SEA ICE CLASSIFICATION FROM HYPERSPECTRAL IMAGES BASED ON SELF-PACED BOOST LEARNING

Dong Wang[1,2], *Feng Gao*[1,2,*], *Junyu Dong*[1,2], *Yang Yang*[3], *Shengke Wang*[1,2]

[1]College of Information Science and Engineering, Ocean University of China
[2] Qingdao Key Laboratory of Mixed Reality and Virtual Ocean, Ocean University of China
[3] School of Law & Politics, Ocean University of China

ABSTRACT

Hyperspectral imagery has evident advantages for sea ice classification due to enormous spectral bands. In this paper, we proposed a novel sea ice classification framework from hyperspectral image based on self-paced boost learning (SPBL). First, the criterion of linear prediction error is used for unsupervised band selection. Then, local binary pattern (LBP) features are extracted from the selected bands. Finally, SPBL is employed as the classifier to provide probability outputs using the extracted features. The proposed framework can capture the intrinsic inter-class discriminative models while ensuring the reliability of the samples involved in learning. The experimental results in real-world dataset demonstrate that the proposed framework is superior to several closely related methods.

***Index Terms*—** Sea ice, hyperspectral image, local binary patterns, self-paced boost learning.

1. INTRODUCTION

Sea ice covers about 25 million square kilometers of the earth. It is a critical component especially in remote polar ocean as it influences climate, wildlife and people who live in the Arctic. Specifically, if gradually warming temperatures melt sea ice over time, surfaces of fewer bright are available to reflect sunlight back into space, and temperatures rise further. This chains of events start a cycle of warming and melting ice, which makes the polar regions the most sensitive areas to climate change on the Earth. In addition, the observation of sea ice is important for safe navigation. In the Arctic, sea ice can be an obstacle to normal shipping routes through the Northern Sea route and Northwest Passage. Therefore, the observations of sea ice have been paid more and more attention recently.

Remote sensing technology can capture large areas of images of sea ice rapidly, and developing robust sea ice classification techniques have been a long-standing goal for operational ice charting services. Hyperspectral images can obtain enormous continues spectrum information, and they are an important resource for sea ice classification. However, hyperspectral image poses big challenges, such as the well-known Hughes phenomenon [1]. Hughes phenomenon means that an increase in dimensions of limited training samples will cause a decrease in classification performance. To solve the problem, feature extraction is considered as one of the most challenging tasks in hyperspectral image applications. Many techniques have been proposed for hyperspectral image classification. Li and Qian [2] presented a hyperspectral classification method based on Gabor features. When combined with nearest regularized subspace, the Gabor features have strong power in spatial information representation. Mirzapour and Ghassemian [3] proposed a feature extraction scheme which combines morphological profiles, global Gabor features, and grey-level co-occurrence matrices. The combined feature set obtains satisfactory results in hyperspectral image classification. Li *et al.* [4] proposed a framework based on local binary patterns (LBP) [8] to extract local image features for hyperspectral image classification. LBP is a simple yet effective operator to describe local spatial information. If combined with more advanced classifiers, the performance of hyperspectral image classification may be further improved.

Recently, self-paced learning [5] has been attracting increasing attention in the filed of machine learning and computer vision. It is inspired by the learning principle underlying the cognitive process of humans. Specifically, humans generally start with learning easier aspects of target task, and then gradually take more complex example into consideration. Pi *et al.* [6] noticed that boosting and self-paced learning are consistent in basic principles and complementary in methodology. Therefore, they proposed self-paced boosting learning (SPBL), which learns a joint manner from weak models to strong model and from easy samples to complex ones. The verified empirically one is that SPBL is effective in computer vision applications. However, in remote sensing applications, SPBL has rarely been considered.

In this paper, we presented a novel sea ice classification

This work was partially supported by the National Natural Science Foundation of China under Grant 41606198 and 41576011, and Natural Science Foundation of Shandong Province under Grant ZR2016FB02. (Email: gaofeng@ouc.edu.cn)

framework from hyperspectral image based on self-paced boost learning. First, a similar feature extraction scheme as mentioned in Li's work [4] is employed. The criterion of linear prediction error (LPE) [7] is used for unsupervised band selection. Then, LBP features are extracted from these bands. Finally, SPBL is employed as the classifier to provide probability outputs using the extracted features.

2. METHODOLOGY

The framework of the proposed method is illustrated in Fig. 1. It consists of two main steps: 1) LBP feature extraction; 2) feature classification based on SPBL. In the remainder of this section, we will describe more details about the feature extraction and classification.

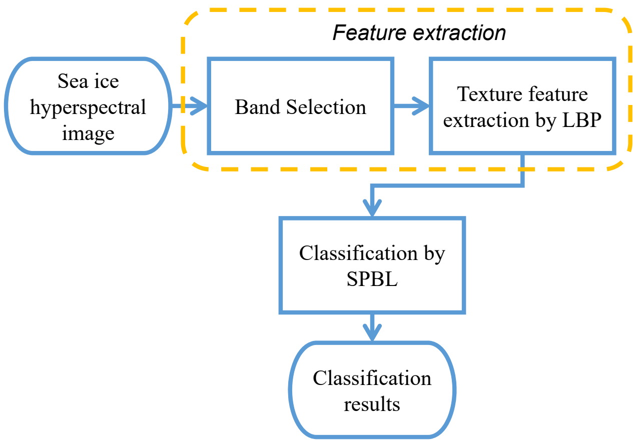

Fig. 1. Framework of the proposed hyperspectral classification method

2.1. Feature Extraction

In the feature extraction, similar to Li's work [4], LPE is first used to select a subset of spectral bands with distinctive and informative features. Considering two spectral bands B_1 and B_2, the third band B can be denoted by $B' = a_0 + a_1 B_1 + a_2 B_2$, where a_0, a_1, a_2 are the parameters that can minimize the linear prediction error. Let the parameter vector $\mathbf{a} = [a_1, a_2, a_3]^T$, and a least square solution can be employed to obtain the parameter vector as follows:

$$\mathbf{a} = (M^T M)^{-1} M^T m_B, \tag{1}$$

where M is an $N \times 3$ matrix whose first column is with all 1s, second column is the B_1-band, and third column is the B_2-band. m_B is the B-band. The band that produces the maximum error is considered as the most dissimilar band to B_1 and B_2. Then, the band will be selected. Similarly, using these three bands, a fourth band can be found, and so on.

After band selection, LBP operator is used to extract contextual features from the selected bands. LBP is a gray-scale and rotation-invariant texture operator [8]. Given a center pixel p_c, each neighbor of a local region is assigned with a binary label, which can be either 1 or 0, depending on whether the scalar value has a larger of intensity value or not. The neighboring pixels are from a set of equally spaced samples over a circle of radius r centered at the scalar value. r determines how far the neighboring pixels can be located away from the center pixel. Along with selected x neighbors $\{p_i\}_{i=0}^{x-1}$, the LBP code for the scalar value p_c is given by

$$\mathrm{LBP}_{x,r}(p_c) = \sum_{i=0}^{x-1} U(p_i - p_c) 2^i, \tag{2}$$

where $U(p_n - p_c) = 0$, if $p_n \le p_c$ and $U(p_n - p_c) = 1$ if $p_n > p_c$. After obtaining the LBP codes, an occurrence histogram is computed over a local patch. A binning procedure is required to guarantee that the histogram features have the same dimension.

2.2. Classification by Self-Paced Boosting Learning

The extracted LBP features are classified by SPBL. Let $\{(x_i, y_i)\}_{i=1}^{n}$ be the training samples, where $x_i \in \mathbb{R}^d$ is the feature of sample i, y_i is the class label of x_i. The robustness of a learning method relies on the loss function to relieve the influence of noisy and confusing data. Instead of directly learning from all the samples, SPBL aims to guide the boosting model to learn asymptotically from the easy samples to complex ones. Therefore, the general objective of SPBL can be formulated as:

$$\begin{aligned} &\min_{W,s} \sum_{i=1}^{n} s_i \sum_{r=1}^{C} L(\rho_{ir}) + \sum_{i=1}^{n} g(s_i;\lambda) + \nu R(W), \\ &s.t.\ \forall i, r, \rho_{ir} = K_{i:} w_{yi} - K_{i:} w_r; W \ge 0; s \in [0,1]^n, \end{aligned} \tag{3}$$

where $K \in \mathbb{R}^{n \times z}$ is the weak classifiers' responses for the training data with $[K_{ij}] = [k_j(x_i)]$, and $K_{i:}$ is the i^{th} row of K; $s_i \in [0,1]$ is the self-paced learning weight of sample x_i that denotes its learning "easiness". $g(\cdot;\lambda) \to \mathbb{R}$ is the self-paced learning function that specifies how samples are selected. A weight s_i is assigned to each sample as a measure of its "easiness". The function $g(s_i;\lambda)$ can dynamically select the easily learned samples that are more discriminative. With joint optimization of parameter s and parameter W, the SPBL model gradually incorporates the training samples from easy ones to complex ones.

An alternating optimization is employed to solve Eq. (3), which optimizes each of the two variables with the other one fixed in an alternating manner. s is optimized as follows:

$$s_i^* = \arg\min_{s_i} s_i l_i + g(s_i;\lambda), \quad s.t.\ \ s_i \in [0,1], \tag{4}$$

where $l_i = \sum_r \ln(1 + e^{-\rho_{ir}})$ denotes the loss of sample x_i. In addition, W is optimized as follows:

$$W^* = \arg\min_W \sum_{i,r} s_i \ln(1 + e^{-\rho_{ir}}) + \nu \|W\|_{2,1} \quad (5)$$
$$s.t.\ \forall i, r, \rho_{ir} = H_{i:}w_{yi} - H_{i:}w_r; W \geq 0.$$

To solve W in the above equation, the column generation method is employed, and the set of active weak classifiers is augmented. Then the optimization continues with a new set of active weak classifiers, until the objective value of boosting reaches a tolerance threshold. Detailed information about the SPBL algorithm can be found in Pi's work [6].

3. EXPERIMENTAL RESULTS AND ANALYSIS

3.1. Experimental Setup

The dataset employed in this paper was acquired on April 12th in 2014, from Hyperion sensor of EO-1. The image covers an area of Baffin Bay in Greenland, and is comprised of 242 spectral bands. The size of the image is 350×300 pixels. There are mainly four classes from the ground truth map: land, sea water, gray ice and white ice. For illustrative purposes, Fig. 2(a) shows a false color composition of the scene, while Fig. 2(b) shows the ground truth map of the scene. Here we randomly select 10% of all labeled pixels in each class for training, and the others are used for testing. More detailed information of the number of training and testing samples is summarized in Table 1.

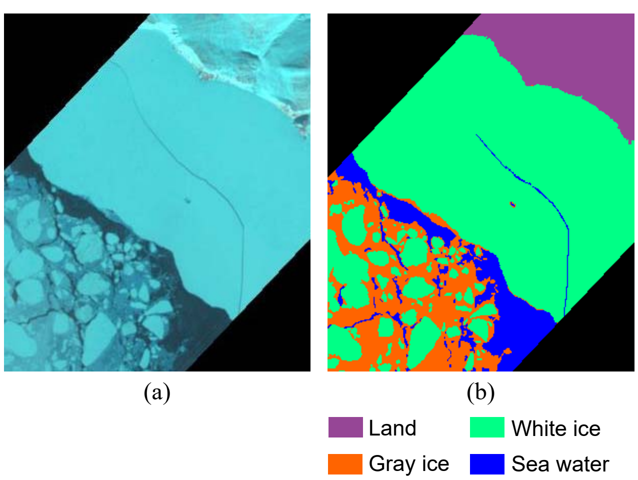

Fig. 2. Baffin Bay dataset. (a) False color composition. (b) Ground truth map.

3.2. Results and Analysis

The performance of the proposed classification methods is shown in Tables 2. We compare the proposed method with two closely related hyperspectral classification methods: LBP-ELM [10] and IFRF [9]. We use the default parameters of the compared methods which were provided in the corresponding references.

Table 1. Class labels and train-test distribution of samples

#	*Class*	*Train*	*Test*
1	Land	1140	10255
2	Sea water	5269	47421
3	Gray ice	1368	12310
4	White ice	750	6754
	Total	8527	76740

Table 2. Classification accuracies of different methods

Class	LBP-ELM	IFRF	Proposed
Land	87.35	98.92	95.18
Sea water	87.60	99.18	97.84
Gray ice	63.60	77.49	95.90
White ice	65.93	81.38	95.85
OA(%)	82.21	93.28	**97.01**
$\kappa \times 100$	67.97	88.54	**94.70**

Fig. 3 presents the classification results by different methods. From visual comparison, it can be observed that there are many noisy pixels in the result generated by LBP-ELM, from the Table. 2, and it can be observed that the proposed method surpasses LBP-ELM by 14.8% in OA. In addition, in the result generated by IFRF, a lot of small white ice is classified incorrectly into gray ice. Therefore, the classification accuracy of IFRF is much lower than the proposed method, and the proposed method surpasses IFRF by 6.2% in OA. The experimental results indicate that the proposed method can achieve good accuracy in sea ice classification by capturing the intrinsic inter-class discriminative patterns while ensuring the reliability of the samples involved in learning.

4. CONCLUSION

In this paper, we proposed a novel sea ice classification framework from hyperspectral image based on self-paced boost learning. First, LPE is employed for unsupervised band selection. Then, LBP features are extracted from the selected bands. Finally, SPBL is employed as the classifier to provide

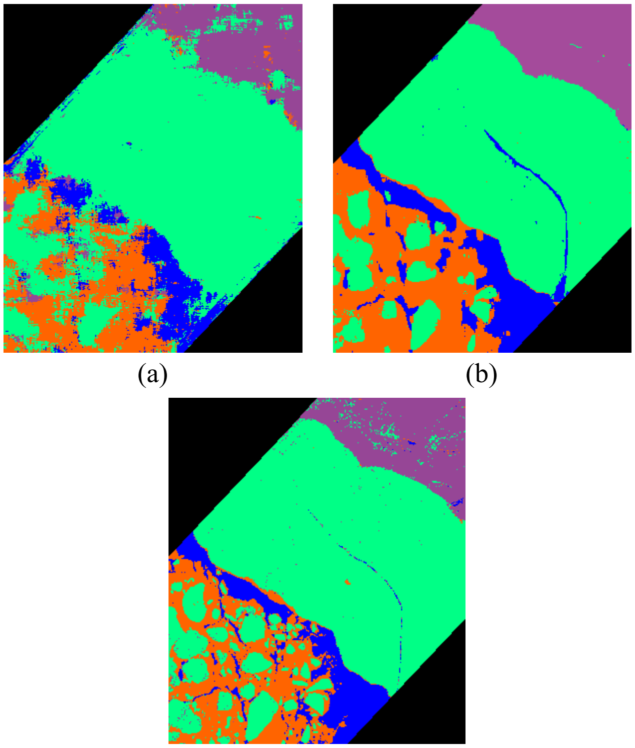
(a) (b) (c)

Fig. 3. Classification results by different methods on the Baffin Bay dataset. (a) LBP-ELM. (b) IFRF. (c) Proposed method.

probability outputs using the extracted features. The proposed framework can capture the intrinsic inter-class discriminative models while ensuring the reliability of the samples involved in learning. The experimental results in real-world dataset show the effectiveness of the proposed method.

In the future, we will focus on learning the more sophisticated feature selection approaches and the corresponding class wise contributions, simultaneously.

5. REFERENCES

[1] M. Gong, M. Zhang, Y. Yuan, "Unsupervised band selection based on evolutionary multobjective optimization for hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 1, pp. 544–557, 2016.

[2] W. Li and Q. Du, "Gabor-filtering based nearest regularized subspace for hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1012–1022, 2014.

[3] F. Mirzapour and H. Ghassemian, "Improving hyperspectral image classification by combining spectral, texture, and shape features," *Int. J. Remote Sens.*, vol. 36, no. 4, pp. 1070–1096, 2015.

[4] W. Li *et al.*, "Local binary patterns and extreme learning machine for hyperspectral imagery classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 7, pp. 3681–3693, 2015.

[5] Q. Zhao *et al.* "Self-paced learning for matrix factorization," *AAAI Conference on Aritificial Intelligence*, pp. 3196–3202, 2015.

[6] T. Pi *et al.*, "Self-paced boost learning for classification," *International Joint Conference on Artificial Intelligence*, pp. 1932–1938, 2016.

[7] Q. Du, H. Yang, "Similarity-based unsupervised band selection for hyperspectral image analysis," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 4, pp. 564–568, 2008.

[8] T. Ojala, M. Pietikainen, T. T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary pattern," *IEEE Trans. Pattern Analysis Mach. Intell.*, vol. 24, no. 7, pp. 971–987, 2002.

[9] X. Kang, S. Li, J. A. Benediktsson, "Feature extraction of hyperspectral images with image fusion and recursive filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, no. 6, pp. 3742–3752, 2014.

[10] W. Li, c. Chen, H. Su, and Q. Du, "Local binary patterns and extreme learning machine for hyperspectral imagery classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 7, pp. 3681-3693, 2015.