deep features generalize from everyday objects to remote sensing and aerial scenes domains? In *2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 44–51, June 2015. doi: 10.1109/CVPRW.2015.7301382.
- [8] Chao Peng, Xiangyu Zhang, Gang Yu, Guiming Luo, and Jian Sun. Large kernel matters—improve semantic segmentation by global convolutional network. *arXiv preprint arXiv:1703.02719*, 2017.
- [9] Liangpei Zhang, Lefei Zhang, and Bo Du. Deep learning for remote sensing data: A technical tutorial on the state of the art. *IEEE Geoscience and Remote Sensing Magazine*, 4(2):22–40, 2016.
- [10] Hengshuang Zhao, Xiaojuan Qi, Xiaoyong Shen, Jianping Shi, and Jiaya Jia. Icnet for real-time semantic segmentation on high-resolution images. *arXiv preprint arXiv:1704.08545*, 2017.
- [11] Xiao Xiang Zhu, Devis Tuia, Lichao Mou, Gui-Song Xia, Liangpei Zhang, Feng Xu, and Friedrich Fraundorfer. Deep learning in remote sensing: a review. *arXiv preprint arXiv:1710.03959*, 2017.