Joint Kernelized Sparse Representation Classification for Hyperspectral Imagery

He Sun, Jinchang Ren, Yijun Yan, Jaime Zabalza, Stephen Marshall
Department of Electronic and Electrical Engineering
University of Strathclyde
Glasgow, United Kingdom
Jinchang.ren@strath.ac.uk

I. Introduction

In recent years, the hyperspectral image (HSI) classification has received much attention due to its importance on the military applications, food quality assessment [1], and land cover analysis [2-5], etc. Multiple classifiers have been adopted to label pixels of HSI images, including support vector machine (SVM), random forest (RF), and recently, the deep learning methods. Considering that HSI pixels belonging to the same class are usually lying in a low-dimensional space, those pixels can be represented by training samples from the same class. Based on that, Sparse Representation Classification (SRC) methods have also introduced in the HSI imagery. For an unlabeled pixel, a few atoms from the constructed training dictionary can sparsely represent it. With the recovered sparse coefficients, the class label can be determined by the residual between the test pixel and its approximation.

With the development of SRC in HIS [5, 6], there is one severe problem during the process of classification. Due to the high dimensions of the HSI data, it may result the Hughes phenomenon. Sufficient training samples are required to overcome the curse of dimensionality. However, sufficient training data are not always available in real application. For example, the ground truth labelling work for remote sensing data is rather inconvenient. Therefore, to solve the above problem, we decide to combine multiple types of features extracted from HSI data, and a joint kernelized SRC will be operated on those extracted features. The aim of our work is to improve the performance of SRC with less training samples.

II. Proposed Method

Our proposed algorithm is comprised of feature extraction and joint kernelized SRC.

A. Feature Extraction

Given the HSI data, three types of spectral and spatial feature descriptors are implemented in our work.

1) *Spectral*: The first one is the original spectral feature, which provides the basic information from the spectral. In this paper, the spectral information is extracted by principle component analysis (PCA) from the original spectral data.

2) *EMP*: With multiple morphological opening and closing with structuring elements, a morphological profile (MP) [3] can be built. After the integration of several MPs, the EMP (Extended MP) feature can be yielded, which can enhance the spatial information.

3) *Gabor Feature*: The Gabor filter is usually applied as a texture feature extractor and it is defined by a Gaussian function multiplying a sine wave.

B. Joint Kernelized SRC

Before the SRC work, we apply ℓ-2 normalization on the acquired features, the ℓ-2 normalization can reinforce the discrimination between different classes. In our work, we consider the test pixel with its neighbor pixels in a 9*9 region together instead of taking the test pixel independently. The label of the test pixel is determined by a joint strategy. By combing different features, the test pixel can be resented by $x^F = \{x^m\}, m = 1,2,3$, where x^m is the m th feature. The corresponding dictionary (training samples) can be also constructed as $D^F = \{D^m\}, m = 1,2,3$.

With the kernel trick k , the kernel matrices of the test pixel and the dictionary can be estimated respectively, which is expressed as $K_{x,D^F} = k(x^m, D^m)$ and $K_{D^F} = k(D^m, D^m)$. The index set Λ_0 is initialized by $argmax(K_{D^F})$. And the correlated matrix C is computed by:

$$C = K_{D^F} - \left(K_{x,D^F}\right)_{:,\Lambda_{a-1}} \left(\left(K_{D^F}\right)_{\Lambda_{a-1},\Lambda_{a-1}} + \lambda I\right)^{-1} * \left(K_{x,D^F}\right)_{\Lambda_{a-1},:} \tag{1}$$

The a is the iteration counter with a maximum value equals to the sparsity level and the λ is the ℓ-2 regularized term. The new index can be selected as $\mu_t = argmax(C)$ and the index set can be updated by $\Lambda_t = \Lambda_{t-1} \cup \{\mu_t\}$. The final index set will determine the class label. And in our work, the kernel function we applied is the Radius Basis Function (RBF) kernel with a kernel parameter γ.

III. Experiment Result

In this section, we apply and evaluate our proposed method on the publicly Pavia University dataset, which is a hyperspectral image dataset acquired by airborne system of NASA. The corrected Pavia University dataset has 103 bands and labelled in nine classes, the resolution of this dataset is 610*340. In our experiments, the number of training samples are 20 per class and the rest samples are assumed as test samples.

In our work, the result of classification is evaluated by the overall accuracy (OA), average accuracy (AA) and Kappa coefficient. To better understand the performance of our SRC algorithm, two state-of-art classification methods are applied. The first one is the composite kernel support vector machine (CK-SVM) which is implemented on multiple features as well [3]. The second one is the multiple feature adaptive sparse representation (MFASR) [4]. All the experiments are repeated ten times to acquire the average result of OA, AA and Kappa. The results are shown in Fig. 1 and Table I.

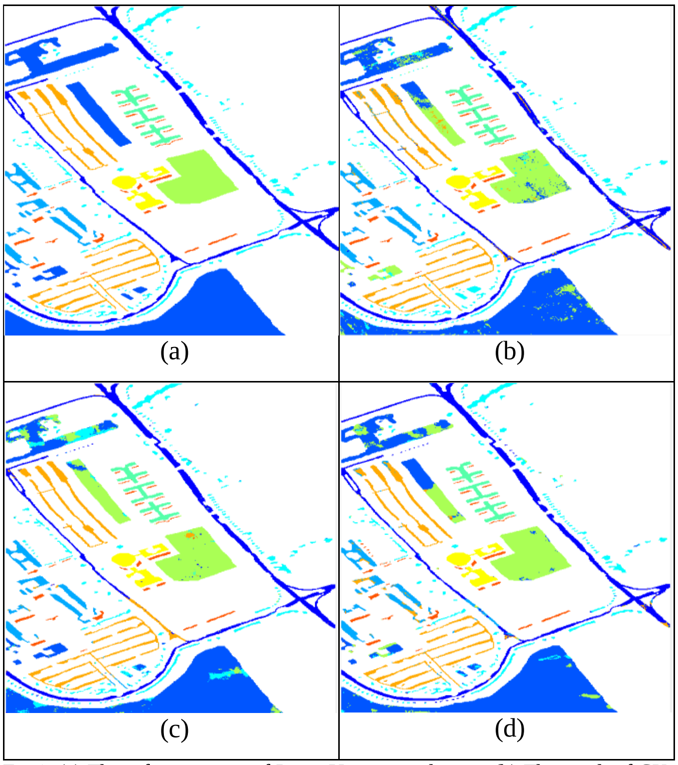

Fig.1. (a) The reference map of Pavia University dataset. (b) The result of CK-SVM. (c) The result of MFASR. (d) The result of our method.

TABLE I

	CK-SVM	*MFASR*	*Ours*
OA(%)	88.74	88.52	91.94
AA(%)	94.81	95.50	89.54
Kappa	85.54	85.40	95.42

IV. Conclusion

In this paper, we proposed a joint kernelized SRC method for multiple features. With multiple features and kernel methods, the performance of SRC is improved when the number of training samples are relatively low. Our experimental results demonstrate that our proposed algorithm can obtain satisfied results and outperform some state-of-art algorithms. In our experiment, the training samples are set as 20 per class, future work will keep improving the performance with less training samples. In addition, the dictionary we used in this paper is acquired directly from the extracted features. Our future work will also focus on applying effective dictionary learning algorithm for constructing better dictionary, relevant techniques like fusion mechanism [7, 8] and some feature extraction techniques [9-11] might be considered as well.

Acknowledgement

Authors would like to thank Prof. P. Gamba for providing the remote sensing dataset. And we would also like to thank Dr L. Fang and Dr J. Li for providing the online code of MFASR and CK-SVM.

References

[1] J. Tschannerl, J. Ren, F. Jack, J. Krause, H. Zhao, W. Huang and S. Marshall, “Potential of UV and SWIR hyperspectral imaging for determination of levels of phenolic flavour compounds in peated barley malt,” *Food Chem.*, vol 270, pp. 105-112, Jan. 2019.

[2] J. Zabalza, J. Ren, J. Zheng, J. Han, H. Zhao, S. Li and S. Marshall, “Novel two-dimensional singular spectrum analysis for effective feature extraction and data classification in hyperspectral imaging,” *IEEE Trans. Geosci. Remote Sens*., vol 53, no. 8, pp. 4418-4433, Aug. 2015.

[3] J. A. Benediktsson, J. A. Palmason, and J. R. Sveinsson. “Classification of hyperspectral data from urban areas based on extended morphological profiles,” *IEEE Trans. Geosci. Remote Sens*., vol 43, no. 3, pp. 480-491, Mar. 2015.

[4] J. Li, P. R. Marpu, A. Plaza, J. M. Bioucas-Dias and J. A. Benediktsson, “Generalized Composite Kernel Framework for Hyperspectral Image Classification”. *IEEE Trans. Geosci Remote Sens*, vol. 51, no. 9, pp. 4816-4829, Sep. 2013.

[5] L. Fang, C. Wang, S. Li, and J. A. Benediktsson. “Hyperspectral image classification via multiple-feature-based adaptive sparse representation,” *IEEE Trans. Instrum.Meas.*, vol 66, no. 7, pp. 1646-1657, Jul. 2017.

[6] T. Qiao, Z. Yang, J. Ren, P. Yuen, H. Zhao, G. Sun, S. Marshall and J. A. Benediktsson. “Joint bilateral filtering and spectral similarity-based sparse representation: a generic framework for effective feature extraction and data classification in hyperspectral imaging,” *Pattern Recognit.*, vol 77, pp. 316-328, May. 2018.

[7] Y. Yan, J. Ren, H. Zhao, G. Sun, Z. Wang, J. Zheng, S. Marshall and J. Soroghan. “Cognitive fusion of thermal and visible imagery for effective detection and tracking of pedestrians in videos,” *Cogn. Comput.*, vol 10, no. 1, pp. 94-104, Feb. 2018.

[8] Y. Yan, J. Ren, Y. Li, J. Windwill, W. Ijomah, K. Chao. “Adaptive fusion of color and spatial features for noise-robust retrieval of colored logo and trademark images,” *Multidim. Syst. Sign. P.*, vol 27, no. 4, pp. 945-968, Oct. 2016.

[9] Q. Liu, Y. Wang, M. Yin, J. Ren and R. Li. “Decontaminate feature for tracking: adaptive tracking via evolutionary feature subset,” *J. Electron Imaging.*, vol 26, no. 6. 2017.

[10] J. Zabalza, J. Ren, Z. Wang, S. Marshall and J. Wang. “Singular spectrum analysis for effective feature extraction in hyperspectral imaging,” *IEEE Geosci Remote Sens Letter*., vol 11, no 11, pp. 1886-1890. Nov. 2011.

[11] J. Zabalza, C. Qing, P. Yuen, G. Sun, H. Zhao and J. Ren. “Fast implementation of two-dimensional singular spectrum analysis for effective data classification in hyperspectral imaging,” *J Franklin Inst.*, vol 355, no 4, pp. 1733-1751. Mar. 2018.

[12] T. Qiao et al. “Effective denoising and classification of hyperspectral images using curvelet transform and singular spectrum analysis,” *IEEE Trans. Geosci. Remote Sens*., vol 55, no 1, pp. 119-133. Jan. 2017.