Computing the Rate of Disappearance of Cropland Using Satellite Images

Sunandan Chakraborty[1], Scot Dalton[1], Yaw Nyarko[2], Lakshminarayanan Subramanian[1]
sunandan@cs.nyu.edu, std5@nyu.edu, yaw.nyarko@nyu.edu, lakshmi@cs.nyu.edu
[1]Dept of Computer Science and [2]Dept of Economics
New York University

1 Introduction

Croplands worldwide are in decline. Degradation of arable land is a cause for concern, especially in developing countries where agriculture, including subsistence farming, makes up a significant percentage of economic output. In developing regions, urban population is increasing, leading to expansion of cities and development of new cities or townships. Often these expansions are done on arable lands. Apart from urban expansion, industrial developments are often done on agricultural land [2012]. All these result into acquisition and loss of arable lands. On many occasions, these acquisitions are unplanned and unauthorized. Such loss of arable land can have huge impact, particularly for agrarian economies. Not only it can affect the lives and livelihoods of the population who are directly dependent on agriculture, it can directly impact food security due to reduced production. Apart from human-led development, changing climate is also leading towards a change in the land pattern. There are reports of Sahara desert expanding southwards in an alarming rate [2012], loss of land in low lying coastal areas due to rising sea level etc. In both the two scenarios described above, a solution to the problem can be a monitoring tool, which can identify the change in land pattern over the years.

In this paper, we present a tool that can monitor this change through satellite images. Google Earth (GE) offers a huge corpus of satellite images across the globe. GE has the current image of a location, as well as an archive of older images of the same places. Given a location or a geographical area, our tool can access the latest available satellite image in addition to earlier images available and classify the images, as cropland, developed, forest or barren. Following this classification process, the tool computes the total amount of change of pattern in the region and also the type of change (e.g. crop land changed to developed land etc.). Figure 1 shows some GE images depicting the loss of vast open land over the years in 2 African cities. This figure also explains how GE images can be used to detect such changes. Similar approaches can be seen in various fields, like famine and agriculture [Quinn et al, 2010][Nivens et al, 2002], environmental changes to detect outbreak of diseases [Ford et al, 2009] etc. In our case, we used raw photographs taken from the satellites from a freely available source with extensive coverage. This makes the approach much more scalable.

Fig. 1. (*from top-left clockwise*) Luanda,Angola [2003]; Luanda,Angola [2011]; Johannesburg, SA [2005]; Johannesburg, SA [2011] (Images captured from Google Earth). The top row images show that how an empty field was completely taken over by housing development in Luanda within 8 years. Similarly, in Johannesburg (bottom row), a green field is totally lost in 6 years due to various developments.

2 System Description

The tool takes a name of a location or its latitude-longitude coordinates as input. Based on the location, the tool uses the Google Earth API to capture all available images, the most recent one as well as the older images in the archive. All images are taken are taken at a uniform elevation of 10 Kms. If the input location is a name, like a province/state or a city, the tool collects the bounding latitude-longitude coordinates of the location and sweeps across the entire region to collect all the images covering the area. The tool is capable of classifying all such images into 4 categories, *Cropland, Developed (urban or industrial), Forest, and Barren.*

Entire image of a location can have sub-regions with one or more of the above categories. Hence, the classification is not done on the entire image of the location. The classifier identifies the regions in the image that fall into each of these categories. The classifier is pre-trained on a manually labeled training set, also obtained from Google Earth. After this classification process, the tool computes the percentage of sub-regions falling under each category, in a particular year. For example, after the classification process for a location, say x, the tool will produce results like, Location: x; year: *2003*; Crop land: *15%*; Developed: *40%*; Forest: *0%*; Barren: *35%.* Finally, comparing these values of the latest image with the earlier images, the tool will compute the rate of change in the land pattern in each category. Figure 2 summarizes the overall architecture of the system.

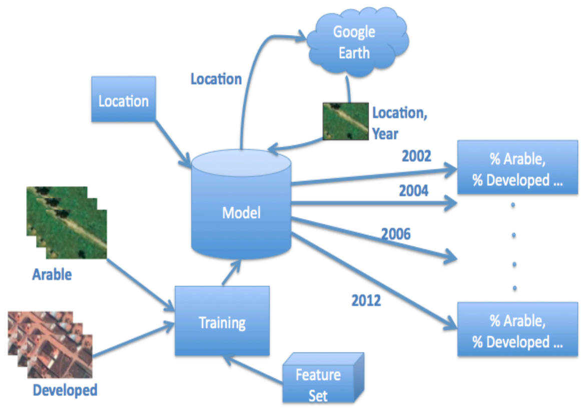

Fig. 2. System Architecture

2.1 Training the System

The tool needs to be trained properly, in order to function in the way it was described above. Training phase involves providing the system with various example images and their corresponding categories. For example, feeding the system with numerous images, which falls under *cropland* category. This way the system will know, how a cropland looks like. Once the system has enough examples of known images, it can take an unknown image and classify it to a proper category.

We collected several images across Africa and this constituted the training set, on which the tool will be trained. The images were collected from countries like, South Africa, Uganda, Nigeria, Malawi, Ghana and 13 other different countries. All the images had uniform scale and elevation of 10 Kms. We divided these sample images into blocks of 100 x 100 pixels. Each such block was treated as a single point in the training data. The blocks were manually tagged with the most representative label among the four categories: cropland, developed, desert and forest. 4000 such blocks (1000 under each category) made up our training set and around 1000 (250 per category) such blocks were kept as the development test set to test the performance of the system.

Several features were extracted from the images to train the system. The features included, mean pixels from the histogram as a measure of color; gradient magnitude convolved using a Sobel gradient mask to measure edges; histogram standard deviations for measuring color variation; Discrete Cosine Transform and SURF (Speeded Up Robust Feature) for measuring texture. Using these features, we trained the model using the K-nearest neighbors algorithm [Dudam et al, 2001].

Once the training phase is over, the system is ready to use. Given a new location, the system can automatically extract the corresponding images from GE and classify the images to compute the percentage change of cropland area in that location.

3 Results

The training set, on which the model was trained had 4000 images. To test the performance of the model we had a separate development test set of 1028 images. The test set had images similar to Figure 1. The tool could detect the changes in the land pattern, as visible in the images of Figure 1. After performing various experiments, the best performance was observed for K-nearest neighbors algorithm and histogram mean pixels as the feature. The accuracy in this case was: **84.43%**. Given the quality of the images, where in some cases, images were blurry, had cloud covers, had variable color shades across different images, classification is a difficult task. Due to such conditions, an accuracy of 84% is quite acceptable and even a slight increase in performance can be a formidable task. We are currently using the system to gather images from entire areas of few African countries and Indian states to compute the changes in cropland areas of these places. We can further test the system by comparing the percentage changes given by the system with the ground truth from official records.

4 Future Work

For further improvement in the model, the feature set can be expanded to include texture and/or edge detection features and enhanced training data, more classification categories etc. Detection of desertification, deforestation or disappearance of water bodies, and also, prediction capabilities in the system can greatly enhance its use.

References

[2012a] http://www.expressindia.com/latest-news/Now-seer-threatens-fast-over-land-acquisition-in-Karnataka/816705/

[2012b] http://www.bbc.co.uk/learningzone/clips/desertification-expansion-of-the-sahara-desert/1498.html

[Quinn et al, 2010] Quinn, J., Okori, W., Gidudu, A.: Increased-specificity famine prediction using satellite observation data. In ACM DEV, London, UK, 2010

[Ford et al, 2009] Ford, T. E., Colwell, R. R., Rose, J. B., Morse, S. S., Rogers, D.J., and Yates, T. L.: Using Satellite Images of Environmental Changes to Predict Infectious Disease Outbreaks. In J. of Emerging Infectious Diseases. V. 15 (9)

[Nivens et al, 2002] Nivens, H. D., Kastens, T. L., Dhuyvetter, K. C., Featherstone, A. M.: Using Satellite Imagery in Predicting Kansas Farmland Values. J. of Agricultural and Resource Economics 27(2):464-480

[Dudam et al, 2001] Duda, R.O., Hart, P.E., Stork Wiley, D.G.: Pattern Classification. Wiley, 2nd edition, 2001