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
Available online 16 May 2018
0924-2716/ © 2018 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights reserved.

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.

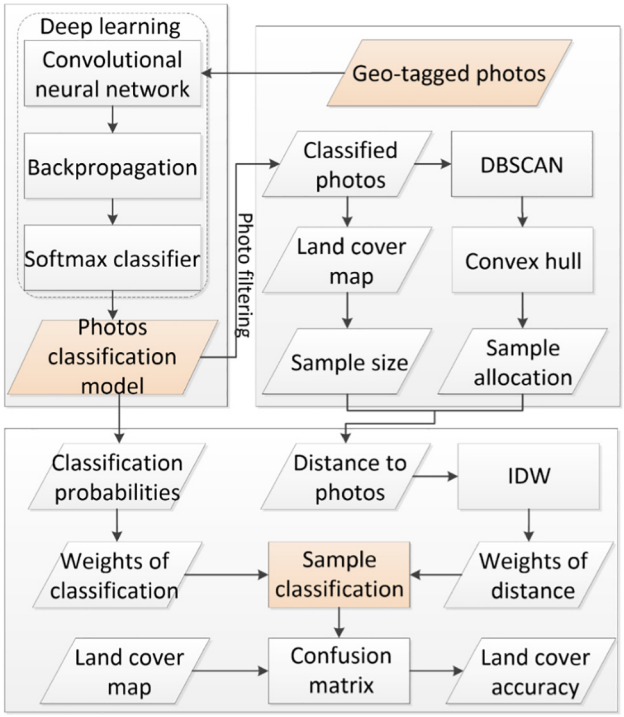


Fig. 1. Methodology workflow for land cover validation using geo-tagged photos.

2.1. Geo-tagged photo classification using the CNN model

An automatic classification model using deep learning is proposed to classify the massive amounts of geo-tagged photos for land cover validation. As the VGG-16 network shows a satisfactory performance in image recognition (Dodge and Karam, 2016), it is utilised to construct the geo-tagged photo classification model. Moreover, other CNN models and their influence on the validation results are discussed in Section 4.2.2.

The overall model framework for geo-tagged photo classification is shown in Fig. 2. Two stages are proposed, including the VGG-16 network construction and model training. Basically, the VGG-16 network is built based on the architecture proposed by Simonyan and Zisserman (2014), and geo-tagged photos with land cover classes are utilised for model training. Based on this model, probabilities corresponding to land cover classes in each geo-tagged photo are calculated. These probabilities are considered as the weights of classification for the sample classification utilised in Section 2.3.

For the VGG-16 network construction, the basic architecture of the VGG-16 network begins with 13 convolutional layers with a padding of 1 for the 3 * 3 pixel window and 5 max-pooling layers performed with a padding of 2 over the 2 * 2 pixel window. It is followed by 3 fully connected layers, in which the first two fully connected layers have 4096 channels, and the channel number in the third layer is determined by the land cover classes that are to be classified. ReLU, which represents the rectified linear function, is used as the activation function of the network. In addition, a softmax classifier is embedded as the final layer of the classification model to predict the land cover class probabilities, defined as follows:

$$P(f_{ik}) = \frac{exp(f_{ik})}{\sum_{k=1}^{K} exp(f_{ik})} \quad (1)$$

where f_{ik} represents the input features of the ith photos belonging to the kth class, while $\sum_{k=1}^{K} exp(f_{ik})$ represents the sum of high-dimensional features that can be classified into each class. Specifically, f_{ik} is a 4096-dimensional vector. $exp(\cdot)$ refers to an expectation function. The softmax classifier is equipped with the loss function, which measures the discrepancy between the predicted class and the labelled one based on the sum of cross entropy. As a result, the probabilities of the ith photo belonging to all land cover classes are defined as $\boldsymbol{P}(f_{ik}) = (P(f_{i1}), P(f_{i2}), P(f_{i3}) \cdots P(f_{ik}) \cdots P(f_{iK}))$, $k \in [1, K]$, where each probability can be calculated via the softmax classifier and utilised as weight of classification, one of the indicators in sample classification.

For model training, all the geo-tagged photos are first resized to 224×224 pixels to fit the pre-trained weights of the VGG-16 model, so the rescaled photos can be used effectively for both model optimisation and land cover prediction. Based on the constructed network, the initialisation of the network weights is proposed by sampling the weights from the normal distribution with mean and variance values of 0 and 0.1, respectively. To optimise the network weights, we adopt back-propagation (LeCun et al., 1989) for both the convolutional layers and fully connected layers. In particular, mini-batch gradient descent, in which the batch size is set to 32 and learning rate is initially set to 0.01, is utilised to minimise the loss function. The number of training iterations is set to 35,000. Moreover, to prevent overfitting, dropout regularisation (Srivastava et al., 2014) is conducted for the first two fully connected layers, with the dropout ratio set to 0.5.

Based on the photo classification model, a photo filtering process is undertaken considering the calculated probabilities before photos are utilised for sample selection and classification. In particular, photos whose maximum probabilities of land cover are lower than a specific threshold are removed from the study. Moreover, photos containing more than one probability higher than this threshold are also excluded from the datasets. This process prevents misclassified photos containing more than one land cover class and keeps the most reliable photos for land cover validation. Threshold selection and photo filtering process evaluation are proposed in Section 3.2.

2.2. Sample selection based on geo-tagged photo distribution

Regarding to the sample selection, both sample size and allocation should be determined. Sample size represents the total number of samples, whereas allocation is the assignment of samples to each land cover class.

Regarding sample size, it is important to ensure that the utilised size is proper, as sizes that are too large or small may lead to insufficient accuracy on the study area. Currently, a variety of approaches have been proposed to estimate the required sample size and are generally based on the empirical determination of sample size (Mayaux et al., 2006; Scepan, 1999), confidence intervals and significance levels (Foody, 2009; Wulder et al., 2006) or acceptance sampling model (Duarte and Saraiva, 2008). Empirical determination is widely used in global land cover datasets, which may lead to biases in regional-scale assessments. Acceptance-sampling-based sample size estimation is a much more complex process because of the producer and consumer risks and the inspection cost (Tong et al., 2011). In terms of considering confidence intervals and significance levels, the approach is based on the proportion of the correctly allocated samples with a certain degree of precision, which has the advantage of being established on statistical theory and facilitating the fundamental requirements of sampling design. Thus, this approach is applied to estimate sample size in accordance to Eq. (2) (Foody, 2009):

$$N = \frac{z_{\alpha/2}^2 P(1-P)}{h^2} \quad (2)$$

where N is the sample size, h is the half width of the desired confidence interval (here, $h = 0.05$) and $z_{\alpha/2}$ is the desired significance level with $\alpha = 0.95$. P is a planning value for the population proportion. The above parameter selection is adopted by research proposed by Schultz et al. (2017).

For the sample allocation, the sampling size in each land cover class is determined based on the approximate regions where geo-tagged

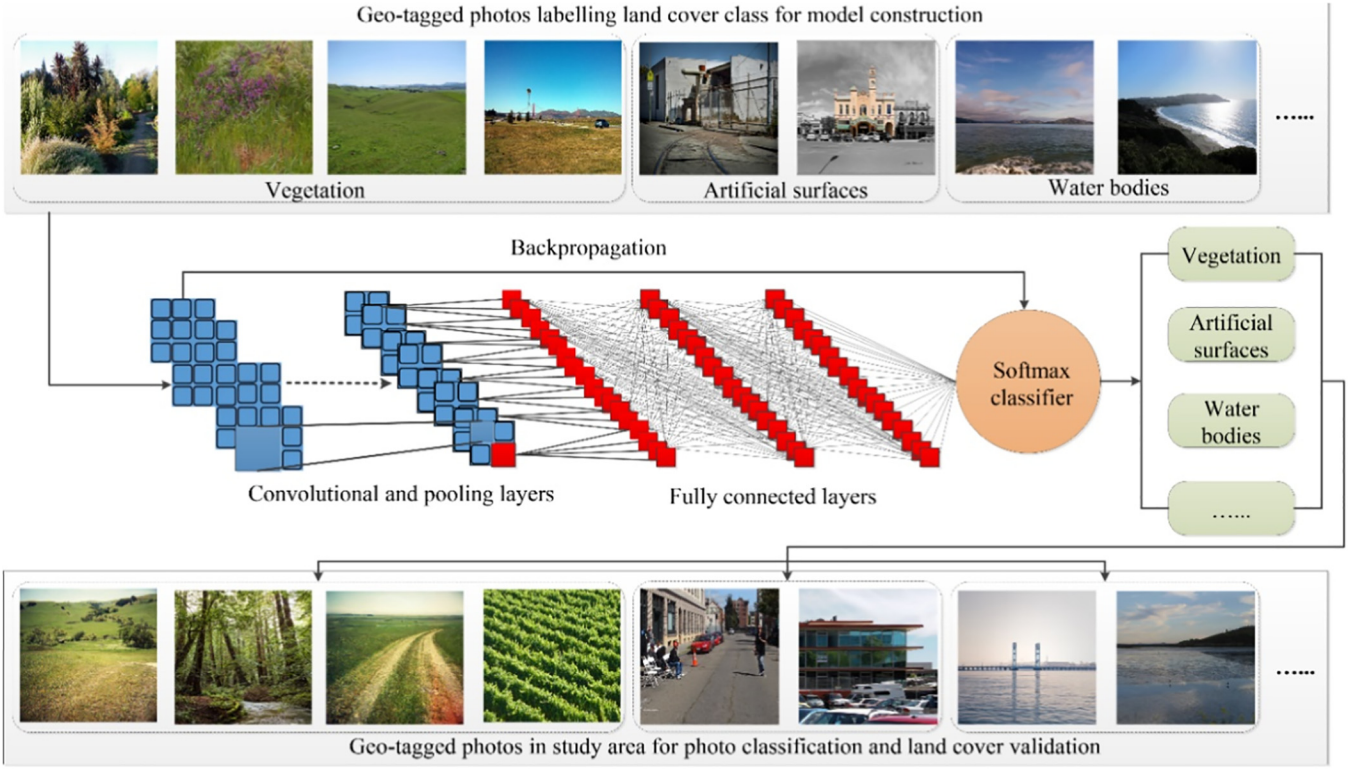


Fig. 2. The overall model framework of CNN model for geo-tagged photo classification

photos are occupied. Specifically, kernel density is applied to estimate the occupied regions of the photos. The sample size in a typical class is defined as:

$$w_k = \frac{A_k}{A} \tag{3}$$

$$N_k = N \bullet w_k \tag{4}$$

where w_k is the proportion of regions classified as one class in the land cover map.A and A_k indicate the total regions where photos are occupied and the regions belonging to a typical class in the land cover maps, respectively. N_k refers to the sample size of each land cover class k.

Based on the determined sample size of each class, samples are supposed to be selected randomly for validation. However, due to spatial heterogeneity, it is necessary to focus on the areas with a high density of photos to reduce the influence of sample misclassification caused by a few misclassified photos. Clustering methods offer the potential to delineate the spatial distribution patterns of geo-tagged photos, such as k-means, hierarchical clustering and spectral clustering. These methods, however, mainly group photos based on the similarity of distance and ignore the photo density in each cluster. Thus, DBSCAN (density-based spatial clustering of applications with noise) (Ester et al., 1996), which aims to delineate clusters with a specific density, is applied to high-density photo clusters. Two parameters are required, including the search radius *Eps* and the minimum number of photos *Minpts* within the search radius which are set as 50 metres and 5, respectively. These clustered photos are then utilised to generate regions by using convex hull. As a result, samples within the determined sample size in each class are selected randomly in these generated regions, which ensures that samples will be classified based on more reliable photos.

2.3. Sample classification with classified geo-tagged photos

Focusing on utilising the selected samples for land cover validation, both the classification results and the distribution of geo-tagged photos should be involved in the sample classification. As shown in Fig. 3, on the one hand, the probabilities of the land cover class in each photo, which are calculated in Section 2.2, are considered as the weights of classification. On the other hand, due to the high spatial heterogeneity of land cover (Chen et al., 2016), the distances between samples and photos are utilised to calculate the weights of distance, which is used to adjust the probabilities that the land cover class in one photo may indicate.

Taking the distance factor into consideration, we adopt an inverse distance weighting (IDW) algorithm to measure the contributions of geo-tagged photos to each sample and define the value of IDW as the weights of distance. IDW is defined as Eq. (5):

$$W_i = \frac{h_i^{-p}}{\sum_{i=1}^{n} h_i^{-p}} \tag{5}$$

where W_i indicates the weight of distance of the ith geo-tagged photo. n refers to the number of photos used to classify the sample. Here, it indicates geo-tagged photos in the same area generated by convex hull. h_i denotes the distance from sample to the ith geo-tagged photo. p refers to the parameter controlling the variation of weights with different distances. Fig. 4 represents the weights of distance with a different parameter p. As 50 metres was set for extracting photos with high density in Section 2.2, weights of distance should be more sensitive within the range of 0-50 metres, and show a tendency of stabilisation beyond 50 metres, which is shown in Fig. 4(c). For $p = 0.01$ in Fig. 4(a) and $p = 0.1$ in Fig. 4(b), weights of distance vary within ranges larger than 50 metres, with range of distances of approximately 200 metres and 150 metres. On the other hand, Fig. 4(d) reveals that the weights of distance could vary greatly within an extremely short distance. Therefore, $p = 0.5$ is adopted in IDW to calculate the weights of distance.

Involving both the weights of classification $P(f_{ik})$ and weights of distance W_i, we modify the weights of land cover classes in each photo as follows:

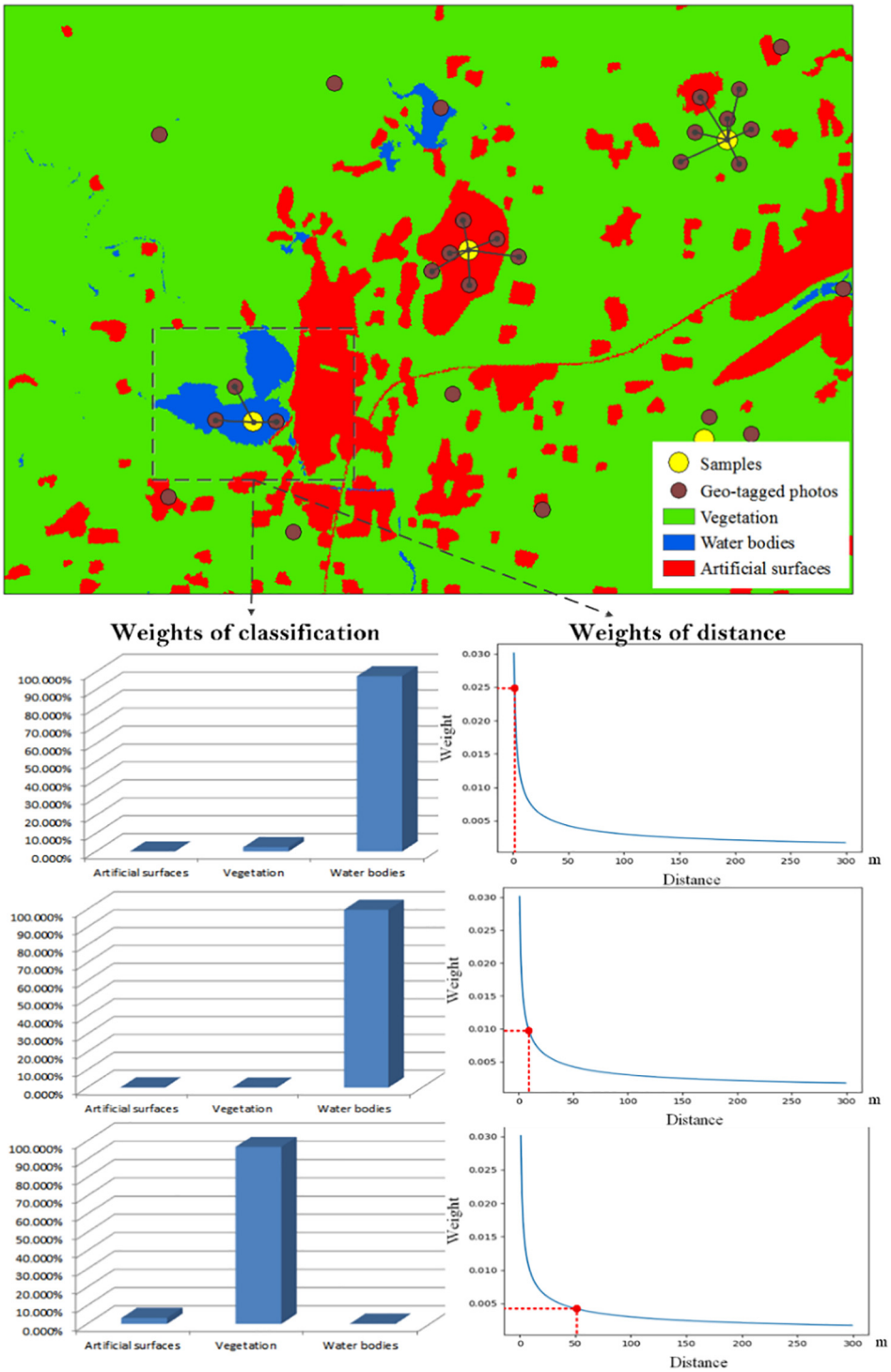


Fig. 3. Sample classification considering the weights of both photo classification and distance.

$$P_{ik} = W_i \times P(f_{ik}) \tag{6}$$

where P_{ik} denotes the modified weights of the kth land cover class in the ith photo.

To assign the weighted geo-tagged photos to the corresponding samples, we compute the probabilities of classes in each sample as Eq. (2):

$$M_{sk} = \sum_{i=1}^{n} P_{ik} \tag{7}$$

where M_{sk} is the probability of belonging to the kth land cover class in the sth sample. n indicates the number of photos used to classify the sth sample. The probabilities of all the land cover classes of the sth sample can be defined as: $\boldsymbol{M}_s = (M_{s1}, M_{s2}, M_{s3} \cdots M_{sk} \cdots M_{sK})$, $k \in [1, K]$. Each sample is classified as the land cover class containing the maximum value of probability, which is then utilised to validate land cover maps.

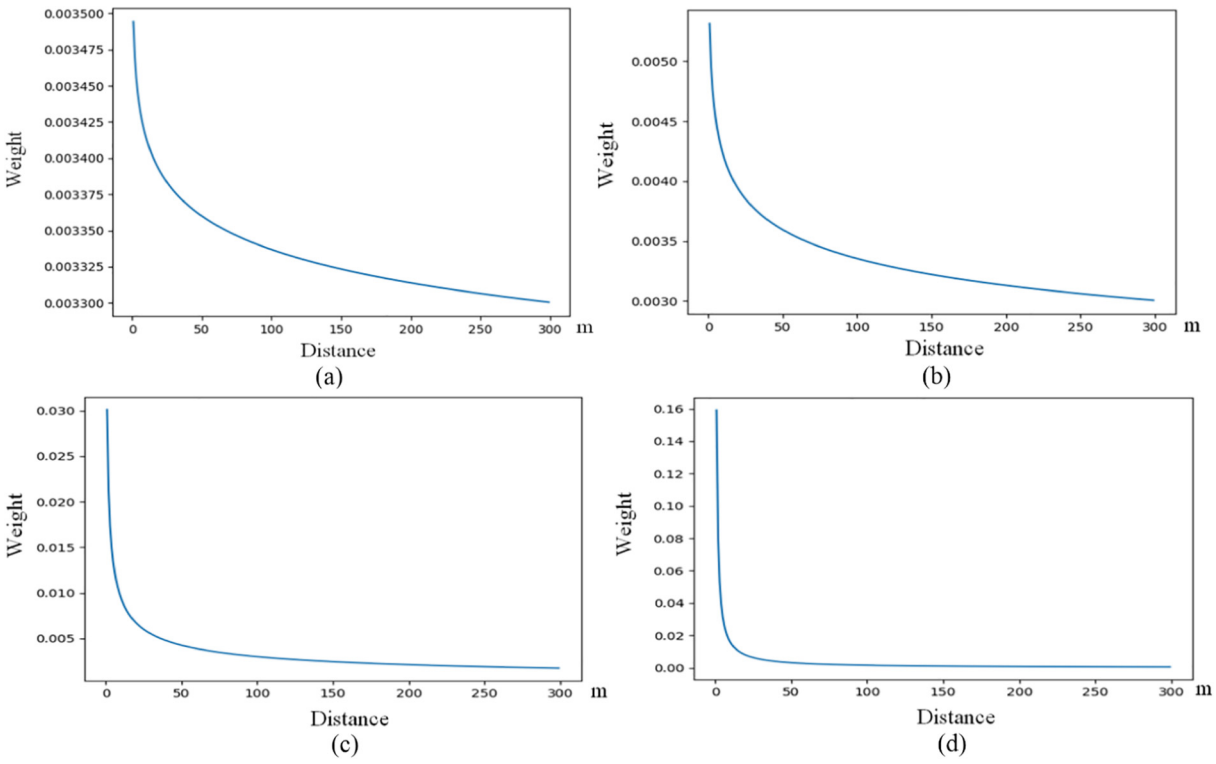


Fig. 4. The weights of distance with different parameter p values; (a) $p = 0.01$, (b) $p = 0.1$, (c) $p = 0.5$, (d) $p = 1$.

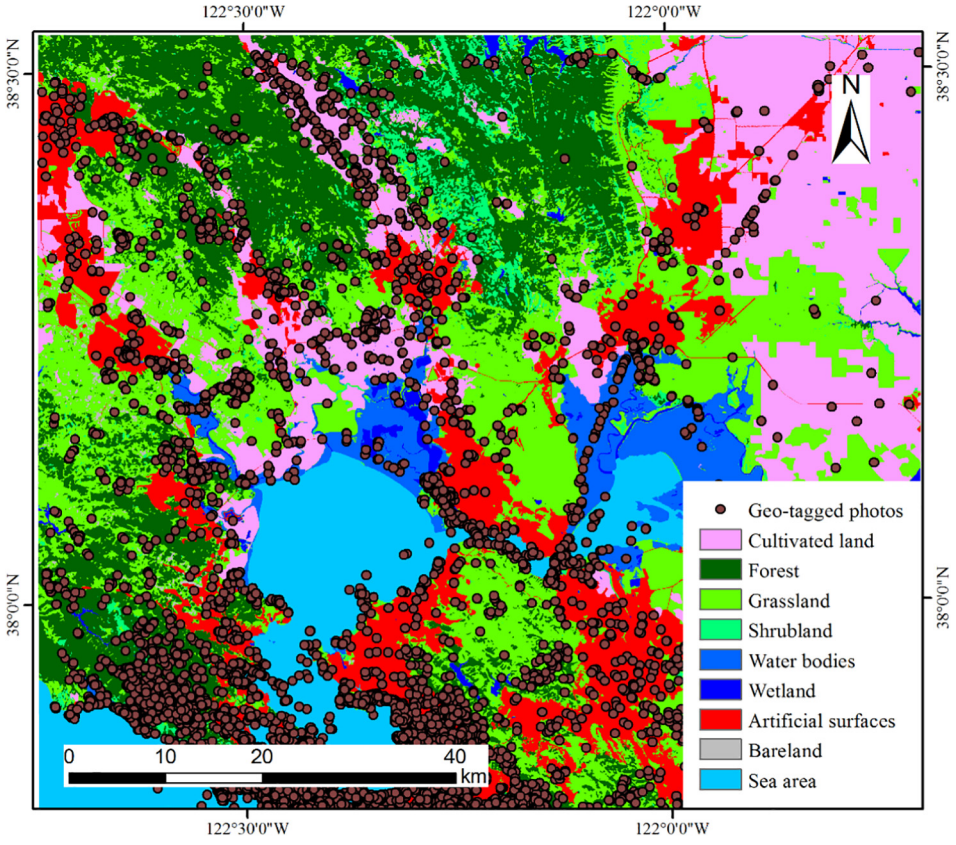


Fig. 5. Study area and geo-tagged photos.

3. Results

3.1. Study area and datasets

The study area, as shown in Fig. 5, is in the western region of California. This area contains variable land cover classes; thus, an accuracy assessment is necessary to validate the land cover maps in this area.

The land cover maps of GlobeLand30, produced by the National Geomatics Centre of China (NGCC), were chosen for validation (Chen et al., 2014). With a resolution of 30 m, 10 land cover classes are classified in a global scale. Two products have been released, GlobeLand30 in 2000 and 2010, and the map in 2010 was selected for the accuracy assessment. In this study, three classes were involved in the validation process, artificial surfaces, water bodies and vegetation. Notably, we included wetland in the water bodies class and merged cultivated land, forest, grassland and shrubland into vegetation, because typical features can hardly be extracted and distinguished with a precise accuracy among these classes based on the currently geo-tagged photos. As a result, the proposed approach was restricted to the validation process of artificial surfaces, water bodies and vegetation. According to previous research (Chen et al., 2015), an overall classification accuracy of over 80% was achieved when conducted by third-party experts. To some extent, this accuracy can be utilised to assess our proposed validation method.

To collect geo-tagged photos for validation, a total of 32367 photos were downloaded from Flickr by searching keywords related to land cover, with artificial surfaces keywords such as “residential”, “building”, “school”, “hospital”, vegetation keywords such as “farmland”, “cropland”, “forest”, “grassland”, “green land”, and water bodies keywords such as “river”, “ocean”, “lake”, “wetland”. Each acquired photo contained basic information, including timestamps, locations, and tags. Since some tags are not the descriptions corresponding to the captured images, only photos with locations and timestamps were used in this study. We divided these photos into two parts, with 10345 global scale photos used to build the classification model and 22,022 photos from the study area used for validation. It should be noted that, while geo-tagged photos were taken from 2010 to 2014, the time gap between the map production and the photos should be considered. Given the detection of land cover changes during this period (details in Section 4.5), only minor changes occurred in the study area. Thus, the major errors caused by this time gap were reduced.

3.2. Geo-tagged photo classification

In the process of geo-tagged photo classification, the VGG-16 model is first constructed and trained using 10,345 photos collected at a global scale. This is followed by classifying 22,022 photos located at the study area based on the VGG-16 model.

Photos for model training were interpreted visually with land cover labels, and typical photos of each land cover class were selected and are shown in Fig. 6. Photos containing built-up areas were considered artificial surfaces. Roads located in other land cover classes were excluded from artificial surfaces, because they cannot meet the requirement of the “minimum spot” in land cover mapping. Vegetation contained forest, grassland, cultivated land, and other types. Water bodies mainly included rivers and lakes occupying the geo-tagged photos.

Based on the photos for model training, the classification accuracy of geo-tagged photos was evaluated through a 10-fold cross-validation. Fig. 7 indicates the validation accuracy during the model construction process within 35,000 iterations. The accuracies varied significantly from 0% to approximately 70% between 0 and around 5000 iterations. It was followed by oscillated variations of accuracies within 70–80%. After around 25000 iterations, a stabilised trend was displayed, resulting in the less than 80% accuracies of the photo classification

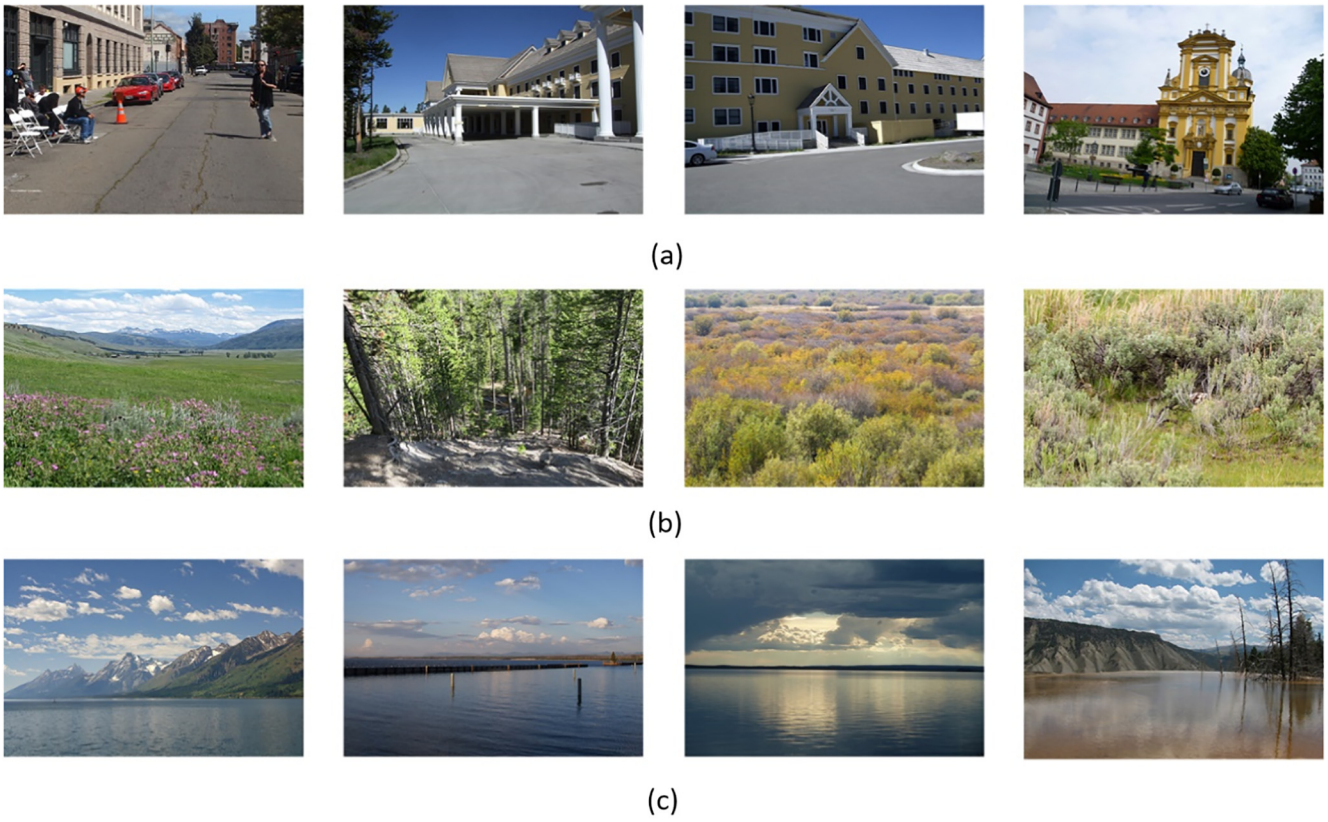


Fig. 6. Geo-tagged photos which are classified as different land cover classes; (a) Artificial surfaces; (b) Vegetation; (c) Water bodies.

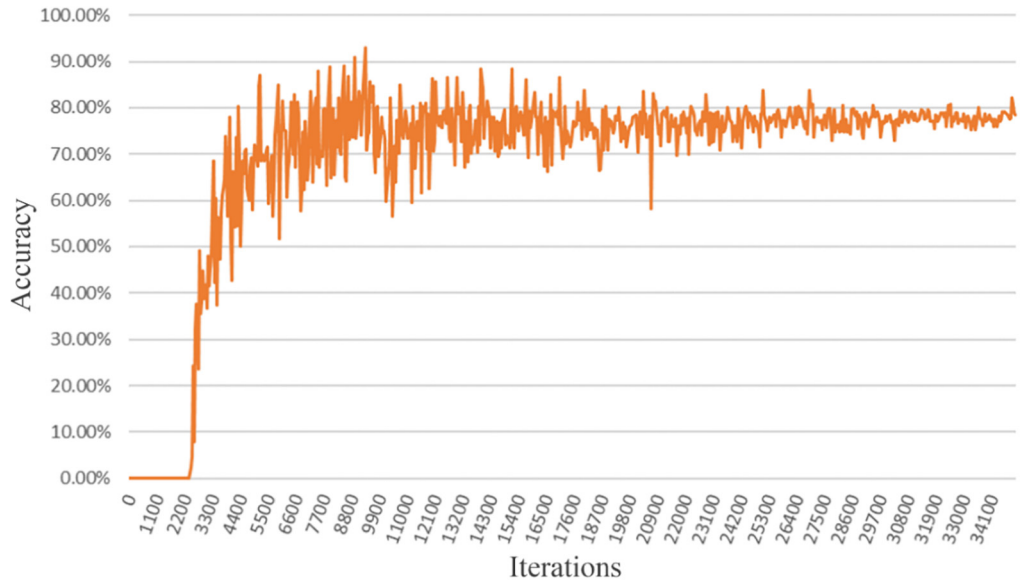


Fig. 7. Validation accuracy of geo-tagged photo classification in each model iteration.

model.

In addition, the classification accuracy for each land cover class is shown in Table 1. The overall accuracy of 76.01% (69.32%, 82.70%) was obtained within a 95% confidence interval. In particular, the producer's accuracies of the artificial surfaces, vegetation and water bodies were 79.64%, 75.74% and 77.74%, respectively. In addition, the user's accuracies corresponding to the three classes were 75.86%, 79.50% and 76.43%. The above analysis indicates a relatively lower accuracy can generate some misclassification results by geo-tagged photos. Thus, effective methods are required to further improve the performance of photo classification.

Utilising the classified geo-tagged photos in model training, the filtering process discussed in Section 2.1 is proposed to improve classification accuracy. To determine the proper threshold for photo filtering, thresholds ranging from 60% to 90% were applied based on the classified photos from the training data, with average accuracies calculated via 10-fold cross-validation. As shown in Table 2, compared with the accuracy of 76.01%% in Table 1, the photo filtering process improved the classification accuracies, which range from 76.76% to 80.54%. In addition, the highest accuracy of 80.54%, which is approximately 4.53% higher than classification accuracy with unfiltered photos, was calculated by utilising the 90% filtering threshold. Although a greater threshold, such as 95%, might result in much higher accuracy, more geo-tagged photos would be excluded, which could lead to insufficient data for land cover validation. Accordingly, a threshold of 90% was selected to filter the classified photos and improve classification accuracy.

Based on the constructed photo classification model, 22,022 photos in the study area were utilised for extracting the land cover information with probabilities. Moreover, by proposing photo filtering with a 90% threshold, 18,798 photos, approximately 85.36% of the total 22,022 photos, remained. The calculated probabilities in these photos are considered as the weights of classification in the process of sample classification in Section 3.4.

3.3. Sample selection

According to the sample size determination in Eq. (2), 179 samples were assigned in the study area. To calculate sample size of each land cover class, a kernel density was applied to estimate the occupied regions of the photos in GlobeLand30. In particular, a 500-metre search radius was set in kernel density, because approximately 97.06% of the total number of geo-tagged photos were distributed within 500 m. As shown in Table 3, 24.58%, 68.16% and 7.26% of all the photos are located in artificial surfaces, vegetation and water bodies of the land cover map, respectively. Accordingly, 44 samples were assigned for validating artificial surfaces, and 122 samples and 13 samples are selected to assess vegetation and water bodies, respectively.

The process of sample selection considering geo-tagged photo distribution is shown in Fig. 8. In particular, Fig. 8(a) displays the

Table 1
Classification accuracy of geo-tagged photos for each land cover class.

Land cover class	Classification accuracy	
Artificial surfaces	Producer's accuracy	79.64%
	User's accuracy	75.86%
Vegetation	Producer's accuracy	75.74%
	User's accuracy	79.50%
Water bodies	Producer's accuracy	77.74%
	User's accuracy	76.43%
Overall accuracy	76.01% (69.32%, 82.70%)[a]	

[a] Data in parentheses are within the 95% confidence interval.

Table 2
Average accuracies of geo-tagged photo classification with different filtering thresholds.

Threshold (%)	Average accuracy (%)	Confidence interval[a]
60	76.76	(70.16%, 83.36%)
65	77.29	(70.62%, 83.96%)
70	77.77	(71.11%, 84.43%)
75	78.34	(71.75%, 84.94%)
80	78.85	(72.37%, 85.37%)
85	79.63	(73.06%, 86.20%)
90	80.54	(74.04%, 87.34%)

* Data in parentheses are within the 95% confidence interval.

Table 3
Number of samples in each class for validating GlobeLand30.

	Artificial surfaces	Vegetation	Water bodies	Total
Area/km^2	324.2	899.01	95.76	1318.97
Percentage	24.58%	68.16%	7.26%	100.00%
Number of samples	44	122	13	179

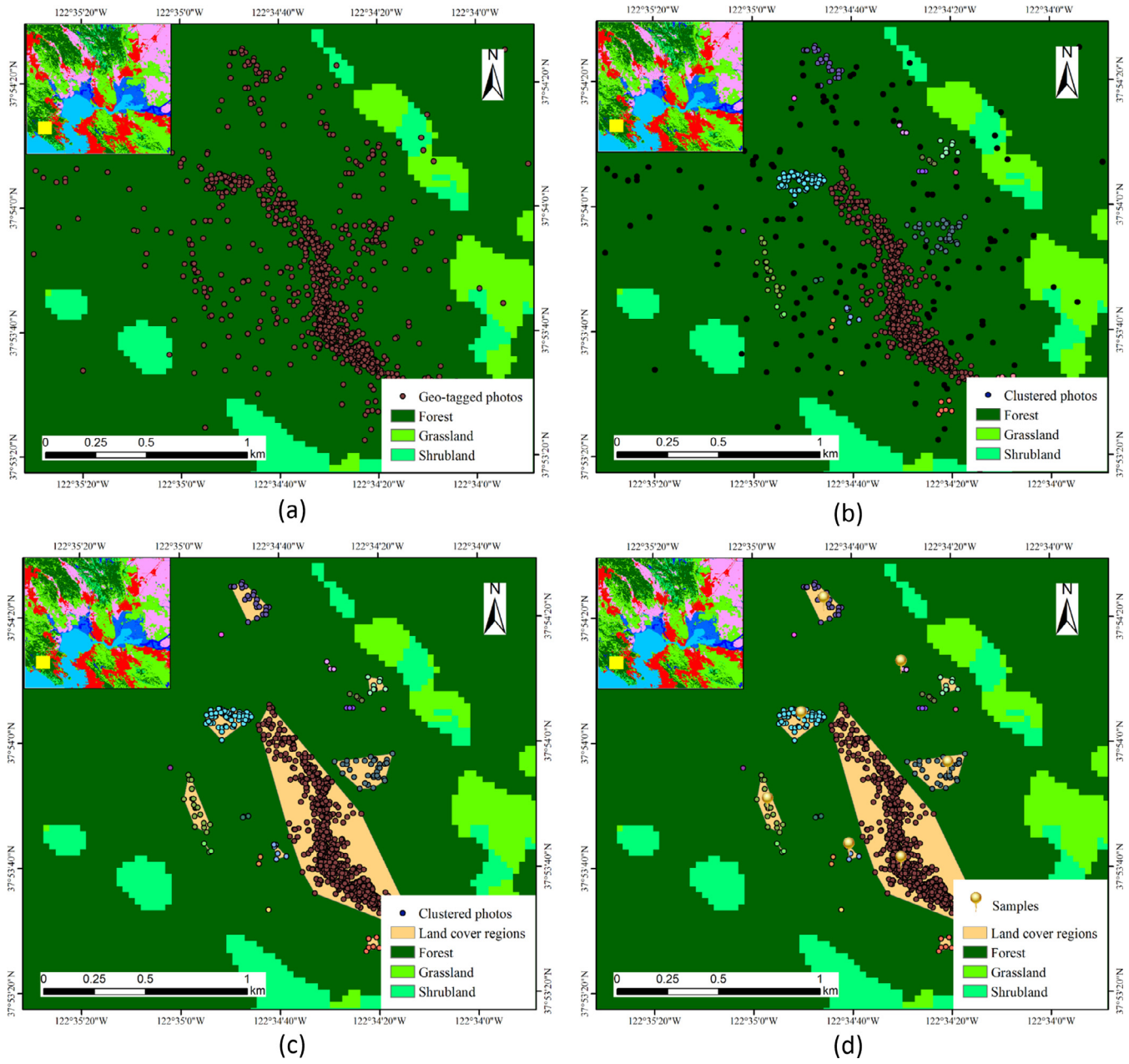


Fig. 8. Process of sample selection in part of the study area; (a) Geo-tagged photos in part of the study area, (b) Clustered geo-tagged photos in part of the study area, (c) Extracted land cover regions in part of the study area, (d) Selection of samples in part of the study area.

distribution of geo-tagged photos in part of the study area. For delineating photos with a high density, the clustered geo-tagged photos are shown in Fig. 8(b). Photos with the same colour were assigned into one cluster, while photos coloured black were considered noise data and were not considered for the land cover region extraction. Fig. 8(c) represents the extracted land cover regions based on the cluster geo-tagged photos. Samples for validation were selected randomly from the extracted land cover regions, which are displayed in Fig. 8(d). Based on the abovementioned process, the selected samples in the study area are shown in Fig. 9.

3.4. Sample classification

Multiple geo-tagged photos located in the generated region in Section 3.3 were involved in sample classification. In particular, the distances between the sample and geo-tagged photos were calculated and then utilised as indicators for determining the weights of distance for each photo. Consequently, sample classification was proposed based on the weights of distance and weights of classification. To ensure the sample quality, these classified samples were assessed via visual interpretation using images from Google Earth. Table 4 represents the accuracy of the sample classification. An overall accuracy of 79.33% was achieved for the 179 samples, with samples classified as artificial surfaces and vegetation containing higher accuracies. The classification accuracy of water bodies was much lower because geo-tagged photos are usually taken along the shore of water bodies instead of inside the water bodies. Fig. 10 displays the distribution of the classified samples.

To further evaluate the strategy of considering multiple geo-tagged photos for sample classification, a comparative experiment was proposed, where only the photos with the shortest distances to each sample were considered. With a value of 72.07%, the overall accuracy based on single classified photos was approximately 7.26% lower than that based on multiple geo-tagged photos (Table 5). Fewer correctly classified samples were obtained for the identification of artificial surfaces,

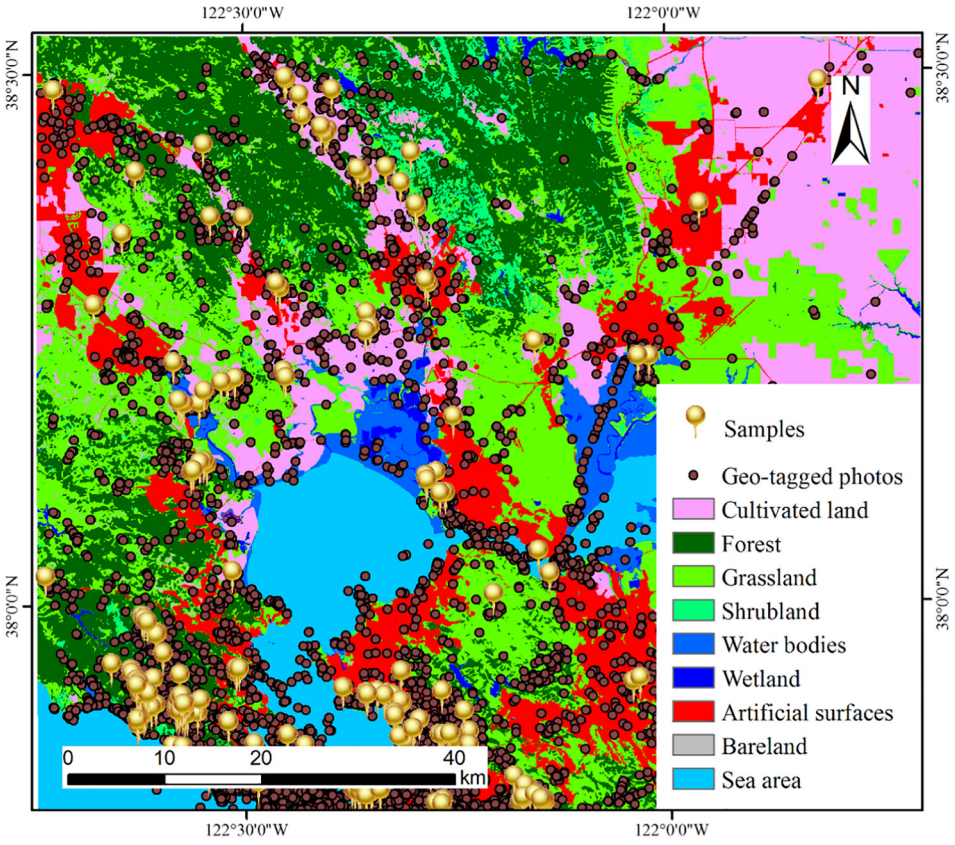


Fig. 9. Samples for land cover validation.

Table 4
Sample classification accuracy considering multiple photos with different distances to each sample.

	Artificial surfaces	Vegetation	Water bodies	Total
Artificial surfaces	32	3	7	42
Vegetation	16	101	2	119
Water bodies	1	8	9	18
Total	49	112	18	179
Overall accuracy	–	–	–	79.33%

vegetation and water bodies. To some extent, this proves the usability of the proposed sample classification approach.

3.5. GlobeLand30 validation

As shown in Table 6, for all 179 samples used for validation, an overall accuracy of 83.80% was obtained, with classified samples taken for validation. Among the three land cover classes, vegetation gave the highest accuracy for both user's accuracy and producer's accuracy, with values of 90.76% and 88.52%, respectively. It was followed by artificial surfaces, with relatively satisfying results of 77.27% and 80.95% for producer's accuracy and user's accuracy, respectively. On the other hand, water bodies achieved much lower accuracies, with a value of 61.54% for producer's accuracy and a value of 44.44% for user's accuracy, and these low accuracies may be caused by the misclassification of samples.

Considering the uncertainty of the classified samples, the land cover map was also validated via visually interpreted samples (Table 7), so there is a guide indicating the accuracy of this map. This assessment results were relatively consistent with the accuracy values achieved using classified samples (Table 6). An overall accuracy of 80.45% was achieved, which is only approximately 3% lower than the result proposed by the classified samples. It should be noted that the assessment of water bodies yielded values of approximately 44% for both user's accuracy and producer's accuracy, similar to the results in Table 6. This proves that the classification accuracy for water bodies is relatively low in GlobeLand30. The inspection of the two assessments indicates that the accuracy obtained by using geo-tagged photos is highly reliable, and it proves that our proposed approach is effective for land cover validation.

4. Discussion

This study represents a novel approach of employing geo-tagged photos for land cover validation with deep learning. Although the proposed approach helps the task of land cover validation, conventional methods of sampling design and classification are still valid and necessary. Several issues are discussed in detail in the following sections.

4.1. The quality of geo-tagged photos

The reliability of crowdsourced data, including geo-tagged photos, has been addressed by recent studies (Zhao et al., 2017). Although geo-tagged photos tend to be widely available on the Internet, in Flickr, Instagram and other websites, not all of these photos correspond to land cover or indicate the typical land cover environment where the photos are located. Thus, given the focus on the massive number of available photos, effective data selection and filtering processes are necessary and inevitable.

Two steps are proposed in our study: the first utilises keywords to search for photos related to land cover, as described in Section 3.1, and the second filters photos considering the probabilities of land cover classification results, as described in Section 3.2. These two processes mainly avoid photos that do not contain land cover or contain an ambiguous land cover environment.

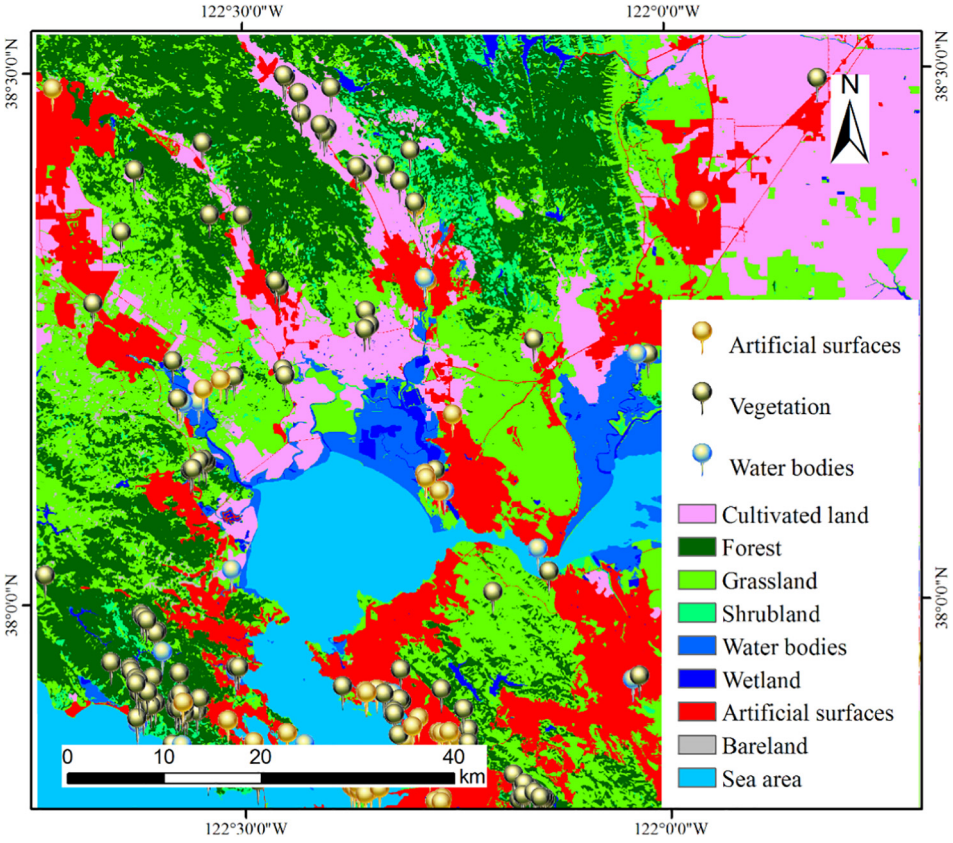


Fig. 10. Classified samples for land cover validation.

Table 5
Sample classification accuracy of the approach using single photos with short distances to each sample.

	Artificial surfaces	Vegetation	Water bodies	Total
Artificial surfaces	25	3	7	35
Vegetation	21	98	5	124
Water bodies	3	11	6	20
Total	49	112	18	179
Overall accuracy	–	–	–	72.07%

Table 6
Land cover validation based on classified samples.

	Artificial surfaces	Vegetation	Water bodies	Total	User's accuracy
Artificial surfaces	34	7	1	42	80.95%
Vegetation	7	108	4	119	90.76%
Water bodies	3	7	8	18	44.44%
Total	44	122	13	179	
Producer's accuracy	77.27%	88.52%	61.54%		83.80%

However, the problem of the current data selection and filtering processes is that photos labelled with land cover related keywords do not entirely indicate land cover environments. For example, a selfie taken in the forest may be labelled "forest", even though the surroundings are not involved in the photo. Moreover, the ambiguity of geo-tagged photos cannot be completely excluded from this study, and focusing on the majority part of this kind of photo may lead to misclassification. For example, a photo located in grasslands likely mostly contains the sky and distant mountains; thus, there is a high probability that the associated photo classification results are not reliable enough

Table 7
Land cover validation based on visually interpreted samples.

	Artificial surfaces	Vegetation	Water bodies	Total	User's accuracy
Artificial surfaces	34	6	4	44	77.28%
Vegetation	14	102	6	122	83.61%
Water bodies	1	4	8	13	44.69%
Total	49	112	18	179	–
Producer's accuracy	69.39%	91.07%	44.44%	–	80.45%

to ensure the quality. It is also important to note that quite a few photos are literally not located where they are taken, leading to more noise data during the photo selection process.

The sample selection strategy proposed in Section 2.2 can reduce the influence of noise data, to some extent. As samples are selected from an area with a high density of photos, the few misclassified photos can be assigned relatively low probabilities. Nevertheless, more effective data selection and filtering methods are still required in future studies.

4.2. The utilisation of the CNN models

4.2.1. Architecture and implementation of the CNN model

This paper utilised the VGG-16 network from CNN models for automatic geo-tagged photo classification. This model was considered due to its satisfactory photo recognition performance and wide applications in many fields. Basically, the model architecture was built following the modelling design proposed by Simonyan and Zisserman (2014). Despite the model reliability, it is possible to investigate other CNN architectures from VGGNet, such as VGG-19, to improve classification accuracy.

In terms of model training, we initialised the network weights by sampling from the normal distribution, with mean and variance values adopted from Simonyan and Zisserman (2014). Strategies of weight initialisation are important, as too large or small weights might lead to insufficient feature extraction from photos. Thus, the evaluation of weight initialisation is necessary in future studies. Moreover, to further optimise the network weights, mini-batch gradient descent with a batch size of 32 and learning rate initially set to 0.01 was involved. The hyperparameters were selected after several experiments with a variety of hyperparameter combinations. However, the regularity of how photo classification accuracy changes with other strategies of hyperparameter optimisation remains an issue; thus, other strategies should be carefully assessed. We will focus on this problem and improve model performance in future studies.

To reduce the overfitting problem, dropout regularisation was considered in model training. However, it did not yield a satisfactory performance in the process of model training, and a relatively lower average accuracy of 76.01% was obtained within the 95% confidence interval (Table 1). To solve this problem, future work will consider the potential of data augmentation to better optimise the classification model.

4.2.2. Impact of the CNN model selection on land cover validation

The utilisation of the CNN model facilitates the automatic classification of geo-tagged photos and further contributes to the sample classification and land cover validation process. Although the accuracy assessment of GlobeLand30-2010 yielded relatively satisfactory results, the performance of different CNN models and their impacts on validation results remain undiscussed. Considering this issue, an additional two CNN models, including ResNet-50 (He et al., 2016) and AlexNet (Krizhevsky et al., 2012), are involved in the proposed approach. Accordingly, analysis will focus on the influence of three CNN models, namely, VGG-16, ResNet-50 and AlexNet, on the changes of photo classification, sample classification and validation results.

For the three separately trained CNN models, Table 8 displays the average geo-tagged photo classification accuracies based on 10-fold cross-validation within the 95% confidence interval. Accuracies range from 76.01% to 81.21%, with ResNet-50 and AlexNet showing slightly better performances than VGG-16. This indicates that both ResNet-50 and AlexNet are more efficient in classifying geo-tagged photos. And thus, impacts of model choice on sample classification and land cover validation results were investigated based on the differences in CNN model accuracies.

After proposing sample selection and classification, the accuracies of the sample classification results of different CNN models were determined and are shown in Table 9. Overall accuracies of 79.33%, 83.24% and 82.12% were obtained using photos classified by VGG-16, ResNet-50 and AlexNet, respectively. These results are consistent with the accuracies of the three CNN models in Table 8, which reveals that the CNN model with better performance has the potential to enhance sample classification results.

Table 8
Average geo-tagged photo classification accuracies of three CNN models.

Deep learning model	Iterations	Average accuracy	Confidence interval[a]
VGG-16	35,000	76.01%	(69.32%, 82.70%)
ResNet-50	35,000	81.21%	(74.92%, 87.49%)
AlexNet	35,000	79.41%	(72.50%, 86.32%)

[a] Data in parentheses are within the 95% confidence interval.

Table 9
Sample classification accuracies based on photos classified by three CNN models.

Deep learning model	Number of samples	Accuracy (%)
VGG-16	179	79.33
ResNet-50	179	83.24
AlexNet	179	82.12

Table 10
Land cover validation based on samples by three CNN models.

		VGG-16 (%)	ResNet-50 (%)	AlexNet (%)
Artificial surfaces	Producer's accuracy	77.27	73.21	78.57
	User's accuracy	80.95	87.23	83.02
Vegetation	Producer's accuracy	88.52	92.45	91.51
	User's accuracy	90.76	90.74	91.51
Water bodies	Producer's accuracy	61.54	88.24	82.35
	User's accuracy	44.44	62.50	70.00
Overall accuracy		83.80	86.03	86.59

Land cover validation was further proposed using classified samples. Table 10 shows the validation results, for each CNN model, of GlobeLand30-2010. Overall accuracies of 83.80%, 86.03% and 86.59% were calculated based on the VGG-16, ResNet-50 and AlexNet models, respectively. Notably, the accuracy based on the visual interpretation of the study area was approximately 80% (Table 7), and the use of ResNet-50 and AlexNet for validation did not result in prominent advantages over the use of VGG-16. Furthermore, the producer's and user's accuracies revealed that the validation of water bodies shows significant discrepancies between VGG-16 and the other two models, as accuracies based on VGG-16 were much lower than those based on ResNet-50 and AlexNet. On the other hand, the artificial surfaces and vegetation accuracies were not notably different among the models.

The above analysis shows the impact of different CNN models on land cover validation results. It indicates that, although one CNN model may have slightly better performance results based on the classification accuracies of both geo-tagged photos and samples, it might not demonstrate a satisfactory accuracy in the validation process. The conducted conclusion might be due to the following three reasons. First, both ResNet-50 and AlexNet do not show distinguished performances, with only 3.4% and 5.2% higher than the accuracy of VGG-16. CNN models with much higher accuracies might improve validation process with efficiency. Second, the quality of geo-tagged photo shows impact on both photo classification and land cover validation, which has been discussed in Section 4.1. Third, apart from CNN model construction, validation result can be also influenced by strategies of sample selection and classification.

4.3. Photo collection and classification with more land cover classes

In this study, three classes, namely, vegetation, water bodies and artificial surfaces, were considered for GlobeLand30-2010 validation. In particular, vegetation contains four land cover classes, cultivated land, forest, grassland and shrubland, whereas wetland is merged into water bodies. This reclassification process was proposed because more detailed land cover classes cannot be distinguished precisely from the existing geo-tagged photos.

This limitation promotes further investigations about geo-tagged photo collection and classification involving cultivated land, forests, grassland, shrubland and wetland. For the geo-tagged photo collection, two effective methods can be applied. The first establishes a corpus containing a variety of keywords related to the land cover classes, which can facilitate the photo searching and collection process. The second approach mainly focuses on selecting photos from land cover related websites. For example, See et al. (2014) proposed a gaming approach in Geo-wiki to capture cropland via geo-tagged photos, while

the Degrees Confluence Project provided forest related photos with detailed latitude and longitude information (Foody and Boyd, 2013). This reveals the potential to integrate geo-tagged photos from land cover related websites in further validation studies. Based on the collection of geo-tagged photos, the construction of the classification model will be further improved. If more features are extracted and determined from photos of detailed land cover classes in the model, the influence of the overfitting problem may be reduced, thus increasing the geo-tagged photo classification accuracy to some extent.

4.4. Sample selection strategy based on photo distribution

Regarding the massive number of distributed geo-tagged photos, we propose a sample selection strategy to determine effective samples for land cover validation and quality evaluation. Unlike the traditional sampling design, which considers the spatial balance and spatial heterogeneity of land cover datasets (Chen et al., 2016; Olofsson et al., 2012), our proposed method mainly concentrates on solving the difficulty of selecting limited samples among photos with a high density. Accordingly, the distribution of geo-tagged photos is our first concern in sample selection.

However, traditional land cover application methods are necessary in regions where crowdsourced data, including geo-tagged photos, are rare or unavailable (Hu et al., 2016; Johnson et al., 2017; Liu et al., 2017). On the one hand, the impacts of a variety of sample size estimation approaches on validation were not investigated in this study. On the other hand, as a properly designed probabilistic sampling design, which is generally proposed in accuracy assessments (Stehman et al., 2012), was not considered in our study, sample biases may occur and thus not represent the larger population.

To combine traditional methods with geo-tagged photos for validation purposes, several questions should be addressed. First, the strategy of sample selection should be reconsidered, with samples allocated based on both land cover variation and photo distribution. In this case, many samples can be automatically classified instead of being interpreted via remotely sensed images or field surveys. Second, the sparsely distributed photos considered noise data in this study still have the potential to assist the traditional sample classification method with visual interpretation. These two issues will be addressed in further studies.

4.5. Time gap between map production and photo selection

Technically, samples used for validation should be from the same year as the map production. As discussed in the current study (Estima and Painho, 2013a), landscapes may change over a long period, leading to land cover variation and uncertainties in different years.

In this study, photos taken from 2010 to 2014 were selected for the validation of GlobeLand30-2010. Considering the uncertainties related to land cover in the time gap between the map production and photo collection, a quantitative change detection was proposed to measure the changes between 2010 and 2014. In particular, land cover classes in 2010 and 2014 were classified separately. For the three land cover classes in 2010, the calculation was based on the land cover information from GlobeLand30. In terms of the land cover distribution in 2014, the target classes were identified based on images from Landsat 8 by a support vector machine (SVM) classifier. Table 11 displays the land cover changes during 2010-2014. Among the three classes, vegetation has the highest percent increase in area, with area and percent changes of 188 km2 and 3.76%, respectively. In addition, artificial surfaces and water bodies show subtle changes of -0.66% and 0.33%, respectively. This analysis reveals minor changes in land cover in this area, which do not depreciate the validation results proposed in this study.

Table 11
Land cover changes in the study area from 2010-2014.

Land cover classes	Changing area/km^2	Percent change/% (%)
Artificial surfaces	−7.82	-0.66
Vegetation	188.00	3.76
Water bodies	1.72	0.33

Although the land cover landscapes of this area showed no significant changes, errors caused by the time gap may still exist in some regions. These errors mostly occur in changed artificial surface areas, where photos classified as artificial surfaces are located in the vegetation region in the land cover map. Moreover, the relatively low validation accuracy of artificial surfaces shown in both Tables 6 and 7 represents another possible source of errors.

To avoid errors caused by landscape changes, geo-tagged photos only taken from the same year as the map production are needed. Considering that photos only from Flickr are not sufficient, multi-sourced geo-tagged photos can be used to improve both data quantity and quality. Furthermore, photos related to some land cover classes, such as water bodies and forests, can be given a greater time tolerance, because these classes are more stable and unchangeable than other classes, such as artificial surfaces.

4.6. Building a common validation database with geo-tagged photos

This study proves the usability of geo-tagged photos for land cover validation. However, one major issue is that only regions containing geo-tagged photos can be assessed with the proposed method. This issue was considered during the sample selection process described in Section 2.2, where regions with a high density were extracted for sample allocation. Consequently, geo-tagged photos served as ancillary data to validate partial regions.

Despite the lack of photos from other areas, those assigned samples with photos can be automatically classified with typical classes, thus reducing the costs of time and labour in the validation process. As the development of the Internet and mobile devices promote the emergence of massive geo-tagged photo resources across the world, the time costing for validation will be dramatically reduced for larger regions, such as countries and continents, or at the global scale. This also reveals the opportunity to build a common validation database with geo-tagged photos to assist the procedure of validation.

Efforts should be focused on constructing this photo database for validation. In fact, several validation databases have been established through traditional data sources (Stehman et al., 2010; Zhao et al., 2014). However, when geo-tagged photos are involved, major concerns involve the construction of proper sampling designs and the collection of effective photos for sample classification. These issues will be explored in future studies.

5. Conclusions

Land cover maps are fundamental parts of different applications, including land use planning and natural resource management, and validation serves as a significant component in map production that ensures data quality. However, the complexity and labour consumption of the conventional validation process creates great challenges in validation process. This study established a novel approach to employ geo-tagged photos for land cover validation with deep learning. As a case study, 32367 photos from Flickr were collected for photo classification using CNN model and to assess the accuracy of GlobeLand30-2010 in western California. Based on the VGG-16 network, the results show an overall accuracy of 76.01% by the geo-tagged photo classification model. Based on the classified photos, 179 samples were classified automatically, with a classification accuracy of 79.33%. Finally, samples utilised to assess GlobeLand30-2010 presented an overall accuracy of 83.80%, which was consistent with the accuracy (80.45%) validated via visual interpretation. These results indicate that geo-tagged photos,

as well as our proposed approach, have the potential to contribute to land cover validation. In addition, CNN models including ResNet-50 and AlexNet were also involved in validation process. Despite of the differences of CNN model accuracies, the validation results, namely 86.03% and 86.59%, showed no significant changes compared with VGG-16. The result proves the usability and reliability of using deep learning for land cover validation.

Future studies will concentrate on the following aspects. Although existing studies have attempted to construct projects with crowdsourced data for land cover applications (Bayas et al., 2017; Fritz et al., 2012; Meng et al., 2017; Xing et al., 2017b), an automatic land cover validation process has not yet been integrated with these collected data. Thus, a geo-tagged photo database will be established and further improved with a large quantity of land cover photos, not only to involve more land cover classes in the validation process but also to increase the classification accuracy of these classes. Aside from this issue, future research could also investigate the potential of combining remotely sensed images and geo-tagged photos for land cover validation. Additionally, the feasibility of automatic land cover validation by other sources of crowdsourced data will be explored.

Acknowledgments

We are grateful for the valuable comments of Prof. Yaolong Zhao, Jinwei Dong and the anonymous reviewers. This work was supported by the National Natural Science Foundation of China (grant numbers 41501420, 41701443); Engineering Research Centre of Geospatial Information and Digital Technology, NASG (Grant number SIDT20170202); and the Project of the Science and Technology Foundation of Guangzhou (Grant numbers 156100021, 201803030034).

The authors declare no conflict of interest.

References

Antoniou, V., Fonte, C.C., See, L., Estima, J., Arsanjani, J.J., Lupia, F., Minghini, M., Foody, G., Fritz, S., 2016. Investigating the feasibility of geo-tagged photographs as sources of land cover input data. ISPRS Int. J. Geo-Inf. 5, 64.

Bayas, J.C.L., Lesiv, M., Waldner, F., Schucknecht, A., Duerauer, M., See, L., Fritz, S., Fraisl, D., Moorthy, I., McCallum, I., 2017. A global reference database of crowdsourced cropland data collected using the Geo-Wiki platform. Sci. Data 4, 170136.

Chen, J., Ban, Y., Li, S., 2014. China: Open access to Earth land-cover map. Nature 514, 434.

Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., He, C., Han, G., Peng, S., Lu, M., 2015. Global land cover mapping at 30m resolution: A POK-based operational approach. ISPRS J. Photogramm. Remote Sens. 103, 7–27.

Chen, F., Chen, J., Wu, H., Hou, D., Zhang, W., Zhang, J., Zhou, X., Chen, L., 2016. A landscape shape index-based sampling approach for land cover accuracy assessment. Sci. China Earth Sci. 59, 2263–2274.

Chen, J., Li, S., Wu, H., Chen, X., 2017. Towards a collaborative global land cover information service. Int. J. Digital Earth 10, 356–370.

Congalton, R.G., Green, K., 2008. Assessing the Accuracy of Remotely Sensed Data: Principles and Practices. CRC Press.

Dodge, S., Karam, L., 2016. Understanding how image quality affects deep neural networks. In: 2016 Eighth International Conference on Quality of Multimedia Experience (QoMEX). IEEE, pp. 1–6.

Duarte, B.P., Saraiva, P.M., 2008. An optimization-based approach for designing attribute acceptance sampling plans. Int. J. Qual. Reliab. Manage. 25, 824–841.

Ester, M., Kriegel, H.-P., Sander, J., Xu, X., 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. Kdd 96, 226–231.

Estima, J., Painho, M., 2013a. Exploratory analysis of OpenStreetMap for land use classification. In: Proceedings of the Second ACM SIGSPATIAL International Workshop on Crowdsourced and Volunteered Geographic Information, ACM, pp. 39–46.

Estima, J., Fonte, C.C., Painho, M., 2014. Comparative study of Land Use/Cover classification using Flickr photos, satellite imagery and Corine Land Cover database. In: Proceedings of the AGILE'2014 International Conference on Geographic Information Science, pp. 3–6.

Estima, J., Painho, M., 2013b. Flickr geotagged and publicly available photos: Preliminary study of Its adequacy for helping quality control of Corine land cover. In: Computational Science and Its Applications – ICCSA. Springer, Berlin Heidelberg, pp. 205–220.

Estima, J., Painho, M., 2014. Photo based Volunteered Geographic Information initiatives: a comparative study of their suitability for helping quality control of Corine Land Cover. Int. J. Agric. Environ. Inf. Syst. (IJAEIS) 5, 73–89.

Feddema, J.J., Oleson, K.W., Bonan, G.B., Mearns, L.O., Buja, L.E., Meehl, G.A., Washington, W.M., 2005. The importance of land-cover change in simulating future climates. Science 310, 1674–1678.

Foley, J.A., DeFries, R., Asner, G.P., Barford, C., Bonan, G., Carpenter, S.R., Chapin, F.S., Coe, M.T., Daily, G.C., Gibbs, H.K., 2005. Global consequences of land use. Science 309, 570–574.

Fonte, C.C., Bastin, L., See, L., Foody, G., Lupia, F., 2015. Usability of VGI for validation of land cover maps. Int. J. Geogr. Inf. Sci. 29, 1269–1291.

Foody, G.M., 2002. Status of land cover classification accuracy assessment. Remote Sens. Environ. 80, 185–201.

Foody, G.M., 2009. Sample size determination for image classification accuracy assessment and comparison. Int. J. Remote Sens. 30, 5273–5291.

Foody, G.M., Boyd, D.S., 2013. Using volunteered data in land cover map validation: Mapping West African forests. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 6, 1305–1312.

Fritz, S., McCallum, I., Schill, C., Perger, C., See, L., Schepaschenko, D., Van der Velde, M., Kraxner, F., Obersteiner, M., 2012. Geo-Wiki: An online platform for improving global land cover. Environ. Modell. Softw. 31, 110–123.

Fritz, S., See, L., Perger, C., Mccallum, I., Schill, C., Schepaschenko, D., Duerauer, M., Karner, M., Dresel, C., Laso-Bayas, J.-C., 2017. A global dataset of crowdsourced land cover and land use reference data. Sci. Data 4, 170075.

Fukushima, K., Miyake, S., 1982. Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition. In: Competition and Cooperation in Neural Nets, Springer, pp. 267–285.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition, In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778.

Hou, D., Chen, J., Wu, H., Li, S., Chen, F., Zhang, W., 2015. Active collection of land cover sample data from geo-tagged web texts. Remote Sens. 7, 5805–5827.

Hu, T., Yang, J., Li, X., Gong, P., 2016. Mapping urban land use by using Landsat images and open social data. Remote Sens. 8, 151.

Iwao, K., Nishida, K., Kinoshita, T., Yamagata, Y., 2006. Validating land cover maps with degree confluence project information. Geophys. Res Lett 33.

Johnson, B.A., Iizuka, K., Bragais, M.A., Endo, I., Magcale-Macandog, D.B., 2017. Employing crowdsourced geographic data and multi-temporal/multi-sensor satellite imagery to monitor land cover change: a case study in an urbanizing region of the Philippines. Comput. Environ. Urban Syst. 64, 184–193.

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems, pp. 1097–1105.

LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W., Jackel, L.D., 1989. Backpropagation applied to handwritten zip code recognition. Neural Comput. 1, 541–551.

LeCun, Y., Bottou, L., Bengio, Y., Haffner, P., 1998. Gradient-based learning applied to document recognition. Proc. IEEE 86, 2278–2324.

Leung, D., Newsam, S., 2012. Exploring geotagged images for land-use classification. In: Proceedings of the ACM Multimedia 2012 Workshop on Geotagging and its Applications in Multimedia, ACM, pp. 3–8.

Leung, D., Newsam, S., 2015. Land cover classification using geo-referenced photos. Multimed. Tools Appl. 74, 11741–11761.

Lillesand, T., Kiefer, R.W., Chipman, J., 2014. Remote Sensing and Image Interpretation. John Wiley & Sons.

Liu, X., He, J., Yao, Y., Zhang, J., Liang, H., Wang, H., Hong, Y., 2017. Classifying urban land use by integrating remote sensing and social media data. Int. J. Geogr. Inf. Sci. 1–22.

Mayaux, P., Eva, H., Gallego, J., Strahler, A.H., Herold, M., Agrawal, S., Naumov, S., De Miranda, E.E., Di Bella, C.M., Ordoyne, C., 2006. Validation of the global land cover 2000 map. IEEE Trans. Geosci. Remote Sens. 44, 1728–1739.

Meng, Y., Hou, D., Xing, H., 2017. Rapid detection of land cover changes using crowdsourced geographic information: a case study of Beijing. China. Sustainability 9, 1547.

Natya, S., Rehna, V., 2016. Land cover classification schemes using remote sensing images: a recent survey. Br. J. Appl. Sci. Technol. 13, 1–11.

Oba, H., Hirota, M., Chbeir, R., Ishikawa, H., Yokoyama, S., 2014. Towards better land cover classification using geo-tagged photographs. In: 2014 IEEE International Symposium on Multimedia, IEEE, pp. 320–327.

Olofsson, P., Stehman, S.V., Woodcock, C.E., Sulla-Menashe, D., Sibley, A.M., Newell, J.D., Friedl, M.A., Herold, M., 2012. A global land-cover validation data set, part I: fundamental design principles. Int. J. Remote Sens. 33, 5768–5788.

Olofsson, P., Foody, G.M., Herold, M., Stehman, S.V., Woodcock, C.E., Wulder, M.A., 2014. Good practices for estimating area and assessing accuracy of land change. Remote Sens. Environ. 148, 42–57.

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., 2015. Imagenet large scale visual recognition challenge. Int. J. Comput. Vision 115, 211–252.

Scepan, J., 1999. Thematic validation of high-resolution global land-cover data sets. Photogramm. Eng. Remote Sens. 65, 1051–1060.

Schultz, M., Voss, J., Auer, M., Carter, S., Zipf, A., 2017. Open land cover from OpenStreetMap and remote sensing. Int. J. Appl. Earth Obs. Geoinf. 63, 206–213.

See, L., Sturn, T., Fritz, S., McCallum, I., Salk, C., 2014. Cropland Capture: A Gaming Approach to Improve Global Land Cover. Agile Digital Editions.

Simonyan, K., Zisserman, A., 2014. Very deep Convolutional Networks for Large-Scale Image Recognition. Also Available at: arXiv preprint arXiv:1409.1556.

Sitthi, A., Nagai, M., Dailey, M., Ninsawat, S., 2016. Exploring land use and land cover of geotagged social-sensing images using Naive Bayes classifier. Sustainability 8, 921.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., 2014. Dropout: A simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. 15,

1929–1958.
Stehman, S.V., 2009. Sampling designs for accuracy assessment of land cover. Int. J. Remote Sens. 30, 5243–5272.
Stehman, S., Olofsson, P., Woodcock, C., Friedl, M., Sibley, A., Newell, J., Sulla-Menashe, D., Herold, M., 2010. Designing a reference validation database for accuracy assessment of land cover. In: Accuracy 2010 Symposium, July, pp. 20–23.
Stehman, S.V., Olofsson, P., Woodcock, C.E., Herold, M., Friedl, M.A., 2012. A global land-cover validation data set, II: augmenting a stratified sampling design to estimate accuracy by region and land-cover class. Int. J. Remote Sens. 33, 6975–6993.
Strahler, A.H., Boschetti, L., Foody, G.M., Friedl, M.A., Hansen, M.C., Herold, M., Mayaux, P., Morisette, J.T., Stehman, S.V., Woodcock, C.E., 2006. Global Land Cover Validation: Recommendations for Evaluation and Accuracy Assessment of Global land Cover Maps. Eur, Communities, Luxembourg, pp. 51.
Tong, X., Wang, Z., Xie, H., Liang, D., Jiang, Z., Li, J., Li, J., 2011. Designing a two-rank acceptance sampling plan for quality inspection of geospatial data products. Comput. Geosci-uk 37, 1570–1583.
Tong, X., Zhang, X., Shan, J., Xie, H., Liu, M., 2013. Attraction-repulsion model-based subpixel mapping of multi-/hyperspectral imagery. IEEE Trans. Geosci. Remote Sens. 51, 2799–2814.
Wulder, M.A., Franklin, S.E., White, J.C., Linke, J., Magnussen, S., 2006. An accuracy assessment framework for large-area land cover classification products derived from medium-resolution satellite data. Int. J. Remote Sens. 27, 663–683.
Xie, H., Tong, X., Meng, W., Liang, D., Wang, Z., Shi, W., 2015. A multilevel stratified spatial sampling approach for the quality assessment of remote-sensing-derived products. IEEE J.Sel. Top. Appl. Earth Obs. Remote Sens. 8, 4699–4713.
Xing, H., Meng, Y., Hou, D., Song, J., Xu, H., 2017b. Employing crowdsourced geographic information to classify land cover with spatial clustering and topic model. Remote Sens. 9, 602.
Xing, H., Meng, Y., Hou, D., Cao, F., Xu, H., 2017a. Exploring point-of-interest data from social media for artificial surface validation with decision trees. Int. J. Remote Sens. 38, 6945–6969.
Xu, G., Zhu, X., Fu, D., Dong, J., Xiao, X., 2017. Automatic land cover classification of geo-tagged field photos by deep learning. Environ. Modell Softw. 91, 127–134.
Yang, Y., Xiao, P., Feng, X., Li, H., 2017. Accuracy assessment of seven global land cover datasets over China. ISPRS J. Photogramm. Remote Sens. 125, 156–173.
Zhao, Y., Gong, P., Yu, L., Hu, L., Li, X., Li, C., Zhang, H., Zheng, Y., Wang, J., Zhao, Y., 2014. Towards a common validation sample set for global land-cover mapping. Int. J. Remote Sens. 35, 4795–4814.
Zhao, Y., Feng, D., Yu, L., See, L., Fritz, S., Perger, C., Gong, P., 2017. Assessing and improving the reliability of volunteered land cover reference data. Remote Sens. 9, 1034.
Zhou, B., Lapedriza, A., Xiao, J., Torralba, A., Oliva, A., 2014. Learning deep features for scene recognition using places database. In: Advances in Neural Information Processing Systems, pp. 487–495.
Zhu, Y., Newsam, S., 2015. Land use classification using convolutional neural networks applied to ground-level images. In: Proceedings of the 23rd SIGSPATIAL International Conference on Advances in Geographic Information Systems, ACM, p. 61.