Technical Report

Using Deep Fully Convolutional Neural Network To Perform Semantic Aerial Image Segmentation, Applied For Potsdam Dataset

Trinh Van Duy [1]
Dinh Viet Sang[1]
[1] Hanoi University of Science and Technology

I use Fully Convolutional Neural Network as describe in the paper [1] to tackle this problem. For faster training, I use VGG pretrained model to fine tune. I use 20 in 24 tiles for training and 4 remainings to validate the model. For each tiles, I crop randomly 4000 images with size (224,224). For testing phase, I perform sliding window and use the model to infer each small (224,224) patch with overlap 112 and then sew it to reform the original patch. No post processing is performed. It takes 36 hours for my model to converge, and about 3 minute to infer to giant patch image on Tesla K20c GPU. Result for validation phase is 87.2375 % overall for 4 giant validate image patches.

This time, I use all 6 channels include IRRGB, DSM, and nDSM to train the model and ensemble models with number of models is 5.

Reference:
[1] Jonathan Long, Evan Shelhamer, Trevor Darrell: Fully Convolutional Networks for Semantic Segmentation. CVPR 2015.