IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 16 June 2023, accepted 31 July 2023, date of publication 15 August 2023, date of current version 30 August 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3305382

RESEARCH ARTICLE

Local Discriminative Embedding Broad Learning System With Graph Convolutional for Hyperspectral Image Classification

WEI LI[1], YUANQUAN SHI[1], LIYUN LI[1], AND XIANGBO MA[2]
[1]College of Computer and Artificial Intelligence, Huaihua University, Huaihua, Hunan 418000, China
[2]Beijing Wanweisheng New Technology Company Ltd., Beijing 102200, China

Corresponding author: Liyun Li (liliyun168@hhtc.edu.cn)

This work was supported by the National Natural Science Foundation of China under Grant 62172182.

ABSTRACT Broad learning system has attracted increasing attention in the hyperspectral image (HSI) classification, due to its universal approximation capability and high efficiency. However, two main challenges remained. One is that the mapping from the original BLS input to the mapped feature (MF) is linear, which is difficult to fully represent the complex spatial-spectral features of HSI. The other is that BLS fails to explore the local geometric structure relationship between samples within HSI. To overcome the limitations mentioned above, we propose a local discriminative embedding broad learning system with graph convolutional (GDEBLS). To address the first challenge, GDEBLS utilizes the graph convolution operation to aggregate the node information in the adjacent graph to learn the context and obtain the rich nonlinear spatial-spectral features in HSI. To deal with the second challenge, our method utilizes a neighborhood selection approach based on manifold structure to calculate the true distances between samples in the manifold space, overcoming the limitations of Euclidean distance measurement. Next, We introduce local manifold structure and discriminative information into BLS. The experimental results show that the proposed method significantly surpasses other state-of-the-art methods.

INDEX TERMS Broad learning systems, hyperspectral image, local geometric structure, graph convolutional.

I. INTRODUCTION

Hyperspectral image (HSI) is obtained by imaging ground objects simultaneously in hundreds of adjacent narrow bands by an imaging spectrometer. Using this image can effectively improve the accuracy of ground object identification. Therefore, HSI is widely used in the fields of ground component analysis, target detection, and environmental monitoring, among which the classification of HSI is a very important basic problem. In order to complete the classification task, researchers have proposed many classical classification methods, such as support vector machine (SVM) [1], multiple logistic regression (MLR) [2], and sparse representation classifier (SRC) [3] et al.

Recently, convolutional neural networks (CNNs) effectively capture local correlations in hyperspectral images through local connections and weight sharing. Li et al. [4] proposed a pixel-pair-based CNN, which constructs new data combinations by pairing any two samples from the available labeled data. This method enables the construction of a large number of labeled pixel pairs, ensuring the thorough training of the deep network. Haut et al. [5] conducted principal component analysis on hyperspectral image data and used spatial patches centered around the pixels to be classified to train a 2D CNN, thereby accomplishing hyperspectral image classification tasks. While this method extracted rich spatial features, it overlooked the spectral information. The 3D-CNN network proposed by He et al. [6] jointly extracts spatial and spectral features by computing multi-scale features. Li et al. [7] proposed using two consecutive 3D convolutional layers for hyperspectral image classification.

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

This network architecture is simple and requires fewer parameters, thereby avoiding overfitting issues that may arise from complex network models. The aforementioned CNN-based methods have achieved good results. However, these methods face challenges such as gradient vanishing and network degradation when the CNN models are trained to deeper depths. To address this issue, Zhong et al. [9] utilized residual networks [8] to develop a deep network model that can continuously learn discriminative features from both spectral and spatial backgrounds. Song et al. [10] further improved the performance by incorporating deep and shallow feature fusion based on the residual network, to further enhance the utilization of each convolutional layer and obtain more effective discriminative features, spatial patches centered around pixels of interest are used to train a 2D CNN for hyperspectral image classification tasks. This approach aims to capture more relevant spatial information and improve the overall performance of the network. In recent years, graph convolutional networks (GCNs) have been widely applied to semi-supervised classification tasks in hyperspectral imaging (HSI). GCNs leverage both the local graph structure and node features to achieve effective information aggregation, making them well-suited for HSI analysis and classification. Qin et al. [11] proposed the spectral-spatial graph convolutional networks (S2GCN). S2GCN combines spectral information with local spatial neighborhood information to perform classification tasks on hyperspectral images. This approach integrates both spectral and spatial features to enhance the discriminative power of the model. Ding et al. [12] proposed the Global Consistent Graph Convolutional Network (GCGCN), the initial graph generation is treated as an optimization variable, allowing for the adaptation and refinement of the graph structure. This approach aims to obtain an adaptive and reliable graph structure for effective information aggregation in the GCN model. Wan et al. [13] proposed a model based on dynamic graphs and graph convolution methods, which were trained collaboratively. They also improved the model accuracy by incorporating multiple scales. Sha et al. [14] introduced a graph attention layer that automatically calculates the similarity between neighboring nodes and uses the generated specific weights as the distances between nodes.

Although the above-mentioned deep learning algorithms have achieved good recognition results, analyzing their internal structures theoretically is very challenging due to the complexity of deep learning network architectures and a large number of hyperparameters. In recent years, an effective machine learning method called the broad learning machine (BLS) [15] has been proposed. The structure of BLS is different from that of deep neural networks. Compared with the ''deep'' structure, BLS is more inclined to construct the network in the direction of ''broad''. BLS firstly performs random feature mapping on the original input and enhances the feature mapping to obtain feature nodes and enhancement nodes respectively; Then the feature node and the enhancement node are merged into the input layer and connected to the output layer; Finally, the connection weights between the output layer and the input layer are obtained by ridge regression inverse. Only the connection weights between the input layer and the output layer need to be calculated, which greatly improves the training speed of the model. Recently, BLS has been widely used in HSI classification [16], [17], [18], [19], [20], [21] once it was proposed. Ma et al. [17] proposed a novel Multiscale Random Convolution Broad Learning System (MRC-BLS). This method combines different-sized kernels to extract multiscale spatial features and then demonstrates its effectiveness through weighted fusion. Chu et al. [18] incorporated discriminative information and the manifold structure of samples into BLS, thereby improving its classification performance in hyperspectral images. Wang et al. [19] proposed a domain-adaptive and manifold-regularized output layer in the Local Adaptive Broad Learning System, ensuring the effectiveness of the classification results. Yao et al. [20] applied local sensitivity discriminant analysis and broad learning to HSI classification.

The above HSI classification methods based on BLS have been proven to be useful, however, some obvious issues have been found in feature representation and local geometric structure preservation. First, the mapping from the original BLS input to the mapped feature (MF) is linear, but the HSI has high-dimensional characteristics and a strong correlation between adjacent bands. At the same time, the presence of foreign objects in the same spectrum and in the same spectrum makes the data structure highly nonlinear. BLS linear sparse features are difficult to fully represent the complex spatial-spectral properties of hyperspectral images. Therefore, the classification of HSI using the original BLS cannot achieve satisfactory results. Second, as a supervised classification method, BLS can only conduct supervised learning on limited labeled samples, resulting in insufficient learning. At the same time, BLS does not fully consider the geometric structure between data points and the potential discriminative information of the data in the learning process, which limits the generalization ability of this method in HSI classification.

Recently, GCN [22] have received extensive attention. GCN provides a powerful and intuitive modeling method, which can effectively solve modeling problems in non-Euclidean spaces. Nodes represent data collected by an object at a specific point in time or data collected in a specific form within a specific time period. GCN can deal with complex pairwise interactions and integrate non-Euclidean spatial data, make full use of the intrinsic relationships between objects, and extract invisible relationships between objects. It has been widely used in supervised learning, unsupervised learning, and clustering. On the other hand, Manifold learning assumes that high-dimensional nonlinear data may be embedded in a low-dimensional manifold. By utilizing the structure of this low-dimensional manifold for data processing, more efficient and accurate results can be obtained. Manifold learning is a method for discovering

the intrinsic geometric structure of sample data. In semi-supervised learning, manifold learning can be used to discover the intrinsic geometric structure of data samples in order to improve learning performance. Melacci and Belkin [23] have achieved good results by introducing the Laplacian matrix, which represents the intrinsic geometric structure of the data, as a penalty term into the regularization framework of machine learning.

Inspired by graph convolutional neural networks, Wang et al. [24] combined graph convolution with broad learning to propose the Graph Convolution Broad Network (GCBN). This method utilizes graph convolution operations to extract rich nonlinear spatial-spectral features from hyperspectral images, instead of relying solely on sparse autoencoder mapping in original broad learning. Then, the discriminative spatial-spectral features are expanded using graph convolution and broadened width, further enhancing the representation capability of features and improving the classification ability of BGCN. However, BGCN does not consider the local manifold of data during the learning process, which means it cannot reveal the underlying geometric structure, especially the local structural and discriminative information of the data. To overcome this limitation, Tuya [25] introduced the concept of manifold learning into broad learning and proposed the graph convolution-enhanced discriminative broad learning system (GCDBLS). This method utilizes graph convolution to extract spatial-spectral features from hyperspectral images and then constructs the local intra-class and inter-class structures of data samples using manifold learning. It effectively reveals the local structural and discriminative information of the data and enhances the discriminative ability of the broad learning system towards data features. In the case of GCDBLS, Euclidean distance is used to represent the similarity between sample points. The smaller the Euclidean distance between two points, the greater their similarity; conversely, the larger the distance, the smaller the similarity. Therefore, the constructed neighborhood graph in GCDBLS reflects the Euclidean distance relationship between the two samples. However, for complex hyperspectral image data, the similarity measure based on Euclidean distance cannot reflect the true distance between data points because complex hyperspectral image data satisfy nonlinear structures [26], while the Euclidean distance measure only satisfies globally linear structures, which cannot meet the actual requirements.

Motivated by the discussion above, we propose a local discriminative embedding broad learning system with graph convolutional (GDEBLS). In view of GCN can describe complex data relations, GDEBLS expresses the intrinsic structure of HSI through graphs, so as to capture the context relations between complex ground objects in HSI. On this basis, GCLBLS takes into account the local geometric structure and local discriminative information of the sample space and introduces a local intra-class graph and a local inter-class graph. Unlike GCDBLS, our method utilizes a manifold-based neighborhood selection approach to compute

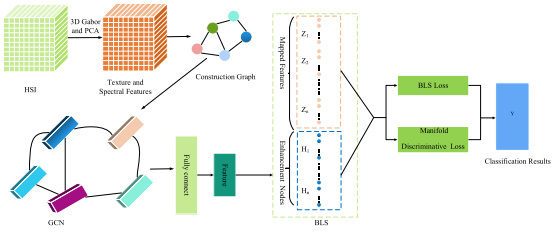

FIGURE 1. Illustration of the GDEBLS structure.

the true distances between samples in the manifold space when constructing the local intra-class graph and inter-class graph. This overcomes the limitations of Euclidean distance measurements. our contribution is three-fold:

(1) A novel hyperspectral image classification method is proposed to learn better feature representation and local geometric structure.

(2) The graph convolutional operation is introduced so that the information between each node is transferred in the graph, and then the adjacency information between each node in the graph is learned to mine the context of HSI.

(3) Introduce local intra-class and inter-class graphs to describe the local geometric structure and local discriminative information of HSI data sample input space.

II. PROPOSED METHOD

The framework of our method is shown in Figure 1. The method mainly consists of three parts: 3D Gabor [27] texture feature extraction, graph convolutional context feature extraction, and local discriminative broad learning system classification. For the original HSI, firstly, texture features of the HSI are extracted using 3D Gabor filters with different directions and frequencies. Next, we use PCA to reduce the dimension of the spatial-spectral feature information after 3D-Gabor filtering; Then, the extracted features are input into the GCN, and the contextual features of the HSI are extracted through the graph convolutional operation; Finally, the extracted context features are input into the discriminative broad learning system for classification.

A. 3D-GABOR

In order to reduce the dependence of the GCN model on training samples and the pressure of feature extraction, we use the Gabor filter to extract spatial information such as edges and textures of images in an unsupervised manner. Gabor filter can effectively extract spatial information including edges and textures, and reduce the difficulty of GCN feature extraction.

The 3D-Gabor filter can directly extract texture features from hyperspectral image data with cubic structure, which is defined as

$$G(x,y,b) = \frac{1}{(2\pi)^{3/2}\sigma_x\sigma_y\sigma_b} \times \exp\left(-\frac{1}{2}\left(\left(\frac{x'}{\sigma_x}\right)+\left(\frac{y'}{\sigma_y}\right)+\left(\frac{b'}{\sigma_b}\right)\right)\right) \times \exp\left(i2\pi\left(xf_x+yf_y+bf_b\right)+\varphi\right) \quad (1)$$

$$f_x = f \sin\phi \cos\theta \tag{2}$$
$$f_y = f \sin\phi \cos\theta \tag{3}$$
$$f_b = f \cos\phi \tag{4}$$

where $G(x, y, b)$ is a 3D-Gabor filter; $(\sigma_x, \sigma_y, \sigma_b)$ is the width of the spatial spectral; θ and φ are filter directions; $[x, y, b]$ is the coordinate of a point $[x', y', b']$ of the original image data after rotation by the angles θ and φ; f is the frequency; (f_x, f_y, f_b) is the component of the frequency in the spatial spectral dimension. θ and φ take values in the set $\left[0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}\right]$, but when $\varphi = 0$ and θ take any value in the set, the center frequency is equal. There are 13 directions in total, representing 13 filters.

Assuming the hyperspectral cube data is $HSI(x, y, b)$, the HSI original cube data is input into the 3D-Gabor Gabor filter, and the real part is extracted after convolution processing. Then a new 3D structural data with texture and spectral features is constructed, namely

$$HSI_{texture}(x, y, b) = HSI(x, y, b) \otimes G(x, y, b) \tag{5}$$

B. GRAPH CONVOLUTIONAL FEATURE EXTRACTION

Since BLS linear sparse features are difficult to fully represent the complex spatial spectral characteristics of hyperspectral images, in this paper, we use GCN to extract contextual features of HSI. GCN is a feature extractor that can directly operate on graph structure data. It can make full use of the dependency between nodes and the feature information of each neighbor node through graph convolutional operation. As a generalization of CNN, GCN can directly convolution graphs. Given a graph $G = (V, E, A)$, where V is the set of nodes in G (the number of nodes is n); E is the set of edges in G; A is adjacency matrix. GCN obtains the representation of hidden layer through the following formula [10]

$$H^{(l+1)} = f\left(H^l\right) = \tau\left(\hat{A}H^l W^l\right); l = 1, \cdots, L \tag{6}$$

where, H^l is the feature matrix of layer l, $H^l \in R^{n \times d_l}$, d_l are the dimensions of each node of layer l; L is the number of GCN layers; W^l is the trainable weight, $W \in R^{d_l \times d_{l+1}}$; $\hat{A}$ is the normalized symmetric adjacency matrix, $\hat{A} = \tilde{D}^{-1/2}\tilde{A}\tilde{D}^{-1/2}$, $\tilde{A}$ is the adjacency matrix with loop, $\tilde{A} \in R^{n \times n}$, $\tilde{A} = A + I$. $\tilde{D}$ is the degree matrix of $\tilde{A}$, $\tilde{D} \in R^{n \times n}$, τ is the nonlinear activation function.

In GCN, a global pooling layer is introduced to aggregate all node features, and the context features are further compressed through the full connection layer. Finally, the context features of hyperspectral images are extracted through GCN.

C. LOCATION DISCRIMINATIVE EMBEDDING BROAD LEARNING SYSTEM CLASSIFICATION

We input the context features of hyperspectral images extracted by 3D-Gabor and graph convolution into the broad learning system for classification. However, we noticed that BLS only pays attention to the separability of various samples when making classification decisions, ignoring the geometric structure between data samples and the potential discriminative information of the data. To some extent, it limits the further improvement of classification performance and generalization ability of BLS in hyperspectral image classification. Cover and Hart [28] emphasized that half of the class information of the sample is hidden in its neighborhood. The manifold learning method proposed in recent years [29] can effectively reveal the local structural information and local discriminative information contained in the sample points. Therefore, in this paper, we introduce the concepts of local intra-class graph and a local inter-class graph, and propose a local discriminative embedding broad learning system (LDEBLS), Unlike GCDBLS, our method utilizes a manifold-based neighborhood selection approach to compute the true distances between samples in the manifold space when constructing the local intra-class graph and inter-class graph.

Suppose $N_k(x_i)$ is the k nearest neighbor set of sample point x_i, G represents the weighted adjacency graph of hyperspectral image feature X_{GCN} extracted by GCN, where the i-th vertex represents the sample point G, which is called a node.

If $x_i \in N_k(x_j)$ or $x_j \in N_k(x_i)$, the vertices in G are connected, and its connection weight value can be defined as W_{ij}.The following thermal kernel function is commonly used to define the weight

$$W_{ij} = \exp\left(-\left\|x_i - x_j\right\|^2 \Big/ t\right) \tag{7}$$

where is $t > 0$ thermonuclear parameter. Then define the corresponding weight matrix of graph G as

$$A_{ij} = \begin{cases} W_{ij}, & x_j \in N_k(x_i) \ \ or \ x_i \in N_k(x_j) \\ 0, & else \end{cases} \tag{8}$$

In formulas (7) and (8), we need to calculate the distances between data samples to determine their neighborhoods. Typically, for complex hyperspectral image data, the neighborhood selection similarity measure relies on metrics based on Euclidean distance. However, the Euclidean distance metric fails to reflect the true distances between data points because complex or high-dimensional data often exhibit nonlinear structures, while the Euclidean distance measure only captures global linear structures and cannot meet practical requirements. To address this issue, we adopt a manifold-based neighborhood selection method to compute the true distances between samples in the manifold space, overcoming the limitations of Euclidean distance measurements. The distance between two points, x_i and x_j, on the manifold in the manifold space is shown in equation (9):

$$L(x_i, x_j) = \exp\left(D(x_i, x_j) \big/ \sigma_{ij}\right) - 1 \tag{9}$$

where $D(x_i, x_j) = d(x_i, x_j) \big/ \max(x_i, x_j)$ represents the normalized Euclidean distance, $d(x_i, x_j)$ represents the Euclidean distance between them, and σ_{ij} represents the adjustable parameter. According to the formula for calculating manifold distance, the proximity distance is calculated. The data points

are viewed as vertices of the graph G (V, E), where V is the set of vertices and E is the set of edges. Let R_{ij} represent the set of all paths connecting data points x_i and x_j in the graph, then the manifold distance between x_i and x_j is given by equation (10):

$$MF\left(x_i, x_j\right) = \min_{q \in Q_{ij}} \sum_{k=1}^{|q|-1} L\left(Q_k, Q_{k+1}\right) \tag{10}$$

where q represents a path connecting x_i and x_j, and $|q|$ represents the number of nodes on path q. From this equation, it can be observed that the manifold distance between two points in space is determined by finding the minimum value among the sum of the manifold distances of all the line segments connecting the nodes between the two points. This minimum value represents the manifold distance between the two points.

According to spectral graph theory, modeling the internal geometric structure of input space with a nearest neighbor graph G can effectively describe the local geometric structure of samples, but only one global graph is not enough to reflect the discriminant structure between samples.

Therefore, the weighted adjacency graph G can be further divided into two complementary weighted adjacency graphs: intra class graph G^w and inter class graph G^b, and satisfy that, $G = G^w \cup G^b, G^w \cap G^b = \phi$ is used to reflect the local intra-class adjacency relationship and the local inter-class adjacency relationship respectively. Defined as

$$A_{ij}^w = \begin{cases} W_{ij}, & x_j \in N_k^w\left(x_i\right) \ or \ x_i \in N_k^w\left(x_j\right) \\ 0, & else \end{cases} \tag{11}$$

$$A_{ij}^b = \begin{cases} W_{ij}, & x_j \in N_k^b\left(x_i\right) \ or \ x_i \in N_k^b\left(x_j\right) \\ 0, & else \end{cases} \tag{12}$$

By using equations (9) and (10), we determine the neighborhoods of $N_k^w\left(x_i\right)$ and $N_k^b\left(x_i\right)$. $N_k^w\left(x_i\right) \in G^w$ is the homogeneous k-nearest neighbor data set of sample point x_i, and $N_k^b\left(x_i\right) \in G^b$ is the heterogeneous k-nearest neighbor data set of sample point x_i.

In order to better describe LDEBLS, the following definitions are given. **Definition 1. scatter matrix of local inter-class graph:**

Let L^w and L^b be Laplace matrices of intra-class graphs G^w and inter-class graphs G^b, respectively, then matrix $H^w = XL^wX^T = X\left(T^w - A^w\right)X^T$ is called local intra-class graph scatter matrix, and matrix $H^b = XL^bX^T = X\left(T^b - A^b\right)X^T$ is called local inter class graph scatter matrix. Where $A_{ij}^{(\cdot)}$ is the weight matrix of graph $G^{(\cdot)}$; $T^{(\cdot)}$ is a diagonal matrix, and its diagonal elements define $t_{ii}^{(\cdot)} = \sum_{j=1}^{l} A_{ij}^{(\cdot)}$, that is, $T_{ii}^{(\cdot)} = \sum_{j=1}^{l} A_{ij}^{(\cdot)}$. H^w and H^b are called local dispersion matrices. In definition 1, the local intra-class graph scatter matrix describes the local structure information of the input sample manifold, and the local inter-class graph scatter matrix describes the local discrimination information of the input sample manifold.

Define 2. local reserved information difference matrix: We call matrix $M = \varsigma H^w\text{-}(1\text{-}\varsigma)H^b$ local information difference matrix. ς is the local information balance parameter. The parameter ς is used to find a satisfactory balance between the local intra-class compactness and the local inter class discreteness of the input sample manifold.

Next, we introduce the basic principles of BLS. The method flow of BLS is described as follows: First, the original data X is mapped to the Mapped Feature (MF) by random weight W^M. The i-th group MF can be expressed as

$$Z_i = \phi\left(XW_i^M + \beta_i^M\right), i = 1, \cdots, d^M \tag{13}$$

where, $Z_i \in R^{n \times G^M}$ and β_i^M are biases, d^M is the number of groups of feature nodes, G^M is the dimension of each group of feature nodes, and $\phi(\cdot)$ is the activation function. Splicing d^M group of feature nodes into $Z = \left[Z_1, Z_2, \cdots, Z_{d^M}\right]$. Then, MF is mapped to the enhancement node (EN) through the random weight W^E to realize the width expansion of the feature

$$H_j = \varphi\left(ZW_j^E + \beta_j^E\right), i = 1, \cdots, d^E \tag{14}$$

where, $H_j \in R^{n \times G^E}$ and β_j^E are biases, d^E is the number of groups of enhancement nodes, G^E is the dimension of each group of enhancement nodes, and $\varphi(\cdot)$ is the activation function.

Splicing d^E group of enhanced nodes into $H = \left[H_1, H_2, \cdots, H_{d^E}\right]$. Let $A = [Z|H]$, the output of BLS is $\hat{Y} = AW^o$, where W^o is the weight from the hidden layer to the output layer. Since W_i^M, β_i^M, W_j^E, and β_j^E are randomly generated and remain unchanged during the training process, the weights that the network needs to learn are only W^o.Therefore, the objective function of BLS is

$$\min_W \left(\|Y - AW\|_2^2 + \frac{\lambda}{2}\left\|W^o\right\|_2^2\right) \tag{15}$$

There are two terms in formula (13), the first is the empirical risk term, and Y is the given supervision information. The function of the first item is to reduce the difference between the output of the BLS and the supervision information Y, where Y represents the label of X.The second term is the structural risk term, which is used to improve the generalization ability of BLS and reduce the risk of overfitting. λ is the regularization parameter.

In the case of keeping the BLS optimization framework unchanged, LDEBLS adds the maximization of the overall inter-class interval of data and the local manifold distribution of data into the BLS objective function. The LDEBLS objective function can be expressed as:

$$\begin{aligned} F &= \|Y - AW\|_2^2 + C_1\|W\|_2^2 + C_2(Tr(M) \\ &= \|Y - AW\|_2^2 + C_1\|W\|_2^2 + C_2(Tr\left(H^w - H^b\right) \\ &= \|Y - AW\|_2^2 + C_1\|W\|_2^2 \\ &\quad + C_2(Tr\left(W^TA\left(L^w - L^b\right)A^TW\right) \end{aligned} \tag{16}$$

the third term $Tr(M)$ is the manifold regularity discriminative term, $Tr(\cdot)$ represents the trace of the matrix; C_1 and C_2 are

TABLE 1. Indian Pines dataset description.

Class NO.	Name	Numbers of samples	Training	Testing
1	Alfalfa	46	5	41
2	Corn-notill	1428	5	1423
3	Corn-mintill	830	5	825
4	Corn	237	5	232
5	Grass-pasture	483	5	478
6	Grass-trees	730	5	725
7	Grass-pasture-mowed	28	5	23
8	Hay-windrowed	478	5	473
9	Oats	20	5	15
10	Soybean-notill	972	5	967
11	Soybean-mintill	2455	5	2450
12	Soybean-clean	593	5	588
13	Wheat	205	5	200
14	Woods	1265	5	1260
15	Buildings-Grass-Trees-Drives	386	5	381
16	Stone-Steel-Towers	93	5	88

TABLE 2. Kennedy Space Center dataset description.

Class NO.	Name	Numbers of samples	Training	Testing
1	Scrub	761	5	756
2	Willow	243	5	238
3	CP hammock	256	5	251
4	CP/Oak	252	5	247
5	Slash pine	161	5	156
6	Oak/Broadleaf	229	5	224
7	Hardwood swamp	105	5	100
8	Graminoid marsh	431	5	428
9	Spartina marsh	520	5	515
10	Cattail marsh	404	5	399
11	Salt marsh	419	5	414
12	Mud flats	503	5	498
13	Water	927	5	922

the regularization coefficients. Solving Equation (14) to get

$$W^o = \left[A^T A + C_1 A \left(L^w - L^b\right) A^T + C_2 I\right]^{-1} A^T Y \quad (17)$$

where I is the identity matrix, A^T is the transpose matrix of A.

III. EXRERIMENTS

In this section, we design comparative experiments and ablation experiments. The overall accuracy (OA), accuracy of each category (AA) and kappa coefficient were used as evaluation metrics. The experiments are implemented using TensorFlow and Keres deep learning frameworks.

A. DATASETS

The experiments use three hyperspectral datasets[1] of Indian Pines (IP), Kennedy Space Center (KSC) and Botswana (BS) to verify the effectiveness of the proposed algorithm. IP dataset, which contains 220 bands, there are 16 categories of main features, and the image size is 145 × 145 pixels, with a spatial resolution of 20m, removing 20 bands with serious noise to obtain data including 200 spectral bands for analysis; KSC dataset, which contains 224 bands, including 13 categories of main features, with an image size of 512 × 614 pixels, after removing the water absorption and low SNR spectral bands, the data including 176 spectral bands are obtained for research and analysis; BS dataset, which contains 14 object categories, the image size is 1476 × 256, removes uncalibrated, water-absorbing areas and noise bands, leaving 145 bands. For the three data sets, we select 5 samples in each class as the training set and the remaining samples as the test set. Table 1-3 provide descriptions of three datasets.

[1]The four HSI benchmarks are available from: http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes

B. EXRERIMENTS SETUPS

The following 8 classification methods are selected for comparison. A method based on CNN: (2D-CNN) [30]; Four GCN based methods: GCN, spatial spectral convolution

TABLE 3. Botswana dataset description.

Class NO.	Name	Numbers of samples	Training	Testing
1	Water	270	5	265
2	Hippo grass	101	5	96
3	Floodplain grass1	251	5	246
4	Floodplain grass2	215	5	210
5	Reeds1	269	5	264
6	Riparian	269	5	264
7	Firescar2	259	5	254
8	Island interior	203	5	198
9	Acacia woodlands	314	5	309
10	Acacia shrublands	248	5	243
11	Acacia grasslands	305	5	300
12	Short mopane	181	5	176
13	Mixed mopane	268	5	263
14	Exposed soils	95	5	90

TABLE 4. Accuracy comparisons for the Indian Pines (%).

Class	SVM	SVMCK	2D-CNN	BLS	SBLS	GCN	SSGCN	MDGCN	GCBN	GCDBLS	GDEBLS
1	95.12	63.41	82.93	96.15	**100.00**	**100.00**	**100.00**	91.30	97.56	**100.00**	**100.00**
2	43.85	32.68	2.62	46.38	36.45	46.52	55.38	72.13	**90.16**	84.37	88.38
3	41.33	49.70	47.76	59.38	64.36	54.91	45.21	73.49	66.06	85.04	**85.77**
4	45.26	98.71	25.43	77.88	**100.00**	97.84	72.41	88.61	**100.00**	99.13	**100.00**
5	84.73	73.85	37.24	85.10	74.90	68.62	69.24	**90.89**	89.96	80.00	81.47
6	53.38	62.62	50.62	81.13	78.76	99.31	50.96	78.63	74.76	77.29	**93.91**
7	91.30	**100.00**	43.48	92.86	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
8	88.58	91.75	56.45	88.43	95.14	39.75	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
9	66.67	80.00	80.02	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
10	31.95	54.19	20.89	71.74	83.66	47.57	70.22	83.64	84.59	**86.41**	76.45
11	46.16	50.00	40.04	53.06	69.27	65.35	76.73	74.54	**76.98**	66.90	69.39
12	22.28	50.68	23.30	61.78	54.42	25.17	88.61	63.41	90.31	**95.38**	87.69
13	76.50	98.50	97.01	99.46	**100.00**	95.50	**100.00**	95.61	**100.00**	**100.00**	99.49
14	51.35	88.17	84.76	84.50	83.33	75.79	84.52	94.23	82.62	**100.00**	**100.00**
15	28.87	50.13	12.34	55.46	84.78	97.11	80.11	83.68	91.86	**98.94**	98.15
16	90.91	**100.00**	90.91	95.89	98.86	**100.00**	**100.00**	**100.00**	**100.00**	98.82	90.59
OA	48.30	59.04	39.99	66.00	71.07	63.62	75.94	80.53	83.56	84.40	**85.40**
Kappa	42.06	54.86	32.50	61.75	68.07	59.10	73.09	78.14	81.52	82.45	**83.49**

neural network (SSGCN) [11], multiscale dynamic GCN (MDGCN) [13], graph convolution broad network (GCBN) [24], GCDBLS [25]; Two BLS based methods: BLS and semi-supervised broad learning system (SBLS) [7]; A traditional machine learning method: support vector machine (SVM) and support vector machine [1] based on composite kernels (SVMCK) [31]. For SVM, BLS, and SBLS, we use grid search to select the optimal parameters. For SSGCN, MDGCN, GCBN and GCDBLS, we directly refer to literature [5], [14], [15], [16], [17] to select the hyperparameters. The parameters of the proposed method are set as: In order to extract various types of texture features, 3D Gabor filter banks with multiple directions and multiple center frequencies are used. The filter bank includes 13 directions, and the center frequency is set {1/2, 1/4}; In the training process of GCN network, the learning rate is set to 0.001, the batch size is set to 32, the number of iterations of the network is 400, the classification accuracy of the model on the test sample set is recorded every 50 times, and the learning rate is dynamically attenuated with a cycle of 50 times. The training loss of the model is optimized by the Adam function. In the LDEBLS network, the number of feature nodes is set to 100, the number of feature node groups is set to 10 groups, and the number of enhanced nodes is set to 1000; The regularization parameters C_1 and C_2 included in GDEBLS are set through 10-fold cross-validation, the range of parameters C_1 and C_2 is $\{2^{-5}, \ldots, 2^5\}$. In order to eliminate the influence of experimental randomness, all experimental results are the mean of 10 results. Tables 4-6 present the experimental results of different algorithms on the three datasets, and Figures 1-3 display the classification result graphs of different algorithms.

TABLE 5. Accuracy comparisons for the KSC (%).

Class	SVM	SVMCK	2D-CNN	BLS	SBLS	GCN	SSGCN	MDGCN	GCBN	GCDBLS	GDEBLS
1	71.16	69.31	20.26	94.71	95.46	85.71	94.72	**100.00**	85.58	96.81	96.81
2	71.85	76.05	36.78	80.67	67.10	78.57	92.08	**100.00**	**100.00**	90.64	93.19
3	63.35	86.06	26.67	96.02	90.16	41.43	75.89	91.24	**100.00**	**100.00**	**100.00**
4	70.85	54.66	32.27	29.55	40.00	38.87	65.46	67.21	**100.00**	96.72	98.77
5	00.64	50.64	46.25	57.05	47.65	61.54	78.48	**100.00**	**100.00**	98.04	98.04
6	58.18	45.09	4.82	44.20	32.26	45.54	98.67	97.77	**100.00**	**100.00**	**100.00**
7	99.00	99.00	25.00	94.00	97.85	93.00	90.20	**100.00**	**100.00**	**100.00**	**100.00**
8	71.13	67.37	35.58	83.10	82.34	50.70	97.19	57.28	91.78	**100.00**	**100.00**
9	88.16	86.60	63.58	89.32	93.11	95.92	99.49	99.42	83.50	**100.00**	**100.00**
10	70.93	81.95	9.43	73.68	95.66	77.19	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
11	88.16	86.96	62.20	87.68	94.59	97.58	**100.00**	97.10	95.89	**100.00**	**100.00**
12	61.85	85.54	53.19	72.99	83.10	88.55	35.80	**100.00**	95.38	62.22	72.73
13	97.94	97.83	86.18	98.26	99.02	**100.00**	**100.00**	98.48	**100.00**	**100.00**	**100.00**
OA	75.59	79.34	45.19	82.41	85.26	79.89	88.61	93.80	94.77	95.22	**96.46**
Kappa	72.90	77.08	39.78	80.41	83.54	77.62	87.35	93.09	94.21	94.68	**96.05**

TABLE 6. Accuracy comparisons for the Botswana(%).

Class	SVM	SVMCK	2D-CNN	BLS	SBLS	GCN	SSGCN	MDGCN	GCBN	GCDBLS	GDEBLS
1	98.14	**100.00**	95.17	95.09	38.87	**100.00**	66.79	47.92	65.28	**100.00**	82.13
2	25.00	**100.00**	17.02	**100.00**	97.92	98.96	96.88	**100.00**	**100.00**	**100.00**	**100.00**
3	36.40	89.02	28.40	96.75	**100.00**	88.62	80.89	80.08	90.24	**100.00**	**100.00**
4	76.64	**100.00**	87.38	97.62	**100.00**	84.76	80.04	91.90	93.33	91.79	97.18
5	39.93	79.17	57.46	80.68	77.27	71.97	94.31	98.86	96.21	86.97	**100.00**
6	29.10	82.58	51.87	65.15	68.94	65.53	83.33	73.48	86.36	99.23	**100.00**
7	91.86	89.76	79.84	98.65	98.43	94.09	84.25	**100.00**	**100.00**	**100.00**	**100.00**
8	76.73	98.48	51.49	86.87	97.47	95.45	96.97	89.39	89.40	**100.00**	**100.00**
9	58.15	98.71	38.66	66.02	99.68	04.21	98.71	99.68	91.59	**100.00**	**100.00**
10	76.11	**100.00**	40.49	86.42	65.02	20.58	95.47	**100.00**	**100.00**	**100.00**	**100.00**
11	38.49	92.33	52.30	91.00	**100.00**	98.67	89.67	89.00	**100.00**	**100.00**	**100.00**
12	93.89	88.64	94.44	86.93	70.45	99.43	98.86	89.77	85.80	**100.00**	**100.00**
13	78.65	**100.00**	7.12	89.73	98.10	42.21	58.94	77.95	79.47	**100.00**	**100.00**
14	43.62	**100.00**	97.87	94.44	72.22	96.67	71.11	95.56	86.67	81.61	**100.00**
OA	62.71	93.58	55.50	86.69	84.80	71.71	85.31	87.04	90.12	97.80	**97.99**
Kappa	59.61	93.05	52.18	85.58	83.52	69.42	84.75	85.95	89.30	97.61	**97.82**

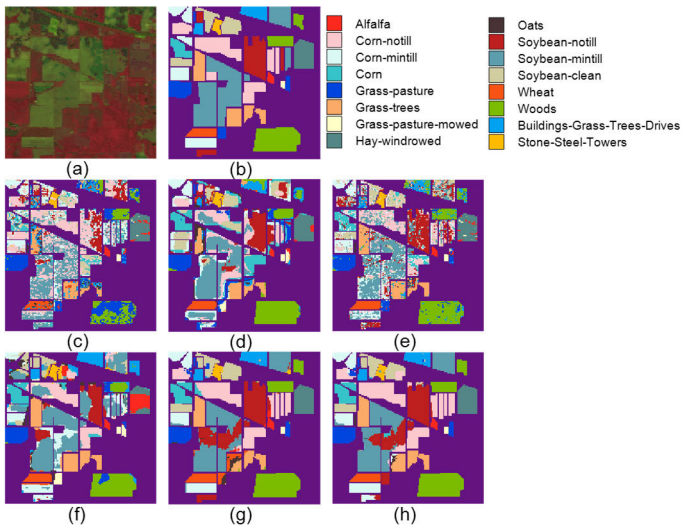

FIGURE 2. Classification results of different methods on Indian Pines. (a) False color image,(b) Ground-truth map (c) SVM, (d) SVMCK, (e) BLS, (f) GCN, (g) GCDBLS, (h) GDEBLS.

It can be seen from table 4 that our method on the Indian pines is superior. Among the 16 categories, 8 categories have the highest classification accuracy, and the OA and kappa coefficients have reached the highest, which are 85.40% and 83.49% respectively. From table 4, the classification effects of SVM, 2D-CNN, BLS and GCN are not ideal,

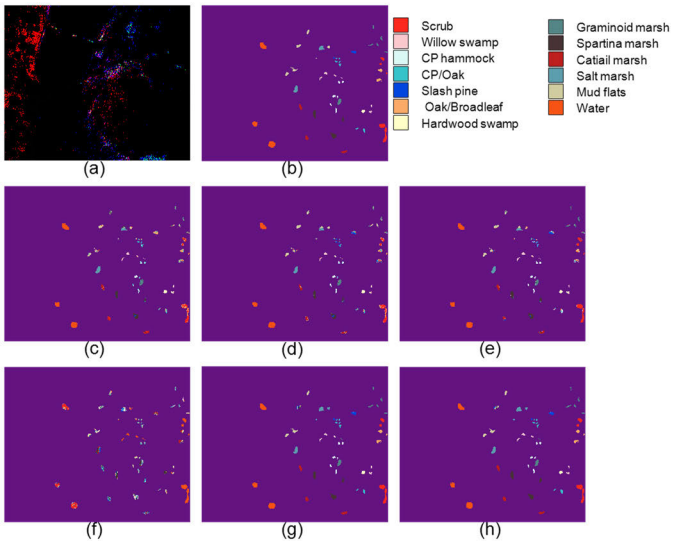

FIGURE 3. Classification results of different methods on KSC. (a) False color image,(b) Ground-truth map (c) SVM, (d) SVMCK, (e) BLS, (f) GCN, (g) GCDBLS, (h) GDEBLS.

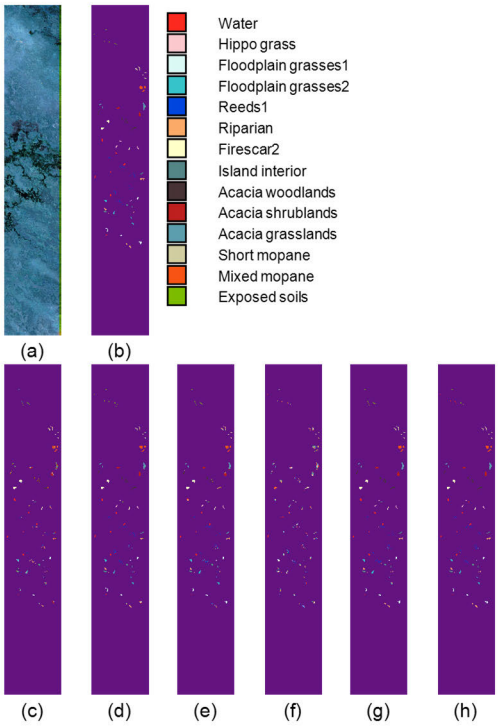

FIGURE 4. Classification results of different methods on Botswana. (a) False color image,(b) Ground-truth map (c) SVM, (d) SVMCK, (e) BLS, (f) GCN, (g) GCDBLS, (h) GDEBLS.

and the OA and kappa are lower than those of other methods. The reason is that SVM and BLS classifiers only consider the spectral information of the pixel and do not use the spatial information. Although the 2DCNN method utilizes the spatial information of HSI, it fails to combine the spectral information for classification. Overall, although the Indian Pines hyperspectral dataset has fewer training samples and complex land cover classes, our method effectively captures both spectral and spatial information of HSI by extracting 3D-Gabor texture features, including direction and scale information. We then use graph convolution to extract deep-level features of hyperspectral images and introduce the local geometric structure information into BLS. This approach achieves high classification accuracy. Visualizing the classification results of the IP dataset with 5 training samples per class and comparing them with the ground truth, our method shows the closest resemblance to the true labels, as shown in Figure 2.

From table 5, the OA of the GDEBLS method is 96.46%, and the Kappa coefficient is 96.05%, which are better than other methods. Compared to SVM and SVMCK, BLS has higher classification accuracy, the reason is that the BLS structure has a good feature broad expansion ability. Compared with BLS, SBLS achieves higher OA and kappa, and is superior to BLS, this is because that SBLS makes full use of the unlabeled sample information. GCN, SSGCN and MDGCN are all graph convolutional network methods, and the classification accuracy of GCN is lower than that of SSGCN, MDGCN and GCBN, the reason is that GCN only considers spectral information. SSGCN and MDGCN not only utilize the spectral information of hyperspectral images but also use spatial information. MDGCN makes full use of the spatial spectral correlation on multiple scales by establishing multiple input maps at different neighborhood scales. when visualizing the classification results of the KSC dataset with 5 training samples per class and comparing them with the ground truth, our method again exhibits the closest resemblance to the true labels, as shown in Figure 3.

It can be seen from table 6 that there are 14 categories in Botswana dataset, and the classification performance of the proposed method on Botswana dataset is superior. Among the 14 categories, 12 categories have the highest

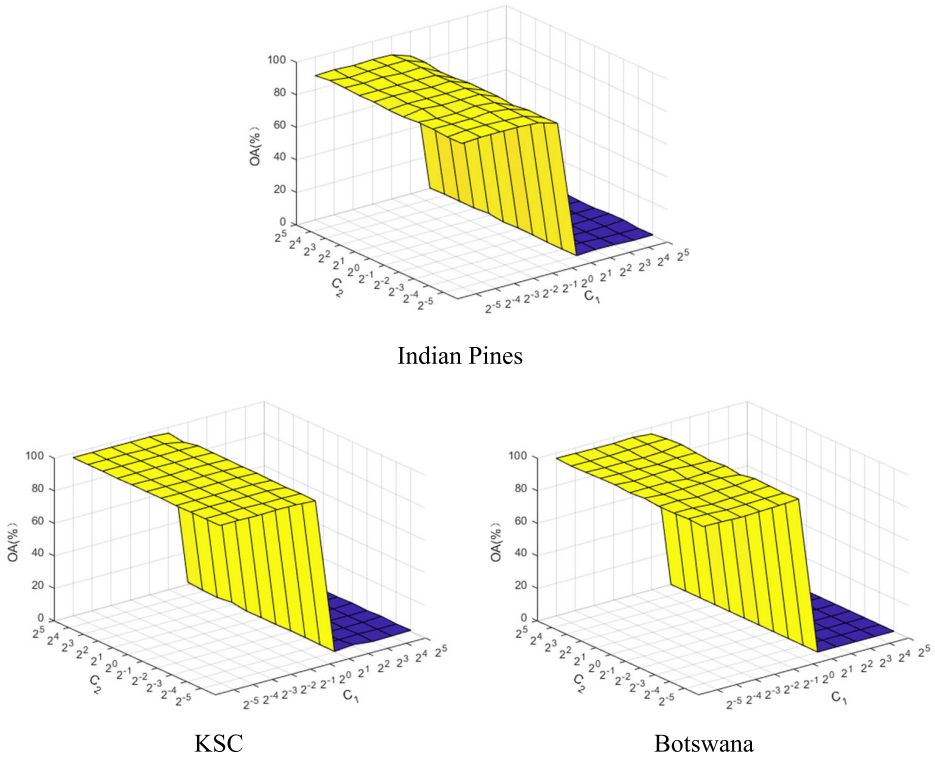

FIGURE 5. Influence of C_1 and C_2 on OA.

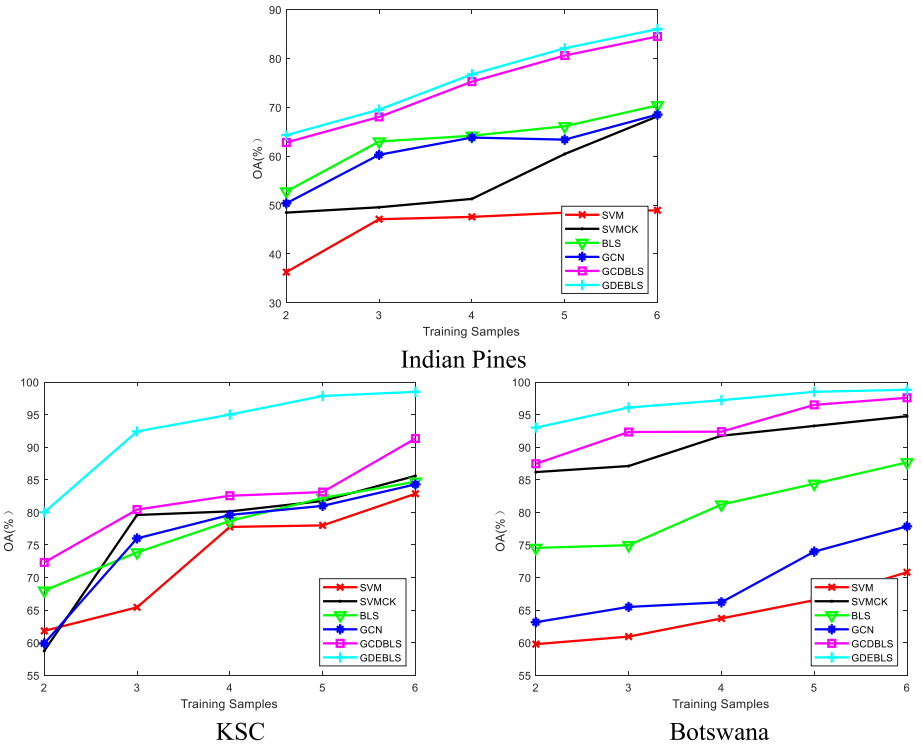

FIGURE 6. Influence of different labeled sample sizes on OA.

classification accuracy, and OA and kappa coefficients have reached the highest, respectively 97.99% and 97.82%. From table 6, the results of GCBN have achieved good results. This is because GCBN also combines the advantages of GCN and BLS. The difference between our method and GCBN is that, our method first uses 3D Gabor to extract the

TABLE 7. Ablation experiment results (%).

	Indian Pines				KSC				Botswana			
	BLS	GBLS	GCBLS	GDEBLS	BLS	GBLS	GCBLS	GDEBLS	BLS	GBLS	GCBLS	GDEBLS
OA	66.00	81.63	83.91	85.40	82.41	94.15	95.13	96.46	86.69	93.68	95.85	97.99
Kappa	61.75	79.26	81.76	83.49	80.41	93.51	94.58	96.05	85.58	93.14	95.51	97.82

local texture features of the image, and uses the K-nearest neighbor method to compose the image to establish the global correlation between pixels, where each node contains rich spatial and spectral information, However, the BGCN method does not consider the local manifold of the data during the learning process, thus failing to reveal the underlying geometric structure and local discriminative information. From table 6, we can observe that the experimental results of GCDBLS are second only to GDEBLS. GCDBLS uses graph convolution to extract the spatial features of hyperspectral images and employs manifold learning to construct the local intra-class and inter-class structures of data samples, effectively revealing the local structural and discriminative information of the data and enhancing the discriminative ability of the wide learning system for data features. However, GCDBLS uses Euclidean distance to represent the similarity between sample points. The smaller the Euclidean distance between two points, the greater their similarity, and vice versa. Therefore, the constructed neighborhood graph of GCDBLS reflects the Euclidean distance relationship between two samples. However, for complex hyperspectral image data, using a similarity measure based on Euclidean distance for neighborhood selection cannot reflect the true distances between data points. This is because complex hyperspectral image data exhibit nonlinear structures, while the Euclidean distance measure only captures global linear structures and cannot meet practical requirements. Unlike GCDBLS, our method overcomes the limitations of Euclidean distance measurements by using a manifold-based neighborhood selection approach to compute the true distances between samples in the manifold space. From Table 6, we can see that our algorithm's experimental results are superior to GCDBLS, thus validating the effectiveness of our approach. Visualizing the classification results of the Botswana dataset with 5 training samples per class and comparing them with the ground truth, our method again exhibits the closest resemblance to the true labels, as shown in Figure 4.

C. ABLATION EXPERIMENTS

GDEBLS first utilizes 3D-Gabor filters to extract directional and size information of texture features from hyperspectral images, enabling the capture of both spectral and spatial information in the HSI. Then, graph convolution is employed to extract deep-level features from the hyperspectral images. Finally, the local geometric structure information of the hyperspectral images is incorporated into the Backpropagation Least Squares (BLS) for classification. To validate the effectiveness of different modules, we refer to the method of using BLS for classification after feature extraction with 3D-Gabor filters as GBLS. Based on the GBLS algorithm, we introduce graph convolutional operations, resulting in GCBLS. On top of GCBLS, we further introduce local geometric structure information, which we refer to as GDEBLS. By comparing the experimental results, it can be observed from Table 7 that GBLS outperforms BLS on the three datasets, thereby validating the effectiveness of the 3D-Gabor module. Furthermore, we can observe that GCBLS yields better results than GBLS, validating the effectiveness of the graph convolution module. GDEBLS demonstrates superior performance compared to GCBLS, thus validating the effectiveness of the local geometric structure module. Through the aforementioned experiments, the effectiveness of different modules in the proposed algorithm has been validated.

D. PARAMETER ANALYSIS

The impact of the regularization parameters C_1 and C_2 on overall accuracy (OA) is observed on the three hyperspectral datasets. In the experiment, different values are chosen for C_1 and C_2, while keeping the parameter settings consistent with table 4. The range of values for C_1 and C_2 is $\{2^{-5}, \ldots, 2^5\}$. Figure 5 presents the effects of different values of C_1 and C_2 on OA. It can be observed that when C1 is less than 2^1, OA increases with the variation of C_2. However, when C_1 exceeds 2^1, the variation of C_2 leads to poorer classification performance in terms of OA. Therefore, based on the analysis of C_1 and C_2 parameters, the optimal values of C_1 and C_2 can be determined to achieve the best classification results.

E. INFLUENCE OF DIFFERENT LABELED SAMPLES SIZES ON OA

In this section, the influence of varying the number of training samples on the performance of different methods is analyzed. For the Indian Pines, KSC, and Botswana datasets, L = {2, 3, 4, 5, 6} samples are randomly selected from each class as training samples. Figure 6 illustrates the trends in overall classification accuracy for different methods under different training sets. It can be observed that increasing the number of training samples has a positive impact on the classification accuracy of all methods. Particularly, the proposed GDEBLS method achieves higher classification accuracy than the five comparative methods, even with a limited number of training samples.

IV. CONCLUSION

In this paper, we propose a local discriminative embedding broad learning system with graph convolutional for hyperspectral image classification. This method aims to solve the problem that BLS does not make full use of the spatial

information of the image in hyperspectral classification and to make full use of the local geometric structure and local discriminative information of the sample space. GDEBLS first uses 3D-Gabor filters to extract directional and size information of texture features from hyperspectral images, allowing for the simultaneous capture of both spectral and spatial information in the HSI. Then, graph convolutional operations are employed to obtain rich nonlinear spatial-spectral features from the hyperspectral image. Finally, based on the idea of manifold learning, intra-class and inter-class graphs are constructed to incorporate these features into BLS. It is worth mentioning that our method utilizes a neighborhood selection approach based on manifold structure to calculate the true distances between samples in the manifold space, overcoming the limitations of Euclidean distance measurement. Experimental results on three widely used hyperspectral datasets show the superiority of our method in HSI classification.

REFERENCES

[1] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 8, pp. 1778–1790, Aug. 2004.
[2] 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.
[3] X. Sun, Q. Qu, N. M. Nasrabadi, and T. D. Tran, "Structured priors for sparse-representation-based hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 7, pp. 1235–1239, Jul. 2014.
[4] W. Li, G. Wu, F. Zhang, and Q. Du, "Hyperspectral image classification using deep pixel-pair features," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 2, pp. 844–853, Feb. 2017.
[5] J. M. Haut, M. E. Paoletti, J. Plaza, A. Plaza, and J. Li, "Visual attention-driven hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 10, pp. 8065–8080, Oct. 2019.
[6] M. He, B. Li, and H. Chen, "Multi-scale 3D deep convolutional neural network for hyperspectral image classification," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2017, pp. 3904–3908.
[7] Y. Li, H. Zhang, and Q. Shen, "Spectral–spatial classification of hyperspectral imagery with 3D convolutional neural network," *Remote Sens.*, vol. 9, no. 1, p. 67, Jan. 2017.
[8] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
[9] Z. Zhong, J. Li, Z. Luo, and M. Chapman, "Spectral–spatial residual network for hyperspectral image classification: A 3-D deep learning framework," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 2, pp. 847–858, Feb. 2018.
[10] W. Song, S. Li, L. Fang, and T. Lu, "Hyperspectral image classification with deep feature fusion network," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 6, pp. 3173–3184, Jun. 2018.
[11] 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. 2019.
[12] Y. Ding, Y. Guo, Y. Chong, S. Pan, and J. Feng, "Global consistent graph convolutional network for hyperspectral image classification," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–16, 2021.
[13] 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.
[14] A. Sha, B. Wang, X. Wu, and L. Zhang, "Semisupervised classification for hyperspectral images using graph attention networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 18, no. 1, pp. 157–161, Jan. 2021.
[15] C. L. P. Chen and Z. Liu, "Broad learning system: An effective and efficient incremental learning system without the need for deep architecture," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 1, pp. 10–24, Jan. 2018.
[16] Y. Kong, X. Wang, Y. Cheng, and C. Chen, "Hyperspectral imagery classification based on semi-supervised broad learning system," *Remote Sens.*, vol. 10, no. 5, p. 685, Apr. 2018.
[17] Y. Ma, Z. Liu, and C. L. P. Chen, "Multiscale random convolution broad learning system for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.
[18] Y. Chu, H. Lin, L. Yang, D. Zhang, Y. Diao, X. Fan, and C. Shen, "Hyperspectral image classification based on discriminative locality preserving broad learning system," *Knowl.-Based Syst.*, vol. 206, Oct. 2020, Art. no. 106319.
[19] H. Wang, X. Wang, C. L. P. Chen, and Y. Cheng, "Hyperspectral image classification based on domain adaptation broad learning," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 3006–3018, 2020.
[20] H. Yao, Y. Zhang, Y. Wei, and Y. Tian, "Broad learning system with locality sensitive discriminant analysis for hyperspectral image classification," *Math. Problems Eng.*, vol. 2020, pp. 1–16, Dec. 2020.
[21] G. Xiao, Y. Wei, H. Yao, W. Deng, J. Xu, and D. Pan, "Hierarchical broad learning system for hyperspectral image classification," *IET Image Process.*, vol. 16, no. 2, pp. 554–566, Feb. 2022.
[22] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," 2016, *arXiv:1609.02907*.
[23] S. Melacci and M. Belkin, "Laplacian support vector machines trained in the primal," *J. Mach. Learn. Res.*, vol. 12, no. 3, pp. 1149–1184, 2011.
[24] H. Wang, Y. Cheng, C. L. P. Chen, and X. Wang, "Semisupervised classification of hyperspectral image based on graph convolutional broad network," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 2995–3005, 2021.
[25] Tuya, "Graph convolutional enhanced discriminative broad learning system for hyperspectral image classification," *IEEE Access*, vol. 10, pp. 90299–90311, 2022.
[26] H. Huang and M. Yang, "Dimensionality reduction of hyperspectral images with sparse discriminant embedding," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 9, pp. 5160–5169, Sep. 2015.
[27] Z. Qian, D. N. Metaxas, and L. Axel, "Extraction and tracking of MRI tagging sheets using a 3D Gabor filter bank," in *Proc. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2006, pp. 711–714.
[28] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Trans. Inf. Theory*, vol. IT-13, no. 1, pp. 21–27, Jan. 1967.
[29] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, Dec. 2000.
[30] Y. Chen, H. Jiang, C. Li, X. Jia, and P. Ghamisi, "Deep feature extraction and classification of hyperspectral images based on convolutional neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 6232–6251, Oct. 2016.
[31] G. Camps-Valls, L. Gomez-Chova, J. Munoz-Mari, J. Vila-Frances, and J. Calpe-Maravilla, "Composite kernels for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 1, pp. 93–97, Jan. 2006.

WEI LI is currently with the College of Computer and Artificial Intelligence, Huaihua University, Hunan. His current research interests include signal processing, image processing, and computer vision.

YUANQUAN SHI is currently with the College of Computer and Artificial Intelligence, Huaihua University, Hunan. His current research interests include signal processing, image processing, and computer vision.

LIYUN LI is currently with the College of Computer and Artificial Intelligence, Huaihua University, Hunan. Her current research interests include signal processing, image processing, and computer vision.

XIANGBO MA is currently with Beijing Wanweisheng New Technology Company Ltd. His current research interests include signal processing, image processing, and computer vision.