Wide-Area Land Cover Mapping With Sentinel-1 Imagery Using Deep Learning Semantic Segmentation Models

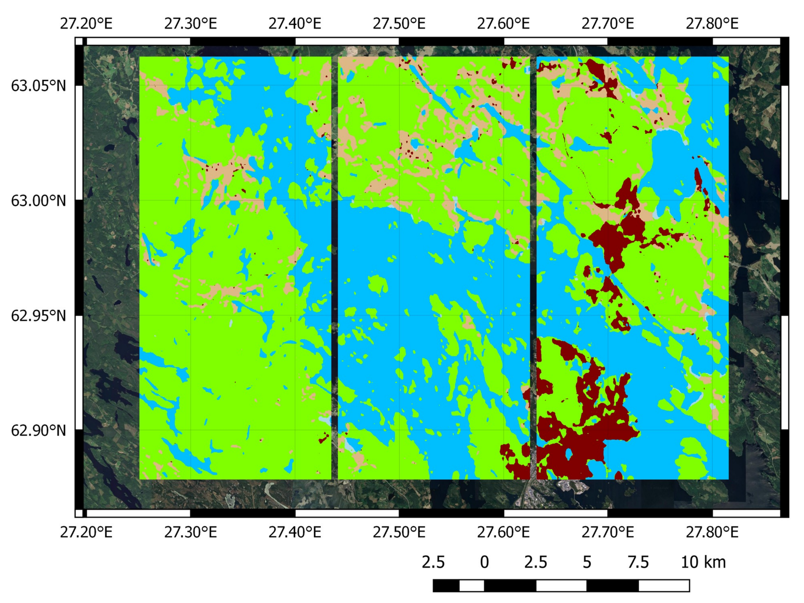
Sanja Šćepanović[1,3*], Oleg Antropov[2], Pekka Laurila[3], Vladimir Ignatenko[3]

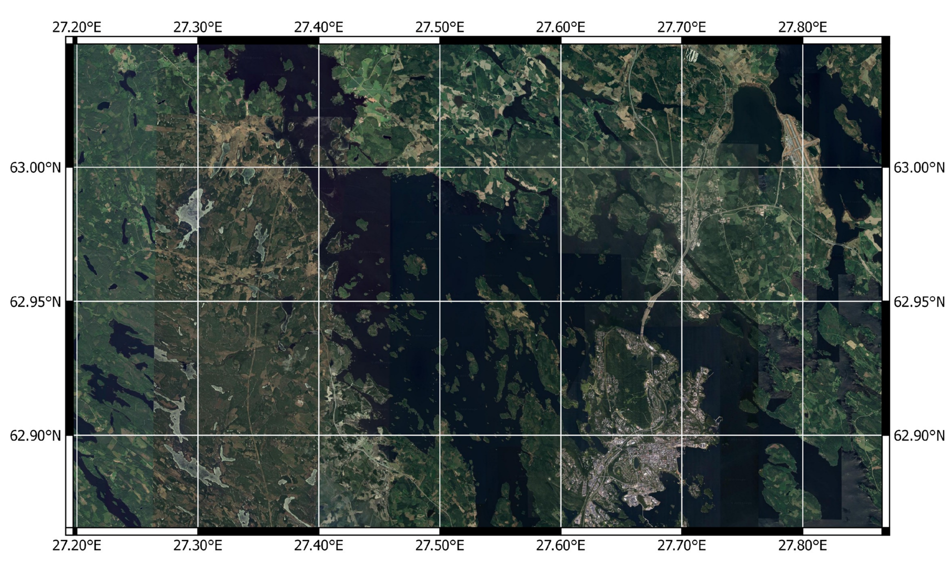
[1]Nokia Bell Labs, Cambridge, UK, [2]VTT, Helsinki, Finland, [3]ICEYE, Espoo, Finland

*the author was with ICEYE when the work was conducted

ABSTRACT

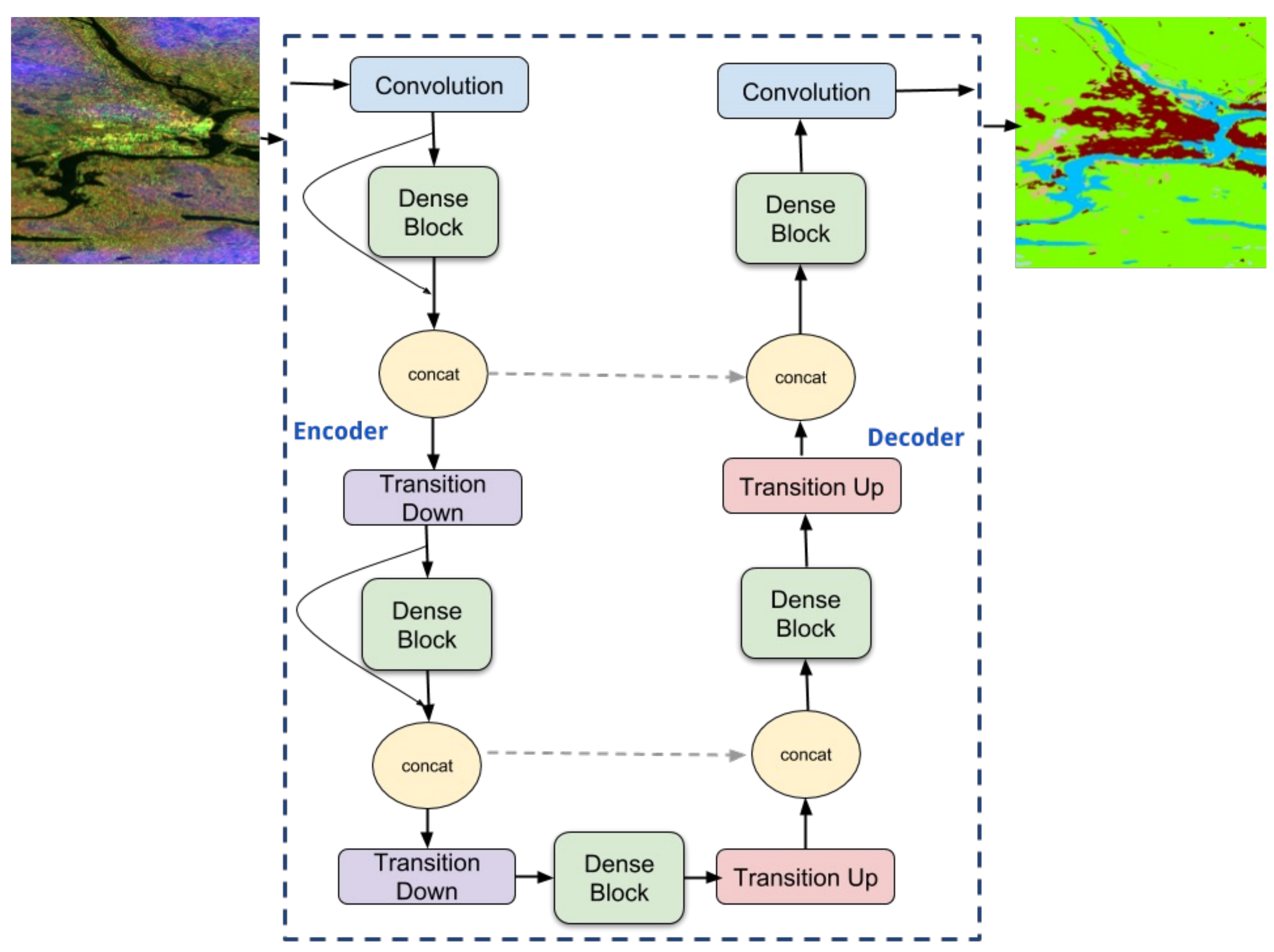
Recent studies demonstrated successful applications of deep learning models to small-scale Land Cover (LC) mapping tasks. However, it is not explicit which of the existing SOTA models for natural images are the best candidates for remote sensing data. We map fundamental LC classes with ESA Sentinel-1 C-band SAR images acquired during the whole summer season of 2018 in Finland. CORINE LC map served as a reference, and we trained the models to distinguish between the five major CORINE-based classes. Seven pretrained semantic segmentation models: U-Net [2], DeepLabV3+ [5], PSPNet [4], BiSeNet [1], SegNet [7], FC-DenseNet [6], and FRRN-B [3] are further fine-tuned and tested. Upon evaluation and benchmarking, all the models demonstrated solid performance with overall accuracy between 87.9% and 93.1%, with solid agreement (Kappa statistic between 0.75 and 0.86). The best models were fully convolutional DenseNets (FC-DenseNet) and SegNet (encoder-decoder-skip). Our results indicate that the semantic segmentation models are suitable for efficient wide-area mapping using satellite SAR imagery.



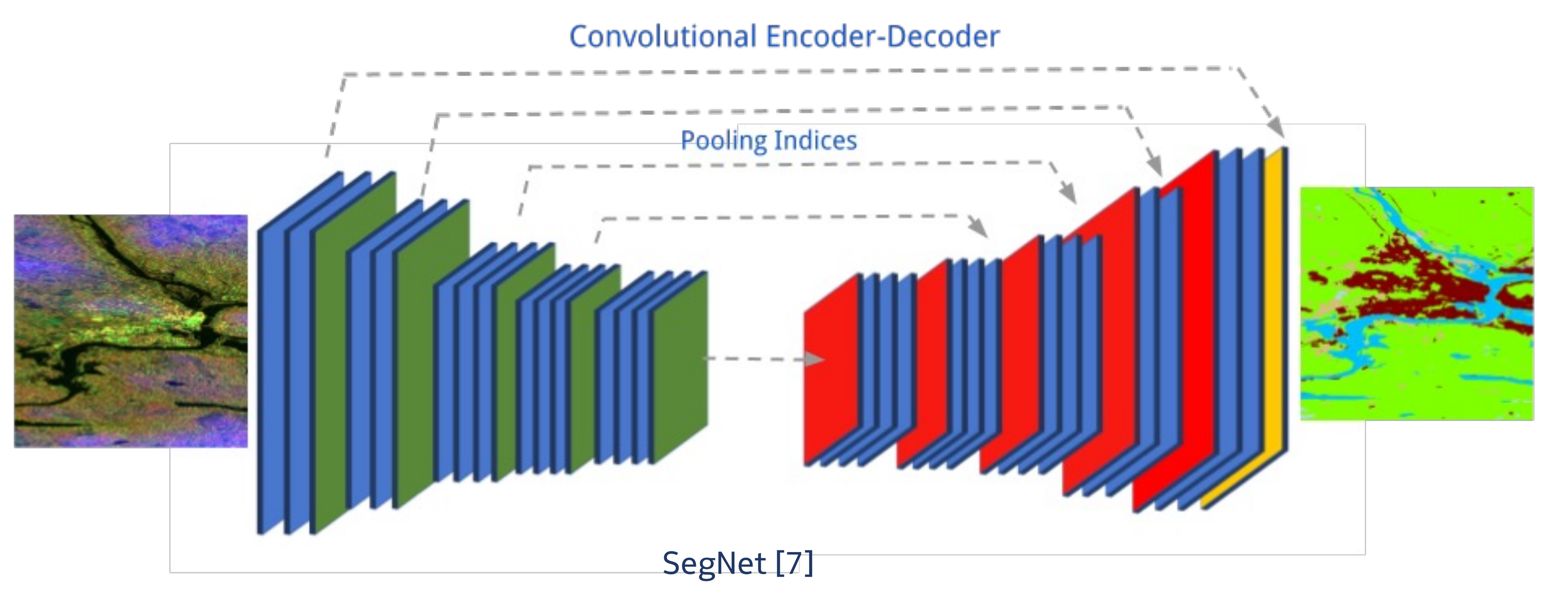


1 OBJECTIVE

Which of the SOTA deep learning semantic segmentation models are most suitable for LC mapping using Sentinel-1 SAR data?

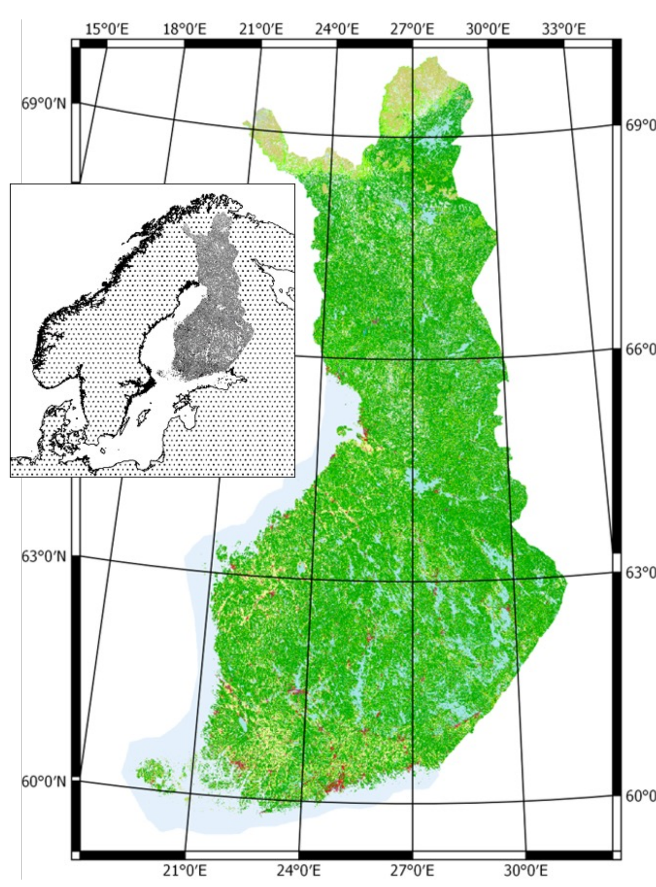


FC-DenseNet [6]



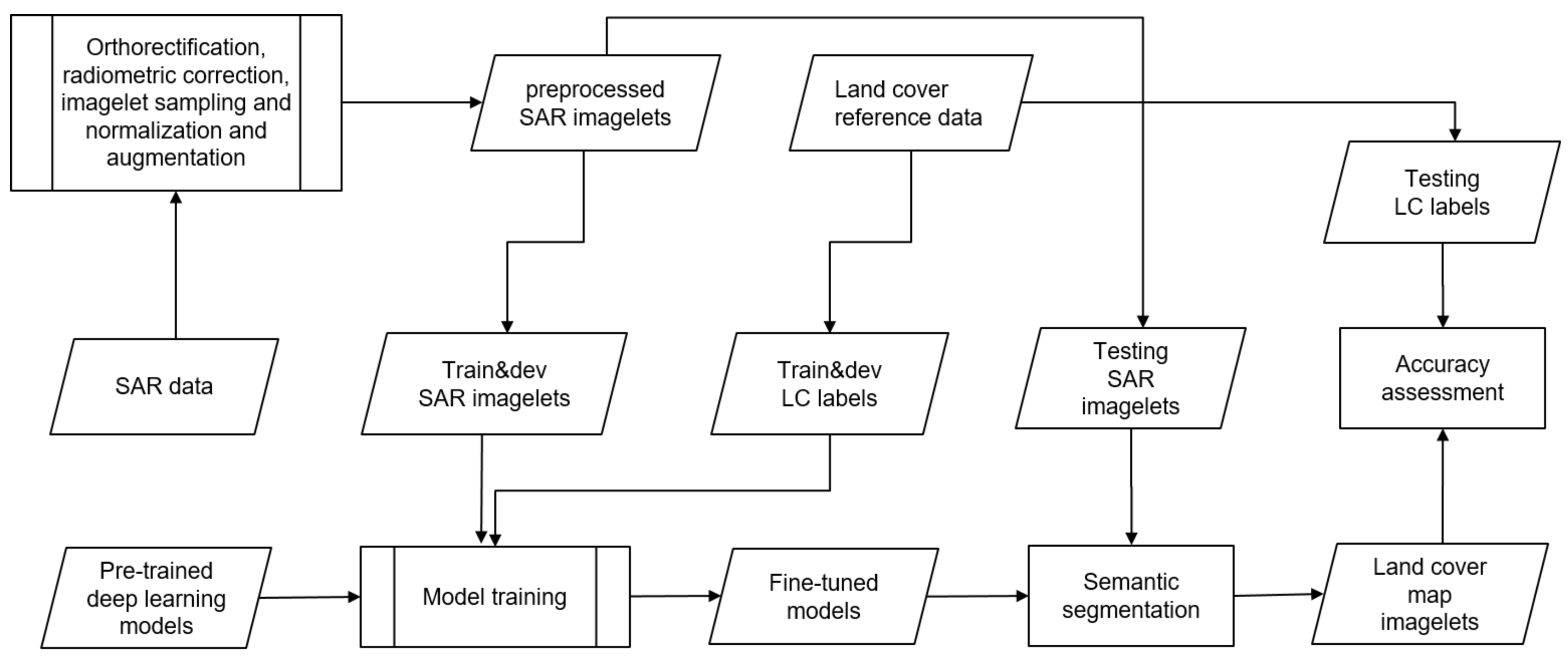
SegNet [7]

Are such models suitable for wide-area land cover mapping in the cloud obscured boreal zone and polar regions, such as in Finland?



2 APPROACH

To improve performance, we employed models whose encoders were pretrained for the ImageNet classification task and fine-tuned them on our SAR dataset.

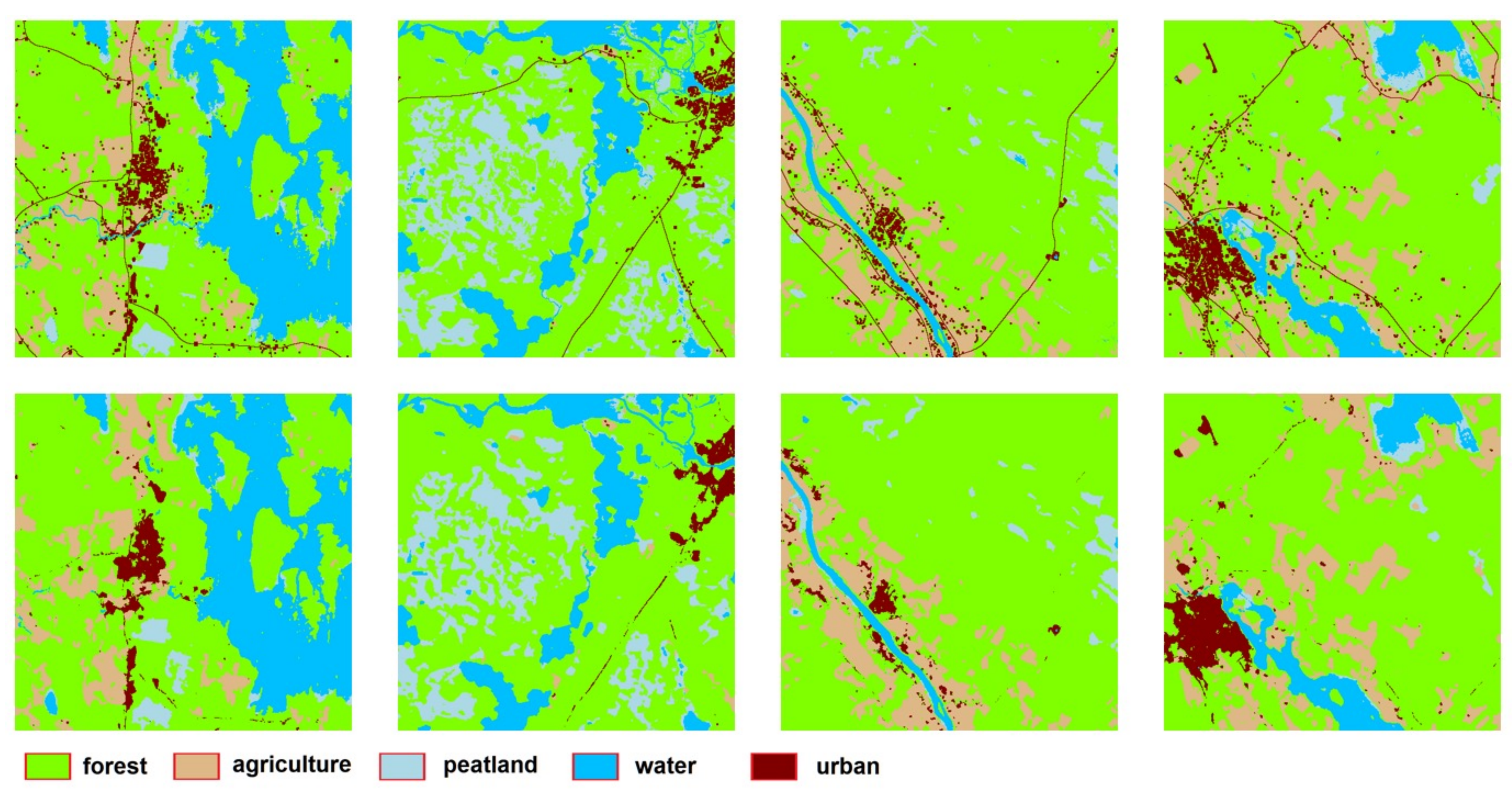


3 RESULTS

All the models performed relatively well on both datasets achieving an overall accuracy above 87% for each model. Two models performed particularly well, achieving an accuracy score nearly 93% on both datasets; those are SegNet, and FC-DenseNet. The advantage for SegNet is that its training and inference times were 2.5 times better compared to the FC-DenseNet of similar accuracy.

TABLE IV
SUMMARY OF THE CLASSIFICATION PERFORMANCE AND EFFICIENCY OF DEEP LEARNING MODELS

LC classes	**Test scale** (km^2)	**Accuracy (UA, PA %)**						
		BiSeNet	DeepLabV3+	SegNet	FRRN-B	U-Net	PSPNet	FC-DenseNet
Urban fabric	7906	59, 29	67, 29	70, 33	68, 33	65, 31	65, 31	69, 34
Agricultural areas	22659	63, 69	76, 70	79, 79	80, 77	69, 77	69, 77	80, 79
Forested areas	199327	91, 96	92, 97	94, 97	94, 97	94, 96	94, 96	94, 97
Peatland, bogs and marshes	10686	77, 48	83, 42	81, 61	77, 65	77, 51	77, 51	78, 65
Water bodies	53022	96, 86	97, 97	98, 98	98, 98	97, 97	97, 97	98, 98
Overall Accuracy (%)		88.48	91.37	**92.78**	92.69	92.25	91.37	**92.78**
Kappa		0.758	0.818	**0.851**	0.849	0.839	0.823	**0.851**
Average inference time (s / image)		0.043	**0.031**	0.073	0.143	0.085	0.053	0.196



4 CONCLUSION

Our study demonstrated the potential for applying state-of-the-art semantic segmentation models to SAR image classification with high accuracy. Several models were benchmarked in a countrywide classification experiment using Sentinel-1 IW-mode SAR data, reaching nearly 93% overall classification accuracy with the best performing models (SegNet and FCDenseNet). This indicates strong potential for using pretrained CNNs for further fine-tuning and seems particularly suitable when the number of training images is. Improvements for future work include multitemporal approaches, employing DEM models, data fusion, very high-resolution SAR imagery, a well as developing models specifically for SAR.

REFERENCES

[0] Šćepanović, Sanja, et al. "Wide-Area land cover mapping with sentinel-1 imagery using deep learning semantic segmentation models." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 14 (2021): 10357-10374.

[1] C. Yu, J. Wang, C. Peng, C. Gao, G. Yu, and N. Sang, “BiseNet: Bilateral segmentation network for real-time semantic segmentation,” in Proc. Eur. Conf. Comput. Vision, 2018, pp. 325–341.

[2] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” in Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervention, 2015, pp. 234–241.

[3] T. Pohlen, A. Hermans, M. Mathias, and B. Leibe, “Full-resolution residual networks for semantic segmentation in street scenes,” in Proc. IEEE Conf. Comput. Vision Pattern Recognit., 2017, pp. 4151–4160.

[4] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, “Pyramid scene parsing network,” in Proc. IEEE Conf. Comput. Vision Pattern Recognit., 2017, pp. 2881–2890.

[5] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, “Encoderdecoder with atrous separable convolution for semantic image segmentation,” 2018, arXiv:1802.02611.

[6] S. Jégou, M. Drozdzal, D. Vazquez, A. Romero, and Y. Bengio, “The one hundred layers Tiramisu: Fully convolutional densenets for semantic segmentation,” in Proc. IEEE Conf. Comput. Vision Pattern Recognit. Workshops, 2017, pp. 1175–1183.

[7] V. Badrinarayanan, A. Kendall, and R. Cipolla, “SegNet: A deep convolutional encoder-decoder architecture for image segmentation,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 12, pp. 2481–2495, Dec. 2017.

[8] Code with documentation: https://github.com/sanja7s/DL_SemSAR_with_docs

ACKNOWLEDGEMENT

This project was conducted as part of the EIT Digital business development experience. https://eit.europa.eu/our-communities/eit-digital