Method Description for Vaihingen:
2D Labeling Challenge
Zhao Xinwei
Institution of Software, Chinese Academy of Science
May, 2018

This model is based on Deeplab v3+, in which Atrous Spatial Pyramid Pooling (ASPP) and global average pooling are combined as the decoder. This model differs from Deeplab v3+ in two aspects: 1) the rates of the atrous convolution in a single block in the encoder are fixed. 2) res101 is taken as the encoder instead of Xception.

This model was trained on 12800 image patches in total, which are acquired by randomly cropping 800 600x600 patches from each of the 16 original training images. Other data augmentation subroutines are carried out as follows:

1. Flip the image horizontally with the probability of 0.5.
2. Scale the image with a factor of 0.5 to 2.0.

This model was trained for 26 epochs with a batch size of 8 using SGD+Momentum, and the learning rate is controlled using polynomial learning rate decay policy with power set to 0.9. Initial learning rate, end learning rate and momentum are set to 7e-3, 1e-6, 0.9 respectively.