Stacked U-Nets with Multi-Output for Road Extraction

同济大学 TONGJI UNIVERSITY

Tao Sun, Zehui Chen, Wenxiang Yang, Yin Wang

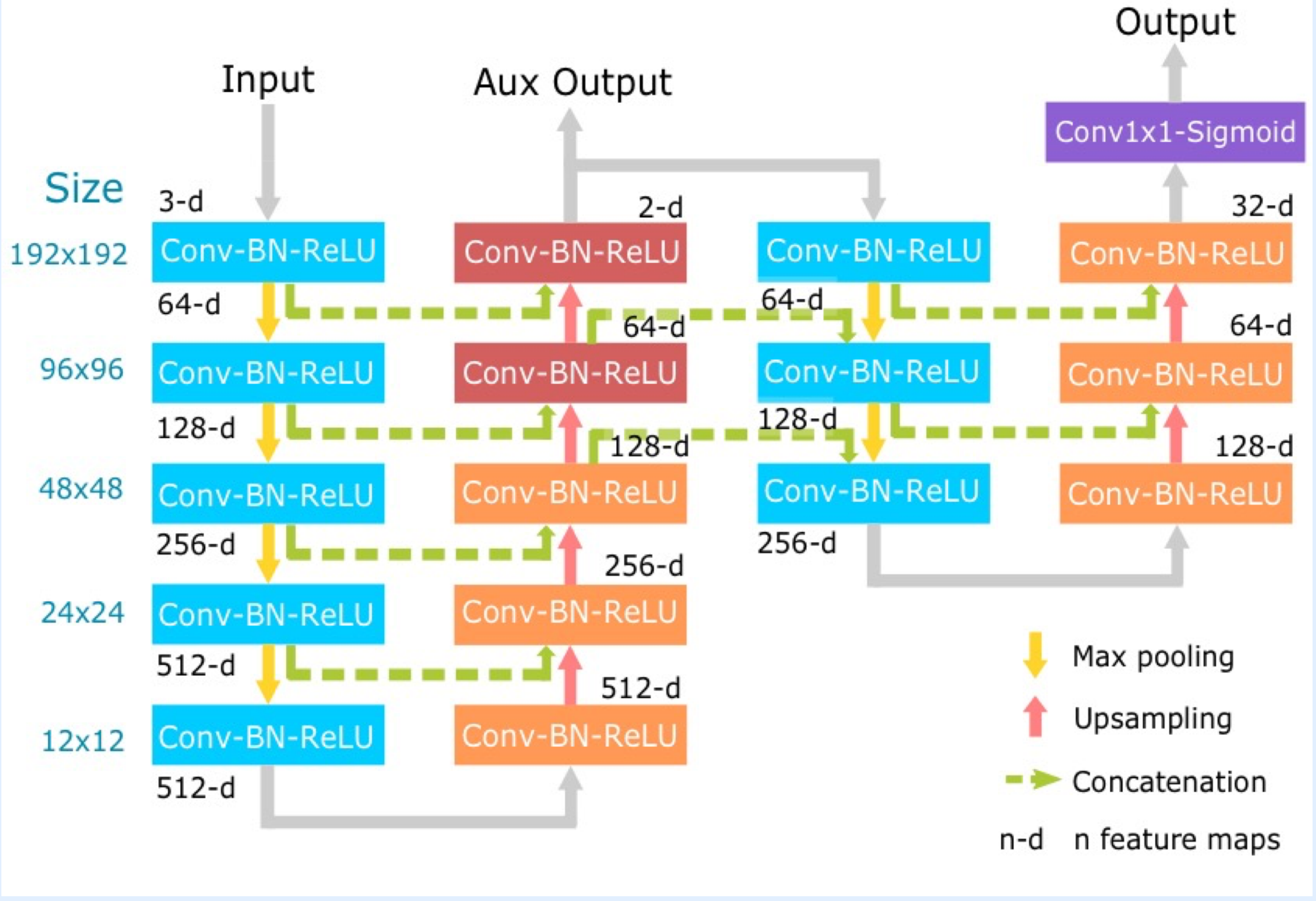
2018 CVPR SALT LAKE CITY • JUNE 18-22

Abstract

We propose a new method for road extraction using stacked U-Nets with multiple outputs. A hybrid loss function handles unbalanced training data. Post-processing methods including road vectorization and connection improve recall. The overall improvement of mean IoU compared to the vanilla VGG network is more than 20%.

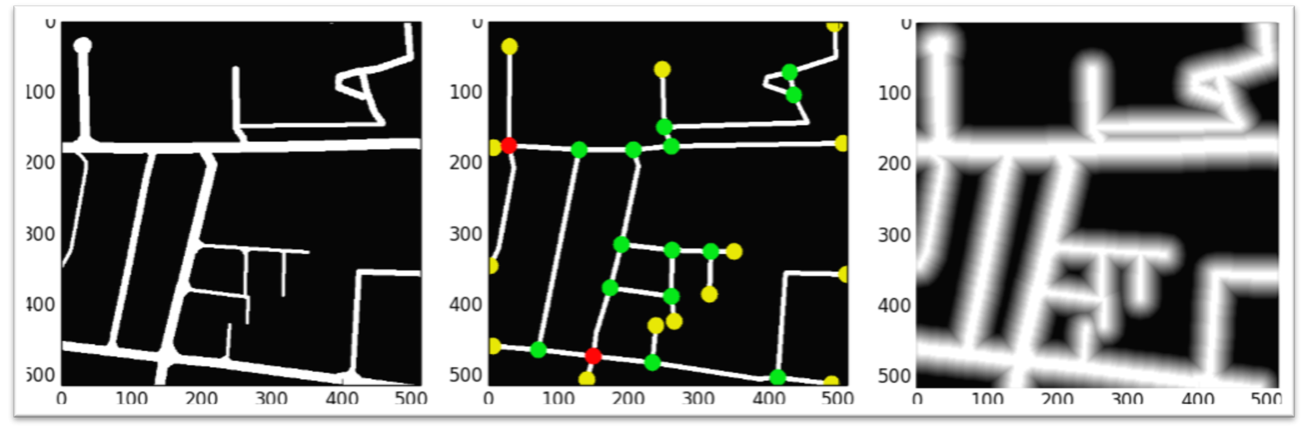
Our Network

- We concatenate two U-Nets to allow multiple outputs. The first U-Net outputs auxiliary information such as road topology and pixel distance to roads. The second U-Net generates road masks by classifying each pixel as road or not.
- We also extend the depth of the first U-Net from three to five to improve model accuracy.

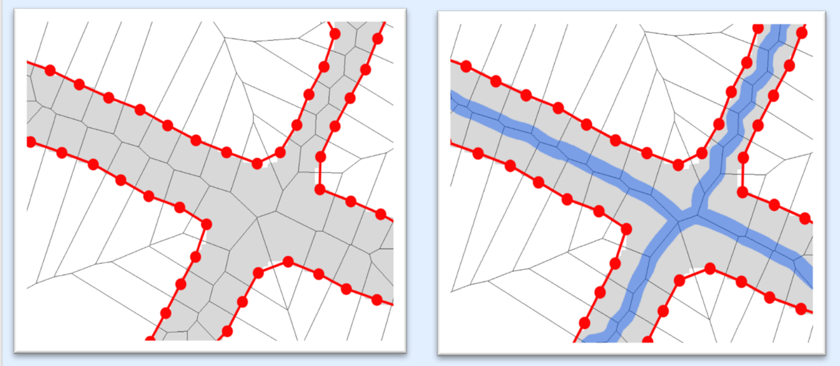


Implementation

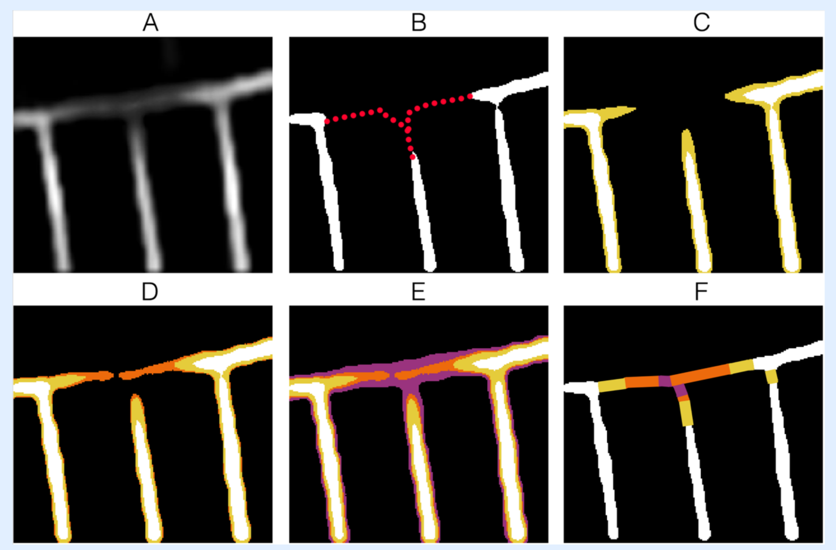
- **Multiple Outputs** force the network to learn structural features, in addition to road classification, we add the following outputs for each pixel:
  - **Outputing degrees**
  - **Distance to the nearest road**



Ground truth Outputing degree Distance to roads

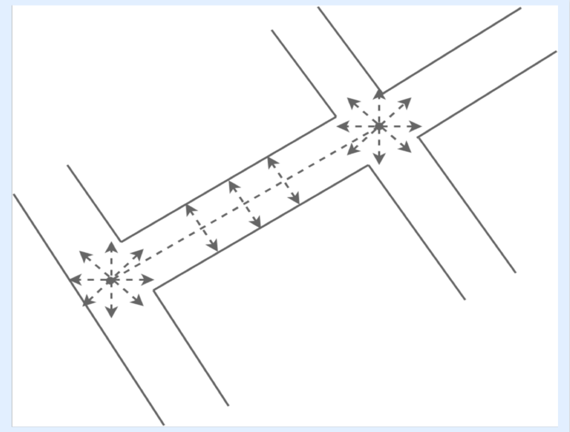


(A) Vectoring road bitmap

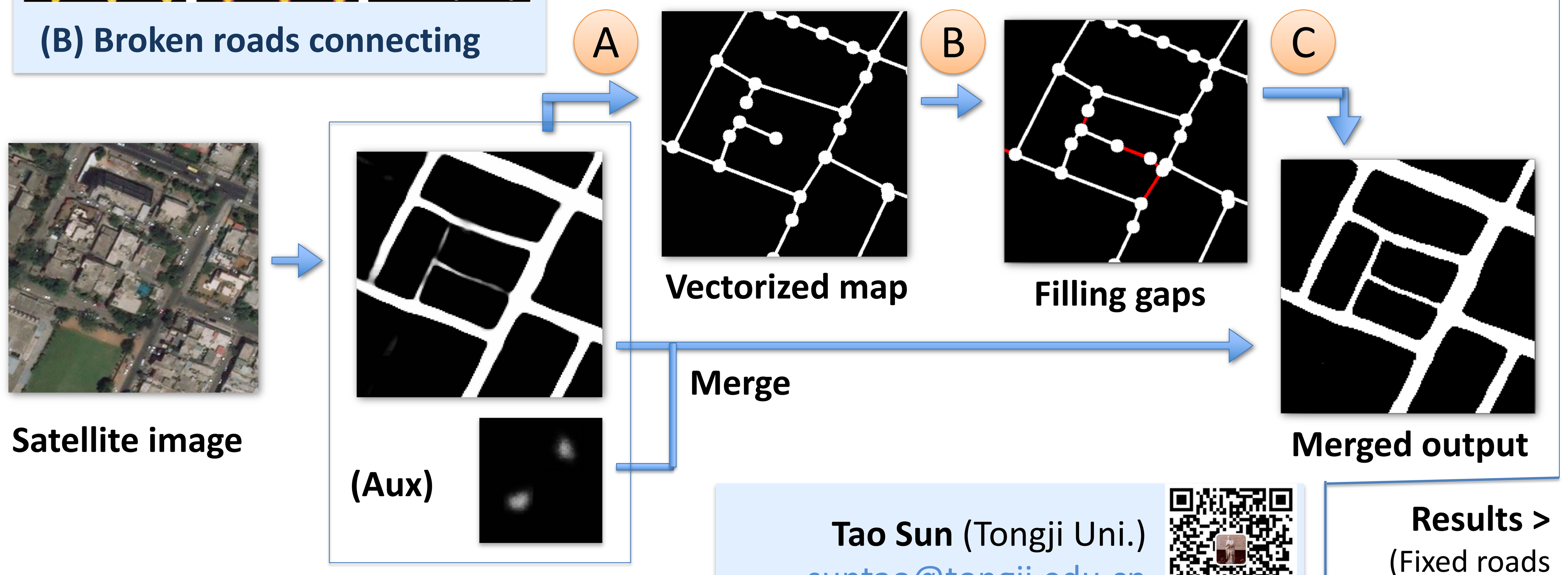


(B) Broken roads connecting

- **Post Processing**
  - **(A) Extracting Road Centerlines** Find the contours of the roads. Then the centerlines are extracted from the facets generated by Voronoi transform of the points on contours.
  - **(B) Connecting Broken Roads** Shortest path search from dead-ends in the vector map, decreasing thresholds helps avoid shortcuts.
  - **(C) Fitting Road Width** Rendering the vector map with appropriate road widths for IoU calculation.



(C) Fit road width



Tao Sun (Tongji Uni.)
suntao@tongji.edu.cn

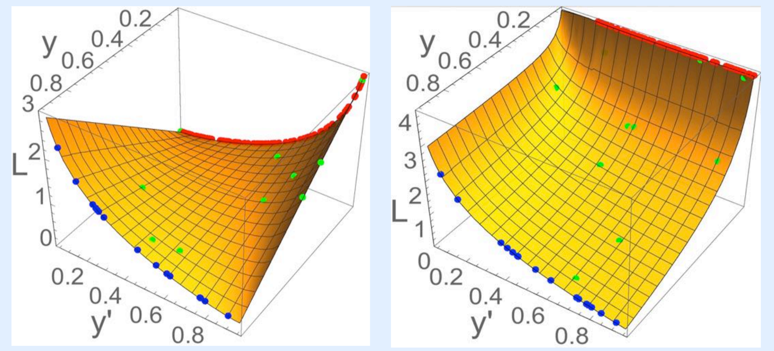
Training & Experiments

$$L_{ce} = -\frac{1}{N}\sum_{i=0}^{N}(y\log y' + (1-y)\log(1-y'))$$

$$L_{jaccard} = \frac{1}{N}\sum_{i=0}^{N}\frac{y_i y_i'}{y_i + y_i' - y_i y_i'}$$

$$L = L_{ce} - \lambda \log L_{jaccard}$$

- **Hybrid Loss.** Cross-entropy loss with unbalanced data leads to slow convergence and low accuracy. We use a hybrid loss of cross-entropy and Jaccard loss, which "lifts up" the back corner, and helps avoid the local optimum.



- Multi-output improves our U-Net about 2% in mIoU. Different λ values in loss function have significant impact. (Tab.1) Averaging overlapped prediction is necessary. Tab.2 shows the overall scores of different settings.

Table 1. Impact of hybrid loss function (on validation set)

λ	mIoU (%)	Precision (%)	Recall (%)
1	42.2	70.8	59.2
10	45.3	65.0	60.9
20	47.1	**71.6**	61.6
30	**49.0**	68.3	64.7
40	48.3	66.0	**65.7**

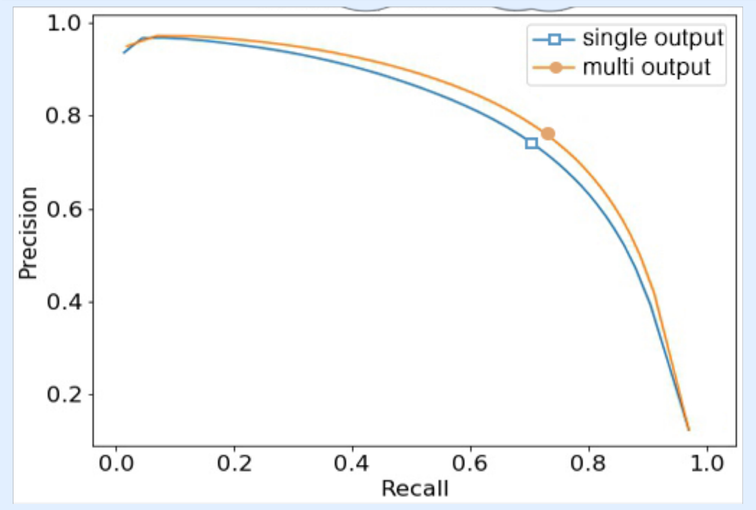


Table 2. Performance of different networks

Networks	mIoU (%)
VGG-19 (w/o multi output)	38.4
VGG-19 (with multi output)	39.3
U-Net (w/o multi output)	42.9
U-Net (with multi output)	44.3
Our U-Net (w/o multi output)	45.6
Our U-Net (with multi output)	**47.1**
Our U-Net (multi + overlap 8x)	53.6
Our U-Net (multi + overlap 16x)	56.8
Our U-Net (multi + overlap 16x + post)	**60.0**

Results >
(Fixed roads shown in red)