TD-Road: Top-Down Road Network Extraction with Holistic Graph Construction

Yang He, Ravi Garg, Amber Roy Chowdhury
Amazon Last Mile

ECCV TEL AVIV 2022

Motivation

- Road network extraction is a fundamental task
- Many previous methods are not end-to-end
  - Predict intermediate results and convert to road graphs
  - Conversion is not precise in complex roads [a]
  - Many of them are bottom-up based on local prediction [a, e, f]
  - Create a graph in a recursive manner [e, f]

- Relation networks have many applications
  - Few-shot learning
  - Learning non-maximum suppression for object detection

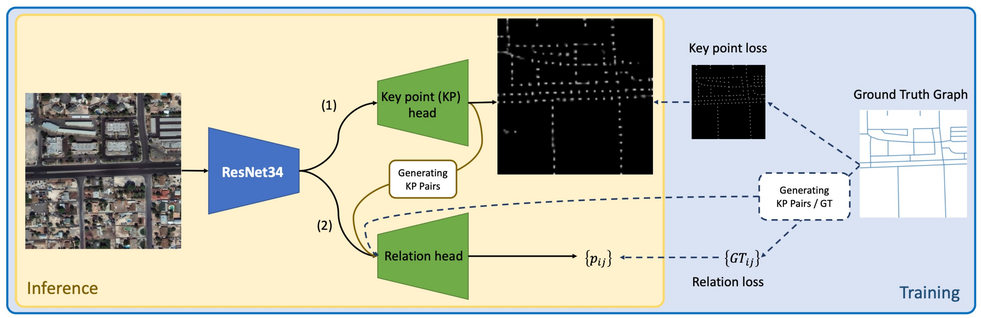
- *Can we train a network to directly output road graphs?*
  - Top-down approach
  - Does not rely on intermediate information prediction
  - Output a graph with one forward
  - Output a graph with neural network

Prior work & Comparison

- Segmentation-based methods
  - DlinkNet [c], DLA [d]
  - Predicting binary masks
  - Leverage segmentation masks as GT
- Graph-based methods
  - Sat2Graph [a], RoadTracer [e], VecRoad [f]
  - Predicting intermediate information
    - Orientations to grow a graph [a, e]
    - Offsets for the next move [f]
  - Require a heuristic processing to convert to final output graphs
  - Leverage orientations/offsets as GT on each node
- Ours
  - *End-to-end model* to produce road graphs
  - *Directly use graph structures* as GT
  - *No post-conversion* required

Method

- Our idea: Road network extraction = Key point prediction + Connectedness prediction



- Key point prediction
  - Binary cross entropy considering segmentation masks

$$\mathcal{L}_{KP} = -\sum_t w_t \cdot (y_t \cdot \log(p_t) + (1 - y_t) \cdot \log(1 - p_t)),\ w_t \in \{w_{kp}, w_r, w_b\}$$

- Connectedness prediction
  - We treat this as a classification problem
  - A relation reasoning module – comparing the key point pairs for all neighbors
  - Neighbor-enhanced relation reasoning

Experimental results

- Datasets
  - City-Scale [a]
  - SpaceNet3 [b]
- Evaluation metrics
  - Topology-based Precision, Recall, F-1
  - APLS (average path length similarity)

Table 1. Comparison results on *City-Scale* dataset.

Method	Backbone	Type	P	Topo R	F-1	APLS
UNet [18]	CNN	Seg.	78.00	57.44	66.16	57.29
DeepRoadMapper [17]	ResNet-50		75.34	65.99	70.36	52.50
Orientation [2]	ResNet-34		75.83	68.90	72.20	55.34
DLinkNet [28]	ResNet-34		78.63	48.07	57.42	54.08
DLA [27,10]	DLA		75.59	72.26	73.89	57.22
RoadTracer [1]	CNN	Graph	78.00	57.44	66.16	57.29
Sat2Graph [10]	DLA		80.70	**72.28**	76.26	63.14
Ours (Naive)	ResNet-34		77.82	68.44	72.54	62.17
Ours (Neighbors)			**81.94**	71.63	**76.27**	**65.74**

Table 2. Comparison results on *SpaceNet* dataset.

Method	Backbone	Type	P	Topo R	F-1	APLS
UNet [18]	CNN	Seg.	68.96	66.32	67.61	53.77
DeepRoadMapper [17]	ResNet-50		82.79	72.56	77.34	62.26
Orientation [2]	ResNet-34		81.56	71.38	76.13	58.82
DLinkNet [28]	ResNet-34		**88.42**	60.06	68.80	56.93
DLA [27,10]	DLA		78.99	69.80	74.11	56.36
RoadTracer [1]	CNN	Graph	78.61	62.45	69.60	56.03
Sat2Graph [10]	DLA		85.93	76.55	80.97	64.43
Ours (Naive)	ResNet-34		82.45	73.54	77.74	60.91
Ours (Neighbors)			84.81	**77.80**	**81.15**	**65.15**

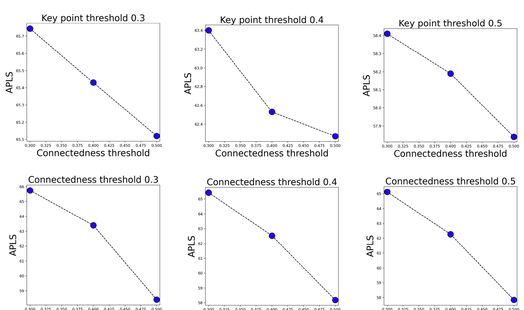


Fig. 8. Results of applying different thresholds on *City-Scale*.

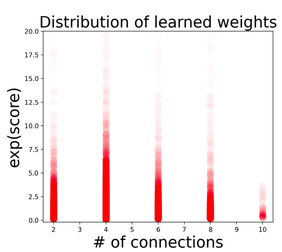


Fig. 9. Scatter plot of learned weights in neighbors-based relation reasoning module.

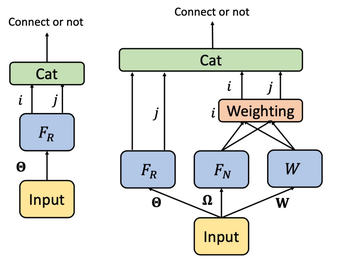


Fig. 4. Illustration of the naive relation reasoning module (left) and the neighbor-enhanced relation module (right).

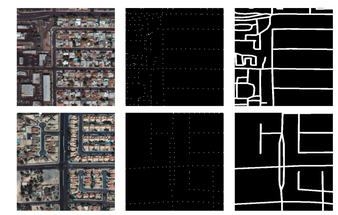
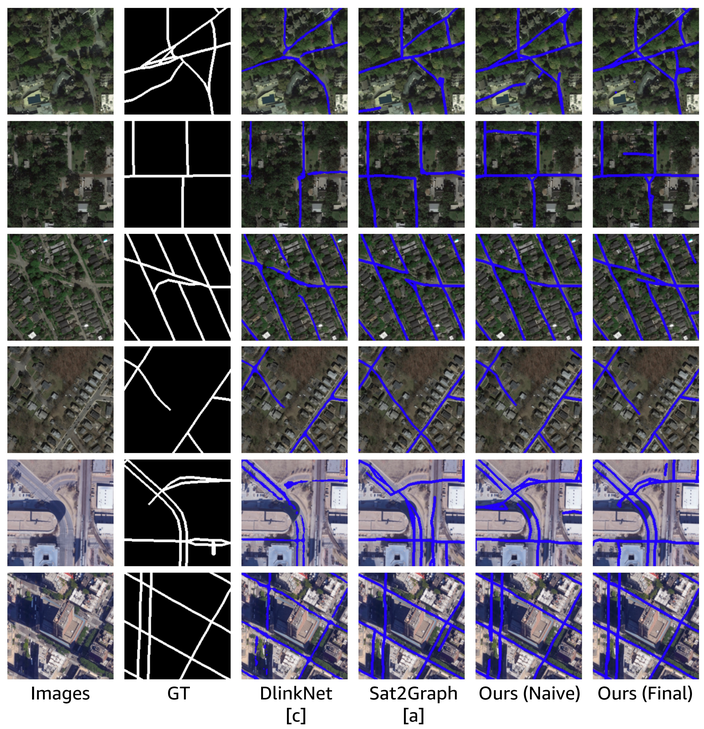


Fig. 5. Ground truth examples in computing loss function. We leverage road mask to help more effective key point prediction.

Fig. 3. Training examples for relation reasoning module. We perform binary classification for each pair between a candidate (blue) and its neighbors (orange). Positive pairs are regarded as the key point pairs, which are connected through a straight line. Otherwise, it will be regarded as negative pairs.



[a] S. He et al., Sat2Graph: Road Graph Extraction through Graph-Tensor Encoding, ECCV, 2022
[b] V. Etten et al., Spacenet: A remote sensing dataset and challenge series, arXiv preprint arXiv:1807.01232, 2018
[c] L. Zhou et al., D-linknet: Linknet with pretrained encoder and dilated convolution for high resolution satellite imagery road extraction, CVPR workshop, 2018
[d] F. Yu et al., Deep layer aggregation, CVPR, 2018
[e] F. Bastani et al., Roadtracer: Automatic extraction of road networks from aerial images, CVPR, 2017
[f] Y.Q. Tan et al., Vecroad: Point-based iterative graph exploration for road graphs extraction, CVPR, 2020