

Multiple Feature-Based Superpixel-Level Decision Fusion for Hyperspectral and LiDAR Data Classification

Sen Jia^{ID}, *Senior Member, IEEE*, Zhangwei Zhan, Meng Zhang, Meng Xu, *Member, IEEE*, Qiang Huang, Jun Zhou^{ID}, *Senior Member, IEEE*, and Xiuping Jia^{ID}, *Senior Member, IEEE*

Abstract—The rapid increase in the number of remote sensing sensors makes it possible to develop multisource feature extraction and fusion techniques to improve the classification accuracy of surface materials. It has been reported that light detection and ranging (LiDAR) data can contribute complementary information to hyperspectral images (HSIs). In this article, a multiple feature-based superpixel-level decision fusion (MFSuDF) method is proposed for HSIs and LiDAR data classification. Specifically, superpixel-guided kernel principal component analysis (KPCA) is first designed and applied to HSIs to both reduce the dimensions and compress the noise impact. Next, 2-D and 3-D Gabor filters are, respectively, employed on the KPCA-reduced HSIs and LiDAR data to obtain discriminative Gabor features, and the magnitude and phase information are both taken into account. Three different modules, including the raw data-based feature cube (concatenated KPCA-reduced HSIs and LiDAR data), the Gabor magnitude feature cube, and the Gabor phase feature cube (concatenation of the corresponding Gabor features extracted from the KPCA-reduced HSIs and LiDAR data), can be, thus, achieved. After that, random forest (RF) classifier and quadrant bit coding (QBC) are introduced to separately accomplish the classification task on the aforementioned three extracted feature cubes. Alternatively, two superpixel maps are generated by utilizing the multichannel simple noniterative clustering (SNIC) and entropy rate superpixel segmentation (ERS) algorithms on the combined HSIs and LiDAR data, which are

then used to regularize the three classification maps. Finally, a weighted majority voting-based decision fusion strategy is incorporated to effectively enhance the joint use of the multisource data. The proposed approach is, thus, named MFSuDF. A series of experiments are conducted on three real-world data sets to demonstrate the effectiveness of the proposed MFSuDF approach. The experimental results show that our MFSuDF can achieve the overall accuracy of 73.64%, 93.88%, and 74.11% for Houston, Trento, and Missouri University and University of Florida (MUFL) Gulport data sets, respectively, when there are only three samples per class for training.

Index Terms—Feature extraction, feature fusion, hyperspectral image (HSI), light detection and ranging (LiDAR), superpixel segmentation.

I. INTRODUCTION

RECENTLY, with the rapid development of remote sensing techniques, it has become feasible to capture multisource images for the same surveyed scene [1]. Concretely, due to technical and cost limitations, each type of remote sensing sensor is designed to operate at given wavelengths, which typically has a few observation purposes and obtains only specific information. By using diverse sensors in the same zone, multiple types of images with different properties can be acquired. The advanced technology makes it possible to collect various aspects of data for the same surveyed area on the earth [2]. The representative and important information includes spectral and spatial information of hyperspectral images (HSIs), height and shape information of light detection and ranging (LiDAR) data, and texture information of synthetic aperture radar (SAR).

In particular, among numerous sources of imagery, HSIs acquired by airborne sensors, with wavelengths that cover visible and near-infrared channels, contain reflected spectra with hundreds/thousands of bands for each pixel on the earth's surface [3]. Compared with other traditional remote sensing data, HSIs with narrower spectral bands and more imaging channels can greatly improve the ability to distinguish objects. Especially, HSIs have piecewise smoothness in both the spectral and spatial domains, leading to highly correlated relationships between neighboring pixels. For the purpose of classifying HSIs, a large number of classification methods have been proposed, including spectral band analysis [4]–[6], support vector machines (SVMs) [7], [8], sparse representation-based classification (SRC) [9]–[11], and random forest (RF) classification [12]–[14].

However, the rich details contained in both the spatial and spectral domains of HSIs not only bring opportunities

Manuscript received December 3, 2019; revised April 30, 2020; accepted May 14, 2020. Date of publication June 5, 2020; date of current version January 21, 2021. This work was supported in part by the National Natural Science Foundation of China under Grant 41971300, Grant 61671307, and Grant 61901278, in part by the Program for Young Changjiang Scholars, in part by the Research and Development Plan in Key Areas of Guangdong Province under Grant 2019B010151001, and in part by the Shenzhen Scientific Research and Development Funding Program under Grant JCYJ20180305124802421 and Grant JCYJ20180305125902403. (Corresponding author: Qiang Huang.)

Sen Jia, Zhangwei Zhan, Meng Zhang, Meng Xu, and Qiang Huang are with the Guangdong Laboratory of Artificial Intelligence and Digital Economy (SZ), Shenzhen University, Shenzhen 518060, China, also with the SZU Branch, Shenzhen Institute of Artificial Intelligence and Robotics for Society, Shenzhen 518060, China, also with the Guangdong Key Laboratory of Urban Informatics, Shenzhen University, Shenzhen 518060, China, and also with the College of Computer Science and Software Engineering, Shenzhen University, Shenzhen 518060, China (e-mail: senjia@szu.edu.cn; 1800271020@email.szu.edu.cn; zhangmeng2016@email.szu.edu.cn; m.xu@szu.edu.cn; jameshq@szu.edu.cn).

Jun Zhou is with the School of Information and Communication Technology, Griffith University, Nathan, QLD 4111, Australia (e-mail: jun.zhou@griffith.edu.au).

Xiuping Jia is with the School of Engineering and Information Technology, The University of New South Wales, Canberra, ACT 2612, Australia (e-mail: x.jia@adfa.edu.au).

Color versions of one or more of the figures in this article are available online at <https://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TGRS.2020.2996599

0196-2892 © 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.

See <https://www.ieee.org/publications/rights/index.html> for more information.

to improve material classification accuracies but also pose a series of challenges in this field [15], [16]. On the one hand, raw HSIs have a large amount of redundant and correlated spectral information. When the number of training samples is fixed, the continued growth of the spectral dimension of HSIs will deteriorate the classification accuracy. This problem is the so-called Hughes phenomenon [17], [18]. Furthermore, in the process of data acquisition, it has been shown that hyperspectral remote sensing images are easily affected by clouds and cloud shadows [19]. Due to these shortcomings, the classifiers that use only the spectral characteristics of training samples inevitably neglect the important spatial clustering distribution information between labeled and unlabeled pixels, which leads to the fact that the classification accuracy is always plagued by the small sample size problem and cannot be greatly improved in practice.

For the abovementioned reasons, taking advantage of multisource data, i.e., images captured from different sensors, can help address the small sample size problem. In particular, the fusion of HSIs and LiDAR data was demonstrated to be beneficial for providing higher classification accuracy than the use of each source individually [20]. Generally, the LiDAR data can contribute potential details for HSIs, such as the height and shape information of land-cover objects, which can improve the HSI classification accuracy [21]. When objects in HSIs are covered by clouds, the LiDAR data can provide supplementary information for the cloudy regions. More importantly, the LiDAR data can indicate the spatial information of the areas in HSIs, which leads to the need for HSIs and LiDAR fusion.

In the literature, many techniques for HSIs and LiDAR data fusion have been developed for the purpose of classification [22]–[27]. Concretely, HSIs and LiDAR data fusion methods can be categorized into pixel-based, feature-based, and decision-based. The pixel-based fusion methods combine information of different images on a pixel-by-pixel basis. Pixelwise details are merged from multisource data, but they are easily affected by noise. The strategy of feature-based fusion methods is to integrate extracted features from different images [28], [29]. Specifically, features extracted from multisource data are combined to create feature vectors for classification. The accuracy of target recognition after feature-based fusion is obviously higher [30], [31]. However, these methods extract image features as fusion information, so many details can be lost. For decision-based fusion methods, different data sources are combined in a higher level of integration [32], [33]. Image data from various individual sensors are first categorized and then merged to incorporate the output of the categorization. In general, the computational complexity of the decision-based fusion methods is minimal, but it is strongly dependent on the categorization results. In summary, the abovementioned methods all have limitations; thus, developing an approach to alleviate these disadvantages is encouraged.

Since image data from multiple platforms carry distinct information, it is important to fuse the features appropriately to improve the classification performance of materials on the surface and consequently increase the quality of the mapping.

For this reason, developing an effective method to extract features from images and establishing multisensor data fusion techniques to further improve the discrimination ability of classifiers is essential [34]. Raw HSIs contain a massive amount of redundant information, which can degrade classification performance. Therefore, an appropriate dimensionality reduction approach must be used to reduce redundancy [35]. To date, there are a variety of dimensional reduction methods, such as low-variance filters, Fisher's linear discriminant analysis (FDA) [36], independent component analysis (ICA) [37], isometric embeddings (ISOMAP) [38], locally linear embedding (LLE) [39], and principal component analysis (PCA) [40]. Among them, kernel principal component analysis (KPCA), as a generalization of PCA, has been proven to be a more effective dimensionality reduction approach for nonlinear data, such as multisource remote sensing data [41], [42]. However, it is unrealistic to directly apply KPCA on HSIs since the computational burden of the covariance matrix is totally unaffordable. Besides, 2-D morphological feature extraction methods have been exploited for HSIs and LiDAR data classification. Nevertheless, HSIs are regarded as 3-D data cubes with rich spectral and spatial information. Therefore, the extracted 2-D morphological features cannot sufficiently represent the joint spectral–spatial structure [43]. Furthermore, Gabor wavelets are powerful tools that can achieve optimal feature representation. Numerous tools have been successfully applied to extract features for face recognition, medical imaging, and so on. It has been found that the Gabor features can present more discriminative image features than other features in these applications [44], [45].

In recent years, with the growing maturity and wide application of superpixel segmentation algorithms in computer vision tasks, the interconnection between superpixel extraction and HSI classification has gradually become an active research field [46]–[48]. The well-known explanation for superpixels is to group together some pixels with similar characteristics to form a more representative, larger element [49]. This new element will be used as the basic unit of other image processing algorithms [50]. More precisely, superpixel segmentation is a process of tagging each pixel in an image. The advantage of superpixel segmentation is that it greatly reduces the dimensions of decision fusion by simultaneously removing abnormal pixels [51].

This article is organized as follows. Section II provides an overview of the proposed multiple feature-based superpixel-level decision fusion (MFSuDF) approach. In Section III, the background of related works, including the Gabor wavelets and superpixel algorithms, is introduced. Section IV describes the proposed methodological framework in detail. Section V presents the experimental setup and the result analysis on three real data sets. Finally, Section VI provides the conclusion and future work for this article.

II. OVERVIEW OF MFSUDF

In this article, an MFSuDF framework that takes full advantage of feature discriminability of different modules and superpixel structure is proposed for HSIs and LiDAR data classification. First, a superpixel-guided KPCA is presented

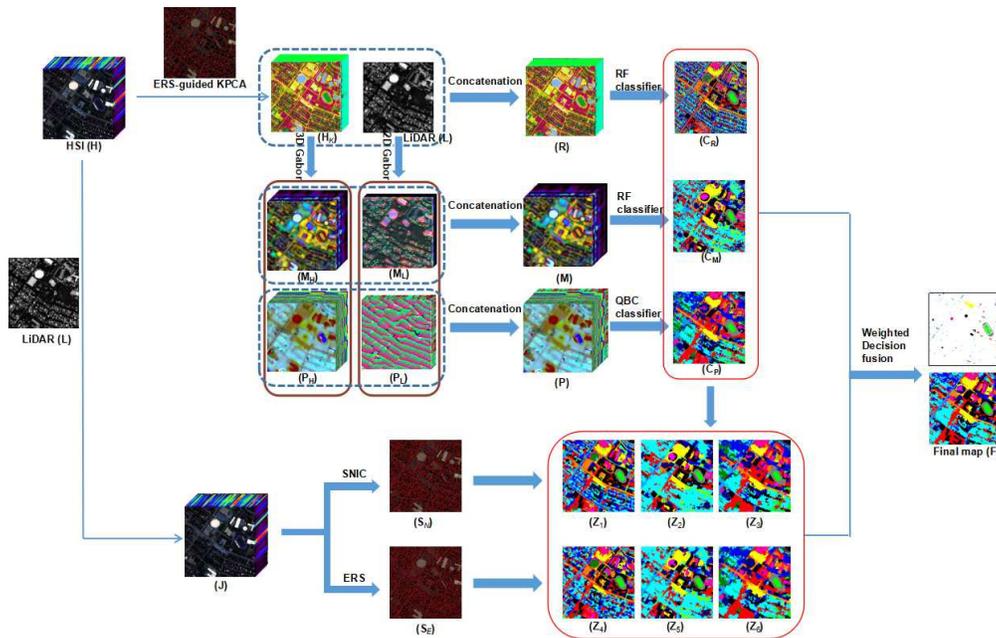


Fig. 1. Flowchart of the proposed MFSuDF method for hyperspectral and LiDAR data classification.

and utilized to reduce the dimensions of HSIs, which optimizes the computing resources and improves the speed of subsequent data processing. Moreover, KPCA with a superpixel guided manner can make the information-focused and irrelevant factors ignorable. Second, the 3-D Gabor filters with different scales and orientations are convolved with the information-focused HSIs features, and correspondingly, the 2-D Gabor wavelets are applied to LiDAR data that possess only a single band. To make full use of the obtained information, the Gabor texture feature is further divided into two parts according to magnitude and phase. Therefore, features relating to three aspects of HSIs and LiDAR data are obtained. Subsequently, the HSIs and LiDAR features of each aspect are concatenated as raw data-based features (i.e., the KPCA-reduced HSIs and LiDAR data), the Gabor magnitude features, and the Gabor phase features. As an effective classifier that could well handle the imbalanced circumstance of the training set, RF is employed on the previous two energy-based features.

Alternatively, the Gabor phase feature would cause additional classification errors when using conventional classification methods due to its instability. Quadrant bit coding (QBC) and Hamming distance metrics provide a novel sample similarity calculation for the Gabor phase. Ultimately, three separate classification maps can be, thus, obtained. In the other part, two algorithms that can control superpixel size, simple noniterative clustering (SNIC), and entropy rate superpixel segmentation (ERS) are extended and adopted to extract superpixel maps by jointly utilizing the HSIs and LiDAR data. Through independently applying the two superpixel maps on the three classification maps, six regularization maps are obtained. Finally, in order to take full advantage of the complementarity among various maps, both the regularization maps and pixelwise classification maps are addressed through a weighted fusion procedure, and thus, the final classification

can be achieved. The proposed MFSuDF method is illustrated in Fig. 1. More precisely, the contributions of our method are listed as follows.

- 1) First, different from previous methods that only considers partial aspects of the Gabor description, the Gabor wavelet-based feature representation is thoroughly exploited in our proposed MFSuDF method. On one hand, the 3-D Gabor filters, which are perfectly in line with the joint spatial-spectral characteristics of HSIs, is adopted. Meanwhile, both 3-D Gabor magnitude and phase features, which, respectively, represent the intensity and variation in the local region, are employed. On the other hand, the LiDAR data containing elevation information have only one band, which could be well characterized by the 2-D Gabor filters, and the corresponding 2-D Gabor magnitude and phase features are taken into account. Evidently, due to the high representative capability of the Gabor-based features and complementary properties of HSIs and LiDAR data, it is very desirable to exploit the extracted features in a unified framework. Therefore, all the four Gabor-based feature modules, including both 3-D and 2-D Gabor magnitude and phase features, are collaboratively investigated, and the classification accuracy can be significantly improved compared with the 2-D Gabor on HSIs and LiDAR band by band or the 3-D Gabor on HSIs and LiDAR fused data.
- 2) Second, the superpixel regularization procedure is specially optimized. In our method, two superpixel maps generated by different strategies (i.e., graph-based ERS and cluster-based SNIC) are incorporated for regularization rather than a single map. Indeed, the number of homogeneous regions in the superpixel map plays a vital role in the regularization procedure, and the

proposed two superpixel map-based regularizing scheme can essentially enhance the resistance to oversegmentation or undersegmentation in superpixel segmentation. Furthermore, the idea of multitask learning is sufficiently exploited. Through separately applying the two superpixel maps on the classification maps achieved from three kinds of features, two of which are Gabor magnitude and phase, while the other one is the combination of KPCA-reduced HSIs and LiDAR data. Confidently, the complementary information contained in the three pixelwise classification maps and six regularized maps is concentrated in the final map through a majority-voting strategy, and the robustness of the proposed approach can be guaranteed.

- 3) Third, a superpixel-guided sample extraction method is proposed to optimize the KPCA calculation process instead of random selection. Due to the high computational requirement of KPCA, it is infeasible to directly apply KPCA on the whole HSIs, and a small number of representative samples should be picked out to reduce calculation. Fortunately, since the homogenous attribute of each superpixel could be well characterized by a superpixel map (here, the ERS method is selected), the superpixel guidance manner can not only guarantee the availability of KPCA computation but also significantly reduce the amount of calculation. Alternatively, it is worth mentioning that our framework is an unsupervised feature manipulation method, which means that prior knowledge is not required in feature extraction and superpixel segmentation, and the adaptability of our method can be ensured. Besides, our framework erodes the influence of artificial factors on the classification results (through multitasking decision fusion), which also reduces the labor and time costs of object identification.

To confirm the effectiveness of our framework, a series of related experiments are designed and conducted. The training samples are selected randomly, and the test samples are derived from the rest of the image. To reduce the risk raised by sample selection, each experiment is run 20 times. The results of three real-data experiments show that the MFSuDF method achieves better performance than the other parallel methods.

III. BACKGROUND

This section introduces the background of the KPCA, Gabor wavelets, and superpixel segmentation algorithm.

A. Kernel Principal Component Analysis (KPCA)

KPCA is an efficient dimensionality reduction approach that has been widely applied to obtain low-dimensional mapping features for high-dimensional data sets [52]. Comparatively, as a nonlinear generalization of PCA, KPCA is proposed to solve the complex structure of the original data. It uses kernel functions to map known low-dimensional data to high-dimensional data and then conducts PCA in high dimensions. The computational process of the whole high-dimensional space involves the inner product, which can be computed by

using the kernel function in the original space. Concretely, for a given nonlinear mapping, the input data $\mathbf{h} = (\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_n)$ (where n represents the number of samples) can be mapped into a higher or infinite dimensional Hilbert space. Assuming that new features are centered and denoted by $\Phi(\mathbf{h}) = (\Phi(\mathbf{h}_1), \Phi(\mathbf{h}_2), \dots, \Phi(\mathbf{h}_n))$, the covariance matrix is defined in the high-dimensional feature space as follows:

$$\bar{\mathbf{C}} = \frac{1}{n} \sum_{i=1}^n \Phi(\mathbf{h}_i) \Phi(\mathbf{h}_i)^T. \quad (1)$$

The eigenvector and eigenvalues can be computed by

$$\bar{\mathbf{C}}\boldsymbol{\beta} = \boldsymbol{\lambda}\boldsymbol{\beta} \quad (2)$$

where $\boldsymbol{\lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ is a diagonal matrix, and $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_n)$ is an $n \times n$ matrix composed of eigenvectors. $\boldsymbol{\beta}$ can be also expressed by

$$\begin{aligned} \boldsymbol{\beta} &= (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_n) \\ &= (\Phi(\mathbf{h})\boldsymbol{\alpha}_1, \Phi(\mathbf{h})\boldsymbol{\alpha}_2, \dots, \Phi(\mathbf{h})\boldsymbol{\alpha}_n) = \Phi(\mathbf{h})\boldsymbol{\alpha} \end{aligned} \quad (3)$$

where $\boldsymbol{\alpha} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \dots, \boldsymbol{\alpha}_n)$ is an $n \times n$ coefficients matrix. Then, by considering the equation

$$\Phi(\mathbf{h})^T \bar{\mathbf{C}}\boldsymbol{\beta} = \Phi(\mathbf{h})^T \boldsymbol{\lambda}\boldsymbol{\beta} \quad (4)$$

and substituting (1) and (3) into (4), we can obtain the following equation:

$$n\boldsymbol{\lambda}\boldsymbol{\alpha} = \mathbf{K}\boldsymbol{\alpha} \quad (5)$$

where \mathbf{K} is the kernel matrix and $\mathbf{K}_{i,j} = \Phi(\mathbf{h}_i)^T \Phi(\mathbf{h}_j)$. Just like PCA, KPCA has similar steps to conduct feature extraction and dimensionality reduction on the kernel matrix.

B. Gabor Wavelets

The 2-D Gabor wavelets of the spatial domain are based on the theory of a 1-D Gabor transform, which was first proposed by Dennis Gabor in 1946 [53]. As an extension of 1-D Gabor wavelets, the 2-D Gabor function is defined as the sinusoidal wave (a kind of plane wave for 2-D Gabor filters) multiplied by a Gaussian function [54]. The characteristics of the 2-D Gabor function are similar to those of the human visual system and have better abilities to represent local structural information. Meanwhile, the 2-D Gabor wavelets formed by Gabor functions have the advantages of optimal localization in both the spatial and frequency domains. Therefore, it could be effective in extracting texture feature representations. More precisely, in the spatial domain (x, y) , the 2-D Gabor wavelets are defined as follows:

$$\begin{aligned} \Psi_{f,\theta}(x, y) &= \frac{1}{2\pi\sigma^2} \exp(j2\pi(xf \cos \theta + yf \sin \theta)) \\ &\quad \times \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \end{aligned} \quad (6)$$

where f and θ represent the central frequency and orientation, respectively. σ is the width of the Gaussian envelope function in the spatial domain (x, y) . Generally, with different scales and orientations, a group of Gabor wavelets is predefined to extract discriminative features for LiDAR data analysis, which is just in a 2-D form.

Similar to the derivation procedure of 2-D Gabor wavelets, the 3-D Gabor wavelets are the 3-D extension of the corresponding 1-D wavelets [55]. Clearly, it can capture discriminative information on local areas in joint spatial–spectral correlations. Suppose that there is a given location (x, y, b) in the 3-D domains, and the 3-D Gabor wavelets can be defined by [56]

$$\begin{aligned} \hat{\Psi}_{\rho, \vartheta, \varphi}(x, y, b) &= \frac{1}{(2\pi)^{2/3} \tau^3} \exp(j2\pi(xu + yv + bw)) \\ &\quad \times \exp\left(-\frac{x^2 + y^2 + b^2}{2\tau^2}\right) \\ u &= \rho \sin \varphi \cos \vartheta, v = \rho \sin \varphi \sin \vartheta, w = \rho \cos \varphi \end{aligned} \quad (7)$$

where ρ and (ϑ, φ) , respectively, denote the central frequency of the sinusoidal wave and the orientation of the 3-D Gabor wavelet in the frequency domain (u, v, w) . Similarly, τ is the width of the Gaussian envelope. Finally, by changing the frequency and orientation, a set of 3-D Gabor wavelets is obtained, which can be utilized to extract features from HSIs with any data transformation.

The extracted Gabor feature can be divided into the magnitude and phase parts. Generally speaking, the magnitude feature is stable and smooth, which actually reflects the local energy spectrum of the image. In fact, the magnitude feature can also be understood as the intensity of the edge in a specific direction, which is of great significance to the matching and recognition of the image. Alternatively, the phase feature is very unstable, which changes sharply even within a short distance. Fortunately, recent advances have revealed that a proper encoding scheme could make good use of the Gabor phase feature.

C. Superpixel Segmentation

ERS segmentation is an efficient and greedy segmentation algorithm for obtaining superpixels of image. Compared with the NCut and watershed superpixel algorithms, ERS possesses both a relatively fast computational speed and satisfactory superpixel segmentation. For a given undirected graph $G = (V, E)$, where V denotes the set of vertices and E includes the set of edges, it is expected that the resulting graph $G' = (V, A)$ consists of T connected subgraph, where $A \subseteq E$ and T is the number of superpixels. Suppose that $w_{i,j}$ denotes the weight of the edge $e_{i,j}$ between points v_i and v_j , and then, the transition probabilities $p_{i,j}$ and the weight of the i th vertex μ_i are, respectively, defined as

$$\begin{aligned} w_i &= \sum_{v_j \in V} w_{i,j}; \quad \mu_i = \frac{w_i}{\sum_{v_j \in V} w_j} \\ p_{i,j}(A) &= \begin{cases} \frac{w_{i,j}}{w_i} & \text{if } i \neq j \text{ and } e_{i,j} \in A \\ 0 & \text{if } i \neq j \text{ and } e_{i,j} \notin A \\ 1 - \frac{\sum_{j: e_{i,j} \in A} w_{i,j}}{w_i} & \text{if } i = j. \end{cases} \end{aligned} \quad (8)$$

Thus, the entropy rate with a random walk $\mathcal{H}(A)$ that determines the cluster compact and homogeneous can be

calculated by

$$\mathcal{H}(A) = - \sum_i \mu_i \sum_j p_{i,j}(A) \log(p_{i,j}(A)). \quad (9)$$

In contrast, let $S_A = \{S_1, S_2, \dots, S_{N_A}\}$ represent the intermediate process, where N_A is the number of connected components, and then, the distribution entropy $H(A)$ is computed by

$$H(A) = - \sum_i p_A(i) \log(p_A(i)) \quad (10)$$

where

$$p_A(i) = \frac{S_i}{|V|}, \quad i = 1, 2, \dots, N_A. \quad (11)$$

Therefore, the balancing term that controls the number of clusters and similar size between clusters can be defined as

$$\mathcal{B}(A) = H(A) - N_A. \quad (12)$$

The ERS algorithm proposes an objective function to linearly associate the entropy rate with a random walk and a balancing term, which is defined by

$$\begin{aligned} \max_A \quad & \mathcal{H}(A) + \gamma \mathcal{B}(A) \\ \text{s.t.} \quad & A \subseteq E \text{ and } N_A \geq T \end{aligned} \quad (13)$$

where $\gamma \geq 0$ is the weight of the balancing term. Although the entropy rate with a random walk and the balancing term is increasing when adding any one edge into a set A , there is no doubt that the compact construction relies on the maximization of the objective function. It is worth mentioning that the value of γ is automatically adjusted based on the number of superpixels.

Alternatively, the SNIC algorithm is an improvement of the simple linear iterative clustering (SLIC) algorithm. It shows great potential in superpixel segmentation because of its low computational complexity and good segmentation results [57]. The definition of distance is the same as that in the SLIC algorithm. The input image consists of the a device-independent color space (CIELAB) color space \mathbf{c} and spatial coordinates \mathbf{x} , and the distance between pixel i and j is calculated by

$$d_{i,j} = \sqrt{\omega_1 \|\mathbf{c}_i - \mathbf{c}_j\|^2 + \omega_2 \|\mathbf{x}_i - \mathbf{x}_j\|^2} \quad (14)$$

where ω_1 and ω_2 represent the weight of the color space and coordinate space, respectively. Similar to SLIC, the SNIC algorithm starts with centroid initialization, and a priority queue Q is used to evaluate the distance between pixels and centroids. Specifically, it begins with adding centroids into the priority queue, and the process of selecting the element that has the smallest distance, updating the value of the centroid, labeling element adjacent to the current pixel, and pushing it into the priority queue is continued until Q is empty.

IV. PROPOSED MFSuDF METHODOLOGY

As shown in Fig. 1, our MFSuDF methodology has three main steps. In the first part, the Gabor wavelets are used to extract representative local texture information from superpixel-guided KPCA features of HSIs and LiDAR data, respectively. Subsequently, classifiers are applied to the Gabor

TABLE I

DEFINITIONS OF THE MATHEMATICAL SYMBOLS USED IN THIS ARTICLE

Symbol	Meaning
\mathbf{H}, \mathbf{L}	HSIs and LiDAR data
X, Y, B	dimensions of the HSIs or LiDAR data
$\mathbf{G}_H, \mathbf{P}_H, \mathbf{M}_H$	3D Gabor, phase and magnitude features of HSIs data
$\hat{\Psi}, \rho, \vartheta, \varphi$	the parameter of the 3D Gabor filter
$\mathbf{G}_L, \mathbf{P}_L, \mathbf{M}_L$	2D Gabor, phase and magnitude features of LiDAR data
Ψ, f, θ	the parameter of 2D Gabor filter
$\mathbf{R}, \mathbf{P}, \mathbf{M}$	concatenation of KPCA, Gabor phase and magnitude
\mathbf{J}, \mathbf{Q}	concatenation of raw data and code matrix of \mathbf{P}
$\mathbf{C}_R, \mathbf{C}_P, \mathbf{C}_M$	classification maps from complementary features
$\mathbf{S}_E, \mathbf{S}_N$	superpixel maps of ERS and SNIC
\mathbf{Z}, \mathbf{F}	regularized maps and final map

(both magnitude and phase) and KPCA features to generate three separate classification maps. In the second part, the ERS and SNIC superpixel segmentation algorithms are, respectively, adopted to generate two superpixel maps from the integrated HSIs and LiDAR data. In the last part, a weighted fusion method is employed to effectively address the three classification maps and six corresponding superpixel regularized ones, and the complementary information inside each map can be sufficiently utilized. To facilitate the description of the following sections, we define some notations in Table I.

A. Generation of Three Classification Maps

1) *Superpixel-Guided KPCA for Dimension Reduction*: Due to the high correlation between neighboring bands of HSIs, the Gabor features extracted and directly data concatenation of raw HSIs and LiDAR data are high-dimensional and redundant. As a nonlinear dimensionality reduction method, KPCA is very suitable for hyperspectral data preprocessing. In particular, there are two methods to select samples for KPCA. When the data set is small, KPCA is generally performed on all data samples. Otherwise, a limited number of samples are randomly picked out from the large data set for feature extraction of KPCA. Clearly, the sample selection strategy should describe the characteristics of the whole data set as far as possible, which is a key factor for KPCA and the following manipulation. Concerning HSIs with a large number of samples, since superpixel segmentation could well reflect the spatial structure of objects, it is natural to introduce superpixel map to guide the sample selection of KPCA, and thus, a superpixel-guided sample selection method is proposed for the KPCA-based dimension reduction (here, the ERS method is utilized due to its efficiency).

Let $\mathbf{H} \in \mathbb{R}^{X \times Y \times B}$ be the raw HSIs, where X, Y, B , respectively, represents the spatial-spectral dimension. Obviously, it is infeasible to conduct all the samples for KPCA since the corresponding kernel matrix $\mathbf{K} \in \mathbb{R}^{XY \times XY}$ is extremely huge. Fortunately, in our superpixel-guided KPCA method, the ERS method is first applied on \mathbf{H} to get the superpixel map (here, the number of superpixel is simply set as $X + Y$). Then, the mean vector of each superpixel is computed, which are collected together to form the sample set $\mathbf{h} = (\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_{X+Y})$. Correspondingly, the size of kernel matrix \mathbf{K} (here, the Gaussian kernel

$\mathbf{K}_{i,j} = \Phi(\mathbf{h}_i)^T \Phi(\mathbf{h}_j) = \exp(-\|\mathbf{h}_i - \mathbf{h}_j\|^2/2)$ is adopted) is only $(X + Y) \times (X + Y)$, which not only greatly decrease the storage requirement and computational burden but also guarantee the representativeness of the sample set. In general, 99% energy contained in the principal components of KPCA are kept, and the resultant KPCA feature cube $\mathbf{H}_K \in \mathbb{R}^{X \times Y \times B_K}$ can be achieved, where B_K is the number of compressed dimension in the spectral domain and much smaller than B .

2) *Gabor Feature Extraction for HSIs and LiDAR Data*: Considering the distinct information between HSIs and LiDAR data, 3-D and 2-D Gabor wavelets are utilized at the same time.

In terms of 3-D data classification, it is desirable to utilize the joint spectral-spatial information of objects, which can present spatial regularities and spectral similarity in a scene [56]. In this part, the 3-D Gabor filters are applied on HSIs for feature extraction. Considering a Gabor wavelet $\hat{\Psi}_{\rho, \vartheta, \varphi}$ (7) with a center frequency ρ and orientation (ϑ, φ) , at location (x, y, b) , the signal variances $\hat{\mathbf{g}}(x, y, b)$ can be expressed as

$$\hat{\mathbf{g}}(x, y, b) = (\mathbf{H}_K \otimes \hat{\Psi}_{\rho, \vartheta, \varphi})(x, y, b) \quad (15)$$

where \otimes denotes the convolution operation. For the purpose of extracting the most discriminative spectral-spatial features from HSIs, only the wavelets with direction that is parallel to the spectral axis are kept, while others are excluded (the Gabor magnitude and phase cubes in this direction possess clear local texture feature and distinct homogeneous region, while the others are less informative) [45], and the parameters of 3-D Gabor filters are set as follows:

$$\begin{aligned} \rho &\in [0.5, 0.25, 0.125, 0.0625] \\ \vartheta &= 0, \quad \varphi = 0. \end{aligned} \quad (16)$$

The feature vectors of all objects, denoted by $\mathbf{G}_H \in \mathbb{R}^{X \times Y \times 4B_K}$, at location (x, y) can be generated by concatenating the complex of the convolutional calculation:

$$\mathbf{G}_H(x, y) = [\hat{\mathbf{g}}_1(x, y, 1 : B_K), \dots, \hat{\mathbf{g}}_4(x, y, 1 : B_K)]. \quad (17)$$

Alternatively, since the LiDAR-derived digital surface model (DSM) only has a single band, its corresponding texture feature of elevation information can be extracted by 2-D Gabor wavelets. Similarly, the 2-D Gabor features can be obtained by convolving the LiDAR data with 2-D Gabor filters (6) as follows:

$$\mathbf{g}(x, y) = (\mathbf{L} \otimes \Psi_{f, \theta})(x, y) \quad (18)$$

where $\mathbf{L} \in \mathbb{R}^{X \times Y}$ denotes the LiDAR data. In order to acquire sufficient discriminative information for classification, the two parameters, f and θ , are set as follows [58]:

$$\begin{aligned} f &\in [0.03589, 0.09473, 0.25, 0.6577] \\ \theta &\in [0, 40, 80, 120, 160, 180]. \end{aligned} \quad (19)$$

As a result, 24 filters with four scales and six orientations are concatenated to create a Gabor feature vector

$$\mathbf{G}_L(x, y) = [\mathbf{g}_1(x, y), \dots, \mathbf{g}_{24}(x, y)]. \quad (20)$$

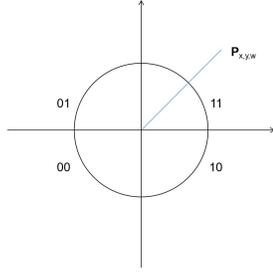


Fig. 2. Working principle of QBC.

After the convolution operator is applied on all samples, the 2-D Gabor feature of LiDAR data $\mathbf{G}_L \in \mathbb{R}^{X \times Y \times 24}$ can be achieved.

3) *Feature Fusion and Classification*: Sections IV-A1 and IV-A2, respectively, describe superpixel-guided KPCA and Gabor feature extraction. In this section, to make full use of the texture information produced by the Gabor wavelets, \mathbf{G}_H and \mathbf{G}_L are divided into two parts according to the phase angle and magnitude, called \mathbf{P}_H and \mathbf{M}_H , and \mathbf{P}_L and \mathbf{M}_L , respectively, [59]. Then, data with different types of information are categorized, and LiDAR features are concatenated with HSIs features. This process is expressed by the following formulation:

$$\begin{aligned} \mathbf{M}_H &= \sqrt{\text{Re}(\mathbf{G}_H)^2 + \text{Im}(\mathbf{G}_H)^2} \\ \mathbf{P}_H &= \arctan(\text{Re}(\mathbf{G}_H), \text{Im}(\mathbf{G}_H)) \\ \mathbf{M}_L &= \sqrt{\text{Re}(\mathbf{G}_L)^2 + \text{Im}(\mathbf{G}_L)^2} \\ \mathbf{P}_L &= \arctan(\text{Re}(\mathbf{G}_L), \text{Im}(\mathbf{G}_L)) \\ \mathbf{R} &= [\mathbf{H}_K, \mathbf{L}]; \mathbf{P} = [\mathbf{P}_H, \mathbf{P}_L]; \mathbf{M} = [\mathbf{M}_H, \mathbf{M}_L] \end{aligned} \quad (21)$$

where Re and Im stands for getting the real and imaginary parts, respectively, of complex numbers. It is worth to point out that except the phase $\mathbf{P} \in \mathbb{R}^{X \times Y \times (4B_K + 24)}$ and magnitude $\mathbf{M} \in \mathbb{R}^{X \times Y \times (4B_K + 24)}$ feature, the KPCA feature \mathbf{H}_K integrated with \mathbf{L} , which is denoted as $\mathbf{R} \in \mathbb{R}^{X \times Y \times (B_K + 1)}$, is also considered for further processing.

Due to the different characteristics of extracted features (the stability of \mathbf{R} and \mathbf{M} is clearly higher than \mathbf{P}), it is essential to select appropriate classifiers to discriminate these features. Concretely, the RF classifier has advantages in many aspects, such as resistance to overfitting and insensitivity to noise [60]. Compared with SVMs, the RF classifier not only has an advantage in computational simplicity and classification time but also achieves a similar accuracy [61]. It is made up of decision trees and bagging sampling techniques. The RF algorithm can be divided into two parts: randomly selecting repeatable training subsets and building multiple decision trees. In our experiments, the RF classifier is applied to \mathbf{R} , which reflects the spectral-spatial characteristics, as well as \mathbf{M} , which contains the Gabor magnitude information. The number of trees is set to 500, and the default subspace is the floor of the logarithmic value of the features.

Besides, QBC is a parameter-free unsupervised classification method [45], [62], which provides a feasible way to encode the discriminative information of the Gabor phase. The working principle of QBC is to encode phase information \mathbf{P}

on the basis of the quadrant that the phase angle locates in. For a Gabor phase feature vector $\mathbf{P}_{x,y,4B_K+24}$ in the spatial position (x, y) , one element $\mathbf{P}_{x,y,z}$ ($z = 1, 2, \dots, 4B_K + 24$) in the vector can be encoded by

$$\mathbf{Q}_{x,y,z} = \begin{cases} 11 & \text{if } \mathbf{P}_{x,y,z} \in \left(0, \frac{\pi}{2}\right] \\ 01 & \text{if } \mathbf{P}_{x,y,z} \in \left(\frac{\pi}{2}, \pi\right] \\ 00 & \text{if } \mathbf{P}_{x,y,z} \in \left(-\pi, -\frac{\pi}{2}\right] \\ 10 & \text{if } \mathbf{P}_{x,y,z} \in \left(-\frac{\pi}{2}, 0\right]. \end{cases} \quad (22)$$

Fig. 2 shows the coding rule of QBC. Subsequently, the Hamming distance D is used to measure the similarity between an unlabeled point (u_1) and a labeled point (u_2) with the relevant coding vector $\mathbf{Q}_{u_1} \in \mathbb{R}^{(4B_K+24)}$ and $\mathbf{Q}_{u_2} \in \mathbb{R}^{(4B_K+24)}$

$$D_{u_1, u_2} = \sum_{i=1}^{2(4B_K+24)} (\mathbf{Q}_{u_1, i} \text{ XOR } \mathbf{Q}_{u_2, i}) \quad (23)$$

where i represents the i th bit in coding vector \mathbf{Q}_{u_1} or \mathbf{Q}_{u_2} , and XOR is the exclusive OR operator. For an unlabeled sample, it would tag the same label as the training sample that corresponds to the minimum D .

Suppose there are m training samples from C classes, in this way, three classification maps, \mathbf{C}_R , \mathbf{C}_M , and \mathbf{C}_P , derived from the KPCA-based features \mathbf{R} , Gabor magnitude \mathbf{M} , and Gabor phase \mathbf{P} , respectively, are generated, which have the same dimensions, i.e., $\mathbb{R}^{X \times Y}$, and the value of each pixel is within $[1, \dots, C]$.

B. Generation of the Superpixel Map

Superpixels, regarded as the gathering of some adjacent pixels within the same class, have been confirmed to improve the classification accuracy. In this article, the ERS and SNIC segmentation algorithms are employed to produce two types of superpixel maps for latter decision fusion.

In order to take advantage of multisource data, HSIs and LiDAR data are joint as concatenated data \mathbf{J} , which can be expressed as $\mathbf{J} = [\mathbf{H}, \mathbf{L}] \in \mathbb{R}^{X \times Y \times (B+1)}$. Since the ERS algorithm formulates the superpixel segmentation problem as an optimization problem presenting an objective function by the graph topology, the concatenated data \mathbf{J} are first transformed into an undirected graph $G = (V, E)$. It is worth to point out that the edge weight is directly computed from \mathbf{J} rather than PCA-reduced data. Then, for a given number of superpixel T , through optimizing the objective function that consists of two elements [the entropy rate of the random walk on G and the balancing term of the cluster distribution, as shown in (13)], the resulting segmentation map \mathbf{S}_E is obtained [63].

SNIC shows great potential for superpixel segmentation because of its low computational complexity and good segmentation results [57]. In order to make SNIC directly applicable in hyperspectral and LiDAR data combination, we extend the three-channel color space given in (14) to the multichannel space for SNIC algorithm, and the distance between pixel i and j is calculated by

$$d_{i,j} = \sqrt{\omega_1 \|\mathbf{J}_i - \mathbf{J}_j\|^2 + \omega_2 \|\mathbf{x}_i - \mathbf{x}_j\|^2} \quad (24)$$

where ω_1 and ω_2 are, respectively, set as $1/(B + 1)$ and 0.5. Eventually, for a given number of superpixel T , the superpixel map \mathbf{S}_N can be generated. It is worth mentioning that more superpixel segmentation algorithms can be imported into the proposed MFSuDF approach, whereas the overall accuracy (OA) is not expected to improve substantially due to less supplementary information of the added superpixel segmentation maps.

C. Superpixel-Based Decision Fusion

In the previous steps, three classification maps (\mathbf{C}_R , \mathbf{C}_M , and \mathbf{C}_P) and two superpixel maps (\mathbf{S}_N and \mathbf{S}_E) are obtained. It can be easily observed that the three classification maps are generally considered for pixelwise classification with a single-type feature. Meanwhile, the superpixel maps (either \mathbf{S}_N and \mathbf{S}_E) can well characterize the spatial structure of objects, which could be utilized to regularize the pixelwise classification maps. Specifically, for simplicity, let \mathbf{C} stand for the classification map (\mathbf{C}_R , \mathbf{C}_P , and \mathbf{C}_M), while \mathbf{S} be the superpixel map (\mathbf{S}_N and \mathbf{S}_E). Accordingly, the regularized map \mathbf{Z} can be computed by

$$\mathbf{Z}_{(t)} = \arg \max_{c=1, \dots, C} \sum_{j \in \mathbf{S}_{(t)}} \mathcal{I}(\mathbf{C}_{(j)}, c), \quad t = 1, \dots, T \quad (25)$$

where C and T are, respectively, the number of classes and superpixels, $\mathbf{S}_{(t)}$ is the t th superpixel in \mathbf{S} , and $\mathbf{C}_{(j)}$ is the predicted class of the j th pixel. Besides, the function \mathcal{I} is defined by

$$\mathcal{I}(t_1, t_2) = \begin{cases} 1 & \text{if } t_1 == t_2 \\ 0 & \text{otherwise} \end{cases} \quad (26)$$

which means that the predicted label of $\mathbf{S}_{(t)}$ is the mode of all the sample labels contained in this superpixel.

After correction by superpixel segmentation, six intermediate classification maps (\mathbf{Z}_l , $l = 1, \dots, 6$) are obtained by pairwise interaction between the three classification maps and two superpixel maps. However, the regularized maps may have overlap error, and each of them performs differently on the peripheral point of the superpixel segmentation. Especially, when the sample number of some classes in the scene is small and the distribution is not concentrated, the superpixel-guided method could have a negative impact on the classification result. Therefore, to further enhance the robustness of the proposed MFSuDF methodology, the three pixelwise classification maps, including \mathbf{C}_R , \mathbf{C}_M , and \mathbf{C}_P , are also taken into account. Finally, the majority voting mechanism is applied on both the six intermediate maps (\mathbf{Z}_l , $l = 1, \dots, 6$) and three pixelwise classification maps, and the final map \mathbf{F} is determined by

$$\mathbf{F}_{(i)} = \arg \max_{c=1, \dots, C} \left[\sum_{l=1}^6 \mathcal{I}(\mathbf{Z}_{l(i)}, c) + 2 \sum_{k=1}^3 \mathcal{I}(\mathbf{C}_{k(i)}, c) \right] \\ i = 1, \dots, XY; \quad \mathbf{C}_k \in \{\mathbf{C}_R, \mathbf{C}_M, \mathbf{C}_P\} \quad (27)$$

where the weight of three pixelwise classification maps is granted as 2 to make the voting procedure more fair. The experiments demonstrate that the decision fusion strategy of

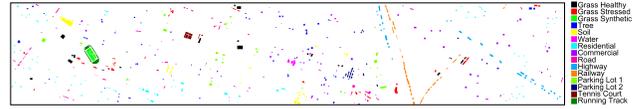


Fig. 3. Ground-truth map of the Houston data set (15 land-cover classes).

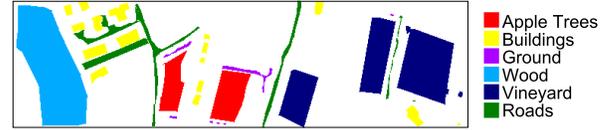


Fig. 4. Ground-truth map of the Trento data set (six land-cover classes).

the various kinds of features can effectively increase the classification accuracy.

The computational complexity of our proposed MFSuDF can be roughly divided into two parts: unsupervised (including superpixel segmentation, ERS-guided KPCA, and Gabor feature extraction) and supervised (including the classification and decision fusion). Specifically, the complexity of two superpixel segmentation methods (ERS and SNIC) is, respectively, $O(XY \log(XY))$ and $O(XY)$. After that, the complexity of ERS-guided KPCA is $O((X + Y)^2 + XY \log(XY))$. Besides, for each Gabor filter, the computational complexity is $O(XYB \log(XYB))$ (it is worth to point out that the Gabor filters with only one orientation are reserved in the proposed approach). Alternatively, for the supervised part, the complexities of the QBC classifier and RF classifier are $O(XYB)$ and $O(XY \log(B))$, respectively, while the decision fusion part is $O(XY)$. Since the unsupervised part can be carried out only once, while the supervised part is linear to the spatial coverage, the computational efficiency of our MFSuDF method can be well guaranteed.

V. EXPERIMENTS

A. Data Sets

1) *Houston Data Set*: The data set captured over the University of Houston campus and its neighboring regions is used in the experiment [20]. The Houston HSI data set contains 144 spectral bands ranging from 380 to 1050 nm, and each band has 349×1905 pixels with 2.5-m spatial resolution. Meanwhile, the corresponding LiDAR data also are the size of 349×1905 with the height information of surface materials; 15 land-cover classes and 15029 labeled samples are given in the ground-truth image, as shown in Table II and Fig. 3.

2) *Trento Data Set*: The second data set is collected over the south of Trento, Italy, consisting of 63 spectral bands that range from 400 to 980 nm [29], [64]. Each band is 600×166 pixels with a spatial resolution of 1 m. Likewise, the LiDAR data only have one band of the same spatial size. The six land-cover classes and 30414 labeled pixels are described in detail in Table III and Fig. 4.

3) *MUUFLL Gulfport Data Set*: The third data set was collected over the Gulf Park Campus of the University of Southern Mississippi [65], [66]. The collected data set includes HIS and LiDAR data. The original HSIs have 325×337

TABLE II
LAND-COVER CLASSES IN THE HOUSTON DATA SET

Class	Land-cover Type	No. of Samples
C1	Healthy Grass	1251
C2	Stressed Grass	1254
C3	Synthetic Grass	697
C4	Trees	1244
C5	Soil	1242
C6	Water	325
C7	Residential	1268
C8	Commercial	1244
C9	Roads	1252
C10	Highways	1227
C11	Railways	1235
C12	Parking Lot 1	1233
C13	Parking Lot 2	469
C14	Tennis Court	428
C15	Running Track	660
	Total	15029

TABLE III
LAND-COVER CLASSES IN THE TRENTO DATA SET

Class	Land-cover Type	No. of Samples
C1	Apple Trees	4034
C2	Buildings	2903
C3	Ground	479
C4	Wood	9123
C5	Vineyard	10501
C6	Roads	3374
	Total	30414

TABLE IV
LAND-COVER CLASSES IN THE MUUFL GULFPORT DATA SET

Class	Land-cover Type	No. of Samples
C1	Trees	23246
C2	Mostly Grass	4270
C3	Mixed Ground	6882
C4	Dirt and Sand	1826
C5	Roads	6687
C6	Water	466
C7	Building Shadows	2233
C8	Buildings	6240
C9	Sidewalks	1385
C10	Yellow Curbs	183
C11	Cloth Panels	269
	Total	53687

pixels with 72 bands. However, eight bands were removed due to severe noise, leaving 64 bands for the experiment. In addition, the scene contains a region of invalid data in the lower-right corner of the image. Therefore, the original HSI was cropped to $325 \times 220 \times 64$ as the new data set. The details are given in Table IV and Fig. 5.

B. Experimental Setup

In this article, we focus on addressing the small sample size problem. Therefore, a limited number of labeled samples (from 3 to 15) are selected from each class to form the training set. The remaining samples are used to compose the test set. Afterward, to evaluate the performance of different methods, the OA and kappa coefficient (κ) are calculated. More precisely, the OA is expressed as a percentage calculated by dividing the number of corrected classified samples by the sum of the test samples. κ is a measure of agreement and is

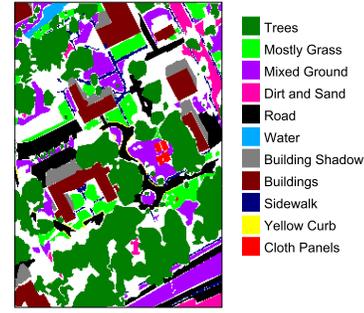


Fig. 5. Ground-truth map of the MUUFL Gulfport data set (11 land-cover classes).

defined as

$$\kappa = \frac{OA - EA}{1 - EA} \quad (28)$$

where EA is the expected accuracy by random chance. The greater the κ value is, the more accurate the classification results.

To verify the effectiveness of the proposed method for classification, a series of experimental tests are carried out. Our proposed approach MFSuDF is compared with 11 state-of-the-art methods.

- 1) *RAW*: The RF classifier is applied to the concatenated HSIs and LiDAR data, i.e., \mathbf{J} .
- 2) *EMAP*: The RF classifier is applied to the extended morphological features of \mathbf{J} [67].
- 3) *GCK*: Generalized composite kernel classifier is applied on \mathbf{J} [68].
- 4) *ERS*: The RF classifier is applied on \mathbf{J} , and ERS is used to regularize the map.
- 5) *3-D-CNN*: 3-D convolutional neural network for HSIs and LiDAR data [69].
- 6) *EPCA*: A novel ensemble classifier is applied on HSIs and LiDAR data [70].
- 7) *GGF*: Generalized graph-based fusion of HSIs and LiDAR data [31].
- 8) *EPFF*: Feature fusion using extinction profiles of HSIs and LiDAR data [25].
- 9) *KPCA-GaborM*: The RF classifier is applied to the Gabor magnitude features of KPCA-based HSIs and LiDAR (i.e., \mathbf{M}), which is just the classification map \mathbf{C}_M .
- 10) *KPCA-ERS*: The RF classifier is applied to the concatenated KPCA-based HSIs and LiDAR (i.e., \mathbf{R}), and ERS is used to regularize the obtained map \mathbf{C}_R , which is just the classification map \mathbf{Z}_4 in Fig. 1.
- 11) *MFDF*: The decision fusion strategy is applied on the three pixelwise classification maps (\mathbf{C}_R , \mathbf{C}_M , and \mathbf{C}_P) without superpixel regularization.

To reduce the effects of random factors, training samples are selected randomly 20 times from each data set. Both the means and standard deviations are calculated in the experiments.

C. Parameter Settings

In our proposed MFSuDF framework, most of the parameters have been set by default. Concretely, the KPCA dimension

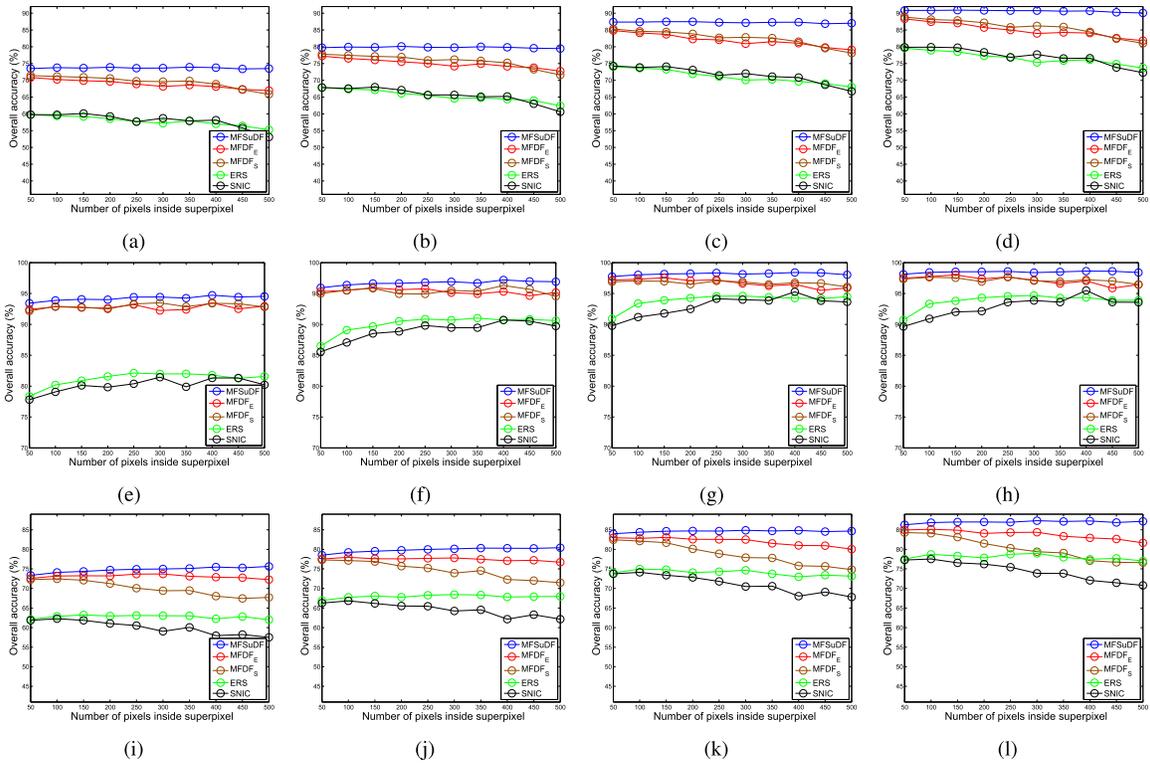


Fig. 6. OA versus the number of pixels inside superpixel with a different number of training samples per class on the Houston data set (first row), the Trento data set (second row), and the MUUFL Gulfport data set (third row). The number of training samples per class is 3 (first column), 5 (second column), 10 (third column), and 15 (last column).

can be decided by the 99% energy contained in the principal components, while the scale and orientation parameters of Gabor wavelets have been presented in (16) and (19).

In this section, let us focus on the last important parameter that is not explicitly given: the number of superpixels in the superpixel map. Since the size of each image is not consistent, it is less meaningful to directly set the same number of superpixels for all data sets. Conversely, some experiments based on the number of pixels inside the superpixel have been conducted, which has shown good performance [71]. At the same time, because our previous experiments are not comparable for the superpixels, some other methods are added for comparison to reinforce the claim of superiority and robustness of MFSuDF.

- 1) *SNIC*: The RF classifier is applied on \mathbf{J} , and SNIC is used to regularize the map.
- 2) *MFDF_S*: The decision fusion strategy is applied on the three SNIC regularized maps (i.e., \mathbf{Z}_1 , \mathbf{Z}_2 , and \mathbf{Z}_3 shown in Fig. 1).
- 3) *MFDF_E*: The decision fusion strategy is applied on the three ERS regularized maps (i.e., \mathbf{Z}_4 , \mathbf{Z}_5 , and \mathbf{Z}_6 shown in Fig. 1).

Besides, the ERS method presented in the last subsection is also included for comparison. The number of pixels inside the superpixel ranges from 50 to 500. Fig. 6 shows that the OA metric varies with the number of pixels inside superpixel on the Houston data set, the Trento data set, and the Missouri University and University of Florida (MUUFL) Gulfport data set, respectively. Here, four circumstances, i.e., 3, 5, 10, and 15 training samples per class, are concerned.

In Fig. 6, the OA of all compared methods generally improves with the increase of training samples (from left to right), and it can be clearly seen that MFSuDF, MFDF_E, and MFDF_S are above the ERS and SNIC, which confirms that the incorporation of Gabor texture information and KPCA-based information can significantly promote the discrimination of objects. Meanwhile, from left subfigures to right subfigures of Fig. 6, which represents an increase in the number of training samples, it can be observed that the curve of MFDF_E and MFDF_S becomes unstable, and even decreases a lot as the size of superpixel increases. This is because the spatial homogeneity of objects can be well described with a small number of pixels inside superpixels. With the increase in the number of pixels inside superpixels, the superpixel map is more likely to be undersegmented, eventually deteriorating the classification performance. Alternatively, since our MFSuDF approach not only includes the regularized maps from both kinds of superpixel segmentation but also imports the pixelwise classification maps of different features (i.e., \mathbf{C}_R , \mathbf{C}_M , and \mathbf{C}_p), MFSuDF is always superior to the compared methods, and the corresponding curve is kept stable with a different number of pixels in superpixel, as shown in Fig. 6. Thus, the robustness of our framework can be guaranteed, and the number of pixels inside superpixel is simply set to 100 by default in our following experiment for all data sets. Correspondingly, the number of superpixel regions is calculated by the spatial size of the scene ($X \times Y$) divided by 100, and Table V lists the number of superpixel regions using ERS and SNIC for the three data sets (it is worth pointing out that the number of superpixel regions of ERS and SNIC is not

TABLE V
NUMBER OF SUPERPIXEL REGIONS USING ERS AND SNIC FOR THE HOUSTON, TRENTO, AND MUUFL GULFPORT DATA SETS

	Houston	Trento	MUUFL Gulfport
ERS	6647	995	714
SNIC	6460	960	704

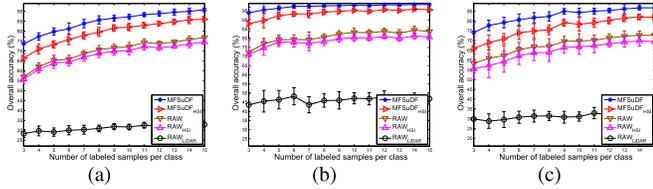


Fig. 7. OA as functions of the number of labeled samples per class on (a) (Left) Houston, (b) (Middle) Trento, and (c) (Right) MUUFL Gulfport data sets.

the same due to different computational procedures of the two segmentation methods).

In our experiments, since 99% energy contained in the principal components of KPCA is kept, the dimensions of KPCA B_K are 18, 10, and 11, respectively, for the Houston, Trento, and MUUFL Gulfport data sets. Alternatively, for the RF classifier, the number of the decision tree is set as 500 by default. Specifically, four 3-D Gabor wavelets are applied on the KPCA-reduced HSIs data, and the dimensions of the Gabor magnitude feature are 72, 40, and 44 for the three data sets, respectively, (the same for the Gabor phase feature). Meanwhile, 24 2-D Gabor filters are applied for LiDAR data, and the dimensions of the achieved Gabor features are 24, 24, and 24 for the three data sets, respectively. Hence, the concatenated feature dimensions are 96, 64, and 68 for the three data sets, and the depth of the trees of the forest is 6, 6, 6, which is computed based on the rule that the feature subspace is the floor of the logarithmic value of feature dimension. Similarly, for the raw data-based feature (concatenated KPCA-reduced HSIs and LiDAR data), the feature dimension is 19, 11, 12 for the three data set, and the depth of the trees of the forest is 4, 3, 3, respectively.

D. Experimental Results

In order to demonstrate the relevance of HSIs and LiDAR data, the RF classifier is applied to the LiDAR and HSIs data, which is, respectively, called RAW_{LiDAR} and RAW_{HSI} . At the same time, the MFSuDF approach applied to the HSIs data, denoted as $MFSuDF_{HSI}$, is also concerned. Fig. 7 shows the OA of the five compared methods (RAW_{LiDAR} , RAW_{HSI} , RAW, $MFSuDF_{HSI}$, and MFSuDF) with a different number of labeled samples per class on the three data sets. Here, the number of labeled samples per class varies from 3 to 15. Each experiment is executed 20 times, and the mean values and standard variation are reported. It can be observed from Fig. 7 that the RAW and MFSuDF methods, which consider both HSIs and LiDAR data, achieve better performance than the rest three methods, including RAW_{HSI} , RAW_{LiDAR} , and $MFSuDF_{HSI}$, validating the indispensability

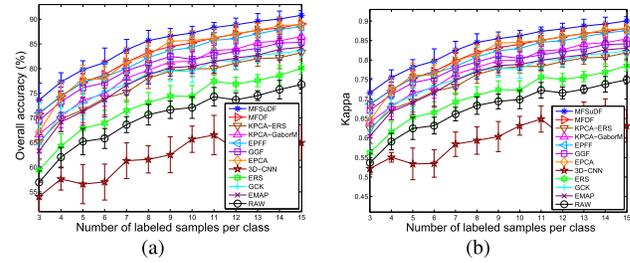


Fig. 8. Houston data set. (a) OA and (b) kappa as functions of the number of labeled samples per class.

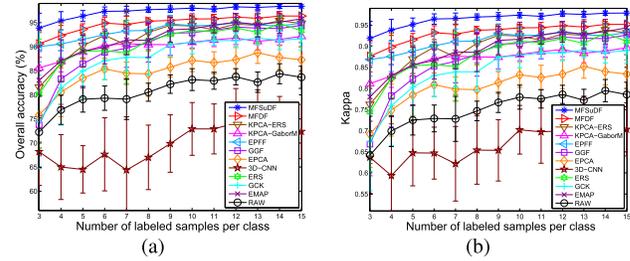


Fig. 9. Trento data set. (a) OA and (b) kappa as functions of the number of labeled samples per class.

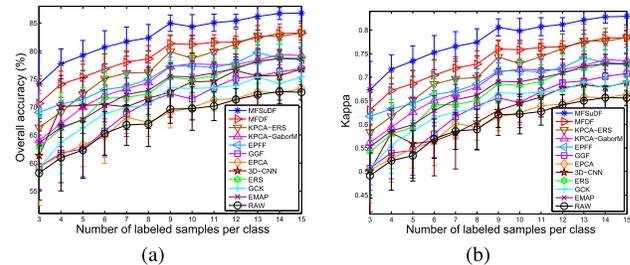


Fig. 10. MUUFL Gulfport data set. (a) OA and (b) kappa as functions of the number of labeled samples per class.

of the LiDAR data. In fact, the HSIs data provide rich discriminative spectral information for material identification, while LiDAR data can contribute potential details for HSIs, including the height and shape information of land-cover objects, which is an important supplement for material classification.

Figs. 8–10 show the OA and κ coefficient of the 12 compared methods (i.e., RAW, EMAP, GSK, ERS, 3-D-CNN, EPCA, GGF, EPFF, KPCA-GaborM, KPCA-ERS, MFDF, and our MFSuDF) when varying the number of training samples per class from 3 to 15. Generally, the classification results are better with the increase in the training sizes for each data set. First, the results of 3-D-CNN (on the Houston and Trento data sets) and RAW without any feature extraction procedure (on the MUUFL Gulfport data set) provide the lowest performance in most cases, implying the instability of deep learning-based method for small sample set issue. Meanwhile, it can also be seen that the ERS method in the Houston data set and EPCA in the other two data sets output the second-worst results, which means that a single feature extraction method can only show the characteristics of one aspect of an object.

TABLE VI
CLASSIFICATION PERFORMANCE USING RAW, EMAP, GCK, ERS, 3-D-CNN, EPCA, GGF, EPPF, KPCA-GABORM, KPCA-ERS, MFDF, AND MFSuDF FOR THE HOUSTON DATA SET WITH THREE LABELED SAMPLES PER CLASS AS TRAINING SET

Class	RAW	EMAP	GCK	ERS	3D-CNN	EPCA	GGF	EPPF	KPCA-GaborM	KPCA-ERS	MFDF	MFSuDF
C1	79.24	80.15	74.42	78.83	58.24	73.74	69.68	85.05	69.16	79.06	83.48	81.94
C2	66.82	61.69	69.92	68.51	78.09	66.21	74.70	78.44	51.73	72.56	77.15	78.94
C3	91.32	99.60	99.67	95.66	95.03	100.00	99.84	99.24	96.99	100.00	98.74	99.71
C4	93.09	73.19	77.45	88.34	57.94	86.94	89.54	89.35	86.73	88.09	93.37	92.79
C5	64.75	71.05	81.87	72.67	85.25	77.28	74.09	86.17	81.78	84.34	89.42	90.46
C6	82.69	88.82	85.67	86.00	58.59	89.72	88.12	75.43	78.83	89.42	82.63	83.98
C7	45.90	57.33	63.75	42.22	27.94	60.24	79.61	74.68	69.68	46.14	63.41	69.38
C8	26.91	33.96	22.84	26.50	32.57	32.94	31.55	34.20	30.03	30.47	35.33	35.61
C9	57.46	60.18	61.21	64.95	32.84	57.80	67.36	63.19	66.73	68.77	71.11	76.89
C10	30.89	45.90	55.33	31.34	37.03	47.66	52.17	41.46	56.37	44.79	56.93	59.28
C11	45.37	62.47	68.38	53.56	46.27	82.19	69.65	68.14	59.04	67.32	60.79	68.13
C12	25.19	34.11	34.22	26.02	33.08	35.66	45.00	40.55	46.52	33.09	39.41	43.35
C13	17.13	53.38	37.35	15.86	33.69	46.89	41.79	60.83	74.84	17.78	61.52	73.37
C14	75.37	95.46	75.99	92.45	61.99	93.84	90.97	98.15	97.52	93.01	95.04	98.87
C15	86.91	97.64	97.42	93.19	85.45	99.98	98.23	97.68	94.41	100.00	97.58	98.61
OA	56.95	63.30	64.59	59.37	54.02	67.12	69.51	71.02	66.67	65.11	70.86	73.64
Kappa	0.54	0.60	0.62	0.56	0.50	0.64	0.67	0.69	0.64	0.62	0.69	0.72

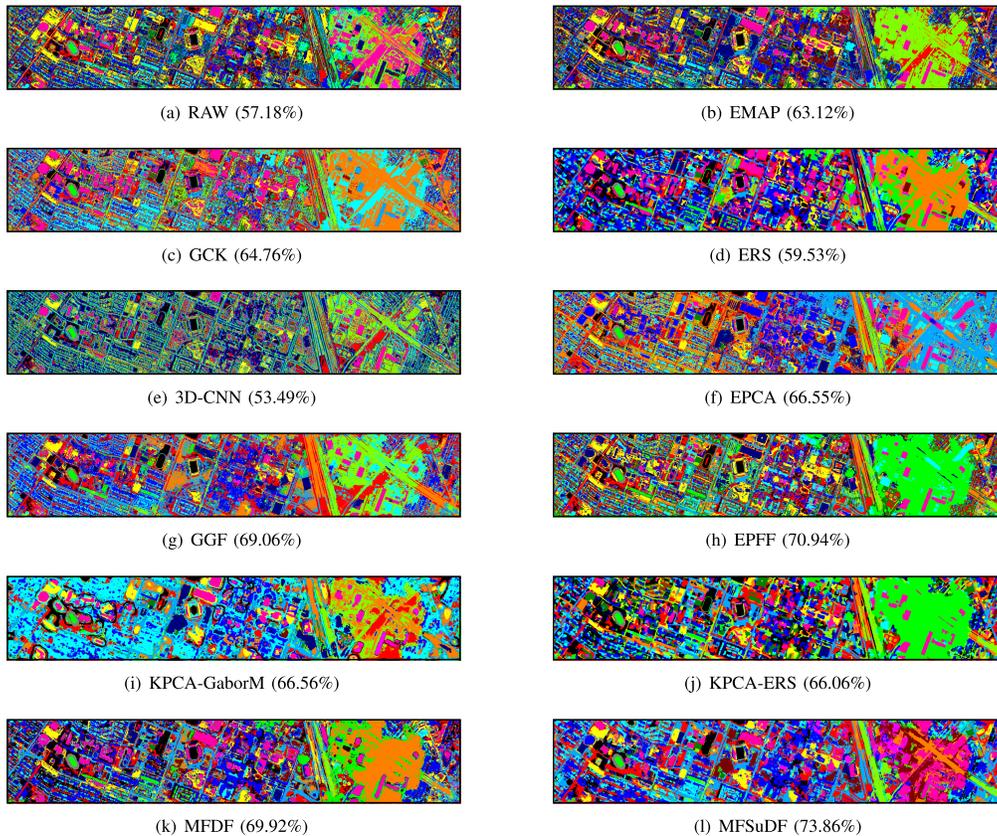


Fig. 11. Houston data set: classification maps obtained by: (a) RAW, (b) EMAP, (c) GCK, (d) ERS, (e) 3-D-CNN, (f) EPCA, (g) GGF, (h) EPPF, (i) KPCA-GaborM, (j) KPCA-ERS, (k) MFDF, and (l) MFSuDF when the number of training samples is three per class (the percentage in the brackets is the corresponding accuracy).

If the discrimination ability of the feature is relatively low, the classification accuracy will be greatly affected. Besides, the accuracies of EPPF are higher than those of EMAP, GCK, GGF, KPCA-GaborM, and KPCA-ERS, implying that a simple spatial filter could have a strong discriminative ability by reducing within-class spectral variations and intrinsic noise.

As expected, the two Gabor feature-based decision fusion methods—MFSuDF and MFDF—have demonstrated better classification performance than the other methods in most cases. This is reasonable since the information complementarity from multiple modules is helpful for material classification, which also validates the effectiveness of decision-level fusion strategy with multiple kinds of features. Furthermore, by

TABLE VII
CLASSIFICATION PERFORMANCE USING RAW, EMAP, GCK, ERS, 3-D-CNN, EPCA, GGF, EPFF, KPCA-GABORM, KPCA-ERS, MFDF, AND MFSuDF FOR THE TRENTO DATA SET WITH THREE LABELED SAMPLES PER CLASS AS TRAINING SET

Class	RAW	EMAP	GCK	ERS	3D-CNN	EPCA	GGF	EPFF	KPCA-GaborM	KPCA-ERS	MFDF	MFSuDF
C1	81.33	92.80	87.66	88.26	86.45	91.42	92.13	91.34	87.41	84.93	91.36	93.00
C2	60.84	70.73	61.73	61.33	61.60	80.50	68.57	83.15	83.58	74.95	92.23	93.55
C3	96.82	97.86	92.43	92.04	95.13	77.40	95.97	96.37	65.41	92.88	91.71	92.13
C4	92.24	97.92	89.41	98.98	70.11	97.70	85.98	97.77	96.33	97.46	98.63	99.59
C5	57.20	74.59	64.46	70.77	65.32	51.60	55.89	87.70	85.12	72.32	88.16	95.01
C6	60.13	64.95	55.98	62.53	51.19	68.72	71.05	74.98	58.87	66.71	72.99	75.37
OA	72.29	83.05	74.38	80.19	68.22	75.82	73.80	90.01	85.59	81.58	90.60	93.88
Kappa	0.64	0.78	0.67	0.74	0.59	0.69	0.67	0.87	0.81	0.76	0.88	0.92

TABLE VIII
CLASSIFICATION PERFORMANCE USING RAW, EMAP, GCK, ERS, 3-D-CNN, EPCA, GGF, EPFF, KPCA-GABORM, KPCA-ERS, MFDF, AND MFSuDF FOR THE MUUFL GULFPORT DATA SET WITH THREE LABELED SAMPLES PER CLASS AS TRAINING SET

Class	RAW	EMAP	GCK	ERS	3D-CNN	EPCA	GGF	EPFF	KPCA-GaborM	KPCA-ERS	MFDF	MFSuDF
C1	63.54	70.09	55.94	70.24	68.58	59.31	57.29	71.30	68.45	70.26	78.17	80.67
C2	65.78	69.15	57.90	66.90	65.09	53.69	63.19	71.06	69.14	62.31	71.41	76.70
C3	45.52	47.07	39.49	48.93	48.90	42.53	54.15	47.50	45.09	54.16	50.08	54.44
C4	46.00	54.88	49.15	47.93	59.26	37.83	48.19	57.73	60.31	46.16	60.98	64.02
C5	71.13	75.04	76.60	78.57	55.10	71.23	76.80	78.78	68.10	86.43	79.03	85.01
C6	84.20	90.74	98.92	94.25	60.92	97.83	95.55	88.06	98.67	94.36	96.88	98.38
C7	63.99	64.85	90.18	70.47	68.18	82.08	78.56	70.30	63.78	86.72	64.09	68.75
C8	35.86	46.26	66.12	36.78	57.60	68.12	46.58	68.89	64.17	50.54	65.90	69.60
C9	40.64	39.81	41.73	36.44	32.34	37.70	42.59	37.20	38.91	35.77	38.16	37.30
C10	50.66	60.22	54.47	9.34	34.26	43.69	53.42	53.33	45.85	12.05	46.39	30.82
C11	90.37	92.30	70.55	89.59	60.89	91.91	85.61	83.22	96.00	85.28	96.26	92.71
OA	58.24	63.65	59.01	62.88	61.33	59.32	59.26	69.04	64.06	66.34	70.65	74.11
Kappa	0.49	0.55	0.51	0.54	0.52	0.50	0.51	0.62	0.56	0.58	0.63	0.67

comparing these two methods that perform decision fusion, it is not difficult to find that the MFSuDF curve is always above the MFDF curve, indicating the necessity of incorporating the superpixel regularization procedure.

Furthermore, when there are only three training samples per class, the results of the 12 methods have significant differences, which are summarized in Tables VI–VIII. Here, the classification performance per class is evaluated by the precision, which is denoted as a percentage calculated by dividing the number of corrected classified samples per class by the number of test samples per class. In these tables, it can be seen that, in most cases, the performance yielded by MFSuDF is better than those yielded by the other methods in the three data sets. Specifically, considering the C5 class (Vineyard) of the Trento data set, it can be found from the ground truth map (see Fig. 4) that the spatial distribution of C5 is very regular; thus, the advantages of feature fusion and superpixel regularization procedure in our MFSuDF method can be fully engaged, and the performance increases from 57.20% of the RAW method to 95.01% of our approach, as shown in Table VII. Alternatively, concerning the C10 class (Yellow Curbs) of the MUUFL Gulfport data set in Table VIII, the classification performances of ERS (9.34%) and KPCA-ERS (12.05%) are much lower than the others. In fact, since there are only 183 labeled samples of C10, and they are scattered in the scene, it is even hard to be seen in Fig. 5. Through integrating the three pixelwise classification maps (i.e., C_R , C_M , and C_P), the accuracy of MFSuDF increases to 30.82%, indicating the necessity and rationality of feature fusion strategy of the proposed method. Of course, the performance of MFSuDF with C10 is still lower than the nonsuperpixel-based methods, which needs

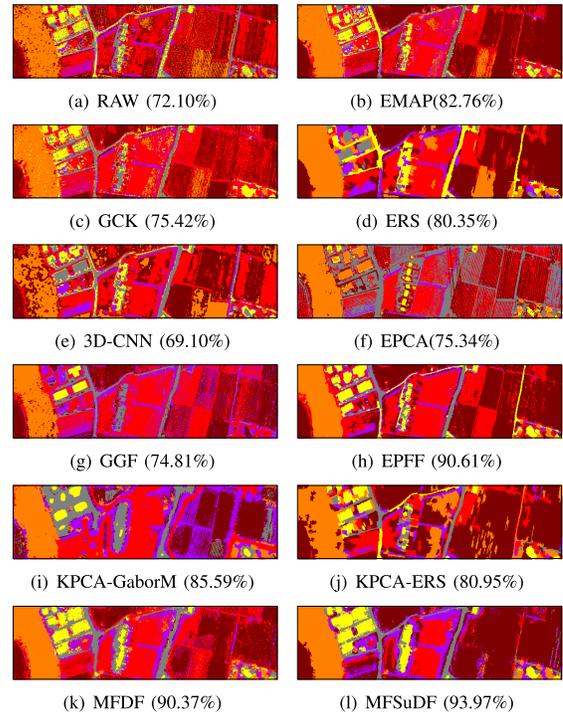


Fig. 12. Trento data set: classification maps obtained by (a) RAW, (b) EMAP, (c) GCK, (d) ERS, (e) 3-D-CNN, (f) EPCA, (g) GGF, (h) EPFF, (i) KPCA-GaborM, (j) KPCA-ERS, (k) MFDF, and (l) MFSuDF when the number of training samples is three per class (the percentage in the brackets is the corresponding accuracy).

further research to incorporate the spatial information more elegant. To illustrate, the complete classification maps for all data of the 12 methods are shown in Figs. 11–13.

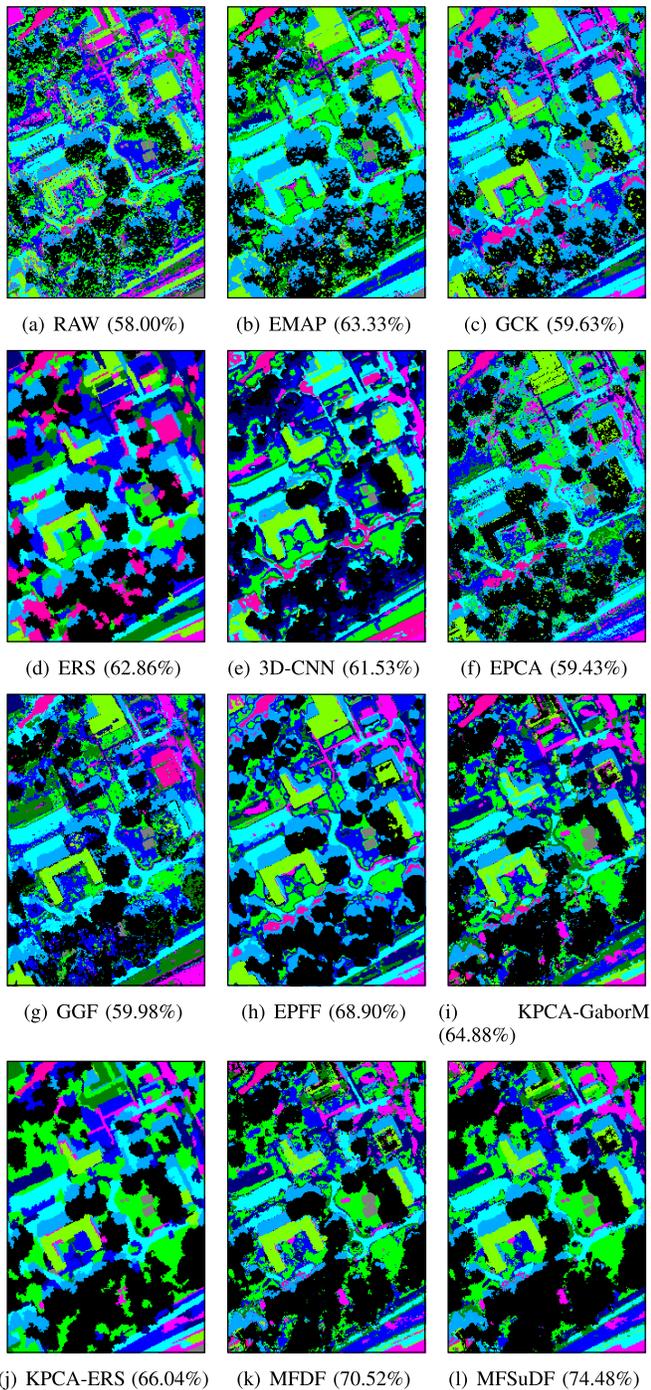


Fig. 13. MUUFL Gulfport data set: classification maps obtained by (a) RAW, (b) EMAP, (c) GCK, (d) ERS, (e) 3-D-CNN, (f) EPCA, (g) GGF, (h) EPFF, (i) KPCA-GaborM, (j) KPCA-ERS, (k) MFDF, and (l) MFSuDF when the number of training samples is three per class (the percentage in the brackets is the corresponding accuracy).

We can easily observe that MFSuDF has the best classification effect.

VI. CONCLUSION

An MFSuDF approach for HSI and LiDAR data classification has been proposed in this article. In detail, after ERS-guided KPCA is applied to HSIs, the 2-D and 3-D Gabor wavelets are employed to extract representative spectral-spatial and elevation information from the

KPCA-reduced HSIs and LiDAR data, respectively. Then, we utilize RF and QBC classifiers to obtain three classification maps containing information from different aspects. Finally, a superpixel-based decision fusion classification method is adopted on all extracted maps (including pixelwise and superpixel regularized) to obtain a final classification map.

In summary, the main contributions of this article are proposing an ERS-guided KPCA, integrating 2-D and 3-D Gabor features of multisource data, exploring the interaction between multiple classification maps and two superpixel segmentations, and designing a decision fusion framework based on all extracted classification maps. Compared with the RAW, EMAP, GCK, ERS, 3-D-CNN, EPCA, GGF, EPFF, KPCA-GaborM, KPCA-ERS, and MFDF methods, the experimental results consistently show that the fusion of HIS and LiDAR data using our proposed MFSuDF method can efficiently improve the classification accuracy. Superpixel-based decision fusion exhibits excellent performance, which proves that decision with different superpixel segmentations is a remarkable tool for classification. As mentioned earlier, the superpixel segmentation may have poor results due to the scattering effects of land-cover classes, which should be paid more attention and is the future direction in our work.

REFERENCES

- [1] J. Richards, *Remote Sensing Digital Image Analysis: An Introduction*. Berlin, Germany: Springer, 2013.
- [2] G. Camps-Valls, D. Tuia, L. Gomez-Chova, S. Jimenez, and J. Malo, *Remote Sensing Image Processing*. San Mateo, CA, USA: Morgan & Claypool, Dec. 2011.
- [3] J. M. Bioucas-Dias *et al.*, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 354–379, Apr. 2012.
- [4] P. Ghamisi, M. Dalla-Mura, and J. Atli-Benediktsson, "A survey on spectral-spatial classification techniques based on attribute profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2335–2353, May 2015.
- [5] G. Zhu, Y. Huang, S. Li, J. Tang, and D. Liang, "Hyperspectral band selection via rank minimization," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 12, pp. 2320–2324, Dec. 2017.
- [6] S. Ozturk, Y. E. Esin, Y. Artan, O. Ozdil, and B. Demirel, "Importance of band selection for ethene and methanol gas detection in hyperspectral imagery," in *Proc. 9th Workshop Hyperspectral Image Signal Processing: Evol. Remote Sens. (WHISPERS)*, Sep. 2018, pp. 1–4.
- [7] L. Gao *et al.*, "Subspace-based support vector machines for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 2, pp. 349–353, Feb. 2015.
- [8] Y. Wang and H. Duan, "Classification of hyperspectral images by SVM using a composite kernel by employing spectral, spatial and hierarchical structure information," *Remote Sens.*, vol. 10, no. 3, p. 441, Mar. 2018.
- [9] Y. Chen, N. M. Nasrabadi, and T. D. Tran, "Hyperspectral image classification via kernel sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 1, pp. 217–231, Jan. 2013.
- [10] Y. M. Ren, Y. N. Zhang, and W. Wei, "Hyperspectral image classification via discriminative sparse representation with extended LBP texture," *Adv. Mater. Res.*, vols. 989–994, pp. 3885–3888, Jul. 2014.
- [11] S. Zhang, S. Li, W. Fu, and L. Fang, "Multiscale superpixel-based sparse representation for hyperspectral image classification," *Remote Sens.*, vol. 9, no. 2, p. 139, Feb. 2017.
- [12] R. Huang and J. Zhu, "Using random forest to integrate LiDAR data and hyperspectral imagery for land cover classification," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2013, pp. 3978–3981.
- [13] R. Hansch and O. Hellwich, "Feature-independent classification of hyperspectral images by projection-based random forests," in *Proc. 7th Workshop Hyperspectral Image Signal Processing, Evol. Remote Sens. (WHISPERS)*, Jun. 2015, pp. 1–4.

- [14] Y. Liu, L. Liu, Y. Gao, and L. Yang, "An improved random forest algorithm based on attribute compatibility," in *Proc. IEEE 3rd Inf. Technol., Netw., Electron. Autom. Control Conf. (ITNEC)*, Mar. 2019, pp. 2558–2561.
- [15] J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. Nasrabadi, and J. Chanussot, "Hyperspectral remote sensing data analysis and future challenges," *IEEE Geosci. Remote Sens. Mag.*, vol. 1, no. 2, pp. 6–36, Jun. 2013.
- [16] M. Kishore and S. B. Kulkarni, "Approches and challenges in classification for hyperspectral data: A review," in *Proc. Int. Conf. Electr., Electron., Optim. Techn. (ICEEOT)*, Mar. 2016, pp. 3418–3421.
- [17] S. Jia, Z. Zhu, L. Shen, and Q. Li, "A two-stage feature selection framework for hyperspectral image classification using few labeled samples," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1023–1035, Apr. 2014.
- [18] Y. Zhou, J. Peng, and C. L. P. Chen, "Dimension reduction using spatial and spectral regularized local discriminant embedding for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 2, pp. 1082–1095, Feb. 2015.
- [19] P. Hartzell, C. Glennie, and S. Khan, "Terrestrial hyperspectral image shadow restoration through LiDAR fusion," *Remote Sens.*, vol. 9, no. 5, p. 421, Apr. 2017.
- [20] C. Debes *et al.*, "Hyperspectral and LiDAR data fusion: Outcome of the 2013 GRSS data fusion contest," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2405–2418, Jun. 2014.
- [21] Y. Gu, Q. Wang, X. Jia, and J. A. Benediktsson, "A novel MKL model of integrating LiDAR data and MSI for urban area classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 10, pp. 5312–5326, Oct. 2015.
- [22] W. Liao, R. Bellens, A. Pizurica, S. Gautama, and W. Philips, "Combining feature fusion and decision fusion for classification of hyperspectral and LiDAR data," in *Proc. IEEE Geosci. Remote Sens. Symp.*, Jul. 2014, pp. 1241–1244.
- [23] M. Khodadadzadeh, J. Li, S. Prasad, and A. Plaza, "Fusion of hyperspectral and LiDAR remote sensing data using multiple feature learning," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 2971–2983, Jun. 2015.
- [24] M. R. Soleimanzadeh, A. Karami, and P. Scheunders, "Fusion of hyperspectral and LiDAR images using non-subsampled shearlet transform," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2018, pp. 8873–8876.
- [25] M. Zhang, P. Ghamisi, and W. Li, "Classification of hyperspectral and LiDAR data using extinction profiles with feature fusion," *Remote Sens. Lett.*, vol. 8, no. 10, pp. 957–966, Oct. 2017.
- [26] B. Bigdeli and P. Pahlavani, "A Dempster Shafer-based fuzzy multisensor fusion system using airborne LiDAR and hyperspectral imagery," *Int. J. Remote Sens.*, vol. 39, no. 21, pp. 7718–7737, 2018.
- [27] C. Ge, Q. Du, W. Li, Y. Li, and W. Sun, "Hyperspectral and LiDAR data classification using kernel collaborative representation based residual fusion," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 6, pp. 1963–1973, Jun. 2019.
- [28] Z. Zhong, B. Fan, K. Ding, H. Li, S. Xiang, and C. Pan, "Efficient multiple feature fusion with hashing for hyperspectral imagery classification: A comparative study," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 8, pp. 4461–4478, Aug. 2016.
- [29] B. Rasti, P. Ghamisi, and R. Gloaguen, "Hyperspectral and LiDAR fusion using extinction profiles and total variation component analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 7, pp. 3997–4007, Jul. 2017.
- [30] W.-S. Chen, X. Dai, B. Pan, and T. Huang, "A novel discriminant criterion based on feature fusion strategy for face recognition," *Neurocomputing*, vol. 159, pp. 67–77, Jul. 2015.
- [31] W. Liao, A. Pizurica, R. Bellens, S. Gautama, and W. Philips, "Generalized graph-based fusion of hyperspectral and LiDAR data using morphological features," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 3, pp. 552–556, Mar. 2015.
- [32] Z. Ye, S. Prasad, W. Li, J. E. Fowler, and M. He, "Classification based on 3-D DWT and decision fusion for hyperspectral image analysis," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 1, pp. 173–177, Jan. 2014.
- [33] K. Schindler, "An overview and comparison of smooth labeling methods for land-cover classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 11, pp. 4534–4545, Nov. 2012.
- [34] X. Huang, X. Liu, and L. Zhang, "A multichannel gray level co-occurrence matrix for multi/hyperspectral image texture representation," *Remote Sens.*, vol. 6, no. 9, pp. 8424–8445, Sep. 2014.
- [35] G. Chen and S.-E. Qian, "Denoising and dimensionality reduction of hyperspectral imagery using wavelet packets, neighbour shrinking and principal component analysis," *Int. J. Remote Sens.*, vol. 30, no. 18, pp. 4889–4895, Aug. 2009.
- [36] Q. Du, "Modified Fisher's linear discriminant analysis for hyperspectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 4, pp. 503–507, Oct. 2007.
- [37] C. S. Dhir and S.-Y. Lee, "Discriminant independent component analysis," *IEEE Trans. Neural Netw.*, vol. 22, no. 6, pp. 845–857, Jun. 2011.
- [38] K. P. Singh, R. Bhai, V. Mishra, P. Nagar, and J. Kasinayal, "Localization in wireless sensor network using LLE-ISOMAP algorithm," in *Proc. IEEE Region 10 Conf. (TENCON)*, Nov. 2017, pp. 393–397.
- [39] H. Cao, H. Zhang, C. Wang, and M. Liu, "Supervised locally linear embedding for polarimetric SAR image classification," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2016, pp. 7561–7564.
- [40] A. Zare, A. Ozdemir, M. A. Iwen, and S. Aviyente, "Extension of PCA to higher order data structures: An introduction to tensors, tensor decompositions, and tensor PCA," *Proc. IEEE*, vol. 106, no. 8, pp. 1341–1358, Aug. 2018.
- [41] M. Fauvel, J. Chanussot, and J. A. Benediktsson, "Kernel principal component analysis for the classification of hyperspectral remote sensing data over urban areas," *EURASIP J. Adv. Signal Process.*, vol. 2009, no. 1, pp. 1–14, Dec. 2009.
- [42] H. Halim, S. M. Isa, and S. Mulyono, "Comparative analysis of PCA and KPCA on paddy growth stages classification," in *Proc. IEEE Region 10 Symp. (TENSYP)*, May 2016, pp. 167–172.
- [43] B. Hou, T. Huang, and L. Jiao, "Spectral-spatial classification of hyperspectral data using 3-D morphological profile," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 12, pp. 2364–2368, Dec. 2015.
- [44] S. Jia, J. Hu, Y. Xie, L. Shen, X. Jia, and Q. Li, "Gabor cube selection based multitask joint sparse representation for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 6, pp. 3174–3187, Jun. 2016.
- [45] S. Jia, L. Shen, J. Zhu, and Q. Li, "A 3-D Gabor phase-based coding and matching framework for hyperspectral imagery classification," *IEEE Trans. Cybern.*, vol. 48, no. 4, pp. 1176–1188, Apr. 2018.
- [46] A. P. Moore, S. J. D. Prince, J. Warrell, U. Mohammed, and G. Jones, "Superpixel lattices," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1–8.
- [47] 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, Nov. 2012.
- [48] S. Jia, B. Deng, J. Zhu, X. Jia, and Q. Li, "Local binary pattern-based hyperspectral image classification with superpixel guidance," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 2, pp. 749–759, Feb. 2018.
- [49] X. Zhang and N. Cahill, "SLIC superpixels for efficient graph-based dimensionality reduction of hyperspectral imagery," *Proc. SPIE*, vol. 9472, May 2015, Art. no. 947209.
- [50] L. Y. Fang, S. T. Li, X. D. Kang, and J. A. Benediktsson, "Spectral-spatial classification of hyperspectral images with a superpixel-based discriminative sparse model," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 8, pp. 4186–4201, Aug. 2015.
- [51] J. Li, H. Zhang, and L. Zhang, "Efficient superpixel-level multitask joint sparse representation for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 10, pp. 5338–5351, Oct. 2015.
- [52] B. Schölkopf, A. Smola, and K. Müller, "Kernel principal component analysis," in *Proc. 7th Int. Conf. Artif. Neural Netw.* Berlin, Germany: Springer-Verlag, 1997, pp. 583–588.
- [53] D. Gabor, "Theory of communication. Part 1: The analysis of information," *J. Inst. Elect. Eng. III, Radio Commun. Eng.*, vol. 93, no. 9, pp. 429–457, Nov. 1946.
- [54] J. G. Daugman, "Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, no. 7, pp. 1169–1179, Jul. 1988.
- [55] L. Shen and L. Bai, "3D Gabor wavelets for evaluating SPM normalization algorithm," *Med. Image Anal.*, vol. 12, no. 3, pp. 375–383, Jun. 2008.
- [56] L. Shen and S. Jia, "Three-dimensional Gabor wavelets for pixel-based hyperspectral imagery classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 12, pp. 5039–5046, Dec. 2011.
- [57] R. Achanta and S. Süsstrunk, "Superpixels and polygons using simple non-iterative clustering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4895–4904.
- [58] S. Jia, B. Deng, J. Zhu, X. Jia, and Q. Li, "Superpixel-based multitask learning framework for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 5, pp. 2575–2588, May 2017.

- [59] S. Jia, B. Deng, H. Xie, and L. Deng, "A Gabor feature fusion framework for hyperspectral imagery classification," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2017, pp. 2394–2397.
- [60] S. Bravo and A. H. Moreno, "A random forest approach for predicting the microwave drying process of Amaranth seeds," in *Proc. IEEE 2nd Int. Conf. Inf. Comput. Technol. (ICICT)*, Mar. 2019, pp. 25–29.
- [61] A. Bosch, A. Zisserman, and X. Munoz, "Image classification using random forests and ferns," in *Proc. IEEE 11th Int. Conf. Comput. Vis.*, Oct. 2007, pp. 1–8.
- [62] J. G. Daugman, "High confidence visual recognition of persons by a test of statistical independence," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 15, no. 11, pp. 1148–1161, Nov. 1993.
- [63] M. Liu, O. Tuzel, S. Ramalingam, and R. Chellappa, "Entropy rate superpixel segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2011, pp. 2097–2104.
- [64] M. Zhang, W. Li, Q. Du, L. Gao, and B. Zhang, "Feature extraction for classification of hyperspectral and LiDAR data using patch-to-patch CNN," *IEEE Trans. Cybern.*, vol. 50, no. 1, pp. 100–111, Jan. 2020.
- [65] P. Gader, A. Zare, R. Close, J. Aitken, and G. Tuell, *Mufl Gulfport Hyperspectral and Lidar Airborne Data Set*. Gainesville, FL, USA: Univ. Florida, 2013.
- [66] X. Du and A. Zare, *Technical Report: Scene Label Ground Truth Map for Mufl Gulfport Data Set*. Gainesville, FL, USA: Univ. Florida, 2017.
- [67] M. Dalla Mura, J. Atli Benediktsson, B. Waske, and L. Bruzzone, "Extended profiles with morphological attribute filters for the analysis of hyperspectral data," *Int. J. Remote Sens.*, vol. 31, no. 22, pp. 5975–5991, Dec. 2010.
- [68] 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.
- [69] Y. Li, H. Zhang, and Q. Shen, "Spectral–spatial classification of hyperspectral imagery with 3D convolutional neural network," *Remote Sens.*, vol. 9, no. 1, p. 67, 2017.
- [70] J. Xia, N. Yokoya, and A. Iwasaki, "Fusion of hyperspectral and LiDAR data with a novel ensemble classifier," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 6, pp. 957–961, Jun. 2018.
- [71] S. Jia, X. Deng, J. Zhu, M. Xu, J. Zhou, and X. Jia, "Collaborative representation-based multiscale superpixel fusion for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 10, pp. 7770–7784, Oct. 2019.



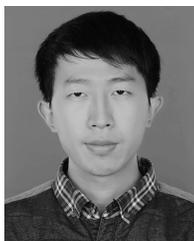
Sen Jia (Senior Member, IEEE) received the B.E. and Ph.D. degrees from the College of Computer Science, Zhejiang University, Hangzhou, China, in 2002 and 2007, respectively.

Since 2008, he has been with the College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China, where he is a Full Professor. His research interests include hyperspectral image processing, signal and image processing, and machine learning.



Zhangwei Zhan received the B.E. degree from the Hefei University of Technology, Hefei, China, in 2017. He is pursuing the master's degree in computer science and technology with the College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China.

His research interests include hyperspectral and LiDAR classification, machine learning, and pattern recognition.



Meng Zhang received the B.E. degree in surveying and mapping engineering from the College of Civil Engineering, Hefei University of Technology, Hefei, China, in 2015, and the M.S. degree in computer technology from the College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China, in 2019.

His research interests include hyperspectral image classification, pattern recognition, and machine learning.



Meng Xu (Member, IEEE) received the B.S. and M.E. degrees in electrical engineering from the Ocean University of China, Qingdao, China, in 2011 and 2013, respectively, and the Ph.D. degree from The University of New South Wales, Canberra, ACT, Australia, in 2017.

She is an Associate Research Fellow with the College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China. Her research interests include cloud removal and remote sensing image processing.



Qiang Huang received the B.S. degree in electrical engineering from Shenzhen University, Shenzhen, China, in 1999, and the Ph.D. degree in electrical engineering and electronics from the University of Liverpool, Liverpool, U.K., in 2004.

He worked as a Research Associate with the University of Leicester, Leicester, U.K., from 2003 to 2004. He was teaching as a Lecturer, an Associate Professor, and a Professor with Shenzhen University. He has published more than 36 research articles. His research interests include distributed embedded

real-time systems, digital image processing, supercomputer systems, and cloud computing.



Jun Zhou (Senior Member, IEEE) received the B.S. degree in computer science and the B.E. degree in international business from the Nanjing University of Science and Technology, Nanjing, China, in 1996 and 1998, respectively, the M.S. degree in computer science from Concordia University, Montreal, QC, Canada, in 2002, and the Ph.D. degree in computing science from the University of Alberta, Edmonton, AB, Canada, in 2006.

He was a Research Fellow with the Research School of Computer Science, The Australian National University, Canberra, ACT, Australia, and a Researcher with the Canberra Research Laboratory, National Information and Communications Technology Australia, Canberra, Australia. In 2012, he joined the School of Information and Communication Technology, Griffith University, Nathan, QLD, Australia, where he is currently an Associate Professor. His research interests include pattern recognition, computer vision, and spectral imaging and their applications in remote sensing, and environmental informatics.



Xiuping Jia (Senior Member, IEEE) received the B.Eng. degree from the Beijing University of Posts and Telecommunications, Beijing, China, in 1982, and the Ph.D. degree in electrical engineering from The University of New South Wales, Canberra, ACT, Australia, in 1996.

Since 1988, she has been with the School of Information Technology and Electrical Engineering, The University of New South Wales, where she is a Senior Lecturer. She is also a Guest Professor with Harbin Engineering University, China, and an

Adjunct Researcher with the China National Engineering Research Center for Information Technology in Agriculture. She has coauthored the remote sensing textbook titled *Remote Sensing Digital Image Analysis* [Springer-Verlag, third edition (1999) and fourth edition (2006)]. Her research interests include remote sensing, image processing, and spatial data analysis.

Dr. Jia is a Subject Editor for the *Journal of Soils and Sediments* and an Associate Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING.