IGARSS 2023 ABSTRACT

Graph Reasoned Multi-Scale Road Segmentation in Remote Sensing Imagery

Andrew Alexander Vekinis

Computational Intelligence Laboratory
Institute of Informatics & Telecommunications,
National Center for Scientific Research "Demokritos", Athens, Greece

The extraction of road networks from remote sensing imagery is a challenging problem due to a multitude of factors such as road surface texture variations, occlusion from clouds, natural canopy and buildings, shadow, as well as road widths and changes in topological complexity. Deviated off-nadir imagery can also compound such issues [1]. However in real-life applications, from designing autonomous navigation planning systems to disaster response scenarios [2], many of these factors may be present. Attempting to segment road sections in an image solely using Convolutional Networks via pixel-wise classification in these cases can lead to unusable fragmented predictions as such networks cannot capture local and distant contextual information that is critical for ensuring road connectivity [3].

In this paper, we propose an efficient multi-task road segmentation model that learns to extract road features at multiple scales, consequently using dual Graph-Convolutional Networks (dGCN) [4] to improve road connectivity. Concretely, the encoder-decoder model inspired by [5] and [6], makes use of Pyramid Pooling Modules (PPM) and Feature Pyramid Networks (FPN) to learn the simultaneous extraction of road segments and their corresponding orientations at scales of increasing fidelity. The predicted road segments at each scale are then fed into a dGCN which is comprised of two parts, a spatial GCN and a feature GCN. The spatial GCN takes into consideration the spatial relationship of every pixel in the image while the feature GCN compares how the predicted road segments correlate with other abstract features located by the spatial GCN. This graph reasoning procedure can identify localized similarities between features which are added to the original feature map which allows for enhanced road connectivity, since local and distant contextual information is now being considered.

The capabilities of the developed model are demonstrated on three diverse road extraction datasets (SpaceNet [7], DeepGlobe [2] and Massachusetts Roads [8]) as can be seen in figure 1. During training, as was done in [3] and [9], all satellite images are cropped to 512^2 for the DeepGlobe [2] and Massachusetts Roads [8] datasets with overlapping regions of 256 pixels. For [7], cropped images of size 650^2 were used with overlapping regions of 215 pixels. From each corresponding cropped image, a random region of 256^2 is extracted from within, augmented by flipping, mirroring and rotation, and then used as the input to the model. During validation, the same cropped image sizes were used corresponding to each dataset, without overlap. Preliminary evaluation metrics for precision, recall, F1 score, IoU^r (relaxed) [8], IoU^a (accurate) [8], and APLS [7] are used to gauge the performance of the developed model as shown in table 1. Computational requirements are also shown in table 2.

As demonstrated in figure 1, in each of the datasets there are cases where there is occlusion from natural canopy as well as buildings that cast shadows of varying intensity. It can be shown however that roads of different sizes are extracted in a non fragmented, continuous manner in spite of the fact that some the imagery is off-nadir. The approach outlined in this paper is able to rapidly extract such road networks from high-resolution satellite and aerial images with a low memory footprint, ensuring that it is useful in real-life applications. It should be noted that this work is continuing with a view towards improving the performance and reducing the computational cost of the developed model even further.

Datasets	Preliminary Result Evaluation Metrics					
	Precision	Recall	F1	IoU^r	IoU^a	APLS
Massachusetts Roads [8]	83.58	85.73	84.64	73.37	57.78	68.23
DeepGlobe [2]	81.17	81.73	81.45	68.71	61.78	69.81
Spacenet [7]	62.13	61.66	62.32	44.42	58.39	60.11

Table 1: Evaluation of our approach in three diverse road extraction datasets.

Compute Requirements	
GFLOPs	Training Memory (GB)
134.72	11.5

Table 2: Computational requirements of our approach

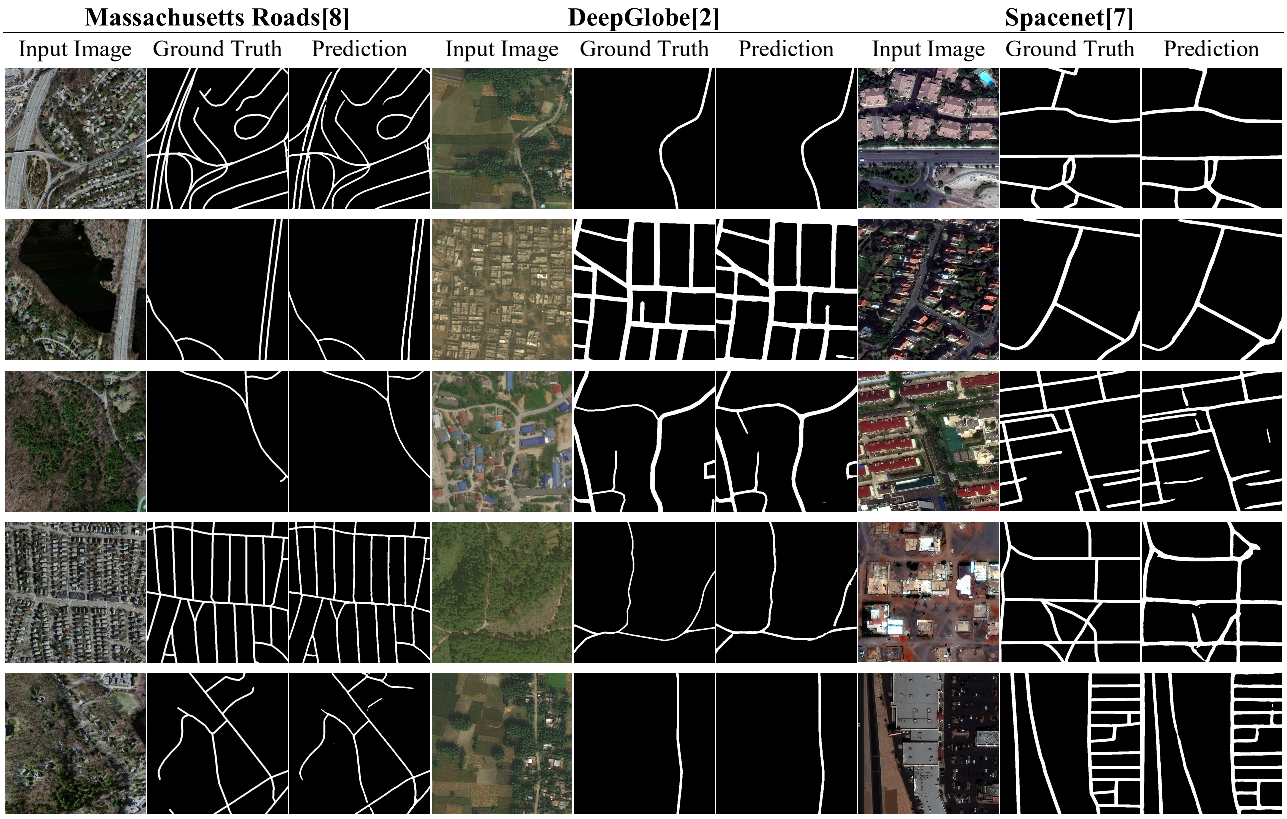

Figure 1: Remote sensing images from the Massachusetts Roads (left), DeepGlobe (centre) and Spacenet (right) test datasets. In each group, the input image is on the left, with ground truth labels in the centre and the predicted roads with our model on the right. Zoom in for better detail.

REFERENCES

[1] N. Weir *et al*., "SpaceNet MVOI: A Multi-View Overhead Imagery Dataset", *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019, pp. 992-1001, doi: 10.1109/ICCV.2019.00108.

[2] I. Demir *et al*., "DeepGlobe 2018: A Challenge to Parse the Earth through Satellite Images", *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2018, pp. 172-17209, doi: 10.1109/CVPRW.2018.00031.

[3] A. Batra, S. Singh, G. Pang, S. Basu, C. V. Jawahar and M. Paluri, "Improved Road Connectivity by Joint Learning of Orientation and Segmentation", *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 10377-10385, doi: 10.1109/CVPR.2019.01063.

[4] L. Zhang et al., "Dual Graph Convolutional Network for Semantic Segmentation", 2019 British Machine Vision Conference (BMVC), 2019, https://doi.org/10.48550/arXiv.1909.06121.

[5] Z. Liu, H. Mao, C.Y. Wu, C. Feichtenhofer, T. Darrell, S. Xie, "A ConvNet for the 2020s", 2022 Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022, pp. 11976-11986

[6] T. Xiao, Y. Liu, B. Zhou, Y. Jiang, J. Sun, "Unified perceptual parsing for scene understanding". In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds.) ECCV 2018. LNCS, vol. 11209, pp. 432–448. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-01228-1_26

[7] A. Etten, D. Lindenbaum, T. Bacastow, "SpaceNet: A Remote Sensing Dataset and Challenge Series", 2018, https://doi.org/10.48550/arXiv.1807.01232

[8] V. Mnih, "Machine Learning for Aerial Image Labeling", PhD Dissertation, University of Toronto, 2013.

[9] W.G.C. Bandara, J.M.J. Valanarasu, V.M .Patel, "Spin road mapper: extracting roads from aerial images via spatial and interaction space graph reasoning for autonomous driving". arXiv preprint arXiv:2109.07701 (2021)