D-LinkNet: LinkNet with Pretrained Encoder and Dilated Convolution for High Resolution Satellite Imagery Road Extraction

Lichen Zhou, Chuang Zhang, Ming Wu, Ruihua Zhang
Beijing University of Posts and Telecommunications

CVPR 2018 SALT LAKE CITY • JUNE 18-22

Objectives

Road Extraction（DeepGlobe Road Extraction Challenge）

- Automatically extracting roads and street networks from satellite images.
- Formulated as a binary segmentation problem to detect all the road pixels in each area.

Task and Approach

- To assign each pixel in satellite images as road or background.
- We introduce a semantic segmentation network, named **D-LinkNet**, for satellite imagery road extraction.

Motivation

Challenge

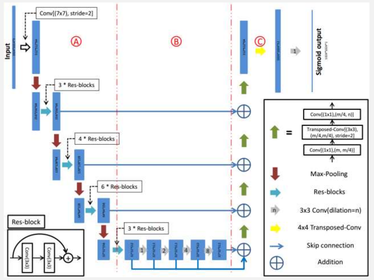
- Roads have natural **connectivity** and **long span**.
- Roads in satellite images are often **slender**, **complex** and cover a small part of the whole image.
- Roads are easily confused with railways and rivers.

What to do

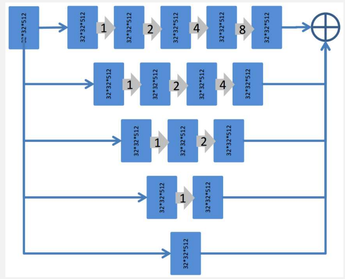
- Design a network with **large receptive field**.
- Preserving the **detailed spatial information**.
- Increasing the network's **recognition accuracy**.

Method

D-LinkNet architecture



D-LinkNet34



Center Dilation

- Encoder: ResNet34/50/101, strong recognition ability.
- Decoder: LinkNet decoder, efficient in computation & memory.
- Dilation: Unrolled as parallel & cascade mode, each path has different receptive field.

D-LinkNet uses Linknet[1] with pretrained encoder as its backbone and has additional stacked dilated convolution layers[2] in the center part. The center dilation part can be unrolled as a combination of cascade mode and parallel mode, enabling the network to combine representations from multi-scale.

Data augmentation

- Color transfer--HSV transfer.
- Spatial transfer—Flip, aspect ratio transfer, scale transfer, shifting.

Test time augmentation

- Vertical/horizontal/diagonal flip. Predict each image 8 times.

Competition Results

Valid-Set	Score
Unet	0.6294
LinkNet34	0.6300
D-LinkNet34	0.6412

Test-Set	Score
D-Unet	0.6194
D-LinkNet34	0.6283
D-LinkNet50	0.6342
D-LinkNet101	0.6237
Final Score	**0.6342**

We won the 1st place in DeepGlobe Road Extraction Challenge.

Conclusion

By enlarging the receptive field and ensembling multi-scale features in the center part while keeping the detailed information at the same time, D-LinkNet can handle roads' properties such as narrowness, connectivity, complexity and long span to some extent.

Code&Paper&PPT

References

[1] Abhishek et al. Linknet: Exploiting encoder representations for efficient semantic segmentation. *arXiv preprint arXiv:1707.03718*, 2017.
[2] Fisher Yu and Vladlen Koltun. Multi-scale context aggregation by dilated convolutions. arXiv preprint arXiv:1511.07122, 2015.