Hyper-Spectral Semantic Segmentation in Varying Satellite Resolutions

Kenan Farmer
Adviser: Adam Finkelstein
Princeton University Department of Computer Science

Introduction

The field of Earth Observation has exploded in the past couple years. The amount of publicly available Earth Observation data has been reaching the Zettabyte scale[1]. Consequently the field has turned to deep learning and computer vision to help aid the analysis of satellite imagery.
This field has grown due to lowered cost of obtaining high resolution imagery and the wide applications automated Earth Observation would have. An incomplete list of applications includes agriculture, mineralogy, surveillance, physics, astronomy, chemical imaging, and environmental sciences[8].

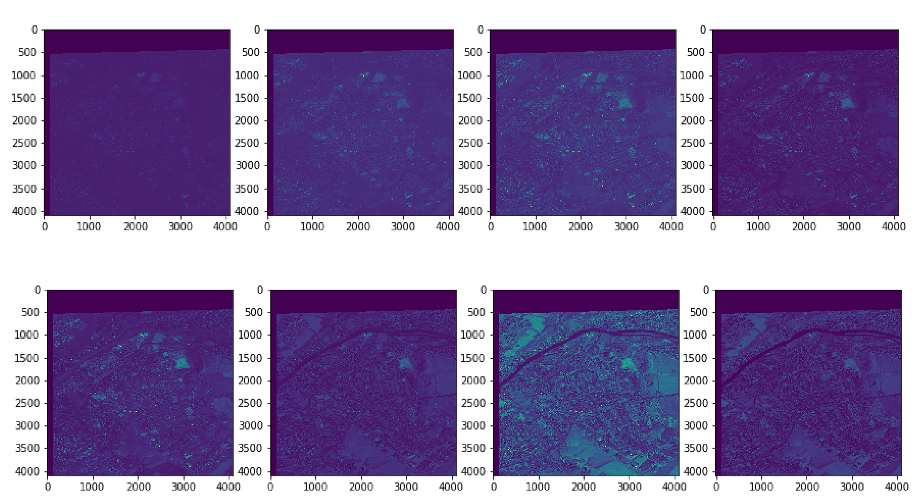

Figure 1: a region of Northern France in 8 different spectrum bands

Problem

We will attempt the direct problem of semantic segmentation for building footprints in hyperspectral satellite images using existing deep learning research. We will evaluate our model using the above task at various satellite resolutions.

Previous Work

Semantic Segmentation for Remote Sensing has explored architectures yielding FCN's as the best family of network architectures [1, 4]. These works have not evaluated performance on varying resolution sizes however. Our models SegNet and UNet are two strong candidates for this project [2, 5]

Data

We draw upon the following datasets for our evaluation of performance in varying resolutions.

- Spacenet WorldView-3 images (.31m)
- Sentinel-2 Images (10-20m)
- Landsat-8 Images (30m)

These are collectively provided in AWS but have different image sizes which we normalize with a **simple sliding window** of geographic area.
Data was fetched with multi-threaded python and cleaned with utilities provided by Spacenet Challenge creators.

Methods

To evaluate our model we will train on high-quality images and then test on differing resolutions to see how our model generalizes. Our evaluation metric will be the jaccard index defined as

$$Jacc(A, B) = \frac{A \cap B}{A \cup B}$$

In the context of polygon sets, this translates to overlapping areas over union of the two set's areas. This metric is easily interpretable from $[0, 1]$ and allows for correct evaluation of segmentation.

Results

Model	Spacenet	Sentinel-2	Landsat-8
U-Net	.78	.21	-
SegNet Basic	.88	.62	.58
SegNet Extended + Dropout	.91	.68	.62

Table 1: Results of each model trained with a subset of Spacenet and evaluated on each dataset

Results were achieved on a p2.xlarge GPU accelerated server on Amazon Web Services

Model Architecture

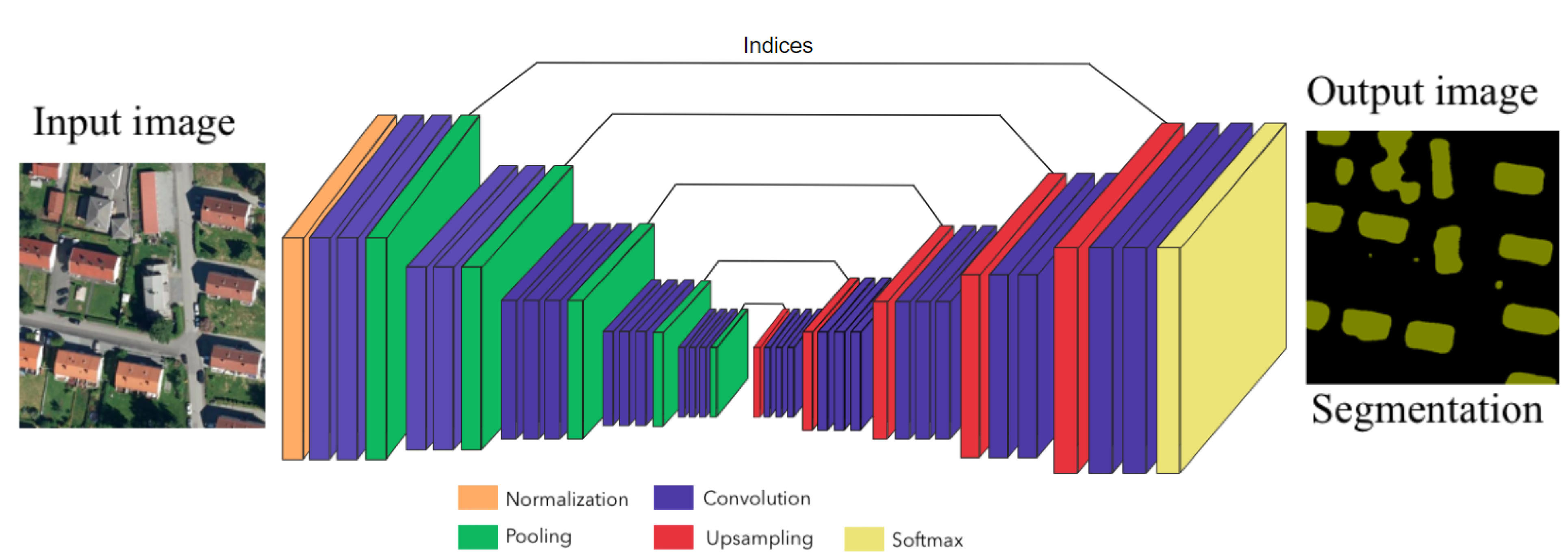

Figure 2: This is the basic SegNet Architecture

SegNet has an encoder-decoder structure with 5 encoders and decoders for the extended model and dropout layers immediately after pooling layers for each encoder.

Conclusion

SegNet with Extended layers and Dropout performs well in high resolution and moderate in lower resolutions

- quarter of performance dropped when resolution changes from $0.31m$ to $10 - 20m$
- U-net is more sensitive to resolution

Future Work

Future Work can be broken into two parts

1. **Extension beyond buildings** - Roads and multi-class segmentation may effect resolution sensitivity
2. **Experimentation with new Deep Learning research** - specifically Capsule Networks implications

References

[1] N. Audebert, B. Le Saux, and S. Lefèvre, "Semantic Segmentation of Earth Observation Data Using Multimodal and Multi-scale Deep Networks," *ArXiv e-prints*, Sep. 2016.

[2] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation," *ArXiv e-prints*, Nov. 2015.

[3] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs," *ArXiv e-prints*, Jun. 2016.

[4] D. Marmanis, J. D. Wegner, S. Galliani, K. Schindler, M. Datcu, and U. Stilla, "Semantic segmentation of aerial images with an ensemble of cnns," 2016.

[5] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *ArXiv e-prints*, May 2015.

[6] S. Sabour, N. Frosst, and G. E Hinton, "Dynamic Routing Between Capsules," *ArXiv e-prints*, Oct. 2017.

[7] A. Sharma, X. Liu, and X. Yang, "Land Cover Classification from Multi-temporal, Multi-spectral Remotely Sensed Imagery using Patch-Based Recurrent Neural Networks," *ArXiv e-prints*, Aug. 2017.

[8] X. X. Zhu, D. Tuia, L. Mou, G.-S. Xia, L. Zhang, F. Xu, and F. Fraundorfer, "Deep learning in remote sensing: a review," *ArXiv e-prints*, Oct. 2017.