# Hyperspectral Image Segmentation based on Graph Processing over Multilayer Networks

Songyang Zhang, Qinwen Deng, and Zhi Ding, *Fellow, IEEE,*

***Abstract*—Hyperspectral imaging is an important sensing technology with broad applications and impact in areas including environmental science, weather, and geo/space exploration. One important task of hyperspectral image (HSI) processing is the extraction of spectral-spatial features. Leveraging on the recent-developed graph signal processing over multilayer networks (M-GSP), this work proposes several approaches to HSI segmentation based on M-GSP feature extraction. To capture joint spectral-spatial information, we first customize a tensor-based multilayer network (MLN) model for HSI, and define a MLN singular space for feature extraction. We then develop an unsupervised HSI segmentation method by utilizing MLN spectral clustering. Regrouping HSI pixels via MLN-based clustering, we further propose a semi-supervised HSI classification based on multi-resolution fusions of superpixels. Our experimental results demonstrate the strength of M-GSP in HSI processing and spectral-spatial information extraction.**

***Index Terms*—Hyperspectral image classification, feature extraction, graph signal processing, spectral clustering.**

## I. Introduction

HYPERSPECTRAL imaging is an analytical technique operating on images at different wavelengths for given geographical areas [1], [2], [3]. Exploiting a wealth of spectral-spatial information, hyperspectral images (HSIs) have seen broad applications in areas such as urban mapping, environment management, crop analysis and food safety inspection [4], [5], [6]. Many such applications often require to label each spatial position shown as an image pixel. Thus, hyperspectral image segmentation has emerged as an important field in HSI analysis [7]. Given the corresponding spectral feature vectors for each position in HSI, HSI segmentation aims to partition image pixels into different feature groups (or regions). Typical types of HSI segmentation include: 1) supervised (semi-supervised) classification given the labels of a training sample set, and 2) unsupervised segmentation without prior knowledge on the labels of pixels. In this work, we investigate the application of multilayer network graph signal processing (M-GSP) in both unsupervised clustering and semi-supervised classification as conceptually illustrated in Fig. 1 and Fig. 3, respectively.

Supervised HSI classification approaches form an important category of HSI analysis in remote sensing and have shown strong performance on standard datasets [8]. One typical approach is to first reduce the feature dimension before applying a classifier, such as support vector machines (SVM) for segmentation [9], [10]. To remove redundancy and to reject noise across the spectral bands, different feature selection [11] and feature extraction [4] methods have shown successes. Among them, we see familiar concepts that leverage principal component analysis (PCA) [12], [13], [14], [15], [16], independent component analysis (ICA) [17], and manifold learning [18], [19], among others. With the recent advances in machine learning, newly proposals such as deep learning [20], active learning [21], and tensor learning [22], have also demonstrated robust performance of HSI classification. Despite their successes, HSI classification algorithms require a significant number of high quality samples for model optimization and can be costly in many real-life scenarios [23]. On the other hand, unsupervised segmentation is also an important task in HSI analysis and has attracted significant coverage recently. Basic Centroid-based clustering methods, such as $k$-means [24] and fuzzy $c$-means [25], aim to minimize intra-cluster distance of samples within the clusters. Other clustering methods, including density-based methods [26] and biological clustering methods [27], have also been adopted for unsupervised HSI segmentation.

In addition to the aforementioned approaches, graph-based methods have been gaining popularity in both supervised (semi-supervised) and unsupervised HSI segmentation, owing to their power in revealing underlying structures among different pixels. Representing HSIs as graphs, spectral clustering [28], [23] can be developed for the unsupervised HSI segmentation. Compared to the centroid-based clustering and density-based methods, graph-based spectral clustering can provide a more general similarity model for pixels and perform well in unsupervised HSI segmentation. Similarly, graph convolutional network (GCN) [29], [30], [31], [32] has also demonstrated its strength in (semi-) supervised HSI segmentation, especially under the condition of limited ground-truth hyperspectral sets [33], by exploring the underlying structures among the pixels.

Despite these reported successes, most of the graph-based methods focus primarily on spatial geometry and consider a rather stationary graph connections for all spectral bands. Such limitation fails to explore different spectrum features of each individual band. For example, Fig. 2 shows that different bands may display different distributions of the pixel volumes, which could mean different graph structures. To compensate, some works would represent each band with an individual graph while neglecting the inter-frame correlations [34]. For these reasons, the unresolved challenge is to understand how

S. Zhang, Q. Deng, and Z. Ding are with Department of Electrical and Computer Engineering, University of California, Davis, CA, 95616. (E-mail: sydzhang@ucdavis.edu, mrdeng@ucdavis.edu, and zding@ucdavis.edu)

This material is based upon work supported by the National Science Foundation under Grants No. 2029848 and 2002937.

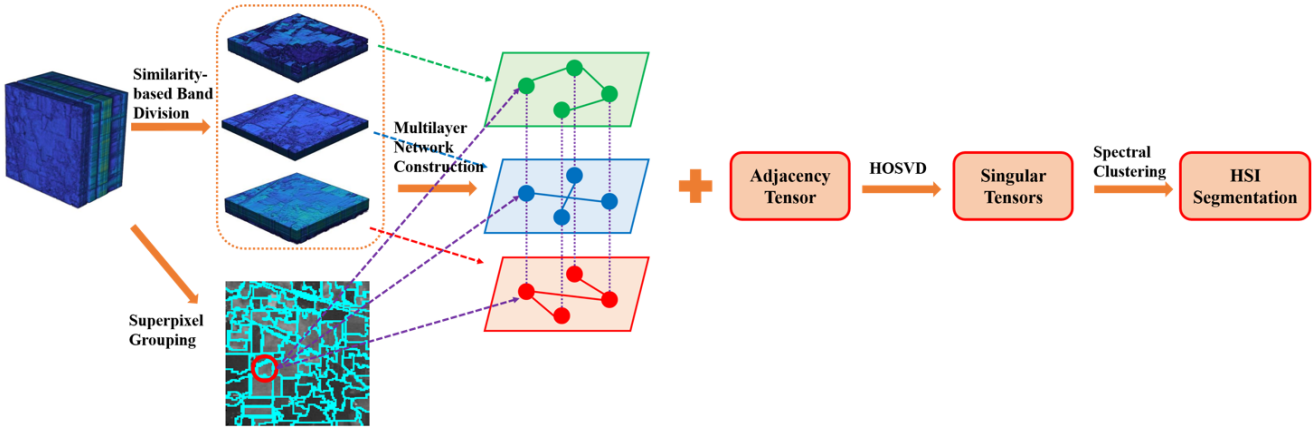
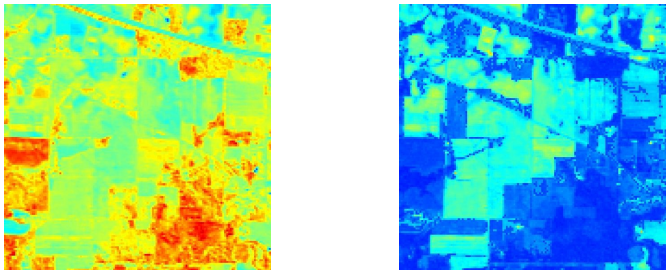



Fig. 1. Scheme of MLN-based Unsupervised HSI Segmentation.



(a) (b)

Fig. 2. Examples of Volumes for Different Bands in the Indian Pines Dataset.

to capture such heterogeneous spectral-spatial structure in an integrative manner jointly instead of individually.

In addition to the limitations of geometric models in general, traditional graph-based methods may also suffer from the large number of HSI pixels to process [23]. Generally, superpixels are constructed to control complexity. However, to achieve a better performance, one needs to determine the resolution of superpixels in advance and also decide how to form superpixels. In [16], different resolutions of superpixels are combined after the individual processing for each of them, which presents the ensuing challenges on the efficient fusion of the segmented results from different resolutions.

In view of the aforementioned challenges, this work introduces a new approach of graph signal processing over multilayer networks (M-GSP) to HSI processing. M-GSP [35], [36] is a tensor-based framework that generalizes traditional graph signal processing (GSP) [37] to process the heterogeneous graph structures across different layers of graphs. Since the spatial positions (pixels) are the same for all hyperspectral bands, we can model HSI data as a multilayer network with the same number of nodes in each layer (also known as multiplex network). The superpixels serve as nodes and each spectrum frame forms one layer as shown in Fig. 1. Next, we develop M-GSP spectral analysis to extract features for HSI segmentation. In this work, we investigate the applications of M-GSP in the feature extraction for both unsupervised and (semi-) supervised HSI segmentation. We summarize our contributions in this work as follows:

- We propose a multilayer network (MLN) model, together with an alternative singular space for the HSI datasets.
- We develop an unsupervised HSI segmentation method based on the M-GSP spectral clustering.
- We develop an MLN-based method for semi-supervised HSI segmentation, which combines multi-resolution information.
- We further propose several novel schemes for decision fusion of the results from different resolutions of superpixels.

We test all the algorithms in the standard *Indian Pines* dataset, *Pavia University* dataset and *Salinas* dataset. The experimental results demonstrate the strength of M-GSP in representing the spectral-spatial structures in HSI, and the efficiency of the proposed HSI segmentation algorithms.

We organize the rest of this manuscript as follows. Starting with a brief introduction of M-GSP in Section II, we introduce the multilayer network construction for HSI datasets together with the introduction of unsupervised HSI segmentation based on M-GSP spectral clustering in Section III. Next, we introduce the M-GSP framework of semi-supervised HSI classification in Section IV, where we also propose several methods for the decision fusion of different resolutions. In Section V, we present details on the experiments and results of the proposed methods in several standard datasets. Finally, we summarize our works in Section VI.

## II. Fundamentals of M-GSP

In this section, we briefly introduce the core concepts of M-GSP used in this work.

A multilayer network $\mathcal{M}$ with $M$ layers and $N$ nodes in each layer can be viewed as projecting $N$ virtual entities (nodes) onto $M$ layers. The resulting $M$ layers form a multilayer "network" $\mathcal{M}$ that be represented by a fourth-order adjacency tensor $\mathbf{A} \in \mathbb{R}^M \times \mathbb{R}^N \times \mathbb{R}^M \times \mathbb{R}^N$ defined as

$$\mathbf{A} = (A_{\alpha i \beta j}), \quad 1 \leq \alpha, \beta \leq M, 1 \leq i, j \leq N, \qquad (1)$$

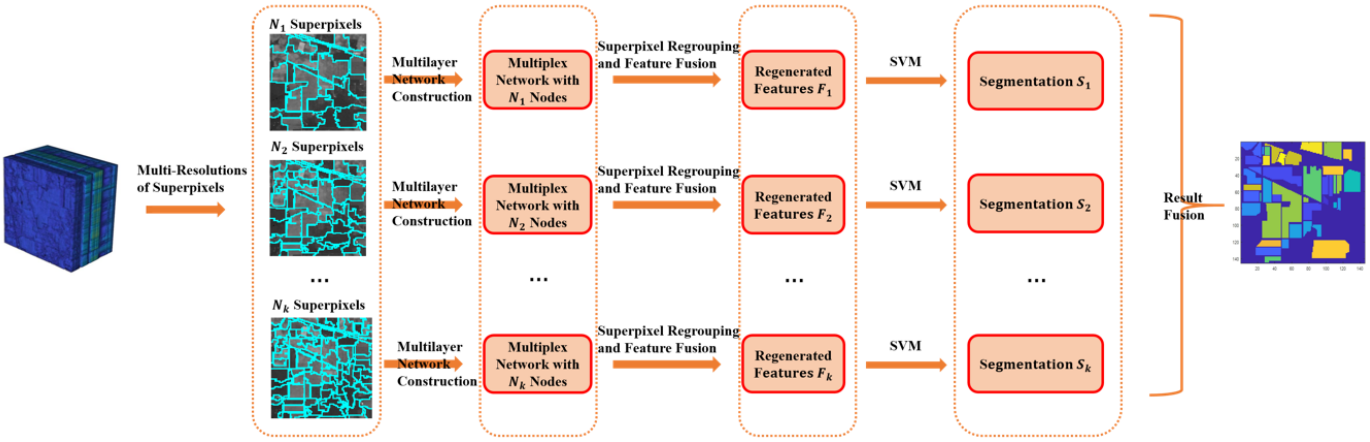



Fig. 3. Scheme of MLN-based Semi-supervised HSI Segmentation.

where each entry $A_{\alpha i \beta j}$ of $\mathbf{A}$ indicates the edge (link) intensity between entity $j$'s projection in layer $\beta$ and entity $i$'s projection in layer $\alpha$ [38]. Generally, we use bold uppercase letters to denote tensors, i.e., $\mathbf{A} \in \mathbb{R}^{I_1} \times \cdots \times \mathbb{R}^{I_Q}$ represents a $Q$th-order tensor with $I_k$ being the dimension of the $k$th order. $A_{i_1 \cdots i_Q}$ denotes the entry of $\mathbf{A}$ in position $(i_1, i_2, \cdots, i_Q)$ with $1 \leq i_k \leq I_k$. In addition to adjacency tensor, Laplacian tensor can be also defined similar to that in normal graphs. Interested reader could refer to [35] for more details.

Since we construct the superpixels from all spectrum images in a HSI dataset, each superpixel can be viewed as an entity to be projected into different spectrum bands. We can then represent the spectral-spatial structure intuitively via the adjacency tensor in Eq. (1). More details of the multilayer construction will be discussed in Section III.

With the adjacency tensor $\mathbf{A} \in \mathbb{R}^M \times \mathbb{R}^N \times \mathbb{R}^M \times \mathbb{R}^N$, it can be decomposed via higher-order singular value decomposition (HOSVD) [39] as

$$\mathbf{A} \approx \mathbf{S} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \times_3 \mathbf{U}^{(3)} \times_4 \mathbf{U}^{(4)}, \tag{2}$$

where $\times_n$ is the n-mode product [40] and $\mathbf{U}^{(n)} = [\mathbf{u}_1^{(n)} \quad \mathbf{u}_2^{(n)} \quad \cdots \quad \mathbf{u}_{I_n}^{(n)}]$ is a unitary $(I_n \times I_n)$ matrix, with $I_1 = I_3 = M$ and $I_2 = I_4 = N$. $\mathbf{S}$ is a $(I_1 \times I_2 \times I_3 \times I_4)$-tensor of which the subtensor $\mathbf{S}_{i_n}$ obtained by freezing the $n$th index to $\alpha$:

- $< \mathbf{S}_{i_n=\alpha}, \mathbf{S}_{i_n=\beta} >= 0$ where $\alpha \neq \beta$.
- $||\mathbf{S}_{i_n=1}|| \geq ||\mathbf{S}_{i_n=2}|| \geq \cdots \geq ||\mathbf{S}_{i_n=I_n}|| \geq 0$.

The Frobenius-norms $\sigma_i^{(n)} = ||\mathbf{S}_{i_n=i}||$ is the $n$-mode singular value, with corresponding singular vectors in $\mathbf{U}^{(i)}$. Since the representing tensor is partial symmetric in the undirected MLN, there are two modes of singular spectrum, i.e., $(\gamma_\alpha, \mathbf{f}_\alpha)$ for mode $1, 3$, and $(\sigma_i, \mathbf{e}_i)$ for mode $2, 4$. More specifically, $\mathbf{U}^{(1)} = \mathbf{U}^{(3)} = (\mathbf{f}_\alpha)$ characterizes the features of layers and $\mathbf{U}^{(2)} = \mathbf{U}^{(4)} = (\mathbf{e}_i)$ characterizes the entities.

If the singular vectors are included in $\mathbf{F}_s = [\mathbf{f}_1 \cdots \mathbf{f}_M] \in \mathbb{R}^M \times \mathbb{R}^M$ and $\mathbf{E}_s = [\mathbf{e}_1 \cdots \mathbf{e}_N] \in \mathbb{R}^N \times \mathbb{R}^N$, the MLN singular transform (M-GST) for a MLN signal $\mathbf{s} \in \mathbb{R}^{M \times N}$ can be defined as

$$\check{\mathbf{s}}_L = \mathbf{F}_s^{\mathrm{T}} \mathbf{s} \mathbf{E}_s \in \mathbb{R}^M \times \mathbb{R}^N. \tag{3}$$

Here, we mainly focus on fundamentals of singular analysis of the undirected multilayer networks. For more details on other concepts, such as MLN Fourier transform, M-GSP filter design, sampling theory and stationary process, readers are referred to [35], [36].

## III. Unsupervised HSI Segmentation Based on M-GSP Spectral Clustering

In this section, we introduce the construction of MLN models for HSI datasets, before proposing an unsupervised segmentation approaches based on M-GSP spectral clustering.

### A. Superpixel Segmentation for HSI

Before venturing into the M-GSP analysis, we first introduce the superpixel segmentation for HSI. In traditional graph-based HSI analysis, image pixels act as nodes and their pairwise distances are calculated to form a graph [23]. However, given a large number of pixels, it becomes inefficient and sometimes impossible to implement full graph-based analysis for pixel-based HSIs. Practically, since pixels within a small region may share similar features, grouping neighboring pixels into superpixels could be a more practical way for graph construction.

In general, a suitable superpixel segmentation algorithm for HSI classification should exhibit low computation complexity and accurate detection of the object boundaries [16]. One category of superpixel segmentation in HSI uses image features such brightness, color and texture cues, to estimate the location of segment boundaries. In [41], a superpixel estimation is developed by adopting the ultrametric contour map (UCM) approaches to the hyperspectral volumes. In [42], the band smoothness is jointly considered with general features to group pixels. Graph-based segmentation approaches are

also common in superpixel segmentation [43]. In [44], an eigen-based solution to normalized cuts (NCuts) is used for superpixel group. However, such eigen-based methods tend to suffer from time-consuming graph construction and matrix decomposition.

For more efficient superpixel segmentations, our work in this manuscript considers the entropy rate superpixel segmentation (ERS) suggested by, e.g., [16]. In ERS [45], a dataset is modeled as a graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, in which the pixels serve as the nodes $\mathcal{V}$ and their pairwise similarities are represented by edges $\mathcal{E}$. Next, a subgraph $\mathcal{A} = \{\mathcal{V}, \mathcal{L}\}$ is formed by choosing a subset of edges $\mathcal{L} \subseteq \mathcal{E}$, such that $\mathcal{A}$ consists of fewer connected components. To obtain such a subgraph, the problem can be formulated as

$$\mathcal{L}^* = \arg_{\mathcal{L}} \max Tr\{H(\mathcal{L}) + \alpha T(\mathcal{L})\} \tag{4}$$

$$s.t. \quad \mathcal{L} \subseteq \mathcal{E}, \tag{5}$$

in which entropy rate term $H(\mathcal{L})$ favors more compact clusters, and regularizing term $T(\mathcal{L})$ punishes large cluster size. Based on the objective function of Eq. (4), a greedy algorithm can be implemented to solve the problem [46].

### *B. Multilayer Network Construction for HSI Datasets*

With superpixel-represented HSI, we now begin multilayer network construction. Consider Fig. 1. An HSI $\mathbf{X} \in \mathbb{R}^K \times \mathbb{R}^N$, containing $K$ spectrum frames and $N$ superpixels, can be modeled by a multilayer network with $M$ layers and $N$ nodes in each layer. Specifically, the MLN consists of the following attributes

- *Layers*: To construct a MLN, we define layers based on the spectrum bands. Since different spectrum frames may share similar features, we first divide the bands into $M$ clusters, i.e., $\mathbf{X}_i \in \mathbb{R}^{K_i} \times \mathbb{R}^N$, $i = 1, \cdots, M$ and $\sum_{i=1}^{M} K_i = K$. Next, each cluster serves as one layer in the multilayer network. Various clustering methods can generate features $\mathbf{X}_i$ for layer $i$. For example, one can divide spectrum band based on a range of wavelength. To capture correlation across different bands more efficiently, the $k$-means clustering is applied for band division.
- *Nodes*: In M-GSP based HSI processing, the superpixels act as virtual entities. By projecting $N$ superpixels into $M$ layers, we form a multilayer network with $M$ layers and $N$ nodes in each layer. We define MLN signals as the divided attributes of each superpixel, i.e., $\mathbf{X}_{i,j} \in \mathbb{R}^{K_i}$ for the superpixel $j$'s projected node in layer $i$.
- *Interlayer connections*: For interlayer connections, each projected node is connected to its counterparts in other layers, i.e., fully connected for all the projected nodes of the same superpixel. As a result, we obtain corresponding entries of the adjacency tensor $\mathbf{A} \in \mathbb{R}^M \times \mathbb{R}^N \times \mathbb{R}^M \times \mathbb{R}^N$ for the superpixel $i$ defined as

$$A_{\alpha i \beta i} = \begin{cases} 1, & \alpha \neq \beta; \\ 0, & otherwise, \end{cases} \tag{6}$$

  where each entry indicates link presence.
- *Intralayer connections*: For the intralayer connections, we calculate the weights between the projected nodes of entities $i$ and $j$ in layer $\alpha$ based on the localized Gaussian distance as follows:

$$A_{\alpha i \alpha j} = \begin{cases} e^{-\frac{||\mathbf{X}_{\alpha,i} - \mathbf{X}_{\alpha,j}||_2^2}{\sigma^2}}, & \text{dis}_1(\mathbf{X}_{\alpha,i}, \mathbf{X}_{\alpha,j}) < p, \\ & \text{dis}_2(p(\alpha,i), p(\alpha,j)) < q; \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

  where $\sigma$, $p$ and $q$ are design parameters and $p(\alpha, i)$ is the position of the superpixel $i$ in layer $\alpha$.

Beyond the traditional Gaussian distance [47], our intralayer connections consider two conditions for determining the presence of links: 1) features between two nodes should be similar; and 2) two connected superpixels should be in a localized region in the HSI. The first condition ensures the similarity of connected nodes while the second condition emphasizes geometric closeness in the HSI. For an initial setup of the parameters, we define $\text{dis}_1$ using $\ell_2$-norm, and define $\text{dis}_2$ as the Euclidean distance between the respective centroids of two superpixels. In terms of design parameters, we set $p$ as the mean of all pairwise similarities and tune the parameters $q$, $\sigma$ based on the specific dataset.

### *C. MLN-based Spectral Clustering for HSI Segmentation*

Spectral clustering is an efficient method for unsupervised HSI segmentation [23]. Modeling HSI by a normal graph before spectral clustering, significant improvement is possible owing to its power in capturing the underlying structures [48], [49]. However, by representing HSI by a single-layer graph, distinction of individual spectrum bands might be overlooked. To capture the heterogeneous spectral-spatial structure in HSI, we propose to segment the HSI based on the M-GSP spectral clustering as follows.

Consider an HSI $\mathbf{X} \in \mathbb{R}^K \times \mathbb{R}^N$ with $K$ spectrum frames and $N$ superpixels. To implement the M-GSP spectral clustering, we first construct an $M$-layer network the with adjacency tensor $\mathbf{A} \in \mathbb{R}^M \times \mathbb{R}^N \times \mathbb{R}^M \times \mathbb{R}^N$ according to Eqs. (6)-(7). We then apply HOSVD to obtain the singular tensors $\mathbf{F}_s = [\mathbf{f}_1 \cdots \mathbf{f}_M] \in \mathbb{R}^M \times \mathbb{R}^M$ and $\mathbf{E}_s = [\mathbf{e}_1 \cdots \mathbf{e}_N] \in \mathbb{R}^N \times \mathbb{R}^N$ to characterize the bands and superpixels, respectively, according to Eq. (2). Since we aim to segment superpixels into meaningful clusters, we focus on the entity-wise spectrum $\mathbf{E}_s$. Arranging $\mathbf{e}_i$ in the descending order of its corresponding singular value $\sigma_i$, i.e.,

$$\sigma_i = ||\mathbf{S}_{i_2=i}||, \tag{8}$$

where $\mathbf{S}_{i_2=i} \in \mathbb{R}^M \times \mathbb{R}^1 \times \mathbb{R}^M \times \mathbb{R}^N$ is the subtensor of the core tensor $\mathbf{S}$ in Eq. (2) by freezing the second order $i_2 = i$, we pick the first $P$ singular vectors to preserve the most critical information for HSI based on a largest gap among the singular values. Clustering based on the $P$ selected singular vectors and labeling each pixel within the superpixel, we can obtain a segmentation of the given HSI. The major process of MLN-based unsupervised segmentation is provided in Algorithm 1.

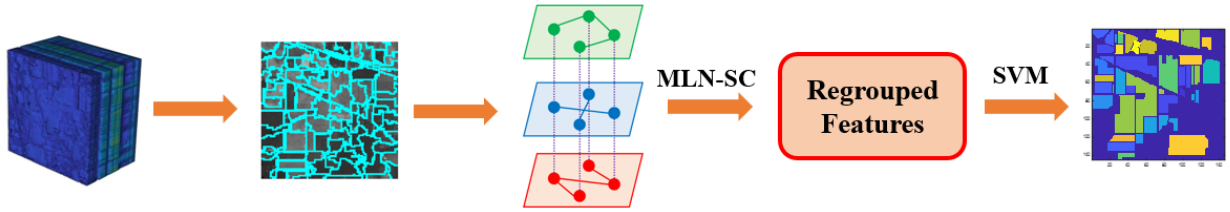



Fig. 4. Scheme of Single-Resolution Segmentation.

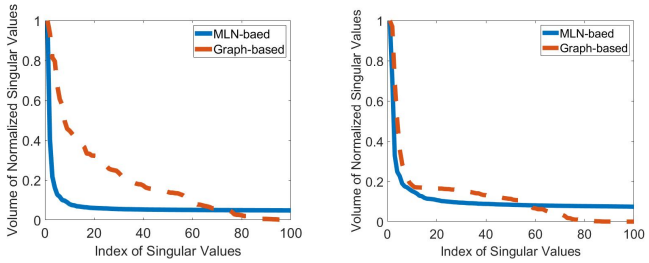
**Algorithm 1** MLN-based Unsupervised HSI Segmentation (MLN-SC)

1: **Input**: HSI $\mathbf{I} \in \mathbb{R}^K \times \mathbb{R}^T$ with $K$ frames and $T$ pixels, and the number of clusters $Q$;
2: Construct $N$ superpixels for HSI based on ERS algorithm as $\mathbf{X} \in \mathbb{R}^K \times \mathbb{R}^N$;
3: Divide bands into $M$ clusters based on $k$-means clustering;
4: Construct an $M$-layer network with adjacency tensor $\mathbf{A} \in \mathbb{R}^M \times \mathbb{R}^N \times \mathbb{R}^M \times \mathbb{R}^N$ as Eq. (6) and Eq. (7);
5: Implement HOSVD on $\mathbf{A}$ to obtain the entity-wise singular tensors $\mathbf{E}_s = [\mathbf{e}_1 \cdots \mathbf{e}_N] \in \mathbb{R}^N \times \mathbb{R}^N$ in the descending order of the 2-mode singular values $\sigma_i$ as Eq. (8);
6: Select the first $P$ singular tensors as $\mathbf{P}_k = [\mathbf{e}_1, \cdots, \mathbf{e}_P] \in \mathbb{R}^N \times \mathbb{R}^P$ based on the largest singular gap;
7: Cluster the rows of $\mathbf{P}_k$ into $Q$ groups based on $k$-means clustering;
8: Cluster the superpixel $i$ into group $j$ if the $i$-th row of $\mathbf{P}_k$ is in group $j$;
9: Label the pixels as the same cluster of its superpixel;
10: **Output**: Segmented HSI.

### D. Discussion

Before we dive further to develop MLN-based semi-supervised HSI segmentation, we provide a short conceptual discussion on the M-GSP singular tensors. Within the context of GSP, how to capture layer-wise and entity-wise information efficiently has become a significant topic of research due to growing interests in spatial-temporal datasets. The authors of [48] have proposed a two-step graph Fourier transform (GFT) to process spatial-temporal graphs, by applying GFT in the spatial domain (intralayer) first and then in the temporal domain (interlayer). However, this approach does not consider the structural correlations between inter- and intra- layers. The authors of [49] developed a joint time-vertex Fourier transform (JFT) by implementing GFT and discrete Fourier transform (DFT) consecutively. However, the DFT block limits the structure of interlayer connections to a path graph. More recently, the authors of [50] proposed a tensor-based multi-way graph signal processing framework (MWGSP) on the product graph. MWGSP constructs separate factor graphs for each mode of a tensor-represented signals and defines a joint spectrum that combines spectra of all factor graphs. One limitation is that the combination process requires a homogeneous structure for each graph layer and does not accommodate potentially different band-wise features in HSI.

By contrast, since we apply HOSVD in M-GSP for achieving tensor-based representation, our M-GSP framework is able to jointly process interlayer and intralayer connections of HSI. Even when focusing on entity-wise singular tensors in the clustering process, we are still able to incorporate layer-wise information. To better understand the property of MLN-based singular tensors, we graphically illustrate the distribution of singular values compared to a graph-based model in Fig. 5. As shown, the energy of MLN-based singular values are more concentrated in the first few dominant singular vectors in low frequency when compared against graph-based singular values. This energy concentration indicates a more convenient and low degradation implementation of spectral clustering within our proposed M-GSP framework.




(a) Indian Pines (b) Salinas

Fig. 5. Distribution of Entity-wise Singular Values.

## IV. Semi-Supervised HSI Classification Based on M-GSP Feature Extraction

In this section, we introduce the semi-supervised HSI segmentation based on M-GSP feature extraction.

### A. Single-Resolution of MLN-based HSI Segmentation

We start with the single-resolution of the superpixels. In the superpixel-based classification, superpixel resolutions affect the final performance: finer resolution could capture more details whereas coarse resolution captures the global information more efficiently. To benefit from both fine and coarse resolutions, we introduce the MLN-based spectral clustering on the fine resolution to regroup superpixels into a coarse resolution (the number of regrouped superpixels should still be larger than the number of classes) and use the regrouped features as classifier inputs. Here, we apply SVM to classify the regrouped superpixels. The concept of our single-resolution HSI segmentation (MLN-SRC) is illustrated by Fig. 4, and the major steps are described in Algorithm 2.

**Algorithm 2** Single-Resolution HSI Segmentation (MLN-SRC)

1: **Input**: HSI $\mathbf{I} \in \mathbb{R}^K \times \mathbb{R}^T$ with $K$ frames and $T$ pixels;
2: Construct $N$ superpixels for HSI based on ERS algorithm as $\mathbf{X} \in \mathbb{R}^K \times \mathbb{R}^N$;
3: Divide bands into $M$ clusters based on $k$-means clustering;
4: Construct an $M$-layer network with adjacency tensor $\mathbf{A} \in \mathbb{R}^M \times \mathbb{R}^N \times \mathbb{R}^M \times \mathbb{R}^N$ as Eq. (6) and Eq. (7);
5: Implement MLN-based spectral clustering to regroup the superpixels into $D$ clusters, and combine the features of pixels within the same cluster as the regrouped features, i.e., $\mathbf{X}_R \in \mathbb{R}^K \times \mathbb{R}^D$;
6: Input $\mathbf{X}_R$ into SVM for the classification of regrouped superpixels;
7: Label the pixels as the same class of its superpixel;
8: **Output**: Segmented HSI.

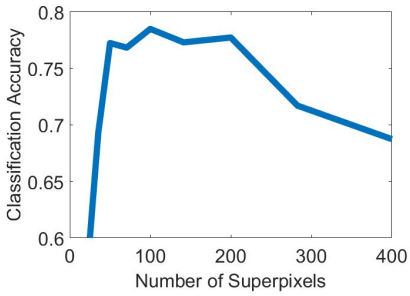
The benefits of the proposed MLN-SRC include:

- Against a singe resolution of coarse superpixels, the MLN-SRC implements an analysis step over a fine resolution and is capable of capturing detailed features. Against a single resolution of fine superpixels, the MLN-SRC substantially reduces pixel number and enhances robustness of the feature inputs to the classifier. Too many superpixels may make the features less distinctive and over-segment the regions, whereas too few superpixels may lead to boundary ambiguity.
- Traditional graph-based superpixel segmentation only captures a single-layer structure. MLN-GSP regrouping could reveal additional feature information across the heterogeneous multi-band structures.
- In traditional superpixel segmentation, the distinct regions are usually labeled as different superpixels. However, in MLN-SRC, superpixels from different regions may have the same labels depending on clustering results. Thus, regrouped features can involve similar pixels that cover a large distance and potentially generate more features.
- MLN-SRC can be easily integrated with other feature extraction or selection algorithms. Dimension reduction techniques such as PCA and ICA can potentially improve the performance when applied on features and feature groups generated by MLN-SRC.

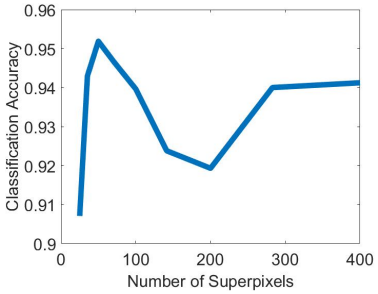
### *B. Multi-Resolution of MLN-based HSI Segmentation*

*1) Multi-Resolution Structure:* Although MLN-based spectral clustering can regroup small superpixels into larger ones and to benefit from both fine and coarse resolutions in MLN-SRC, the initial resolution setting of superpixels still affects final performance. As Fig. 6 shows, different initial resolutions can lead to different levels of accuracy. It is practically difficult to determine the optimal initial number of superpixels.

Similar to [16], we consider a multi-resolution structure of classification (MLN-MRC) shown as Fig. 3. In this framework, we examine several different initial resolutions of superpixels. MLN-SRC is applied to each initial resolution to regroup the superpixels with a same reduction ratio in group numbers, i.e., $70\%$ of initial superpixels. Applying SVM to classifying the multiple regrouped superpixels, we fuse the results from different initial resolutions in final segmentation. The algorithm is described in Algorithm 3.




(a) Indian Pines




(b) Salinas

Fig. 6. Accuracy of MLN-SRC over Different Initial Resolutions.

**Algorithm 3** Multi-Resolution HSI Segmentation (MLN-MRC)

1: **Input**: HSI $\mathbf{I} \in \mathbb{R}^K \times \mathbb{R}^T$ with $K$ frames and $T$ pixels;
2: Construct multiple resolutions of superpixels for HSI;
3: Construct different multilayer networks for each resolution of HSI;
4: Implement MLN-SRC for each resolution to obtain the sub-results;
5: Fuse all the sub-results for the final segmentation;
6: **Output**: Segmented HSI.

Although the multi-resolution structures have been considered in literature, MLN-MRC exhibits two major distinctions. First, we apply a novel MLN-based clustering algorithm to regroup the superpixels and generate new features for classification. Second, we provide several novel decision fusion strategies, based on both confidence score and graph structures to be discussed below.

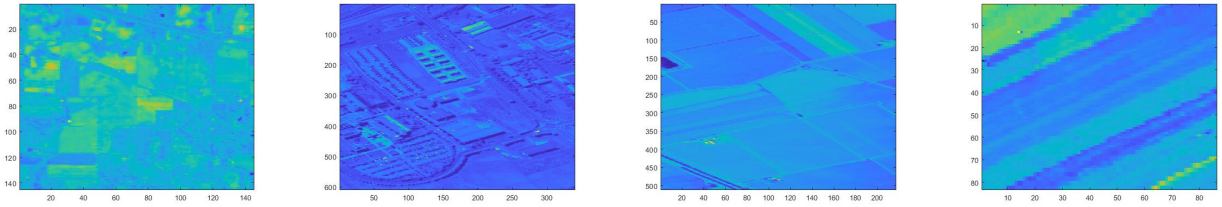
*2) Decision Fusion:* In [16], a decision fusion scheme based on majority voting (MV) is proposed. In this method, the label $l$ of a specific pixel is determined by

$$l = \arg\max_i \sum_{j=1}^{C} w_j \cdot \delta(l_j), \tag{9}$$

where $C$ is the number of distinct resolutions, $l_j$ is the label of the pixel in resolution $j$, $w_j$ is the voting strength, and $\delta(l_j) = 1$ if $l_j = i$; otherwise, $\delta(l_j) = 0$. Note that a basic majority voting based on equal strength $w_j = C^{-1}$ applies

TABLE I
STATISTICS OF DIFFERENT HSI DATASETS

| HSI | Pixel Size | # of Spectrum Bands | # of Classes | # of Labeled Samples |
|---|---|---|---|---|
| IndianP | 145×145 | 200 | 16 | 10249 |
| PaviaU | 610×340 | 103 | 9 | 42776 |
| Salinas | 512×217 | 204 | 16 | 54129 |
| SalinasA | 83×86 | 204 | 6 | 5348 |



(a) Indian Pines (b) Pavia University (c) Salinas (d) SalinasA

Fig. 7. Mean of HSI over Spectrum Dimension.

the same strength to different resolutions but would ignore the difference of multiple resolutions. To improve decision fusion, we introduce several novel strategies for the decision fusion.

- *Validation Accuracy* (VA): One intuitive way to design decision strength is based on validation accuracy. Here, we can apply the validation accuracy directly as the weighting strength $w_j$ for resolution $j$ to fuse the decision according to Eq. (9).
- *Decision Value* (DV): As one alternative, the decision probability for each class of pixels can be used as the weight. In multi-class SVM, the predicted label is determined according to the decision value $\mathbf{p} \in \mathbb{R}^C$, where $C$ is the number of classes [53]. Let $\mathbf{p}_{ij}$ be the decision value of pixel $i$ in $j$th resolution. We set the weight of $l_{ij}$ to

$$w_{ij} = \max_k \mathbf{p}_{ij}(k). \tag{10}$$

  Unlike validation accuracy which is the same for all pixels in each resolution, this weight based on decision value may vary even for pixels at the same resolution.
- *Graph Total Variation* (TV): Graph-based metrics can serve as weights. For a robust setup of superpixels, signals should be smooth and exhibit stable underlying graph structure. To this end, we introduce graph-based total variation to measure smoothness. Given a superpixel segmentation $j$ of a HSI with $N$ superpixels and $K$ spectrum frames, we regenerate the features of each superpixel by averaging all pixels within. We then construct a single-layer graph based on Gaussian distance to measure similarity between different superpixels. Defining a Laplacian matrix by $\mathbf{L} = \mathbf{D} - \mathbf{A}$ where $\mathbf{D}$ is the degree matrix and $\mathbf{A}$ is the adjacency matrix, the total variation of the feature signal $\mathbf{s}_p \in \mathbb{R}^N$ for the $p$th band frame over $\mathbf{L}$ is

$$\mathrm{TV}_p = ||\mathbf{s}_p - \frac{1}{|\lambda_{\max}|}\mathbf{L}\mathbf{s}_p||_2^2, \tag{11}$$

  where $\lambda_{\max}$ is the largest eigenvalue of $\mathbf{L}$. Total variation describes the propagation differences between two steps. A smaller total variation indicates a more smooth signal. With $K$ frames in total, final smoothness for resolution $j$ is defined as $\mathrm{SM}_j = \frac{1}{K}\sum_p \mathrm{TV}_p$. Since we prefer a larger weight for the smooth signal, the final weight of the resolution $j$ is defined as

$$w_j = e^{-SM_j}. \tag{12}$$

- *Von Neumann Entropy* (VN): The stability of the underlying graph structure can also indicate the confidence level of a specific superpixel resolution. In quantum theory [54], a pure state leads to a zero Von Neumann entropy. The entropy is larger if there are more mixed states in the system. Similarly, in our HSI analysis, since we prefer a stable system or a pure state on the underlying graph, the weight should be larger if the Von Neumann entropy is smaller. Consider the Von Neumann entropy introduced to evaluate the graph stability [54]. Similar to total variation, a Laplacian matrix $\mathbf{L}$ can be defined with adjacency matrix $\mathbf{A} = (a_{pq})$ for the $j$th resolution. First, define $c = 1/(\sum_{p,q} a_{pq})$ and rescale the Laplacian matrix

$$\mathbf{L}_G = c \cdot (\mathbf{D} - \mathbf{A}), \tag{13}$$

  We can define the weight for the $j$th resolution as

$$w_j = e^{-h_j}. \tag{14}$$

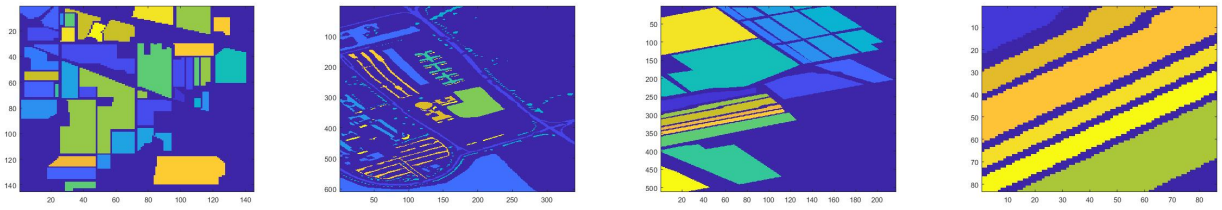
  based on the Von Neumann entropy

$$h_j = -Tr[\mathbf{L}_G \log_2 \mathbf{L}_G]. \tag{15}$$

Note that, here we provided several possible alternatives for the weights of decision fusion. The performances of the various proposed fusion strategies will be presented in Section V-C. We plan to investigate more M-GSP based approaches to result fusion in future works.

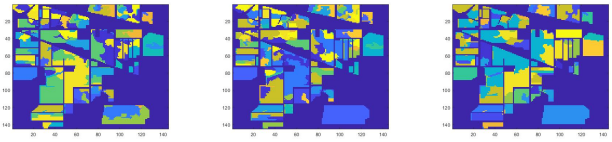
## V. EXPERIMENTAL RESULTS

We now test the performance of the proposed unsupervised and semi-supervised segmentation approaches in several well-known datasets to demonstrate the efficacy of M-GSP in



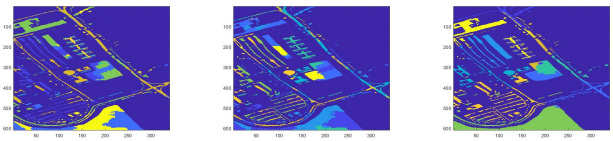
(a) Indian Pines (b) Pavia University (c) Salinas (d) SalinasA

Fig. 8. Ground Truth of Class Labels



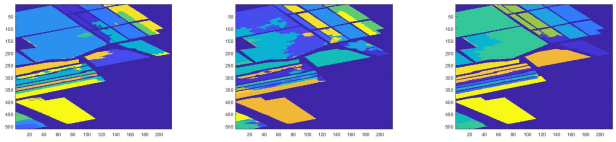
(a) K-means (b) GSP (c) M-GSP

Fig. 9. Segmented Results of Indian Pines.



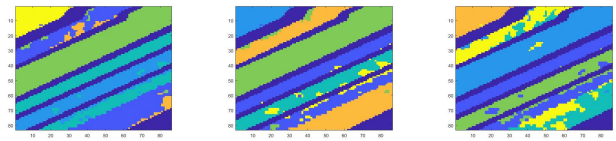
(a) K-means (b) GSP (c) M-GSP

Fig. 10. Segmented Results of Pavia University.



(a) K-means (b) GSP (c) M-GSP

Fig. 11. Segmented Results of Salinas.



(a) K-means (b) GSP (c) M-GSP

Fig. 12. Segmented Results of Salinas A.

HSI analysis. We also comparatively test the performance of various different fusion decisions.

### A. Dataset

We test the performances of the proposed methods based on four public HSI datasets accessible from website [1]. The first HSI is *Indian Pines* (IndianP) scene originally gathered by AVIRIS sensors over an agricultural field. The second HSI dataset is the *University of Pavia* (PaviaU) acquired by ROSIS sensor. Note that, some of the samples in PaviaU contain no information and have to be discarded before analysis. Two other HSIs used in the experiments are the *Salinas Scene* (Salinas) and *Salinas-A Scene* (SalinasA) datasets, which were collected by the 224-band AVIRIS sensor over Salinas Valley, California, and exhibit high spatial resolution. For each dataset, we have groundtruth classes for part of samples.

For these HSIs, Table I provides vital statistics and we provide visual illustration of the geometric plots in Fig. 7 and Fig. 8. Note that, Fig. 8 treats the unlabeled groundtruth samples as backgrounds with the same class label. Interested readers can find more information on the HSI datasets at the website[1].

### B. Unsupervised HSI Segmentation

In this part, we first test the performance of unsupervised HSI segmentation. Although our proposed MLN-based method is amenable to various sophisticated clustering approaches, we find it easier to demonstrate it by using a basic spectral clustering scheme. In this section, we mainly use comparison of basic methods, such as $k$-means clustering and GSP-based spectral clustering [35], to demonstrate the power of M-GSP in processing the HSI datasets. We will illustrate the benefits of multilayer network models and M-GSP in comparison with other more advanced methods in semi-supervised segmentation in Section V-C.

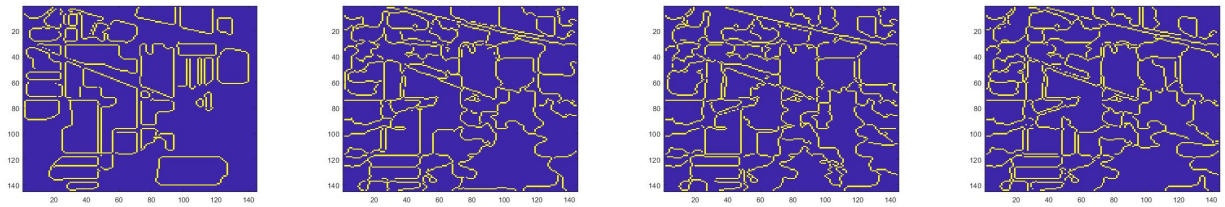
To validate the performance of different methods, we consider two experimental setups. In the first scenario, we carry out unsupervised segmentation on all data samples and evaluate the overall visualization results of labeled samples. In the second stage, we process all data samples but focus on detecting boundaries (edges) of each cluster, in terms of both visualization results and numerical accuracy.

For fair comparison, we segment the HSI into $N$ superpixels first before applying respective clustering algorithms thereupon. For the GSP-based method, we construct the graph $\mathbf{W} \in \mathbb{R}^{N\times N}$ using Gaussian distance

$$W_{ij} = \begin{cases} e^{-\frac{||\mathbf{s}_i - \mathbf{s}_j||_2^2}{\sigma^2}}, & ||\mathbf{s}_i - \mathbf{s}_j||_2^2 \leq \tau, \\ 0, & otherwise, \end{cases} \tag{16}$$

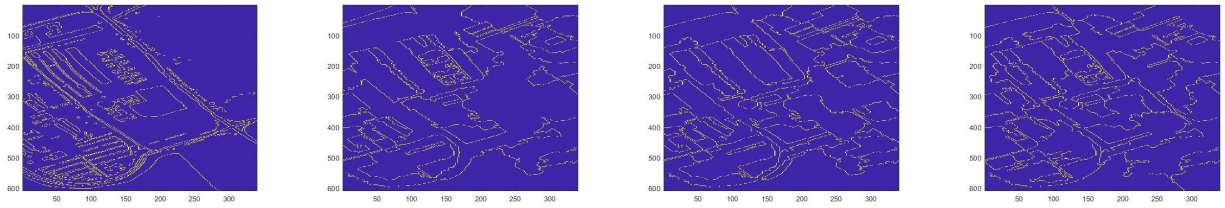
where $\mathbf{s}_i$ denotes the feature of $i$th superpixel. The threshold $\tau$ is set to the statistical mean of all pairwise distances among superpixels, and $\sigma$ is tunable according to specific datasets. For M-GSP based methods, we construct the multilayer network with $M = 10$ layers, and calculate the distance based on Eq. (6) and Eq. (7). The parameter $p$ in Eq. (7) is also set to the

[1]http://www.ehu.eus/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes



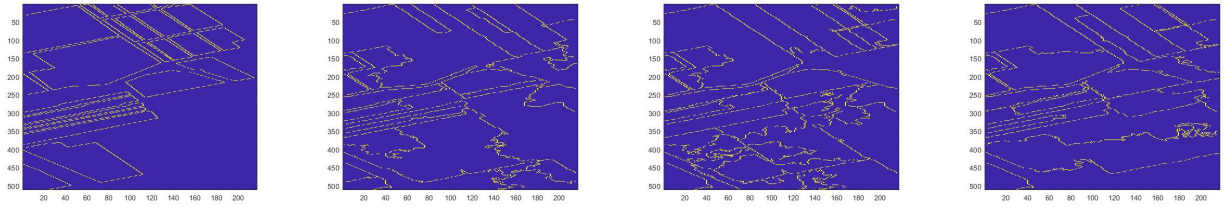
(a) Ground Truth (b) K-means (c) GSP (d) M-GSP

Fig. 13. Boundaries in Indian Pines.



(a) Ground Truth (b) K-means (c) GSP (d) M-GSP

Fig. 14. Boundaries in Pavia University.



(a) Ground Truth (b) K-means (c) GSP (d) M-GSP

Fig. 15. Boundaries in Salinas.

statistical mean of all pairwise intralayer feature distances, and $q = 100$ is used. We select the number of key spectra based on the largest gap of singular values. We summarize the results below.

*1) Visualization of HSI Segmentation:* By setting $N = 500$ for all tested HSIs, Fig. 9 - Fig. 12 present the visualization results of the three HSI segmentations using different algorithm in comparison to the groundtruth data. These results show that, in general, the MLN-based spectral clustering (M-GSP) displays more stable segmentations than the single layer GSP-based spectral clustering (GSP) as well as the $k$-means algorithm. Since it is harder to evaluate the details for too many classes, we can focus more on the exemplary *Salinas A* dataset. In particular, Fig. 8(d) shows six different groundtruch classes. However, both $k$-means and GSP-based method failed to detect the class No.6, marked by orange in Fig. 12(a) and marked by yellow in Fig. 12(b), respectively. However, in Fig. 12(c), the MLN-GSP method successfully identified all six classes and delivered results that are closer to groundtruth. Recall that, from Fig. 5, the MLN-based singular values are more concentrated, which provides the benefits of robustness in spectral clustering. Results from "Indian Pines" and "Pavia University" also show similarly stronger performance for M-GSP. These results collectively demonstrate improved efficiency of M-GSP in unsupervised HSI segmentation.

TABLE II
ACCURACY OF SEGMENTATION BOUNDARIES

| Data | K-means | GSP | M-GSP |
|---|---|---|---|
| IndianP | 0.8257 | 0.8298 | **0.8441** |
| Salinas | 0.9208 | 0.9285 | **0.9409** |
| PaviaU | 0.9070 | 0.9088 | **0.9255** |

*2) Boundary Segmentation:* Although unsupervised methods may generate meaningful segmentation different from groundtruth, we are still interested in how far the segmented results are from the true labels. Since it is inefficient to match all the clusters to the corresponding true labels, we focus on the boundaries of the segmentation. In the boundary detection, we set $N = 100$ and define the accuracy as

$$Acc = \sum_{i=1}^{T} 1(L_i = \hat{L}_i)/T, \tag{17}$$

where $T$ is the number of pixels in the HSI, $L_i$ is the true labels of edges for the $i$th pixel, $\hat{L}_i$ is the estimated labels, and $1(\cdot)$ denotes the indicator function. The boundary results are visualized in Fig. 13 - Fig. 15. Since we consider the unlabeled samples in the clustering process, there can be more details for the HSIs than their corresponding groundtruth, especially

TABLE III
OVERALL ACCURACY OF DIFFERENT HSI SEGMENTATION

| Data | TS/C | Raw | PCA | ICA | LPP | NPE | LPNPE | LDA | LFDA | SPCA | MSPCA | MLN-SRC | MLN-MRC |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| IndianP | 5 | 0.4488 | 0.4637 | 0.4521 | 0.5358 | 0.5368 | 0.6725 | 0.5995 | 0.5962 | 0.7734 | 0.7868 | 0.7432 | **0.7936** |
| | 10 | 0.5577 | 0.5572 | 0.5712 | 0.7041 | 0.7049 | 0.7645 | 0.6930 | 0.6491 | 0.8576 | 0.8712 | 0.8604 | **0.8773** |
| | 20 | 0.6381 | 0.6297 | 0.6441 | 0.8026 | 0.7987 | 0.8351 | 0.7656 | 0.7401 | 0.9390 | **0.9569** | 0.9226 | 0.9545 |
| PaviaU | 5 | 0.6459 | 0.6526 | 0.6658 | 0.7086 | 0.6835 | 0.7612 | 0.7243 | 0.7467 | 0.7439 | 0.7849 | 0.7561 | **0.8236** |
| | 10 | 0.7022 | 0.7015 | 0.7139 | 0.8129 | 0.8063 | 0.8255 | 0.8124 | 0.7895 | 0.8342 | **0.9167** | 0.8398 | 0.8896 |
| | 20 | 0.7585 | 0.7591 | 0.7665 | 0.8600 | 0.8569 | 0.8856 | 0.8500 | 0.8698 | 0.8938 | **0.9537** | 0.8697 | 0.9432 |
| Salinas | 5 | 0.8179 | 0.8187 | 0.8175 | 0.8523 | 0.8486 | 0.9209 | 0.8903 | 0.8883 | 0.9442 | 0.9500 | 0.9499 | **0.9588** |
| | 10 | 0.8524 | 0.8528 | 0.8574 | 0.8860 | 0.8899 | 0.9452 | 0.9146 | 0.8277 | 0.9678 | 0.9815 | 0.9614 | **0.9863** |
| | 20 | 0.8785 | 0.8779 | 0.8808 | 0.9061 | 0.9069 | 0.9589 | 0.9372 | 0.9356 | 0.9837 | 0.9904 | 0.9840 | **0.9915** |

TABLE IV
PEFORMANCE UNDER THE SAME SETS OF SUPERPIXEL RESOLUTIONS

| Data | IndianP | | | Salinas | | | PaviaU | | |
|---|---|---|---|---|---|---|---|---|---|
| TS/C | 5 | 10 | 20 | 5 | 10 | 20 | 5 | 10 | 20 |
| MSPCA | 0.6663 | 0.7476 | 0.9406 | 0.8521 | 0.9662 | **0.9934** | 0.7078 | 0.8447 | **0.9420** |
| MLN-MRC | **0.7094** | **0.8062** | **0.9455** | **0.9365** | **0.9738** | 0.9931 | **0.8218** | **0.8590** | 0.9334 |

for *Pavia University* and *Indian Pines*. However, we can still derive some benefits of M-GSP in boundary detection. As Fig. 15 shows, M-GSP can generate clearer edges, obviously on the top half image, while GSP and K-means appear to over-segment. We also present the accuracy defined in Eq. (17) as Table II. These results show that M-GSP performs better than K-means and GSP. They demonstrate the efficiency of MLN models in HSI analysis.

It is important to note that

- We do not claim M-GSP to be the best approaches for all scenarios. Without impractically requiring hindsight to fine-tune various parameters for each HSI dataset to generate "the best results", M-GSP delivers consistently strong and stable segmentation results for various HSI datasets by relying on some basic guidelines for selecting parameters.
- We present M-GSP as an alternative approach to HSI segmentation. Its scalability and practical tuning allow easier integration with other graph-based algorithms and deep learning neural networks.

### C. Semi-supervised HSI Segmentation

We next test M-GSP in (semi)-supervised HSI segmentation.

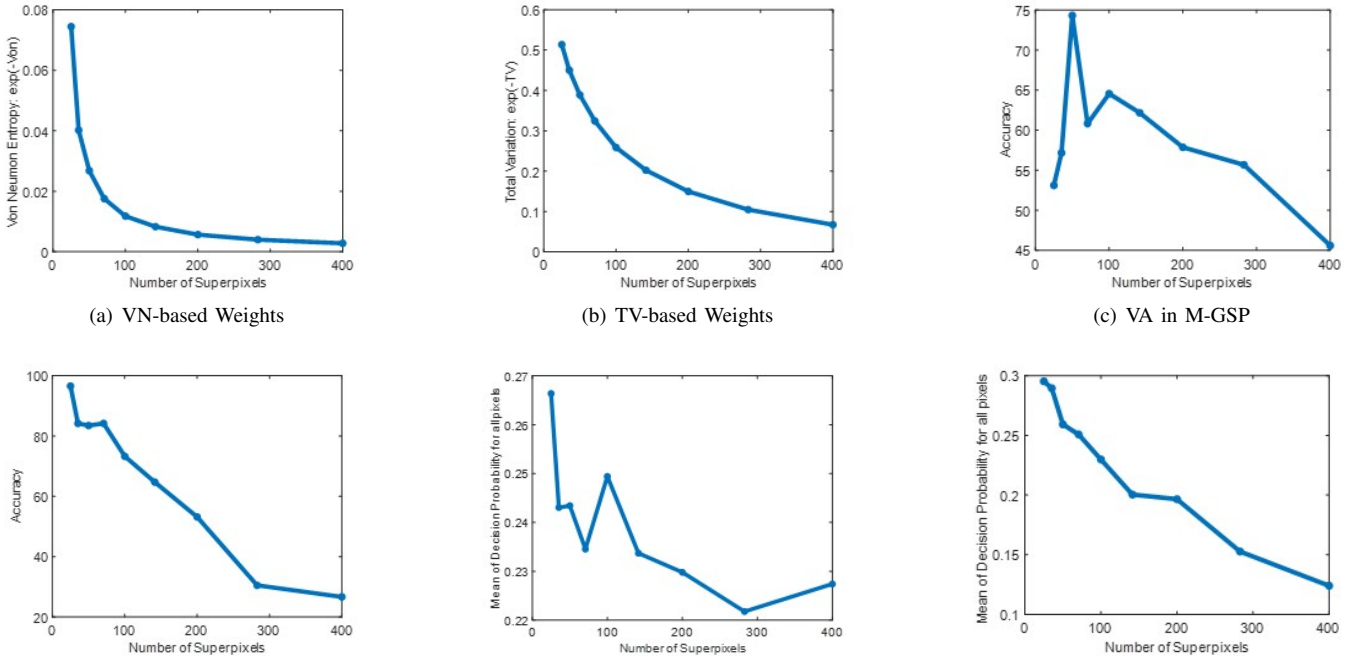
*1) Overall Accuracy:* Applying M-GSP based spectral clustering as feature extraction of HSI, we compare the proposed algorithms with several well known feature extraction algorithms, including PCA [55], ICA [17], LPP [56], NPE [57], LP-NPE [58], LDA [15], LFDA [59], SPCA [16] and MSPCA [16]. Notably, MSPCA is also an algorithm integrating multiple resolutions of superpixels. For the proposed MLN-SRC and MLN-MRC, we regroup the superpixels to $70\%$ of the original superpixel number, before extracting features based on M-GSP. Here, we show the results of MLN-MRC according to different decision values (fusion weights). More analysis of different fusion strategies will be illustrated further in Section V-C2.

The overall accuracy under different numbers of training samples per class (TS/C) is shown in Table III. In this experiment, parameters of multiple resolutions are tuned for different HSIs. From the test results, the proposed MLN-MRC exhibits a superior overall performance, especially for those scenarios with fewer classes. MLN-SRC is marginally better than the SPCA. Note that, since our MLN-based methods are easily integrable with various dimension reduction algorithms, we can combine MLN-SRC with SPCA to further improve the performance. Because of page limitation, we leave the integration of MLN-based methods and other dimension reduction algorithms in future studies.

TABLE V
PERFORMANCE OF DIFFERENT FUSION STRATEGIES

| | TS/C | MV | VA | DV | VN | TV |
|---|---|---|---|---|---|---|
| Indian Pines | | | | | | |
| MS-PCA | 5 | 0.6663 | 0.7241 | **0.7334** | 0.7263 | 0.7265 |
| | 10 | 0.7476 | 0.7778 | **0.7829** | 0.7661 | 0.7661 |
| MLN-MRC | 5 | 0.7164 | **0.7338** | 0.7260 | 0.7271 | 0.7295 |
| | 10 | 0.8220 | 0.8253 | **0.8253** | 0.7886 | 0.7823 |
| Pavia University | | | | | | |
| MS-PCA | 5 | 0.7078 | 0.7440 | 0.7317 | **0.7798** | 0.7712 |
| | 10 | 0.8447 | 0.8541 | 0.8553 | **0.8588** | 0.8500 |
| MLN-MRC | 5 | 0.8162 | 0.8140 | **0.8240** | 0.7977 | 0.8085 |
| | 10 | 0.8549 | 0.8586 | **0.8605** | 0.8459 | 0.8435 |
| Salinas | | | | | | |
| MS-PCA | 5 | 0.8521 | 0.9256 | 0.9349 | **0.9655** | 0.9634 |
| | 10 | 0.9662 | 0.9774 | 0.9718 | **0.9789** | 0.9765 |
| MLN-MRC | 5 | 0.9469 | 0.9456 | 0.9485 | **0.9844** | 0.9622 |
| | 10 | 0.9872 | 0.9874 | 0.9873 | 0.9857 | **0.9942** |

In real/practical scenarios, tuning parameters is always important for performance improvement. Since the prior knowledge of the optimal parameter choice is not always available, we also test the performance of MSPCA and MLN-MRC under the same set of input superpixel numbers to facilitate fair comparison. In IV, we fuse the results from 9 resolutions, i.e., $N_i \in [25, 35, 50, 70, 100, 140, 200, 280, 400]$, for all comparative methods. From these numerical results, we see that MLN-MRC provides an explicit improvement over MSPCA given fewer numbers of training samples since MLN-based methods can better exploit the underlying correlation among all pixels. When given enough TS/C, MSPCA exhibits performance similar to that of MLN-MRC.




(a) VN-based Weights (b) TV-based Weights (c) VA in M-GSP

(d) VA in MSPCA (e) Mean DV in M-GSP (f) Mean DV in MSPCA

Fig. 16. Decision Strengths over Different Numbers of Superpixels.

TABLE VI
OVERALL ACCURACY IN NOISY ENVIRONMENT

| **(Uniform/Gaussian)** | | Indian Pines | | Pavia University | | Salinas | |
|---|---|---|---|---|---|---|---|
| TS/C | Noise Level | MPCA | MLN-MRC | MSPCA | MLN-MRC | MSPCA | MLN-MRC |
| Setup 1 | | | | | | | |
| 5 | 5% | 0.7140/0.7012 | **0.7377/0.7493** | 0.7219/0.7160 | **0.8230/0.8126** | 0.8592/0.8670 | **0.9231/0.9159** |
| | 10% | 0.7110/0.6702 | **0.7135/0.7276** | 0.7080/0.7031 | **0.8085/0.7873** | 0.8341/0.8421 | **0.8956/0.9023** |
| | 15% | 0.6868/0.6507 | **0.7133/0.7215** | 0.6831/0.6817 | **0.7932/0.7640** | 0.8107/0.8215 | **0.8815/0.8975** |
| 10 | 5% | 0.7613/0.7753 | **0.8033/0.8056** | **0.8499**/0.8374 | 0.8414/**0.8475** | 0.9509/0.9647 | **0.9632/0.9684** |
| | 10% | 0.7522/0.7498 | **0.7788/0.7984** | 0.8481/0.8436 | **0.8597/0.8461** | 0.9408/0.9431 | **0.9445/0.9491** |
| | 15% | 0.7317/0.7382 | **0.7544/0.7845** | 0.8350/0.7848 | **0.8467/0.8333** | 0.9381/**0.9368** | **0.9421**/0.9317 |
| Setup 2 | | | | | | | |
| 5 | 5% | 0.6839/0.6740 | **0.7023/0.6898** | 0.7328/0.7093 | **0.8422/0.8390** | 0.8591/0.8679 | **0.9240/0.9050** |
| | 10% | 0.6640/0.6501 | **0.6764/0.6647** | 0.7052/0.6972 | **0.8023/0.7837** | 0.8471/0.8512 | **0.9045/0.8878** |
| | 15% | 0.6413/0.6247 | **0.6687/0.6427** | 0.6754/0.6789 | **0.7635/0.7530** | 0.8317/0.8421 | **0.8915/0.8775** |
| 10 | 5% | 0.7579/0.7408 | **0.8022/0.7810** | **0.8649**/0.8437 | 0.8573/**0.8632** | 0.9682/0.9621 | **0.9735/0.9633** |
| | 10% | 0.7409/0.7333 | **0.7800/0.7762** | 0.8442/0.8354 | **0.8469/0.8501** | 0.9521/0.9450 | **0.9547/0.9532** |
| | 15% | 0.7208/0.7235 | **0.7695/0.7584** | **0.8410/0.8309** | 0.8368/0.8258 | **0.9437/0.9416** | 0.9305/0.9241 |

*2) Analysis of Different Fusion Strategies:* Now, we analyze the performance of different fusion strategies. Since MSPCA [16] also fuses multiple resolutions of superpixels, we investigate our proposed fusion methods in combination with both MSPCA and MLN-MRC.

We follow the same setup of 9 resolution as Table IV. The test accuracy is shown in Fig. 16. As shown, our proposed weight approaches lead to significant improvement over the basic majority voting (VT) for MSPCA. Compared to SVM-based weights, these graph-based weights are better for MSPCA since additional geometric information is considered. For MLN-MRC, the proposed weights show a slight improvement, which suggests that MLN-MRC is less sensitive to different decision strengths. To better understand the effect of different weights, we also present decision strengths over different numbers of superpixels in Fig. 16. In Figs. 16(a)-16(b), the graph-based weights favor low resolution to form baselines and use high resolutions to interpolate details. Thus, for HSIs with larger area of segmented groups, such as *Salinas*, the graph-based weights generate superior performances. Using SVM-based weights, we find no consistent trend among different superpixels, and the results vary for different datasets. Since the MLN-based methods have already incorporated the underlying geometric structures, they continue to display robust results even when using SVM-based weights.

*3) Robustness in Noisy Datasets:* We further evaluate the robustness of proposed methods in noisy environment. More specifically, we consider two types of noise models: 1) pixel-dependent noise where pixel noise variance depending on corresponding pixel data value; and 2) non-pixel dependent noise where noise variance is defined by mean of all pixel values. These two different noise models describe two different

practical sensing noises. We also test both uniform noise and Gaussian noise. From the test performances shown in Table VI, we find the newly proposed MLN methods to be less sensitive to various types of sample noises.

*4) Complexity:* Since our proposed M-GSP processing can be flexibly integrated with other dimension reduction methods to reduce complexity, we find it unnecessary to provide evaluation of computation complexity for various setups. In general, the original MLN-MRC has a similar runtime as MSPCA. For example, under the same settings of multi-resolutions as Section V-C1 in *Indian Pines* HSI with tuning parameters of SVM among 15 sets, the runtimes for MLN-MRC and MSPCA are $28.21$ seconds and $25.37$ seconds, respectively. These and other tests indicate similar computation complexity for methods based on M-GSP and PCA.

## VI. Conclusions

This work introduces the use of M-GSP in hyperspectral imaging processing. To capture heterogeneous underlying structures within different but highly correlated spectrum frames in hyperspectral images (HSIs), we propose to represent HSIs via multilayer graph networks. Analyzing singular spectra of adjacency tensor for the multilayer network, we first develop a MLN-based spectral clustering for unsupervised HSI segmentation. Extracting features based on MLN-GSP, we then propose two algorithms for the semi-supervised HSI segmentation. We also consider several novel decision fusion strategies for multiple resolution superpixel analysis. Our experimental results demonstrate the robustness and efficiency of the proposed methods, successfully showcasing the power of M-GSP in HSI analysis.

Since the proposed MLN-based methods is amenable to integration with various dimension reduction and feature extraction methods, future works may consider M-GSP in connection with other advanced HSI processing approaches for performance improvement and for robustness. Another interesting future direction is to extract more valuable features via MLN Fourier transform (M-GFT) beyond only spectrum decomposition. Additionally, we also find the challenge of processing unlabeled samples (background) an exciting future research direction.

## References


[1] W. Ma et al., "A signal processing perspective on hyperspectral unmixing: insights from remote sensing," in *IEEE Signal Processing Magazine*, vol. 31, no. 1, pp. 67-81, Jan. 2014.

[2] M. Fauvel, Y. Tarabalka, J. A. Benediktsson, J. Chanussot, and J. C. Tilton, "Advances in spectral-spatial classification of hyperspectral images," *Proceedings of the IEEE*, vol. 101, no. 3, pp. 652–675, Mar. 2013.

[3] L. Zhang, L. Zhang, and B. Du, "Deep learning for remote sensing data: A technical tutorial on the state of the art," *IEEE Geoscience and Remote Sensing Magazine*, vol. 4, no. 2, pp. 22–40, 2016.

[4] W. Zhao and S. Du, "Spectral–spatial feature extraction for hyperspectral image classification: a dimension reduction and deep learning approach," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 8, pp. 4544-4554, Aug. 2016.

[5] Y. Z. Feng, and D. W. Sun, "Application of hyperspectral imaging in food safety inspection and control: a review," *Critical Reviews in Food Science and Nutrition*, vol. 52, no. 11, pp. 1039-1058, Oct. 2011.

[6] H. Z. Shafri, E. Taherzadeh, S. Mansor, and R. Ashurov, "Hyperspectral remote sensing of urban areas: an overview of techniques and applications," *Research Journal of Applied Sciences, Engineering and Technology*, vol. 4, no. 11, pp. 1557-1565, Jun. 2012.

[7] D. A. Landgrebe, *Signal Theory Methods in Multispectral Remote Sensing*. Hoboken, NJ: Wiley, 2003.

[8] H. Zhang, H. Zhai, L. Zhang and P. Li, "Spectral–spatial sparse subspace clustering for hyperspectral remote sensing images," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 6, pp. 3672-3684, Jun. 2016.

[9] M. Fauvel, J. A. Benediktsson, J. Chanussot and J. R. Sveinsson, "Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 11, pp. 3804-3814, Nov. 2008.

[10] F. Melgani, and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 8, pp. 1778-1790, Aug. 2004.

[11] Q. Wang, J. Lin, and Y. Yuan, "Salient band selection for hyperspectral image classification via manifold ranking," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 6, pp. 1279–1289, Jun. 2016.

[12] C. Rodarmel, and J. Shan, "Principal component analysis for hyperspectral image classification," *Surveying and Land Information Science*, vol. 62, no. 2, pp. 115-122, 2002.

[13] V. Laparra, J. Malo, and G. Camps-Valls, "Dimensionality reduction via regression in hyperspectral imagery," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 6, pp. 1026–1036, Sep. 2015.

[14] J. Ma, Y. Ma, and C. Li, "Infrared and visible image fusion methods and applications: A survey," *Inf. Fusion*, vol. 45, pp. 153–178, Jan. 2019.

[15] S. Prasad and L. M. Bruce, "Limitations of principal components analysis for hyperspectral target recognition," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 4, pp. 625–629, Oct. 2008.

[16] J. Jiang, J. Ma, C. Chen, Z. Wang, Z. Cai and L. Wang, "SuperPCA: A superpixelwise pca approach for unsupervised feature extraction of hyperspectral imagery," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 8, pp. 4581-4593, Aug. 2018.

[17] J. Wang and C.-I. Chang, "Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 6, pp. 1586–1600, Jun. 2006.

[18] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, 2000.

[19] J. Ma, J. Jiang, C. Liu, and Y. Li, "Feature guided gaussian mixture model with semi-supervised em and local geometric constraint for retinal image registration," *Inf. Sci*, vol. 417, pp. 128–142, Nov. 2017.

[20] Y. Chen, Z. Lin, X. Zhao, G. Wang and Y. Gu, "Deep learning-based classification of hyperspectral data," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 6, pp. 2094-2107, Jun. 2014.

[21] J. Li, J. M. Bioucas-Dias and A. Plaza, "Semisupervised hyperspectral image segmentation using multinomial logistic regression with active learning," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 11, pp. 4085-4098, Nov. 2010.

[22] X. Guo, X. Huang, L. Zhang, L. Zhang, A. Plaza and J. A. Benediktsson, "Support Tensor Machines for Classification of Hyperspectral Remote Sensing Imagery," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 6, pp. 3248-3264, Jun. 2016.

[23] Y. Zhao, Y. Yuan, and Q. Wang, "Fast spectral clustering for unsupervised hyperspectral image classification," *Remote Sensing*, vol. 11, no. 4, pp. 399, Jan. 2019.

[24] F. Nie, X. Wang, M. Jordan, and H. Huang, "The constrained laplacian rank algorithm for graph-based clustering," in *Proceedings of the AAAI Conference on Artificial Intelligence*, Phoenix, Arizona, USA, Feb. 2016.

[25] W. Peizhuang, "Pattern recognition with fuzzy objective function algorithms," *SIAM Review*, vol. 25, no. 3, pp. 442, Jul. 1983.

[26] S. Vijendra *Efficient clustering for high dimensional data: Subspace based clustering and density based clustering*, *Information Technology Journal*, vol. 10, no. 6, pp. 1092-1105, 2011.

[27] Y. Zhong, L. Zhang, and W. Gong, "Unsupervised remote sensing image classification using an artificial immune network," *International journal of remote sensing*, vol. 32, no. 19, pp. 5461-5483, Aug. 2011.

[28] L. Zhang, and J. A. You, *A spectral clustering based method for hyperspectral urban image,* in *Proceedings of the 2017 Joint Urban Remote Sensing Event (JURSE)*, Dubai, UAE, March 2017; pp. 1–3.

[29] D. Hong, L. Gao, J. Yao, B. Zhang, A. Plaza and J. Chanussot, "Graph convolutional networks for hyperspectral image classification," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 7, pp. 5966-5978, Jul. 2021.

[30] S. Wan, C. Gong, P. Zhong, B. Du, L. Zhang and J. Yang, "Multiscale dynamic graph convolutional network for hyperspectral image classification," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 5, pp. 3162-3177, May 2020.

[31] B. Liu, K. Gao, A. Yu, W. Guo, R. Wang, and X. Zuo, "Semisupervised graph convolutional network for hyperspectral image classification," *Journal of Applied Remote Sensing*, vol. 14, no. 2, p. 026516, Jun. 2020.

[32] A. Qin, Z. Shang, J. Tian, Y. Wang, T. Zhang and Y. Y. Tang, "Spectral–spatial graph convolutional networks for semisupervised hyperspectral image classification," in *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 2, pp. 241-245, Feb. 2019.

[33] J. Nalepa, M. Myller, Y. Imai, K. I. Honda, T. Takeda,and M. Antoniak, "Unsupervised segmentation of hyperspectral images using 3-D convolutional autoencoders," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 11, pp. 1948-1952, Nov. 2020.

[34] S. Zhang, H. Zhang, H. Li and S. Cui, "Tensor-based spectral analysis of cascading failures over multilayer complex systems," in *Proceedings of 56th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Monticello, USA, Oct. 2018, pp. 997-1004.

[35] S. Zhang, Q. Deng, and Z. Ding, "Graph signal processing over multilayer networks–part i: foundations and spectrum analysis," arXiv:2108.13638, 2021.

[36] S. Zhang, Q. Deng, and Z. Ding, "Graph signal processing over multilayer networks–part ii: useful tools and practical applications," arXiv:2108.13639, 2021.

[37] A. Ortega, P. Frossard, J. Kovačević, J. M. F. Moura and P. Vandergheynst, "Graph signal processing: overview, challenges, and applications," in *Proceedings of the IEEE*, vol. 106, no. 5, pp. 808-828, May 2018.

[38] M. De Domenico, A. Sole-Ribalta, E. Cozzo, M. Kivela, Y. Moreno, M. A. Porter, S. Gomez, and A. Arenas, "Mathematical formulation of multilayer networks," in *Physical Review X*, vol. 3, no. 4, p. 041022, Dec. 2013.

[39] L. De Lathauwer, B. De Moor, and J. Vandewalle, "A multilinear singular value decomposition," *SIAM Journal on Matrix Analysis and Applications*, vol. 21, no. 4, pp. 1253-1278, Jan. 2000.

[40] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Review*, vol. 51, no. 3, pp. 455-500, Aug. 2009.

[41] P. Massoudifar, A. Rangarajan and P. Gader, "Superpixel Estimation for Hyperspectral Imagery," *2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops*, Columbus, OH, USA, Jun. 2014, pp. 287-292.

[42] Y. R. Fan, and T. Z. Huang, "Hyperspectral image restoration via superpixel segmentation of smooth band," *Neurocomputing*, vol. 455, pp. 340-352, Sep. 2021.

[43] Q. Yan, L. Xu, J. Shi, and J. Jia, "Hierarchical saliency detection," in CVPR, Portland, OR, USA, Jun. 2013, pp. 1155–1162.

[44] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 8, pp. 888–905, Aug. 2000.

[45] M.-Y. Liu, O. Tuzel, S. Ramalingam, and R. Chellappa, "Entropy rate superpixel segmentation," in *CVPR*, Colorado Springs, CO, USA, Jun. 2011, pp. 2097–2104.

[46] G. L. Nemhauser, L. A. Wolsey, and M. L. Fisher, "An analysis of approximations for maximizing submodular set functions-I," *Mathematical Programming*, vol. 14, no. 1, pp. 265–294, 1978.

[47] S. Chen, D. Tian, C. Feng, A. Vetro and J. Kovačević, 'Fast resampling of three-dimensional point clouds via graphs," in *IEEE Transactions on Signal Processing*, vol. 66, no. 3, pp. 666-681, Feb., 2018.

[48] A. Y. Ng, M. I. Jordan, and Y. Weiss, "On spectral clustering: analysis and an algorithm," in *Advances in Neural Information Processing Systems*, Vancouver, Canada, Dec. 2002, pp. 849-856.

[49] U. V. Luxburg, "A tutorial on spectral clustering," *Statistics and Computing*, vol. 17, no. 4, pp. 395-416, Aug. 2007.

[50] P. Das and A. Ortega, "Graph-based skeleton data compression," *2020 IEEE 22nd International Workshop on Multimedia Signal Processing (MMSP)*, Tampere, Finland, Sep. 2020, pp. 1-6.

[51] F. Grassi, A. Loukas, N. Perraudin and B. Ricaud, "A time-vertex signal processing framework: scalable processing and meaningful representations for time-series on graphs," in *IEEE Transactions on Signal Processing*, vol. 66, no. 3, pp. 817-829, Feb. 2018

[52] J. S. Stanley, E. C. Chi and G. Mishne, "Multiway graph signal processing on tensors: integrative analysis of irregular geometries," in *IEEE Signal Processing Magazine*, vol. 37, no. 6, pp. 160-173, Nov. 2020.

[53] J. Weston, and C. Watkins, "Multi-class support vector machines," *Technical Report*, Department of Computer Science, Royal Holloway, University of London, May, 1998.

[54] M. De Domenico, V. Nicosia, A. Arenas, and V. Latora, V. "Structural reducibility of multilayer networks.," *Nature communications*, vol. 6, no. 1, pp. 1-9, Apr. 2015.

[55] B. Scholkopf, A. Smola, and K. R. Muller, "Nonlinear component analysis as a kernel eigenvalue problem," *Neural Computation*, vol. 10, no. 5, pp. 1299–1319, 1998.

[56] X. He and P. Niyogi, "Locality preserving projections," in *Advances in Neural Information Processing Systems*, Vancouver, Canada, Dec. 2004, pp. 153–160.

[57] X. He, D. Cai, S. Yan, and H. J. Zhang, "Neighborhood preserving embedding," in *ICCV*, Beijing, China, Oct. 2005, pp. 1208–1213.

[58] Y. Zhou, J. Peng and C. L. P. Chen, "Dimension reduction using spatial and spectral regularized local discriminant embedding for hyperspectral image classification," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 2, pp. 1082-1095, Feb. 2015.

[59] W. Li, S. Prasad, J. E. Fowler and L. M. Bruce, "Locality-preserving dimensionality reduction and classification for hyperspectral image analysis," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 4, pp. 1185-1198, Apr. 2012.