Crop Classification

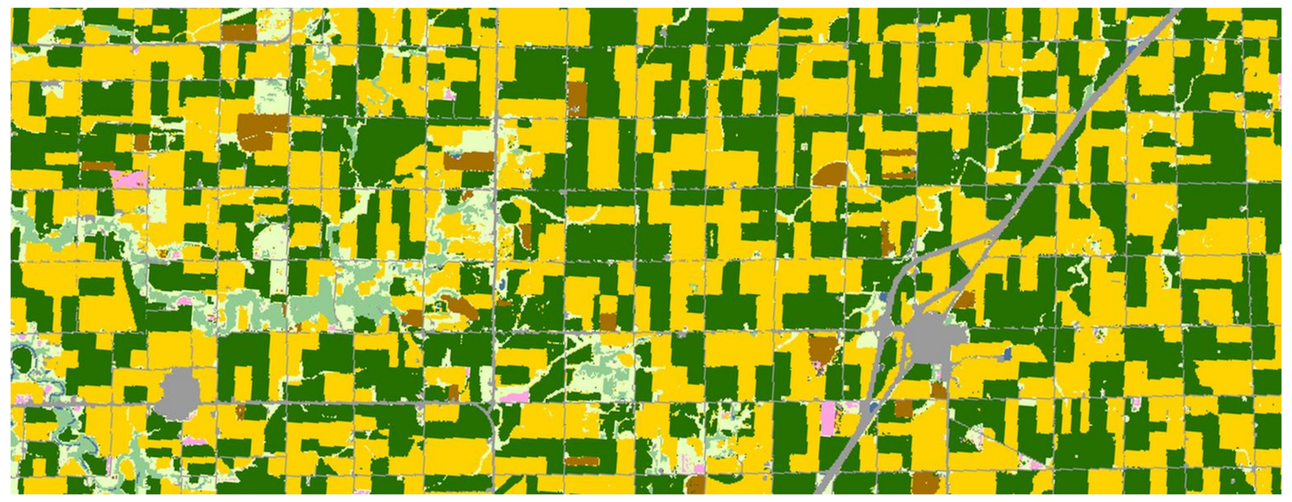
Spencer Barton and Darrin Willis
{sebarton, dswillis}@andrew.cmu.edu

Overview

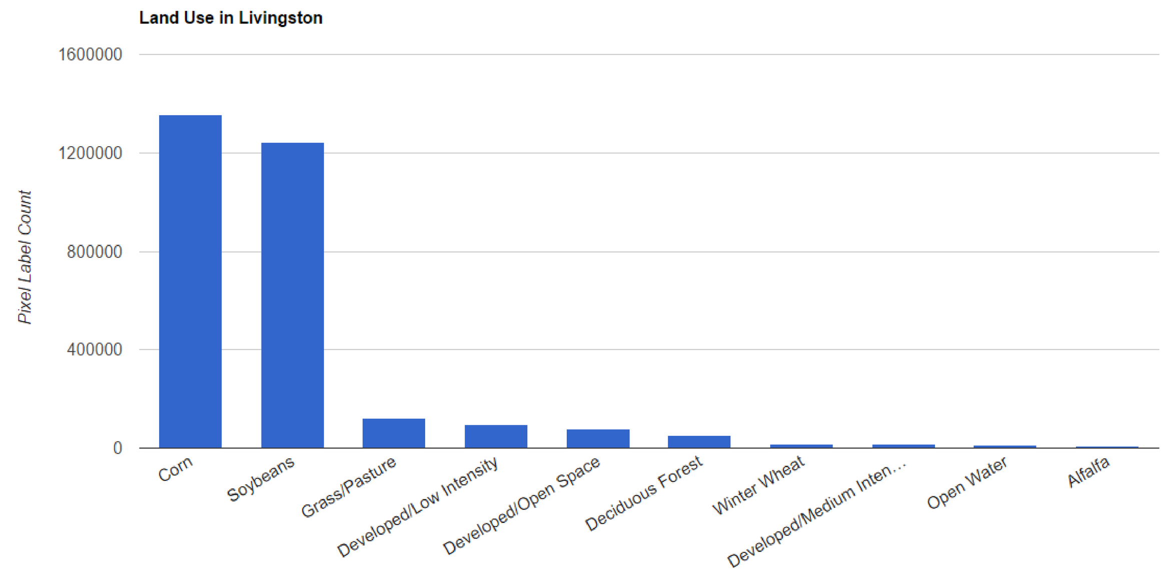
This project aims to classify crops via satellite images. There are two fundamental objectives within this: the crop fields must be segmented out from each other, and each segmented field must be classified. This lends itself well to a two stage pipeline, in which first a super-pixel clustering algorithm is employed to find the fields and then some classifier is used to determine the field.

Datasets

We chose to focus on Livingston Co., IL which is home to mostly soybeans and corn. Our data labels came from CropScape, a USDA project.



Our satellite images came from Google Maps. We utilized the API to select specific map regions. The land in Livingston is mostly corn and soybeans:



Data Cleaning

We face two key challenges early on: splicing maps and aligning the training data. The Maps API only allowed us to extract small images which we needed to splice together. This splicing did not work out well:

For image/label alignment we utilized ImageJ to manually match the two images. The labels were noisy so we used SLIC to find similar regions which we relabeled with the dominant label.

Algorithm: Segmentation

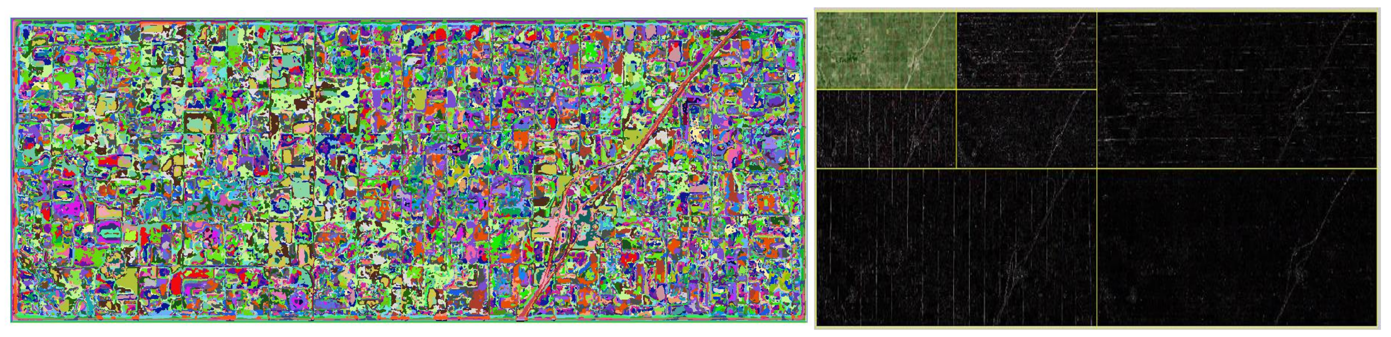
We primarily used the SLIC segmentation algorithm for the first half of our pipeline, and experimented with several different algorithms for the classification step.

SLIC worked well as evidenced by this image with pixels labeled with the average pixel value:

Algorithm: Classification

K-means worked moderately well, but suffered from the problem of mapping k-means clusters to crops. Taking the most common k-means labeling for each ground truth label ended up over-labeling everything as a single crop. K-nearest-neighbor was also used, but required more labeled data than we had access to in order to be effective. K-means result below:

Our feature vectors for KNN and K-means were the main part of the algorithm which was tweaked with. We experimented with a bag-of-words style histogram, as well as a set of wavelet responses.



Results

Our results were mixed. We tried a number of classifiers and feature vectors but were not able to get great classification results. Our correct classification rates:

Algorithm	Features	Accuracy
Kmeans	Histograms of haar wavelets on LAB space	0.19
Kmeans	LAB Histograms	0.26
ANN, 10 nodes, 1 layer	RGB, LAB and wavelets histograms with only 3 labels (corn, soy, neither)	0.28
KNN, K = 1, euclidean distance	LAB Histograms	0.30
Kmeans	Histograms of Haar wavelets on RGB space	0.31
KNN, K = 5, euclidean distance	Wordmap histograms	0.34
Kmeans	RGB Histograms	0.36
KNN, K = 3, euclidean distance	RGB, LAB and wavlets histograms	0.46

References

- A. Laine, J. Fan, "Texture classification by wavelet packet signatures," Pattern Analysis and Machine Intelligence, IEEE Transactions on (http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=244679&tag=1)
- A. Vedaldi and B. Fulkerson, "VLFeat: An Open and Portable Library of Computer Vision Algorithms," 2008. http://www.vlfeat.org/
- Google Maps. (2015). [Livingston County, IL] [Earth Map]. Retrieved fromhttps://www.google.com/maps/place/Livingston+County,+IL/@40.8653884,-88.5832004,10z/data=!3m1!4b1!4m2!3m1!1s0x880c64feed7dacb7:0x84dd08edd462cd64?hl=en
- Rasband, W.S., ImageJ, U. S. National Institutes of Health, Bethesda, Maryland, USA, http://imagej.nih.gov/ij/, 1997-2014.
- USDA National Agricultural Statistics Service Cropland Data Layer. {YEAR}. Published crop-specific data layer [Online]. Available at http://nassgeodata.gmu.edu/CropScape/ (accessed 2015; verified 2014). USDA-NASS, Washington, DC.