

Lewis Fishgold
Azavea

Our entry uses a simplified version of U-Net which is an encoder-decoder architecture for semantic segmentation. The model was trained from scratch without the use of pretraining, and was not combined with any other models as part of an ensemble. We trained the model for 100 epochs with 4096 samples per epoch with a batch size of 8 using the Adam optimizer. However, this was excessive as the validation loss stabilized after around 40 epochs. The learning rate started at 0.001 and was divided by 10 after the validation loss did not improve for ten epochs. The training process took ~12 hrs on an NVIDIA Tesla K80 GPU. The input to the network included red, green, infrared, NDVI, and depth channels which were concatenated together. The training data was randomly augmented using 90 degree rotations and horizontal and vertical flips.

The code for the project is open source and is located at:
<https://github.com/azavea/keras-semantic-segmentation>.

The experiment configuration file used to generate these results is located at:
https://github.com/azavea/keras-semantic-segmentation/blob/develop/src/experiments/3_28_17/vaihingen_unet_all.json