Engineering, Economics, and Regulation for Energy Access in Developing Countries

15.017/6.934J

Session 13. Monday October 22, 2018

Machine Learning to Inform Large Scale Planning

Stephen J. Lee
leesj@mit.edu
Contents

1. GIS-Based Planning Comments
2. Satellite Imagery Overview
3. Automatic Infrastructure Mapping
   • Intro to deep learning
   • Building extraction
   • Load localization
4. Electrification Status Estimation
5. Demand Characterization
   • REM Studies
6. Synergies with Other Sectors
7. Adaptive Planning
GIS-based vs Traditional Planning

- Relative to traditional static electrification master planning, GIS-based approaches are easier to analyze, adapt, visualize, and share.

- Traditional plans can take two to three years and cost > $2 MM to prepare. In contrast, GIS-based plans for Rwanda and Kenya each cost about $1 MM and took one year.

Reference Electrification Model

Large-scale electricity infrastructure planning for:

1. Government and distribution company planners - comprehensive planning
2. Electricity regulators and policymakers - setting tariffs allocating subsidies
3. Commercial firms, including developers, - seeking to compete for construction tenders

http://universalaccess.mit.edu/rem-web-demo/
Data Requirements

Building Extraction

Load Localization and Characterization

Electrification Status Estimation and Demand Characterization

Reference Electrification Model Core (clustering, RNM, etc.)

Visualization
Contents

1. GIS-Based Planning Comments
2. Satellite Imagery Overview
3. Automatic Infrastructure Mapping
   • Intro to deep learning
   • Building extraction
   • Load localization
4. Electrification Status Estimation
5. Demand Characterization
   • REM Studies
6. Synergies with Other Sectors
7. Adaptive Planning
Imagery Types

**Tradeoff**
- Spatial resolution v.s. temporal resolution + coverage

**Government (Free)**

**Landsat**
- Since 1972, 30m res, 2 weeks

**Sentinel**
- Radar, 30-120m res, radar, 3 days

**MODIS**
- Since 1999, daily, 36 bands, 1-2 day
  - 250 m (bands 1-2)
  - 500 m (bands 3-7)
  - 1000 m (bands 8-36)

**NAIP**
- Since 2003, continental US only, 1m res, annually

**DMSP-OLS/VIIRS**
- Nighttime lights, Since 1992, whole world, 1 km, nightly

**Shuttle Radar Topography Mission (SRTM)**
- 30 m res

**Commercial**

**DigitalGlobe:** [http://www.digitalglobe.com](http://www.digitalglobe.com)
- Sub-meter imagery back to 1999
- Down to 30 cm

**Planet:** [https://www.planet.com/](https://www.planet.com/)
- 3-5m resolution, 175+ satellites in orbit
- Imaging entirety of Earth’s landmass every single day
- Acquired Terra Bella from Google to compete with high res

**Airbus:** [http://www.intelligence-airbusds.com/](http://www.intelligence-airbusds.com/)

**Satellogic:** [https://www.satellogic.com/](https://www.satellogic.com/)
Satellite Imagery

Satellite Imagery

https://spacenews.com/39883planet-labs-secures-funding-to-launch-another-72-satellites/
Contents

1. GIS-Based Planning Comments
2. Satellite Imagery Overview
3. Automatic Infrastructure Mapping
   • Intro to deep learning
   • Building extraction
   • Load localization
4. Electrification Status Estimation
5. Demand Characterization
   • REM Studies
6. Synergies with Other Sectors
7. Adaptive Planning
General classification problems

input

images/video

audio

text

output

Label: “Motorcycle”
Suggest tags
Image search
...

Speech recognition
Music classification
Speaker identification
...

Web search
Anti-spam
Machine translation
...

Source: MIT 6.819/6.869 Advances in Computer Vision, Fall 2015
Why is this hard?

You see this:

But the camera sees this:

<p>| | | | | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>194</td>
<td>210</td>
<td>201</td>
<td>212</td>
<td>199</td>
<td>213</td>
<td>215</td>
<td>195</td>
<td>178</td>
<td>158</td>
<td>182</td>
<td>209</td>
<td></td>
</tr>
<tr>
<td>180</td>
<td>189</td>
<td>190</td>
<td>221</td>
<td>209</td>
<td>205</td>
<td>191</td>
<td>167</td>
<td>147</td>
<td>115</td>
<td>129</td>
<td>163</td>
<td></td>
</tr>
<tr>
<td>114</td>
<td>126</td>
<td>140</td>
<td>188</td>
<td>176</td>
<td>165</td>
<td>152</td>
<td>140</td>
<td>170</td>
<td>106</td>
<td>78</td>
<td>88</td>
<td></td>
</tr>
<tr>
<td>87</td>
<td>103</td>
<td>115</td>
<td>154</td>
<td>143</td>
<td>142</td>
<td>149</td>
<td>153</td>
<td>173</td>
<td>101</td>
<td>57</td>
<td>57</td>
<td></td>
</tr>
<tr>
<td>102</td>
<td>112</td>
<td>106</td>
<td>131</td>
<td>122</td>
<td>138</td>
<td>152</td>
<td>147</td>
<td>128</td>
<td>84</td>
<td>58</td>
<td>66</td>
<td></td>
</tr>
<tr>
<td>94</td>
<td>95</td>
<td>79</td>
<td>104</td>
<td>105</td>
<td>124</td>
<td>129</td>
<td>113</td>
<td>107</td>
<td>87</td>
<td>69</td>
<td>67</td>
<td></td>
</tr>
<tr>
<td>68</td>
<td>71</td>
<td>69</td>
<td>98</td>
<td>89</td>
<td>92</td>
<td>98</td>
<td>95</td>
<td>89</td>
<td>88</td>
<td>76</td>
<td>67</td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>56</td>
<td>68</td>
<td>99</td>
<td>63</td>
<td>45</td>
<td>60</td>
<td>82</td>
<td>58</td>
<td>76</td>
<td>75</td>
<td>65</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>43</td>
<td>69</td>
<td>75</td>
<td>56</td>
<td>41</td>
<td>51</td>
<td>73</td>
<td>55</td>
<td>70</td>
<td>63</td>
<td><strong>44</strong></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>50</td>
<td>57</td>
<td>69</td>
<td>75</td>
<td>75</td>
<td>73</td>
<td>74</td>
<td>53</td>
<td>68</td>
<td>59</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td>72</td>
<td>59</td>
<td>53</td>
<td>66</td>
<td>84</td>
<td>92</td>
<td>84</td>
<td>74</td>
<td>57</td>
<td>72</td>
<td>63</td>
<td>42</td>
<td></td>
</tr>
<tr>
<td>67</td>
<td>61</td>
<td>58</td>
<td>65</td>
<td>75</td>
<td>78</td>
<td>76</td>
<td>73</td>
<td>59</td>
<td>75</td>
<td>69</td>
<td>50</td>
<td></td>
</tr>
</tbody>
</table>
Pixel-based representation

Learning algorithm

Source: MIT 6.819/6.869 Advances in Computer Vision, Fall 2015
What we want

Input

Raw image

Feature representation

E.g., Does it have Handlebars? Wheels?

Motorbikes

“Non”-Motorbikes

Learning algorithm

Sources:

- MIT 6.819/6.869 Advances in Computer Vision, Fall 2015
ImageNet Challenge

- 1000-way image classification
- ImageNet Large Scale Visual Recognition Challenge (ILSVRC) held since 2010
- Organized by the Stanford Computer Vision Lab
Object recognition

Source: MIT 6.819/6.869 Advances in Computer Vision, Fall 2015
The brain: a potential motivation for deep learning

Auditory cortex learns to see!

Source: MIT 6.819/6.869 Advances in Computer Vision, Fall 2015
The brain adapts!

Seeing with your tongue

Human echolocation (sonar)

Haptic belt: Direction sense

Implanting a 3rd eye

[BrainPort; Welsh & Blasch, 1997; Nagel et al., 2005; Constantine-Paton & Law, 2009]
Neural networks
Activation functions

**Sigmoid activation function**

\[
\frac{1}{1 + e^{-x}}
\]

**tanh(x)**

\[f(x) = \max(0, x)\]

**ReLU**
Neural networks: architectures

“2-layer neural net,” or “1-hidden-layer neural net”

“3-layer neural net,” or “2-hidden-layer neural net”
Convolutional neural networks aka ConvNets aka CNNs aka Computer vision Savior
ConvNets: architecture

Convolutional Neural Networks are just Neural Networks BUT:
1. Local connectivity

- a hidden neuron in next layer

image: 32x32x3 volume
before: full connectivity: 32x32x3 weights

now: one neuron will connect to, e.g. 5x5x3 chunk and only have 5x5x3 weights.

note that connectivity is:
- local in space (5x5 inside 32x32)
- but full in depth (all 3 depth channels)
ConvNets: architecture

Convolutional Neural Networks are just Neural Networks BUT:

1. Local connectivity

Multiple neurons all looking at the same region of the input volume, stacked along depth.
ConvNets: architecture

Convolutional Neural Networks are just Neural Networks BUT:

1. Local connectivity

These form a single [1 x 1 x depth] “depth column” in the output volume
Fast-forward to today

[From recent Yann LeCun slides]

Low-Level Feature → Mid-Level Feature → High-Level Feature → Trainable Classifier

Feature visualization of convolutional net trained on ImageNet from [Zeiler & Fergus 2013]
Going deep

Training set: Aligned images of faces.

object models

object parts (combination of edges)

edges

pixels
ConvNets perform Classification

< 1

1000-dim vector

"tabby cat"

end-to-end learning

Source: MIT 6.819/6.869 Advances in Computer Vision, Fall 2015

Credit: Long et al
< 1/5 second

end-to-end learning

Source: MIT 6.819/6.869 Advances in Computer Vision, Fall 2015

Credit: Long et al
End-to-end, Pixels-to-pixels network

convolution

conv, pool, nonlinearity

upsampling

pixelwise output + loss

Source: MIT 6.819/6.869 Advances in Computer Vision, Fall 2015

Credit: Long et al
Long and Shelhamer et al., Fully Convolutional Neural Networks for Semantic Segmentation

We are adopting an implementation of this model

Source: Long and Shelhamer et al. “Fully Convolutional Networks for Semantic Segmentation”
Contents

1. GIS-Based Planning Comments
2. Satellite Imagery Overview
3. Automatic Infrastructure Mapping
   • Intro to deep learning
   • Building extraction
   • Load localization
4. Electrification Status Estimation
5. Demand Characterization
   • REM Studies
6. Synergies with Other Sectors
7. Adaptive Planning
Our manual annotation tool is compatible with Amazon Mechanical Turk and can also work as a standalone web application.

- Based off of Prof. Jianxiong Xiao’s DrawMe project (Princeton, MIT)
Labeled ground truth (Uganda SST, Google Maps API)
Building Extraction (Uganda SST, Google Maps API)
Inferences and Building Localization
Dr. Yuan, Oak Ridge National Lab

- 30 cm resolution images (much higher quality)
- Not developing countries
- Model outlines probable borders on houses
- Nice way to visualize results

Commoditization of building extraction: Google Earth Engine

Source: Prof. Roy’s Group, Media Lab
Commercialization of building extraction: Ecopia, DigitalGlobe Data
Commoditization of building extraction: Facebook, Columbia, DigitalGlobe Data

Connectivity Lab population mapping
Posted by Facebook Engineering
686 Views

Facebook AI enhanced

DigitalGlobe satellite image of Naivasha, Kenya (left) and results of our analysis of the same area (right).
Jean et al.’s Convolutional Neural Networks for Predicting Poverty

Jean et al.’s Convolutional Neural Networks for Predicting Poverty (Cont.)

LandScan Data

-LandScan data obtained, Oak Ridge National Lab
  -Community standard for global population distribution.
  -Approximately 1 km resolution (30" X 30")
  -Represents ambient population (average over 24 hours)
  -Uses spatial data and imagery analysis technologies and a multi-variable dasymetric modeling approach to disaggregate census counts within an administrative boundary
  -Currently performing correlation studies for validation
Google Maps imagery in Uganda is limited and of low quality

- This is a major risk for Uganda SST master planning study

Higher frequency of:

- Clouds
- Color changes and poor contrast
- Missing data (~25%)

Some areas are still good (~50%)
We identify more rocks and fewer houses in Uganda
Missing Google Maps Data at Zoom 18
(60 cm resolution)
DigitalGlobe GBDX Partnership

- Largest catalog of imagery available for the globe
- Started collecting in 1992
- Operator of civilian remote sensing spacecraft
- Changing business model for monetizing their imagery
Contents

1. GIS-Based Planning Comments
2. Satellite Imagery Overview
3. Automatic Infrastructure Mapping
   • Intro to deep learning
   • Building extraction
   • Load localization
4. Electrification Status Estimation
5. Demand Characterization
   • REM Studies
6. Synergies with Other Sectors
7. Adaptive Planning Planning
Inferences and Building Localization
Load Localization and Classification

Random Pixel Sampling without replacement

Claudio’s algorithm using dilations. “Taking bites out of footprint masks”
Contents

1. GIS-Based Planning Comments
2. Satellite Imagery Overview
3. Automatic Infrastructure Mapping
   • Intro to deep learning
   • Building extraction
   • Load localization
4. Electrification Status Estimation
5. Demand Characterization
   • REM Studies
6. Synergies with Other Sectors
7. Adaptive Planning
Coarse electrification Inference by Doll et al.

- Uses nighttime satellite imagery and the Global Rural Urban Mapping Project (GRUMP) gridded population dataset

- These were then compared against statistics on electricity access compiled by the IEA at the country level

- Shows that population density information is very important

---

C. N. Doll, S. Pachauri, Estimating rural populations without access to electricity in developing countries through night-time light satellite imagery. In Energy Policy, pages 5661-5670. 2010
Min et al.’s logistic regressions using nighttime lights and census exclusively

- Compares night-time light output from the DMSP-OLS against ground-based survey data on electricity use in 232 electrified villages and additional administrative data on 899 unelectrified villages in Senegal and Mali.
Estimating Electrification Status

Population density

Existing grid and transformer locations

Night time lights data

55
Estimating Electrification Status, Vaishali District, India

2011 Census Data at the Village-Level with:

• Power Supply For Domestic Use (Status A(1)/NA(2))
• Power Supply For Domestic Use Summer (April-Sept.) per day (in Hours)
• Power Supply For Domestic Use Winter (Oct.-March) per day (in Hours)
• Power Supply For Agriculture Use (Status A(1)/NA(2))
• Power Supply For Agriculture Use Summer (April-Sept.) per day (in Hours)
• Power Supply For Agriculture Use Winter (Oct.-March) per day (in Hours)
• Power Supply For Commercial Use (Status A(1)/NA(2))
• Power Supply For Commercial Use Summer (April-Sept.) per day (in Hours)
• Power Supply For Commercial Use Winter (Oct.-March) per day (in Hours)
• Power Supply For All Users (Status A(1)/NA(2))
• Power Supply For All Users Summer (April-Sept.) per day (in Hours)
• Power Supply For All Users Winter (Oct.-March) per day (in Hours)
2012 building-level and 2014 sub-county-level data for all of Uganda. 2016 building-level data for the Uganda SST

- 241,038 km²
- >10,000 individual houses, businesses, health centers, and schools sampled in 2012
- Sub-county level information for the full 2014 census

- 10,914 km²
  (Massachusetts is 27,336 km²)
- ~500 households surveyed
Nighttime Lights
Building Locations
Survey data and district maps

Lee, S., Adaptive Electricity Access Planning, 2018
(Incomplete) Transformer locations
Preliminary model: Gaussian process classification using

Likelihood: \( \sigma_{probit}(t) := \int_{-\infty}^{t} \mathcal{N}(\tau|0, 1) d\tau \)

Mean function: \( m(x) = c \)

Covariance function: Squared exponential with automatic relevance determination \( k(x, z) = \sigma_f^2 \exp(-1/2(x - z)^T \Lambda^{-2}(x - z)) \) specifying hyperparameters \( \psi \lambda_1, \lambda_2, \) and \( \sigma_f \).

Fully Independent Training Conditional (FITC) approximation for approximating the covariance function. Uses a low-rank diagonal matrix \( \tilde{K} = Q + \text{diag}(K - Q) \) where \( Q = K_u^T K_{uu}^{-1} K_u \), instead of the exact covariance matrix \( K \). \( K_u \) and \( K_{uu} \) contain covariances and cross-covariances for and between inducing points \( u^i \) and data points \( x^j \).

Laplace Approximation (LA) is used. LA approximates the posterior by a Gaussian centered at the mode with an appropriate curvature.
GP Classification on 2012 ERT Uganda Electrification Survey Data
with Mean Constant = -3.981, Longitude Length Scale = 0.016,
Latitude Length Scale = 0.019, and Signal Standard Deviation = 3.480
Using the Laplace Approximation

Lee, S., Adaptive Electricity Access Planning, 2018
GP Classification on 2016 GIZ Uganda Electrification Survey Data
with Mean Constant = -0.734, Longitude Length Scale = 0.088,
Latitude Length Scale = 0.075, and Signal Standard Deviation = 2.096
Using the Laplace Approximation
Electrification Status Estimates

region-level properties

region-level observations

cell-level properties

cell-level observations

Lee, S., Adaptive Electricity Access Planning, 2018
Electrification Status Estimates
Estimating Electrification Uncertainty (Entropy)
Contents

1. GIS-Based Planning Comments
2. Satellite Imagery Overview
3. Automatic Infrastructure Mapping
   • Intro to deep learning
   • Building extraction
   • Load localization
4. Electrification Status Estimation
5. Demand Characterization
   • REM Studies
6. Synergies with Other Sectors
7. Adaptive Planning
C&I vs Residential Consumption for African Countries
Consumption growth in Kenya

REM results for different demand assumptions

<table>
<thead>
<tr>
<th>Demand (GWh/yr)</th>
<th>Low case</th>
<th>Central case</th>
<th>High case</th>
</tr>
</thead>
<tbody>
<tr>
<td>103</td>
<td>$35.9 million</td>
<td>$0.37/kWh</td>
<td>$0.18/kWh</td>
</tr>
<tr>
<td>369</td>
<td>$63.6 million</td>
<td>$0.13/kWh</td>
<td>$109.3 million</td>
</tr>
<tr>
<td>915</td>
<td>$0.18/kWh</td>
<td>$0.37/kWh</td>
<td>$109.3 million</td>
</tr>
</tbody>
</table>

Index (1 = central demand case)

Total annual cost ($ million)
Cost per unit generated ($/kWh)
Grid Share
Contents

1. GIS-Based Planning Comments
2. Satellite Imagery Overview
3. Automatic Infrastructure Mapping
   • Intro to deep learning
   • Building extraction
   • Load localization
4. Electrification Status Estimation
5. Demand Characterization
   • REM Studies
6. Synergies with Other Sectors
7. Adaptive Planning Planning
Coordinated clean cooking and electrification planning

Residential Electric Cookstove Penetration

IEA – WEO 2018, Lee et al contributions
Prospects for coordinated electrification and agricultural sector planning
Contents

1. GIS-Based Planning Comments
2. Satellite Imagery Overview
3. Automatic Infrastructure Mapping
   • Intro to deep learning
   • Building extraction
   • Load localization
4. Electrification Status Estimation
5. Demand Characterization
   • REM Studies
6. Synergies with Other Sectors
7. Adaptive Planning
Open loop vs. closed loop approaches
Decisionmaking under uncertainty
Thanks! Questions?
Addendum
<table>
<thead>
<tr>
<th>Survey Year</th>
<th>Data Set</th>
<th>Method</th>
<th>GP Weight</th>
<th>Acc.</th>
<th>Prec.</th>
<th>Rec</th>
<th>F1</th>
<th>Obj. Score</th>
<th>Val.</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012</td>
<td>Train</td>
<td>All Neg.</td>
<td>0</td>
<td>87.65</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>2012</td>
<td>Train</td>
<td>LR (Unreg.)</td>
<td>0</td>
<td>89.21</td>
<td>0.63</td>
<td>0.3</td>
<td>0.41</td>
<td>1778.64</td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td>Train</td>
<td>Gaus. Proc.</td>
<td>1</td>
<td>93.31</td>
<td>0.75</td>
<td>0.69</td>
<td>0.72</td>
<td>1484.85</td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td>Train</td>
<td>Comb. By Acc.</td>
<td>0.51</td>
<td>90.32</td>
<td>0.69</td>
<td>0.4</td>
<td>0.5</td>
<td>1465.42</td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td>Train</td>
<td>Comb. By F1.</td>
<td>0.64</td>
<td>91.69</td>
<td>0.74</td>
<td>0.51</td>
<td>0.6</td>
<td>1456.61</td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td>Train</td>
<td>Comb. By Opt.</td>
<td>0.68</td>
<td>91.92</td>
<td>0.74</td>
<td>0.53</td>
<td>0.62</td>
<td>1456.11</td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td>Test</td>
<td>All Neg.</td>
<td>0</td>
<td>87.64</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>2012</td>
<td>Test</td>
<td>LR (Unreg.)</td>
<td>0</td>
<td>88.99</td>
<td>0.63</td>
<td>0.26</td>
<td>0.37</td>
<td>796.65</td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td>Test</td>
<td>Gaus. Proc.</td>
<td>1</td>
<td>88.47</td>
<td>0.53</td>
<td>0.54</td>
<td>0.54</td>
<td>804.18</td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td>Test</td>
<td>Comb. By Acc.</td>
<td>0.51</td>
<td>89.06</td>
<td>0.61</td>
<td>0.33</td>
<td>0.42</td>
<td>719.57</td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td>Test</td>
<td>Comb. By F1.</td>
<td>0.64</td>
<td>88.99</td>
<td>0.58</td>
<td>0.39</td>
<td>0.47</td>
<td>732.68</td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td>Test</td>
<td>Comb. By Opt.</td>
<td>0.68</td>
<td>88.91</td>
<td>0.57</td>
<td>0.4</td>
<td>0.47</td>
<td>738.02</td>
<td></td>
</tr>
<tr>
<td>2016</td>
<td>Train</td>
<td>All Neg.</td>
<td>0</td>
<td>55.76</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>2016</td>
<td>Train</td>
<td>LR (Unreg.)</td>
<td>0</td>
<td>57.58</td>
<td>1</td>
<td>0.04</td>
<td>0.08</td>
<td>226.08</td>
<td></td>
</tr>
<tr>
<td>2016</td>
<td>Train</td>
<td>Gaus. Proc.</td>
<td>1</td>
<td>88.48</td>
<td>0.88</td>
<td>0.86</td>
<td>0.87</td>
<td>99.85</td>
<td></td>
</tr>
<tr>
<td>2016</td>
<td>Train</td>
<td>Comb. By Acc.</td>
<td>0.61</td>
<td>88.48</td>
<td>0.91</td>
<td>0.82</td>
<td>0.86</td>
<td>139.63</td>
<td></td>
</tr>
<tr>
<td>2016</td>
<td>Train</td>
<td>Comb. By F1.</td>
<td>0.92</td>
<td>88.48</td>
<td>0.89</td>
<td>0.84</td>
<td>0.87</td>
<td>107.38</td>
<td></td>
</tr>
<tr>
<td>2016</td>
<td>Train</td>
<td>Comb. By Opt.</td>
<td>1.42</td>
<td>88.48</td>
<td>0.88</td>
<td>0.86</td>
<td>0.87</td>
<td>71.41</td>
<td></td>
</tr>
<tr>
<td>2016</td>
<td>Test</td>
<td>All Neg.</td>
<td>0</td>
<td>55.63</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>2016</td>
<td>Test</td>
<td>LR (Unreg.)</td>
<td>0</td>
<td>57.04</td>
<td>1</td>
<td>0.03</td>
<td>0.06</td>
<td>97.17</td>
<td></td>
</tr>
<tr>
<td>2016</td>
<td>Test</td>
<td>Gaus. Proc.</td>
<td>1</td>
<td>85.92</td>
<td>0.84</td>
<td>0.84</td>
<td>0.84</td>
<td>51.95</td>
<td></td>
</tr>
<tr>
<td>2016</td>
<td>Test</td>
<td>Comb. By Acc.</td>
<td>0.61</td>
<td>82.39</td>
<td>0.85</td>
<td>0.73</td>
<td>0.79</td>
<td>65.4</td>
<td></td>
</tr>
<tr>
<td>2016</td>
<td>Test</td>
<td>Comb. By F1.</td>
<td>0.92</td>
<td>85.92</td>
<td>0.85</td>
<td>0.83</td>
<td>0.84</td>
<td>54.33</td>
<td></td>
</tr>
<tr>
<td>2016</td>
<td>Test</td>
<td>Comb. By Opt.</td>
<td>1.42</td>
<td>85.92</td>
<td>0.84</td>
<td>0.84</td>
<td>0.84</td>
<td>47.35</td>
<td></td>
</tr>
</tbody>
</table>
WV2 vs Google Maps

Learning curves shown for FCNs fine-tuned with (a) the WorldView-2 images with flip, (b) the WorldView-2 images alone, and (c) the Google Maps images. *Please note that (c) was run using incorrect accuracy formula parameters. Similar to (a) and (b) it actually maintained >99% pixel accuracy throughout fine-tuning as determined by later testing.*
Error Metrics for Semantic Segmentation:

Pixel Accuracy: \( \sum_i \frac{n_{ii}}{\sum_i t_i} \)

Mean Accuracy: \( \frac{1}{n_{cl}} \sum_i \frac{n_{ii}}{t_i} \)

Intersection over Union (IU): \( \sum_i \frac{n_{ii}}{t_i + \sum_j (n_{ji} - n_{ii})} = \frac{TP}{TP + FP + FN} \)

Mean IU:
\( \frac{1}{n_{cl}} \sum_i \frac{n_{ii}}{t_i + \sum_j (n_{ji} - n_{ii})} \)

Frequency Weighted IU:
\( \frac{1}{\sum_k t_k} \left( \frac{1}{n_{cl}} \sum_i t_i n_{ii} / (t_i + \sum_j (n_{ji} - n_{ii})) \right) \)

<table>
<thead>
<tr>
<th>Fine-tuning Set</th>
<th>Epochs</th>
<th>Pixel Acc.</th>
<th>Mean Acc.</th>
<th>Background IU</th>
<th>Building IU</th>
<th>Mean IU</th>
</tr>
</thead>
<tbody>
<tr>
<td>WorldView-2 w/ Flip</td>
<td>35</td>
<td>99.96%</td>
<td>78.74%</td>
<td>99.96%</td>
<td>45.29%</td>
<td>72.62%</td>
</tr>
<tr>
<td>WorldView-2 w/ Flip</td>
<td>100</td>
<td>99.96%</td>
<td>78.37%</td>
<td>99.96%</td>
<td>46.65%</td>
<td>73.30%</td>
</tr>
<tr>
<td>WorldView-2</td>
<td>35</td>
<td>99.95%</td>
<td>75.96%</td>
<td>99.95%</td>
<td>38.65%</td>
<td>69.30%</td>
</tr>
<tr>
<td>WorldView-2</td>
<td>100</td>
<td>99.96%</td>
<td>79.14%</td>
<td>99.96%</td>
<td>47.19%</td>
<td>73.58%</td>
</tr>
<tr>
<td>Google Maps</td>
<td>35</td>
<td>99.94%</td>
<td>52.01%</td>
<td>99.94%</td>
<td>3.98%</td>
<td>51.96%</td>
</tr>
<tr>
<td>Google Maps</td>
<td>100</td>
<td>99.95%</td>
<td>65.40%</td>
<td>99.95%</td>
<td>26.67%</td>
<td>63.31%</td>
</tr>
</tbody>
</table>

Error metrics for FCNs fine-tuned with various training sets for 35 and 100 epoch durations. Our FCN fine-tuned with WorldView-2 data for 100 epochs obtained the most promising metrics overall.