MAPPING AGRICULTURAL FIELDS IN SUB-SAHARAN AFRICA

Debats S., Fuchs T., Estes L., Caylor, K.
sdebats@princeton.edu

Abstract

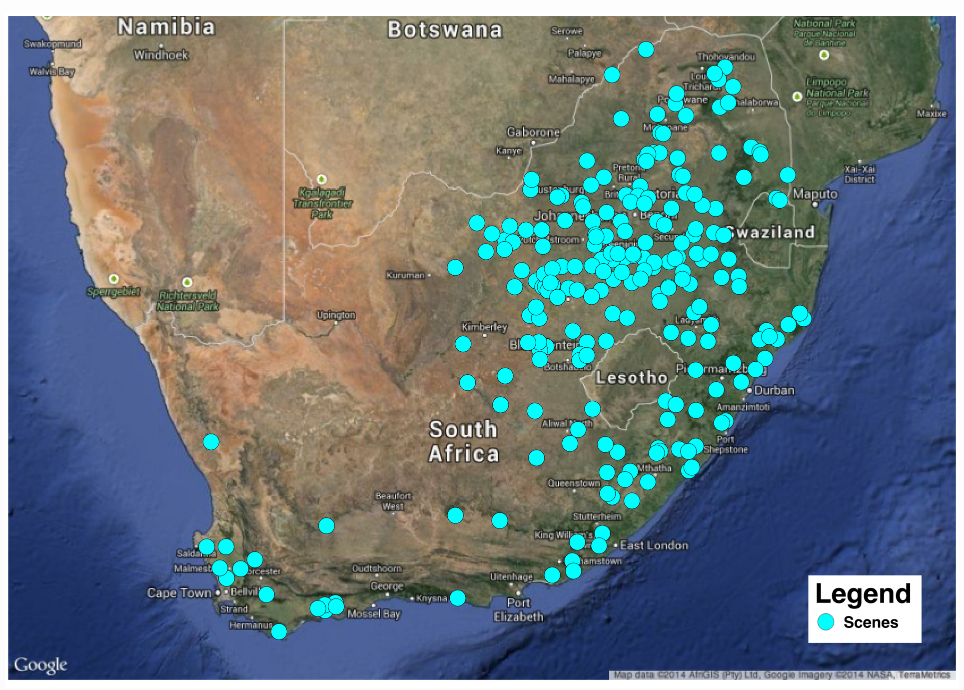
Agriculture is a main driver of land-use change, especially as population increases. 25% of the world's undernourished live in Sub-Saharan Africa, which is dominated by smallholder agriculture with field sizes less than 2 hectares (141 m x 141 m). Gridded land cover data sets provide overall extents of agriculture, but no information on smallholders' fields.

Our ongoing research focuses on mapping boundaries of individual agricultural fields across Sub-Saharan Africa, using a joint framework of a random forest and graph cuts [1],[2],[3]. The random forest classifies remote sensing image pixels, based on labeled training data. Graph cuts enforce labeling consistency by obtaining a maximum *a posteriori* estimate of a Markov random field by solving a multiway minimum cut problem, thereby smoothing classifications into coherent regions corresponding to agricultural fields.

Labeled Data

Government-issued hand labels of field boundaries of South Africa. 222 scenes (1 km x 1 km) were extracted and paired with Google Maps satellite images ($\sim$ 2.5 m resolution).

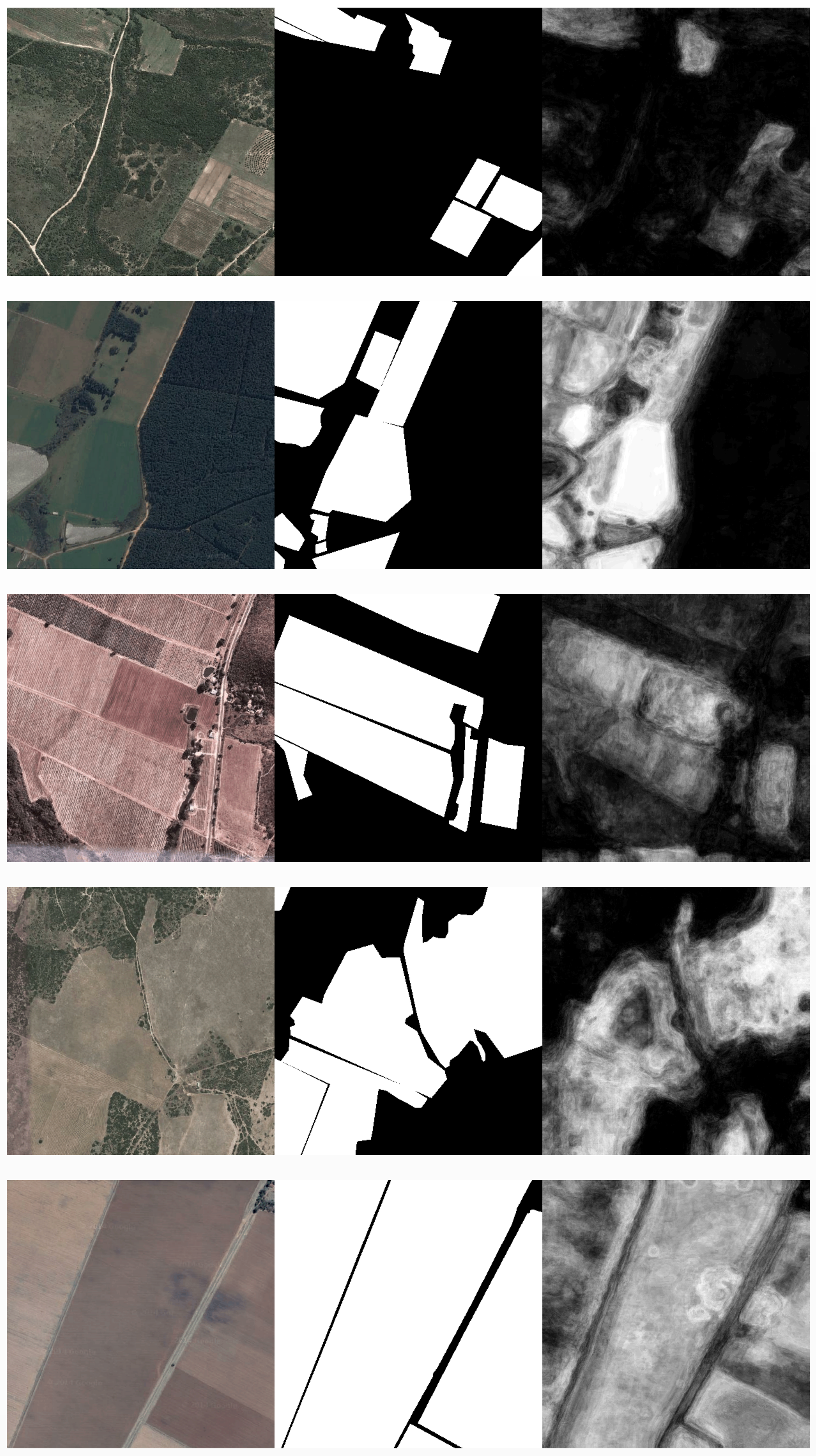
South African data set of hand-labeled field boundaries



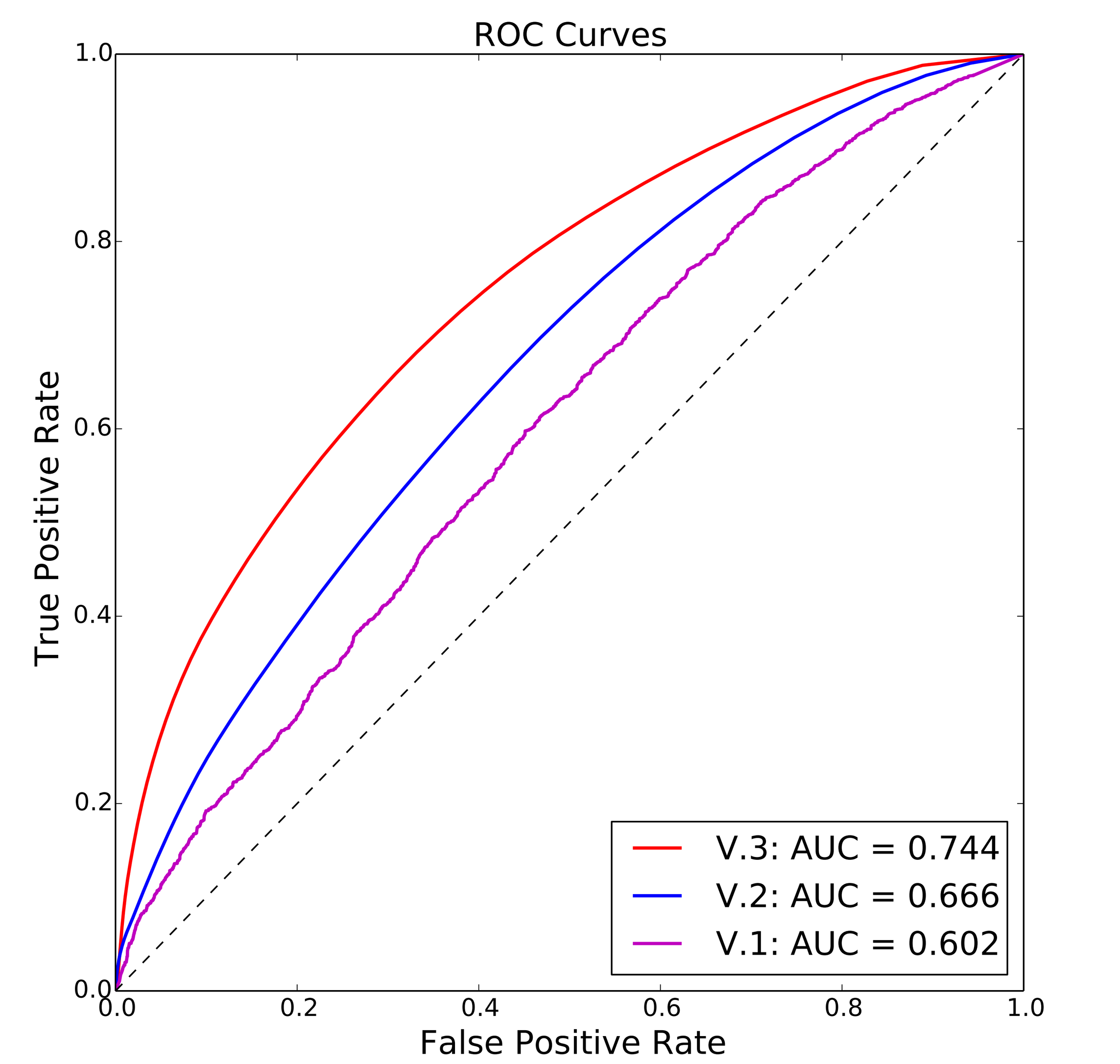
222 scenes used in analysis

The Princeton Mapping Africa Project (`http://mappingafrica.princeton.edu`) produces additional hand labels using the crowd-sourcing platform, Amazon Mechanical Turk. Collecting additional training data will enable the expansion of the algorithm's geographic range to areas in Sub-Saharan Africa outside of South Africa.

Random Forest

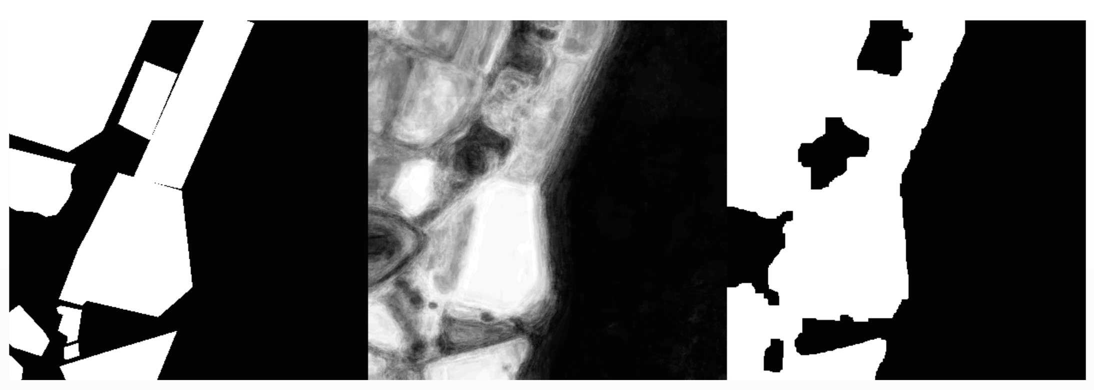


From left to right: Google Maps images, hand labels, and random forest class probabilities



- Feature channels include various image filters applied to the RGB images, while features consist of averaging windows.
- Classifier performs well on larger, uniformly textured agricultural fields.
- Classifier struggles with very small, isolated fields and in-field variability.

Graph Cuts



From left to right: hand labels, RF class probabilities, and graph cut segmentation

After an appropriate accuracy is obtained with the random forest classifications, graph cuts will segment the images by applying a min-cut/max-flow algorithm to minimize the energy function [2]:

$$E(L) = \sum_{p \in P} D_p(L_p) + \sum_{(p,q) \in \mathcal{N}} V_{p,q}(L_p, L_q)$$

Future Work

- Increase random forest accuracy with more features, such as additional RGB features, textons and single-histogram class models, or histogram of oriented gradients (HOG) features [4].
- Addition of multi-spectral remote sensing imagery (e.g. Landsat) in addition to RGB.
- Expand geographic range of classifier with crowd-sourced training data sets.
- Implement graph cuts and polygonize field boundaries for land cover maps.
- Apply algorithm to time series of images to track agricultural land-use changes.

References

[1] L. Breiman, Random forests, in *Machine Learning*, 2001

[2] Y. Boykov, V. Kolmogorov, An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision, in *IEEE TPAMI*, 2004

[3] V. Kaynig, T. Fuchs, J. Buhmann, Neuron geometry extraction by perceptual grouping in ssTEM images, in *CVPR*, 2010

[4] F. Schroff, A. Criminisi, A. Zisserman, Object Class Segmentation Using Random Forests, in *BMVC*, 2008

Acknowledgements

This research is supported by National Science Foundation (NSF) Grant No. 1026776; NASA Jet Propulsion Laboratory Student Internship Program (JPLSIP); and Princeton Environmental Institute's Science, Technology, and Policy Fellowship (PEI-STEP).