

LETTER

Extreme Learning Machine with Superpixel-Guided Composite Kernels for SAR Image Classification

Dongdong GUAN^{†a)}, Xiaolan TANG[†], Li WANG[†], *Nonmembers*, and Junda ZHANG[†], *Student Member*

SUMMARY Synthetic aperture radar (SAR) image classification is a popular yet challenging research topic in the field of SAR image interpretation. This paper presents a new classification method based on extreme learning machine (ELM) and the superpixel-guided composite kernels (SGCK). By introducing the generalized likelihood ratio (GLR) similarity, a modified simple linear iterative clustering (SLIC) algorithm is firstly developed to generate superpixel for SAR image. Instead of using a fixed-size region, the shape-adaptive superpixel is used to exploit the spatial information, which is effective to classify the pixels in the detailed and near-edge regions. Following the framework of composite kernels, the SGCK is constructed base on the spatial information and backscatter intensity information. Finally, the SGCK is incorporated an ELM classifier. Experimental results on both simulated SAR image and real SAR image demonstrate that the proposed framework is superior to some traditional classification methods.

key words: *extreme learning machine (ELM), superpixel, composite kernels (CK), SAR image classification*

1. Introduction

SAR image classification plays a very important role in different applications, such as land use analysis, pollution monitoring, and military applications [1]. In the last decades, many methods have been developed for SAR image classification [2]. Among these methods, the support vector machine (SVM) has played a dominant role in the field of SAR image classification. Although the SVM is superior in terms of accuracy, and robustness to noise, its computational cost is usually high. Recently, a fast and effective machine learning method called ELM has been proposed, which is competitive with SVMs in terms of accuracy and has very high computational efficiency. In addition, ELM can directly perform multiclass classification using multi-output nodes, while SVMs need to employ one-against-all (OAA) or one-against-one (OAO) method to implement multiclass classification.

Due to these remarkable advantages, ELM has been widely used in the fields of hyperspectral images (HSIs) processing [3]. Unfortunately, the use of the ELM for SAR image classification has not yet been extensively explored. In this work, we study and evaluate the application of ELM to SAR image classification. Because SAR image is heavily contaminated by speckle noise, directly applying the ELM

cannot always obtain satisfactory classification results. Inspired by the success of CK framework [4], [5], we consider introducing local spatial information under the framework of composite kernels to reduce the effect of speckle noise. However, the traditional CK methods extract the spatial information in a fixed-size square window, which may damage the structure and edge details. To solve the problem, we introduce the adaptive superpixel-guided window to extract spatial information. Since superpixels adhere well to edges, the extracted spatial information can capture the spatial structure, and thus the corresponding classification result can preserve most structure details. In addition, the proposed method (hereafter referred to as ELM-SGCK for simplicity) inherits the high computational efficiency of ELM.

The main contributions of this article are summarized as followings: 1) a new superpixel generation algorithm based on the GLR similarity is proposed for SAR image; 2) the SGCK is designed to jointly utilize the backscatter intensity information and the adaptive spatial information for SAR image classification; 3) both the simulated and real SAR image are used to evaluate the proposed ELM-SGCK. The experimental results demonstrate the superiority of ELM-SGCK over several traditional CK classification approaches, in terms of classification accuracy, visual quality and computational speed.

2. Proposed Method

2.1 Extreme Learning Machine

This subsection briefly reviews the ELM classifiers. ELM is a generalized single hidden layer feedforward neural network (SLFN), where its hidden node parameters, namely both weight connecting vectors and bias are randomly assigned and thus the output weights can be analytically computed with a least squares solution [3]. Similar to SVM, the inner product operation involved in the computation can be replaced by a kernel matrix. Thus, we can obtain a kernel-based ELM (KELM). Given N training samples $\{\mathbf{x}_i, \mathbf{y}_i\}_{i=1}^N$, where $\mathbf{x}_i \in R^d$ is the feature vector and $\mathbf{y}_i = [y_{i1}, \dots, y_{iM}]^T \in R^M$ is the label vector. For M classes, if $y_{ik} = 1$ and the other elements in \mathbf{y}_i are -1 , then the sample \mathbf{y}_i belongs to the k th class. According to [5], the output function of the KELM can be expressed as

$$f(\mathbf{x}) = \mathbf{K}_x \left(\frac{\mathbf{I}}{C} + \mathbf{K} \right)^{-1} \mathbf{Y} \quad (1)$$

Manuscript received December 28, 2017.

Manuscript revised March 1, 2018.

Manuscript publicized March 14, 2018.

[†]The authors are with College of Electronic Science, National University of Defense Technology, Changsha, 410073, China.

a) E-mail: guandd@hotmail.com (Corresponding author)

DOI: 10.1587/transinf.2017EDL8281

where $\mathbf{K} = [\mathbf{K}(x_i, x_j)]_{i,j=1}^N$ is the kernel matrix between training samples, $\mathbf{K}_x = [\mathbf{K}(x, x_1), \dots, \mathbf{K}(x, x_N)]$ is the kernel matrix between the test sample and training samples, $\mathbf{Y} = [y_1, \dots, y_N]^T \in R^{N \times m}$ is the label matrix, $\mathbf{I} \in R^{N \times N}$ is the identity matrix, and C is a regularization parameter introduced for better stability and generalization. During the prediction phase, the label of each test sample x_i is finally determined according to the index of the output node with the largest value in $f(x_i)$ calculated by (1). To calculate (1), we need to define how to compute the kernel matrix. In this paper, the proposed SGCK, as described in the following sections, is used to compute the kernel matrix.

2.2 Generation of Superpixels in SAR Image

SLIC is a widely used superpixel generation algorithm in the field of computer vision [6]. However, the original SLIC used Euclidean distance to measure the pixel intensity similarity, which is not robust for speckle noise. In recent, some studies have shown that the GLR similarity is very powerful to measure the pixel intensity similarity for SAR image [7]. Therefore, in this paper, we proposed to use more reasonable GLR criterion instead of Euclidean distance to generate superpixels for SAR image. Assuming that multilook intensity SAR image obeys the gamma distribution, based on [8], the GLR similarity between two pixels in SAR image is defined as

$$S_{GLR}(i, j) = \log \left(\sqrt{\frac{I_i}{I_j}} + \sqrt{\frac{I_j}{I_i}} \right) \quad (2)$$

where I_i and I_j are intensity values of the two pixels. Similar to SLIC, the final similarity is defined as

$$S(i, j) = S_{GLR}(i, j) + m \cdot S_{XY}(i, j) \quad (3)$$

where $S_{XY}(i, j)$ is the pixel location similarity as described in [6], m is a parameter that allows us to weigh the relative importance between pixel intensity and location similarities. Once the new similarity is defined, we can use it to generate superpixels for SAR image by performing the same procedures as SLIC. The superpixel generating results obtained by the proposed modified SLIC are shown in Fig. 1. It can be seen that superpixels adhere well to image boundaries.

2.3 Constitution of the SGCK

After the superpixel is generated, the spatial information

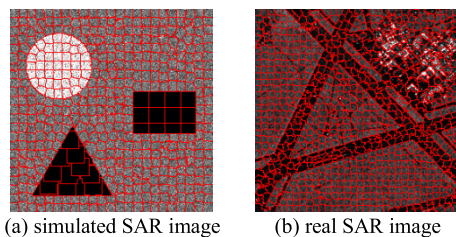


Fig. 1 Superpixel generating results by the proposed modified SLIC.

within the superpixel can be extracted. Here, the simple mean filtering is used to explore the spatial information within each superpixel. Then, the mean value is assigned to all pixels in each superpixel, and all the filtered superpixels can constitute a mean feature image I^{Mean} . Given a pixel x_i , its spatial feature, namely the pixel value in the I^{Mean} , is denoted as x_i^s and its backscatter intensity is denoted as x_i^b . By means of the Gaussian radial basis function (RBF), two kernels, the backscatter intensity kernel K_b and the spatial feature kernel K_s , are defined as follows:

$$K_b(x_i, x_j) = \exp \left(-\frac{\|x_i^b - x_j^b\|^2}{2\sigma_b^2} \right) \quad (4)$$

$$K_s(x_i, x_j) = \exp \left(-\frac{\|x_i^s - x_j^s\|^2}{2\sigma_s^2} \right) \quad (5)$$

where σ_b and σ_s are the bandwidth of the RBF kernels. Then, the composite kernels can be constructed

$$K_c(x_i, x_j) = (1 - \mu)K_b(x_i, x_j) + \mu K_s(x_i, x_j) \quad (6)$$

where μ is the weight, which controls the relative proportion of the backscatter intensity and spatial information in the composite kernels. Finally, the composite kernels are incorporated into (1) to solve the KELM model and implement classification.

3. Experiments

In this section, both simulated SAR and real SAR images are used to test the performance of the proposed ELM-SGCK method. The classification results of the proposed method are visually and quantitatively compared with some widely used classification methods, such as KELM, ELM with CK (ELM-CK). In addition, SVM with the proposed SGCK (SVM-SGCK) is also included for comparison. For ELM-CK, the spatial feature is extracted within a fixed-size

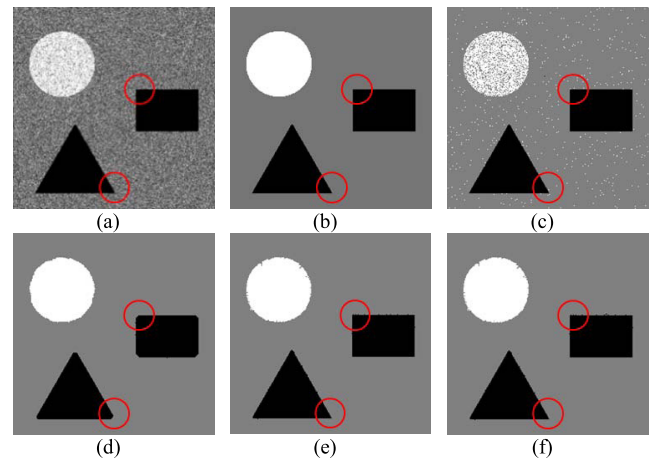


Fig. 2 Classification results for the simulated SAR image. (a) The simulated SAR image. (b) Ground truth map and (c)-(f) classification maps generated by different methods: KELM, ELM-CK, SVM-SGCK, and the proposed ELM-SGCK.

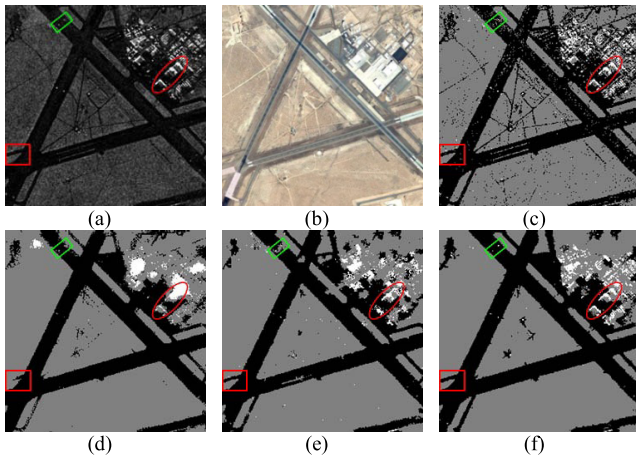


Fig. 3 Classification results for the real SAR image of China Lake Airport. (a) The original SAR image. (b) Google earth image and (c)-(f) classification maps generated by different methods: KELM, ELM-CK, SVM-SGCK, and the proposed ELM-SGCK.

local region as described in [5]. The SVM-SGCK is implemented using the LIBSVM toolbox [9]. All experiments have been developed in MATLAB R2014b and executed on a computer with an Intel 2.7-GHz CPU and 8-GB RAM. All data have been normalized to the range [0, 1].

3.1 Data Set Description

The simulated SAR image is shown in Fig. 2(a), which is generated by adding three-look multiplicative speckle noise to a test image. The test image, as presented in Fig. 2(b), has the size of 255×255 pixels and contains three classes with three gray values (255, 117, and 0). The real SAR image is a 3-m spatial resolution Ku-band SAR image (256×256 pixels) of the China Lake Airport, California, as shown in Fig. 3(a). It consists of three land covers: runway (dark), vacancy area (gray), and airport buildings (bright). The corresponding Google earth image is shown in Fig. 3(b). For the quantitative comparison, the overall accuracy (OA), average accuracy (AA), and Kappa coefficient are adopted as the metrics to evaluate the classification results [10].

3.2 Parameter Selection

The penalty parameter C and the RBF kernel parameter σ involved in the four methods are selected by performing a threefold cross-validation on the training set. The parameter σ varies in the range $\{2^{-3}, \dots, 2^3\}$, and C ranges from 10^0 to 10^5 . As indicated in [5], the spatial kernel should be assigned with larger weight to obtain smoother classification results. Therefore, for ELM-CK, SVM-SGCK, and the proposed ELM-SGCK, the kernel weight is set to 0.8. The square window size is set to 11×11 for ELM-CK. For the superpixel generating algorithm in Sect. 2.2, the number of superpixels is set to 500 for the simulated SAR image and 1000 for the real SAR image. The parameter m is always set to 0.1 for the two images so that the generating superpixels

Table 1 Classification accuracy of the simulated SAR image obtained by various classification methods.

	KELM	ELM-CK	SVM-SGCK	ELM-SGCK
OA (%)	96.90	99.84	99.86	99.86
AA (%)	89.68	99.48	99.65	99.64
Kappa	0.9106	0.9955	0.9961	0.9960

adhere well to image boundaries.

3.3 Classification Results and Analysis

The first experiment was performed on the simulated SAR image. In this experiment, 1% of the total pixels were randomly selected as the training samples and the remaining 99% of data as the test set. As shown in Fig. 2(c)-(f), due to speckle noise, KELM produces a noisy “salt-and-pepper” classification map. By introducing the spatial information, the ELM-CK can produce smoother classification maps. However, the spatial information extracted from the fixed-size square-window results in that ELM-CK cannot accurately preserve the edges (e.g., red circle regions in Fig. 2). Thanks to the proposed SGCK, both the SVM-SGCK and ELM-SGCK can not only produce a smoother appearance but also preserve more accurate edge details. The quantitative results in terms of OA, AA, and the Kappa coefficient for various classification methods are shown in Table 1. The reported accuracy values are the average results over five runs with randomly selected training samples. As can be seen, the SVM-SGCK and ELM-SGCK obtain almost the same classification accuracy, which is higher than the traditional KELM and ELM-CK.

The second experiment was conducted on the real SAR image of China Lake Airport. Due to the lacking of the ground truth, we manually labeled some easily recognizable pixels for each class according to the corresponding Google image and the related geographic information. The labeled pixels are used as training samples, which make up about 1% of the total pixels in the SAR image. Because the ground truth corresponding to the China Lake Airport real SAR image is absent, the evaluation of the proposed method is based on visual inspection. According to [11], the accurate location of runway contour and airport buildings and the regional consistency of the vacancy region are concerned in the paper. Figure 3(c) shows the classification result of KELM. We can see that the classification result is noisy. That is because the spatial information is not considered in KELM. Figure 3(d) presents the classification result of ELM-CK. It is observed that the label is more consistent in the homogeneous region, but the contour of the runway (e.g., red rectangle region) and buildings structures (e.g., red ellipse region) are not well preserved. This is because the fixed-size window will inevitably damage the edge and structure details in the near-edge regions. Figure 3(e) shows the classification result of SVM-SGCK. Although the narrow runway in red rectangle region can be detected, the edges of the runway in other region (e.g., blue rectangle region) and some buildings (e.g., red ellipse region) cannot

Table 2 Computing time (in seconds) for the classification of the two test images obtained by various classification methods.

	KELM	ELM-CK	SVM-SGCK	ELM-SGCK
<i>Simulated SAR image</i>	1.51	4.65	7.25	4.62
<i>Real SAR image</i>	1.55	4.85	105.81	5.04

be located accurately. In addition, some pixels belonging to the vacancy area are misclassified as airport buildings. As shown in Fig. 3 (f), the classification result of the proposed ELM-SGCK gives clearer contour of the runway, more precise location of airport buildings, and better regional consistency of the vacancy region than other methods.

In order to compare the efficiency of different classification methods, the running times of various classification methods for classifying the two images are reported in Table 2. SVM-SGCK always consumes more time than ELM-CK and ELM-SGCK. Particularly, it takes about 106s for the real SAR image. That is because SVM-SGCK involves in solving the numerical optimization problem, and its convergence speed is related to characteristics of the data. In contrast, the proposed ELM-SGCK can be analytically computed, and therefore have higher computational efficiency than SVM-SGCK. It takes about 5s for the two images. Noting that the superpixel generating algorithm is very fast, its computing time can be negligible. Overall, the proposed ELM-SGCK can not only improve the classification performance, but also has very high computational efficiency.

4. Conclusion

In this paper, a new classification method named ELM-SGCK has been proposed. Firstly, a modified SLIC superpixel generating algorithm based on the GLR similarity criterion is developed for SAR image. Then, the SGCK is designed to jointly utilize the backscatter intensity information and the spatial information within each superpixel. Finally, the SGCK is incorporated into the ELM to implement classification. Experimental results on both simulated and real SAR images demonstrate that the proposed method outperforms traditional CK classification approaches, in terms of label consistency, detail preservation, and computational speed. The future work will investigate more effective techniques to exploit the spatial information within the superpixel.

Acknowledgments

The authors would like to thank Sandia National laboratories for providing the “China Lake Airport” real SAR image, Prof. C. Lin for providing LIBSVM toolbox, and Prof. G. Huang for providing ELM codes. They would also like to thank the editors and all of the anonymous reviewers for their constructive comments.

References

- [1] C.J. Oliver, “Rain forest classification based on SAR texture,” *IEEE Trans. Geosci. Remote Sens.*, vol.38, no.2, pp.1095–1104, 2000.
- [2] Ö. Aytekin, M. Koç, and İ. Ulusoy, “Local primitive pattern for the classification of SAR images,” *IEEE Trans. Geosci. Remote Sens.*, vol.18, no.9, pp.1549–1562, 2012.
- [3] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, “Extreme learning machines for regression and multiclass classification,” *IEEE Trans. Syst., Man, Cybern.-Part B: Cybernetics*, vol.42, no.2, pp.513–529, 2012.
- [4] G. Camps-Valls, L. Gomez-Chova, J. Munoz-Mari, J. Vila-Frances, and J. Calpe-Maravilla, “Composite kernels for hyperspectral image Classification,” *IEEE Geosci. Remote Sens. Lett.*, vol.3, no.1, pp.93–97, 2006.
- [5] Y.C. Zhou, J.T. Peng, and C.L.P. Chen, “Extreme Learning Machine With Composite Kernels for Hyperspectral Image Classification,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol.8, no.6, pp.2351–2360, 2015.
- [6] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, “SLIC superpixels compared to state-of-the-art superpixel methods,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.34, no.11, pp.2274–2282, 2012.
- [7] C.-A. Deledalle, L. Denis, and F. Tupin, “How to compare noisy patches? Patch similarity beyond Gaussian noise,” *International Journal of Computer Vision*, vol.99, no.1, pp.86–102, 2012.
- [8] K. Conradsen, A.A. Nielsen, J. Schou, and H. Skriver, “A test statistic in the complex Wishart distribution and its application to change detection in polarimetric SAR data,” *IEEE Trans. Geosci. Remote Sens.*, vol.41, no.1, pp.4–19, 2003.
- [9] C.-C. Chang and C.-J. Lin, “LIBSVM: a library for support vector machines,” *ACM Transactions on Intelligent Systems and Technology*, vol.2, no.3, pp.1–27, 2011.
- [10] C. Oliver and S. Quegan, *Understanding Synthetic Aperture Radar Images*, Artech House, Norwood, MA, 1997.
- [11] J. Gu, L.C. Jiao, S.Y. Yang, F. Liu, B. Hou, Z.Q. Zhao, “A Multi-kernel Joint Sparse Graph for SAR Image Segmentation,” *IEEE Journal of Selected Topics In Applied Earth Observations and Remote Sensing*, vol.9, no.3, pp.1265–1285, 2016.