Supplementary Material of Foreground-Aware Relation Network for Geospatial Object Segmentation in High Spatial Resolution Remote Sensing Imagery

Zhuo Zheng   Yanfei Zhong   Junjue Wang   Ailong Ma

Wuhan University, Wuhan, China

{zhengzhuo, zhongyanfei, kingdrone, maailong007}@whu.edu.cn

1. Qualitative analysis

For segmentation task, the visual performance is important. Fig.1 shows the a few visualization results of object segmentation over these methods. We can find that deeper backbone network is significant for object segmentation in the HSR remote sensing imagery. U-Net without deep backbone failed to recognize the roundabout, while other methods with deep backbone were easy to recognize it. FCN-8s with VGG-16 was unable to completely segment the roundabout, while the methods with deeper backbone can recognize it without "hole". Among these methods with deeper backbone, FarSeg has obviously lower false alarm ratio than other methods for roundabout. Such as DenseASPP, Deeplab v3 and PSPNet, their predicted mask of roundabout are inaccurate with many false alarms around the roundabout. Benefited from F-S relation module, FarSeg leverages the latent semantics of this scenario to augment the discrimination of roundabout features, obtaining better results. We also can observe that the soccer ball field is not complete and its proportion is small in this scenario. Only RefineNet and FarSeg can recognize it relatively well. Furthermore, soccer ball field mask predicted by FarSeg has more clear edge, while one from RefineNet is over smooth. It suggests that foreground-aware modeling is important for object segmentation in the HSR remote sensing imagery.
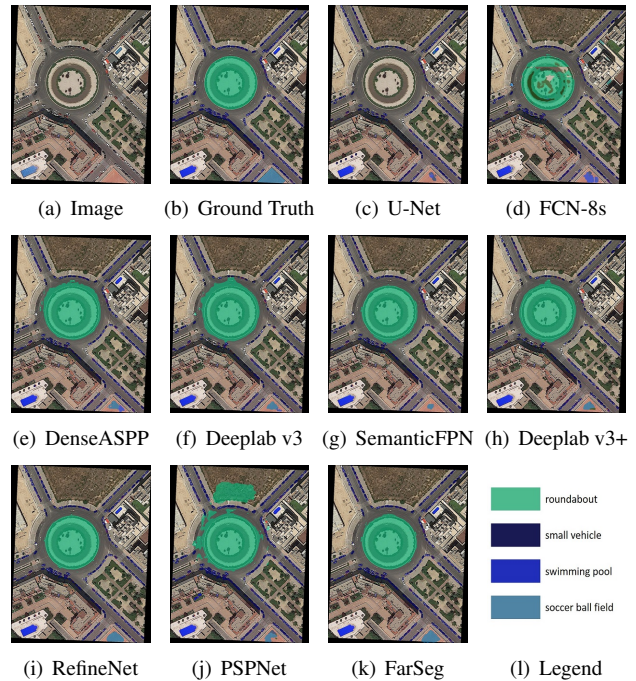

Figure 1. Visualization of object segmentation results on an image in the *val* set of iSAID dataset. (a) original image. (b) Ground truth. (c) U-Net. (d) FCN-8s. (e) DenseASPP. (f) Deeplab v3. (g) Semantic FPN. (h) Deeplab v3+. (i) RefineNet. (j) PSPNet. (k) FarSeg. (l) Legend for categories happened in this image.

2. Visualization

There are many visualization results on large area HSR remote sensing images.

(a) 5028 × 2331 pixels. Legend: harbor, ship, small vehicle, large vehicle
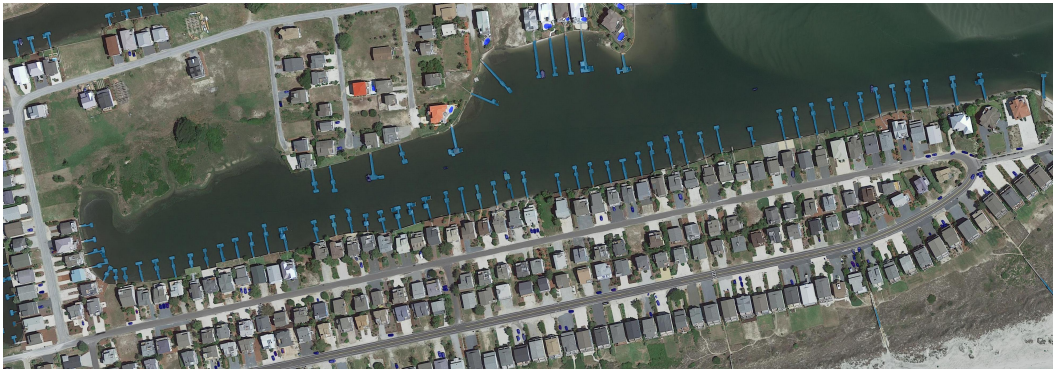
(b) 3964 × 1367 pixels. Legend: harbor, swimming pool, small vehicle

(c) 1818 × 1162 pixels. Legend: roundabout, small vehicle, large vehicle

(d) 1080 × 626 pixels. Legend: tennis court, baseball diamond, small vehicle

(e) 2081 × 1904 pixels. Legend: harbor, ship

(f) 6635 × 6508 pixels. Legend: plane, storage tank, small vehicle, large vehicle

(g) 2086 × 2451 pixels. Legend: small vehicle, large vehicle

Figure 2. Visualization of object segmentation results on large area HSR remote sensing images.