

Article

Saliency Preprocessing Locality-Constrained Linear Coding for Remote Sensing Scene Classification

Lipeng Ji ^{1,*}, Xiaohui Hu ² and Mingye Wang ¹

¹ School of Automation Science & Electrical Engineering, Beihang University, Beijing 100083, China; marcel0829@126.com

² Institute of Software Chinese Academy of Science, Beijing 100190, China; hxh@iscas.ac.cn

* Correspondence: jlp_1987@buaa.edu.cn; Tel.: +86-010-8231-7751

Received: 30 July 2018; Accepted: 26 August 2018; Published: 30 August 2018



Abstract: Locality-constrained Linear Coding (LLC) shows superior image classification performance due to its underlying properties of local smooth sparsity and good construction. It encodes the visual features in remote sensing images and realizes the process of modeling human visual perception of an image through a computer. However, it ignores the consideration of saliency preprocessing in the human visual system. Saliency detection preprocessing can effectively enhance a computer's perception of remote sensing images. To better implement the task of remote sensing image scene classification, this paper proposes a new approach by combining saliency detection preprocessing and LLC. This saliency detection preprocessing approach is realized using spatial pyramid Gaussian kernel density estimation. Experiments show that the proposed method achieved a better performance for remote sensing scene classification tasks.

Keywords: saliency preprocessing LLC; saliency detection; image processing; scene classification

1. Introduction

Over recent decades, an overwhelming amount of high-resolution (HR) remote sensing images have become available. Since remote sensing images have abundant structural patterns and spatial information that are difficult to be fully applied directly, we need to correctly classify them by senses before further processing. Therefore, remote sensing scene classification is a central issue in remote sensing applications [1,2].

Various methods have been proposed to classify remote sensing scenes over the years. Bag-of-Features (BoF) [3,4] is a classical method in whole-image categorization tasks. This method first forms a histogram based on a remote sensing image's local features, and then uses the histogram to represent the remote sensing image. However, this method lacks consideration of the spatial layout information of features in remote sensing images. There are some improvements based on the BoF method, such as those shown in References [5,6], and the Spatial Pyramid Matching (SPM) method [7], which is a successful method. The SPM method divides a remote sensing image into different scale spatial sub-regions. Then, histograms of local features from each sub-region are computed. Usually, $2^l \times 2^l$ sub-regions (where $l = 0, 1, 2$) are used. SPM has shown a better performance than BoF in most image classification tasks; however, the traditional SPM approach also has limitations. It requires nonlinear classifiers to complete classification. To improve SPM, a new coding algorithm, named Locality-constrained Linear Coding (LLC), was proposed [8]. The LLC method is widely applied in image classification tasks. It considers both the locality constraints and global sparsity when coding remote sensing image. LLC shows a state-of-the-art classification accuracy. In this paper, the proposed remote sensing scene classification approach is based on LLC. The proposed approach

will improve the codebook technology in traditional LLC by combining saliency detection technology, which is good for remote sensing scene classification.

LLC has achieved state-of-the-art performances in several image classification tasks; however, it is still based on modeling the cognitive mechanism of human vision to complete the image classification task. Physiological and psychophysical evidence indicates that the human's visual system has evolved a specialized focus processing treatment called "attentive" mode, which is directed to particular locations in the visual field [9]. Based on the "attentive" mode, researchers have proposed several visual saliency detection methods [10]. LLC does not take into account visual saliency detection. Thus, we can improve LLC by combining it with visual saliency detection.

Saliency detection was introduced into the field of computer vision in the late 1990s. According to the computational view, saliency detection can be grouped into different algorithms. One category is the center-surround thought. This type of algorithm assumes that a local window exists, which can divide an image into a center containing an object and a surround [10–13]. Another category of saliency detection algorithms is based on frequency domain methods [14–16]. The other category is relying on information theory concepts [17–19]. In Reference [13], Fast and Efficient Saliency Detection (FESD) was proposed, which belongs to the first saliency detection algorithm category. In the FESD method, a saliency map is built by computing the kernel density estimation, which has a faster computing performance compared with several other density estimations [20,21]; however, most remote sensing images, in which saliency maps are obtained using FESD, are not always the best. This is because remote sensing images usually have a wide field of view and the size of the kernel used for FESD, chosen by experience, is not necessarily optimal. However, we found that FESD can be improved using a simple technique. The technique is named spatial pyramid Gaussian kernel density estimation (SPGKDE), which is suitable for saliency detection of remote sensing images. In our research, SPGKDE is used for the preprocessing of remote sensing images. Then, by combining LLC encoding, the classification accuracy of remote sensing images can be improved.

The main contribution of this paper is to propose a new kind of remote sensing scene classification method by combining SPGKDE saliency detection preprocessing and LLC to improve remote sensing scene classification accuracy. This paper is structured as follows. Section 2 describes the remote sensing image scene classification proposed method in detail. In particular, this section outlines SPGKDE saliency detection preprocessing, and explains how and why remote sensing scene classification accuracy can be improved by SPGKDE preprocessing. Section 3 shows a comparison of the experimental results based on traditional LLC classification and the proposed method, followed by discussions. Section 4 concludes the paper.

2. Methodology

As mentioned above, this paper proposes a saliency preprocessing locality-constrained linear coding method for remote sensing scene classification. Figure 1 shows the realization of this method. The core technologies are SPGKDE and saliency preprocessing LLC. More specifically, descriptions of these two techniques are as follows.

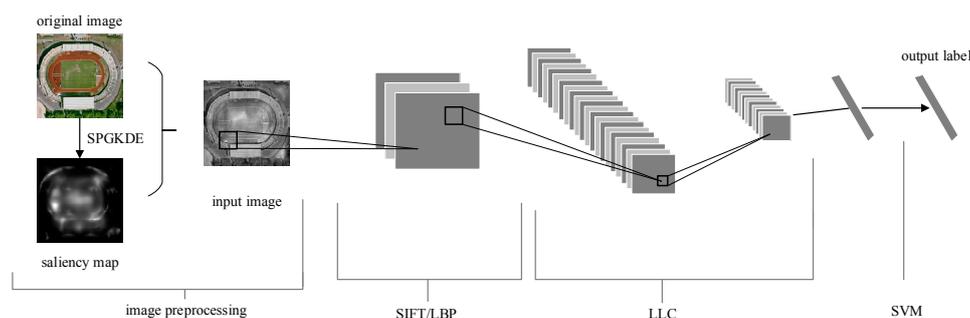


Figure 1. The flow chart of the proposed method.

2.1. Spatial Pyramid Gaussian Kernel Density Estimation Saliency Detection Preprocessing

FESD builds a saliency map of a remote sensing image using Gaussian kernel density estimation [13]. This method also implicitly considers sparse sampling and center bias. Since the human eye is more focused on the center of an image and is accustomed to taking a photo with the subject in the center, this method works well for most images. However, this method is not universal for remote sensing images. In some remote sensing images, the scene occupies the entire image. Center bias can lead to the loss of some salient information, which is detrimental to the classification of remote sensing images. We therefore need to improve FESD and propose SPGKDE. Figure 2 shows the weakness of the FESD method, which is the loss of salient information.

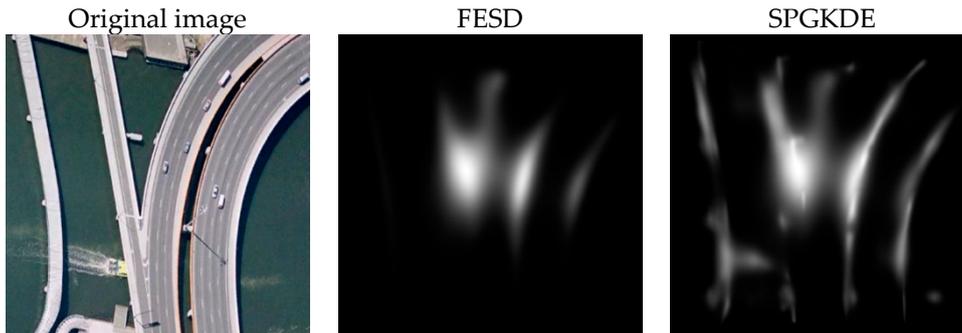


Figure 2. Saliency maps comparison between FESD (fast and efficient saliency detection method, which is proposed by reference [13]) and SPGKDE (spatial pyramid Gaussian kernel density estimation).

In this paper, a spatial pyramid Gaussian kernel density estimation (SPGKDE) saliency detection based on FESD is proposed. It requires minor changes but offers a great improvement on FESD. It is proposed for obtaining a remote sensing image’s saliency map more effectively, which is used for saliency preprocessing locality-constrained linear coding.

The spatial pyramid method (SPM [7]) is a simple and practical method in computer vision. This method divides a remote sensing image, from coarse to fine, by level. Then, local features in each level are aggregates later. Usually, $2^l \times 2^l$ sub-regions (where $l = 0, 1, 2$) are used. The proposed method, SPGKDE, is based on this thought. Figure 3 shows the realization of SPGKDE.

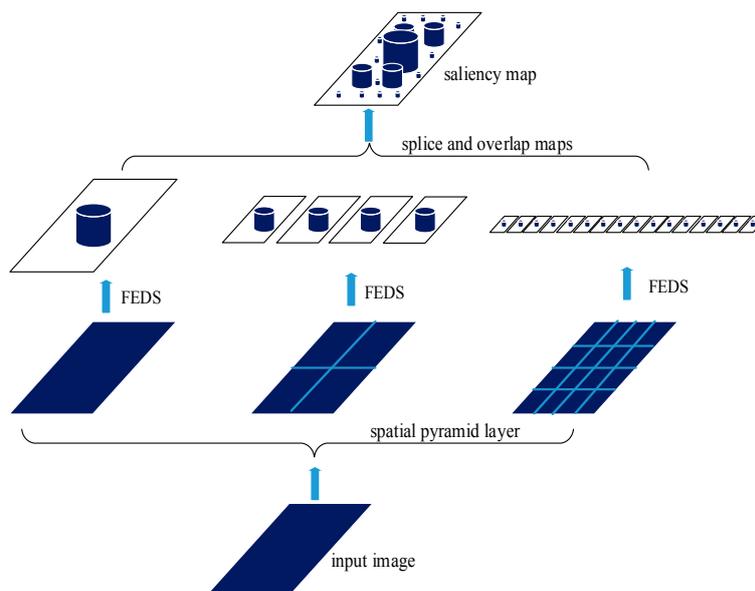


Figure 3. The flow chart of the SPGKDE.

Assume that there is a remote sensing image I and R_i^l is one of its sub-regions. For each sub-region image R_i^l , each pixel exists as $x = (\bar{x}, f)$, where f is a feature vector extracted from R_i^l , and \bar{x} denotes the coordinate of pixel x in R_i^l . For each sub-region R_i^l , we can get its corresponding saliency map S_i^l using the FESD method proposed in Reference [13].

$$S_i^l(x) = \mathfrak{A} * [P_r^m(1|f, \bar{x})]^\alpha \tag{1}$$

where $*$ is convolution operator, \mathfrak{A} is a circular averaging filter, $P_r^m(1|f, \bar{x})$ is a calculated probability for characterizing the pixels in the saliency areas, and $\alpha \geq 1$ is a factor that affects high probability areas.

Then, the saliency map S of remote sensing image I can be calculated as follows:

$$S = d_0 S_1^0 + d_1 \sum_{i=1}^4 S_i^1 + d_2 \sum_{i=1}^{16} S_i^2 = \sum_{l=0}^2 \sum_{i=1}^{4^l} d_l S_i^l \tag{2}$$

where $d_l = (l = 0, 1, 2)$ means a weight, and it can be defined as $d_l = \frac{1}{2^l}$, and S_i^l is acquired by Equation (1). When the level l is increasing, the weight d_l is decaying to prevent the collection of too much useless salient information.

Finally, a preprocessed image is obtained by:

$$I' = I + \zeta \cdot S \tag{3}$$

where I' is the preprocessed image, and $\zeta \in (0, 1]$ is used in order to avoid adding invalid details caused by the saliency map.

SPGKDE has a stronger nature of saliency detection than FESD because of its re-aggregation of salient information from different image space scales. After saliency detection preprocessing, the salient information can increase the inter-class variations between different remote sensing scenes. This is beneficial for improving the accuracy of classification tasks. In this paper, the Gaussian kernel is fixed as a 9×9 size, and the proportional control coefficient ζ is fixed at 0.5.

2.2. Saliency Preprocessing Locality-Constrained Linear Coding

Different kinds of coding algorithms have been proposed in the past few decades [5,8], most of which usually consist of feature extraction and feature coding. Experimental results have shown that, with a certain visual codebook, utilizing different coding schemes will directly affect the remote sensing scene classification accuracies [22,23]. Meanwhile, sparsity is less essential than locality under certain assumptions, as pointed out in Reference [24]. Therefore, Locality-constrained Linear Coding (LLC) was proposed [8].

LLC is widely applied in image classification tasks. It adds local restrictions to achieve global sparsity. Based on a given visual codebook, LLC provides analytical solutions. This coding method also has a fast coding speed. In this paper, LLC is selected as the basic method for remote sensing image feature coding.

Let F be a set of L -dimensional local descriptors extracted from the remote sensing image, i.e., $F = [f_1, f_2, \dots, f_P] \in R^{L \times P}$. Given a codebook $B = [b_1, b_2, \dots, b_Q] \in R^{L \times Q}$ with Q entries, LLC obeys the following criteria:

$$\min_C \sum_{i=1}^P \|f_i - Bc_i\|^2 + \lambda \|d_i \odot c_i\|^2, \quad s.t. \mathbf{1}^T c_i = 1, \forall i \tag{4}$$

where $d_i \in R^Q$ is the locality adaptor, and \odot is the element-wise multiplication. Usually, we have:

$$d_i = \exp\left(\frac{dist(f_i, B)}{\sigma}\right) \tag{5}$$

where $dist(f_i, B) = [dist(f_i, b_1), \dots, dist(f_i, b_q)]^T$, and $dist(f_i, b_i)$ represent the distance between f_i and b_i . σ is the weight used for adjusting the locality adaptor decay speed. Further, we can get the following solutions:

$$\tilde{c}_i = (C_i + \lambda diag^2(d_i))\mathbf{1} \tag{6}$$

$$c_i = \tilde{c}_i / \mathbf{1}^T \tilde{c}_i \tag{7}$$

$$C_i = (B^T - \mathbf{1}f_i^T)(B^T - \mathbf{1}f_i^T)^T \tag{8}$$

Unlike most coding methods, LLC can provide analytical solutions. It is of great benefit to computation. It also can be seen from Equation (8) that the quality of the given codebook B directly affects the coding results in LLC, and, further, indirectly affects the classification accuracy of remote sensing image scene categories. To better complete the task of remote sensing scene classification, we make full use of the advantages of the saliency detection information. Thus, we propose saliency preprocessing LLC.

For each remote sensing image I , we can obtain its preprocessed image I' by the method proposed in Section 2.1. We use image I and I' to get the corresponding codebook B and B' . Then, the given codebook B in Equation (8) can be replaced by B_{final} , which uses the following formula:

$$B_{final} = \frac{1}{2}(B + B') \tag{9}$$

where codebook B represents the original remote sensing images features codebook, and B' comes from the computations of corresponding saliency preprocessed remote sensing images. Figure 4 shows the differences between the traditional LLC codebook and the proposed method codebook. Obviously, the features used to generate the new codebook B_{final} are more prominent than those of codebook B constructed from traditional LLC features, and this new codebook B_{final} will improve the performance of remote sensing scene classification.

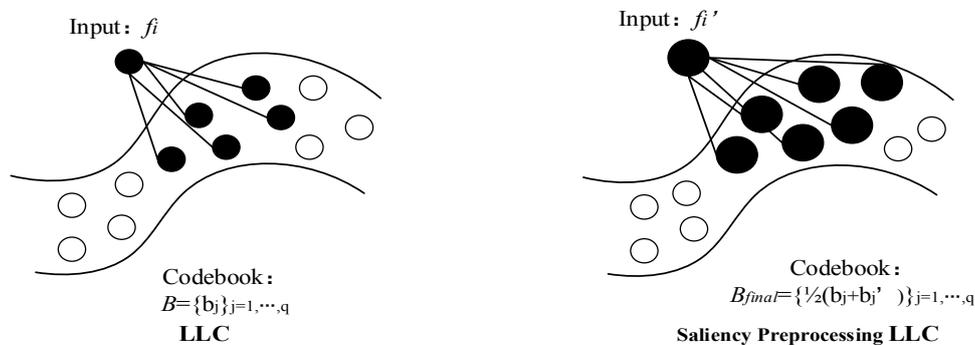


Figure 4. Comparison between LLC (Locality-constrained Linear Coding) and proposed method.

Saliency preprocessing LLC only solves the problem of feature coding. In this paper, Scale Invariant Feature Transform (SIFT) [25] and Local Binary Pattern (LBP) [26] are used to extract image features and a support vector machine (SVM) [27,28] is used as the training classifier. Of course, to prove the effectiveness of the proposed method, this paper uses a public 19-class remote sensing scene dataset to conduct experiments [29,30]. The dataset is proposed by Dengxin Dai and Wen Yang, and named as WHU-RS dataset. The images in this dataset are a fixed size of 600×600 pixels. All the images are collected from Google Earth. In the early days, the dataset contained 12 categories of physical scenes in the satellite imagery [29]; later, it was expanded to 19 classes [30]. The 19 classes of the dataset are airport, beach, bridge, river, forest, farmland, meadow, mountain, pond, parking, port, park, viaduct, desert, football field, railway station, residential area, industrial area and commercial area.

The dataset has higher intra-class variations and smaller inter-class dissimilarities. The experimental results and discussions are given in the next section.

3. Experiments and Discussion

In this section, we will focus on experiments based on a WHU-RS dataset, demonstrate the advantage of SPGKDE preprocessing for remote sensing images, and then show the performance of both the traditional LLC and the proposed method. Thereafter, the results will be briefly discussed.

3.1. SPGKDE Preprocessing

Figure 5 shows several sample images of the WHU-RS dataset and the corresponding saliency detection results computed by SPGKDE (proposed in this paper).

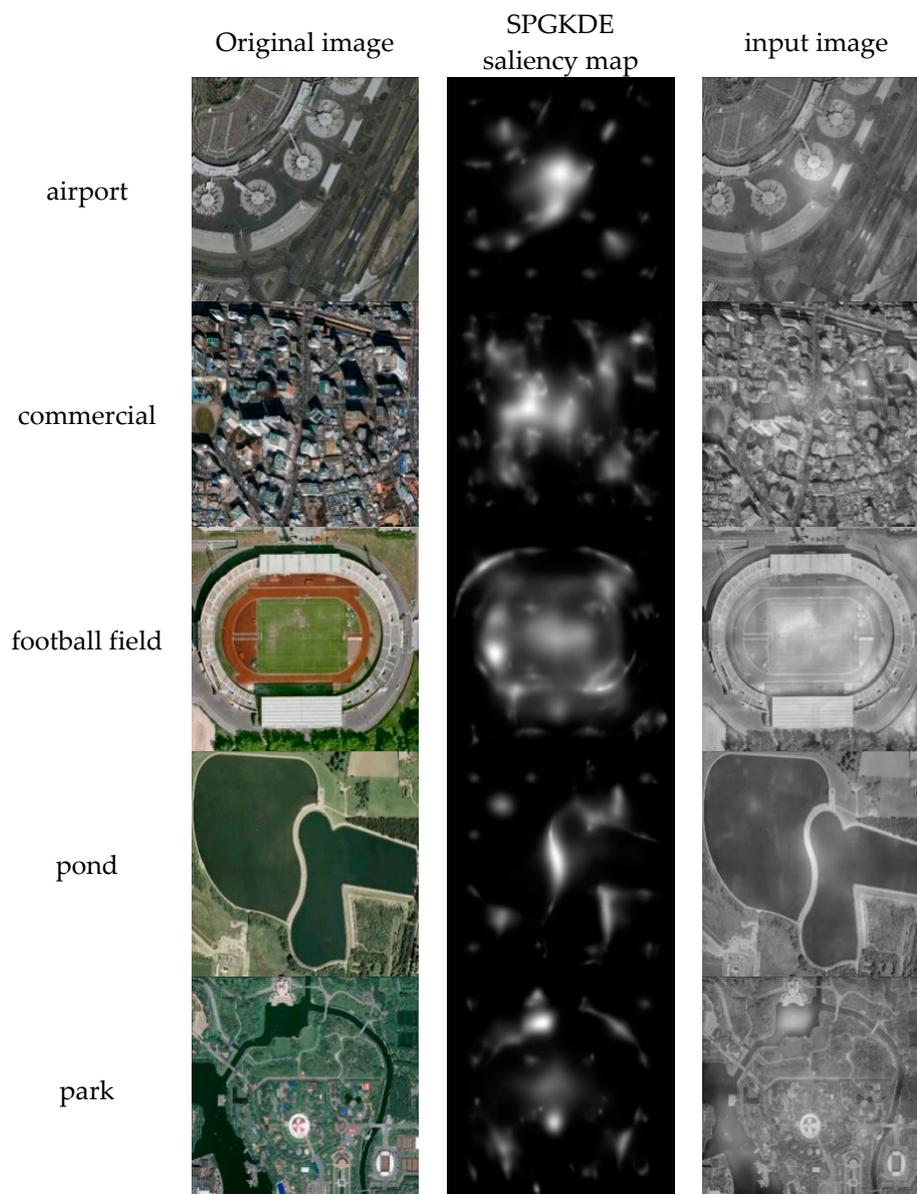


Figure 5. Cont.

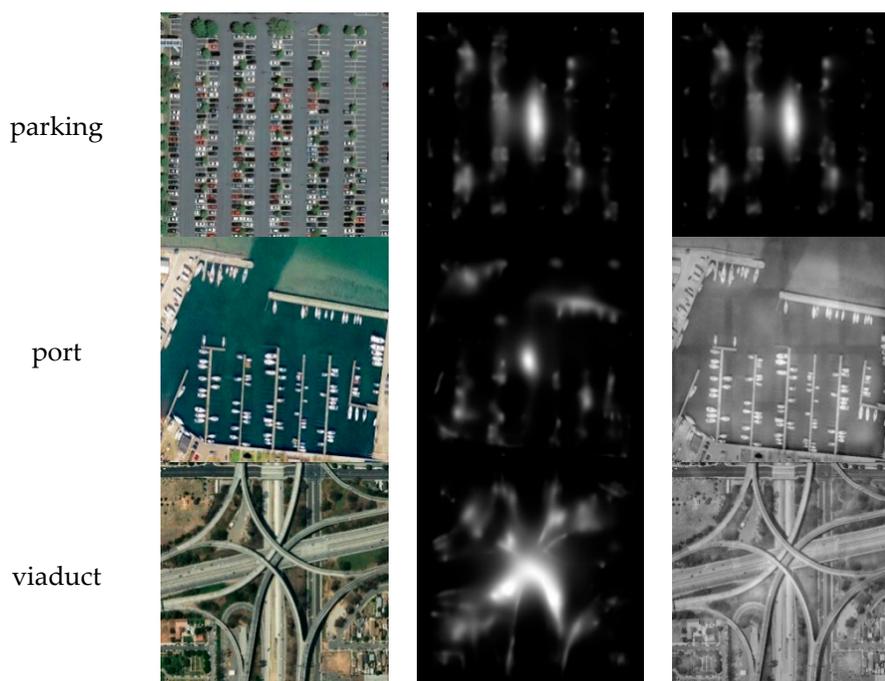


Figure 5. Samples of a 19-class dataset, SPGKDE saliency detection, and the final preprocessed images for LLC input.

As shown in Figure 5, most of the images can obtain useful and expected saliency maps via the proposed SPGKDE saliency detection, and after detection preprocessing, key parts of the images can be made more prominent. The preprocessing technique can increase inter-class variations and reduce inter-class dissimilarities. Though, inevitably, there are some unsatisfactory saliency maps for images that may lead to classification confusion; for instance, the pond scene image in Figure 5, where the features of the pond are vague, whereas the bridge becomes rather prominent after saliency detection preprocessing and would be wrongly classified into the park scene class, though this phenomenon is very rare. In fact, a very high percentage of images can correctly receive saliency detection preprocessing.

3.2. Performance of Traditional LLC and Proposed Method

In this paper, two kinds of low-level feature are used to form a codebook for LLC separately; namely, Scale Invariant Feature Transform (SIFT) and Local Binary Pattern (LBP).

3.2.1. Performance Based on SIFT Feature

The SIFT vector has a dimension of 128. Half of the dataset is used as the training set and the other half is used for the test. The accuracy of classification of traditional LLC and the proposed method based on SIFT is shown in Table 1, where the BoF performance is also shown as a baseline method.

Table 1. Classification accuracies based on Scale Invariant Feature Transform (SIFT).

Descriptors	Methods			
	Traditional BoF	Traditional LLC	Proposed Method	
SIFT	72.87%	73.27%	79.01%	

Figure 6 shows the confusion matrices generated by the traditional LLC and the proposed method based on SIFT. It is benefit for observing more differences in detail between the two methods. We can observe that more scene image categories, such as airport, bridge and commercial area were correctly

classified by the proposed method. Although the classification accuracies of meadow, parking and residential area categories were reduced, especially the classification accuracy of meadow, which dropped sharply by 12%, the method improved the classification accuracy of the entire public dataset by about 6%.

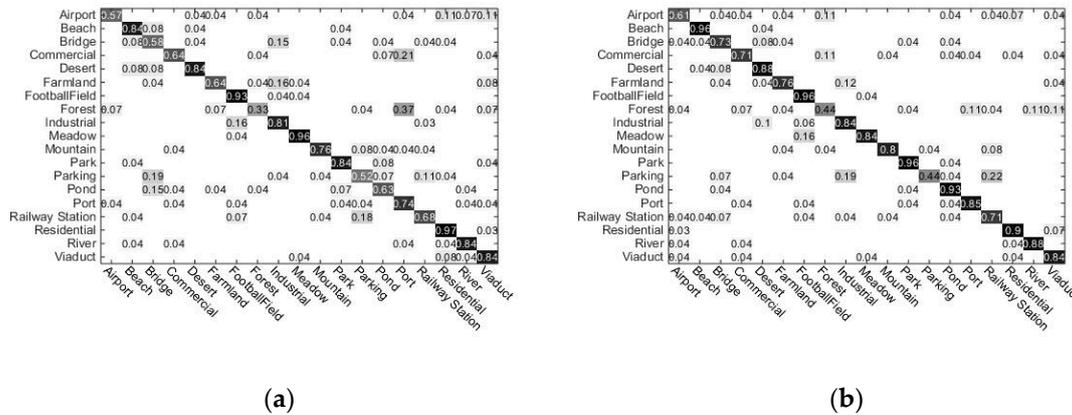


Figure 6. Confusion matrices based on SIFT; (a) Traditional LLC (SIFT); (b) Proposed method (SIFT).

To investigate the impact of saliency detection preprocessing, different proportions of training and test are used for the experiments. The training proportion is ranged from 20% to 80%. The experimental results are shown in Figure 7. The classification accuracies are clearly improved with our proposed method.

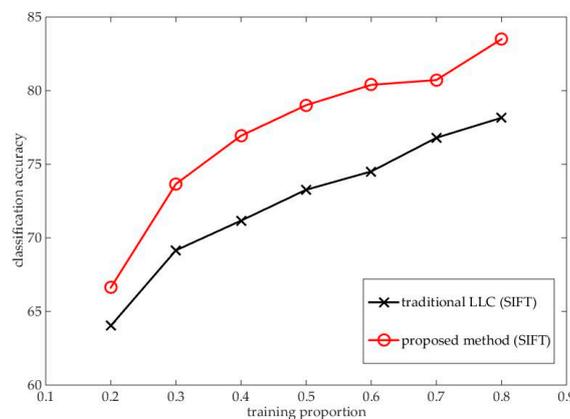


Figure 7. Different ratios of training set experiments based on SIFT.

3.2.2. Performance Based on LBP Feature

LBP vector is a 256-dimension feature vector. Half of the dataset is used as the training set and the other half is used for the test. The accuracy of classification of the traditional LLC and the proposed method based on LBP is shown in Table 2, where the BoF performance is shown as a baseline method as well.

Table 2. Classification accuracies based on Local Binary Pattern (LBP).

Descriptors	Methods		
	Traditional BoF	Traditional LLC	Proposed Method
LBP	68.71%	72.87%	78.22%

To observe more classification details of the proposed method, we gave the confusion matrices of traditional LLC and proposed method by Figure 8. The confusion matrices are generated by the traditional LLC and the proposed method based on LBP. We can see that the classification accuracies of the four categories is reduced, namely football field, industrial, meadow and mountain, accounting for 21% of all categories, whereas the classification accuracy of the entire public dataset is improved by approximately 5.4%.

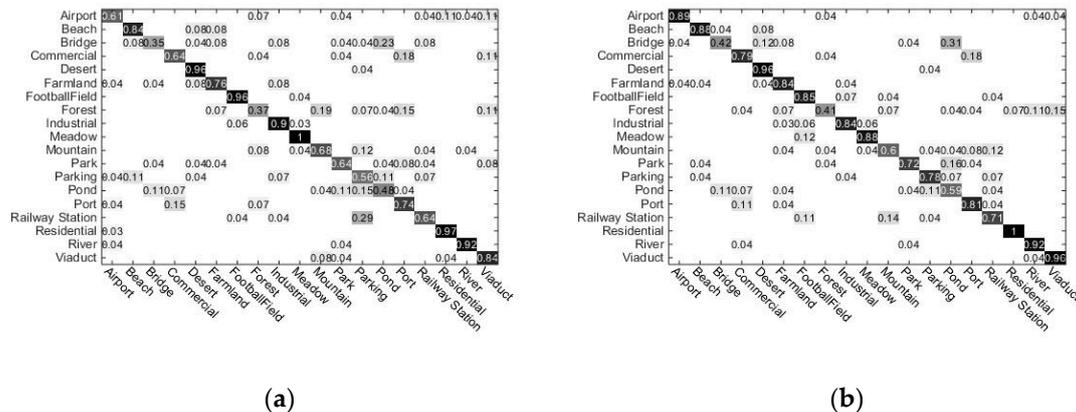


Figure 8. Confusion matrices based on LBP. (a) Traditional LLC (LBP); (b) Proposed method (LBP).

To investigate the impact of saliency detection preprocessing, different proportions of training and test are used for the experiments. The training proportion ranged from 20% to 80%. The experimental results are shown in Figure 9. As can be seen, the classification accuracies are improved with our proposed method.

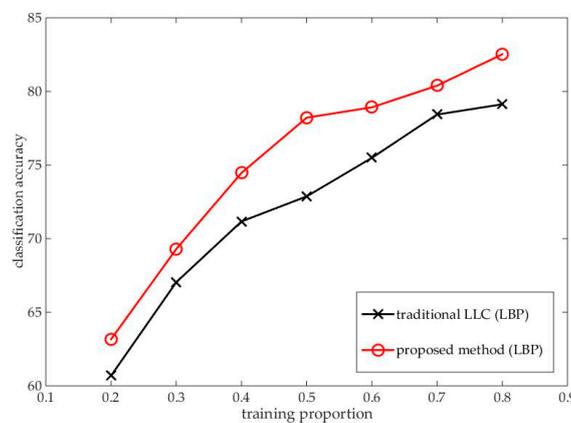


Figure 9. Different ratios of training set experiments based on LBP.

Through the above experiments, we can see that the proposed method can improve the scene classification accuracies of remote sensing images, regardless of the features extracted by the SIFT or LBP methods. However, the classification accuracies of the meadow category are reduced using the proposed method. To determine why the classification of the meadow scene images is more difficult to identify using the proposed method, we reviewed the entire process of processing meadow category images. We found that almost all the meadow-category images in the WHU-RS dataset did not have an obvious saliency region, as shown in Figure 10. The most striking feature of the meadow category images was the color. When we used SPGKDE to preprocess meadow category images, not only did we get little color information, but there was also additional untrue saliency region information. Furthermore, weakening the color information and forcibly mining saliency region

information, can lead to generating the wrong codebook for LLC, which is detrimental for classification based on LLC. This may explain why the proposed method reduced the classification accuracies of the meadow category. Fortunately, there are saliency regions in most remote sensing images and this happens only for meadow category images.

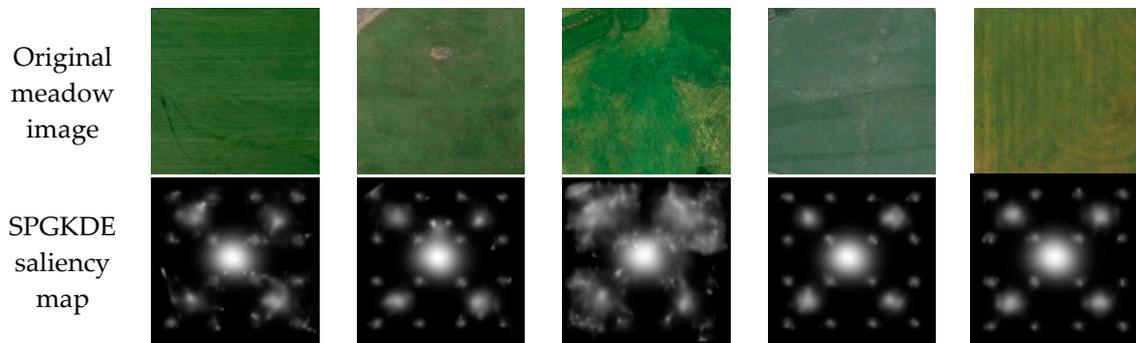


Figure 10. Meadow category images and corresponding saliency maps.

From the above experimental results, we have good reason to believe that the proposed method can improve remote sensing scene classification performance.

4. Conclusions

In this paper, an improvement saliency detection method named SPGKDE is proposed based on the existing saliency detection FESD for remote sensing image preprocessing. The new saliency detection method re-aggregates salient information from different image space scales and is obviously more applicable to remote sensing images. SPGKDE has a wider field of view than FESD.

Thus, a new kind of remote sensing scene classification method that combines SPGKDE saliency detection preprocessing and LLC is proposed. The method is easy to operate. Visual saliency detection plays an important role in remote sensing image analysis and the traditional LLC classification technology ignores this technology, resulting in a limited accuracy for image classification. This paper proposes integrating an improved saliency detection method—SPGKDE— into the LLC classification. This method can increase inter-class variations and reduce inter-class dissimilarities. In fact, the preprocessing method improves the codebook technology in traditional LLC. It is a core technology that directly determines the classification result. This method achieves a better simulation of the human visual system than traditional LLC. In this paper, both SIFT and LBP features were used for experiments. The experiments show that the proposed method is useful and can improve remote sensing scene classification accuracy.

Author Contributions: Conceptualization, L.J.; Data curation, L.J. and M.W.; Formal analysis, L.J.; Funding acquisition, X.H.; Investigation, M.W.; Methodology, L.J. and X.H.; Software, L.J. and M.W.; Writing—original draft, L.J.

Funding: This research was funded by the National Natural Science Foundation of China (U1435220).

Acknowledgments: The authors acknowledge the National Natural Science Foundation of China (U1435220).

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

References

1. Cheng, G.; Han, J.; Guo, L.; Liu, Z.; Bu, S.; Ren, J. Effective and efficient midlevel visual elements-oriented land-use classification using VHR remote sensing images. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 4238–4249. [[CrossRef](#)]
2. Anwer, R.M.; Khan, F.S.; van de Weijer, J.; Molinier, M.; Laaksonen, J. Binary patterns encoded convolutional neural networks for texture recognition and remote sensing scene classification. *ISPRS J. Photogramm. Remote Sens.* **2018**, *138*, 74–85. [[CrossRef](#)]
3. Sivic, J.; Zisserman, A. Video google: A text retrieval approach to object matching in videos. In Proceedings of the IEEE International Conference on Computer Vision, Nice, France, 13–16 October 2003; Volume 2, pp. 1470–1477.
4. Csurka, G.; Dance, C.R.; Fan, L.; Bray, C.; Csurka, G. Visual categorization with bags of keypoints. *Workshop Statist. Learn. Comput. Vis. Eccv.* **2004**, *44*, 1–22.
5. Bosch, A.; Zisserman, A.; Muñoz, X. Scene classification using a hybrid generative/discriminative approach. *IEEE Trans. Pattern Anal. Mach. Intell.* **2008**, *30*, 712–727. [[CrossRef](#)] [[PubMed](#)]
6. Yang, L.; Jin, R.; Sukthankar, R.; Jurie, F. Unifying discriminative visual codebook generation with classifier training for object category recognition. In Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, AK, USA, 23–28 June 2008.
7. Lazebnik, S.; Schmid, C.; Ponce, J. Beyond bags of features: spatial pyramid matching for recognizing natural scene categories. In Proceedings of the Conference on Computer Vision and Pattern Recognition, New York, NY, USA, 17–22 June 2006; pp. 2169–2178.
8. Wang, J.; Yang, J.; Yu, K.; Lv, F.; Huang, T.; Gong, Y. Locality-constrained linear coding for image classification. In Proceedings of the Computer Vision and Pattern Recognition, San Francisco, CA, USA, 13–18 June 2010; pp. 3360–3367.
9. Koch, C.; Ullman, S. Shifts in selective visual attention: Towards the underlying neural circuitry. *Human Neurobiol.* **1985**, *4*, 219–227.
10. Achanta, R.; Estrada, F.; Wils, P.; Süsstrunk, S. Salient region detection and segmentation. In *Computer Vision Systems*; Springer: Berlin, Germany, 2008; pp. 66–75.
11. Seo, H.J.; Milanfar, P. Training-free, generic object detection using locally adaptive regression kernels. *IEEE Trans. Pattern Anal. Mach. Intell.* **2009**, *32*, 1688–1704.
12. Rahtu, E.; Heikkilä, J. A Simple and efficient saliency detector for background subtraction. In Proceedings of the IEEE International Conference on Computer Vision Workshops, Kyoto, Japan, 27 September–4 October 2009; pp. 1137–1144.
13. Tavakoli, H.R.; Rahtu, E. Fast and efficient saliency detection using sparse sampling and kernel density estimation. In *Image Analysis*; Springer-Verlag: Berlin, Germany, 2011; pp. 666–675.
14. Guo, C.; Ma, Q.; Zhang, L. Spatio-temporal saliency detection using phase spectrum of quaternion fourier transform. In Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, AK, USA, 23–28 June 2008.
15. Hou, X.; Zhang, L. Saliency detection: A spectral residual approach. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Minneapolis, MN, USA, 17–22 June 2007.
16. Achanta, R.; Hemami, S.; Estrada, F.; Süsstrunk, S. Frequency-tuned salient region detection. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 1597–1604.
17. Bruce, N.D.B.; Tsotsos, J.K. Saliency based on information maximization. In Proceedings of the NIPS'05 Proceedings of the 18th International Conference on Neural Information Processing Systems, Cambridge, MA, USA, 5–8 December 2005; pp. 155–162.
18. Lin, Y.; Fang, B.; Tang, Y. A computational model for saliency maps by using local entropy. In Proceedings of the National Conference on Artificial Intelligence, Atlanta, GA, USA, 11–15 July 2010; Volume 2, pp. 967–973.
19. Mahadevan, V.; Vasconcelos, N. Spatiotemporal saliency in dynamic scenes. *IEEE Trans. Pattern Anal. Mach. Intell.* **2009**, *32*, 171–177. [[CrossRef](#)] [[PubMed](#)]
20. Aiazzi, B.; Alparone, L.; Baronti, S.; Garzelli, A.; Zoppetti, C. Nonparametric change detection in multitemporal SAR images based on mean-shift clustering. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 2022–2031. [[CrossRef](#)]

21. Liu, J.; Tang, Z.; Cui, Y.; Wu, G. Local competition-based superpixel segmentation algorithm in remote sensing. *Sensors* **2017**, *17*, 1364. [[CrossRef](#)] [[PubMed](#)]
22. Chen, J.; Li, Q.; Peng, Q.; Wong, K.H. CSIFT based locality-constrained linear coding for image classification. *Pattern Anal. Appl.* **2015**, *18*, 441–450. [[CrossRef](#)]
23. Yu, K.; Zhang, T.; Gong, Y. Nonlinear learning using local coordinate coding. In Proceedings of the Advances in Neural Information Processing Systems 22, Vancouver, BC, Canada, 7–10 December 2009; pp. 2223–2231.
24. Wang, S.; Ding, Z.; Fu, Y. Marginalized denoising dictionary learning with locality constraint. *IEEE Trans. Image Process.* **2018**, *27*, 500–510. [[CrossRef](#)] [[PubMed](#)]
25. Lowe, D.G.; Lowe, D.G. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **2004**, *60*, 91–110. [[CrossRef](#)]
26. Ojala, T.; Pietikainen, M.; Maenpaa, T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *Pattern Anal. Mach. Intell. IEEE Trans.* **2002**, *24*, 971–987. [[CrossRef](#)]
27. Boser, B.E.; Guyon, I.M.; Vapnik, V.N. Training algorithm for optimal margin classifiers. In Proceedings of the Fifth Annual ACM Workshop on Computational Learning Theory, New York, NY, USA, 27–29 July 1992; pp. 144–152.
28. Fan, R.E.; Chang, K.W.; Hsieh, C.J.; Wang, X.-R.; Lin, C.-J. LIBLINEAR: A library for large linear classification. *J. Mach. Learn. Res.* **2008**, *9*, 1871–1874.
29. Dai, D.; Yang, W. Satellite image classification via two-layer sparse coding with biased image representation. *IEEE Geosci. Remote Sens. Lett.* **2011**, *8*, 173–176. [[CrossRef](#)]
30. Sheng, G.; Yang, W.; Xu, T.; Sun, H. High-resolution satellite scene classification using a sparse coding based multiple feature combination. *Int. J. Remote Sens.* **2012**, *33*, 2395–2412. [[CrossRef](#)]



© 2018 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/>).