IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 31, 2021, accepted February 13, 2021, date of publication February 23, 2021, date of current version March 26, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3061489

Stack Attention-Pruning Aggregates Multiscale Graph Convolution Networks for Hyperspectral Remote Sensing Image Classification

NA LIU[1], BIN ZHANG[2], QIUHUAN MA[1], QINGQING ZHU[1], AND XIAOLING LIU[1]
[1]School of Intelligent Manufacturing, Qingdao Huanghai University, Qingdao 266427, China
[2]Operation Branch of Qingdao Metro Group Company Ltd., Qingdao 266000, China

Corresponding author: Na Liu (214884817@qq.com)

This work was supported by the Science and Technology Program of colleges and universities in Shandong Province under Grant J18KB163.

ABSTRACT The hyperspectral remote sensing images are classified by traditional neural networks methods can achieve promising performance, but only operate on regular square regions with fixed. This will lead to between neighborhood pixels have limitations in achieve long-distances joint interaction modeling and cross-spacetime information flow for capturing complex spatial-temporal dependencies. Meanwhile ignoring importance detail information and improved utilization of irrelevant information. In the work, we propose a stack attention-pruning multiscale aggregates graph convolution framework (SAP-MAGACN). The framework can automatically learn and selectively attend to the relevant subspace structure by stack attention-pruning module, can effectively disentangle the complex space structure of remote sensing images and capture the rich structural semantics. Meanwhile a refine graph of neighborhood pixels are constructures. Then we adopt the aggregation manner for multiscale graph convolution of pixels nodes in different neighborhood for effective long-range joint interaction modeling. Finally, we leverage dense cross-spacetime edges to completion propagation of multiscale spatial-temporal information, and gradually produce the discriminative embedded features and effectively distinguish the categories of boundary pixels. The experimental results shown the propose SAP-MAGCN outperformance all others state-of-the-art methods on Indian Pines and Salinas public benchmark datasets. Such as the OA, AA and Kappa of our propose SAP-MAGCN frameworks is 96.75%, 95.73% and 97.33%, respectively, on Indian Pines datasets.

INDEX TERMS Hyperspectral remote sensing image classification, stack attention-pruning, multiscale graph convolution networks, longdistances joint interaction, multiscale spatial-temporal information, cross-spacetime edges.

I. INTRODUCTION

Image classification is one of the main tasks of remote sensing image processing. In recent years, influenced by the continuous development of remote sensing technology, hyperspectral remote sensing image has been playing an increasingly important role in many fields such as ecological environmental protection, land resource survey, military monitoring, agriculture and Marine research [1].

Up to now, researchers have been proposed many different methods to distinguish the attributes of the image elements in hyperspectral images. The early stage mainly based on traditional pattern recognition and representation of image content by means of manual marking. Such as BOW sparse representation, scale invariance (SITF). etc. feature representation, simple linear classifier. Among these traditional methods, Random Forest (RF) [2], decision-making Tree (DMT) [3] or simple Logistic Regression (LR) [5] has good robustness and satisfactory classification effect due to its strong processing ability for high-dimensional data. Nevertheless, With the considering that spatially neighboring pixels usually carry correlated information in a smooth spatial domain, it is difficult to distinguish the categories of pixels in an im-age with spectral information [6]. Therefore, many researchers combine spectral and spatial information to propose a multi-source feature joint classification method.

The associate editor coordinating the review of this manuscript and approving it for publication was Jiankang Zhang.

For example, the morphological contour method proposed by Fauvel *et al.* [7] and Song *et al.* [8] to effectively combine spatial and spectral information.

However, these methods are based on the simple features of manual participation setting and screening, which largely depend on prior knowledge, and the classification accuracy is seriously affected. To address these problems, deep learning technology has been widely used in hyperspectral image classification tasks due to its powerful presentation capabilities [9]. Deep learning method utilizes the activation state of neurons to gradually gather low-level features and automatically obtain abstract higher-order representations, avoiding complex feature engineering [10]. For example, [11]–[15] use the convolutional neural network to obtain the local deep abstract features of ground objects in remote sensing images, classify and identify them, and achieve a better classification effect. Chen *et al.* [16] constructed a stacked automatic encoder for high-level feature extraction to classify hyperspectral images. [16], [17], [19] *et al.* used cyclic neural network to further improve the utilization rate of semantic information of ground objects in remote sensing images, and enhance the representation ability of semantic features and the accuracy of classification and recognition. Lee *et al.* [20] designed a deep CNN, which is able to optimally explore contextual interactions by exploiting local spectral-spatial relation-ship among spatially neighboring pixels. To further im-prove the flexibility of convolutional neural network and cyclic neural network in remote sensing image content representation, the attention mechanism model is introduced into remote sensing image classification tasks [21]–[23]. Although the traditional deep learning method has achieved a better classification effect to some extent, it can only obtain the higher-order features of ground objects in the regular distribution region, and cannot adaptively capture the geometric changes of different object regions in the hyperspectral remote sensing image. In addition, in the feature extraction process, the boundary pixel category attribute information may be lost, affecting the final classification accuracy. In other words, the traditional neural network cannot capture the high order feature information of all the ground objects in the hyperspectral remote sensing image. Therefore, how to enhance the utilization rate of boundary pixel attribute information and capture the higher-order feature information of ground objects in irregular distributions areas so as to improve the classification accuracy of hyperspectral remote sensing images has become a re-search hotspot.

In recent years, graph convolution network (GCN) has been paid more and more attention [24], [25]. GCN can run directly on the graph structure and aggregate and transfer neighbor node information. At the same time, it can also capture higher-order semantic information of irregularly distributed data. For example, Sheng *et al.* [26] applied the image convolutional neural network to the classification of hyperspectral remote sensing im-ages for the first time, which improved the classification accuracy. Kang *et al.* [27] propose a deep metric learning based on scalable neighborhood components for remote sensing scene characterization, which aims at discovering the neighborhood structure in the metric space and preserving the class discrimination capability. Although this method captures the property information of boundary pixel effectively, it ignores the difference between different scale feature graphs and the correlation of graph nodes in spatial substructure. At the same time, it does not select the key information and discard the redundant information. In other words, it increases the use of redundant information.

In this paper, we proposed a stack attention-pruning multiscale aggregates graph convolution network (SAP-MAGCN) for hyperspectral remote sensing image classification. In SAP-MAGCN framework, we first prune the pixels that are weakly correlated to each other and delete this irrelevant redundant information by stack attention. And constructure the refine graph of neighborhood pixels. Then take many factors into consideration such as hyperspectral remote sensing data is often contaminated by redundant noise, and traditional graph convolution network (GCN) is limited by the initial graph structure when captures neighborhood nodes information, namely, the initial input graph structure may not accurate. Therefore, to effectively solve the limitation brought by predefine initial graph structure, we bring the graph convolution networks (GCN) into processing the graph of neighborhood pixels. Secondly, since hyperspectral remote sensing image obtain rich space structure information while the multiscale structure information is proven useful for improve the accuracy of image classification. we construction the multiscale graph by adaptive changes to the numbers of neighborhood nodes, and design a novel fusion manner called masked fusion, in this process, the adaptive changes of graph nodes and embedding features can effectively reinforce interaction of between neighborhood nodes, and establish dependency and complementary of the different scaled structure information. Finally, we adopt aggregates manner to aggregate the multiscale graph convolution blocks which facilitates direct information flow across spacetime and further boosts framework performance.

To sum up, the key contributions of the work are summarized as follows:

(1) We proposed a stack attention-pruning multiscale aggregates graph convolution (SAP-MAGCN) framework that removes the limitation of predefine graph structure, and removes redundant dependencies between node features from different neighborhoods and scales, which allows powerful multiscale aggregators to effectively modeling long-distance joint interaction and cross spatial-temporal dependencies on neighborhood nodes, and further refine multiscale graph.

(2) Multiscale graph convolution is utilized to extensively exploit and disintegrate the complex spatial structure information. Meanwhile, can achieve better features representation by masked fusion and aggregates manner, the interaction and spatial-temporal dependencies are enhanced of neighborhood nodes.

(3) The stack attention-pruning scheme is involved in our propose SAP-MAGCN framework, which significantly reduces the using of irrelevant redundant information and further refine the graph of neighborhood nodes. Then the experimental results on Indian pines and Salinas public benchmark datasets show that our proposed SAP-MAGACN acquire best performance when compared with others state-of-the-art methods.

The remaining of the paper is organized structured as follows: The related works on hyperspectral remote sensing image classification. Are summarized in Section II. We elaborate the proposed SAP-MAGCN framework in Section III. Section IV describes the experimental parameters and classification results. Finally, Section V draws the conclusion with future research discussion.

II. RELATED WORK

In this section, we elaborate existing method of hyperspectral remote sensing image classification.

With the development of remote sensing technology, the classification of hyperspectral remote sensing image has been studied deeply. For example, the classification of remote sensing images by support vector machines (SVM) [28] with kernel function shows a good classification effect. However, the correlation between adjacent pixels cannot be effectively utilized. In order to solve these defects, spectral space joint information was designed, For example, Li *et al.* assumed that adjacent pixels in the space have the same labels, and used Markov random fields (MRF) [28] to capture the contextual information in the image space to realize the classification of hyperspectral remote sensing images, which achieved encouraging results. Zhu *et al.* [29] captured HSV and ISH features of hyperspectral remote sensing images and classified them, achieving a good classification effect.

These methods use manually captured spectral and spatial features to represent the complex contents of the image, which are highly dependent on the prior knowledge of experts and have poor applicability. The classification method of hyperspectral remote sensing image based on deep learning can automatically capture the high-level discriminant features of image and effectively improve the classification accuracy of image. For example, Li *et al.* [30] used deep confidence network to capture high-level abstract features of images and image classification. Shi *et al.* [31] used a recursive neural network (RNN)to capture the multi-scale spectral spatial features of images, and learned the spatial dependence between non-adjacent pixels in a two-dimensional space domain. Yang *et al.* [32] proposed a dual-channel convolutional neural network to capture spectral and spatial joint features of hyperspectral remote sensing images, in which different channels are used to learn spectral and spatial features respectively. Jiang *et al.* [33] introduced the concept of multi-scale into the residual convolutional neural network to capture the multi-scale spatial information of images, further proving the effectiveness of multi-scale information. Zhou *et al.* [34] proposed a rotational invariant feature learning and joint decision method based on Siamese convolution neural network and combined recognition and validation model due to the limited size of remote sensing scene data set and the lack of label information. Duan *et al.* [35] proposed a multi-scale full variational method to extract structural features from hyper-spectral remote sensing images, tested and verified on multiple benchmark data sets, and proved that this method can well solve the problems of poor feature representation of small sample data. Xu *et al* [37], [38] propose multiscale octave 3D CNN with channel and spatial attention and faster multiscale capsule network with octave convolution for hyperspectral image classification. Although these deep learning methods can automatically extract the spectral and spatial features of images, they cannot effectively consider the geometric appearance of each local area, resulting in classification errors.

In recent years, the graph neural network has received wide attention. The compared to the convolutional neural network, recursive neural network and the mechanisms of attention, graph neural network (GNN) [36], [37] is effective in handling the irregular distribution of data, it can be through the feature information gathering and transferring node, capture figure higher-order abstract characteristics of data structure, and the figure structure data embedded in the first low dimensional discriminant space, realize the node classification and related tasks. In general, the traditional graph neural network is inefficient in the operation of largescale graph structured data. Therefore, in order to solve these problems, Bruna *et al.* [38] proposed the spectral convolutional network, which convolved the neighborhood nodes of graph structured data and obtained the higher-order abstract features of graph structured data. With the continuous development of graph convolutional neural network, GCN has been widely used in many fields such as natural language processing and Skeleton-based Action Recognition, and has achieved good results [39]–[43]. For example, Kipf and Welling [44] proposed semi-supervised graph convolutional neural network. Marcheggiani *et al.* [45] proposed a depth map convolutional encoder for text generation of structured data. In these works, GCN was simplified by a first-order approximation of graph spectral convolution, which leads to more efficient run operations.

Although GCN has made great achievements in many fields, it has only been used for hyper-spectral image classification in the previous two works [25], [26]. Among them, [25] only uses a fixed graph in the process of node convolution, so it cannot accurately reflect the internal relationship between pixels. Reference [26] effectively solved the influence brought by the graph structure itself, it did not capture category attribute information of boundary pixels to the maximum extent and did not strengthen the dependence between different scale features. Therefore, in order to solve these problems, we propose SAP-MAGCN framework, it can strengthen the dependencies between different scales characteristics, characteristics of the different scales of graph thinning, highlight the differences of characteristics between,

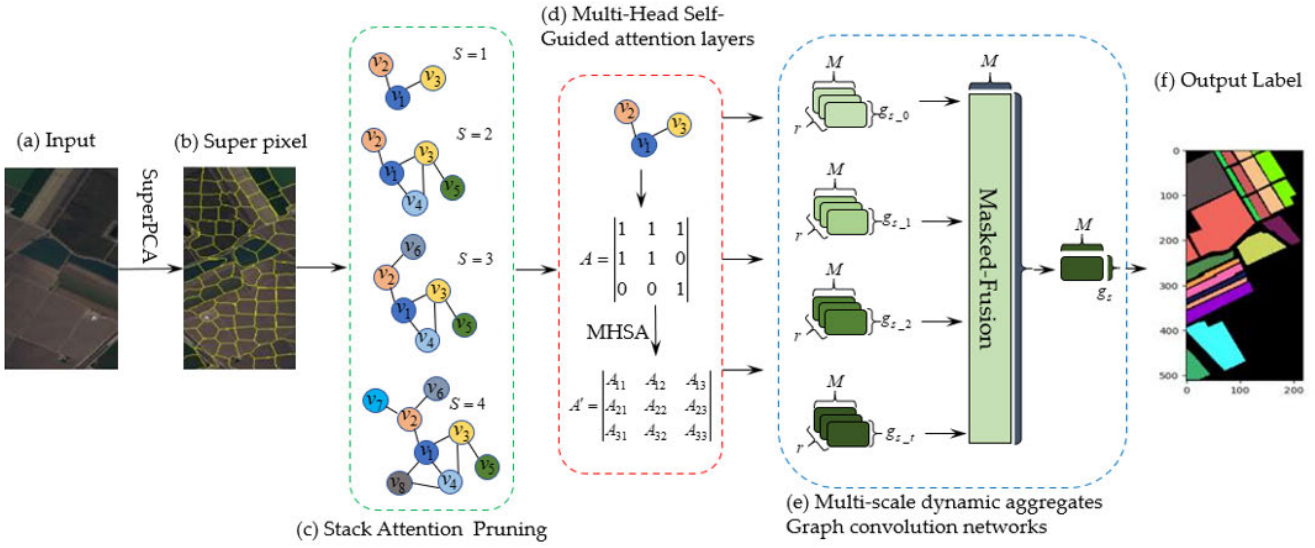


FIGURE 1. The framework of our propose SAP-MAGCNs.

and capture the maximum boundary pixels, the category of the attribute information, ensure good classification effect.

III. THE PROPOSED METHOD

In this section, we will present the basic components used for constructing our SAP-MAGCN model. refer to the Fig.1 for details. First, in order to reduce the running complexity of the model, the hyperspectral remote sensing image was preprocessed by using simple stack attention algorithm and divided into several subspace pixel blocks. Secondly, the pixel that constitutes the subspace pixel block is treated as a fully connect, and the whole fully connect pixel blocks is pruned into several subpixel block of different scales, and the adjacency matrix is constructed on this basis.

Among them, the edge weight of nodes not in the subpixel block is zero, and the edge weight of nodes in the subpixel block is allocated through the stack-attention pruning layer, which can effectively remove irrelevant redundant information, strengthen the correlation between relevant pixels, and retain the category attribute information of boundary pixels to the greatest extent. Then, the multi-scale aggregates graph convolutional network is used to perform convolution operation on these subpixel blocks of different scales, aggregate the multi-scale spectral spatial features of images, and gradually refine the input graph to further strengthen the dependence between the features of different scales. Finally, the mask fusion layer is used to splicing the feature maps of different scales to realize the accurate classification of hyperspectral remote sensing images. In figure 1, the SAP-MAGCN algorithm we proposed is mainly composed of three parts: the input of hyperspectral remote sensing image (see parts (a)). The backbone of SAP-MAGCNs, which includes stack attention pruning (part b and (c) in figure 1 for details), self-attention guided construction of adjacency matrices (part d), multiscale aggregates graph operations, and mask fusion layers (see parts e and f). Classification results (see section h).

In figure 1, (a) is the original hyperspectral remote sensing image. (b) subspace pixel blocks of hyperspectral remote sensing images, each of which contains several pixels. (c) to use the clipped graph structure data of the stack attention, which contains nodes and edges; (d) the adjacency matrix constructed for the use of the attention guidance module is the self-attention networks. Where represents the original adjacency matrix, and if there is an edge connection between nodes, it is 1; otherwise, it is 0. is the reconstructed adjacency matrix, and represents the attentional edge weights between nodes. (e) graph convolution black of multi-scale dense connection, where represents adjacency matrices constructed by multi-head self-attention networks; For the mask fusion layer, which integrates feature information of nodes of different scales and realizes classification. (f) for the classification results of hyperspectral remote sensing image.

Next, we will detail the key steps of our proposed SAP-MAGCNs algorithm.

A. PRETREATMENT BY SUPERPCA

Normally, hyperspectral remote sensing images contain hundreds of pixels, which may increase the running complexity of our proposed SAP-MAGCNs frameworks if used directly for classification operations. In order to effectively solve this defect, we used the SuperPCA [46] algorithm was carried out on the original remote sensing image segmentation preprocessing operations such as super pixels, reducing proposed frameworks complexity and strengthening of the remote sensing image spectrum space correlation, then helped to keep the local spatial structure of hyperspectral remote sensing image information, and to further highlight the correlation between similar pixels.

B. MULTI-SCALE PRUNING BY STACK ATTENTION

To further eliminate subspace pixel block of different pixel between the redundant information, we use the stack attention module to pruning the fully connect pixel blocks and according to the different scale, will be the same pixel block to build the fully connect pixel cut to contain different number of nodes of the subpixel block, different scales is completed construction of graph nodes. In the process of constructing subpixel blocks [47], we need to calculate the information gain of both the global and local attribute to ensure the optimal partitioning. The subpixel block construction is shown in equation 1.

$$\alpha_{t,x} = \lambda\alpha_{t,x}^{global} + \gamma\alpha_{t,x}^{local}, \quad \gamma \approx \frac{1}{2} - \lambda. \tag{1}$$

where, γ, λ indicates factor of weights. $\alpha_{t,x}^{global}$ and $\alpha_{t,x}^{local}$ indicates the probability of global and local attention, respectively. $\alpha_{t,x}^{global}$ and $\alpha_{t,x}^{local}$ as:

$$\begin{cases} \alpha_{t,x}^{global} = \dfrac{exp(score(h_t, \bar{h}_x))}{\sum_{x'=1}^{T_x} exp(score(h_t, \bar{h}_{x'}))} \\ \alpha_{t,x}^{local} = \dfrac{exp(score(h_t, \bar{h}_x))}{\sum_{T_x\cdot\delta(v_p^T tanh(W_p h_t))-D}^{T_x\cdot\delta(v_p^T tanh(W_p h_t))+D} exp(score(h_t, \bar{h}_{x'}))} \cdot \\ exp(-\dfrac{(x - T_x \cdot \delta(v_p^T tanh(W_p h_t)))^2}{8D^2}) \end{cases} \tag{2}$$

where, δ indicates activation function.

When each local attribute is used as the dependent attribute of other global attributes, the original image graph structure data can be divided into N different subspace pixel block and further pruned to improve the difference between of different scales. Thus, N multi-scale graphs with different nodes are constructed. The specific build process is shown in figure 2.

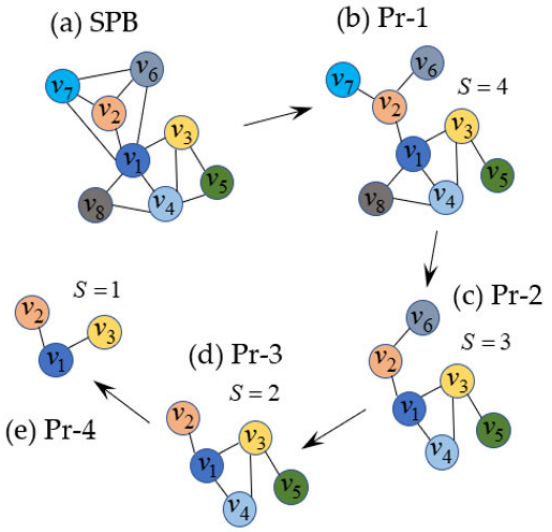


FIGURE 2. The pruning process of multi-scale pixel blocks by stack attention. "Pr-X" indicates pruning; "SPB" indicates super pixel fully connect blocks.

In figure 2, (a) is the initial graph data constructed after fully connect pixel block segmentation; (b) the subpixel block after the first pruning of the dependency attribute, the irrelevant connection edges are removed, that is, the large-scale figure $S = 4$; (c) is the dependent decision tree after the second pruning. It can be seen that v_7 and v_8 pixel are pruning, and irrelevant connection edges are pruning at the same time; (d) is the subpixel blocks after the third pruning, and it can be seen that the pixel v_6 is pruning to form a small scale figure $S = 2$; (e) is the subpixel block after the fourth pruning. It can be seen that the pixel v_4 and v_5 are pruning to form a small-scale figure $S = 1$.

Multiscale information has been widely applied in hyperspectral remote sensing image classification task [33], [48], [49], because the surface object in the hyperspectral remote sensing image often has a number of different geometric shape, the different scales of structure information can be from different levels, different angles describe the content of the complex image, this method can not only through different neighborhood scale figure to get multi-scale structure on optical spatial information, at the same time rely on stack attention pruning the original figure data structure, can effectively reduce the redundant information on graph nodes and edges, meanwhile, also reduces the SAP-MAGCNs frameworks complexity, and further highlights the relevance between adjacent pixel block.

Specifically, through the multi-scale graph after pruning by relying on the decision tree, on the s-scale, each pixel x_i is connected with other adjacent fully connect pixel blocks, while a certain fully connect pixel block V_i is pruned for many times, and the node set of the multi-scale graph formed is shown in equation 3.

$$\gamma_s(v_i) = \gamma_{s-1}(v_i) \cup \gamma_1(\gamma_{s-1}(v_i)). \tag{3}$$

where $\gamma_0(v_i) = v_i$,and $\gamma_1(v_i)$ is the set l-order of v_i in the subtree.

C. GRAPH BULID BY SELF-GUIDED ATTENTION

The stack attention is to pruning the whole fully connect pixel block super pixel block into a subspace pixel block and construct an adjacency matrix. However, this pruning is predefined in that the connectionless edge weight is directly defined as 0 during pruning. Such a pruning strategy may remove the part relevant information from the original fully connect super pixel block. Therefore, to effectively solve this problem, we designed a multi-head self-attention guide module to redistribute the edge weight to the trimmed subspace pixel block, further highlighting the correlation and interaction between graph nodes, namely, building stronger multi-scale graph structure.

For self-guided attention module, we adopt the strategy of multi-head self-attention [48]–[50] to transform the multi-scale subspace pixel block into a fully connected graph. And constructs the adjacency matrix A' from attention guidance, meanwhile, the weights have been improved of edges. Then, each A' corresponds to a certain fully connected

graph and each entry A'_{ij} is the weight of the edge going from node i to node j. The self-attention machine can capture the interactions between two arbitrary positions of a single sequence, and attention for inducing relations between nodes, (See part d in figure 1). The calculation of A' is shown in equation 4.

$$\alpha_{mhsa_{ij}} = \frac{exp(LeakyReLu(a^T[w\vec{v}_i||w\vec{v}_j]))}{\sum_{k\in N_i} exp(LeakyReLu(\vec{a}^T[w\vec{v}_i||w\vec{v}_k]))} \tag{4}$$

where T indicates the transpose of matrix; w indicates the weights of nodes. N_i indicates the neighbor nodes of node i, $LeakyReLu(\cdot)$ indicates [51]–[53] activate function. Namely, M matrices are constructed, where N is a hyper-parameter. The four heads were used for our SAP-MAGCNs models.

The pruning subspace pixel block adjacency matrix as the initialization input, the dependencies between nodes can be further captured by the multi-head self-attention guidance module. Meanwhile, the category attribute information of boundary pixel is captured to the maximum extent and enhance the ability of interaction between pixel node and subpixel block, and the classification accuracy of boundary pixel is improved.

D. DYNAMIC AGGREGATES MULTI-SCALE GCNS

Graph convolutional network (GCN) [10], [41]–[44], [54] is also a kind of direct effect on graph structure data, and each node in the graph is associated with a corresponding label, and the label is transferred through transfer and sink node information to obtain unlabeled node labels. Formally, the definition of an undirected graph is shown in equation 5.

$$\varsigma = (v, \varepsilon) \tag{5}$$

where $v = \{v_1, v_2, \ldots, v_n\}$ indicates the set of nodes, $\varepsilon = \{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n\}$ indicates a set of edge. The notation A' the adjacency matrix of ς which indicates whether each pair of nodes is connected and can be calculated as.

$$A'_{ij} = \begin{cases} \alpha_{mhsa_{ij}}, & v_i \neq v_j \\ 1, & v_i = v_j \\ e^{-\kappa||v_i - v_j||^2}, & otherwise \end{cases} \tag{6}$$

where, $\alpha_{mhsa_{ij}}$ indicates the weights of neighbor nodes by multi-head self-guided attention coefficient; the parameter κ is empirically set to 0.2 in the experiments [26], [47].

However, our propose SAP-MAGCNs frameworks including multiple multi-scale dynamic aggregates GCNs module (in FIGURE 1 (e) dynamic aggregates graph convolution networks), and each dynamic aggregates GCNs module including multiple (r) aggregates graph convolution layers. Thus, the module of adjacency matrix can be dynamic update as.

$$A'^{(r)} \leftarrow A(A'^{(r-1)} + \alpha h^{(r-1)}h^{(r)T})A^T + \beta I \tag{7}$$

where $h^{(r-1)}$ indicates the output features of $(r-1)^{th}$ layers; α, β indicates the coefficient of association. $h^{(r)}$ Can indicates as.

$$h^{(r)} = \delta(\widetilde{A}h^{(r-1)}W^{(r)}) \tag{8}$$

where δ indicates activate function of $softplus(\cdot)$; and $\widetilde{A}$ can be as show in equation (9).

$$\begin{cases} I + D^{-\frac{1}{2}}A'D^{-\frac{1}{2}} \rightarrow \widetilde{D}^{-\frac{1}{2}}\widetilde{A}\widetilde{D}^{-\frac{1}{2}} \\ \widetilde{D}_{ij} = \sum_j \widetilde{A}_{ij} \end{cases} \tag{9}$$

where I indicates the identity matrix.

Since we have M different attention guided adjacency matrices, M separate densely connected layers are required. Accordingly, we modify the computation of each layer as follows (for the l^{th} matrix $\widetilde{A}^t$) in equation 10.

$$H_{ti}^{(l)} = \delta(\sum_{j=1}^{n} \widetilde{A}_{ij}^t W_t^{(l)} Z_j^l + b_t^{(l)}) \tag{10}$$

where $t = 1, \ldots, M$ and t selects the weight matrix and bias term associated with the attention guided adjacency matrix $\widetilde{A}^t$. $Z_j^{(l)}$ as.

$$Z_j^{(l)} = [X_j, h_j^{(l)}, \ldots, h_j^{(l)}] \tag{11}$$

where, each densely connected layer has L sub-layers.

For masked Fusion module, our can intuitive to represent the classification network as a including masked function fully-connected layer, namely, we can concatenate the output of the g_{st} GCNs along the column dimension, and the calculation as.

$$FCN_{g_{st}} = SoftMax([GCN_{g_{s0}}, \ldots, GCN_{g_{st}}]W_{fcn}) \tag{12}$$

where $st = 0, 1, 2, 3$, W_{fcn} indicates the weights of fully connected layers.

The dynamic aggregates and graph constructure process of SAP-MAGCNs is summarized in algorithom 1.

Algorithm 1: The Dynamic Aggregates and Graph Constructure Process of SAP-MAGCNs

Input:input fully connect pixel block X; the number of pruning S by stack attention; the initial adjacency matrices A ;
1. for $s = 1$ to S
2. calculate the subspace pixel block x_i,and $x_i = SAP(X)$, $i = 1, 2, 3, 4$ (in FIGURE 2)
3. Update the multi-scale graph g_{st} and adjacency matrix $A_\prime$ (in equation (6))
4. input the SAP-MAGCNs, and achieve the masked fusion $FCN_{g_{st}}$ (in equation (12))
5. Optimizer the graph structure by $\tau(g_{st})$
output:optimization structure of the multi-scale graph

In algorithm 1, $\tau(g_{st})$ indicates the loss function, and can be calculated as.

$$\tau(g_{st}) = \sum_{i-1}^{s} \tau(g_{st}^i), \quad i = 1, 2, 3, 4 \tag{13}$$

where $\tau(g_{st})$ indicates the cross-entropy error loss, namely, the cross-entropy error is adopted to penalize the difference between the network output and the labels of the original labeled examples.

What has been discussed above, we proposed the SAP-MAGCNs hyperspectral remote sensing image classification method, the correlation between nodes are emphasized, utmost ground to capture and retain the boundary pixels, the category of the attribute information, at the same time, this paper introduced a concept of stack attention pruning multi-scale, capture the surface object in the remote sensing image of local and global spatial semantics, and use the multi-head self-guided attention modules on the multi-scale characteristic figure gradually refinement, the increase of capability of characterization of feature maps.

IV. EXPERIMENTAL PARAMETERS AND RESULTS

Next, we conduct exhaustive experiments to validate the effectiveness and dependability of the proposed SAP-MAGCNs method, and also provide the corresponding algorithm analyses. To be specific, we first compare SAP-MAGCNs with other state-of-the-art approaches on two publicly available hyperspectral image reference datasets, where five metrics including per-class accuracy, overall accuracy (OA), average accuracy (AA) and kappa coefficient are adopted. Then, we demonstrate that both the multi-scale manipulation and dynamic graph design in our SAP-MAGCNs are beneficial to obtaining the promising performance. After that, we validate the effectiveness of our method in dealing with the boundary regions.

A. PUBLICLY AVAILABLE DATASETS

1) INDIAN PINES [57]

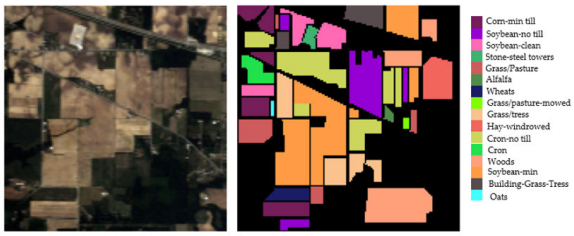
This dataset was collected from the Indian Pines experimental base in Indiana in the United States in 1992 using AVIRIS. The image size was 145 * 145, that is, 21,025-pixel blocks. The spatial resolution is 20 meters and the number of spectral bands is 220.In order to reduce the impact of noise on classification accuracy, 20 bands with high noise were removed, and the remaining 200 bands were tested. The image contains 16 different types of surface objects, including corn fields, grassland, wheat, wood, soybeans and stone steel towers. FIGURE 3 exhibits the false color image and ground-truth map of the Indian Pines dataset. As can be seen from table 1, there are 16 categories and the number of labeled pixels is 10249.

The amounts of labeled of various classes are listed in TABLE 1.

2) SALINAS [58]

The total spectral band of this data set is 224, after removing the redundant band, it is 204. The spatial resolution of the image is 3.7 meters and the size is 512*217. There are a total of 111104 pixels, among which 54129 pixels are used for classification (the remaining pixels are background pixels, that is, unavailable pixels) and include 16 surface objects such as arable land and celery. Figure 4 exhibits the false color image and ground-truth map of the Indian Pines dataset. As can be seen from table 2, there are 16 categories and the number of labeled pixels is 54129.

The amounts of labeled of various classes are listed in TABLE 2.



(a) false color image of Indian Pines (b) The ground-truth map of Indian Pines

FIGURE 3. The datasets of Indian pines.

TABLE 1. The labeled of indian pines.

ID	Class	Samples
1	Alfalfa	46
2	Corn-no till	1428
3	Cron-min til	830
4	Cron	237
5	Grass/Pasture	483
6	Grass/tress	730
7	Grass/pasture-mowed	28
8	Hay-windrowes	478
9	Oats	20
10	Soybean-no till	972
11	Soybean-min till	2455
12	Soybean-clean till	593
13	Wheats	205
14	Woods	1265
15	Buliding-Grass-Tress	386
16	Stone-steel towers	93

TABLE 2. The labeled of salinas.

ID	Class	Samples
1	Brocoli-green-weeds-1	2009
2	Brocoli-green-weeds-2	3726
3	Fallow	1976
4	Fallow_rough_plow	1394
5	Fallow_smooth	2678
6	Stubble	3959
7	Celery	3579
8	Grapes_untrained	11271
9	Soil_vinyard_develop	6203
10	Corn_senesc_green_weeds	3278
11	Lettuce_romaine_4wk	1068
12	Lettuce_romaine_5wk	1927
13	Lettuce_romaine_6wk	916
14	Lettuce_romaine_7wk	1070
15	Vinyard-untrained	7268
16	Vinyad_vertical_trellis	1807

B. EXPERIMENTAL PARAMETERS

In our experiments, the proposed algorithm is implemented via keras with SGD optimizer. During training, 80% of the labeled examples are used to learn the network parameters

TABLE 3. The classification results of indian pines.

ID	R2DCNN [54]	DRCNN [55]	JSDF [58]	MSDC [56]	MDGCN [26]	nonlocal GCN [2]	mini-batch GCN [3]	SAP-MAGCNs
1	55.49	81.07	91.24	85.29	100.00	100.00	100.00	100.00
2	72.33	91.20	90.76	94.12	88.11	78.32	83.21	97.90
3	69.81	92.21	88.92	84.22	97.59	96.38	90.36	97.59
4	97.84	100.00	100.00	100.00	87.50	91.67	95.83	100.00
5	87.75	97.32	88.90	97.57	95.84	97.91	93.75	95.83
6	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
7	97.35	100.00	98.63	99.51	100.00	100.00	100.00	100.00
8	100.00	100.00	100.00	92.49	100.00	100.00	100.00	100.00
9	75.89	83.29	79.67	85.90	100.00	100.00	100.00	50.00
10	89.46	99.67	97.15	94.58	93.81	91.75	91.75	94.85
11	89.22	92.45	97.09	97.05	86.59	74.79	77.64	95.53
12	57.13	94.59	78.82	85.70	93.22	94.91	98.30	100.00
13	88.76	99.52	98.16	87.22	100.00	100.00	100.00	100.00
14	100.00	100.00	100.00	99.97	96.03	94.44	96.82	100.00
15	99.56	99.69	100.00	98.87	97.44	97.43	97.43	100.00
16	92.74	99.02	98.62	98.16	100.00	100.00	100.00	100.00
OA	73.29	87.16	89.39	89.41	92.88	88.48	88.39	97.65
AA	85.83	95.63	94.24	93.79	96.01	94.85	95.32	95.73
Kappa	69.82	84.95	87.30	88.29	91.92	93.26	94.31	97.33

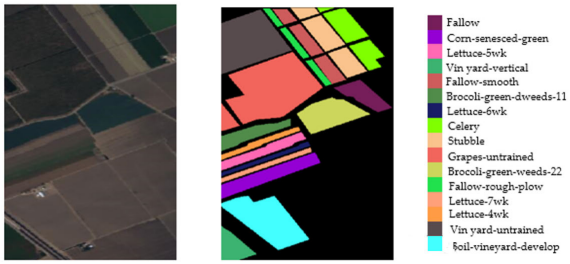


FIGURE 4. The datasets of Salinas (a) False color image. (b) Ground-truth map.

and 20% are used as validation set to tune the hyperparameters. Meanwhile, all the unlabeled examples are used as the test set to evaluate the classification performance. The network architecture of our proposed SAP-MAGCNs is kept identical for all the datasets. Specifically, three neighborhood scales, namely $s = 1$, $s = 2$, $s = 3$ and $s = 4$, are respectively employed for graph construction to incorporate multi-scale spectral-spatial information into our model. For each scale, we employ three graph convolutional layers with 32 hidden units, the learning rate and the number of training epochs are set to 0.005 and 500, respectively. The self-attention head number is set to 4.

To evaluate the classification ability of our proposed method, other recent state-of-the-art hyperspectral image classification methods are also used for comparison. Specifically, we employ same CNN-based methods, i.e., R-2D-CNN [53]. Diverse Region based deep CNN [54] and MS-DCNN [56]. Joint collaborative representation and SVM with Decision Fusion (JSDF-SVM) [58]. By contrast, GCN-based methods such as multi scale dynamic graph convolution networks (MDGCN) [26].

C. EXPERIMENTAL RESULTS

To show the effectiveness of our proposed SAP-MAGCNs, here we evaluate the classification performance by comparing SAP-MAGCNs with the aforementioned baseline methods.

1) THE CLASSIFICATION RESULTS OF INDIAN PINES

The results are as follows in table 3. In table 3, we present the SAP-MAGCNs model classification results in indian pines dataset.

FIGURE 5 exhibits a visual comparison of the classification results generated by different methods on the Indian Pines dataset, and the ground-truth map is provided in Figure.5(b).

In table 4, S represents the scale Figure 6 shows a visual comparison of the classification results generated by the different parameters of the method we proposed on the Indian pine datasets, and the ground truth graph is provided in figure 6 (b).

The classification results obtained by different methods on the Indian Pines dataset are summarized in Table 3 and Table 4. We observe that the traditional convolution neural networks (CNN) based methods including R-2D-CNN and DR-CNN achieve relatively low classification accuracy, which is due to the reason that they can only conduct the convolution on a regular image grid, so the specific local spatial information cannot be captured.

By contrast, Multi Scale-based methods such as JSDF and MSDC achieve relatively good classification accuracy. which combines spectral and spatial information in diverse scales, ranks in the second place. This implies that the multiscale spectral-spatial features are quite useful to enhance the classification performance. The GCN-based methods such as MSGCNs, SAGGCNs and MDGCN are capable of adaptively aggregating the features on irregular non-Euclidean regions, so they can yield better performance than others methods such as R-2D-CNN, DR-CNN, JSDF and MSDC.

For the nonlocal GCN and min-batch GCN methods, the performance of classification outperformance others methods such as MSDC and DRCNN *et al.* The mainly reason is improving the interaction of neighborhood nodes by GCN layers and mini-batch strategy. In the table 5 and Figure 6, we can observe SAP-MAGCNs methods outperformance the

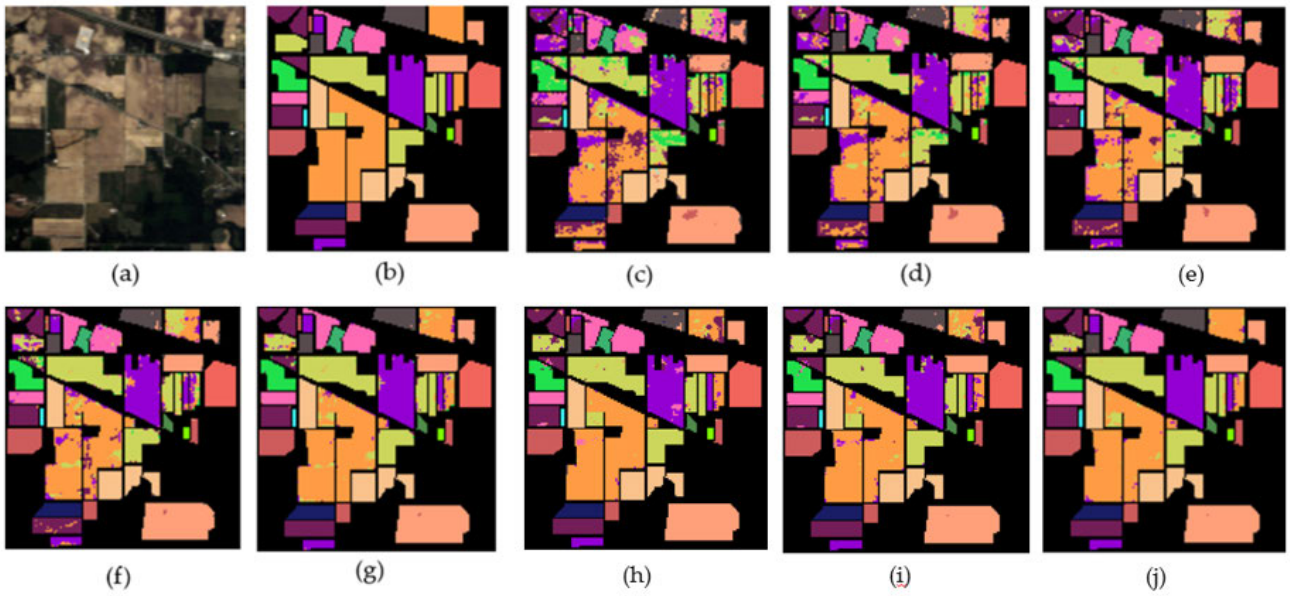


FIGURE 5. The classification results generated by different methods. (a) False color images. (b) ground truth map; (c) R-2D-NN; (d) DRCNN; (e) JSDF; (f) MSDC; (g) MDGCN; (h) nonlocal GCN; (i) mini-batch GCN; (j) SAP-MAGCNs.

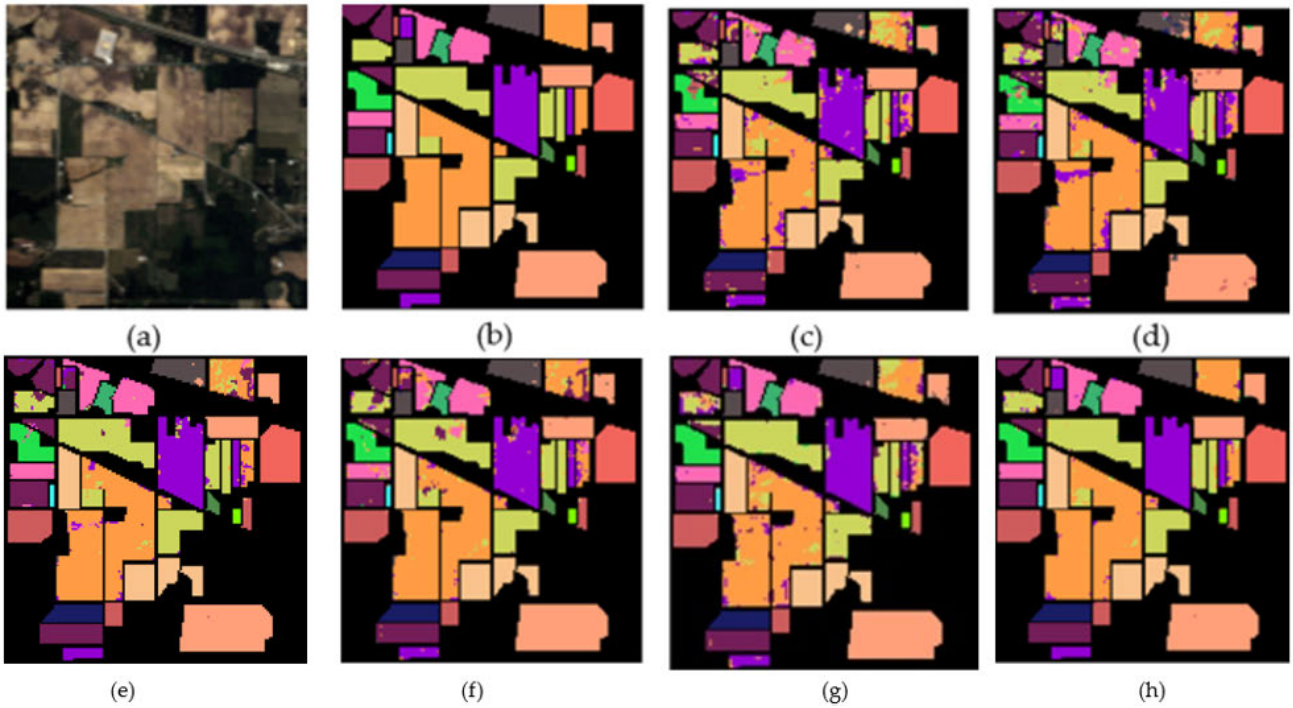


FIGURE 6. The classification results generated by different methods. (a) False color images. (b) ground truth map; (c) S = 1,2; (d) S = 3,4; (e) MSGCNs; (f) SAGGCNs; (g) SAP-MAGCNs (Non) indicates the frameworks not including the SuperPCA block. (h) SAP-MAGCNs; Among them, MSGCNs stands for no self-attention guidance layer.

SAP-MAGCNs (Non) methods by 6.92%, 0.21% and 7.84%, respectively.

Furthermore, we observe that the proposed SAP-MAGCNs achieves the best performance among all the methods in terms of OA, AA, and Kappa coefficient, and the standard deviations are also very small, which reflects that the proposed SAP-MAGCNs models is more stable and effective than the compared methods.

2) THE CLASSIFICATION RESULTS OF SALINAS DATASETS

The results are as follows:

In table 5, we present the SAP-MAGCNs model algorithm and others methods.

Figure 7 exhibits a visual comparison of the classification results generated by different methods on the Indian Pines dataset, and the ground-truth map is provided in Figure.7(b).

TABLE 4. The classification results of our models.

ID	S=1,2	S=3,4	MSGCNs	SAGGCNs	SAP-MAGCNs (Non)	SAP-MAGCNs
1	100.00	100.00	100.00	100.00	100.00	100.00
2	90.90	83.21	93.00	93.71	84.61	97.90
3	95.18	96.38	98.80	100.00	93.97	97.59
4	100.00	95.83	100.00	91.67	91.66	100.00
5	91.67	95.83	97.91	97.92	95.83	95.83
6	100.00	100.00	100.00	100.00	100.00	100.00
7	100.00	100.00	100.00	66.67	100.00	100.00
8	100.00	100.00	100.00	100.00	100.00	100.00
9	50.00	50.00	50.00	50.00	100.00	50.00
10	94.85	93.81	95.87	98.97	91.75	94.85
11	85.77	78.46	91.46	93.90	80.48	95.53
12	91.53	91.53	96.61	93.22	96.61	100.00
13	100.00	100.00	100.00	100.00	100.00	100.00
14	98.41	96.03	99.20	98.41	96.03	100.00
15	97.44	97.44	97.44	97.44	97.43	100.00
16	100.00	100.00	100.00	100.00	100.00	100.00

TABLE 5. The classification results of salinas.

ID	R2DCNN [54]	DRCNN [55]	JSDF [58]	MSDC [56]	MDGCN [26]	SAP-MAGCNs
1	99.25	100.00	100.00	100.00	100.00	100.00
2	100.00	100.00	99.73	99.87	100.00	100.00
3	97.72	99.74	99.49	100.00	100.00	100.00
4	94.27	98.20	98.57	98.92	99.64	100.00
5	93.10	97.20	98.32	98.32	98.88	99.25
6	99.87	100.00	100.00	100.00	100.00	100.00
7	93.30	99.86	99.86	100.00	100.00	100.00
8	66.33	76.44	78.17	80.12	87.84	91.84
9	99.92	100.00	100.00	100.00	100.00	100.00
10	93.29	95.42	95.88	96.95	98.62	98.48
11	98.59	99.06	98.59	99.53	99.06	99.07
12	89.35	100.00	99.74	99.48	100.00	100.00
13	87.98	99.40	100.00	100.00	100.00	100.00
14	90.65	99.06	99.53	99.53	99.53	100.00
15	75.45	79.02	81.15	83.90	89.40	94.36
16	96.12	96.68	97.78	98.34	99.44	99.45
OA	87.70	91.64	92.39	93.30	95.85	97.38
AA	92.57	96.26	96.67	97.18	98.28	98.90
Kappa	86.34	90.70	91.54	92.56	95.38	97.08

TABLE 6. The classification results of sap-MAGCNS in salinas datasets.

ID	S=1,2	S=3,4	MSGCNs	SAGGCNs	SAP-MAGCNs
1	99.50	99.04	99.50	100.00	100.00
2	100.00	100.00	100.00	100.00	100.00
3	100.00	98.98	100.00	100.00	100.00
4	100.00	100.00	100.00	100.00	100.00
5	95.15	96.27	98.13	98.51	99.25
6	100.00	100.00	100.00	100.00	100.00
7	99.44	99.16	99.44	99.44	100.00
8	77.46	70.54	74.09	84.47	91.84
9	100.00	99.35	100.00	99.84	100.00
10	95.43	91.77	94.51	96.04	98.48
11	98.13	99.06	99.07	98.13	99.07
12	100.00	100.00	100.00	100.00	100.00
13	100.00	98.90	100.00	100.00	100.00
14	99.06	100.00	99.07	100.00	100.00
15	84.31	88.17	93.26	94.09	94.36
16	99.45	97.79	99.45	99.45	99.45
OA	92.61	91.28	93.16	95.55	97.38
AA	96.77	96.19	97.28	98.12	98.90

Figure 8 shows a visual comparison of the classification results generated by the different parameters of the method we proposed on the Salinas dataset, and the ground truth map is provided in figure 8 (b).

From the quantitative results in table 4 and 5, the following conclusions can be obtaining:

(1) R-2D-CNN, DR-CNN and other classification algorithms are the lowest in multiple evaluation indexes (OA, AA and Kappa), because when convolutional neural network captures the deep information of surface objects in images, it will ignore a large amount of global information.

(2) MS - DCNN due to capture the surface of the object in the image classification algorithm of multi-scale spatial structure information, make its classification accuracy is R-2D-CNN and DR-CNN classification algorithms such as CNN (CNN) is high, because this method captures the different dimensions of spatial structure information, strengthen the correlation between different scale features at the same time, in addition, a further sign of multi-scale spatial structure information, can effectively improve the classification accuracy of the surface of the object in the image.

(3) the classification algorithm proposed in this chapter is most effective in terms of overall accuracy (OA), average accuracy (AA), Kappa coefficient. Because, SAP-MAGCNs classification algorithm not only can effectively capture the

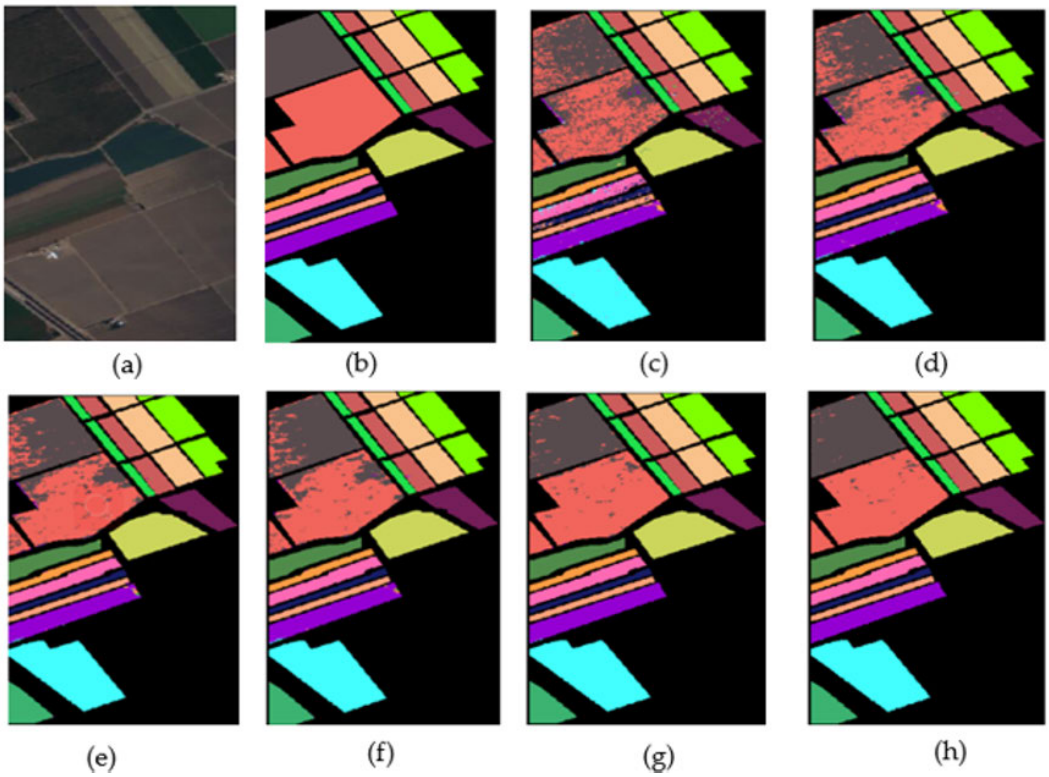


FIGURE 7. The classification results generated by different methods. (a) False color images. (b) ground truth map; (c) R-2D-NN; (d) DRCNN; (e) JSDF; (f) MSDC; (g) MDGCN; (h) SAP-MAGCNs.

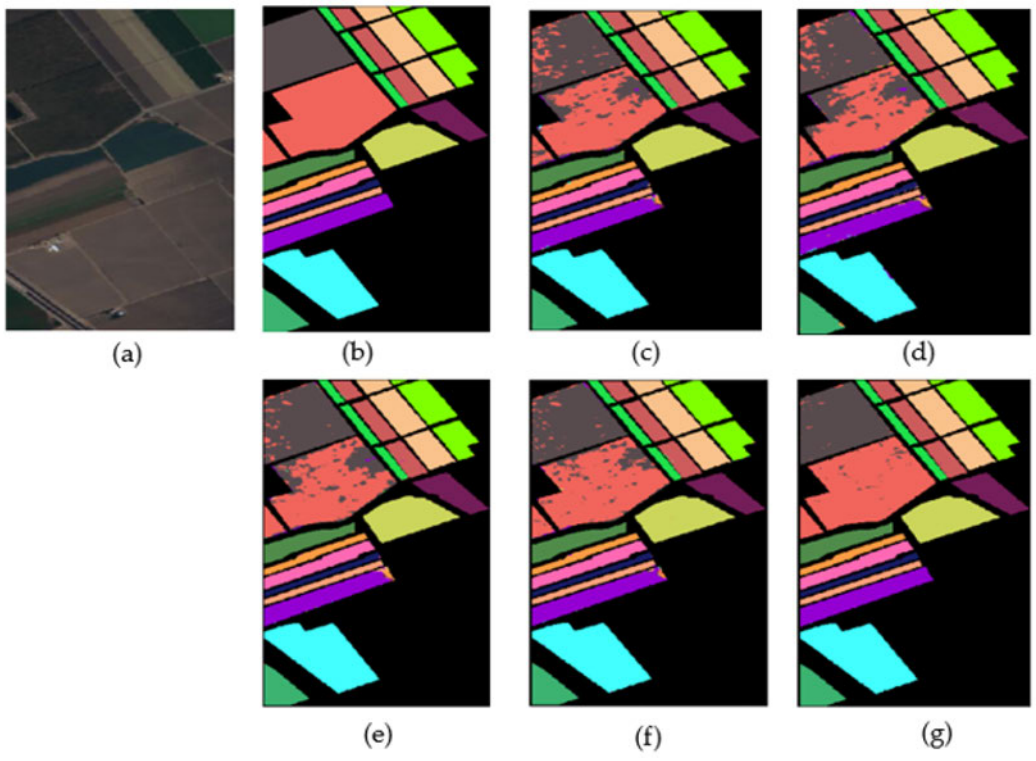


FIGURE 8. The classification results generated of Salinas datasets by different methods. (a) False color images. (b) ground truth map; (c) S = 1,2; (d) S = 3,4; (e) MSGCNs; (f) SAGGCNs; (g) SAP-MAGCNs; Among them, MSGCNs stands for no self-attention guidance layer.

surface of the object in the image of the deep semantic information, at the same time through information gathering node could improve the ability of features of characterization, and maximize the extraction of the boundary pixels,

TABLE 7. The classification results of different quantities salinas.

ID	Train = 0.1	Train = 0.2	Train = 0.3	Train = 0.4	Train = 0.5	SAP-MAGCNs
1	96.26	100.00	99.93	99.81	99.50	100.00
2	98.79	79.30	99.42	99.73	99.56	100.00
3	2.34	31.30	82.22	81.53	89.56	100.00
4	1.34	84.75	98.47	95.60	97.67	100.00
5	9.57	77.59	34.59	94.54	97.71	99.25
6	97.09	99.81	99.33	99.93	100.00	100.00
7	94.37	99.47	100.00	99.93	99.79	100.00
8	87.61	13.67	62.67	61.55	75.68	91.84
9	60.46	78.09	90.44	99.94	100.00	100.00
10	1.29	69.37	84.05	87.19	91.42	98.48
11	24.94	93.68	97.54	95.67	95.90	99.07
12	0.00	1.49	42.15	82.75	99.68	100.00
13	19.50	88.67	87.58	98.09	99.31	100.00
14	89.36	83.88	88.31	94.28	98.83	100.00
15	5.53	89.68	78.48	79.88	83.70	94.36
16	54.42	82.50	98.47	97.03	98.06	99.45
OA	54.70	66.06	80.66	86.49	91.41	97.38
AA	46.43	73.33	83.98	91.73	95.39	98.90
Kappa	48.57	62.89	78.60	85.02	90.47	97.08

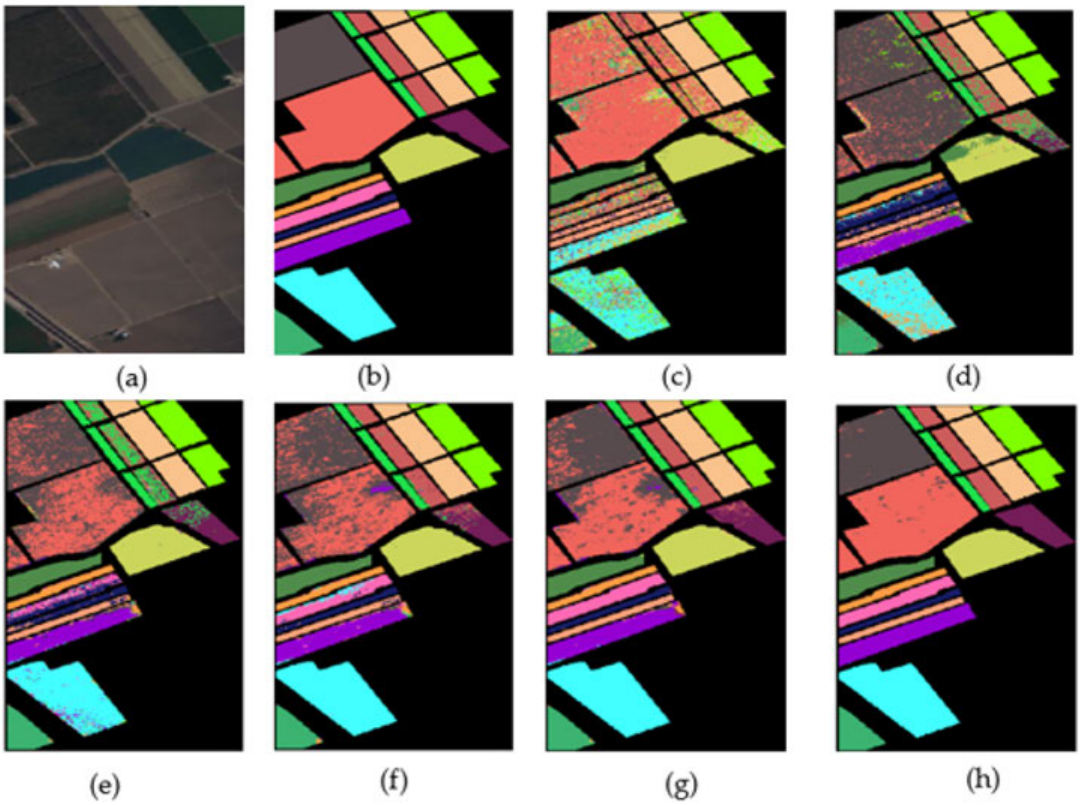


FIGURE 9. The classification results generated of salinas datasets by different training datasets. (a) False color images. (b) ground truth map; (c) train = 0.1(d) train = 0.2; (e) train = 0.3; (f) train = 0.4; (g)train = 0.5; (h) SAP-MAGCNs; Among them," Train" is training datasets of salinas.

the category of the attribute information, such as further illustrates the advantages of SAP-MAGCNs classification algorithm, compared with other classification algorithms, this method is stable and good robustness.

3) THE CLASSIFICATION RESULTS OF DIFFERENT QUANTITIES OF SALINAS DATASETS ON THE SAP-MAGCNS MODEL

In order to further verify the effectiveness of the SAP-MAGCNs model method, we will use different training sets to conduct training tests on the model. Among them, the Salinas data sets for training are 10%, 20%, 30%, 40%, 50%, etc. The results of classification on the Salinas data set are shown in table 7. In table 7, we present the SAP-MAGCNs model algorithm.

As can be seen from table 7: As the amount of training data increases, the classification accuracy of SAP-MAGCNs method increases. Moreover, when the training data is 10%, a better classification effect is obtained. This is because the self-attentional guidance layer in SAP-MAGCNs enhances

the subtlety while removing irrelevant redundant information. It is further proved that SAP-MAGCNs method can also achieve good classification effect on small sample data sets. Figure 9 shows a visual comparison of the classification results generated by the different parameters of the method we proposed on the salinas dataset, and the ground truth map is provided in figure 9 (b).

It is obvious from figure 9 that with the increase of the training data set, the classification effect on SAP-MAGCNs method gradually gets better. And it further proves that the number of training samples will affect the accuracy of classification. The more training samples, the higher the accuracy.

V. CONCLUSION

In this paper, a hyperspectral image classification method based on stack-attention-pruning aggregates multiscale graph convolutional network (SAP-MAGCNs) is proposed. The dependent decision tree by stack attention is used to prune the graph data, which reduces the computational complexity, strengthens the relevance and interaction between different nodes of the subtree species, and simplifies and reconstructs the subtree with different dependent attributes. Pair then design the boot module, since the attention in the tree node distribution of the weight and the maximum retained the boundary pixels, the category of the attribute information, as well as to the different scales of subtree are realized, finally, using multi-scale connection diagram convolution dense network of arithmetic operation, the capture of hyperspectral remote sensing image multi-scale local and global semantic information, at the same time to strengthen the dependency relationship between-n multi-scale feature maps. The experimental results on three widely-used hyperspectral image datasets demonstrate that the proposed SAP-MAGCNs is able to yield better performance when compared with the state-of-the-art methods.

Next work we will continue to explore the semantic information of remote sensing image of classification tasks. Namely, we will further design a simply and powerful semantic networks-based graph and attention pruning.

ACKNOWLEDGMENT

Under the careful guidance of my laboratory partners and my mentors, the authors successfully completed the experiment in this article and got satisfactory results. They would like to thank my sincere wishes and thanks to you for your good health and happiness.

REFERENCES

[1] X. Han, J. Yu, J. Luo, and W. Sun, "Reconstruction from multispectral to hyperspectral image using spectral library-based dictionary learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 3, pp. 1325–1335, Mar. 2019.

[2] M. Zhou, J. Shu, and Z. Chen, "Classification of hyperspectral remote sensing image based on genetic algorithm and SVM," *Proc. SPIE*, vol. 7809, no. 1, Aug. 2010, Art. no. 78090A.

[3] C. G. Xu and A. Anwar, "Based on the decision tree classification of remote sensing image classification method application," in *Proc. Appl. Mech. Mater.* Trans Tech Publications, vol. 316, 2013, pp. 193–196.

[4] F. Chen, R. Ren, T. Van De Voorde, W. Xu, G. Zhou, and Y. Zhou, "Fast automatic airport detection in remote sensing images using convolutional neural networks," *Remote Sens.*, vol. 10, no. 3, p. 443, Mar. 2018.

[5] B.-C. Kuo, C.-S. Huang, C.-C. Hung, Y.-L. Liu, and I.-L. Chen, "Spatial information based support vector machine for hyperspectral image classification," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2010, pp. 832–835.

[6] R. Hang, Q. Liu, H. Song, and Y. Sun, "Matrix-based discriminant subspace ensemble for hyperspectral image spatial–spectral feature fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 2, pp. 783–794, Feb. 2016.

[7] M. Fauvel, J. A. Benediktsson, J. Chanussot, and J. R. Sveinsson, "Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 11, pp. 3804–3814, Nov. 2008.

[8] B. Song, J. Li, M. D. Mura, P. Li, A. Plaza, J. M. Bioucas-Dias, J. A. Benediktsson, and J. Chanussot, "Remotely sensed image classification using sparse representations of morphological attribute profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 8, pp. 5122–5136, Aug. 2014.

[9] P. Zhong, Z. Gong, S. Li, and C.-B. Schonlieb, "Learning to diversify deep belief networks for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 6, pp. 3516–3530, Jun. 2017.

[10] X. Yang, Y. Ye, X. Li, R. Y. K. Lau, X. Zhang, and X. Huang, "Hyperspectral image classification with deep learning models," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 9, pp. 5408–5423, Sep. 2018.

[11] X. Dai, X. Wu, B. Wang, and L. Zhang, "Semisupervised scene classification for remote sensing images: A method based on convolutional neural networks and ensemble learning," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 6, pp. 869–873, Jun. 2019.

[12] J. Wang, Y. Wang, J. Dang, and L. Wei, "Target detection based on cascade network and densely connected network in remote sensing image," in *Proc. IEEE 16th Int. Conf. Dependable, Auton. Secure Comput.*, Aug. 2018, pp. 19–25.

[13] L. Wang, J. Zhang, P. Liu, K.-K. R. Choo, and F. Huang, "Spectral–spatial multi-feature-based deep learning for hyperspectral remote sensing image classification," *Soft Comput.*, vol. 21, no. 1, pp. 213–221, Jan. 2017.

[14] G. Cheng, C. Yang, X. Yao, L. Guo, and J. Han, "When deep learning meets metric learning: Remote sensing image scene classification via learning discriminative CNNs," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 5, pp. 2811–2821, May 2018.

[15] S. Piramanayagam, "Classification of remote sensed images using random forests and deep learning framework," *Proc. SPIE*, vol. 10004, Oct. 2016, Art. no. 100040L.

[16] Y. Chen, Z. Lin, X. Zhao, G. Wang, and Y. Gu, "Deep learning-based classification of hyperspectral data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2094–2107, Jun. 2014.

[17] X. Zhang, Y. Sun, K. Jiang, C. Li, L. Jiao, and H. Zhou, "Spatial sequential recurrent neural network for hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 11, pp. 4141–4155, Nov. 2018.

[18] B. Liu, X. Yu, A. Yu, P. Zhang, and G. Wan, "Spectral–spatial classification of hyperspectral imagery based on recurrent neural networks," *Remote Sens. Lett.*, vol. 9, no. 12, pp. 1118–1127, Dec. 2018.

[19] J. Luo, J. Wu, S. Zhao, L. Wang, and T. Xu, "Lossless compression for hyperspectral image using deep recurrent neural networks," *Int. J. Mach. Learn. Cybern.*, vol. 10, no. 10, pp. 2619–2629, Oct. 2019.

[20] H. Lee and H. Kwon, "Going deeper with contextual CNN for hyperspectral image classification," *IEEE Trans. Image Process.*, vol. 26, no. 10, pp. 4843–4855, Oct. 2017.

[21] Q. Wang, S. Liu, J. Chanussot, and X. Li, "Scene classification with recurrent attention of VHR remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 2, pp. 1155–1167, Feb. 2019.

[22] L. Li, X. Tingbao, and C. Yun, "Fuzzy classification of high resolution remote sensing scenes using visual attention features," *Comput. Intell. Neurosci.*, vol. 2017, Jul. 2017, Art. no. 9858531.

[23] G. Sumbul, R. G. Cinbis, and S. Aksoy, "Multisource region attention network for fine-grained object recognition in remote sensing imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 4929–4937, Jul. 2019.

[24] A. Qin, Z. Shang, J. Tian, Y. Wang, T. Zhang, and Y. Y. Tang, "Spectral–spatial graph convolutional networks for semisupervised hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 2, pp. 241–245, Feb. 2018.

[25] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," 2016, *arXiv:1609.02907*. [Online]. Available: https://arxiv.org/abs/1609.02907

[26] S. Wan, C. Gong, P. Zhong, B. Du, L. Zhang, and J. Yang, "Multiscale dynamic graph convolutional network for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 5, pp. 3162–3177, May 2020.

[27] J. Kang, R. Fernandez-Beltran, Z. Ye, X. Tong, P. Ghamisi, and A. Plaza, "Deep metric learning based on scalable neighborhood components for remote sensing scene characterization," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 12, pp. 8905–8918, Dec. 2020.

[28] J. Zhao, W. Gao, Z. Liu, G. Mou, L. Lu, and L. Yu, "A classification of remote sensing image based on improved compound kernels of SVM," in *Proc. Int. Conf. Comput. Comput. Technol. Agricult.*, Beijing, China, Oct. 2009, pp. 15–20.

[29] J. Li, J. M. Bioucas-Dias, and A. Plaza, "Spectral–spatial hyperspectral image segmentation using subspace multinomial logistic regression and Markov random fields," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 3, pp. 809–823, Mar. 2012.

[30] X. Zhu and W. Bao, "Comparison of remote sensing image fusion strategies adopted in HSV and IHS," *J. Indian Soc. Remote Sens.*, vol. 46, no. 3, pp. 377–385, 2018.

[31] T. Li, J. Zhang, and Y. Zhang, "Classification of hyperspectral image based on deep belief networks," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2014, pp. 5132–5136.

[32] C. Shi and C.-M. Pun, "Multi-scale hierarchical recurrent neural networks for hyperspectral image classification," *Neurocomputing*, vol. 294, pp. 82–93, Jun. 2018.

[33] J. Yang, Y. Zhao, J. C.-W. Chan, and C. Yi, "Hyperspectral image classification using two-channel deep convolutional neural network," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2016, pp. 5079–5082.

[34] H. Jiang and N. Lu, "Multi-scale residual convolutional neural network for haze removal of remote sensing images," *Remote Sens.*, vol. 10, no. 6, p. 945, Jun. 2018.

[35] Y. Zhou, X. Liu, J. Zhao, D. Ma, R. Yao, B. Liu, and Y. Zheng, "Remote sensing scene classification based on rotation-invariant feature learning and joint decision making," *EURASIP J. Image Video Process.*, vol. 2019, no. 1, pp. 1–11, Dec. 2019.

[36] P. Duan, X. Kang, S. Li, and P. Ghamisi, "Noise-robust hyperspectral image classification via multi-scale total variation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 6, pp. 1948–1962, Jun. 2019.

[37] Q. Xu, Y. Xiao, D. Wang, and B. Luo, "CSA-MSO3DCNN: Multi-scale octave 3D CNN with channel and spatial attention for hyperspectral image classification," *Remote Sens.*, vol. 12, no. 1, p. 188, Jan. 2020.

[38] Q. Xu, D. Wang, and B. Luo, "Faster multiscale capsule network with octave convolution for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 18, no. 2, pp. 361–365, Feb. 2021.

[39] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, "A comprehensive survey on graph neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 1, pp. 4–24, Jan. 2021.

[40] J. Zhou, G. Cui, Z. Zhang, C. Yang, Z. Liu, L. Wang, C. Li, and M. Sun, "Graph neural networks: A review of methods and applications," 2018, *arXiv:1812.08434*. [Online]. Available: https://arxiv.org/abs/1812.08434

[41] J. Bruna, W. Zaremba, A. Szlam, and Y. LeCun, "Spectral networks and locally connected networks on graphs," 2013, *arXiv:1312.6203*. [Online]. Available: http://arxiv.org/abs/1312.6203

[42] H.-J. Yoon, J. Gounley, M. T. Young, and G. Tourassi, "Information extraction from cancer pathology reports with graph convolution networks for natural language texts," in *Proc. IEEE Int. Conf. Big Data*, Dec. 2019, pp. 4561–4564.

[43] R. Liu, C. Xu, T. Zhang, W. Zhao, Z. Cui, and J. Yang, "Si-GCN: Structure-induced graph convolution network for skeleton-based action recognition," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2019, pp. 1–8.

[44] M. Rashid, H. Kjellström, and Y. Lee, "Action graphs: Weakly-supervised action localization with graph convolution networks," in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis.*, Dec. 2020, pp. 615–624.

[45] L. Yang, J. Zhuang, H. Fu, K. Zhou, and Y. Zheng, "SketchGCN: Semantic sketch segmentation with graph convolutional networks," 2020, *arXiv:2003.00678*. [Online]. Available: https://arxiv.org/abs/2003.00678

[46] D. Marcheggiani and L. Perez-Beltrachini, "Deep graph convolutional encoders for structured data to text generation," 2018, *arXiv:1810.09995*. [Online]. Available: https://arxiv.org/abs/1810.09995

[47] A. Qin, Z. Shang, J. Tian, Y. Wang, T. Zhang, and Y. Y. Tang, "Spectralspatial graph convolutional networks for semisupervised hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 2, pp. 241–245, Feb. 2019.

[48] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Sásstrunk, "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.

[49] H. S. Dhiman, D. Deb, and V. E. Balas, "Decision tree ensemble-based regression models," in *Supervised Machine Learning in Wind Forecasting and Ramp Event Prediction*. 2020, pp. 61–73.

[50] A. Vaswani *et al.*, "Attention is all you need," 2017, *arXiv:1706.03762*. [Online]. Available: https://arxiv.org/abs/1706.03762

[51] Z. Lin, M. Feng, C. Nogueira dos Santos, M. Yu, B. Xiang, B. Zhou, and Y. Bengio, "A structured self-attentive sentence embedding," 2017, *arXiv:1703.03130*. [Online]. Available: http://arxiv.org/abs/1703.03130

[52] Y. Zhang, Z. Guo, and W. Lu, "Attention guided graph convolutional networks for relation extraction," 2019, *arXiv:1906.07510*. [Online]. Available: https://arxiv.org/abs/1906.07510

[53] Z. Guo, Y. Zhang, Z. Teng, and W. Lu, "Densely connected graph convolutional networks for graph-to-sequence learning," *Trans. Assoc. Comput. Linguistics*, vol. 7, pp. 297–312, Nov. 2019.

[54] M. Zhang, W. Li, and Q. Du, "Diverse region-based CNN for hyperspectral image classification," *IEEE Trans. Image Process.*, vol. 27, no. 6, pp. 2623–2634, Jun. 2018.

[55] C. Zhang, G. Li, S. Du, W. Tan, and F. Gao, "Three-dimensional densely connected convolutional network for hyperspectral remote sensing image classification," *J. Appl. Remote Sens.*, vol. 13, no. 1, p. 1, Feb. 2019.

[56] L. Yutong, L. Zhiqing, and Y. Xiaoling, "Application of improved convolution neural network in remote sensing image classification," *J. Comput. Appl.*, vol. 25, pp. 1–12, Apr. 2018.

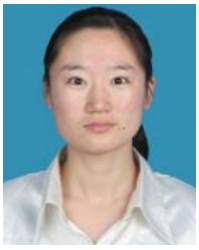
[57] C. Bo, H. Lu, and D. Wang, "Hyperspectral image classification via JCR and SVM models with decision fusion," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 2, pp. 177–181, Feb. 2016.

[58] S. Liu and Q. Shi, "Multitask deep learning with spectral knowledge for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 12, pp. 2110–2114, Dec. 2020.

NA LIU was born in Binzhou, Shandong, China, in 1984. She received the bachelor's degree in electronic information science and technology from Qingdao Agricultural University, in 2007, and the master's degree in control engineering from the Qingdao University of Science and Technology, in 2014. She has been teaching with Qingdao Huanghai University since 2007. She has published 14 articles, presided over and participated in six provincial projects, applied for one invention patent, and authorized seven utility model patents. Her research interests include detection technology and automation device, and intelligent control.

BIN ZHANG was born in Qingdao, Shandong, China, in 1985. He received the bachelor's degree in electronic information science and technology from Qingdao Agricultural University, in 2007. He has been involved in research and application of industrial automation control and won invention patents and Qingdao science and technology progress awards.

QIUHUAN MA was born in Heze, Shandong, China, in 1983. She received the master's degree in examine technology and automatization equipment from the Shandong University of Science and Technology, in 2010, She has been teaching with Qingdao Huanghai University since 2010. She has published more than 11 articles. Three Utility model patents, the research interests include electronic information and auto-control.



QINGQING ZHU was born in Heze, Shandong, China, in 1984. She received the bachelor's degree in automation from Weifang University, in 2006, and the master's degree in control theory and control engineering from the China University of Petroleum (East China), in 2009. Since 2010, she has been teaching with Qingdao Huanghai University and published more than 20 articles. Her research interests include the application of intelligent sensor and network technology.

XIAOLING LIU was born in Chifeng, Neimenggu, China, in 1984. She received the bachelor's degree in electronic information engineering and the master's degree in information and communication engineering from Dalian Maritime University, in 2007 and 2009, respectively. She has been teaching in electronics engineering with Qingdao Huanghai University since 2009. She has published more than ten articles. Her research interest includes information processing of offshore marine environment.