Graph Neural Sparsification for Hyperspectral Image Classification with Local and Global Consistency

Haojie Hu[1], Minli Yao[1], Fang He[1] and Fenggan Zhang[1]

[1]Xi'an Research Institute of Hi-Tech, Xi'an, 710025, China.

Abstract

Recently, graph neural network (GNN) has drawn increasing attention for hyperspectral image (HSI) classification, and there are many related works have been proposed. However, previous works are focus on either the superpixel-level local spatial information or the pixel-level global spectral information. In this paper, a novel Graph Neural Sparsification Network (GNSN) is proposed for HSI classification. In our method, a fully connected graph is adaptively constructed to make full use of local spatial information and global spectral information. Besides, we apply a neural sparsification technique to remove potentially task-irrelevant edges in case of misleading message propagation. The resulting network (GNSN) is end-to-end trainable. We conduct experiments on three popular benchmarks, including Indian Pines, Pavia University, and Kennedy Space Center, and achieve state-of-the-art performance on all three datasets.

Keywords

Graph neural network (GNN), hyperspectral image (HSI), superpixel, sparsification

1. Introduction

HSI classification is a fundamental and challenging problem in HSI processing [1], which aims to assign a specific label to each pixel in the image. It has been widely applied to many scenarios, such as military target detection, vegetation monitoring, and disaster prevention and control.

Convolutional neural networks (CNNs) have proven their effectiveness in image processing by automatically extract discriminative features. Many CNN-based variants have been designed for HSI classification from the 1D-CNN to the 3D-CNN, from the single CNN to the hybrid CNN. Although the existing CNN-based methods have achieved good performance to some extent, they need many training parameters and tend to be overfitting due to the scarcity of the training samples. Meanwhile, CNNs tend to blur the classification boundary using a fixed shape kernel around the central pixel to extract features. Therefore, the accurate classification of HSI is still challenging.

To build more flexible models, many GNN-based methods have been proposed for HSI classification. Generally, there can be two categories according to the types of adjacent graph. One is pixel-level adjacent graph[2, 3]. This type of graph can directly propagate information between nearby and distant regions, while it takes each pixel as a node of a graph, leading to a massive amount of computation and limits its applicability. Superpixel-level adjacent graph is another commonly used graph for HSI classification [4, 5], which is usually accomplished by a superpixel segmentation algorithm. However, this graph only focuses on the local structure of the data. In order to capture the long-range contextual information, multiple convolution layers are required to be stacked, which is highly inefficient and will bring potential concerns of over-smoothing.

Aiming to address the above issues of the superpixel-level adjacent graph, we propose a novel GNN-based HSI classification algorithm (GNSN), in which a fully connected graph is adaptively constructed to make full use of local spatial and global spectral information. Furthermore, in order to avoid potentially misleading message propagation, we apply the neural sparsification technique to remove potentially task-irrelevant edges. The main contributions of this paper are summarized as follows:

- We incorporate both local and global information by a learnable graph, which will automatically decide whether the model will likely pay more attention to the nodes near it or pay more attention to the nodes far away from it;
- GNSN applies a sparsification strategy that favors subsequent classification task by removing potentially task-irrelevant edges;
- Experiments on three benchmark HSI datasets show that GNSN outperforms other compared HSI classification methods.

CDCEO 2021: 1st Workshop on Complex Data Challenges in Earth Observation, November 1, 2021, Virtual Event, QLD, Australia.

haojiehu705@gmail.com (H. Hu); yaominli66@163.com (M. Yao); fanghe1107@gmail.com (F. He); zfg417@163.com (F. Zhang)

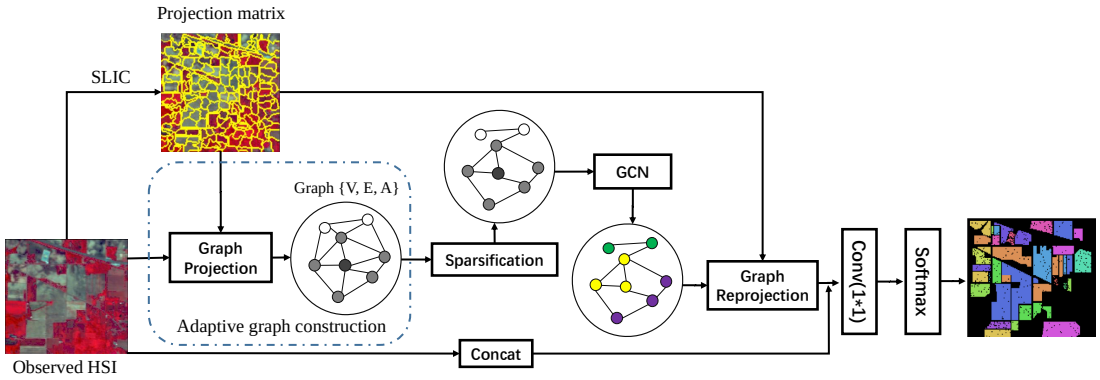

Figure 1: Schematic depiction of the proposed Graph Neural Sparsification Network. GNSN first assigns nearby pixels with similar features to the same vertex after graph projection. And then, a learned adjacent matrix that captures both local and global structure is obtained through adaptive graph construction. Under the guidance of sparsification, the graph removes most task-irrelevant edges and is subsequently fed into a two-layer GCN to get the high-level feature map. After the graph reprojection, the corresponding coarse segmentation result is obtained. Finally, the refined feature map is fused with the original feature to get the final refined classification result.

2. Approach

The task for hyperspectral classification aims to predict the labels of all the pixels in the HSI by a classification model given the supervised data set. As illustrate in Fig. 1, we present a Graph Neural Sparsification Network for the HSI classification model to adaptively capture both local and global contextual information. GNSN takes the original image data $\boldsymbol{X} \in \mathbb{R}^{H\times W\times B}$ and the association matrix $\boldsymbol{Q} \in \mathbb{R}^{HW\times N}$ produced by a superpixel segmentation algorithm as input, where H, W B and N represent the height, width, the number of bands, and the number of superpixels respectively. By treating superpixels as the vertexes in the graph, we transform the image to the graph representation $\boldsymbol{V} \in \mathbb{R}^{N\times B}$. Applying the adaptive graph construction discussed below, we can obtain the similarity matrix $\boldsymbol{A} \in \mathbb{R}^{N\times N}$ that captures both local and global structure in the image data. Subsequently, the sparsification operation removes most task-irrelevant edges in the graph. Following the paradigm of GCN, a two-layer GCN network is applied on the graph to further propagate information, resulting in a high-level feature map $\widetilde{\boldsymbol{V}} \in \mathbb{R}^{N\times C}$, where C is the dimension of the feature map. After the graph reprojection, the corresponding coarse segmentation result $\widetilde{\boldsymbol{X}} \in \mathbb{R}^{H\times W\times C}$ is obtained in the original grid form. Finally, the refined feature map concatenated with the original feature passes through a conventional 1×1 convolutional layer to get the final classification result. GNSN consists of four main blocks: Graph Projection and Reprojection, Graph Construction, Sparsification, and GCN. In the following, we will describe each block in detail.

Graph Projection and Reprojection. To reduce the computational complexity and meanwhile preserve the local structure of HSI, the GNN-based models usually work on superpixel-based nodes but not the pixel-based nodes. Similarly, we first apply the simple linear iterative clustering (SLIC) method to partition the entire image into many spatially connected superpixels. By treating each superpixel as a graph node and taking the average spectral signatures of each superpixel as the node feature, the graph projection assigns $\boldsymbol{X}$ to a set of vertices and gets the feature matrix $\boldsymbol{V}$ of the vertices in the graph by the form of matrix multiplication as follows

$$\boldsymbol{V} = \hat{\boldsymbol{Q}}^T \text{Flatten}(\boldsymbol{X}), \tag{1}$$

where $\hat{\boldsymbol{Q}}$ is the normalized $\boldsymbol{Q}$ by column, i.e., $\hat{\boldsymbol{Q}}_{i,j} = \boldsymbol{Q}_{i,j} / \sum_m \boldsymbol{Q}_{m,j}$. Flatten$(\cdot)$ denotes flattening the HSI data by the spatial dimension, and the association matrix $\boldsymbol{Q} \in \mathbb{R}^{HW\times N}$ is introduced by SLIC that is defined as

$$\boldsymbol{Q}_{i,j} = \begin{cases} 1, & if \quad \boldsymbol{x}_i \in \mathcal{S}_j, \\ 0, & \text{otherwise}, \end{cases} \tag{2}$$

where $\mathcal{S}_i$ is the i-th superpixel that consists of several homogeneous pixels.

We perform graph reprojection to transform the resulting features back to the original coordinate space. The graph reprojection operation takes the inputs of transformed vertex features $\widetilde{\boldsymbol{V}}$ and the assignment matrix $\boldsymbol{Q}$, and produces the corresponding 2D feature map $\widetilde{\boldsymbol{X}}$. This operation is defined as

$$\widetilde{\boldsymbol{X}} = \text{Reshape}(\boldsymbol{Q}\widetilde{\boldsymbol{V}}), \tag{3}$$

where Reshape(·) denotes restoring the spatial dimension of the flattened data.

Graph Construction. The graph (which is the adjacency matrix in our formulation) is constructed according to the similarity between different nodes. Denote $\boldsymbol{A}_{ij}$ as the (i, j)-element of the adjacency matrix $\boldsymbol{A}$, we have:

$$\boldsymbol{A}_{ij} = \boldsymbol{w}^T[\boldsymbol{W}^T\boldsymbol{v}_i||\boldsymbol{W}^T\boldsymbol{v}_j] + b_{(\boldsymbol{v}_i,\boldsymbol{v}_j)}, \tag{4}$$

where $b_{(v_i,v_j)} = 1$ if the i-th region and the j-th region are spatially related, and otherwise $b_{(v_i,v_j)} = 0$. Compared to the adjacency graph described in [5, 6], where the receptive field is restricted to the neighbors, we can see that Eq. (4) provides both global and local information. And the learnable graph will automatically decide whether the model will likely pay more attention to the nodes near it or pay more attention to the nodes far away from it.

Sparsification Network. We focus on k-neighbor graph for graph neural sparsification. The k-neighbor graph is obtained by repeatedly sampling k edges for each node in the original graph. To make samples differentiable, we apply Gumbel-Softmax to generate differentiable discrete samples. Firstly, softmax function is employed to compute the probability to sample the edge, which is

$$\pi_{v_i,v_j} = \frac{\exp\left(\boldsymbol{A}_{v_i,v_j}\right)}{\sum_{w\in\mathbb{N}_{v_i}}\exp\left(\boldsymbol{A}_{v_i,w}\right)} \tag{5}$$

Then we can generate differentiable samples using Gumble-Softmax as follows

$$x_{v_i,v_j} = \frac{\exp\left(\left(\log\left(\pi_{v_i,v_j}\right) + \epsilon_v\right)/\tau\right)}{\sum_{w\in\mathbb{N}_{v_i}}\exp\left(\left(\log\left(\pi_{v_i,w}\right) + \epsilon_w\right)/\tau\right)}, \tag{6}$$

where x_{v_i,v_j} is a scalar that represents weather to sample the edge between v_i and v_j, $\epsilon_{v_j} = -\log(-\log(s))$ with s randomly drawn from Uniform (0,1), and τ is a hyperparameter that controls the interpolation between discrete distribution and continuous categorical densities.

Graph Convolution Network (GCN).

We perform two-layer GCN from [7] to further propagate information on the graph. Specifically, for a single graph convolution with its parameter $\boldsymbol{W} \in \mathbb{R}^{d\times d'}$, the operation is defined as

$$\widetilde{\boldsymbol{V}} = f(\boldsymbol{A}\boldsymbol{V}\boldsymbol{W}), \tag{7}$$

where f can be any nonlinear activation function. We can see that GCN calculates new features of a vertex as a weighted average of features of its neighbors on a graph, which allows the vertices in the same cluster to have similar features, making the subsequent classifications much easier.

3. Experiments

In this section, we conduct experiments to validate the effectiveness of our proposed method GNSN on three real-world benchmark datasets, namely Indian Pines (IP), Pavia University (PU), and Kennedy Space Center (KSC). The Indian Pines dataset is captured over the agricultural test site in the Indian Pines, which contains 200 spectral bands with a size of 145×145 and 16 terrain classes for research and analysis. The Pavia University dataset is the continuous imaging of 115 spectral bands with a size of 610×340. This dataset includes nine land-cover classes. The KSC data contains 176 bands with a size of 512×614. For classification purposes, 13 classes representing the various land cover types that occur in this environment were defined for the site.

3.1. Experimental settings

To quantitatively evaluate different models for hyperspectral image classification tasks from various aspects, three metrics, including overall accuracy (OA), average accuracy (AA), and kappa coefficient, are used to evaluate the performance of all the compared methods. For all the datasets, we randomly select 30 labeled pixels in each class for training and choose 15 labeled pixels if the corresponding class has less than 30 samples. Besides, the learning rate and the number of training epochs are set to 0.001 and 2000, respectively. All the reported results are calculated based on the average of ten training sessions to obtain stable results, and the best results are highlighted in bold.

3.2. Classification results

In our experiments, we compare our method with several HSI classification methods: SVM, Multilayer Perception (MLP), 3DCNN [8], NLGCN [9] and GSAGE [10]. The quantitative results of different methods on all the three datasets are shown in Table 1. From the results, we can see that GNN-based methods, including NLGCN, GSAGE, and the proposed GNSN, show better performance than traditional machine learning and deep learning models (i.e., SVM, MLP, 3DCNN), demonstrating that GNN is more helpful for extracting discriminative information. Note that NLGCN and GSAGE neglect fusing the information of both local and global information, thus resulting in lower performance than our GNSN.

4. Conclusion

In this article, we propose a novel Graph Neural Sparsification Network (GNSN) for hyperspectral image classification. Different from prior way of constructing the

Table 1
Classification results (in percent) of defferent methods on three real datasets

Methods		SVM	MLP	3DCNN	NLGCN	GSAGE	GNSN
IP	OA	69.72	70.07	79.19	83.83	85.36	**94.08**
	AA	81.01	80.78	87.88	91.17	91.13	**96.20**
	Kappa	66.00	66.31	76.55	81.70	83.35	**93.24**
PU	OA	77.54	80.59	83.35	87.11	89.41	**97.72**
	AA	85.16	86.40	88.45	91.95	92.62	**97.33**
	Kappa	71.49	75.25	78.59	83.35	86.21	**96.97**
KSC	OA	86.70	87.72	91.47	93.76	96.78	**99.61**
	AA	81.92	83.81	88.11	90.78	95.00	**99.50**
	Kappa	85.16	86.31	90.48	93.02	96.40	**99.56**

adjacent graph, the proposed GNSN constructs a learnable graph that reveals the local and glocal structure of HSI. Besides, a differentiable sparsification technique is applied to favor subsequent classification task by removing potentially task-irrelerant edges. Experimental results reveal that our method significantly outperforms its counterparts in terms of OA, AA, and Kappa.

References

[1] D. Hong, X. Wu, P. Ghamisi, J. Chanussot, N. Yokoya, X. X. Zhu, Invariant attribute profiles: A spatial-frequency joint feature extractor for hyperspectral image classification, IEEE Transactions on Geoscience and Remote Sensing 58 (2020) 3791–3808. doi:10.1109/TGRS.2019.2957251.

[2] A. Qin, Z. Shang, J. Tian, Y. Wang, T. Zhang, Y. Y. Tang, Spectral spatial graph convolutional networks for semisupervised hyperspectral image classification, IEEE Geoscience and Remote Sensing Letters 16 (2019) 241–245. doi:10.1109/LGRS.2018.2869563.

[3] D. Hong, L. Gao, N. Yokoya, J. Yao, J. Chanussot, Q. Du, B. Zhang, More diverse means better: Multimodal deep learning meets remote-sensing imagery classification, IEEE Transactions on Geoscience and Remote Sensing 59 (2021) 4340–4354. doi:10.1109/TGRS.2020.3016820.

[4] D. Hong, L. Gao, J. Yao, B. Zhang, A. Plaza, J. Chanussot, Graph convolutional networks for hyperspectral image classification, IEEE Transactions on Geoscience and Remote Sensing 59 (2021) 5966–5978. doi:10.1109/TGRS.2020.3015157.

[5] Y. Ding, X. Zhao, Z. Zhang, W. Cai, N. Yang, Graph sample and aggregate-attention network for hyperspectral image classification, IEEE Geoscience and Remote Sensing Letters (2021) 1–5. doi:10.1109/LGRS.2021.3062944.

[6] Y. Ding, X. Zhao, Z. Zhang, W. Cai, N. Yang, Multiscale graph sample and aggregate network with context-aware learning for hyperspectral image classification, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 14 (2021) 4561–4572. doi:10.1109/JSTARS.2021.3074469.

[7] T. N. Kipf, M. Welling, Semi-supervised classification with graph convolutional networks, arXiv preprint arXiv:1609.02907 (2016).

[8] Y. Li, H. Zhang, Q. Shen, Spectral–spatial classification of hyperspectral imagery with 3d convolutional neural network, Remote Sensing 9 (2017) 67.

[9] L. Mou, X. Lu, X. Li, X. X. Zhu, Nonlocal graph convolutional networks for hyperspectral image classification, IEEE Transactions on Geoscience and Remote Sensing 58 (2020) 8246–8257. doi:10.1109/TGRS.2020.2973363.

[10] P. Yang, L. Tong, B. Qian, Z. Gao, J. Yu, C. Xiao, Hyperspectral image classification with spectral and spatial graph using inductive representation learning network, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 14 (2021) 791–800. doi:10.1109/JSTARS.2020.3042959.