

Population Density Estimation with Deconvolutional Neural Networks

Amy Zhang¹, Xianming Liu², Tobias Tiecke², Andreas Gros¹

¹Core Data Science, ²Connectivity Lab

facebook

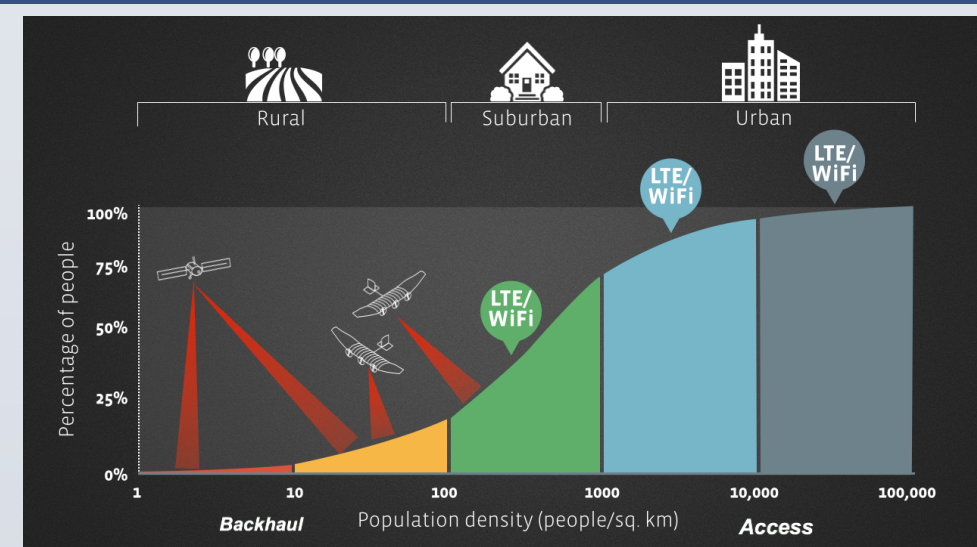
INTRODUCTION

This work is part of the Internet.org initiative to provide connectivity all over the world. Population density data is helpful in driving a variety of technology decisions, but currently, a microscopic dataset of population doesn't exist. Current state of the art population density datasets are at ~1000km² resolution. To create a better dataset, we have obtained 1PB of satellite imagery at 50cm/pixel resolution to feed through our building classification pipeline.

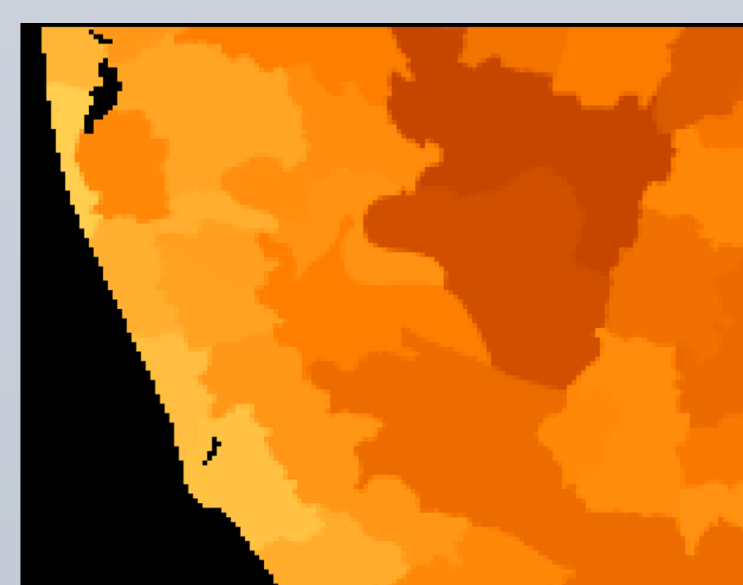
Our final data set has a spatial resolution of 30m x 30m and thereby improves over previous countrywide data sets by multiple orders of magnitude. We've analyzed 35 countries and counting.

We also worked with partners in the humanitarian community to validate the data and understand other use cases including improved healthcare delivery, infrastructure development, and crisis recovery.

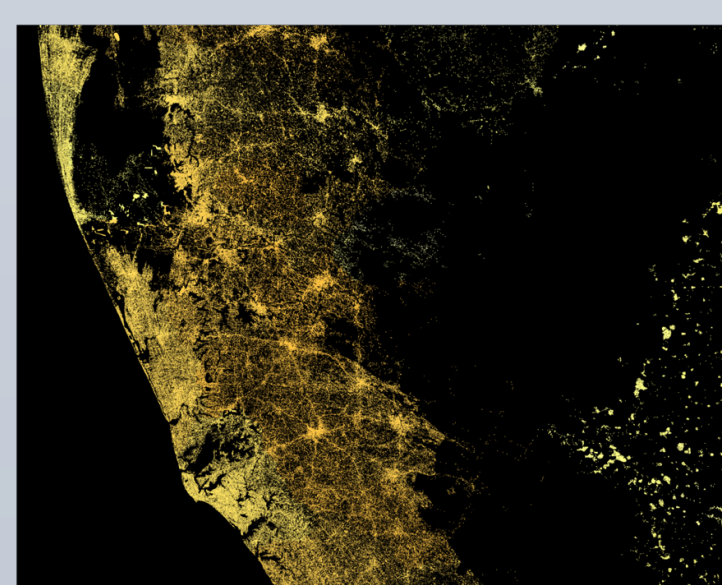
MOTIVATION & BACKGROUND



Our contribution is a fast building classification pipeline, that can run through a country in 8 hours on Facebook infrastructure. It is composed of an edge detection method for fast bounding box proposals, and a weakly supervised deconvolutional neural network that is trained for pixel-level classification, then mean pooled over the bounding box to output a probability of a building or buildings present in the bounding box. We train a global model and obtain precision and recall of > 90% in most countries. Countries with poorer results we use active learning techniques to re-sample data and fine-tune a new model. We also developed a weakly supervised footprint segmentation model that processes larger images with more context and produces a mask of location and shape of each building and a denoising network to clean up poorer quality source data.



Previous state of the art



Our data

PIPELINE

Bounding box proposals: We use a straight edge detection algorithm to generate bounding box proposals (64x64 masks centered around each straight edge)

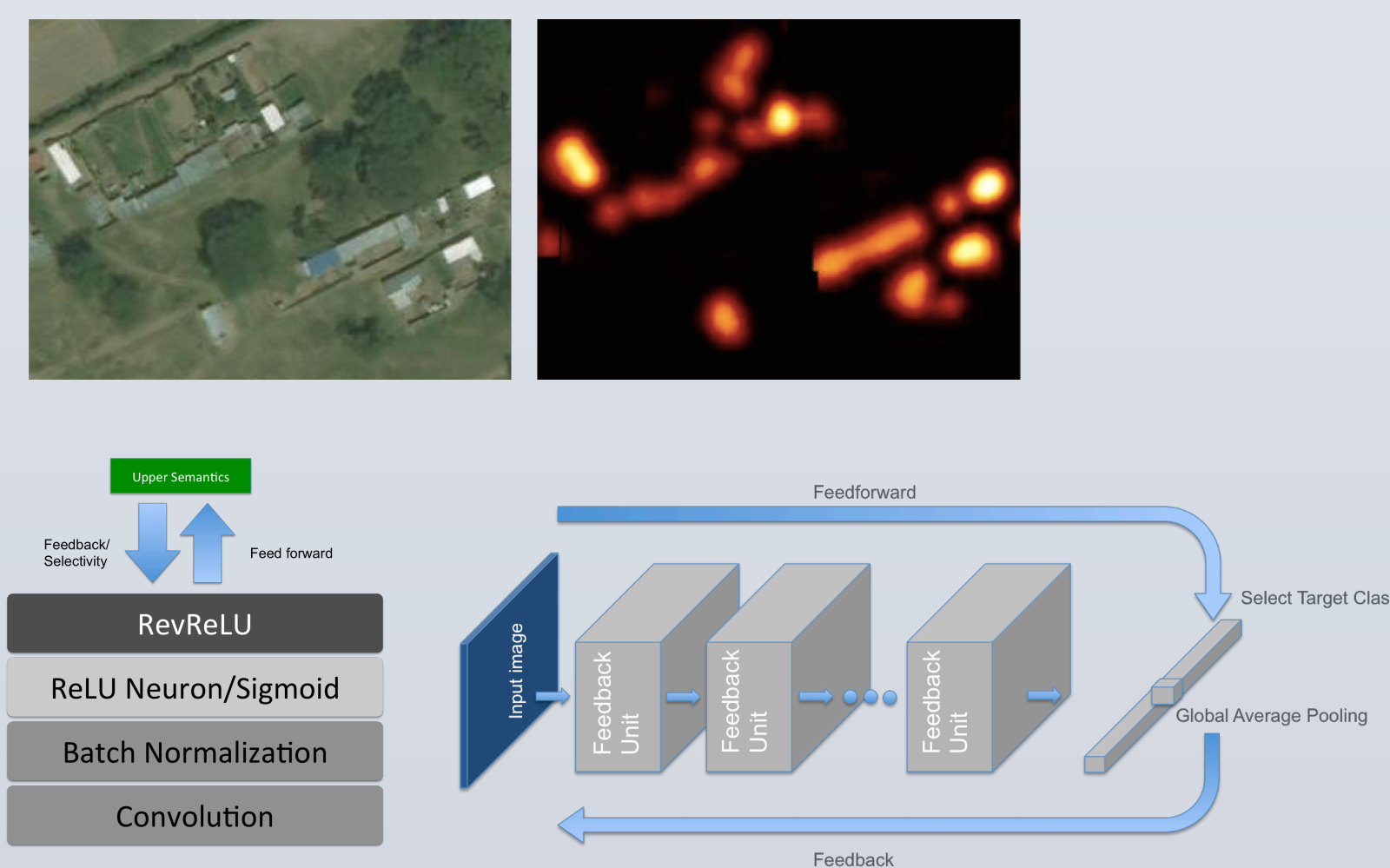
Reduces number of samples to feed through the network and reduces skew of underlying distribution.

Building classification: We use a weakly supervised SegNet architecture for building classification.

SegNet Architecture [1]:

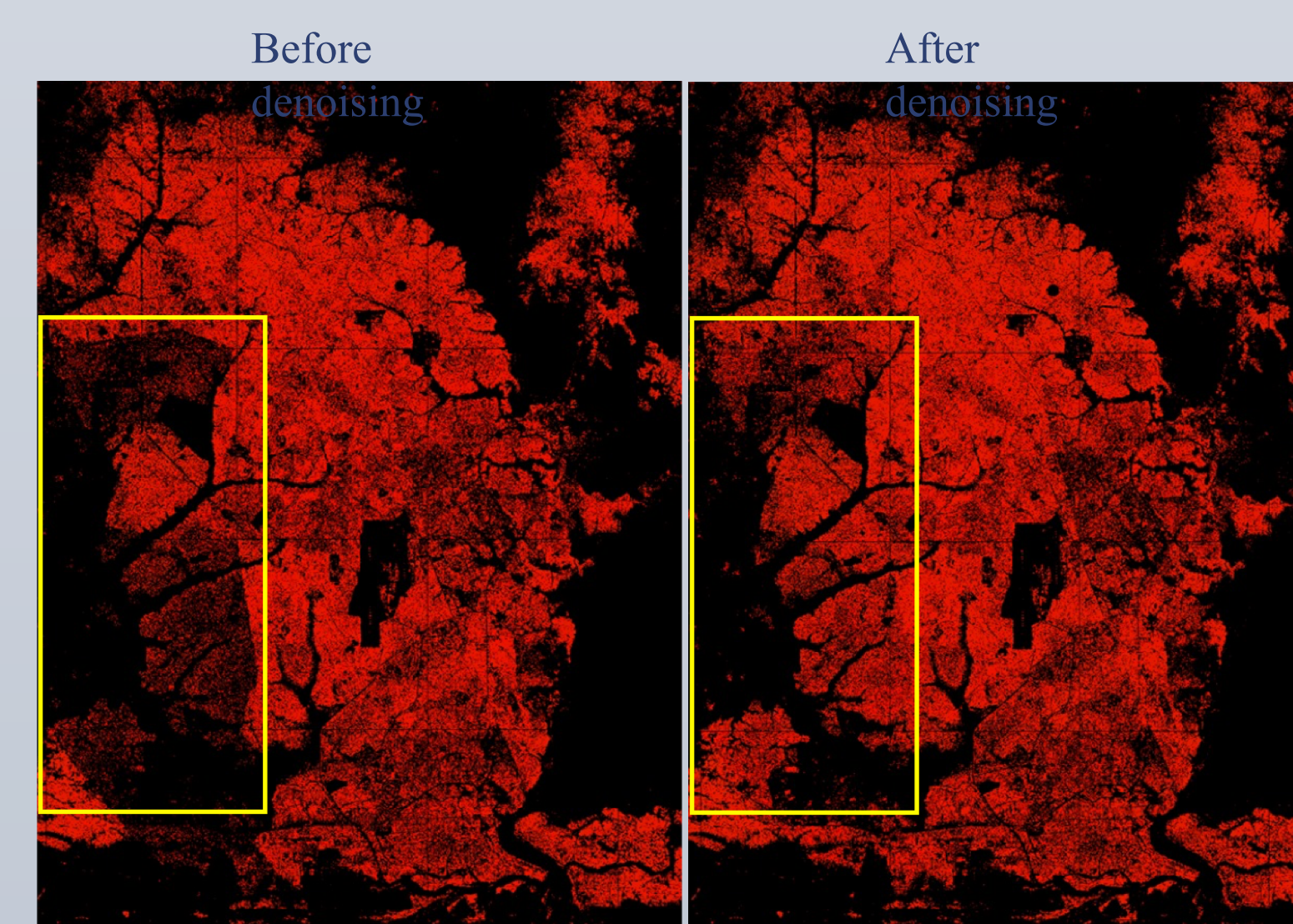
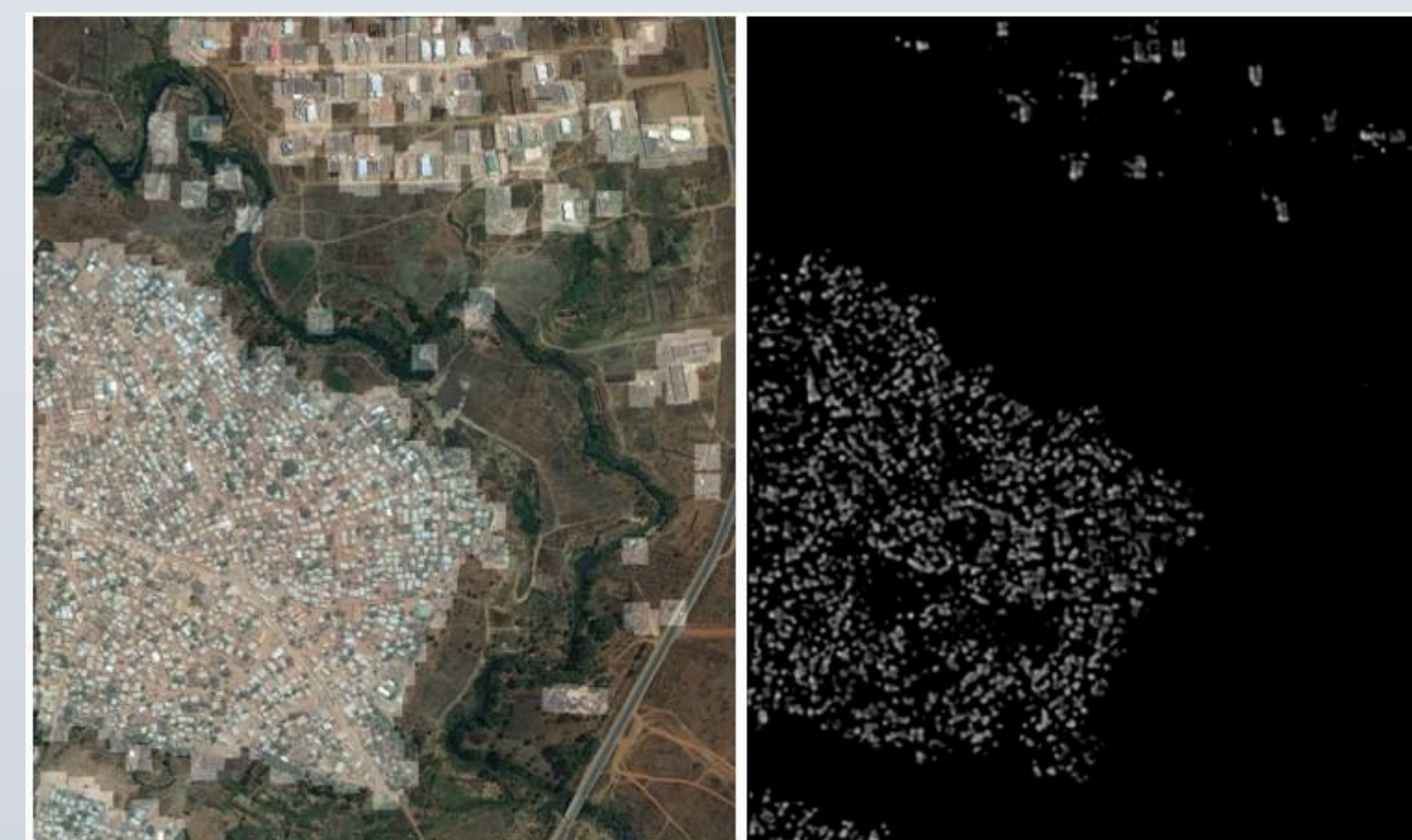
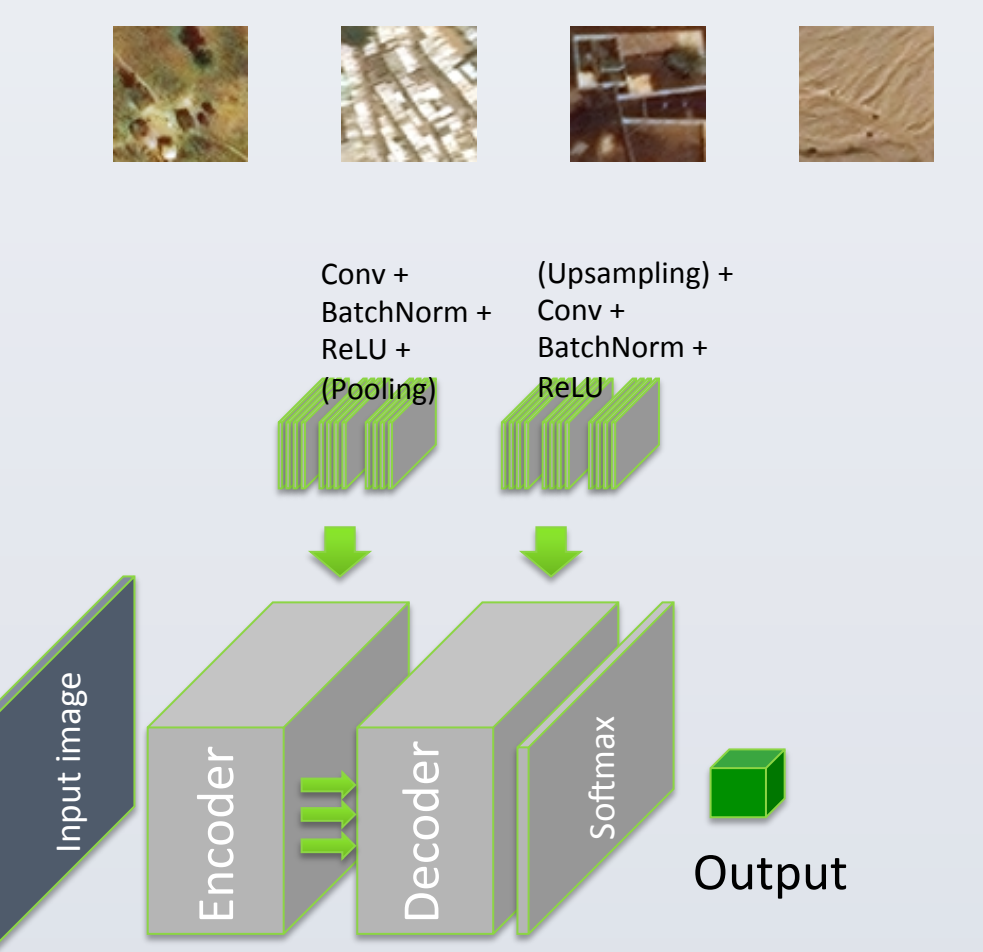
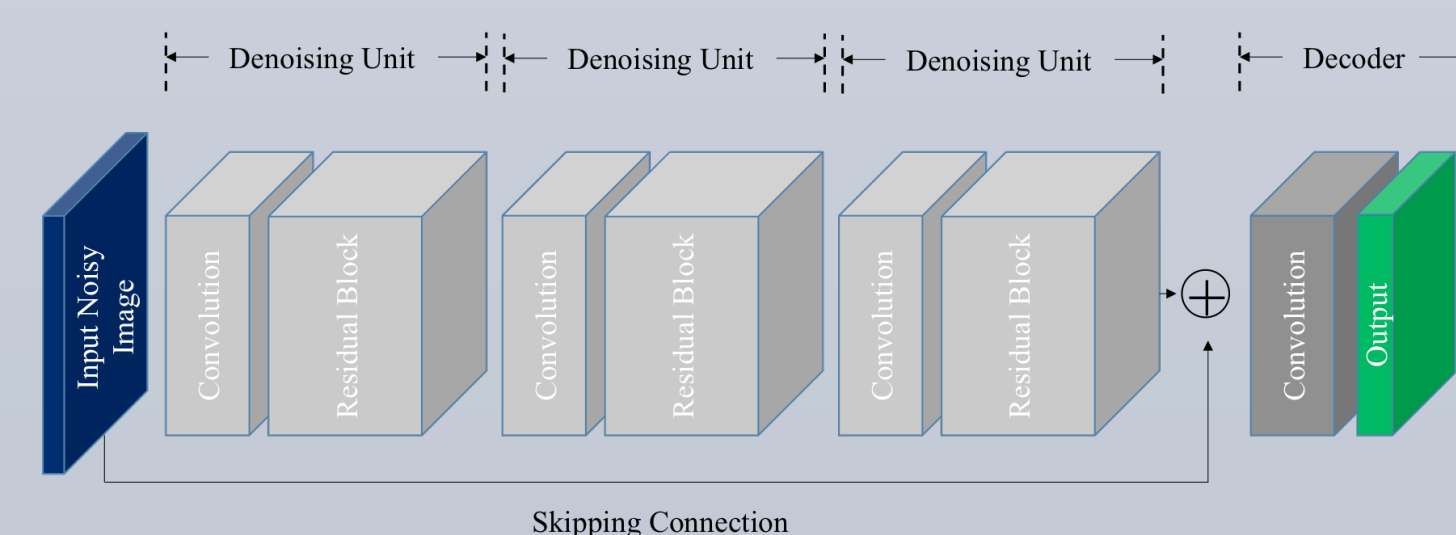
- Encoder network and Decoder network
- Encoder network is VGG16 without the fully connected layers
- Decoder network has a decoder for every encoder, upsamples feature maps according to saved max pool indices

Building segmentation: We found that models that performed well at pixel level classification for building segmentation did not perform well on an aggregate level at classification, and vice versa. Therefore we developed and trained a separate neural network for pixel-level segmentation (which tends towards false positives) which we suppress by combining with the classification results.



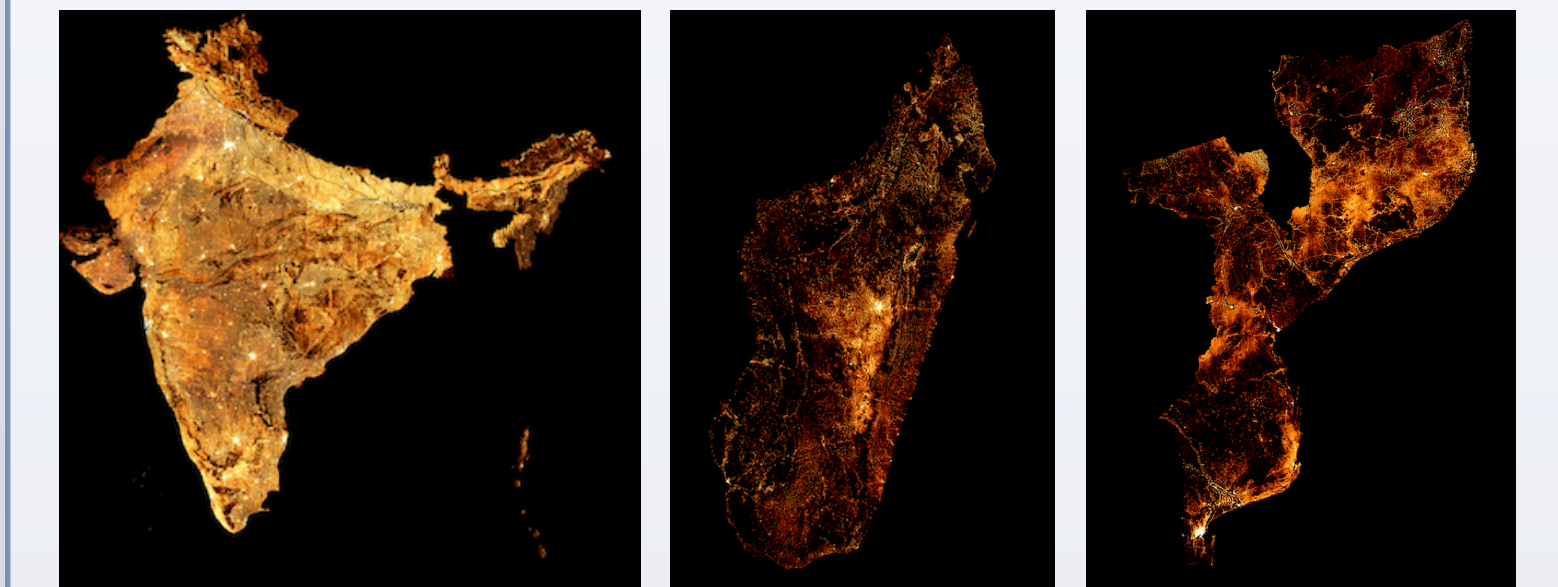
Denoising: There was quality variation in the satellite imagery that caused systematic errors (mostly false negatives).

We designed a network to denoise the source imagery before running through our classification pipeline, which dramatically improved results.



RESULTS

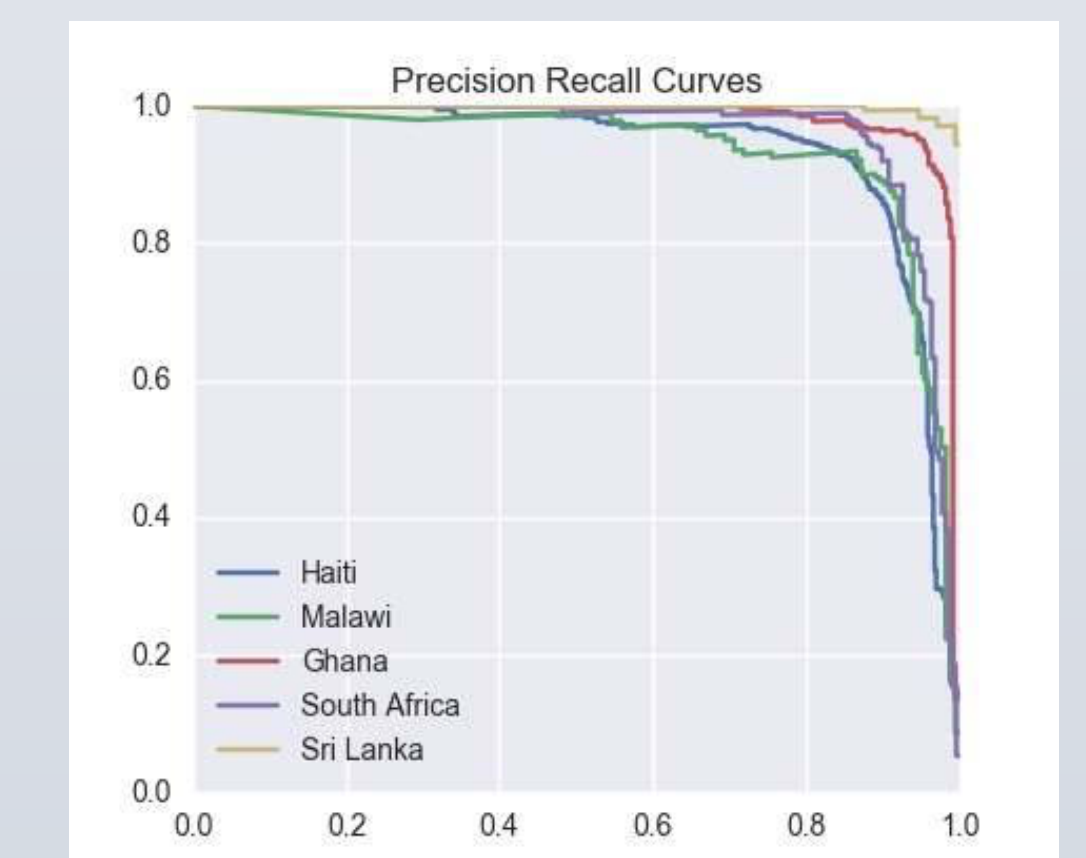
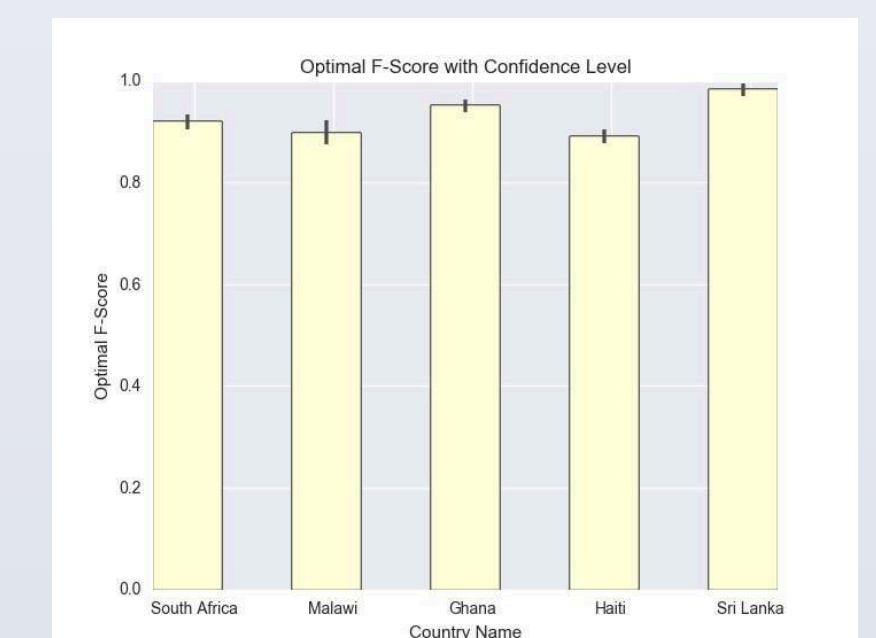
Heat maps of building density:



India Madagascar Mozambique

We've released datasets for 5 countries so far:

<https://ciesin.columbia.edu/data/hrsl/>



FUTURE WORK

- Improving robustness of models to systematic errors caused by noise and new terrain
- Increasing land coverage with new datasets
- Performing population estimates without census data

REFERENCES

- Vijay Badrinarayanan, Alex Kendall, and Roberto Cipolla. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *CoRR*, abs/1511.00561, 2015.
- Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3431–3440, 2015.
- Chunshui Cao, Xianming Liu, Yi Yang, Yinan Yu, Jiang Wang, Zilei Wang, Yongzhen Huang, Liang Wang, Chang Huang, Wei Xu, et al. Look and think twice: Capturing top-down visual attention with feedback convolutional neural networks. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 2956–2964, 2015.
- Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. Image denoising with block-matching and 3d filtering. In *Electronic Imaging 2006*, pages 606414–606414. International Society for Optics and Photonics, 2006.