Random Forest Classification of Remote Sensing Images Based on Superpixel Statistics (Postprint)

Authors: Shi Caixia, Zhao Chuangang, Pang Lei

Date: 2018-05-20T00:00:00+00:00

Abstract

To address the insufficient utilization of spatial distribution information in remote sensing image land cover classification methods, this paper proposes a random forest classification method based on superpixel statistics. The study area is Haidian District, Beijing, with Landsat-8 satellite data employed as the primary data source. The SLIC superpixel segmentation algorithm is improved to adapt it for superpixel segmentation in multispectral remote sensing images. Six common statistical measures of superpixels (minimum, maximum, mean, standard deviation, upper quartile, and lower quartile) are extracted for random forest classification of remote sensing images. Experimental results demonstrate that the proposed method achieves an overall classification accuracy of 89.01% for the study area, significantly reducing misclassification and omission errors of ground objects, and can be generalized to land cover classification of Landsat-8 remote sensing images.

Full Text

Preamble

Title: Random Forest Remote Sensing Image Classification Based on Superpixel Statistics

Authors: Shi Caixia[1], Zhao Chuangang[1], Pang Lei[2]
(1. School of Information Science & Technology, Beijing Forestry University, Beijing 100083, China; 2. School of Geomatics and Urban Information, Beijing University of Civil Engineering and Architecture, Beijing 100044, China)

Abstract: To address the insufficient utilization of spatial distribution information in remote sensing image land cover classification methods, this paper proposes a random forest remote sensing image classification method based on superpixel statistics. Taking Haidian District, Beijing as the study area and Landsat-8 satellite imagery as the primary data source, we adapt the SLIC superpixel segmentation method for multispectral remote sensing images and extract six common statistical measures (minimum, maximum, mean, standard deviation, upper quartile, and lower quartile) from superpixels for random forest classification. Experimental results demonstrate that the proposed method achieves an overall classification accuracy of 89.01% for the study area, significantly reducing commission and omission errors and demonstrating potential for broader application in Landsat-8 remote sensing image land cover classification.

Keywords: Landsat-8; random forest; superpixel; land cover; simple linear iterative cluster (SLIC)

Funding: This research was supported by the Architecture Foundation (61501019) and the General Project of the Beijing Municipal Education Commission' s Science and Technology Program (SQKM201610016008).

Author Biographies: Shi Caixia (1991–), female, from Lüliang, Shanxi, Master's student, research interests include computer applications, image processing, and machine learning (scx_fighting@163.com); Zhao Chuangang (1971–), male, associate professor, Ph.D., research interests include information retrieval and machine learning; Pang Lei (1971–), female, associate professor, Ph.D., research interests include multi-baseline InSAR data processing and tomographic SAR remote sensing.

0 Introduction

Remote sensing image classification represents a critical technology for remote sensing information extraction. Over the past several decades, researchers worldwide have dedicated substantial effort to improving classification accuracy. Remote sensing image classification methods generally fall into two categories: pixel-based classification and object-based classification. Pixel-based approaches treat individual pixels as the fundamental processing unit, neglecting the spatial relationships among neighboring pixels and thereby limiting potential improvements in classification accuracy. In contrast, object-based classification methods fully exploit both spectral and spatial distribution information, yielding superior performance compared to pixel-based methods.

Image segmentation constitutes a key step in object-based classification, with segmentation quality directly impacting classification accuracy. Superpixel segmentation has emerged as a rapidly developing approach that facilitates the extraction of local image features while effectively reducing computational complexity for subsequent processing. Superpixel methods can be broadly classified into two categories: graph-based approaches such as Normalized Cuts (NC), Graph-Based Segmentation (GS), and Superpixel Lattice (SL); and gradient descent-based methods including Watersheds (WS), Mean-Shift (MS), Quick

Shift (QS), Turbopixels (TP), and Simple Linear Iterative Clustering (SLIC). Among these, the SLIC algorithm offers advantages in segmentation speed, computational efficiency, edge adherence, compactness, and controllable superpixel quantity. While existing studies have applied SLIC to remote sensing image segmentation, most have focused on color imagery rather than multispectral data.

The Random Forest (RF) algorithm, an ensemble learning method based on classification and regression decision trees, has garnered widespread attention in remote sensing classification. Feature selection significantly influences RF classification accuracy. Previous research has integrated spectral, shape, and texture features for high-resolution imagery classification, incorporated Getis statistics to enhance RF performance, and combined Moment Distance Index (MDI) with spectral and topographic features for vegetation classification in the Pamir region. However, no universally applicable feature selection method exists for diverse remote sensing image types, leaving considerable room for improvement.

This paper adapts the SLIC superpixel segmentation algorithm for multispectral remote sensing images and extracts six statistical measures from superpixels as features for RF-based land cover classification, thereby effectively utilizing both spectral and spatial distribution information.

1.1 Random Forest

Random Forest is an ensemble classifier composed of decision trees $\{T(x, \Theta_k), k = 1, 2, ..., K\}$, where the final classification result is determined by majority voting among base classifiers. The basic RF workflow is illustrated in [Figure 1: see original paper].

The algorithm proceeds as follows:
a) Using Bootstrap sampling to randomly draw K training subsets from the original dataset.
b) Constructing decision tree models from each of the K training subsets, yielding K classification results.
c) Determining the final classification through majority voting among the K results.

In step (b), the input variables for each decision tree are randomly selected from M features, typically with $m = \sqrt{M}$ (where m represents the number of features). Node splitting criteria commonly employ the Gini index and Entropy, calculated as:

$$Gini(D) = 1 - \sum_{i=1}^{m} p_i^2$$

$$Entropy(D) = -\sum_{i=1}^{m} p_i \log(p_i)$$

where m denotes the number of classes in training dataset D, and p_i represents the probability of data belonging to class i.

1.2.1 SLIC Superpixel Segmentation Algorithm

The SLIC algorithm is a gradient descent-based superpixel segmentation method that clusters pixels according to spatial and color distance similarity, requiring only a single parameter K to control the number of superpixels generated.

The basic SLIC workflow comprises:
a) Initializing seed points. On an image with N pixels, seed points are uniformly distributed with step size $S = \sqrt{N/K}$. Within the 3×3 neighborhood of each original seed point, the pixel with minimum gradient value becomes the new seed point, and each seed receives a unique label.
b) Assigning labels to pixels. Within a $2S \times 2S$ search window centered at each seed point, distances between the seed and all pixels in the search range are computed. Each pixel receives the label of its nearest seed point.
c) Updating seed points. The centroid of each superpixel is calculated and becomes the new seed location.
d) Repeating steps (b) and (c) until convergence.
e) Removing multiply-connected and excessively small superpixels.

In step (b), distance similarity is measured in 5-dimensional space (CIELAB color space $[l, a, b]^T$ and coordinate space $[x, y]^T$), computed as:

$$d_c = \sqrt{(l_i - l_j)^2 + (a_i - a_j)^2 + (b_i - b_j)^2}$$

$$d_s = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$$

$$D = \sqrt{\left(\frac{d_c}{m}\right)^2 + \left(\frac{d_s}{S}\right)^2}$$

where d_c represents color difference between pixels i and j in color space; d_s denotes spatial distance; S is the seed spacing; m is a constant balancing color similarity and spatial proximity; and D measures overall similarity, with smaller values indicating greater similarity.

1.2.2 Improved SLIC Superpixel Segmentation Algorithm

The original SLIC algorithm processes color images, whereas our data comprises multispectral remote sensing imagery with seven bands. We extend the distance measurement dimension from 5 to 9 (color space: bands 1-7 $[b_1, b_2, b_3, b_4, b_5, b_6, b_7]^T$ and coordinate space $[x, y]^T$), yielding the modified distance formula:

$$d_{band} = \sqrt{\sum_{k=1}^{7} (b_{ik} - b_{jk})^2}$$

$$D = \sqrt{\left(\frac{d_{band}}{m}\right)^2 + \left(\frac{d_s}{S}\right)^2}$$

where d_{band} represents color difference between pixels i and j in the 7-band space, with b_{ik} and b_{jk} denoting pixel values at band k for pixels i and j, respectively.

1.2.3 Superpixel Statistics

Using the improved SLIC algorithm, we compute superpixels and extract their features as classification vectors for the RF model, thereby leveraging spatial distribution information. From a statistical perspective and based on the boxplot in [Figure 2: see original paper], a dataset' s distribution can be characterized by six statistical measures: minimum (min), maximum (max), mean (mean), upper quartile (upperQ), lower quartile (lowerQ), and standard deviation (stdDev). These six statistics collectively describe superpixel characteristics.

1.3.1 Algorithm Description

This paper integrates the improved SLIC superpixel segmentation with RF classification, using superpixel statistical features as input for remote sensing image land cover classification. The technical workflow is shown in [Figure 3: see original paper].

The main processing steps are:
a) Preprocess Landsat-8 OLI multispectral imagery, including radiometric calibration, atmospheric correction, mosaicking, and clipping to the study area.
b) Apply the improved SLIC algorithm to segment the image into superpixels, clustering similar pixels within local regions. Extract six statistical measures (min, max, mean, stdDev, upperQ, lowerQ) for each superpixel across each band as statistical features representing the superpixel region.
c) Combine multispectral bands, texture features, and superpixel statistics as

input variables to construct an RF classification model and generate the land cover map.
d) Validate classification accuracy.

In step (b), superpixel segmentation using the improved SLIC algorithm is illustrated in [Figure 4: see original paper]. For each superpixel block, we extract 42 statistical features across 7 multispectral bands ($7 \times 6 = 42$), as shown in [Figure 5: see original paper]. Texture features in step (c) are computed from the Gray-Level Co-occurrence Matrix (GLCM), including mean, variance, homogeneity, contrast, dissimilarity, entropy, angular second moment, and correlation—eight features calculated for each of Landsat-8' s seven multispectral bands.

1.3.2 Algorithm Complexity Analysis

This study employs SLIC superpixel segmentation and random forest classification for remote sensing image analysis. Assuming an image contains N pixels, the RF algorithm uses K decision trees with M classification features. The SLIC segmentation algorithm has time complexity $O(N)$, while the RF classification algorithm has approximate complexity $O(KM\sqrt{M}\log(N))$.

2.1.1 Study Area Overview and Classification System

This study focuses on Haidian District, Beijing (39°53 –40°09 N, 116°03 –116°23 E) ([Figure 6: see original paper]). The district features a temperate humid monsoon climate with cold winters and hot, rainy summers. Covering 430.77 km² with terrain sloping from high in the west to low in the east, the western mountainous area is forested while the eastern and southern plains contain mixed farmland and scattered residential areas. Ten rivers and several reservoirs (including Yuyuantan and Shangzhuang) create a typical urban land cover mosaic. Based on the national land use/land cover classification system and the characteristics of Landsat-8 data, we define five land cover classes: water, cultivated land, forest land, built-up area, and bare land.

2.1.2 Data Source

The remote sensing data were downloaded from the United States Geological Survey (USGS) website as L1T terrain-corrected Landsat-8 OLI multispectral imagery. One scene (path/row: 123/032, acquired October 3, 2013) covers the study area after preprocessing. High-resolution Google Earth imagery served as auxiliary data for training sample selection and accuracy validation.

2.2.1 Software

We used ENVI 5.3.1 for image preprocessing and IDL as the development language, combined with the EnMAP-Box tool for experimental validation. The experimental platform employed a 2.10 GHz processor, 6 GB RAM, and Windows 7 64-bit operating system.

2.2.2 Sample Selection

Training and validation samples were selected through visual interpretation of Landsat-8 false-color composite imagery ([Figure 7: see original paper]) with reference to high-resolution Google Earth imagery, ensuring uniform distribution across the study area. Stratified random sampling was employed, with details provided in (note: training and validation samples are independent).

2.2.3 SLIC Superpixel Segmentation and Feature Extraction

Segmentation scale critically affects statistical feature extraction when using superpixels. For a segmentation scale K in an image with N pixels, the expected superpixel size is N/K. Larger K values produce smaller superpixels, clustering similar land cover types more effectively, but may lead to over-segmentation and reduced computational efficiency beyond an optimal point. Conversely, smaller K values generate larger superpixels that may group different land cover types, causing under-segmentation. Therefore, selecting an appropriate scale requires balancing segmentation quality and speed.

Given the 30 m spatial resolution and 987 × 999 pixel image size, we tested K values of 10,000, 30,000, 50,000, 70,000, and 90,000. [Figure 8: see original paper] shows local segmentation results for these scales visualized using mean values from bands 5, 4, and 3. We extracted six statistical measures (min, max, mean, stdDev, upperQ, lowerQ) for each superpixel across seven bands (42 features), plus eight GLCM texture features (mean, variance, homogeneity, contrast, dissimilarity, entropy, angular second moment, correlation) computed in a 5×5 window for each band (56 features). These were combined with the seven multispectral bands to create a 105-dimensional feature vector for each pixel. Based on the experimental results in Section 3.2, we selected $K = 50,000$ as the optimal segmentation scale.

2.2.4 RF Classification Model Construction

The RF classification model parameters were configured as follows: 100 decision trees, feature subset size of 10 ($m = \sqrt{M}$, where $M = 105$ candidate features), and Gini index as the splitting criterion. The feature vector comprising 105 elements (7 multispectral bands, 42 superpixel statistics, and 56 texture features) served as input to the RF model.

3.1 Classification Results

[Figure 9: see original paper] compares RF classification results using only multispectral and texture features versus those supplemented with superpixel statistical features (at $K = 50,000$). The inclusion of superpixel statistics substantially improves visual classification quality.

3.2 Accuracy Verification

[Figure 10: see original paper] illustrates classification accuracy for each land cover type across the five segmentation scales. Forest land shows consistent accuracy across scales due to distinct superpixel statistical characteristics, while water, cultivated land, built-up area, and bare land exhibit greater sensitivity to scale. The $K = 50,000$ scale yields highest extraction accuracy for these classes. Average extraction accuracies across the five scales are 84.34% for water, 87.01% for cultivated land, 89.25% for forest land, 85.47% for built-up area, and 89.23% for bare land.

compares classification accuracy with and without superpixel statistical features (at $K = 50,000$). The proposed method improves overall accuracy by approximately 5.2% and Kappa coefficient by 7% compared to using only multispectral and texture features.

presents confusion matrices for both approaches. The superpixel-enhanced RF algorithm improves classification accuracy for water, cultivated land, forest land, built-up area, and bare land by 5.94%, 3.91%, 1.55%, 7.35%, and 8.28%, respectively. The baseline method shows lower accuracy for water, bare land, and built-up area, partly due to spectral confusion between built-up areas and bare land. Superpixel statistical features at $K = 50,000$ significantly reduce commission errors between these classes and improve discrimination of water, forest, and cultivated land, though forest land improvement is modest.

3.3 Variable Importance Analysis

Feature importance quantifies a variable’ s impact on classification accuracy. In an RF model with N decision trees, for variable F_i, the i-th tree computes OOB error OOB_1. After randomly permuting variable F_i in OOB data, the OOB error becomes OOB_2. Variable importance is calculated as:

$$Importance(F_i) = \frac{1}{N}\sum_{i=1}^{N}(OOB_2 - OOB_1)$$

Larger $Importance(F_i)$ values indicate greater variable significance.

[Figure 11: see original paper] ranks variable importance for the 42 superpixel statistics (at $K = 50,000$) and 56 GLCM texture features, showing the top 30. Among these, 24 are superpixel statistics while only 6 are GLCM texture features. The top three features are GLCM mean values, indicating their importance. However, other GLCM texture features rank much lower and show substantially less importance than superpixel statistics, demonstrating lower discriminative power. This confirms that superpixel statistical features significantly enhance classification accuracy.

---

4 Conclusion

This paper proposes an RF classification method for Landsat-8 multispectral imagery based on superpixel statistics. We extend and improve the SLIC algorithm to generate superpixels for multispectral data and extract six statistical measures as classification features, effectively leveraging spatial relationships among pixels to enhance feature discriminability. The method substantially improves overall classification accuracy and Kappa coefficient while reducing commission and omission errors. Variable importance analysis further confirms the critical role of superpixel statistical features in improving classification performance.

References

[1] Myint S, Gober P, Brazel A, et al. Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery [J]. Remote Sensing of Environment, 2011, 115 (5): 1145-1161.

[2] Yu Q, Gong P, Clinton N, et al. Object-based detailed vegetation classification with airborne high spatial resolution remote sensing imagery [J]. Photogrammetric Engineering and Remote Sensing, 2006, 72 (7): 799-811.

[3] Liu D, Xia F. Assessing object-based classification: advantages and limitations [J]. Remote Sensing Letters, 2010, 1 (4): 187-194.

[4] 宋熙煜, 周利莉, 李中国, 等. 图像分割中的超像素方法研究综述 [J]. 中国图象图形学报, 2015, 20 (5): 599-608.

[5] Ren Xiao-feng, Malik J. Learning a classification model for segmentation [C]// Procs of the 9th IEEE International Conference on Computer Vision. Washington DC: IEEE Computer Society, 2003: 10-17.

[6] Shi J, Malik J. Normalized cuts and image segmentation [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2000, 22 (8): 888-905.

[7] Felzenszwalb P, Huttenlocher D. Efficient graph-based image segmentation [J]. International Journal of Computer Vision, 2004, 59 (2): 167-181.

[8] Moore A, Prince S, Warrell J, et al. Superpixel lattices [C]// Proc of IEEE Conf. on Computer Vision and Pattern Recognition. 2008.

[9] Vincent L, Soille P. Watersheds in digital spaces: an efficient algorithm based on immersion simulations [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 1991, 13 (6): 583-598.

[10] Comaniciu D, Meer P. Mean shift: a robust approach toward feature space analysis [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2002, 24 (5): 603-619.

[11] Vedaldi A, Soatto S. Quick shift and kernel methods for mode seeking [C]. Proc. European Conf. Computer Vision, 2008.

[12] Levinshtein A, Stere A, Kutulakos K, et al. Turbopixels: fast superpixels using geometric flows [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2009, 31 (12): 2290-2297.

[13] Achanta R, Shaji A, Smith K, et al. SLIC superpixels compared to state-of-the-art superpixel methods [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2012, 34 (11): 2274–2282.

[14] Yi Yuan, Xiangyu Hu. Random forest and objected-based classification for forest pest extraction from UAV aerial imagery [C]// Proc of ISPRS International Archives of the Photogrammetry Remote Sensing and Spatial Information Sciences, 2016: 1093-1098.

[15] 董志鹏, 梅小明, 陈杰, 等. 简单线性迭代聚类的高分辨率遥感影像分割 [J]. 遥感信息, 2016, 31 (6): 83-88.

[16] Breiman L. Random forests [J]. Machine learning, 2001, 45 (1): 5-32.

[17] Lowe B, Kulkarni A. Multispectral image analysis using Random Forest [J]. International Journal on Soft Computing, 2015, 6 (1): 1-13.

[18] Rodriguez-Galiano V F, Ghimire B, Rogan J, et al. An assessment of the effectiveness of a random forest classifier for land-cover classification [J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2012, 67 (1): 93-104.

[19] 刘毅, 杜培军, 郑辉, 等. 基于随机森林的国产小卫星遥感影像分类研究 [J]. 测绘科学, 2012, 37 (4): 194-196.

[20] 刘海娟, 张婷, 侍昊, 等. 基于 RF 模型的高分辨率遥感影像分类评价 [J]. 南京林业大学学报: 自然科学版, 2015, 39 (1): 99-103.

[21] Ghimire B, Rogan J, Miller J. Contextual land-cover classification: incorporating spatial dependence in land-cover classification models using random forests and the Getis statistic [J]. Remote Sensing Letter, 2010, 1 (1): 45-54.

[22] Eric A L S, Kenneth G B, Raul V. Multispectral and texture feature application in image-object analysis of summer vegetation in eastern Tajikistan Pamirs [J]. Remote Sensing, 2016, 8 (78): 1-20.

[23] Haralick R M, Shanmugam K, Dinstein I. Textural features for image classification [J]. IEEE Trans on Man and Cybernetics, 1973, 3 (6): 610-621.

[24] 姚登举, 杨静, 詹晓娟. 基于随机森林的特征选择算法 [J]. 吉林大学学报: 工学版, 2014, 44 (1): 137-141.

[25] Van Der Linden S, Rabe A, Held M, et al. The EnMAP-Box—A toolbox and application programming interface for EnMAP data processing [J]. Remote Sensing, 2015, 7 (9): 11249-11266.

Note: Figure translations are in progress. See original paper for figures.