Method Description for the Potsdam 2D Semantic Labeling Contest

Wenli Tang

School of Remote Sensing and Information Engineering, Wuhan University

In this work, we use ResNet101 as our baseline model. To classify small objects such as cars, we employed atrous convolution for dense feature extraction and field-of-view enlargement, which is inspired by DeepLab.

During training, we only use the IRRG data of Potsdam offered by the organizer, no any other data. We perform multi-scale and flipped patches as data augmentation to get more information from the original images. Also, we use "poly" learning strategy for training rather than fixed learning rate. Different from the former method, we make full use of GPU by maximum the image patch(720x720) on *Nvidia TITAN 1080Ti* during inference, and use IRRG data instead of RGB data.

Reference:

Chen L C, Papandreou G, Kokkinos I, et al. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs.[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2016, PP(99):1-1.

Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]// Computer Vision and Pattern Recognition. IEEE, 2015:3431-3440.