

Article



# A Deep Neural Networks Approach for Augmenting Samples of Land Cover Classification

# Chuanpeng Zhao <sup>1,2</sup> and Yaohuan Huang <sup>1,2,\*</sup>

- State Key Laboratory of Resources and Environmental Information System, Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China; zhaocp.15s@igsnrr.ac.cn
- <sup>2</sup> College of Resource and Environment, University of Chinese Academy of Sciences, Beijing 100049, China
- \* Correspondence: huangyh@lreis.ac.cn; Tel.: +86-10-6488-9681

Received: 15 July 2020; Accepted: 11 August 2020; Published: 13 August 2020



Abstract: Land cover is one of key indicators for modeling ecological, environmental, and climatic processes, which changes frequently due to natural factors and anthropogenic activities. The changes demand various samples for updating land cover maps, although in reality the number of samples is always insufficient. Sample augment methods can fill this gap, but these methods still face difficulties, especially for high-resolution remote sensing data. The difficulties include the following: (1) excessive human involvement, which is mostly caused by human interpretation, even by active learning-based methods; (2) large variations of segmented land cover objects, which affects the generalization to unseen areas especially for proposed methods that are validated in small study areas. To solve these problems, we proposed a sample augment method incorporating the deep neural networks using a Gaofen-2 image. To avoid error accumulation, the neural network-based sample augment (NNSA) framework employs non-iterative procedure, and augments from 184 image objects with labels to 75,112 samples. The overall accuracy (OA) of NNSA is 20% higher than that of label propagation (LP) in reference to expert interpreted results; the LP has an OA of 61.16%. The accuracy decreases by approximately 10% in the coastal validation area, which has different characteristics from the inland samples. We also compared the iterative and non-iterative strategies without external information added. The results of the validation area containing original samples show that non-iterative methods have a higher OA and a lower sample imbalance. The NNSA method that augments sample size with higher accuracy can benefit the update of land cover information.

**Keywords:** sample augment; deep neural network; small size samples; land cover; object-based image analysis

# 1. Introduction

Land cover is the physical material (e.g., grass, trees, bare ground, and water) at the surface of the earth, which is an essential parameter for global change, crop production estimation, and terrestrial water cycle [1–4]. Land cover is changing as a result of both natural factors and human activities, increasing difficulties and uncertainties in updating land cover maps [5]. Fortunately, remote sensing technology provides images that models reality of land cover as an indispensable data source for updating land cover maps [6]. A sample usually contains the label information of a location in image, and it is used to label other locations in the image. The sample is undoubtedly crucial for updating land cover maps by remote sensing classification, as it impacts the accuracy and quality of the end product.

The sample size (i.e., the number of samples) affects the accuracy of remote sensing classification, and reducing the number of samples will produce lower classification accuracy in general [7–11], especially for object-based image analysis (OBIA). For a remote sensing image, the classification can be

2 of 17

conducted on each pixel or a bunch of neighboring pixels (i.e., image object) [12,13]. Compared with pixels that provides spectral information, image objects contain additional information on spectra, geometry, context, and texture. Thus, OBIA leads to sample sparsity in high dimensional data space, which increases the needs for larger sample size [14]. Although some machine learning algorithms, which are popular in supervised remote sensing classification, are tolerable to insufficient sample size in high dimensions, studies show that the sample size leads to larger variations in accuracy than the algorithms themselves [10,15]. A large sample size demands a lot of manpower and financial resources, which may seem unrealistic for updating land cover maps when the land-surface elements are continuously changing both temporally and spatially. A small sample set is obviously more efficient in manpower and financial resources, and one that is interpreted by experts is especially more representative than random sampling results, since an expert can generalize the characteristics of some land-surface elements successfully from few examples or their early experience [16]. However, the small sample size may highly reduce the land cover classification accuracy for current computer-based classification algorithms. Thus, it is necessary to augment sample size from small to large to ensure the accuracy of land cover maps.

To augment a sample size, a number of techniques that utilize an existing small sample set have been developed [17]. These can be categorized into two basic types [11]: (1) active learning and (2) semisupervised learning. Also, classification methods with prediction probability can be used in sample augment [18–20]. Active learning will query unlabeled samples in the training data set for their labels, and thus requires a lower number of samples to classify land covers. Much research has been conducted on active learning-based sample size augmented for remote sensing classification [21–23]. Although active learning can reduce the amount of sample required, the process of labeling new samples still demands a large amount of manpower, especially in a complicated study area. Meanwhile, the samples queried by active learning may be uninformative or indistinguishable by humans. Semisupervised learning uses unlabeled samples with the help of labeled samples. The label propagation (LP) algorithm is one of the popular semisupervised methods, which exploits labeled and unlabeled samples in constructing a graph model to predict the labels of unlabeled samples by the similarity between two samples [24–26]. For example, Shi et al. [27] used LP to predict the unlabeled samples in remote sensing image classification, and Wang, Hao, Wang, and Wang [25] propagated labels to unlabeled samples using LP with the help of spatial-spectral graph. However, the feature vector of segmented land cover image objects varies with segmentation parameters, which brings difficulty in selecting a representative sample set. The variation in land cover inherent from the dynamics and complexity of land covers [28] also affects the generalization of LP derived graph under small sample size, which is a big challenge for sample augmentation. One possible way to alleviate the effects is to utilize the generalization power of deep neural networks (DNN) [29], which have achieved remarkable practical success in various application domains [30–33]. Though there are sample augment methods incorporating current development of neural networks on hyperspectral images [34–37], the methods working on features derived from multispectral images are still lacking.

To alleviate the effects of variations and to improve the size and accuracy of sample augmented results, we developed a sample augmented framework that incorporates DNN. The proposed neural network-based sample augment (NNSA) framework can be described in four steps: (1) Select optimal features for identifying each land cover category; (2) measure similarities between image objects and samples belonging to a certain land cover category; (3) feed DNN with the similarity measurement results; and (4) cluster to refine sample augmented results by DNN. To quantitatively evaluate the proposed NNSA, we compared the method results with those of LP and DNN in reference to expert interpretation results. We compared the generalization capacities of the three methods on another unseen coastal validation area, which is different from samples that only contain inland land cover characteristics. Furthermore, we compared the iterative and non-iterative strategies for sample augment in the inland validation area.

# 2. Materials and Methods

# 2.1. Data and Study Area

The Gaofen-2 image data acquired on 26 May 2015, were used as a major data source. The Gaofen-2 satellite scans a swath of ~45 km and provides 1-m panchromatic images and 4-m multispectral images with 4 bands that belong to a series of civilian high-resolution optical satellites of the China National Space Administration. Both multispectral and panchromatic images were orthorectified and corrected to surface reflectance with the FLAASH (Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes) algorithm [38] and then were fused together using the Gram-Schmidt pan-sharpen method [39].

The study area is approximately 518 km<sup>2</sup> over the coast of the Bohai Sea and covers more than 90 villages and 3 towns (Figure 1). This area belongs to the Beijing-Tianjin-Hebei region, where unprecedented coastal development has led to fragmented, complicated and fast-changing land covers [40]. The land cover type of the study area consists of six classes, including water, forestland, grass land, crop land, bare land, and residential and built-up land, which involve in the entire National Land Resource Classification System of China. Furthermore, high-resolution satellites provide more detailed information on land covers and greater separability between further subcategories, which leads to increased intraclass variations. The residential and built-up land mainly includes residential districts, industrial areas, roads, and vegetable greenhouses. The forestland is comprised of sparse forest alongside roads, dense forest along the river and close to the sea, and juvenile woodland. Crop land has different colors, such as brown, light brown, dark brown, and green, which depend on the crops and soil moisture. The large intraclass variability poses challenges for sample augmentation. For example, crop lands that are in green color are easily confused with the grass land during sample augmentation.



**Figure 1.** The study area used in this study. The study area is shown in red in (**a**). The image is shown in a true-color composite of red, green and blue bands in (**b**). The two testing areas interpreted by experts on land cover are shown in (**c**,**d**). Samples manually selected with a total of 184 instances are shown in (**c**).

The size of Gaofen-2 image data of our study area is 26,631 rows by 27,407 columns. It is segmented to 154,667 image objects using eCognition Developer 9.0 software (Trimble Inc., Munich, Germany) and an automated parameterization algorithm [41]. A total of 28 attributes were calculated on these image objects (Table 1). Then, a total of 184 instances for all six land cover categories was selected as initial samples for augmentation (Figure 1c). To verify the sample augmentation method, we collected the

expert-interpreted land cover results (Table 2), which consists of two parts (Figure 1c,d), with a total of 23,484 image objects.

Features Category	Object Features		
Spectral	Mean (4), Standard deviation (4), Skewness (4), Brightness	13	
Geometry	Border index, Compactness, Shape index	3	
Texture	Gray level co-occurrence matrix (GLCM) Homogeneity (all direction), GLCM Contrast (all direction), GLCM Dissimilarity (all direction), GLCM Entropy (all direction), GLCM Ang. 2nd moment (all direction), GLCM Mean (all direction), GLCM Standard Deviation (all direction), GLCM Correlation (all direction)	8	
Customized	Normalized difference vegetation index (NDVI), Normalized difference water index of McFeeters (NDWIF), Soil adjusted vegetation index (SAVI), Optimized soil adjusted vegetation index (OSAVI)	4	
	Total	28	

Table 1. Features that were calculated to identify image objects in this study.

Table 2. Samples of six land cover categories for feature selection and sample augmentation testing.

	Number of Samples		
Land Cover	Training Objects	<b>Testing Objects</b>	
Water	29	1504	
Forest land	36	3788	
Grass land	32	3266	
Crop land	33	8722	
Bare land	22	1063	
Residential and built-up land	32	5141	

#### 2.2. Feature Selection from Small-Size Samples

Feature selection can improve the performance of object-based image classification [11]. To select features under low sample size and high dimensionality, we adapted a revised version of our previous work [14], which is designed to provide a solution to this problem. The previous work, namely the group-corrected partial least squares generalized linear regression (PLSGLR) method, can be described in three steps: (1) Group features based on Pearson's correlation coefficient; (2) rank features by PLSGLR and remove insignificant features; (3) reconstruct categories when the features are added one-by-one to calculate the Bayesian information criterion.

Compared to our previous work, we improved the stability of feature selection against random sampling uncertainty by incorporating a co-occurrence matrix and voting strategy. Given the feature group result G that results from one of the total sampling numbers, N, the co-occurrence matrix P can be defined as

$$P(i,j) = \frac{1}{N} \sum \begin{cases} 1, & G(i,j) = 1\\ 0, & G(i,j) = 0 \end{cases}$$
(1)

where *i* and *j* are the *i*-th feature and the *j*-th feature, respectively. When the value of the co-occurrence between a pair of features is greater than threshold, *th*, the group between two features is retained; otherwise it is discarded. With regard to the ranking matrix, *R*, with *m* rows and *N* columns, the feature of the *i*-th position, *f<sub>i</sub>*, can be expressed as

$$f_i = \max\{R_{ik}\}, \quad k = 1, 2, \dots, N$$
 (2)

where *m* is the number of features. For the given ranking vector  $\{R_1, R_2, \dots, R_n\}$ , the BIC matrix *B* with *n* rows and *N* columns, the final feature number,  $n_f$ , is defined as

$$n_f = \frac{1}{N} \sum b_k, \quad k = 1, 2, \dots, N$$
 (3)

where *n* is the number of ranking features ( $n \le m$ ), and  $b_k$  is the number of optimal features at the *k*-th sampling result.

#### 2.3. Similarity Measurement of Image Objects

Nonparametric methods make fewer preliminary hypotheses and are more powerful for describing the nonlinear and complex relationships [42]. To model the relationship between unlabeled image objects and samples that are marked as belonging to a certain category, a nonparametric method, kernel density estimation (KDE), which had been applied to land cover classification [18], is used to extract the relationships. For samples of a category, we use KDE to extract the curve on one of the selected features without assuming the relationship in advance. The curve f(x) can be described as

$$f(x) = \frac{1}{Mh} \sum_{i=1}^{M} K\left(\frac{x - x_i}{h}\right) \tag{4}$$

where *M* is the number of samples of a category, *x* is the value to be estimated,  $x_i$  is value of the *i*-th sample, *K* is the kernel function, and *h* is the bandwidth that controls smoothness of the estimated curve. In this study, we used the Gaussian kernel and determined the bandwidth with the normal reference rule [43].

To improve the performance and stability, we employed repeated sampling with a replacement strategy and calculated the normalized relationship each time. The curve F(x) can be represented as the average of a set of curves that result from the repeated sampling process.

$$F(x) = \frac{1}{N} \sum_{k=1}^{N} f_k(x)$$
(5)

After the curve F(x) of a feature on a category is determined, the similarity of the selected features of a certain category is calculated. For an unlabeled image object, the similarities corresponding to values of the selected features are calculated by the interpolation method, and similarities of different features are weighted to generate the final similarity. The similarity *S* can be expressed as follows

$$S = \sum_{j=1}^{n} \omega F_j(x) \tag{6}$$

where *j* is the *j*-th selected feature and  $\omega$  is user-defined weight vector. As the results of feature selection are sorted in descending order of importance, a weight vector (0, 0, ..., 1) means the limiting factor principle; a weight vector (1, 0, ..., 0) represents the dominant factor principle; and a weight vector (1, 1, ..., 1) indicates the average principle.

#### 2.4. Sample Augmentation by Neural Networks

After feature selection and similarity measurement, the image object that has a similarity over 0.5 for a certain land cover category is feed to a DNN. The structure of DNN is displayed in Figure 2. Given an image object *X* with a label *y*, the feature vector  $\{f_1, f_2, \dots, f_m\}$  is fed to the input layer of a DNN. For the hidden neuron  $h_j$  that is activated with a rectified linear unit (ReLU), the output value can be written as

$$h_j = \max\left\{0, \sum_{k=1}^m w_k f_k\right\} \tag{7}$$

In view of the output neuron  $o_k$ , which incorporates the results of multiple hidden layers,  $S_k$  can be described as

$$o_k = \frac{1}{1 + e^{-s_k}}$$

$$s_k = \sum_j h_j w_{jk}$$
(8)

where *j* is the *j*-th hidden node, and *k* is the *k*-th output node. The parameters of the DNN are optimized to minimize the binary cross entropy error:

$$E = -\sum_{k} [y_k \log(o_k) + (1 - y_k) \log(1 - o_k)]$$
(9)

where  $y_k$  is the label of an image object, and  $o_k$  is the prediction result. In addition, dropout is employed to prevent the DNN from overfitting [44].



Figure 2. Workflow of sample augment by the proposed framework.

In this study, we used dropout layer with a fixed drop rate of 0.25 between each hidden layer, which could prevent overfitting. We used RMSprop optimizer with adaptive learning rate to train the model, and cross entropy to evaluate predicted classes. The implementation was based on Keras in python environment [45].

The data set is then randomized and split into a training set accounting for 80% and a testing set accounting for 20%. A DNN with the same structure is trained using the training set. The DNN used in this study has four hidden layers. In addition, the number of neurons of each layer in the DNN is determined by 80% decreasing in number, empirically.

## 2.5. Postprocessing by Clustering

The similarity measurement can introduce an error into the sample augmentation process, and the error may be further amplified with supervised learning. To eliminate these error samples, a clustering method is used can discover the potential structure of samples. It is feasible in practice to manually specify the cluster centers or to automatically determine the cluster numbers with some criteria. In this study, we employed the spectral clustering method [46,47] for post processing. Given a set of augmented samples, *X*, with the calculated image objects features, the similarity  $\omega_{ij}$  between two samples (image objects),  $x_i$  and  $x_j$ , can be written as

$$w_{ij} = \begin{cases} 0, & x_i \notin \Omega(x_j) \text{ or } x_j \notin \Omega(x_i) \\ e^{-\frac{\|x_i - x_j\|_2^2}{2\delta^2}}, & x_i \in \Omega(x_j) \text{ and } x_j \in \Omega(x_i) \end{cases}$$
(10)

where  $x_i \in \Omega(x_j)$  indicates that  $x_i$  is among the k-nearest neighbors of  $x_j$ , and  $\delta$  controls the width of the neighborhoods. Then, the normalized Laplacian matrix  $L_{rm}$  of the similarity matrix W is constructed as follows

$$L_{rm} = E - D^{-1}W \tag{11}$$

where *E* is an identity matrix, and *D* is a diagonal matrix with diagonal elements written as  $D_{ii} = \sum \omega_{ij}$ .

The eigenvectors of the Laplacian matrix are sorted according to the eigenvalues, and the top k eigenvectors (e.g., a number of cluster) are delivered into the k-means algorithm. The cluster number k is determined by Bartlett's test [48].

As a summary, Figure 2 shows the detailed workflow of sample augment by the proposed framework.

### 2.6. Validation

To test thoroughly the sample augment results, we compared the results of NNSA with the results of the LP and DNN. As the core part of NNSA, DNN has the potential to augment samples directly without the help of the proposed framework. Two regions with experts interpreted land covers are employed to compare the augment accuracy. One region (Figure 1c) contains all 184 training and testing samples, while the other region (Figure 1d) does not contain any samples used in this study. The two regions show different land cover characteristics. One covered with typical inland land covers has large areas of crop land, and the other in coastal areas has marine farms and more forest land. We compared the overall accuracy (OA) and sample imbalance of these methods in each region. The sample imbalance is described by the maximum value of the sample size ratio between different land cover categories.

We also compared the iterative and non-iterative strategies in sample augment. In the iterative scenario, we iteratively improved the number of samples of LP and support vector machine (SVM). The radial basis kernel function was used in SVM with the gamma parameter of 5 and the cost parameter of 10 after tuning. We iterated 300 times over the region that contains original samples for each scenario, adding 10 samples to training dataset each time for iterative scenario, and assessing the results by OA and sample imbalance. In the non-iterative scenario, we just trained from the original samples and selected equal sample size by classification probabilities.

#### 3. Results

#### 3.1. Sample Augmentation Results

The accuracy and loss curve for training DNN in NNSA are shown in Figure 3. The accuracy values for were converged at about 96% and 98% from the epoch number of 1500. The training stopped at the epoch number of 3442, after the model was not improved after trying 500 epochs.



**Figure 3.** Accuracy and loss curve for training deep neural networks (DNN) in neural network-based sample augment (NNSA).

The sample augmentation results by the NNSA are shown in Figure 4. There are 48.56% image objects assigned to a certain category by NNSA. Among these augmented samples, 43.76% image objects, which were distributed in the oceans, rivers, and fishery farms, were identified as water. In view of the crop land, 22.47% image objects were mainly scattered around the residential land in the eastern part of the study area. A further 16.95% image objects that were concentrated in residential areas, roads, and fisheries farms were identified as the residential and built-up land. There were 9.57% image objects that were labeled as the forestland, which was mainly distributed near the sea and in the vicinity of the rivers. The 2.01% image objects that were assigned to the grass land were located around the forestland and water. The remaining 5.23% image objects that were recognized as bare land were located by the sea and rivers. The augmented samples are consistent with the distribution of different land cover categories, although there are inevitably some errors, such as some crop land in the green being labeled as the grass land.



Figure 4. The sample augment results from NNSA framework.

## 3.2. Comparison and Validation

# 3.2.1. Performance in the Similar Land Cover Region

The validation area with original 184 training samples has typical inland land cover characteristics. In this validation area, the OAs of NNSA, LP, and DNN are 83.85%, 61.16%, and 88.52%, respectively. Figure 5 demonstrates further accuracy assessment in metrics of UA and PA. The absence of the grass land and the bare land in DNN results from their prediction probabilities being lower than 0.9 (to acquire similar sample size in the validation regions, we adopt this threshold). In these two land cover categories, NNSA and LP both suffer from low accuracy. As to the remaining land cover categories, the DNN has higher accuracy than those of LP, which demonstrates the reasonability of the introduction DNN to sample augment area.



**Figure 5.** The accuracy assessment of sample augment results from NNSA in validation area with original samples including (**a**) user's accuracy and (**b**) producer's accuracy. The letters near the *x*-axis represent land cover categories. The land cover categories water, forestland, grass land, crop land, bare land, and residential and built-up land are denoted as W, F, G, C, B, and R, respectively.



**Figure 6.** The validation area with original samples. The expert interpreted land cover result is shown in (**a**). The sample augment results of NNSA, label propagation (LP), and DNN are shown in (**b**–**d**), respectively.

Figure 6 shows the spatial distribution of the augmented samples. All the three methods (i.e., NNSA, LP, and DNN) perform well in the crop land with a PA of over 90%. For water and forest land, DNN has similar accuracy with NNSA, while LP labels some wet crop land as water, and some buildings as the forest land. In view of grass land and bare land, DNN has lower probability in identifying these image objects. NNSA and LP both suffer from low accuracy in these two land cover categories, especially in the grass land. The grass land is very similar to crop land in green, since the segmentation scale invalid the shape difference between crop land and grass land. The clustering process of NNSA alleviates the problem to some extent. NNSA and DNN achieved better results than LP by numeric accuracy assessment and visual inspection in this validation area.

#### 3.2.2. Performance in Dissimilar Land Cover Region

In this validation area, the OAs of NNSA, LP, and DNN are 75.80%, 50.53%, and 72.45%, respectively. The OA values are about 10% lower than the above validation region. The Figure 7 demonstrates the UAs and PAs of these three methods. The DNN has high accuracy in forest land and crop land, except it mislabeled some wet crop land as water. For the grass land, LP performs better in this validation area since there are more grass lands. LP suffered from low accuracy in crop land and water, as it also identified wet crop land as water. NNSA performs better in these land cover categories except grass land.



**Figure 7.** The accuracy assessment of sample augment results from NNSA in validation area without original samples including (**a**) user's accuracy and (**b**) producer's accuracy. The letters near the *x*-axis represent land cover categories. The land cover categories water, forestland, grass land, crop land, bare land, and residential and built-up land are denoted as W, F, G, C, B, and R, respectively.

The spatial distribution of the augmented samples is shown in Figure 8. The consistency between these method results and expert interpretation decreased, compared to the previous validation area. The reduction in consistency may result from the complex land covers in coastal area. This area has both inland land covers and marine land covers. The original samples only contain the characteristics of inland land covers. Thus, the reduction in accuracy is inevitable. The ditches scatter in crop land increase the soil moisture, which confuses the LP method. NNSA achieved better results by visual inspection in this validation area.



**Figure 8.** The validation area without original samples. The expert interpreted land cover result is shown in (**a**). The sample augment results of NNSA, LP, and DNN are shown in (**b**–**d**), respectively.

# 4. Discussion

# 4.1. Effects of Segmented Land Cover Image Objects and Intraclass Variability

The results of this paper show that all the three methods (i.e., NNSA, LP, and DNN) have a reduction in accuracy in the validation area without samples, though NNSA incorporates the generalization capacity of DNN up to four hidden layers. We interpret the mechanism as follows: as illustrated in Figure 9, the crop land has at least four manifestations in the study area, and the samples collected from a small part of the area take no consideration of the variations of each land cover category. Also, the segmentation scheme converting images from pixels to objects affects the results. Inevitably, the objects corresponding to ground entities and patches of surface cover depend on the segmentation parameters [13,49]. The parameters in this study area are selected by optimizing the mean of local variance in a global search scheme [41]. The segmented objects therefore do not guarantee the fitness on each land cover category. For example, reflectivity of crop land is affected by many factors such as humidity, crop variety, growth cycle, and temperature in reality; experts identify it by regular boundaries. When the segmentation results on crop land is fragmentized, the information on geometry describing the boundary will be invalidated. Segmentation based on certain sub-category land cover objects may relieve this problem, which needs further studies.



**Figure 9.** Illustration of intraclass variation of the crop land in the study area. Four patches of crop land with different appearance are shown in (**a**–**d**).

In addition to the fragmentized segmentation results, intraclass variability of a land cover category also affect the augmentation results. In this study, we used a classification system derived from Data Center for Resources and Environmental Sciences, Chinese Academy of Sciences [50]. Take crop land as an example: the category consists of paddy field and dry farming field, which was defined by the way of land use. Since it is unreasonable to define subcategories for land covers in detail, visual inspection as in [14] is inevitable. As a complementary, an unsupervised clustering method as in Section 2.5 is also recommended. With a better segmentation result with a refined classification system, the efficiency NNSA will be improved. The refined classification system, which decreases the intraclass variability by subcategories, needs further exploration.

## 4.2. Comparisons with Other Sample Augment Methods

The NNSA extends the previous studies that augments small samples to large samples. NNSA utilizes classification methods with predication probability for sample augmentation and has higher OA than LP in reference to expert interpreted results. The NNSA uses non-iterative sample augment approach, while LP employs iterative sample augment procedure. We compared the iterative and non-iterative sample augment procedure without external information added, even active learning and semisupervised learning favor to iteratively improve the number of samples. Figure 10 shows the accuracy and sample imbalance of LP and SVM in two scenarios. In our study area, the non-iterative LP had higher accuracy and lower sample imbalance. The accuracy reached the highest at the 36th iteration (OA = 89.45%), and then gradually decreased to around 70%. In view of the results of SVM, the non-iterative had a little bit lower accuracy than the iterative version. Considering the severe imbalance of the iteration SVM results (Table 3), its reasonable that the non-iterative SVM is more suitable for sample augmentation.

Non-iterative procedure avoids the accumulation of error labels, which is critical for object-based image analysis. For multispectral image, there are often dozens of features after segmentation. As to hyperspectral data, there may be hundreds of features available. Only some of these features describe the characteristics of an image object, since the segmentation process does not guarantee the validity of these features. The iterative procedure may suffer from accumulation and amplification of errors during the rolling.



**Figure 10.** The overall accuracy and sample imbalance of: (**a**) LP and (**b**) support vector machine (SVM) under iterative and noniterative scenarios.

**Table 3.** Summary of overall accuracy and sample imbalance of LP and SVM under iterative and noniterative scenarios.

	Itera	tive	Noniterative	
	Overall Accuracy	Imbalance	Overall Accuracy	Imbalance
LP	min: 65.65%	min: 3.23	min: 84.50%	min: 3.23
	max: 89.45%	max: 38.80	max: 88.77%	max: 31.13
	mean: 74.47%	mean: 30.77	mean: 87.64%	mean: 20.06
	deviation: 8.32%	deviation: 6.59	deviation: 0.61%	deviation: 5.59
SVM	min: 84.50%	min: 3.31	min: 78.30%	min: 2.96
	max: 96.60%	max: 127.52	max: 92.58%	max: 40.85
	mean: 94.09%	mean: 72.20	mean: 86.65%	mean: 15.68
	deviation: 2.67%	deviation: 40.19	deviation: 2.57%	deviation: 6.37

The sample augment procedure is highly needed for land cover classification over very large areas [28], as there are many unseen areas that lack samples in this scenario. To cope with the unseen areas, these methods employ different strategies. The active learning queries humans about the true labels of a series of samples to migrate to unseen areas [21,22,51], thus taking the variations of unseen areas into consideration. Compared to active learning, NNSA avoids feeding method with sequences of samples, since selecting and interpreting samples are not error-free and potentially cause biased classification results [52,53]. The iterative methods fuse samples with high similarities into the original sample set. As mentioned above, the segmented objects may not describe reality well, and this fusion will introduce high uncertainty in each rolling procedure. The NNSA also incorporates similar samples but refines by robust neural networks and cluster scheme without the rolling procedure, which reduces the uncertainty and avoids the error accumulation.

There is still work to be done for the development of NNSA. One part comes from the DNN itself, which remain unsolved such as hyperparameter selection and network structure determination. The other part involves the descriptions of segmented image objects. The image objects can be described as feature vectors or images in two dimensions. The proposed NNSA uses feature vectors as inputs. Recent advances in deep convolutional neural network hold the promise of describing the image objects by themselves [54], thus reducing the uncertainty of the NNSA framework.

# 4.3. Sample Augmentation for Remote Sensing with an Insufficient Sample Size

Recent development of remote sensing sensors supplies a large volume of fine-scale spatial-temporal data [55]. Most applications of big data methods in geography focus on social behaviors due to the sufficiency of mobile phone data, microblog data, and traffic data [56]. The big data methods are always limited by insufficient samples in land cover classification [57]. Field surveys require more investment; thus, a large sample size is always inaccessible. The NNSA framework, designed for jumping from a small sample set to a huge sample set, can alleviate the problem to some extent.

Both the sample size and sample quality are crucial [58,59], though big data methods concentrate more on sample size. The balance between sample size and sample quality such as representativeness is important for sample augmentation. The proposed NNSA currently focuses on the sample size, and the representativeness is implicit in the cluster procedure. To reduce computational overhead over large areas, representative samples can be selected from each cluster.

## 5. Conclusions

In this paper, we presented a sample augmentation framework that incorporates DNN for object-based image analysis to augment samples with high accuracy. The proposed framework was applied to a scene of Gaofen-2 image to augment 184 samples to a large sample set, and achieved an improvement on the accuracy in comparison with LP and DNN. NNSA achieved an overall accuracy about 20% higher than that of LP in both validation areas. In detail, the NNSA method achieved an overall accuracy of 83.85% in validation area with original samples and an overall accuracy of 75.80% in validation area without original samples. To prove the advantages of the non-iterative strategy used in NNSA, we compared the iterative and non-iterative sample augment procedures in the validation area when no external information was added, and found the non-iterative procedure has lower sample imbalance and higher overall accuracy. The non-iterative procedure avoids error accumulations since the segmented image objects may deviate from the ground entities in global segmentation parameter optimization procedures. The NNSA incorporates similar samples but refines them by robust neural networks and cluster scheme without the rolling procedure, which reduces the uncertainty and avoids the error accumulation.

The proposed NNSA framework can be used to generate a big sample set from a relatively small one for object-based image analysis in land cover classification. The NNSA merely utilizes the original 184 samples and avoids error accumulations by a non-iterative procedure, which can be applied in sample augment in large area. As there is a gap between insufficient sample sizes and big data for remote sensing, the proposed framework may be helpful for land cover classification by small size samples. Future work will focus on the description of image objects.

**Author Contributions:** Conceptualization, Y.H.; methodology, C.Z.; validation, C.Z. and Y.H.; resources, Y.H.; data curation, Y.H.; writing—original draft preparation, C.Z.; writing—review and editing, Y.H.; visualization, C.Z.; supervision, Y.H.; funding acquisition, Y.H. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Key Research and Development Program of China, Grant number 2016YFC0401404 and grant number 2017YFB0503005; and the Strategic Priority Research Program of the Chinese Academy of Sciences, grant number XDA23100301.

Conflicts of Interest: The authors declare no conflict of interest.

# References

- 1. 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. Global consequences of land use. *Science* **2005**, *309*, 570–574. [CrossRef] [PubMed]
- 2. Ahl, D.E.; Gower, S.T.; Mackay, D.S.; Burrows, S.N.; Norman, J.M.; Diak, G.R. Heterogeneity of light use efficiency in a northern wisconsin forest: Implications for modeling net primary production with remote sensing. *Remote Sens. Environ.* **2004**, *93*, 168–178. [CrossRef]
- 3. Sterling, S.M.; Ducharne, A.; Polcher, J. The impact of global land-cover change on the terrestrial water cycle. *Nat. Clim. Chang.* **2013**, *3*, 385. [CrossRef]
- Findell, K.L.; Berg, A.; Gentine, P.; Krasting, J.P.; Lintner, B.R.; Malyshev, S.; Santanello, J.A.; Shevliakova, E. The impact of anthropogenic land use and land cover change on regional climate extremes. *Nat. Commun.* 2017, *8*, 989. [CrossRef]
- 5. Verburg, P.H.; Neumann, K.; Nol, L. Challenges in using land use and land cover data for global change studies. *Glob. Chang. Biol.* **2011**, *17*, 974–989. [CrossRef]
- 6. Yifang, B.; Gong, P.; Gini, C. Global land cover mapping using earth observation satellite data: Recent progresses and challenges. *ISPRS* **2015**, *103*, 1–6.
- 7. Ma, L.; Cheng, L.; Li, M.; Liu, Y.; Ma, X. Training set size, scale, and features in geographic object-based image analysis of very high resolution unmanned aerial vehicle imagery. *ISPRS* **2015**, *102*, 14–27. [CrossRef]
- 8. Li, M.; Ma, L.; Blaschke, T.; Cheng, L.; Tiede, D. A systematic comparison of different object-based classification techniques using high spatial resolution imagery in agricultural environments. *Int. J. Appl. Earth. Obs. Geoinf.* **2016**, *49*, 87–98. [CrossRef]
- Rogan, J.; Franklin, J.; Stow, D.; Miller, J.; Woodcock, C.; Roberts, D. Mapping land-cover modifications over large areas: A comparison of machine learning algorithms. *Remote Sens. Environ.* 2008, 112, 2272–2283. [CrossRef]
- 10. Li, C.; Wang, J.; Wang, L.; Hu, L.; Gong, P. Comparison of classification algorithms and training sample sizes in urban land classification with landsat thematic mapper imagery. *Remote Sens.* **2014**, *6*, 964–983. [CrossRef]
- 11. Ma, L.; Li, M.; Ma, X.; Cheng, L.; Du, P.; Liu, Y. A review of supervised object-based land-cover image classification. *ISPRS* **2017**, *130*, 277–293. [CrossRef]
- 12. Blaschke, T. Object based image analysis for remote sensing. ISPRS 2010, 65, 2–16. [CrossRef]
- 13. Benz, U.C.; Hofmann, P.; Willhauck, G.; Lingenfelder, I.; Heynen, M. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for gis-ready information. *ISPRS* **2004**, *58*, 239–258. [CrossRef]
- 14. Huang, Y.; Zhao, C.; Yang, H.; Song, X.; Chen, J.; Li, Z. Feature selection solution with high dimensionality and low-sample size for land cover classification in object-based image analysis. *Remote Sens.* **2017**, *9*, 939. [CrossRef]
- 15. Qian, Y.; Zhou, W.; Yan, J.; Li, W.; Han, L. Comparing machine learning classifiers for object-based land cover classification using very high resolution imagery. *Remote Sens.* **2015**, *7*, 153–168. [CrossRef]
- 16. Lake, B.M.; Salakhutdinov, R.; Tenenbaum, J.B. Human-level concept learning through probabilistic program induction. *Science* **2015**, *350*, 1332–1338. [CrossRef] [PubMed]
- 17. Persello, C.; Bruzzone, L. Active and semisupervised learning for the classification of remote sensing images. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 6937–6956. [CrossRef]
- 18. Duong, P.C.; Trung, T.H.; Nasahara, K.N.; Tadono, T. Jaxa high-resolution land use/land cover map for central vietnam in 2007 and 2017. *Remote Sens.* **2018**, *10*, 1406. [CrossRef]
- 19. Richards, J.A.; Jia, X. Using suitable neighbors to augment the training set in hyperspectral maximum likelihood classification. *IEEE Geosci. Remote Sens. Lett.* **2008**, *5*, 774–777. [CrossRef]
- Du, L.; Wang, Y.; Xie, W. A Semi-supervised Method for Sar Target Discrimination Based on Co-training. In Proceedings of the IGARSS 2019–2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019; IEEE: New York, NY, USA, 2019.
- 21. Pasolli, E.; Melgani, F.; Tuia, D.; Pacifici, F.; Emery, W.J. Svm active learning approach for image classification using spatial information. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 2217–2233. [CrossRef]
- 22. Tuia, D.; Pasolli, E.; Emery, W.J. Using active learning to adapt remote sensing image classifiers. *Remote Sens. Environ.* **2011**, *115*, 2232–2242. [CrossRef]

- 23. Bruzzone, L.; Marconcini, M. Toward the automatic updating of land-cover maps by a domain-adaptation svm classifier and a circular validation strategy. *IEEE Trans. Geosci. Remote Sens.* **2009**, 47, 1108–1122. [CrossRef]
- 24. Zhu, X.; Lafferty, J.; Rosenfeld, R. Semi-supervised Learning with Graphs. Ph.D Thesis, Carnegie Mellon University, Pittsburgh, PA, USA, 2005.
- 25. Wang, L.; Hao, S.; Wang, Q.; Wang, Y. Semi-supervised classification for hyperspectral imagery based on spatial-spectral label propagation. *ISPRS* **2014**, *97*, 123–137.
- 26. Chapelle, O.; Scholkopf, B.; Zien, A. Semi-supervised Learning (Chapelle, O. Et al., Eds.; 2006) [book reviews]. *IEEE Trans. Neural Netw.* **2009**, *20*, 542. [CrossRef]
- Shi, Q.; Du, B.; Zhang, L. Domain adaptation for remote sensing image classification: A low-rank reconstruction and instance weighting label propagation inspired algorithm. *IEEE Trans. Geosci. Remote Sens.* 2015, 53, 5677–5689.
- 28. Gómez, C.; White, J.C.; Wulder, M.A. Optical remotely sensed time series data for land cover classification: A review. *ISPRS* **2016**, *116*, 55–72.
- 29. Zhang, C.; Bengio, S.; Hardt, M.; Recht, B.; Vinyals, O. Understanding deep learning requires rethinking generalization. *arXiv preprint* **2016**, arXiv:1611.03530.
- 30. Hinton, G.E.; Salakhutdinov, R.R. Reducing the dimensionality of data with neural networks. *Science* **2006**, *313*, 504–507.
- 31. Bahdanau, D.; Cho, K.; Bengio, Y. Neural machine translation by jointly learning to align and translate. *arXiv preprint* **2014**, arXiv:1409.0473.
- 32. Huang, B.; Zhao, B.; Song, Y. Urban land-use mapping using a deep convolutional neural network with high spatial resolution multispectral remote sensing imagery. *Remote Sens. Environ.* **2018**, *214*, 73–86. [CrossRef]
- 33. Wang, S.; Quan, D.; Liang, X.; Ning, M.; Guo, Y.; Jiao, L. A deep learning framework for remote sensing image registration. *ISPRS* **2018**, *145*, 148–164. [CrossRef]
- 34. Wang, C.; Zhang, L.; Wei, W.; Zhang, Y. Hyperspectral image classification with data augmentation and classifier fusion. *IEEE Geosci. Remote Sens. Lett.* **2019**, *17*, 1420–1424. [CrossRef]
- 35. Cao, X.; Yao, J.; Xu, Z.; Meng, D. Hyperspectral image classification with convolutional neural network and active learning. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 4604–4616. [CrossRef]
- 36. Hu, Y.; Zhang, Q.; Zhang, Y.; Yan, H. A deep convolution neural network method for land cover mapping: A case study of Qinhuangdao, China. *Remote Sens.* **2018**, *10*, 2053.
- 37. Gaetano, R.; Ienco, D.; Ose, K.; Cresson, R. A two-branch cnn architecture for land cover classification of pan and ms imagery. *Remote Sens.* **2018**, *10*, 1746.
- 38. Felde, G.; Anderson, G.; Cooley, T.; Matthew, M.; Berk, A.; Lee, J. Analysis of Hyperion Data with the FLAASH Atmospheric Correction Algorithm. In Proceedings of the IGARSS 2003 IEEE International Geoscience and Remote Sensing Symposium, Toulouse, France, 21–25 July 2003; IEEE: New York, NY, USA, 2003.
- Laben, C.A.; Brower, B.V. Process for Enhancing the Spatial Resolution of Multispectral Imagery Using Pan-Sharpening. Available online: https://patentimages.storage.googleapis.com/f9/72/45/c9f1fffe687d30/ US6011875.pdf (accessed on 3 August 2020).
- 40. Murray, N.J.; Clemens, R.S.; Phinn, S.R.; Possingham, H.P.; Fuller, R.A. Tracking the rapid loss of tidal wetlands in the yellow sea. *Front. Ecol. Environ.* **2014**, *12*, 267–272. [CrossRef]
- 41. Drăguţ, L.; Csillik, O.; Eisank, C.; Tiede, D. Automated parameterisation for multi-scale image segmentation on multiple layers. *ISPRS* **2014**, *88*, 119–127. [CrossRef]
- 42. Zhu, X.; Liu, D. Improving forest aboveground biomass estimation using seasonal landsat ndvi time-series. *ISPRS* **2015**, *102*, 222–231. [CrossRef]
- 43. Racine, J.S. Nonparametric econometrics: A primer. Found. Trends Econom. 2008, 3, 1–88. [CrossRef]
- 44. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* **2014**, *15*, 1929–1958.
- 45. Gulli, A.; Pal, S. Deep Learning with Keras; Packt: Birmingham, UK, 2017; pp. 1–318.
- 46. Von Luxburg, U. A tutorial on spectral clustering. Stat. Comput. 2007, 17, 395–416. [CrossRef]
- Ng, A.Y.; Jordan, M.I.; Weiss, Y. On spectral clustering: Analysis and an algorithm. In Proceedings of the NIPS'01: Proceedings of the 14th International Conference on Neural Information Processing Systems: Natural and Synthetic, Vancouver, BC, Canada, 3–8 December 2001; MIT Press: Cambridge, MA, USA, 2001.

- Bruneau, P.; Parisot, O.; Otjacques, B. A heuristic for the automatic parametrization of the spectral clustering algorithm. In Proceedings of the 2014 22nd International Conference on Pattern Recognition, Stockholm, Sweden, 24–28 August 2014; IEEE: New York, NY, USA, 2014.
- Dronova, I.; Gong, P.; Wang, L. Object-based analysis and change detection of major wetland cover types and their classification uncertainty during the low water period at poyang lake, China. *Remote Sens. Environ.* 2011, 115, 3220–3236. [CrossRef]
- 50. Liu, J.; Liu, M.; Zhuang, D.; Zhang, Z.; Deng, X. Study on spatial pattern of land-use change in china during 1995–2000. *Sci. China Ser. D* 2003, *46*, 373–384.
- 51. Tuia, D.; Ratle, F.; Pacifici, F.; Kanevski, M.F.; Emery, W.J. Active learning methods for remote sensing image classification. *IEEE Trans. Geosci. Remote Sens.* 2009, 47, 2218. [CrossRef]
- 52. McIver, D.; Friedl, M. Using prior probabilities in decision-tree classification of remotely sensed data. *Remote Sens. Environ.* **2002**, *81*, 253–261. [CrossRef]
- 53. Pal, M.; Mather, P. Some issues in the classification of dais hyperspectral data. *Int. J. Remote Sens.* 2006, 27, 2895–2916. [CrossRef]
- Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet Classification with Deep Convolutional Neural Networks. Available online: http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neuralnetworks.pdf (accessed on 1 August 2020).
- 55. Goodchild, M.F.; Guo, H.; Annoni, A.; Bian, L.; de Bie, K.; Campbell, F.; Craglia, M.; Ehlers, M.; van Genderen, J.; Jackson, D. Next-generation digital earth. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 11088–11094. [CrossRef]
- 56. Liu, J.; Li, J.; Li, W.; Wu, J. Rethinking big data: A review on the data quality and usage issues. *ISPRS* **2016**, *115*, 134–142. [CrossRef]
- 57. Bioucas-Dias, J.M.; Plaza, A.; Camps-Valls, G.; Scheunders, P.; Nasrabadi, N.; Chanussot, J. Hyperspectral remote sensing data analysis and future challenges. *IEEE Geosci. Remote Sens. Mag.* 2013, 1, 6–36. [CrossRef]
- 58. Bruzzone, L.; Demir, B. *Land Use and Land Cover Mapping in Europe*; Springer: Berlin/Heidelberg, Germany, 2014; pp. 127–143.
- 59. Shao, Y.; Lunetta, R.S. Comparison of support vector machine, neural network, and cart algorithms for the land-cover classification using limited training data points. *ISPRS* **2012**, *70*, 78–87.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).