

Article

Improvement of Moderate Resolution Land Use and Land Cover Classification by Introducing Adjacent Region Features

Longlong Yu ^{1,2,*} , Jinhe Su ^{1,2}, Chun Li ^{1,2}, Le Wang ^{1,2}, Ze Luo ¹ and Baoping Yan ¹

¹ Computer Network Information Center, Chinese Academy of Sciences, Beijing 100190, China; sujinh@cnic.cn (J.S.); lichun@cnic.cn (C.L.); wongle@cnic.cn (L.W.); luoze@cnic.cn (Z.L.); ybp@cnic.cn (B.Y.)

² University of Chinese Academy of Sciences, Beijing 100190, China

* Correspondence: yulonglong@cnic.cn; Tel.: +86-182-016-89838

Received: 22 December 2017; Accepted: 7 March 2018; Published: 8 March 2018

Abstract: Landsat-like moderate resolution remote sensing images are widely used in land use and land cover (LULC) classification. Limited by coarser resolutions, most of the traditional LULC classifications that are based on moderate resolution remote sensing images focus on the spectral features of a single pixel. Inspired by the spatial evaluation methods in landscape ecology, this study proposed a new method to extract neighborhood characteristics around a pixel for moderate resolution images. 3 landscape-metric-like indexes, i.e., mean index, standard deviation index, and distance weighted value index, were defined as adjacent region features to include the surrounding environmental characteristics. The effects of the adjacent region features and the different feature set configurations on improving the LULC classification were evaluated by a series of well-controlled LULC classification experiments using K nearest neighbor (KNN) and support vector machine (SVM) classifiers on a Landsat 8 Operational Land Imager (OLI) image. When the adjacent region features were added, the overall accuracies of both the classifiers were higher than when only spectral features were used. For the KNN and SVM classifiers that used only spectral features, the overall accuracies of the LULC classification were 85.45% and 88.87%, respectively, and the accuracies were improved to 94.52% and 96.97%. The classification accuracies of all the LULC types improved. Highly heterogeneous LULC types that are easily misclassified achieved greater improvements. As comparisons, the grey-level co-occurrence matrix (GLCM) and convolutional neural network (CNN) approaches were also implemented on the same dataset. The results revealed that the new method outperformed GLCM and CNN approaches and can significantly improve the classification performance that is based on moderate resolution data.

Keywords: land use and land cover; classification; scale; adjacent region feature; remote sensing; landscape ecology

1. Introduction

Land use and land cover (LULC) information is one of the most essential inputs for environmental monitoring tasks and numerous interdisciplinary studies, including research on climate change and nature conservation, since LULC information is crucial to understanding the complex underlying patterns and mechanisms among natural processes and human activities [1–5]. Remote sensing is capable of providing large scale and long time series information of earth surface. LULC classification based on remote sensing data, which is a basic issue in geographical information system (GIS) fields, is playing an increasingly important role at present [1,4,6,7]. Landsat data are free on the United States Geological Survey (USGS) website for downloading and relevant analysis. Landsat-like moderate spatial resolution images are capable of providing global-scale information on the earth surface and

have been the major data source of LULC classification, especially at large scales [4,7–9]. Furthermore, Landsat data have a remarkable temporal range of over 40 years and have great potential for LULC classification, change detection, and relevant analysis [7,10,11].

Numerous effective methods and advanced classifiers have been applied to improve the performance of LULC classification that is based on moderate resolution data, and most of these methods have been implemented in the feature generation step of LULC classification. Neighborhood characteristics around a pixel provide spatial information that can potentially discriminate LULC types with similar spectral characteristics. For high or very high resolution LULC classification, neighborhood characteristics in different forms have been proven to be beneficial for improving the classification performance. The most popular method is using a grey level occurrence matrix (GLCM) to obtain the texture features such as the mean, contrast, homogeneity, and angular second moment [12–16]. In recent years, numerous studies on high-resolution LULC classification have used powerful state-of-the-art techniques of deep learning methods, such as convolutional neural network (CNN) for object-level classification in a limited area [17–20]. The CNN approach uses convolutional windows and local connections to effectively extract the spatial information. In regard to LULC classification that is based on moderate or coarser resolution images, such as Landsat and MODIS, the CNN approach and texture features may still be advantageous in some special cases, such as cropland classifications or in the extraction of certain LULC types in special regions [21–23]. However, for moderate resolution LULC classification with high thematic resolution, very few researches involved neighborhood characteristics using methods like GLCM and CNN. The improvement of classification accuracy appears to be dependent on the resolution level applied. Chen, Stow, and Gong [24] found that textural features were more effective in improving the classification accuracy of land use classes at finer resolution levels, and when the spatial resolution exceeded a certain level, adding texture did not lead to higher classification accuracy. Thus these methods, which have been proved very effective in high or very high resolution LULC classification, may face challenges at a much coarser resolution. In this situation, most of efforts have concentrated on using multitemporal or multiseasonal data, i.e., more than one remote sensing image is selected to obtain multitemporal features, or phenological characteristics are extracted based on a series of fusion data generated from models [10,25–32]. However, multitemporal methods are still based on the spectra information of a single pixel, i.e., the environment surrounding the pixel, which may be crucial for LULC classification, is not evaluated.

Landscape ecology, which has been widely recognized as a highly interdisciplinary science of spatial heterogeneity, is a newly developed and one of the most active disciplines that involves many novel methods and concepts that differ from traditional ecology [33–36]. The core concepts of landscape ecology consist of spatial pattern, heterogeneity, and scale. Unlike traditional ecology, landscape ecology focuses on the global spatial patterns or the structure features of a region instead of those of a single site to evaluate the conditions or ecological functions. Spatial heterogeneity and the relationships between patterns and process, are popular research aspects in landscape ecology [35,37,38]. In landscape ecology, landscape metrics are defined as quantitative indexes to describe the spatial structures and patterns of a given region at a specific scale. Many landscape metrics in different forms, such as shape metrics, diversity metrics, and area metrics, have been proposed by ecologists to measure the relevant features of a region, and relevant research has demonstrated that scale changing influences the evaluation of landscape metrics [39–46]. The Landscape metrics of a certain region are usually calculated from categorical maps and serve for relevant ecological analysis [39,42,44,45]. Landscape-metric-like indexes may have potential to capture useful neighborhood characteristics for improving the performance of moderate resolution LULC classification.

The objective of this study is to include the neighborhood characteristics around a pixel from a landscape ecology perspective to help improve the performance of LULC classification based on moderate resolution remote sensing images, such as Landsat. The researchers used basic landscape-metric-like indexes, i.e., mean index, standard deviation index, and distance weighted value index to evaluate the surrounding environment. These indexes are in the same or similar forms

with the mean and variation filters, which has been used for decades and proved to be effective for different purposes in digital image processing [47–50]. In this study, we redefined and understood them from a novel perspective of landscape ecology for neighborhood characteristic extraction. As comparisons, the grey-level co-occurrence matrix (GLCM) and convolutional neural network (CNN) approaches were also implemented on the same dataset.

2. Methods

2.1. Study Area

An entire Landsat 8 OLI image (path/row: 123/32, date: 10 July 2017) was selected from the USGS website, and the analysts have prior knowledge of the LULC in the coverage area of the image. The image was utilized as the data source for the control experiments over different feature set configurations. The selected Landsat OLI image was the latest Landsat image with a zero cloud cover rate, i.e., the image is almost not affected by clouds. One single image consists of more than 6×10^6 pixels, which is already extensive data level for machine learning methods in LULC classification. In addition to the simplification, a single image with information from all of the pixels collected under the same atmospheric conditions at same time does not need further processing, such as atmosphere correction, to make digital number (DN) values that are comparable among different remote sensing images [51]. Then, the DN values can be directly used as the feature generation source since they are a specific transformation of the real surface reflectance for all of the pixels in the image, and this method will avoid any possible disturbance to the control experiments. As shown in Figure 1, the image covers the majority of Beijing city, part of Hebei Province, and part of Tianjin city. The typical LULC types, including forest, shrub, grassland, waterbody, cropland, bare land, and urban areas in a temperate climate region make the coverage ideal for the LULC classification experiments.

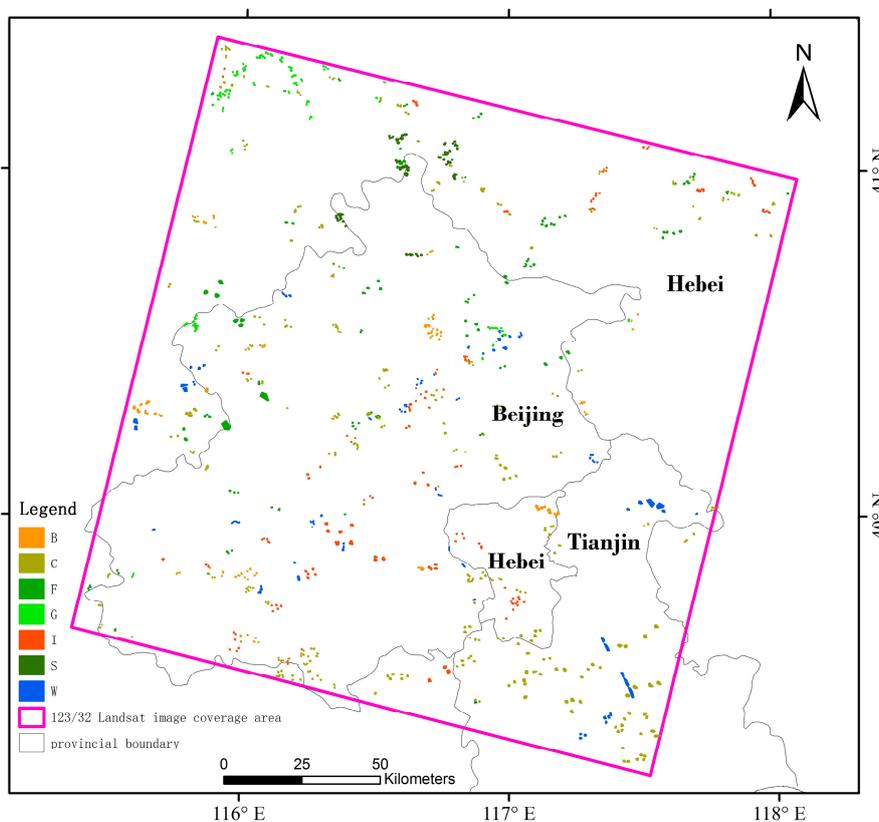


Figure 1. 123/32 Landsat image coverage area and sample polygons for each LULC type. (Note: B: bare land, C: cropland, F: forest, G: grass, I: impervious, S: shrub, W: water).

2.2. Dataset

The classification categories were built based on the previous knowledge of the distribution of LULC types in the study area, including forest, grass, shrub, water, cropland, bare land, and impervious. Sample polygons of each type were built according to the classification categories and their presented visual characteristics in the remote sensing images with the help of experiential knowledge of researchers, high-quality Google Earth images, and the ArcGIS environment. Google Earth, which is a contemporary high-resolution archive, represents a significant, rapidly expanding, cost-free, and largely unexploited resource for scientific inquiry. High-resolution Google Earth imagery has been widely used for assessing moderate resolution remote sensing products [52,53]. All the sample polygons of each type are shown in Figure 1.

Pixels from the same sample polygon tend to have similar large-scale features in adjacent regions, not only because they are of the same classification category, but also because they are in the same site with the same large-scale surroundings. Thus, the classification accuracy may be artificially high if a sample polygon provides both training data and validation data. Thus, for each LULC type, we divided the sample polygons into two groups, half for producing training data and half for validation data. This process indeed generated a stricter validation standard for the generalization capacities of the classifiers.

The cover percentages of each land cover type vary greatly. Some types, such as forest, water, crops, and impervious surfaces are widely distributed in relatively larger patches, while other types are rare and in smaller patches, such as grass and bare land. Thus, the total number of pixels of the different LULC types vary considerably, as shown in Figure 1. To balance the amounts of training and validation data of the different LULC types, 3000 pixels for each LULC type were randomly selected as the training data for the supervised classifiers, and another 3000 pixels were selected as validation data. There are seven LULC types, so there was a total of 42,000 pixels in the sample polygons and 21,000 were used as training data, and the other 21,000 were used for the accuracy assessment. The training data and validation data for each LULC type are from different polygons, as mentioned above. We used the same training data and validation data for all of the classifiers and methods to maintain a strict and absolutely contrasting effect.

All the above processes, including sampling from different groups of polygons for training and validation data, balancing the dataset, and using the same data for different classifiers, are conducive for evaluating the performance of the classifiers with different feature set configurations. Thus, the contributions of the newly included feature sets will be tested more reasonably.

2.3. Adjacent Region Feature Extraction

From a landscape ecology perspective, the attributes of a pixel are strongly related to the surrounding environment. Thus, information on the spatial patterns and heterogeneity of adjacent regions at different scales may be beneficial to the LULC classification of the central pixel. Inspired by the use of landscape metrics/indexes to evaluate the attributes of spatial patterns or the spatial heterogeneity in landscape ecology, we defined 3 basic adjacent region indexes—mean index, standard deviation index, and distance weighted value index—to evaluate the environment surrounding a focus pixel from a landscape perspective. Moving windows were used to extract the features of adjacent regions, and the scale must be taken into consideration when evaluating a region. Relevant research has found that scale may affect the evaluation of a region, i.e., moving windows of different sizes may capture different attributes or potential spatial patterns, which may be useful to distinguish different LULC types [43,45,46]. As an example, Figure 2 shows four adjacent region feature extraction windows surrounding adjacent regions of incremental scales centred on the red focus pixel. Several definitions will be given as follows for unequivocal description in this paper.



Figure 2. An example of moving windows for the extraction of adjacent region feature at different scales of the focus pixel. (Note that this is just a schematic, and not all the feature extraction windows are shown here).

Definition 1. (*Adjacent region*): An adjacent region Z of a pixel is the area in the moving square feature extraction window with a certain size. Note that the given pixel is the central pixel of the moving window.

Definition 2. (*Scale*): The scale S of the adjacent region is denoted by the length of the relative side of the square feature extraction window, which is measured by the number of pixels on one side. S has to be an odd number to ensure that the focus pixel is the geometric centre of the extraction window. The value range of scale S was set to an odd number of the sequence $(1, 2, 3, \dots, 2n - 1)$. Thus, when $S = 1$, the adjacent region is limited to the central pixel itself.

For instance, as shown in Figure 2, we demonstrate four adjacent regions of the central pixel at the scales of 1, 5, 9, and 13.

Definition 3. (*Mean index*): The mean index (MI) of an adjacent region z measures the mean reflectance level in the square feature extraction window with the scale s .

$$MI(z) = \frac{1}{S^2} \sum_{p_i \in z} v_{p_i} \quad (1)$$

where p_i is a pixel in an adjacent region z , and v_{p_i} is the DN value of pixel p_i .

Definition 4. (*Standard deviation index*): The standard deviation index measures the variability of the DN values of all pixels in a square feature extraction window. The standard deviation index is defined to evaluate the degree of variation among the reflectance levels of the adjacent region of a focus pixel with a scale of s .

$$SDI(z) = \sqrt{\frac{1}{S^2} (v_{p_i} - MI(z))^2} \quad (2)$$

Definition 5. (*Distance weighted value index*): The distance weighted value index (DWVI) of an extracted adjacent region is defined based on the principle that the reference value of a pixel in an adjacent region to the central pixel declines when the distance between the two pixels increases. DWVI is defined as follow:

$$DWVI(z) = \frac{\sum_{v_i \in z} d(p_i, c_z) v_{p_i}}{\sum_{v_i \in z} d(p_i, c_z)} \quad (3)$$

Similar to landscape metrics, the three basic adjacent region indexes are defined to capture the component characteristics and spatial information of a region at a certain scale. Landscape metrics are calculated with categorical maps of a certain region, while the three basic adjacent region features can be extracted from the original remote sensing images of the adjacent region.

2.4. Feature Set Configuration and LULC Classification

Scale or moving window size selection is an issue that is worth considering [24,43,46,54]. Higher scales may lead to higher accuracies as well as higher computation costs, so the selection of scale depends on the tradeoffs in different situations. In this study, each Landsat OLI band was used for each pixel in the image. Window sizes from 3 pixels \times 3 pixels to 23 pixels \times 23 pixels were used in this study. With the resolution of 30 m, the length of the largest feature extraction window is 690 m. The largest window covers a 47.61 ha region that is adjacent to the central pixel and is large enough to capture the features of a region adjacent to a site. Moreover, the results that are presented in the next section show that the performances of the classifiers change only slightly once such a high scale is reached.

Among the Landsat OLI data bands, some are designed for special use. The coastal aerosol band is mostly used for coastal water monitoring. The cirrus band is mainly used for cloud detection. These bands are less useful for LULC classification tasks and can always be removed since they contain limited land surface information. The thermal infrared bands are at a much lower resolution of 100 m. Finally, bands 2, 3, 4, 5, 6, and 7 of the Landsat OLI data were selected as the data sources. All processes will be repeated on each selected band for every pixel.

Six raw spectral features from the Landsat OLI bands were first selected as the feature set for the LULC classification, and the performances of the classification were then contrasted with the following classifiers that utilized the features from adjacent regions. As shown in Table 1, to evaluate the effect of different types of adjacent region features at different scales, four experimental sequences with different feature set configurations were designed. For the first sequence, each type of adjacent region feature (MI, or SI, or DWVI) at a certain scale s (1, 2, 3, . . . , 23) were successively added to the feature set. For the second experimental sequence, all of the adjacent region features of a certain type with scales no larger than a given maximum were added to the feature set as a whole. For the third experimental sequence, the three adjacent region features (MI, SI, and DWVI) at a certain scale s were successively added to the feature set. Finally, for the fourth experimental sequence, the three features of the adjacent region (MI, SI, and DWVI) with scales no higher than a given maximum were added to the feature set as a whole. It is worth noting that all of the feature set configurations contain six basic spectral features, since when $S = 1$, there are six spectral features in the feature set. Therefore, in any configuration, the classification performance with $S = 1$ becomes the basic comparative references.

Table 1. Different feature set configurations used in this study.

Experimental Sequences	Feature Dimensions	Feature Set Configurations
Spectral features only	6	6 spectral features ($S = 1$)
Sequence 1	12	6 spectral + MI or SDI or DWVI features ($S = s$)
Sequence 2	$6 + (s + 1) \times 3$	6 spectral + MI or SDI or DWVI features ($S \leq s$)
Sequence 3	24	6 spectral + MI & SDI & DWVI features ($S = s$)
Sequence 4	$6 + (0.5 \times (s + 1) - 1) \times 18$	6 spectral + MI & SDI & DWVI features ($S \leq s$)

A KNN classifier based on the K-nearest neighbour algorithm and a support vector machine (SVM) classifier based on the support vector machine algorithm were selected for LULC classification. The KNN classifier is one of the most fundamental but effective non-parametric classification methods [55]. SVM aims to find an optimal hyperplane between classes and addresses if the data are linearly separable or not. An SVM classifier usually performs well and is one of the most widely used

advanced discrimination models that has been used in numerous classification tasks containing LULC classification [56,57].

2.5. Comparable Methods

To compare the new feature extraction method with other methods, GLCM and CNN approaches were also implemented using the same data.

Specifically, for the GLCM approach, the mean, contrast, homogeneity, and angular second moment were extracted from the GLCM. These texture features have proved to be effective in classification. Additionally, mean, homogeneity, and angular second moment are least correlated with each other [58]. GLCM texture features with different configurations on different scales were evaluated in the similar way in Table 1 to find the best-performed configuration.

For CNN approach, we referenced the architecture of GoogLeNet Inception that achieves the new state of the art for classification and detection in the ImageNet Large-Scale Visual Recognition Challenge 2014 (ILSVRC14) [59]. The “Inception modules” combines multiscale convolution kernels with a parallel pooling path in each such stage, which has an additional beneficial effect. It is worthy to note that there two different type of moving window. One type is the same with GLCM texture window and adjacent region defined above in this study. Another type is the convolution window used in CNN architecture. We proposed a similar strategy with GoogLeNet Inception version 1. Two “Inception modules” were stacked in the final CNN architecture which we designed for pixel based LULC classification, as is shown in Figure 3. As the biggest convolution window is 5×5 , the smallest adjacent region scale we used in CNN approach will be 5×5 as well. Thus, to maintain a contrasting effect, the scale range 5 to 23 is evaluated by CNN approach.

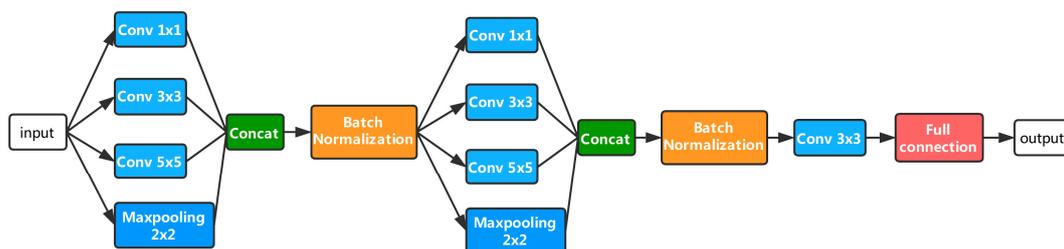


Figure 3. Convolutional neural network (CNN) architecture used in this study.

2.6. Accuracy Assessment

The accuracies of the LULC classification were evaluated using a confusion matrix to calculate the overall accuracy, kappa coefficient, producer’s accuracy, and user’s accuracy [60]. All of the classifiers with different feature set configurations used the same training and testing datasets. Hence, we implemented a series of standard control experiments to evaluate the effect of adjacent region features. Furthermore, in order to access statistical differences between the accuracy measurements of classification results using different approaches, a Z-test was performed to see if they were significantly different [61].

3. Results

Figure 4 shows how the overall accuracy changed with different feature set configurations. The results show that, in any cases, when adjacent region features were added, both classifiers (KNN and SVM) achieve considerable performance improvements. In addition, when higher scale adjacent region features were added, the improvement continued until a relative gently change range was reached when the accuracy of the LULC classification was very high. For the KNN classifier, the overall accuracy of the LULC classification with only spectral features was 85.45% and the highest overall accuracy was 94.52%, which was achieved when all of the adjacent region features were

introduced in an accumulated-scale with the largest scale of 19. At the same time, the kappa coefficient improved from 0.83 to 0.94. For the SVM classifier, the improvement was from 88.87% to the highest overall accuracy of 96.97% when the configuration was the combination of all there adjacent region features at a scale of 17. The kappa coefficient improved from 0.87 to 0.96. Figure 4 also indicates that, in general, when the three adjacent region metrics were evaluated, the DWVI and MI outperformed the SDI.

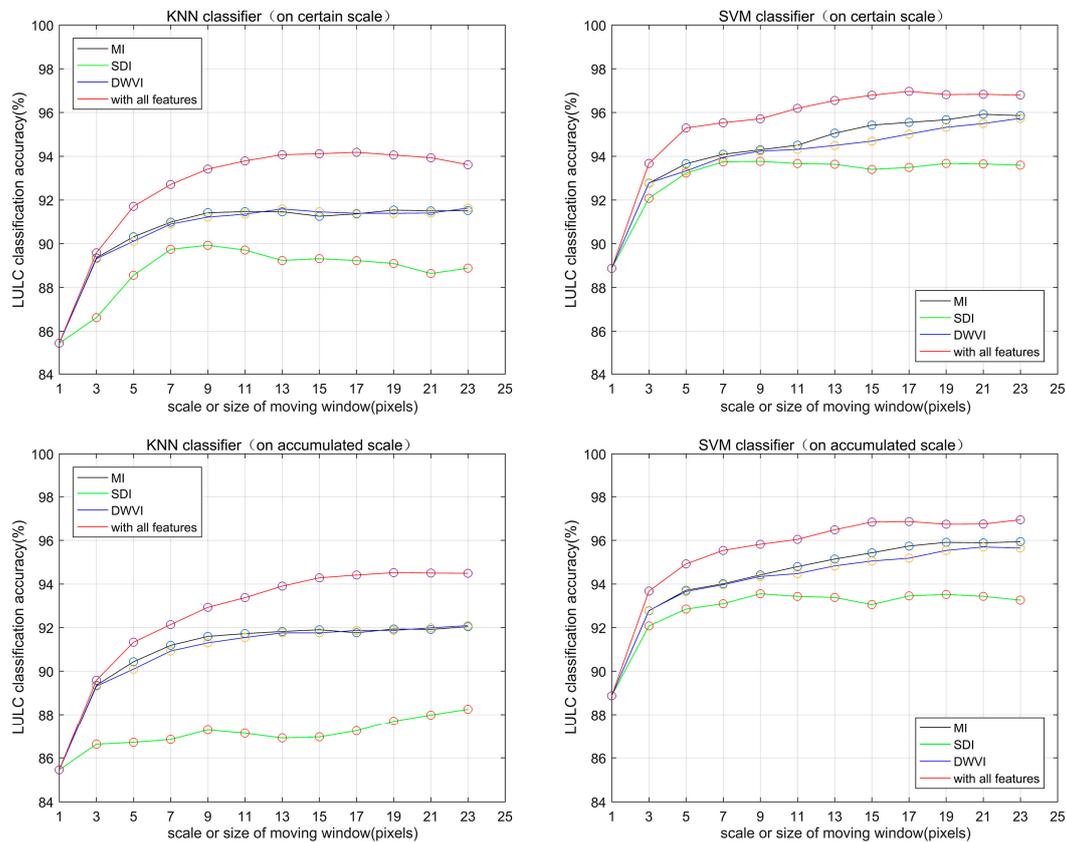


Figure 4. Overall accuracies of the land use and land cover (LULC) classifications with different feature set configurations.

To show more details, the confusion matrixes of the classification results using only spectral features and the best performing feature set configurations with both KNN and SVM classifiers were shown in Tables 2 and 3. The large improvements indicated the effectiveness of utilizing adjacent region features, especially for the LULC types that had lower accuracy when only spectral features were used, such as crop land, forest, and shrub, as shown in Tables 2 and 3. For instance, with the KNN and SVM classifiers, the user's accuracy of crop land improved from 76.7% to 92.5% and from 72.4% to 84.6%, respectively; the producer's accuracy improved from 79.7% to 91.4%, and from 77.6% to 92.6%, respectively. The results indicated that these two types are quite challenging for even an advanced classifier using only spectral features. These LULC types, such as crop land, are highly heterogeneous since numerous crop species show quite different optical characteristics. But, as a whole, one LULC type has global spatial pattern features that will not be captured when only one pixel is evaluated. As clearly shown in the results, these features may be crucial for improving the performance of LULC classification based on Landsat-like moderate resolution data.

Table 2. Confusion matrixes of LULC Classification using K nearest neighbor (KNN) classifiers with only spectral information and the configuration that achieved the highest overall accuracy (94.52%). (Note: B: bare land, C: cropland, F: forest, G: grass, I: impervious, S: shrub, W: water, UA: user's accuracy, PA: producer's accuracy).

	Ground Truth (Pixels)							UA (%)	PA (%)	
	B	C	F	G	I	S	W			Total
Mapped Classes (Pixels)										
<i>Using Only Spectral Features with KNN</i>										
B	2648	141	2	1	139	12	0	2943	90.0	88.3
C	151	2172	21	82	160	200	47	2833	76.7	72.4
F	34	49	2610	17	2	562	1	3275	79.7	87.0
G	2	80	34	2724	2	47	0	2889	94.3	90.8
I	126	63	0	0	2662	1	0	2852	93.3	88.7
S	39	419	331	176	16	2178	0	3159	69.0	72.6
W	0	76	2	0	19	0	2952	3049	96.8	98.4
<i>Using All Adjacent Region Features (MI+SDI+DWVI) with KNN on Accumulated Scales (s ≤ 19)</i>										
B	2944	46	0	5	65	0	0	3060	96.2	98.1
C	11	2537	16	48	77	61	23	2773	92.5	84.6
F	0	82	2863	102	0	63	1	3111	92.0	95.4
G	0	95	98	2815	0	4	0	3012	93.5	93.8
I	35	18	0	0	2843	0	0	2896	98.2	94.5
S	10	212	23	30	0	2872	0	3147	91.3	95.7
W	0	10	0	0	15	0	2976	3001	99.2	99.2
Total	3000	3000	3000	3000	3000	3000	3000	21,000		

Table 3. Confusion matrixes of LULC Classification using support vector machine (SVM) classifiers with only spectral information and the configuration that achieved the highest overall accuracy (96.97%). (Note: B: bare land, C: cropland, F: forest, G: grass, I: impervious, S: shrub, W: water, UA: user's accuracy, PA: producer's accuracy).

	Ground Truth (Pixels)							UA (%)	PA (%)	
	B	C	F	G	I	S	W			Total
Mapped Classes (Pixels)										
<i>Using Only Spectral Features with SVM</i>										
B	2734	159	6	0	95	12	0	3006	91.0	91.1
C	149	2329	10	92	111	93	38	2822	79.7	77.6
F	13	25	2667	1	0	496	0	2914	91.5	88.9
G	1	57	1	2804	3	20	0	3202	87.6	93.5
I	88	50	0	1	2789	0	0	2928	95.3	93.0
S	15	336	496	102	2	2379	0	3122	76.2	79.3
W	0	44	0	0	0	0	2962	3006	98.5	98.7
<i>Using All Adjacent Region Features (MI+SDI+DWVI) with SVM on Certain Scale (s = 17)</i>										
B	2973	91	0	14	49	0	0	3127	95.1	99.1
C	20	2779	13	37	15	159	18	3041	91.4	92.6
F	0	22	2955	2	0	30	0	3009	98.2	98.5
G	0	61	15	2941	0	11	0	3028	97.1	98.0
I	7	2	0	2	2936	0	2	2949	99.6	97.9
S	0	45	17	4	0	2800	0	2866	97.7	93.3
W	0	0	0	0	0	0	2980	2980	100	99.3
Total	3000	3000	3000	3000	3000	3000	3000	21,000		

Regarding the performances of the different classifiers that were applied in this study, Figure 5 compares the overall accuracies of the LULC classifications using KNN and SVM classifiers. With different feature set configurations, the advanced SVM classifier outperformed KNN classifier in all

cases. Generally, the overall improvements of the different feature set configurations were similar as the scale increased.

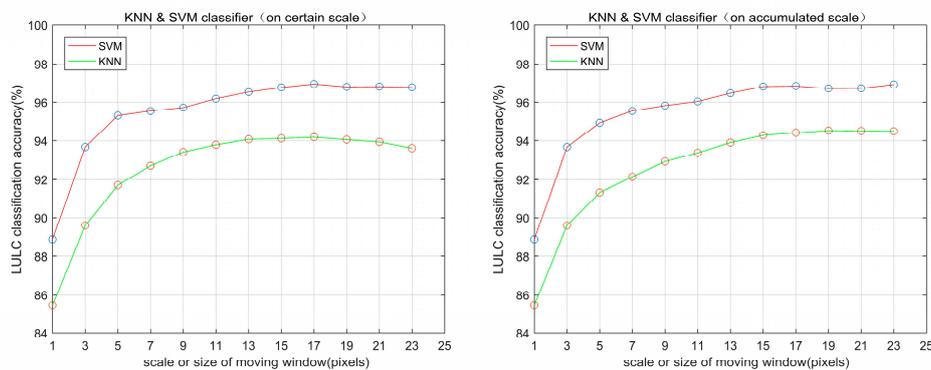


Figure 5. Overall accuracies of the LULC classification with different feature set configurations in the SVM and KNN classifiers.

The LULC classification results of both the KNN and SVM classifiers using only spectral features and using the best performing feature set configurations containing adjacent region features are shown in Figure 6. All of the results reflect the distribution of different LULC types in the Landsat image coverage area based on experts' knowledge. In the northwest mountainous areas, forests, shrubs, and grasses account for a large proportion. Cropland and impervious surfaces are primarily distributed in the southeast rural and urban areas.

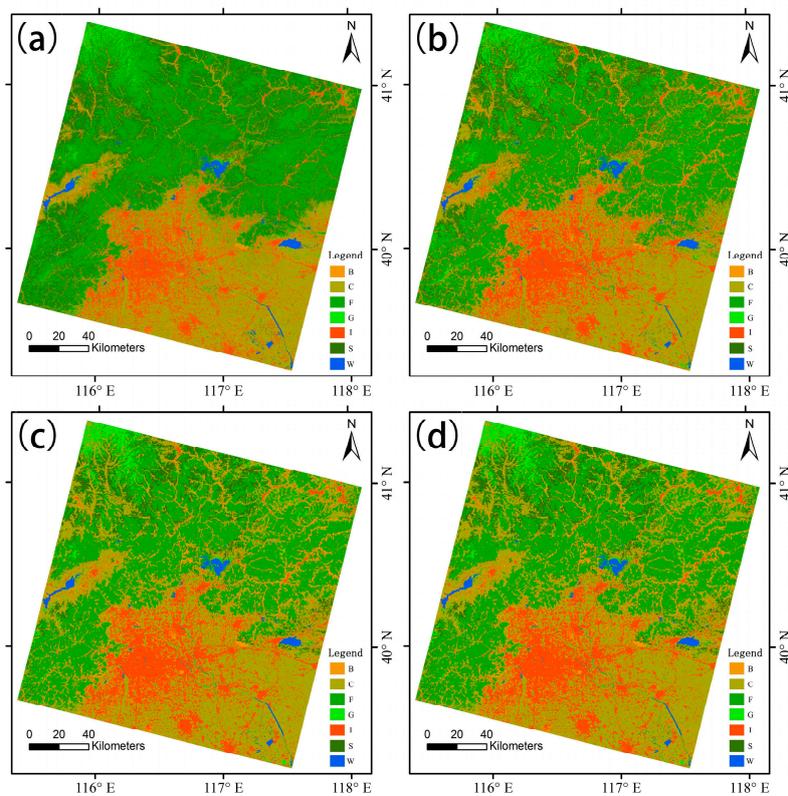


Figure 6. LULC classification maps. (Note: B: bare land, C: cropland, F: forest, G: grass, I: impervious, S: shrub, W: water; (a) KNN + spectral features only; (b) KNN + spectral features + all adjacent region features ($S \leq 19$); (c) SVM + spectral features only; and, (d) SVM + spectral features + all adjacent region features ($S = 17$)).

To demonstrate the effect of adjacent region features in a more visible way, the classification results of a representative region are shown in Figure 7 as a typical example. The maps show that when adjacent region features are included, the classifier tended to yield more aggregated class objects, and fragmented patches were reduced. The reason may be that the classifier captured neighborhood characteristics and became more tolerant to the variance within classes. Thus, the classifier was less likely to misclassify and split a single-class region, especially for those LULC classes with high degree of heterogeneity, such as shrub and crop land. The overall accuracy of LULC classification obtained substantial improvements, and more importantly, the spatial distribution and the pattern of the classification results are more similar to the sample polygons that are based on expert knowledge, which is a more reasonable result that is desired by the LULC classification task.

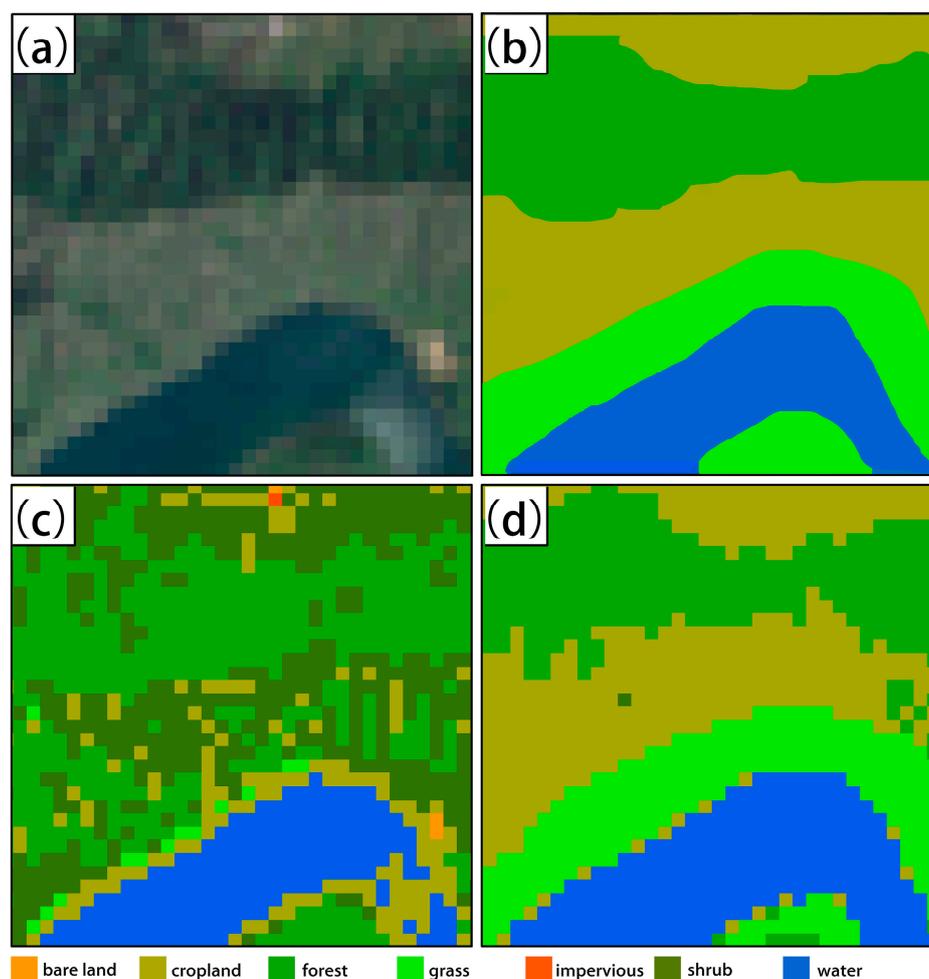


Figure 7. LULC classification of a representative region: (a) Landsat image in true color; (b) sample polygons based on expert knowledge; (c) SVM with spectral features only; (d) SVM + spectral features + all adjacent region features ($S = 17$).

Figure 8 shows the overall accuracy curves with the best performed configurations using different feature extraction methods (GLCM, CNN, and adjacent region features). For the GLCM approach, the highest overall accuracy was 92.03%, which was achieved on SVM classifier when angular second moment features were introduced at a scale of 17. For the CNN approach, the highest overall accuracy was 94.58%, which was achieved at a scale of 19. As previously stated, for the newly proposed method using adjacent region features, the highest overall accuracy was 96.97%, which was achieved on SVM classifier when the configuration was the combination of all three adjacent region features at a scale

of 17. The result indicated that when it comes to moderate resolution LULC classification, the newly proposed method outperformed GLCM and CNN approaches in all cases that we have evaluated.

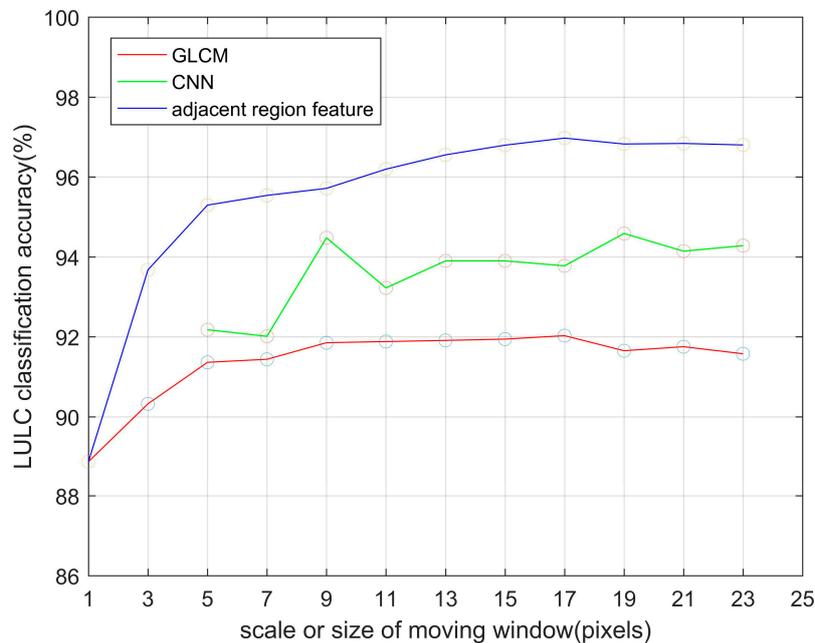


Figure 8. Performances of grey level occurrence matrix (GLCM), CNN, and adjacent region features with feature set configurations that achieved the highest overall accuracies.

To determine whether the classification accuracies of different approaches were significantly different, the Z-test was used to compare the confusion matrixes. $Z > 1.96$ or $Z < -1.96$ would indicate the difference of the two confusion matrixes being significant at the 5% significance level [61]. As shown in Table 4, the Z-test value for comparison between the confusion matrixes of classification result using the best performed configurations using different feature extraction methods (GLCM, CNN and adjacent region features) are all larger than 1.96. The results indicate that GLCM, CNN, and the proposed method could significantly improve the land cover classification accuracy of using spectral features only. These findings also indicate that the classification accuracy improvement achieved by adding adjacent region features is greater than that of GLCM and CNN, and the performances were significantly different.

Table 4. Z-test values for comparison between confusion matrixes of LULC classification using different approaches with the best performed configurations.

Classification Methods	Spectral Features Only	GLCM	CNN	With Adjacent Region Features
Spectral features only	-	-	-	-
GLCM	7.76	-	-	-
CNN	12.94	5.52	-	-
With adjacent region features	13.77	8.32	4.37	-

4. Discussion

Neighbourhood characteristics around a pixel are important for LULC classification using remote sensing data [12,14–18,20,22–24]. However, neighbourhood characteristics extraction methods, which have proved very effective in high resolution classification, have not been commonly used in moderate

resolution LULC classification. This study proposed a novel method to evaluate the adjacent region around a pixel inspired by the landscape metrics.

In the extraction of neighbourhood characteristics, moving window size, which is defined as scale in this study, is an important factor influencing classification accuracy. How to select the optical window size or scale is a notable issue. A number of studies have been conducted to discuss this problem in both image classification area and landscape ecology [24,43,46,54]. A bigger window may include more useful features for classification, but it also may include more redundant information that are not beneficial or even have negative effect. As is shown in Figure 8, the performances of all the approaches that were used in this study were affected by changing window size. For GLCM, CNN and our new proposed methods, the window sizes or the scales, where the highest accuracies were achieved, were 17, 9 and 17, respectively.

The results that were obtained using different classifiers applied in this study reveal that the advanced non-parametric classifier SVM could achieve a more satisfactory classification result than the basic KNN classifier. Regarding different approaches applied in this study for neighbourhood characteristic extraction, the adjacent region features performed better than the GLCM and CNN approaches. How can these basic statistical indices outperform more complicated texture feature methods and powerful state-of-the-art techniques of CNN? The author's viewpoint is that at moderate or coarser resolution, the spectral features of a much bigger pixel may be the mixture of different reflection levels. The object-level texture features, which can be effectively captured by GLCM and CNN, tend to be smoothed and the higher scale texture features often reflect the topographical change, which has little reference value to the classification of a single pixel. However, Neighbourhood characteristics in other forms may be still beneficial. In landscape ecology, patch level and regional level statistical features are crucial to describe a landscape type.

In general, the DWVI and MI outperformed the SDI when the three adjacent region indexes were evaluated. The MI and DWVI are more like component indexes for measuring the component characteristics of an adjacent region or the surrounding environment, whereas the SDI is more like a texture feature for evaluating the variation pattern. The results also indicated that at a much coarser resolution, texture-like features may not lead to satisfactory improvements on LULC classification accuracy.

This study is designed to investigate the effects of adjacent region features, which are extracted from multiscale regions around a pixel from a landscape ecology perspective, on improving LULC classification using moderate resolution remote sensing data such as Landsat. Three newly defined basic adjacent region features were evaluated and compared with GLCM and CNN approaches. However, as there are numerous landscape metrics, more forms of adjacent region features will be evaluated on other advanced classifiers in future research.

5. Conclusions

At a much coarser resolution, LULC classification based on moderate resolution remote sensing data like Landsat images has limitations in terms of using neighborhood characteristics for accuracy improvements when compared to similar methods that were designed for very high or very high resolution remote sensing images, such as the GLCM and deep CNN approaches. Inspired by the concepts and methods in landscape ecology, three landscape-metric-like indexes were defined to include adjacent region features into the moderate resolution LULC classification. This study implemented a series of well-controlled experiments of LULC classification with different configurations to investigate the efficacy of adjacent region features to improve the LULC classification performances of Landsat-like moderate resolution remote sensing data. The study indicated that adjacent region features that contain environmental information around pixels greatly improved the LULC classification accuracy with different classifiers. For KNN and SVM classifiers, the best overall classification accuracies that were achieved were 94.52% and 96.97%, and the improvement were over 9% and 8%, respectively. Moreover, when compared to the GLCM and CNN approaches, the proposed

method achieve much better improvements. The results reveal that adjacent region features are beneficial for moderate resolution LULC classification and can significantly improve the classification performance. The proposed methods have great potential for improving LULC classification based on Landsat-like moderate resolution remote sensing data, especially for large temporal and spatial scale analyses.

Acknowledgments: Three anonymous referees gave the authors constructive comments that lead to great improvements in the presentation of this article. This research was partially supported by Open Research Fund Program of State key Laboratory of Hydrosience and Engineering (sklhse-2017-B-03), the Natural Science Foundation of China under Grant No. 61361126011 and No.90912006, and the Special Project of Informatization of Chinese Academy of Sciences in “the Twelfth Five-Year Plan” under Grant No. XXH12504-1-06. The authors also thank USGS project (<https://www.usgs.gov/>) for providing free Landsat data.

Author Contributions: Longlong Yu and Jinhe Su designed and performed the experiments; all authors wrote the paper together.

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

References

- Justice, C.O.; Vermote, E.; Townshend, J.R.G.; Defries, R.; Roy, D.P.; Hall, D.K.; Salomonson, V.V.; Privette, J.L.; Riggs, G.; Strahler, A.; et al. The Moderate Resolution Imaging Spectroradiometer (MODIS): Land remote sensing for global change research. *IEEE Trans. Geosci. Remote Sens.* **1998**, *36*, 1228–1249. [[CrossRef](#)]
- Pearson, R.G.; Dawson, T.P.; Liu, C. Modelling species distributions in Britain: A hierarchical integration of climate and land-cover data. *Ecography* **2004**, *27*, 285–298. [[CrossRef](#)]
- Pielke, R.A. Land use and climate change. *Science* **2005**, *310*, 1625–1626. [[CrossRef](#)] [[PubMed](#)]
- Souza, C.M., Jr.; Siqueira, J.V.; Sales, M.H.; Fonseca, A.V.; Ribeiro, J.G.; Numata, I.; Cochrane, M.A.; Barber, C.P.; Roberts, D.A.; Barlow, J.; et al. Ten-year Landsat classification of deforestation and forest degradation in the Brazilian Amazon. *Remote Sens.* **2013**, *5*, 5493–5513. [[CrossRef](#)]
- Ren, G.; Young, S.S.; Wang, L.; Wang, W.; Long, Y.; Wu, R.; Li, J.; Zhu, J.; Yu, D.W. Effectiveness of China’s National Forest Protection Program and nature reserves. *Conserv. Biol.* **2015**, *29*, 1368–1377. [[CrossRef](#)] [[PubMed](#)]
- 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](#)]
- Roy, D.P.; Wulder, M.A.; Loveland, T.R.; Woodcock, C.E.; Allen, R.G.; Anderson, M.C.; Helder, D.; Irons, J.R.; Johnson, D.M.; Kennedy, R.; et al. Landsat-8: Science and product vision for terrestrial global change research. *Remote Sens. Environ.* **2014**, *145*, 154–172. [[CrossRef](#)]
- Hansen, M.C.; Potapov, P.V.; Moore, R.; Hancher, M.; Turubanova, S.A.; Tyukavina, A.; Thau, D.; Stehman, S.V.; Goetz, S.J.; Loveland, T.R.; et al. High-resolution global maps of 21st-century forest cover change. *Science* **2013**, *342*, 850–853. [[CrossRef](#)] [[PubMed](#)]
- Powell, S.L.; Pflugmacher, D.; Kirschbaum, A.A.; Kim, Y.; Cohen, W.B. Moderate resolution remote sensing alternatives: A review of Landsat-like sensors and their applications. *J. Appl. Remote Sens.* **2007**, *1*. [[CrossRef](#)]
- Yuan, F.; Sawaya, K.E.; Loeffelholz, B.C.; Bauer, M.E. Land cover classification and change analysis of the Twin Cities (Minnesota) Metropolitan Area by multitemporal Landsat remote sensing. *Remote Sens. Environ.* **2005**, *98*, 317–328. [[CrossRef](#)]
- Xian, G.; Homer, C.; Fry, J. Updating the 2001 National Land Cover Database land cover classification to 2006 by using Landsat imagery change detection methods. *Remote Sens. Environ.* **2009**, *113*, 1133–1147. [[CrossRef](#)]
- Zhang, Q.; Wang, J.; Gong, P.; Shi, P. Study of urban spatial patterns from SPOT panchromatic imagery using textural analysis. *Int. J. Remote Sens.* **2003**, *24*, 4137–4160. [[CrossRef](#)]
- Pacifici, F.; Chini, M.; Emery, W.J. A neural network approach using multi-scale textural metrics from very high-resolution panchromatic imagery for urban land-use classification. *Remote Sens. Environ.* **2009**, *113*, 1276–1292. [[CrossRef](#)]
- Shaban, M.; Dikshit, O. Improvement of classification in urban areas by the use of textural features: The case study of Lucknow city, Uttar Pradesh. *Int. J. Remote Sens.* **2001**, *22*, 565–593. [[CrossRef](#)]
- Haralick, R.M.; Shanmugam, K. Textural features for image classification. *IEEE Trans. Syst. Man Cybern.* **1973**, *6*, 610–621. [[CrossRef](#)]

16. Smits, P.C.; Annoni, A. Updating land-cover maps by using texture information from very high-resolution space-borne imagery. *IEEE Trans. Geosci. Remote Sens.* **1999**, *37*, 1244–1254. [[CrossRef](#)]
17. Berberoglu, S.; Curran, P.J.; Lloyd, C.D.; Atkinson, P.M. Texture classification of Mediterranean land cover. *Int. J. Appl. Earth Obs. Geoinf.* **2007**, *9*, 322–334. [[CrossRef](#)]
18. Agüera, F.; Aguilar, F.J.; Aguilar, M.A. Using texture analysis to improve per-pixel classification of very high resolution images for mapping plastic greenhouses. *ISPRS J. Photogramm. Remote Sens.* **2008**, *63*, 635–646. [[CrossRef](#)]
19. Guidici, D.; Clark, M. One-Dimensional Convolutional Neural Network Land-Cover Classification of Multi-Seasonal Hyperspectral Imagery in the San Francisco Bay Area, California. *Remote Sens.* **2017**, *9*, 629. [[CrossRef](#)]
20. Scott, G.J.; Curran, P.J.; Lloyd, C.D.; Atkinson, P.M. Training Deep Convolutional Neural Networks for Land-Cover Classification of High-Resolution Imagery. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 549–553. [[CrossRef](#)]
21. Kussul, N.; Lavreniuk, M.; Skakun, S.; Shelestov, A. Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 778–782. [[CrossRef](#)]
22. Yu, L.; Wang, Z.; Tian, S.; Ye, F.; Ding, J.; Kong, J. Convolutional Neural Networks for Water Body Extraction from Landsat Imagery. *Int. J. Comput. Intell. Appl.* **2017**, *16*, 1750001. [[CrossRef](#)]
23. Ishii, T.; Nakamura, R.; Nakada, H.; Mochizuki, Y.; Ishikawa, H. Surface object recognition with CNN and SVM in Landsat 8 images. In Proceedings of the IAPR International Conference on Machine Vision Applications, Tokyo, Japan, 18–22 May 2015.
24. Chen, D.; Stow, D.A.; Gong, P. Examining the effect of spatial resolution and texture window size on classification accuracy: An urban environment case. *Int. J. Remote Sens.* **2004**, *25*, 2177–2192. [[CrossRef](#)]
25. Langley, S.K.; Cheshire, H.M.; Humes, K.S. A comparison of single date and multitemporal satellite image classifications in a semi-arid grassland. *J. Arid Environ.* **2001**, *49*, 401–411. [[CrossRef](#)]
26. Feng, G.; Masek, J.; Schwaller, M.; Hall, F. On the blending of the Landsat and MODIS surface reflectance: Predicting daily Landsat surface reflectance. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 2207–2218. [[CrossRef](#)]
27. Jia, K.; Liang, S.; Zhang, N.; Wei, X.; Gu, X.; Zhao, X.; Yao, Y.; Xie, X. Land cover classification of finer resolution remote sensing data integrating temporal features from time series coarser resolution data. *ISPRS J. Photogramm. Remote Sens.* **2014**, *93*, 49–55. [[CrossRef](#)]
28. Jia, K.; Liang, S.; Wei, X.; Yao, Y.; Su, Y.; Jiang, B.; Wang, X. Land Cover Classification of Landsat Data with Phenological Features Extracted from Time Series MODIS NDVI Data. *Remote Sens.* **2014**, *6*, 11518–11532. [[CrossRef](#)]
29. Jia, K.; Liang, S.; Wei, X.; Yao, Y.; Su, Y.; Jiang, B.; Wang, X. Land cover classification using Landsat 8 Operational Land Imager data in Beijing, China. *Geocarto Int.* **2014**, *29*, 941–951. [[CrossRef](#)]
30. Zhu, Z.; Woodcock, C.E. Continuous change detection and classification of land cover using all available Landsat data. *Remote Sens. Environ.* **2014**, *144*, 152–171. [[CrossRef](#)]
31. Senf, C.; Leitão, P.J.; Pflugmacher, D.; van der Linden, S.; Hostert, P. Mapping land cover in complex Mediterranean landscapes using Landsat: Improved classification accuracies from integrating multi-seasonal and synthetic imagery. *Remote Sens. Environ.* **2015**, *156*, 527–536. [[CrossRef](#)]
32. Kussul, N.; Skakun, S.; Shelestov, A.; Lavreniuk, M.; Yailymov, B.; Kussul, O. Regional scale crop mapping using multi-temporal satellite imagery. *ISPRS Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2015**, *XL-7/W3*, 45–52. [[CrossRef](#)]
33. Forman, R.T.T.; Godron, M. Patches and Structural Components for A Landscape Ecology. *BioScience* **1981**, *31*, 733–740.
34. Forman, R.T.; Godron, M. *Landscape Ecology*; Jhon Wiley & Sons: New York, NY, USA, 1986; 619p.
35. Forman, R.T.T. Some general principles of landscape and regional ecology. *Landsc. Ecol.* **1995**, *10*, 133–142. [[CrossRef](#)]
36. Wu, J. Landscape Ecology, Cross-disciplinarity, and Sustainability Science. *Landsc. Ecol.* **2006**, *21*, 1–4. [[CrossRef](#)]
37. Wu, J.; Hobbs, R. Key issues and research priorities in landscape ecology: An idiosyncratic synthesis. *Landsc. Ecol.* **2002**, *17*, 355–365. [[CrossRef](#)]
38. Turner, M.G. Landscape Ecology: What Is the State of the Science? *Annu. Rev. Ecol. Evol. Syst.* **2005**, *36*, 319–344. [[CrossRef](#)]

39. Herold, M.; Menz, G. Landscape metric signatures (LMS) to improve urban land use information derived from remotely sensed data. In *A Decade of Trans-European Remote Sensing Cooperation*; CRC Press: Boca Raton, FL, USA, 2001; pp. 251–256.
40. Wu, J.; Shen, W.; Sun, W.; Paul, T.; Tueller, P.T. Empirical patterns of the effects of changing scale on landscape metrics. *Landscape Ecol.* **2002**, *17*, 761–782. [[CrossRef](#)]
41. McGarigal, K.; Cushman, S.A.; Neel, M.C.; Ene, E. *FRAGSTATS: Spatial Pattern Analysis Program for Categorical Maps*; Umass Landscape Ecology Lab: Amherst, MA, USA, 2002.
42. Chen, L.; Fu, B.-J.; Xu, J.Y.; Gong, J. Location-weighted landscape contrast index: A scale independent approach for landscape pattern evaluation based on Source-Sink ecological processes. *Acta Ecol. Sin.* **2003**, *23*, 2406–2413.
43. Wu, J. Effects of changing scale on landscape pattern analysis: Scaling relations. *Landscape Ecol.* **2004**, *19*, 125–138. [[CrossRef](#)]
44. Sofia, G.; Marinello, F.; Tarolli, P. A new landscape metric for the identification of terraced sites: The Slope Local Length of Auto-Correlation (SLLAC). *ISPRS J. Photogramm. Remote Sens.* **2014**, *96*, 123–133. [[CrossRef](#)]
45. McGarigal, K. FRAGSTATS Help. Available online: <https://www.umass.edu/landeco/research/fragstats/documents/fragstats.help.4.2.pdf> (accessed on 8 March 2018).
46. Turner, M.G.; O'Neill, R.V.; Gardner, R.H.; Milne, B.T. Effects of changing spatial scale on the analysis of landscape pattern. *Landscape Ecol.* **1989**, *3*, 153–162. [[CrossRef](#)]
47. Hall, E. *Computer Image Processing and Recognition*; Elsevier: Amsterdam, The Netherlands, 1979.
48. Lee, J.-S. Refined filtering of image noise using local statistics. *Comput. Graph. Image Process.* **1981**, *15*, 380–389. [[CrossRef](#)]
49. Roesser, R. A discrete state-space model for linear image processing. *IEEE Trans. Autom. Control* **1975**, *20*, 1–10. [[CrossRef](#)]
50. Tanimoto, S.; Pavlidis, T. A hierarchical data structure for picture processing. *Comput. Graph. Image Process.* **1975**, *4*, 104–119. [[CrossRef](#)]
51. Kawata, Y.; Ohtani, A.; Kusaka, T.; Ueno, S. Classification accuracy for the MOS-1 MESSR data before and after the atmospheric correction. *IEEE Trans. Geosci. Remote Sens.* **1990**, *28*, 755–760. [[CrossRef](#)]
52. Potere, D. Horizontal positional accuracy of Google Earth's high-resolution imagery archive. *Sensors* **2008**, *8*, 7973–7981. [[CrossRef](#)] [[PubMed](#)]
53. Knorn, J.; Rabe, A.; Radeloff, V.C.; Kuemmerle, T.; Kozak, J.; Hostert, P. Land cover mapping of large areas using chain classification of neighboring Landsat satellite images. *Remote Sens. Environ.* **2009**, *113*, 957–964. [[CrossRef](#)]
54. Liu, J.; Liu, H.; Janne, H.; Matti, M. Posterior probability-based optimization of texture window size for image classification. *Remote Sens. Lett.* **2014**, *5*, 753–762. [[CrossRef](#)]
55. Peterson, L.E. K-nearest neighbor. *Scholarpedia* **2009**, *4*, 1883. [[CrossRef](#)]
56. Pal, M.; Mather, P. Support vector machines for classification in remote sensing. *Int. J. Remote Sens.* **2005**, *26*, 1007–1011. [[CrossRef](#)]
57. Poursanidis, D.; Chrysoulakis, N.; Mitraka, Z. Landsat 8 vs. Landsat 5: A comparison based on urban and peri-urban land cover mapping. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *35*, 259–269. [[CrossRef](#)]
58. Murray, H.; Lucieer, A.; Williams, R. Texture-based classification of sub-Antarctic vegetation communities on Heard Island. *Int. J. Appl. Earth Obs. Geoinf.* **2010**, *12*, 138–149. [[CrossRef](#)]
59. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Boston, MA, USA, 7–12 June 2015.
60. Foody, G.M. Status of land cover classification accuracy assessment. *Remote Sens. Environ.* **2002**, *80*, 185–201. [[CrossRef](#)]
61. Foody, G.M. Classification accuracy comparison: Hypothesis tests and the use of confidence intervals in evaluations of difference, equivalence and non-inferiority. *Remote Sens. Environ.* **2009**, *113*, 1658–1663. [[CrossRef](#)]

