Remote Sensing of Environment 217 (2018) 144–157

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

A novel co-training approach for urban land cover mapping with unclear Landsat time series imagery

Ting Hu[a], Xin Huang[a,b,*], Jiayi Li[a], Lefei Zhang[c]

[a] School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, PR China
[b] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, PR China
[c] Computer School, Wuhan University, Wuhan 430072, PR China

ARTICLE INFO

Keywords:
Land-cover mapping
Unclear observations
Landsat
Co-training
Collaborative representation classifier
Matrix completion

ABSTRACT

Landsat time-series (LTS) imagery shows potential for dynamic mapping in urban areas. However, unclear observations (clouds, cloud shadows, snow/ice, and SLC-off data) in the dataset inevitably restrict the efficacy of many of the state-of-the-art classifiers. In this work, we present a novel co-training classification approach consisting of two steps to cope with the unclear observations. Firstly, we develop a method called the MCCR classifier that deals with the unclear observations in an error-recoverable way. The clear observations are utilized to recover the unclear ones in the training samples by the use of the matrix completion (MC) algorithm, and the collaborative representation (CR) classifier is exploited to handle unclear observations in the unlabeled data. Secondly, considering that the random forest (RF) classifier is able to cope with contaminated data in an error-tolerant way, a co-training approach (CotrRM) based on the RF and MCCR classifiers is also proposed to further improve the classification efficacy. The CotrRM method is executed by iteratively constructing semi-labeled training sets based on the crisp and soft predictions of the two individual classifiers on the unlabeled data. To validate the effectiveness of the proposed MCCR classifier and CotrRM method, LTS imagery of the city of Wuhan (a metropolitan city of China) from four years (11 images from 2000, 16 images from 2005, 13 images from 2010, and 15 images from 2015) was adopted. The experiments showed that the MCCR classifier performs as well as the RF classifier for the mapping of urban land cover with contaminated LTS imagery. Moreover, the proposed CotrRM method has the ability to further improve the classification performance. The proposed approach can not only work effectively in the classification, but can also recover the unclear observations in the LTS imagery, courtesy of the MC algorithm. The overall accuracies of the land-cover changes between each two adjacent periods are all over 85%. Given the effectiveness and flexibility, the proposed method could also be applied in other unclear data classification.

1. Introduction

Urban areas host more than half of the worlds' population and play a central role in efforts to mitigate and adapt to the effects of climate and other ecosystem changes (United Nations, 2014). As human-dominated habitats, urban areas have developed at an unprecedented rate in recent decades, especially in developing countries such as China (Long et al., 2009). Urban land-cover maps are one of the most fundamental datasets used in many scientific fields, e.g., urban heat island effects (Estoque and Murayama, 2017), air pollution (Lin et al., 2015), urban ecosystem service (Haase et al., 2014) and local climate zone (Middel et al., 2014).

Remotely sensed imagery of various spatial resolutions has been widely used to produce land-cover maps. High resolution (HR) imagery has the ability to provide fine spatial detail for urban mapping (Huang et al., 2017). However, the sparse coverage, limited access, and absence of historical data impede the use of such data (Huang et al., 2014b; Zhu, 2017). Land-cover maps generated from coarse spatial resolution images, such as Moderate Resolution Imaging Spectroradiometer (MODIS) and Advanced Very High Resolution Radiometer (AVHRR), have been reported to show a limited mapping accuracy (Fritz et al., 2010), especially in urban areas with complex distributions of many combinations of materials (Chen et al., 2016). Considering the issues of spatial detail, data availability, and areal coverage, medium resolution imagery such as Landsat is more appropriate at present for urban land-cover mapping and change detection. Furthermore, the Landsat satellite archive has a long and continuous record stretching over 40 years and

* Corresponding author at: School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, PR China.
E-mail addresses: huang_whu@163.com, xhuang@whu.edu.cn (X. Huang).

https://doi.org/10.1016/j.rse.2018.08.017
Received 19 November 2017; Received in revised form 26 June 2018; Accepted 14 August 2018

has been open access since 2008, which has resulted in it being widely used in the monitoring of land-surface dynamics (Gómez et al., 2016). There have been numerous applications that have adopted Landsat imagery as the main data source, e.g., forest disturbance and recovery surveillance (Senf et al., 2015; Grogan et al., 2015; Matthew et al., 2016; White et al., 2017), agricultural expansion and intensification (Kontgis et al., 2015; Qin et al., 2015), and impervious surface cover characterization (Lu et al., 2011; Zhang and Weng, 2016).

Single-date images only reflect static land surfaces, but the multi-temporal signature extension provides dynamic observations of land cover. Moreover, it has been verified that using multi-temporal images as input can help to improve the classification accuracy (Bhandari et al., 2012), especially for the land-cover types that have similar spectral characteristics in a single-date image, such as cropland and forest (Schneider, 2012). Landsat can visit the same location every 16 days. However, due to the existence of clouds as well as their shadows, the frequency of clear observations for a specific location is generally much < 16 days. What is worse, the scan-line corrector (SLC) of Landsat 7 malfunctioned in May 2003, which causes wedge-like data gaps in the ETM+ scenes (SLC-off data). These unclear observations (i.e., clouds, cloud shadows, snow/ice, SLC-off data) inevitably influence the availability of multi-temporal images and their mapping accuracy. In order to address this problem, Grinand et al. (2013) used images with the lowest cloud cover in the study period to estimate every-five-year deforestation. The time interval of mapping results obtained by the use of such an approach is restricted to the frequency of the clear satellite observations. Zhu and Woodcock (2014) fit a time-series model by using all the clear observations for each pixel. But this algorithm requires sufficient clear observations to initialize the model. Beckschäfer (2017) produced annual best-available-pixel (BAP) composites, e.g., the least cloudy pixels, from various acquisition dates to cope with the data quality problem, but the experimental results showed that careful selection of images was mandatory in BAP compositing. The previous studies using time-series Landsat imagery are usually based on clear observations, but in Schneider (2012), where dense time series Landsat data, regardless of data quality, were stacked for urban land-cover classification. It was demonstrated that this approach could achieve a better accuracy than simply discarding the unclear datasets. In their experiments, the random forest (RF) classifier was also proved to outperform the maximum likelihood classifier and support vector machine classifier when handling contaminated Landsat data. Some other studies (Breiman, 2001; Zhu et al., 2016b) have also shown that the RF classifier can work effectively, even when the dataset contains some noise. The RF classifier is an ensemble learning method that combines K binary CART trees (Classification And Regression Trees). Each tree is created by selecting a random subset of the features or predictive variables at each node with replacement and the trees grow without pruning (Rodriguez-Galiano et al., 2012; Pelletier et al., 2016). The random and bootstrapped manner enables RF to be a robust and error-tolerant classifier when unclear observations are included (Belgiu and Drăguţ, 2016). However, this method neglects the degree of the reliability between the clear and unclear observations, and may suffer from the negative effect of the unclear ones (Rodriguez-Galiano et al., 2012).

Therefore, it is worth investigating new methods that can focus more on the utilization of both the clear and unclear observations. The improvement of the usability of unclear observations is proved to be conducive to the classification procedure (Schneider, 2012). Training samples from the same class are correlated in both the spectral and temporal dimensions. Therefore, the matrix, constituted by stacking the feature vectors of each training sample from the same class, should be approximately low rank (Chen and Yang, 2014; Cabral et al., 2015). The clear observations of training samples can then be utilized to recover the unclear ones by the matrix completion (MC) algorithm (Candès and Recht, 2009; Cai et al., 2010). To avoid the negative impact of unclear observations in the unlabeled data, the collaborative representation (CR) classifier (Zhang et al., 2012) is an appropriate method to combine with the MC algorithm since its classification hyperplane would not be seriously affected even when only partial observations are available, inferring the CR classifier has the potential to flexibly maintain the discriminative ability when discarding the unclear observations (Li et al., 2014a; Waqas et al., 2013). Therefore, we propose the MCCR classifier that exploits the merits of both MC and CR to cope with poor-quality LTS data classification.

In the MCCR classifier, the clear observations are utilized to recover the unclear ones by the MC algorithm and to then construct the CR classifier. It should be noted that the error-recoverable manner of the MCCR classifier is quite different from the error-tolerant way of the RF classifier. Co-training is a semi-supervised learning paradigm where two basic classifiers are iteratively retrained with the additional semi-labeled samples based on the predictions of either classifier of unlabeled samples (Blum and Mitchell, 1998; Xu et al., 2012). The efficacy of co-training classification method has also been validated in remote sensing imagery, including multispectral image, hyperspectral image, and high spatial resolution image (Persello and Bruzzone, 2014; Zhang et al., 2014). The predictions of RF and MCCR in the unlabeled samples could be complementary to each other, making it possible to exploit the co-training paradigm to provide a better decision than each separate classifier (Zhang and Zhou, 2011; Zhu et al., 2016a). Therefore, we further embed the two classifiers into the co-training paradigm and develop a novel approach (CotrRM), which has the potential to enhance the classification. To the best of our knowledge, this is the first time that an error-recoverable approach with a co-training scheme has been explored to handle contaminated data classification of LTS imagery.

Specifically, in this study, we attempt to investigate the following two research questions:

1) Can the MCCR classifier effectively deal with contaminated data in LTS imagery classification?
2) Is it possible to fuse the MCCR and RF classifiers in order to further raise the performance of their individual use?

2. Study site and datasets

In our study, the city of Wuhan was chosen as the study site (Fig. 1) as it is a typical city that has experienced high-speed development and urbanization in recent decades. Wuhan is the capital of Hubei province, and had a population of over 10 million in 2012 (Wuhan Municipal Statistics Bureau, 2013). The mean annual temperature ranges from 15.8 °C to 17.5 °C (Han et al., 2009), with annual average rainfall of 1050 mm to 2000 mm (Wang et al., 2015). Wuhan lies in the middle reaches of the Yangtze River, and its unique locational characteristics have made it one of the biggest metropolises in central China.

The city of Wuhan is almost entirely covered by the Landsat scene of WRS-2 Path 123 and Row 39, except for about 5% of the study area in the north of Huangpi and Xinzhou districts, as shown in Fig. 1. Since our objective was to investigate a classification method, we selected only this scene to represent the study area. All the available Level 1 Terrain (Corrected) (L1T) Landsat 5,7, and 8 surface reflectance (SR) products acquired in 2000, 2005, 2010, and 2015 with < 60% unclear observations (i.e., pixels with clouds, cloud shadows, snow/ice, and SLC-off data) were downloaded from the U.S. Geological Survey (USGS). Fig. 2 demonstrates the date distribution of the selected images for each year. The Fmask algorithm with its default setting was applied to each image to detect clouds, cloud shadows, and snow/ice. The average producer's accuracy of Fmask in cloud detection is reported to be 92.1% and user's accuracy is as high as 89.4% (Zhu and Woodcock, 2012). This Fmask layer and SR product can provide us whether the observation is clear or not.

Similar to Ying et al. (2017), in this study, we selected time-series Landsat images within a specific given year to produce an annual land-cover map, which can characterize the temporal characteristics for the