ISPRS Journal of Photogrammetry and Remote Sensing 141 (2018) 237–251

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

Exploring geo-tagged photos for land cover validation with deep learning

Hanfa Xing[a,b,c], Yuan Meng[c,*], Zixuan Wang[c], Kaixuan Fan[c], Dongyang Hou[c,*]

[a] School of Geography, South China Normal University, Guangzhou, Guangdong, China
[b] Guangdong Research Center for Smart Land, School of Geography, South China Normal University, Guangzhou, Guangdong, China
[c] College of Geography and Environment, Shandong Normal University, Jinan, Shandong, China

ARTICLE INFO

Keywords:
Land cover
Accuracy assessment
Crowdsourced photos
Convolutional neural network
Sample classification

ABSTRACT

Land cover validation plays an important role in the process of generating and distributing land cover thematic maps, which is usually implemented by high cost of sample interpretation with remotely sensed images or field survey. With an increasing availability of geo-tagged landscape photos, the automatic photo recognition methodologies, e.g., deep learning, can be effectively utilised for land cover applications. However, they have hardly been utilised in validation processes, as challenges remain in sample selection and classification for highly heterogeneous photos. This study proposed an approach to employ geo-tagged photos for land cover validation by using the deep learning technology. The approach first identified photos automatically based on the VGG-16 network. Then, samples for validation were selected and further classified by considering photos distribution and classification probabilities. The implementations were conducted for the validation of the GlobeLand30 land cover product in a heterogeneous area, western California. Experimental results represented promises in land cover validation, given that GlobeLand30 showed an overall accuracy of 83.80% with classified samples, which was close to the validation result of 80.45% based on visual interpretation. Additionally, the performances of deep learning based on ResNet-50 and AlexNet were also quantified, revealing no substantial differences in final validation results. The proposed approach ensures geo-tagged photo quality, and supports the sample classification strategy by considering photo distribution, with accuracy improvement from 72.07% to 79.33% compared with solely considering the single nearest photo. Consequently, the presented approach proves the feasibility of deep learning technology on land cover information identification of geo-tagged photos, and has a great potential to support and improve the efficiency of land cover validation.

1. Introduction

Land cover is an indispensable variable of biophysical materials on the earth surface (Chen et al., 2017; Fritz et al., 2017), and it has served as an essential variable in environmental monitoring, the management of natural resources, urban planning and many other applications (Feddema et al., 2005; Foley et al., 2005). The land cover maps are usually produced by the automatic classification of remotely sensed images (Natya and Rehna, 2016). Because of landscape variability and the uncertainty of map production methods, validating the accuracy of land cover map production is of vital importance in the process of map generation (Foody, 2002; Yang et al., 2017). Conventionally, land cover maps are validated by visual interpretations, in which ground truth is selected and classified based on remotely sensed images or field surveys and then compared with land cover maps to obtain validation accuracy (Strahler et al., 2006). However, evaluating land cover maps by classifying all samples manually is a complex, laborious process, especially for areas that are not reachable or interpretable (Fonte et al., 2015).

Increasingly available geo-tagged photos generated from crowdsourced data, such as photos from Geo-Wiki, Panoramio and Flickr, have been utilised widely in land cover/land use applications. Leung and Newsam (2012) utilised bag of visual words (BoW) and probabilistic latent semantic analysis (pLSA) to understand image and text features in geo-tagged photos for land use classification. In terms of land cover analysis, a study proposed by Estima et al. (2014) extracted useful information for land cover classification through the expert classification of Flickr photos. In addition, Oba et al. (2014) improved the classification efficiency of land cover by automatically extracting the image features and textual information of photos. As the aforementioned studies have revealed the usability of geo-tagged photos, efforts have been made to incorporate these photos into land cover accuracy assessment. A programme named the Degree Confluence Project provides geo-tagged photos at all of the degree confluences, which helps the validation process of existing land cover maps (Foody

* Corresponding authors.
E-mail addresses: xinghanfa@sdnu.edu.cn (H. Xing), mengyuan@stu.sdnu.edu.cn (Y. Meng), wangzixuan@stu.sdnu.edu.cn (Z. Wang), fankaixuan@stu.sdnu.edu.cn (K. Fan), houdongyang1986@cumt.edu.cn (D. Hou).

https://doi.org/10.1016/j.isprsjprs.2018.04.025
Received 7 December 2017; Received in revised form 28 April 2018; Accepted 30 April 2018

and Boyd, 2013; Iwao et al., 2006). Moreover, photos from Flickr and Panoramio have also been investigated with both distribution and quality evaluations for land cover assessments (Estima and Painho, 2013b, 2014). Despite these progresses, challenges still remain in applying geo-tagged photos for land cover validation. On the one hand, strategies are required to effectively extract information from geo-tagged photos for their utilisation in accuracy assessments (Strahler et al., 2006). On the other hand, due to the spatial heterogeneity of geo-tagged photos, approaches of sample selection validation and classification should be proposed by considering the image features and spatial distributions of photos.

For geo-tagged photo classification, visual interpretations by volunteers have been utilised in previous studies. Antoniou et al. (2016) proved the usability of photos in land cover applications by visually interpreting geo-tagged photos from Flickr, Panoramio and Geograph. On the other hand, they also discussed the high time requirements and difficulties associated with manual classification, as well as the feasibility of proposing automatic approaches. This hypothesis has been supported and proved by other studies. Leung and Newsam (2015) and Sitthi et al. (2016) extracted colour, edge and other features from photos to build an automatic photo classification model. Furthermore, deep learning technologies reveal significant advances in the field of image recognition. Specifically, the convolutional neural network (CNN) (Fukushima and Miyake, 1982) has become the most promising algorithm and has resulted in satisfactory performances in a range of image classification problems. Handwritten digital recognition has made significant achievements by learning high-level features from the MNIST dataset (available at http://yann.lecun.com/exdb/mnist/) (LeCun et al., 1998). In addition, CNN model training based on the ImageNet (available at http://www.image-net.org/) and Places (available at http://places.csail.mit.edu/) databases have improved the accuracy and efficiency of object and scene recognition (Russakovsky et al., 2015; Zhou et al., 2014). The state-of-the-art deep learning technology of photo recognition offers an alternative opportunity for the automatic classification of ground information. In fact, studies proposed by Zhu and Newsam (2015) and Xu et al. (2017) have investigated the usability of different CNN models for the identification of land cover classes from geo-tagged photos. Although deep learning provides an efficient approach for extracting land cover information, the classified photos are still unable to be directly applied to validation. As classification probabilities are usually obtained for each geo-tagged photo, selecting and classifying samples for validation by considering both photo locations and classified probabilities remains a problem.

For sample selection, strategies are usually considered to generate unbiased and representative samples for validation (Congalton and Green, 2008; Tong et al., 2011; Xie et al., 2015). However, the uncertainty and unreliability of the assessed accuracy may occur when samples are constructed with variable methods, which may lead to a more time-consuming and expensive process of selecting high-quality reference data (Fonte et al., 2015). When selecting samples based on crowdsourced data, including geo-tagged photos, for accuracy assessment, current studies mainly focus on the following two aspects. On the one hand, geo-tagged photos from pre-defined sample locations have been collected in several studies to validate existing land cover maps (Foody and Boyd, 2013; Iwao et al., 2006). As samples were selected regardless of the photo distribution, more efforts were required to identify those samples for which effective photos were difficult to obtain. On the other hand, pixel-by-pixel comparisons were performed in many of the studies, with the total data as validation samples (Hou et al., 2015; Xing et al., 2017a). As a result, too many inefficient samples were selected, especially in regions with a high density of photos. Accordingly, effective sample selection strategies for existing photos are inevitable. However, sample selection cannot be fully achieved by conventional sampling design methods, as the methods usually select samples based on land cover variety and thus ignore photo availability (Olofsson et al., 2012, 2014; Stehman, 2009). In fact, sample selection can be extremely affected by the spatial distribution of geo-tagged photos, because accuracy is more convincible in regions containing effective photos for validation. Consequently, sample selection, especially sample allocation, should be applied based on the distribution of geo-tagged photos.

For sample classification, the land cover class should be determined for each location where samples are selected. The identified land cover class in each sample plays a vitally important role in validation, as it is considered as ground truth and determines the quality of the existing land cover maps. In the conventional validation process, it is usually regarded as an interpretation issue, in which land cover information corresponding to each sample is directly extracted through remotely sensed images or field surveys (Lillesand et al., 2014; Tong et al., 2013). Meanwhile, land cover information extracted from crowdsourced data, including geo-tagged photos, can facilitate the sample classification process. Currently, approaches proposed by existing studies mainly concentrate on classifying independent land cover samples based on isolated crowdsourced data, such as the nearest photo to one sample, regardless data distributions (Hou et al., 2015; Xing et al., 2017a). However, because of the spatial heterogeneity of land cover landscapes, samples are usually unable to be correctly delineated based on limited geo-tagged photos. For example, a sample in a water body may be surrounded by several geo-tagged photos which are located in artificial surfaces. Given this tendency for sample misclassification, the distance between the sample and the photos should be taken into consideration to measure the confidence of the classification result. Moreover, image features also indicate the probability of one land cover class, as the surrounding photos in artificial surfaces may reflect the existence of water bodies. Accordingly, both the image features and photo distribution are significant indicators for sample classification.

The aforementioned three stages indicate the major challenges in employing geo-tagged photos for land cover validation; the first challenge is constructing an automatic photo classification model and extracting effective information for the accuracy assessment, the second challenge is selecting effective samples via fitting the geo-tagged photo distribution, and the third challenge is sample classification based on land cover information extracted from the photos. To address these issues, this paper proposes a novel method for land cover validation using geo-tagged photos. The remainder of the paper is organised as follows. Section 2 elaborates the proposed framework of employing geo-tagged photos for land cover validation. The experiment performance with Flickr photos collected in western California is described in Section 3. Section 4 discusses the quality of the geo-tagged photos, the utilisation of the CNN models, photo collection and classification with more land cover classes, the sample selection strategy based on photo distribution, the time gap between the map production and photo selection, and building a common validation database with geo-tagged photos. Section 5 concludes this study and presents the future work of our research.

2. Methodology

The exploration of utilising geo-tagged photos for land cover validation is split into three processes (shown in Fig. 1). First, a photo classification model is built using deep learning to automatically classify the land cover of geo-tagged photos. The probabilities of the land cover classes are calculated in the model and are considered as the weights of classification in each photo. Second, a sample selection strategy considering the photo distributions is proposed to determine both the sample size and allocation in the next step. In the third stage, the distance between the samples to the geo-tagged photos is extracted and utilised to calculate weights of distance using an inverse distance weighting (IDW) algorithm. Both weights of classification and weights of distance are involved in sample classification. These samples are utilised to assess the accuracy of land cover maps.