PaRK-Detect: Towards Efficient Multi-Task Satellite Imagery Road Extraction via Patch-Wise Keypoints Detection

Supplementary Material

Shenwei Xie
xieshenwei@bupt.edu.cn
Wanfeng Zheng
zhengwanfeng@bupt.edu.cn
Zhenglin Xian
2019213264@bupt.edu.cn
Junli Yang
yangjunli@bupt.edu.cn
Chuang Zhang
zhangchuang@bupt.edu.cn
Ming Wu ✉
wuming@bupt.edu.cn

Beijing University of Posts and Telecommunications, China

1 Overview

In this supplementary material, we present more experimental results and analyses to support our proposed patch-wise road keypoints detection scheme and multi-task framework.

- We present samples of datasets we used for experiments in section 2.
- We provide the summary of the APLS metric in section 3.
- We provide the implementation details of graph optimization strategies in section 4.
- We conduct extensive experiments on the joint pyramid upsampling (JPU) module within the framework architecture in section 5.
- We show more visualization results of constructed road graphs in section 6.

2 Experimental Datasets

Fig. 1 shows several representative satellite imagery in DeepGlobe Road Extraction Dataset [2], Massachusetts Roads Dataset [3], and RoadTracer Dataset [1].

JPU →	position locating	probability predicting	link predicting	DeepGlobe			Massachusetts Roads		
				P-F1	APLS	IoU	P-F1	APLS	IoU
				77.44	69.07	65.24	74.36	62.80	56.43
	✓			**78.04**	**69.83**	**65.53**	**74.80**	**63.37**	**56.80**
	✓	✓		76.85	68.54	64.97	73.49	62.15	56.38
	✓		✓	77.32	69.38	65.23	73.80	62.46	56.61
	✓	✓	✓	76.21	67.96	64.91	72.17	60.87	55.89

Table 1: Experimental results of introducing JPU for probability and link predicting network.

3 APLS Metric

To measure the similarity between ground truth and proposal road graphs, Van *et al.* [5] proposed Average Path Length Similarity (APLS). Based upon Dijkstra's shortest path algorithm, APLS sum the differences in optimal paths between ground truth and proposal graphs. Graph augmentation, node snapping, and symmetric comparisons are utilized in evaluating the APLS metric.

4 Graph Optimization Details

Connecting adjacent but unconnected endpoints. We traverse every two patches that are adjacent (1×2 rec-patch) and find out all rec-patches that two patches within are also road patches. For each rec-patch, we count the number of links of two patches in it respectively. If the numbers are both less than or equal to 1, then it means the road keypoints of these two patches are both endpoints of the predicted road graph. For two adjacent endpoints in a rec-patch, if they are unconnected, we add a link between them.

Removing triangle and quadrilateral. We traverse all 2^2 patches (2×2 super-patch) out of all 64^2 patches and find out all interconnected three patches that are also road patches within each super-patch. Then we remove the diagonal link between these three patches and remain the other two links. Besides, we also find out all super-patches that four patches within each of them are all road patches and connected as a quadrilateral. Then we remove the longest link of these four links based on the distance between road keypoints in these patches.

5 Framework Architecture

We conduct extensive experiments to verify the effectiveness of multi-scale feature fusion with the joint pyramid upsampling (JPU) [6] module for patch-wise keypoint locating instead of probability or link predicting. As shown in Tab. 1, introducing JPU for probability or link predicting network decreases overall road graph construction performance with pixel-based F1 score and APLS, which implies low-level detailed information is unnecessary and even detrimental for learning patch-wise road probability and adjacent relationships.

6 More Visualization Results

As shown in Fig. 2, we present more visualization results of constructed road graphs based on different road extraction approaches for comparison, including D-LinkNet [7], VecRoad [4], and our proposed PaRK-Detect scheme.

Figure 1: Examples of satellite imagery in DeepGlobe Road Extraction Dataset (**Top**), Massachusetts Roads Dataset (**Middle**), and RoadTracer Dataset (**Bottom**).

Figure 2: **More visualization results of constructed road graphs. Up Left:** original satellite imagery. **Up Right:** road extraction results based on D-LinkNet. **Down Left:** road extraction results based on VecRoad. **Down Right:** road extraction results based on PaRK-Detect scheme.

References

[1] Favyen Bastani, Songtao He, Sofiane Abbar, Mohammad Alizadeh, Hari Balakrishnan, Sanjay Chawla, Sam Madden, and David DeWitt. Roadtracer: Automatic extraction of road networks from aerial images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4720–4728, 2018.

[2] Ilke Demir, Krzysztof Koperski, David Lindenbaum, Guan Pang, Jing Huang, Saikat Basu, Forest Hughes, Devis Tuia, and Ramesh Raskar. Deepglobe 2018: A challenge to parse the earth through satellite images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 172–181, 2018.

[3] Volodymyr Mnih. *Machine learning for aerial image labeling*. University of Toronto (Canada), 2013.

[4] Yong-Qiang Tan, Shang-Hua Gao, Xuan-Yi Li, Ming-Ming Cheng, and Bo Ren. Vecroad: Point-based iterative graph exploration for road graphs extraction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8910–8918, 2020.

[5] Adam Van Etten, Dave Lindenbaum, and Todd M Bacastow. Spacenet: A remote sensing dataset and challenge series. *arXiv preprint arXiv:1807.01232*, 2018.

[6] Huikai Wu, Junge Zhang, Kaiqi Huang, Kongming Liang, and Yizhou Yu. Fastfcn: Rethinking dilated convolution in the backbone for semantic segmentation. *arXiv preprint arXiv:1903.11816*, 2019.

[7] Lichen Zhou, Chuang Zhang, and Ming Wu. D-linknet: Linknet with pretrained encoder and dilated convolution for high resolution satellite imagery road extraction. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 182–186, 2018.