Environmental Modelling & Software 80 (2016) 41–53

Contents lists available at ScienceDirect

Environmental Modelling & Software

journal homepage: www.elsevier.com/locate/envsoft

A platform for crowdsourcing the creation of representative, accurate landcover maps

L.D. Estes [a, b, *, 1], D. McRitchie [c, 1], J. Choi [a], S. Debats [a], T. Evans [d], W. Guthe [a], D. Luo [a], G. Ragazzo [a], R. Zempleni [a], K.K. Caylor [a]

[a] Civil and Environmental Engineering, Princeton University, Princeton, NJ 08544, USA
[b] Woodrow Wilson School, Princeton University, Princeton, NJ 08544, USA
[c] Computational Science and Engineering Support, Office of Information Technology, Princeton University, Princeton, NJ 08544, USA
[d] Department of Geography, Indiana University, Bloomington, IN 47405, USA

ARTICLE INFO

Article history:
Received 1 May 2015
Received in revised form
10 November 2015
Accepted 30 January 2016
Available online xxx

Keywords:
Remote sensing
Landcover
Crowd-sourcing
Accuracy assessment
Representative sampling
Object extraction

ABSTRACT

Accurate landcover maps are fundamental to understanding socio-economic and environmental patterns and processes, but existing datasets contain substantial errors. Crowdsourcing map creation may substantially improve accuracy, particularly for discrete cover types, but the quality and representativeness of crowdsourced data is hard to verify. We present an open-sourced platform, DIYlandcover, that serves representative samples of high resolution imagery to an online job market, where workers delineate individual landcover features of interest. Worker mapping skill is frequently assessed, providing estimates of overall map accuracy and a basis for performance-based payments. A trial of DIYlandcover showed that novice workers delineated South African cropland with 91% accuracy, exceeding the accuracy of current generation global landcover products, while capturing important geometric data. A scaling-up assessment suggests the possibility of developing an Africa-wide vector-based dataset of croplands for $2–3 million within 1.2–3.8 years. DIYlandcover can be readily adapted to map other discrete cover types.

Availability

DIYlandcover's source code will be made available free of charge for suitable non-commercial purposes under a GPLv3 license, upon consultation with the authors. For those interested in commercial applications, the prospective licensee should contact Princeton University's Office of Technology Licensing. The details of a specific application of the software for delineating crop fields in sub-Saharan Africa can be found at mappingafrica.princeton.edu, together with associated information about participating in the project, including digitizing rules and links for accessing the mapping interface.

1. Introduction

Regional maps of landcover provide critical information on food security estimates (e.g. Monfreda et al., 2008; Licker et al., 2010; See et al., 2015; Lobell, 2013), models of land–atmosphere interactions (e.g. Liang et al., 1994), and calculations of carbon stocks (e.g. Ruesch and Gibbs, 2008), greenhouse gas emissions (e.g. Searchinger et al., 2015), and habitat change (e.g. Gibbs et al., 2010). These maps are particularly important in developing regions, such as sub-Saharan Africa, where government land use data are often lacking, error-prone, and inconsistent (Ramankutty et al., 2008; See et al., 2015). These developing regions are also experiencing rapid land use changes (Gibbs et al., 2010; Rulli et al., 2013) that pose pressing development challenges (e.g. how to feed people at substantially lower environmental cost Searchinger et al., 2015).

Unfortunately, landcover datasets derived from medium to coarse resolution satellite sensors are particularly inaccurate (Fritz et al., 2010; Fritz and See, 2008). One major reason for poor accuracy is the fact that land use patterns in these regions are dominated by smallholder farming. Smallholder fields are typically smaller ($\leq$2 ha) than the resolution ($\leq$ 6 ha) of the most commonly

* Corresponding author. Woodrow Wilson School and Civil and Environmental Engineering, Princeton University, Princeton, NJ 08544, USA.
E-mail address: lestes@princeton.edu (L.D. Estes).
[1] Equal contributors.

http://dx.doi.org/10.1016/j.envsoft.2016.01.011
1364-8152/

used satellite imagery (Jain et al., 2013). Furthermore, smallholders often plant diverse mixtures of crops, which further increases within-pixel heterogeneity (Jain et al., 2013), and their fields often contain remnant trees and have irregular boundaries, which makes them spectrally harder to distinguish from the surrounding vegetation (See et al., 2015; Lobell, 2013).

New techniques for merging multiple landcover products are helping to substantially improve map accuracy (Fritz et al., 2011, 2015). However, these approaches cannot overcome the mismatch between sensor resolution and smallholder field size. High resolution satellite imagery ($\sim$ 5 m) is becoming increasingly available—and presumably will become more affordable—so the resolution problem should be solved in the near future (See et al., 2015; Lobell, 2013). But high resolution comes at the expense of higher spectral variability; centimeter-scale data require lower orbits, narrower swaths, and greater communication bandwidth, which combine with clouds to greatly limit the area that can be imaged under contemporaneous environmental conditions, and from comparable viewing angles. This means that high resolution image mosaics covering large areas contain substantial and largely uncorrectable spectral differences caused by variations in atmospheric conditions, vegetation phenology, and bidirectional reflectance. This variability propagates error in automated classifications over large regions, which can already be substantial when there is high within-cover variability (Debats et al., 2015), or high heterogeneity among cover types (Gross et al., 2013).

It remains a major challenge to develop algorithms that can accurately classify landcover in the face of both increased image variability and substantial spatial heterogeneity. Promising methods are emerging, which draw on advances in computer vision and machine learning, such as semantic segmentation (e.g. Schroff et al., 2008) and Randomized Quasi-Exhaustive feature selection (Tokarczyk et al., 2015), to find optimal classifiers within complex urban environments Frhlich et al. (2013) and highly variable smallholder fields (e.g. Debats et al., 2015). However, these advances are primarily in pixel-wise classification. Accurate, automated methods for extracting individual objects within a single cover type, particularly over wide areas, is arguably even more difficult. Object delineation is an important goal of landcover mapping, as cover geometries encode critical social and environmental information (Fritz et al., 2015), and can play an important role in improving environmental monitoring systems. For example, in agroecosystems, field boundaries can provide a filter for extracting “pure”, crop-specific time series of satellite-derived vegetation indices, which helps to improve the accuracy of remotely sensed yield estimates (Estes et al., 2013a, b). Some limited progress has been made with automated approaches, but these have been demonstrated mainly for small areas where the cover objects have regular geometries and sharp boundaries (e.g. commercial agricultural fields Yan and Roy, 2014; Ozdarici-Ok and Akyurek, 2014; Ozdarici-Ok et al., 2015). Such methods are not yet proven over large areas with more complex, less distinct cases.

An alternative approach is to employ humans, who are very adept at recognizing patterns in noisy images (Biederman, 1987). The superiority of human over machine pattern recognition provides the motivation for CAPTCHA (Ahn et al., 2003), which secures websites by requiring human users to recognize fuzzy or irregular letters and numbers that are too difficult for automated algorithms to identify. Human-interpreted landcover maps are thus likely to be consistently more accurate than automated classifiers. Unfortunately, since humans are much slower at data processing than computers, human-generated landcover maps covering large areas will require much more time and expense to create. However, this problem is being alleviated by the growth of the internet, which makes it increasingly feasible to turn pattern recognition problems into many small tasks that are undertaken by a large number of online workers—the human equivalent of parallel processing. This ability to “crowdsource” (Howe, 2006) such work supports projects ranging from galactic classification (Lintott et al., 2008) to ornithological surveys (Sullivan et al., 2009). Crowdsourcing of landcover is already being used in the Geo-wiki project, which uses online volunteers to correct landcover data based on their own interpretations of high resolution satellite imagery (Fritz et al., 2009, 2012, 2015). Recently, these data have been used to create the most accurate (82%) global cropland map (Fritz et al., 2011, 2015).

While the use of crowdsourcing is an extremely promising development for landcover mapping, and is being increasingly used for this and other environmental monitoring applications (Jacobson et al., 2015; Fraternali et al., 2012; Schellekens et al., 2014), many existing projects (e.g. OpenStreetMap (openstreetmap.org)) are geared towards users who create content according to their personal interests, thus the resulting maps are unlikely to be geographically representative (Fraternali et al., 2012). Furthermore, verifying the accuracy of crowdsourced data is a challenge (Allahbakhsh and Benatallah, 2013; Flanagin and Metzger, 2008; See et al., 2015) that remains largely unaddressed by existing platforms. In terms of using crowdsourcing to improve landcover data, prior efforts have focused primarily on validating pixel-based classifications, and less on delineating individual cover objects, which is arguably one of the greatest advantages that people have over machines. Indeed, recognizing and digitizing individual, discrete cover types such as crop fields is considered fairly “straightforward” for humans (Yan and Roy, 2014).

In this paper, we describe *DIYlandcover* (or “Do-it-Yourself” landcover), a new platform for creating crowdsourced landcover data that addresses the three aforementioned limitations. DIYlandcover was designed for mapping discrete, but “noisy”, cover types, where object extraction is of primary interest. Specifically, our platform provides online workers with tools to 1) delineate landcover objects within 2) representatively selected locations, while the resulting maps are subjected to 3) periodic quality assessments that provide estimates of individual worker and overall map accuracy. We provide an overview of DIYlandcover's design and mechanics, and report on the results of a trial mapping crop fields in South Africa, which suggest that DIYlandcover allows inexperienced online workers to generate high accuracy ($>90\%$), geometrically rich, and geographically representative landcover data at a much faster rate than is usually possible with human-based mapping.

2. System design

The inspiration for DIYlandcover came from GeoTerraImage, a company that mapped South Africa's arable cropland by manually digitizing fields visible in high resolution satellite imagery (GeoTerraImage, 2008). The resulting map set is 97% accurate in distinguishing cropped from uncropped areas at a 4 ha resolution (see detailed accuracy assessment in Appendix S1), and provides rich detail on field type and geometry. However, making these maps was an expensive and lengthy process; the estimated labor cost for digitizing was \$5 km^{-2}, and the project took approximately 2.5 years to complete (Ferreira, pers. comm.).

We developed DIYlandcover to help overcome these constraints of cost and production time, while retaining the advantages of human image interpretation skill demonstrated by GeoTerraImage. Our platform connects workers in an online job marketplace to a map application programming interface (API) that hosts high resolution satellite imagery. DIYlandcover currently works with Amazon's Mechanical Turk (Services, 2012) and the Google Maps API,