Multi-feature Fusion Based Deep Forest for Hyperspectral Image Classification

Xiaobo Liu[1,2,3](✉), Mostofa Zaman Mohammad[1,2,3], Chaochao Zhang[1,2,3], Xin Gong[1,2,3], and Zhihua Cai[4]

[1] School of Automation, China University of Geosciences, Wuhan 430074, China
xbliu@cug.edu.cn
[2] Hubei Key Laboratory of Advanced Control and Intelligent Automation for Complex Systems, Wuhan 430074, China
[3] Engineering Research Center of Intelligent Technology for Geo-Exploration, Ministry of Education, Wuhan 430074, China
[4] School of Computer Science, China University of Geosciences, Wuhan 430074, China

Abstract. Multi-feature fusion is a useful way to improve the classification of hyperspectral image (HSI). But the multi-feature fusion is usually at the decision level of classifier, which causes less link between features or poor extensibility of feature. In this paper, we propose a multi-feature fusion based deep forest method for HSI classification, named mfdForest. In mfdForest, the morphological features, saliency features, and edge features are extracted, then the three deep multi-grained scanning branches in dgcForest (one of improved deep forest) are used to extract and fuse the extracted features deeply, and the fused features are sent into cascade forest in dgcForest for classification. Experimental results indicate that the proposed framework consumes less training time and has better performance on two HSI data sets.

Keywords: Deep forest · Hyperspectral image classification · Multi-feature fusion

1 Introduction

Hyperspectral images (HSI) are widely used in many fields, such as agriculture, military, medicine, due to the rich information it contains [1–4]. The increase in spatial resolution and spectral resolution makes HSI contain more fine-grained feature information. As the amount of information increases, stronger feature extraction capabilities are needed, multi-feature extraction and fusion are proven to be an effective way to improve feature extraction [5–7].

To make full use of information contained in hyperspectral data, multi-feature based strategy has been widely applied to current popular method. He et al. [8] proposed a multi-scale 3D deep convolutional neural network to jointly learn both 2D multi-scale spatial feature and 1D spectral feature from HSI data in an end-to-end approach, achieved better results with large-scale data set. However, Deep Neural Network (DNN) is hard

to fuse multi-feature obtained by different methods limited by the structure of neural network. Some methods use traditional extractor to extract multiple shallow features to improve utilization of HSI information by enhancing the diversity of features, such as Extended Morphological Profile (EMP), Gabor, Local Binary Patterns (LBP), Scale-invariant Feature Transform (SIFT), etc. [9–12]. Li et al. [13] used linear divergence analysis (LDA) to extract spectral features, and adaptive weighted filters (AWFs) to extract spatial information. After multiple iterations, they were fused with LBP features. The experimental results proved that the method can extract feature information further effectively. Zhang et al. [14] used Principal Component Analysis (PCA), Extended morphological profile (EMP), Differential Morphological Profiles (DMP), and Gabor filters to extract different features, and used Support Vector Machine (SVM) and Gabor filters for fusion and classification, achieved good performance. The aforementioned methods combine traditional feature extractor and extracted features have no ability to present well. Liu et al. [15] firstly extracted EMP features from HSI that was reduced by PCA, and then extracted Boolean Map based Saliency (BMS) visual features and fused them. While data redundancy is reduced, feature extraction is improved. However, BMS can not extract edge information well due to the less focus on edge of object. The above methods do not fully extract the spatial features.

After multi-feature extraction provides guarantee for HSI classification, the use of high-performance classifiers is also crucial in HSI classification. Deep forest is a new deep model of the alternative DNN proposed by Zhou et al. [11]. The deep forest used non-differentiable modules to build deep model, and the modules can be decision trees, random forest, et al. As the deep forest inherits the advantages of decision-tree ensemble approach, with fewer hyper-parameters than deep neural networks, and its model complexity can be automatically determined in a data-dependent way. The deep forest includes multi-grained scanning and cascade forest, the multi-grained scanning further enhanced the representational learning ability, potentially enabling deep forest to be contextually or structurally aware; the cascade levels can be automatically determined such that the model complexity can be determined in a data-dependent way. Yin et al. [12] applied deep forest to HSI classification, making it possible for deep forest to be used for HSI classification. Liu et al. [13] proposed a deep multi-grained scanning cascade forest (dgcForest), and improved the deep forest to make it more suitable for HSI classification by deep multi-grained scanning. But the dgcForest extracts features just using different sizes of sliding windows, not making full use of spatial information on HSI.

Extended morphological profile (EMP) is a simple and commonly used feature map method that can denoise images [19]. Saliency detection can highlight the spatial scene features [20]. However, saliency detection will weaken the boundary information of each area in HSI. Edge detection can detect the edge feature of feature map, which makes up for the deficiency of saliency detection. Thus, the combination of three features has a certain rationality.

In this paper, we proposed a multi-feature fusion based deep forest method for HSI classification. The main contributions of this paper are summarized as follows.

1) Three deep multi-grained scanning branches in dgcForest were used to deeply extract EMP features, saliency features and edge features, which can supplement each other, to make full use of spatial information of HSI.

2) In order to enhance features, voting fusion method was used to fuse three deeply extracted features, which can make the link between features closer than decision-level fusion method.

2   Proposed Method

Reasonable multi-feature selection, efficient multi-feature extraction and fusion are essential for promoting HSI classification. In this part, a novel framework is proposed to extract multiple complementary features in depth and fuse them effectively to improve the classification of HSI. The framework is shown in Fig. 1. Firstly, in order to decrease the redundancy of HSI in the spectral channel, PCA is introduced. Secondly, following the role played by EMP and Visual Saliency Detection (VSD) in previous studies, EMP operation is used on the global PCA image to remove the noise and smooth the image. And then VSD is used on EMP image to extract the more salient information, thereby reducing the impact of irrelevant background on HSI classification. However, VSD is difficult to detect the edge of ground objects, which will cut down the classification of pixels in the edge area. Therefore, edge detection (ED) is introduced to compensate for the loss of edge information caused by VSD, which is used to determine the edge by detecting the dramatic changes in the gray value around the pixel and calculating the reciprocal of the gray value change, and locate the edge by detecting the dramatic changes in the gray value around the pixel, that is, calculating the reciprocal of the gray value change. In the process of extracting edge information from EMP images, first, the EMP feature map is denoised by Gaussian blur. Second, a discrete differential operator called the Sobel operator is used to calculate the approximate gradient of the image, and Non-Maximum Suppression (NMS) is used to obtain the point with the largest gradient. Finally, the edge point is obtained by setting the threshold, and the edge maps are obtained by calculating all the edge points. The calculation process is shown in Eq. 1.

$$\begin{cases} C(E) = DUAL\_THRESHOLD\big(NMS\big(G_f\big)\big) \\ \quad G_f = \big(arctan\big(\frac{G_x}{G_y}\big), \sqrt{G_x^2 + G_y^2} \\ \quad\quad G_x, G_y = \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \end{cases} \tag{1}$$

In the Eq. 1, C(E) represents the operation of edge detection on the EMP feature map, DUAL_THRESHOLD indicates dual threshold filtering in CED, NMS is non-maximum suppression, G_f is the gradient of the pixel (x, y), including the horizontal gradient G_f and the vertical gradient G_f.

Although VSD and ED can extract semantic features, which is high-level feature compared to texture features. But in the face of HSI, traditional methods are difficult to extract enough useful information. To extract multi-feature effectively, our proposed method uses the efficient feature extractor of dgcForest, deep multi-grain scanning, to extract these three features separately. Firstly, each pixel of HSI is set as the center pixel and the sub-block is obtained by choosing a neighbor region of this pixel with size of $(2*w-l)*(2*w-l)$, where w and l is the size and step length of multi-grained scanning window of dgcForest, which ensures that each neighbor region obtained by deep multi-grained scanning contains the central pixel uniquely. Next, these sub-pixels

blocks containing global spatial and local spatial information are input in deep multi-grained scanning, and three multi-grained scanning will output three class probability vectors in parallel (random forest, the core component of deep multi-grain scanning, can output a probability vector of class that this sample belongs to). Then, the three feature vectors are simply added to obtain the final deep feature vector containing global and local information. Finally, the fused vector was sent into cascade forest to obtain final prediction.
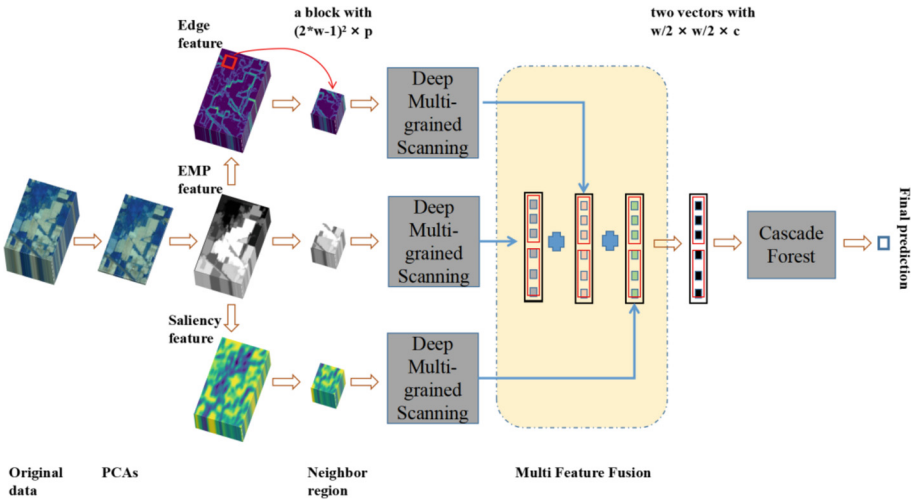

Fig. 1. The flowchart of our proposed method: mfdForest.

3   Experimental Results

This section presents the classification accuracy of several states-of-the-art methods and three cases that our method used different types of features. The performance of our proposed algorithm will be demonstrated through extensive experiments on two well-known HSI datasets, using three evaluative metrics including average accuracy (AA), overall accuracy (OA) and Kappa coefficient (Kappa). 10% of each category is selected as the training set on the two datasets, and the rest data as testing data, which are described in Table 1. The effects of several contributions of the proposed approach also be summarized in this section. All experiments were implemented on ubuntu16.04, with the CPU of Xeon E5-1603 and GPU of GeForce RTX1080Ti.

Table 1. The detail of two data sets.

Data set	Class	Pixel	Band	Resolution	Sample
Indian Pine	16	145*145	200	20 m	10249
Salinas	16	512*217	200	3.7 m	43980

To ensure the validity of the experiment, the parameters involved in our methods are consistent with comparative methods: ACGAN [18], ResNet [19], dgcForest [14]. The parameters involved are number of PCA, the number of morphological opening and closing operations on EMP, the size of neighbor region in dgcForest, the number of decision tree in random forest and completely random forest, the number of random forest and completely random forest in deep multi-grained scanning and cascade forest. In order to ensure the consistency of the experiment, the above parameters are set to the same. It's just that the layer of cascading forest is increased to 2. Table 2 shows the parameters setting. The number of features selected through the first layer of the cascading forest will be different according to the datasets. Except for the experiment about selecting the different number of features, other experiments uniformly use the optimal number corresponding to the datasets.

Table 2. Parameters setting of proposed method.

	PCA	3
Multi-feature extraction branches	Number of morphological opening and closing operations	4
	Neighbor region size	7
	Number of Random Forest	1
	Number of Completely Random Forest	1
	Number of Decision Tree	40
	Number of k-Fold Cross-Validation	5
Cascade Forest	Number of Levels	1
	Number of Random Forest	4
	Number of Completely Random Forest	4
	Number of Decision Tree	80

In order to evaluate the contribution of edge detection, two sets of experiments are prepared. It can be seen from in Fig. 2, Fig. 3, Table 3, Table 4, that the addition of feature extraction branches improves OA, AA, and KAPPA on three datasets. The EMP, EMP + Saliency, mfdForest means that our algorithm uses different type of feature extraction branch, the EMP just uses extended morphological profile feature for mfdForest, the EMP + Saliency uses extended morphological profile and Boolean Map Saliency Detection features for mfdForest, mfdForest uses extended morphological profile, Boolean Map Saliency Detection, and edge detection features together.

In Fig. 2 and Fig. 3, our algorithm has a better performance than other methods in OA in two data sets. Although our method has a lower accuracy than ResNet when using less than three branches in Indian Pines data set, but three branches can still have a best performance. What's more, in the case of our method using different number of extraction branches, the accuracy was increased with more branches. As can be seen from Table 3 and Table 4, with the addition of the morphological features, saliency features, and edge features extracted, the classification accuracy increases. It shows that each feature we chosen and feature fusion method we used has played a significant role. In Table 3 and Table 4, our method also has superiority in OA, AA and Kappa comparing to other methods, and AA has a subtle change in the case of using two extraction branches in Indian Pines data set. It shows that saliency detection hasn't improve classification in some classes, which is the point we need to solve later.

In terms of running time in Table 3 and Table 4, which is the total time of training and testing, although our improved algorithm is slower than dgcForest, the reduction in speed has not been presented as multiple, the accuracy of both is higher than the dgcForest, indicating that the three extracted features can improve the classification accuracy and complement each other.
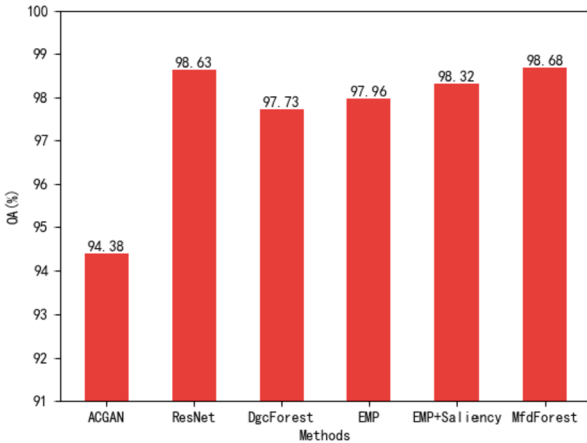

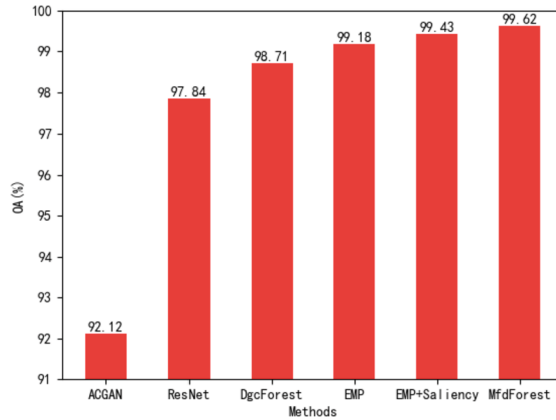**Fig. 2.** OA of different methods in Indian Pines

Fig. 3. OA of different methods in Salinas.

Table 3. Accuracy comparison of different algorithms in Indian Pines datasets.

Method	ACGAN	ResNet	DgcForest	EMP	EMP + Saliency	mfdForest
OA	94.38 ± 1.88	98.63 ± 0.25	97.73 ± 0.23	97.96 ± 0.3	98.32 ± 0.14	**98.68 ± 0.15**
AA	76.29 ± 3.28	90.33 ± 0.29	95.92 ± 0.51	94.62 ± 2.62	94.59 ± 2.02	**96.8 ± 1.54**
K*100	93.58 ± 2.15	98.44 ± 0.28	97.41 ± 0.18	96.23 ± 0.77	97.37 ± 1.01	**98.55 ± 0.71**
1	57.14 ± 46.78	**100.00 ± 0.00**	89.63 ± 2.30	96.97 ± 2.93	96.10 ± 1.95	97.56 ± 1.95
2	90.41 ± 4.26	**98.81 ± 00.46**	96.58 ± 0.66	96.45 ± 0.91	97.70 ± 0.92	98.44 ± 0.69
3	**99.41 ± 0.40**	98.35 ± 1.33	96.32 ± 0.90	96.78 ± 2.05	98.39 ± 1.14	97.72 ± 1.75
4	91.97 ± 1.58	98.64 ± 1.92	**99.06 ± 0.39**	95.21 ± 1.44	97.84 ± 0.97	98.6 ± 1.25
5	97.11 ± 1.46	98.06 ± 1.81	94.08 ± 0.51	96.14 ± 2.30	96.05 ± 1.88	**98.39 ± 2.13**
6	98.32 ± 0.60	98.59 ± 1.29	**99.32 ± 0.25**	98.6 ± 1.07	98.00 ± 1.32	99.09 ± 0.86
7	20.00 ± 40.00	66.67 ± 27.14	84.00 ± 1.73	67.2 ± 19.82	64.8 ± 14.4	**90.00 ± 13.24**
8	94.82 ± 2.94	99.39 ± 00.54	**99.88 ± 0.10**	99.81 ± 0.17	99.67 ± 0.35	99.07 ± 0.35
9	0.00 ± 0.00	0.00 ± 0.00	94.44 ± 0.00	85.56 ± 20.96	81.11 ± 19.12	**99.53 ± 0.16**
10	93.88 ± 2.50	97.34 ± 2.31	96.46 ± 0.09	97.17 ± 1.00	97.44 ± 1.17	**97.49 ± 0.06**
11	98.94 ± 0.97	98.87 ± 0.28	98.89 ± 0.24	99.24 ± 0.88	99.15 ± 0.61	**99.82 ± 0.17**
12	82.54 ± 29.04	**99.56 ± 0.09**	96.86 ± 0.07	97.45 ± 1.14	98.84 ± 0.89	98.13 ± 0.32
13	97.61 ± 2.96	97.78 ± 2.14	98.38 ± 0.00	98.6 ± 1.21	98.27 ± 0.99	**98.92 ± 1.00**
14	98.64 ± 0.41	99.24 ± 0.22	99.60 ± 0.16	**99.82 ± 0.12**	99.75 ± 0.23	99.74 ± 0.13
15	**99.78 ± 0.20**	98.19 ± 1.11	99.50 ± 0.36	99.77 ± 0.16	99.42 ± 0.73	97.7 ± 0.90
16	0.00 ± 0.00	95.86 ± 2.96	91.67 ± 0.00	90.00 ± 11.82	90.95 ± 8.54	**98.96 ± 0.96**
Running time(s)	836.15	923.48	**94.6**	114.37	139.93	152.11

Table 4. Accuracy comparison of different algorithms in Salinas datasets.

Method	ACGAN	ResNet	DgcForest	EMP	EMP + Saliency	mfdForest
OA	92.12 ± 1.42	97.84 ± 0.12	98.71 ± 0.11	99.18 ± 0.22	99.43 ± 0.28	**99.62 ± 0.21**
AA	86.59 ± 2.96	98.63 ± 0.27	98.63 ± 0.09	99.18 ± 0.28	99.43 ± 0.49	**99.62 ± 0.32**
K*100	86.59 ± 2.96	97.59 ± 0.13	98.68 ± 0.12	99.09 ± 0.33	99.36 ± 0.24	**99.57 ± 0.15**
1	91.18 ± 1.59	100.0 ± 0.00	98.96 ± 0.12	99.8 ± 0.12	98.53 ± 0.15	**100.00 ± 0.00**
2	44.71 ± 37.94	99.94 ± 0.05	**99.99 ± 0.03**	99.78 ± 0.6	99.38 ± 0.31	99.87 ± 0.13
3	96.00 ± 4.38	99.22 ± 0.01	98.60 ± 0.61	99.3 ± 1.39	98.8 ± 1.01	**100.00 ± 0.00**
4	95.02 ± 2.70	99.22 ± 0.01	97.86 ± 0.14	99.25 ± 0.19	98.57 ± 0.17	**99.28 ± 0.53**
5	32.29 ± 39.59	99.17 ± 0.22	**99.87 ± 0.09**	97.9 ± 1.96	99.44 ± 0.26	99.48 ± 0.49
6	94.38 ± 1.62	99.78 ± 0.07	**100.00 ± 0.00**	99.32 ± 0.25	99.42 ± 0.4	99.87 ± 0.12
7	95.15 ± 3.89	**99.98 ± 0.03**	95.98 ± 0.03	99.83 ± 0.11	99.83 ± 0.14	99.89 ± 0.01
8	97.58 ± 2.02	**99.77 ± 0.27**	98.55 ± 0.15	98.92 ± 0.12	99.52 ± 0.1	99.36 ± 0.19
9	97.16 ± 1.68	96.63 ± 0.21	97.99 ± 0.03	99.51 ± 0.29	**99.87 ± 0.22**	99.65 ± 0.27
10	98.06 ± 1.25	99.73 ± 0.05	98.88 ± 0.13	98.57 ± 1.08	99.69 ± 0.13	**100.00 ± 0.00**
11	98.61 ± 0.96	99.15 ± 0.27	97.77 ± 0.35	99.62 ± 0.2	99.25 ± 0.44	**100.00 ± 0.00**
12	94.33 ± 6.11	96.24 ± 2.09	97.74 ± 0.34	99.69 ± 0.27	**100.00 ± 0.00**	99.84 ± 0.13
13	99.22 ± 0.70	98.84 ± 1.13	98.71 ± 0.29	**100.00 ± 0.0**	**100.00 ± 0.00**	98.68 ± 0.21
14	90.50 ± 4.79	**99.16 ± 0.91**	98.50 ± 0.71	97.59 ± 3.28	98.33 ± 1.37	98.61 ± 0.05
15	73.42 ± 36.72	91.72 ± 0.42	99.07 ± 0.63	99.04 ± 0.35	99.12 ± 0.28	**99.56 ± 0.23**
16	94.21 ± 5.26	**100.0 ± 0.00**	97.92 ± 0.12	99.28 ± 0.16	99.61 ± 0.31	99.34 ± 0.32
Running time(s)	1020.11	1861.4	368.9	437.21	472.98	531.29

4   Conclusion

In this paper, we introduce a multi-feature fusion based deep forest to extract and fuse three feature maps of the HSI for improving classification accuracy. For a certain pixel, even the classification improved by certain feature map is not good, but one of the other two features may have a better effect, so it can work together for each pixel to get the best results. Although we use three branches to increase the size of the model, which result in the running time longer than dgcForest. This problem can be solved by parallel computing, and we will set three branches of mfdForest in parallel in the future work.

References

1. Tang, Y., et al.: Apple bruise grading using piecewise nonlinear curve fitting for hyperspectral imaging data. IEEE Access **8**, 147494–147506 (2020)
2. Shimoni, M., Haelterman, R., Perneel, C.: Hypersectral imaging for military and security applications: combining myriad processing and sensing techniques. IEEE Geosci. Remote Sens. Mag. **7**(2), 101–117 (2019)

3. Pike, R., Lu, G., Wang, D., Chen, Z.G., Fei, B.: A Minimum spanning forest-based method for noninvasive cancer detection with hyperspectral imaging. IEEE Trans. Biomed. Eng. **63**(3), 653–663 (2016)

4. Zhang, L., Zhang, L., Du, B.: Deep learning for remote sensing data: a technical tutorial on the state of the art. IEEE Geosci. Remote Sens. **4**(2), 22–40 (2016)

5. Yang, C., Li, Y., Peng, B., Cheng Y., Tong, L.: Road material information extraction based on multi-feature fusion of remote sensing image. In: 2019 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). pp. 3943–3946, IEEE (2019)

6. Lu, J., Ma, C., Zhou, Y., Luo, M., Zhang, K.: Multi-feature fusion for enhancing image similarity learning. IEEE Access **7**, 167547–167556 (2019)

7. Zhao, S., Nie, W., Zhang, B.: Multi-feature fusion using collaborative residual for hyperspectral palmprint recognition. In: 2018 IEEE 4th International Conference on Computer and Communications (ICCC), Chengdu, China, pp. 1402–1406, IEEE (2018)

8. He, M., Li, B., Chen, H.: Multi-scale 3D deep convolutional neural network for hyperspectral image classification. In: 2017 IEEE International Conference on Image Processing (ICIP). pp. 3904–3908. IEEE (2018)

9. Gu, Y., Liu, T., Jia, X., et al.: Nonlinear multiple kernel learning with multiple-structure-element extended morphological profiles for hyperspectral image classification. IEEE Trans. Geosci. Remote Sens. **54**(6), 3235–3247 (2016)

10. Jia, S., Lin, Z., Deng, B., et al.: Cascade superpixel regularized gabor feature fusion for hyperspectral image classification. IEEE Trans. Neural Netw. Learn. Syst. **31**(5), 1638–1652 (2019)

11. Jia, S., Deng, B., Zhu, J., et al.: Local binary pattern-based hyperspectral image classification with superpixel guidance. IEEE Trans. Geosci. Remote Sens. **56**(2), 749–759 (2017)

12. He, Z., Li, J., Liu, K., et al.: Kernel low-rank multitask learning in variational mode decomposition domain for multi-hyperspectral classification. IEEE Trans. Geosci. Remote Sens. **56**(7), 4193–4208 (2018)

13. Li, F., Wang, J., Lan, R., et al.: Hyperspectral image classification using multi-feature fusion. Opt. Laser Technol. **110**, 176–183 (2019)

14. Zhang, C., Han, M., Xu, M.: Multi-feature classification of hyperspectral image via probabilistic SVM and guided filter. In: 2018 International Joint Conference on Neural Networks (IJCNN). pp. 1–7. IEEE (2018)

15. Liu, X., Yin, X., Cai, Y., et al.: Visual saliency-based extended morphological profiles for unsupervised feature learning of hyperspectral images. IEEE Geosci. Remote Sens. Lett. **17**(11), 1963–1967 (2019)

16. Zhou, Z., Feng, J.: Deep forest: towards an alternative to deep neural networks. In: 2017 International Joint Conference on Artificial Intelligence Organization (IGCAI). pp. 3553–3559

17. Yin, X., Wang, R. et al.: Deep forest-based classification of hyperspectral images. In: Chinese Control Conference, pp. 10367–10372, IEEE (2018)

18. Liu, X., Wang, R., Cai, Z., et al.: Deep multigrained cascade forest for hyperspectral image classification. IEEE Trans. Geosci. Remote Sens. **57**(10), 8169–8183 (2019). https://doi.org/10.1109/TGRS.2019.2918587

19. Luo, H., Tang, Y.Y., Yang, X., et al.: Autoencoder with extended morphological profile for hyperspectral image classification. In: 2017 IEEE International Conference on Cybernetics (CYBCONF). pp. 293–296. IEEE (2017)

20. Zhang, J., Sclaroff, S.: Saliency detection: a Boolean map approach. In: 2013 IEEE International Conference on Computer Vision (ICCV). pp. 153–160. IEEE (2013)

21. Guo, G., Neagu, D., Huang, X., Bi, Y.: An effective combination of multiple classifiers for toxicity prediction. In: Wang, L., Jiao, L., Shi, G., Li, X., Liu, J. (eds.) Fuzzy Systems and

Knowledge Discovery, pp. 481–490. Springer Berlin Heidelberg, Berlin, Heidelberg (2006). https://doi.org/10.1007/11881599_56

22. Zhu, L., Chen, Y., Ghamisi, P., Benediktsson, J.A.: Generative adversarial networks for hyperspectral image classification. IEEE Trans. Geosci. Remote Sens. **56**(9), 5046–5063 (2018)

23. Zhong, Z., Li, J., Luo, Z., Chapman, M.: Spectral-spatial residual network for hyperspectral image classification: A 3-D deep learning framework. IEEE Trans. Geosci. Remote Sens. **56**(2), 847–858 (2018)