



Article

Deep Spatial Graph Convolution Network with Adaptive Spectral Aggregated Residuals for Multispectral Point Cloud Classification

Qingwang Wang ^{1,2} , Zifeng Zhang ^{1,2}, Xueqian Chen ^{1,2}, Zhifeng Wang ³, Jian Song ^{1,2} and Tao Shen ^{1,2,*}

- ¹ Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650500, China; wangqingwang@kust.edu.cn (Q.W.); 20212104057@stu.kust.edu.cn (Z.Z.); 20222204177@stu.kust.edu.cn (X.C.); songjian@kust.edu.cn (J.S.)
- ² Yunnan Key Laboratory of Computer Technologies Application, Kunming University of Science and Technology, Kunming 650500, China
- ³ Beijing Anlu International Technology Co., Ltd., Beijing 100043, China; wangzhifeng@anlu-int.com
- * Correspondence: shentao@kust.edu.cn; Tel.: +86-0871-65916593

Abstract: Over an extended period, considerable research has focused on elaborated mapping in navigation systems. Multispectral point clouds containing both spatial and spectral information play a crucial role in remote sensing by enabling more accurate land cover classification and the creation of more accurate maps. However, existing graph-based methods often overlook the individual characteristics and information patterns in these graphs, leading to a convoluted pattern of information aggregation and a failure to fully exploit the spatial–spectral information to classify multispectral point clouds. To address these limitations, this paper proposes a deep spatial graph convolution network with adaptive spectral aggregated residuals (DSGCN-ASR). Specifically, the proposed DSGCN-ASR employs spatial graphs for deep convolution, using spectral graph aggregated information as residuals. This method effectively overcomes the limitations of shallow networks in capturing the nonlinear characteristics of multispectral point clouds. Furthermore, the incorporation of adaptive residual weights enhances the use of spatial–spectral information, resulting in improved overall model performance. Experimental validation was conducted on two datasets containing real scenes, comparing the proposed DSGCN-ASR with several state-of-the-art graph-based methods. The results demonstrate that DSGCN-ASR better uses the spatial–spectral information and produces superior classification results. This study provides new insights and ideas for the joint use of spatial and spectral information in the context of multispectral point clouds.

Keywords: multispectral point clouds; land cover classification; spatial–spectral information; deep spatial graph convolution network; adaptive spectral residuals



Citation: Wang, Q.; Zhang, Z.; Chen, X.; Wang, Z.; Song, J.; Shen, T. Deep Spatial Graph Convolution Network with Adaptive Spectral Aggregated Residuals for Multispectral Point Cloud Classification. *Remote Sens.* **2023**, *15*, 4417. <https://doi.org/10.3390/rs15184417>

Academic Editor: Pablo Rodríguez-González

Received: 30 July 2023

Revised: 30 August 2023

Accepted: 5 September 2023

Published: 7 September 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Navigation is a field of immense importance in modern society, requiring the integration of various disciplines, such as cartography, geography, remote sensing technology, and computer science. Geographic information systems (GISs) and points of interest (POIs) are indispensable for effective navigation. In this context, remote sensing plays a vital role by providing essential components such as accurate base maps and precise land cover models. Thus, land cover classification has emerged as a fundamental research direction within the realm of remote sensing.

Since the early 2000s, the use of laser detection and ranging (LiDAR) technology has substantially contributed to the field of remote sensing. LiDAR has emerged as a valuable tool for collecting high-quality data, offering a rich and detailed data foundation for accurate and refined land cover classification. As an active remote sensing method, LiDAR provides distinct advantages in land cover analysis; for example, it is unaffected

by environmental factors such as illumination, allowing for consistent data collection regarding the spatial distribution of land cover. This capability makes LiDAR a valuable tool for high-resolution and accurate land cover classification. However, because the point cloud is formed using non-Euclidean data with an irregular distribution, point cloud processing became a new challenge. Many studies have achieved notable success with LiDAR point clouds, of which the most classic is the Pointnet series [1–3].

1.1. Data Description

As an evolutionary LiDAR technology, airborne multispectral LiDAR systems can capture the spatial information of land cover while acquiring the spectral intensity of the corresponding points. Teledyne Optech unveiled the inaugural airborne multispectral LiDAR system in 2014, which operates across three channels. Channel 1 operates at a mid-infrared (MIR) wavelength of 1550 nm with a forward-looking angle of 3.5 degrees. Channel 2 operates at a near-infrared (NIR) wavelength of 1064 nm with a nadir-looking angle of zero degrees. Lastly, Channel 3 operates in the green spectrum with a wavelength of 532 nm and a forward-looking angle of seven degrees. Two datasets captured by the system are shown in Figure 1: the Harbor of Tobermory (HT) and the University of Houston (UH).

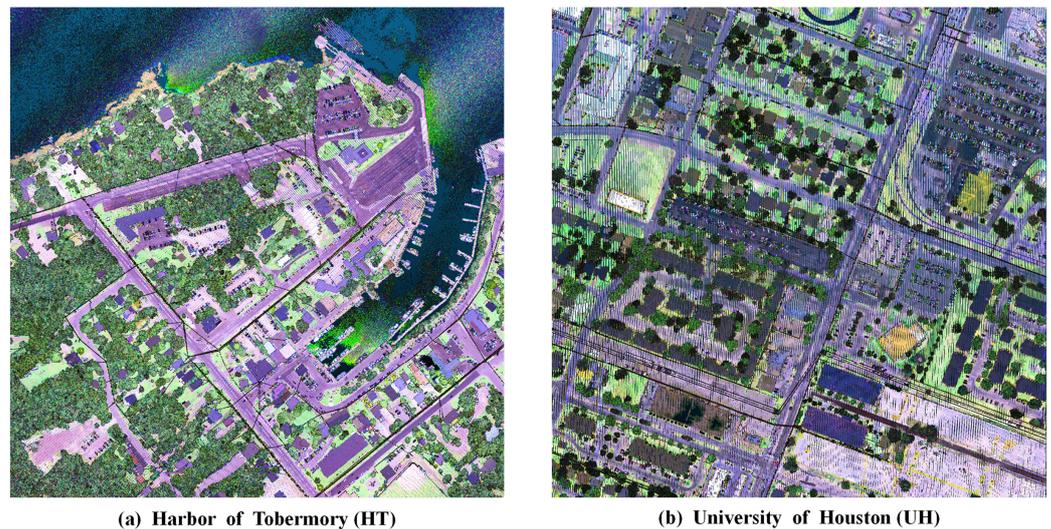


Figure 1. Visualization of two scenes of multispectral point cloud datasets: (a) Harbor of Tobermory (HT) and (b) University of Houston (UH).

1.2. Related Literature

The emergence of multispectral LiDAR has enriched the information dimension of point cloud data. The multispectral point cloud inherits the ability of the traditional point cloud to characterize the spatial distribution of land cover while collecting corresponding spectral information for each point. With the increase in the information dimension, researchers have been faced with a new dilemma of how to effectively and jointly use the rich spatial–spectral information in multispectral point clouds.

1.2.1. Image-Oriented Methods

Several researchers have transformed 3D multispectral point clouds into 2D images to employ traditional image-oriented methods, such as support vector machine (SVM) [4], Adaboost [5], random forest [6], Markov random field [7], and conditional random field [8].

Other researchers have proposed deep learning models specifically designed for point clouds. Yu et al. introduced the CapViT model, a cross-context capsule vision transformer, for land cover classification using multispectral LiDAR data. It uses three streams of capsule transformer encoders to capture long-range global feature interactions at different context scales and effectively fuses cross-context feature semantics for accurate land cover

type inferences [9]. ESA-CapsNet uses a novel capsule encoder–decoder architecture and a capsule-based attention module to extract informative feature semantics and enhance feature saliency and robustness [10]. Wang et al. proposed a neural network architecture for learning with point clouds that captures semantically similar structures in deeper layers despite the long distance between them in the original input space. The network utilizes a dynamic graph convolutional neural network (DGCNN) approach, which combines global shape structure with local neighborhood information to improve the learning process [11]. Liu et al. proposed RS-CNN [12], a relation–shape convolutional neural network that extends a regular-grid CNN to irregular configurations for point cloud analysis by learning from the geometric topology constraint among points. Shape awareness and robustness are achieved by learning a high-level relationship expression from predefined geometric priors, leading to contextual shape-aware learning for point cloud analysis. [13–19].

However, the transformations performed by these methods result in the loss of the original information of the multispectral point cloud.

1.2.2. Point-Oriented Methods

Traditional Methods Jing et al. proposed SE-Pointnet++ by embedding the squeeze-and-excitation block (SE block) into the Pointnet++ network to improve the performance of multispectral LiDAR point cloud classification by modeling the interdependence between channels. They utilized Pointnet++, DGCNN, GACNet, and RSCNN as comparison models to demonstrate the superiority of SE-Pointnet++ in accomplishing multispectral LiDAR point cloud feature classification [1–3,11,12]. Hu et al. proposed RandLA-Net, a lightweight neural architecture that uses random point sampling and a novel local feature aggregation module to efficiently perform semantic segmentation on large-scale 3D point clouds [20]. Wang et al. proposed a TMDE algorithm for extracting discriminative geometric–spectral features from multispectral point cloud data. The algorithm preserves the intraclass sample distribution and maximizes the distance between different classes [21–24].

Graph-Based Methods Graph neural networks have received increasing attention from researchers due to their inherent ability to accurately characterize non-Euclidean data [25]. Some examples of these networks include GAC [26], FR-GCNet [27], GACNN [28], and MaSGCN [29]. For graph-based methods, the most immediate challenge is effectively measuring the similarity between points to represent a multispectral point cloud as a graph. Once a suitable similarity metric is found, state-of-the-art graph neural networks such as GCN [30], GAT [31], GCBNet [32], and GCNII [33] can be used to classify multispectral point clouds.

Despite these advances, effective methods are still needed to utilize the spatial–spectral information contained in multispectral point clouds without losing valuable information. Further research in this area holds promise for advancing the field of multispectral point cloud analysis and classification.

1.3. Motivation and Contributions

Researchers have either used the spatial distance to construct a graph or spectral similarity or have simply combined the similarities of the two in equal proportions to produce a joint graph. To verify the advantages and disadvantages of these technical routes, we used a simple GCN to classify the two previously mentioned multispectral point cloud datasets using each of the above three methods to construct a graph. The classification results are visualized in Figure 2.

Upon visualizing the classification results, we found that the spatial graph tends to assign neighboring land covers to the same class, resulting in a contiguous distribution pattern. Conversely, the spectral graph exhibits superior capability in capturing long-range dependencies and effectively delineating boundaries between land cover types. However, the spectral graph also demonstrates a greater tendency to incorporate irrelevant land cover information and is more susceptible to interference from complex spectral signatures, such as those originating from water bodies. Ideally, the strengths of both should be combined

to achieve finer classification while maintaining the robustness of the spatial graph and taking advantage of the high-quality performance of the spectral graph on the boundary. The visualization results (Figure 2c,f) show that simply combining the two does not achieve the ideal state, so achieving the reasonable joint use of spatial–spectral information is an important problem to be solved.

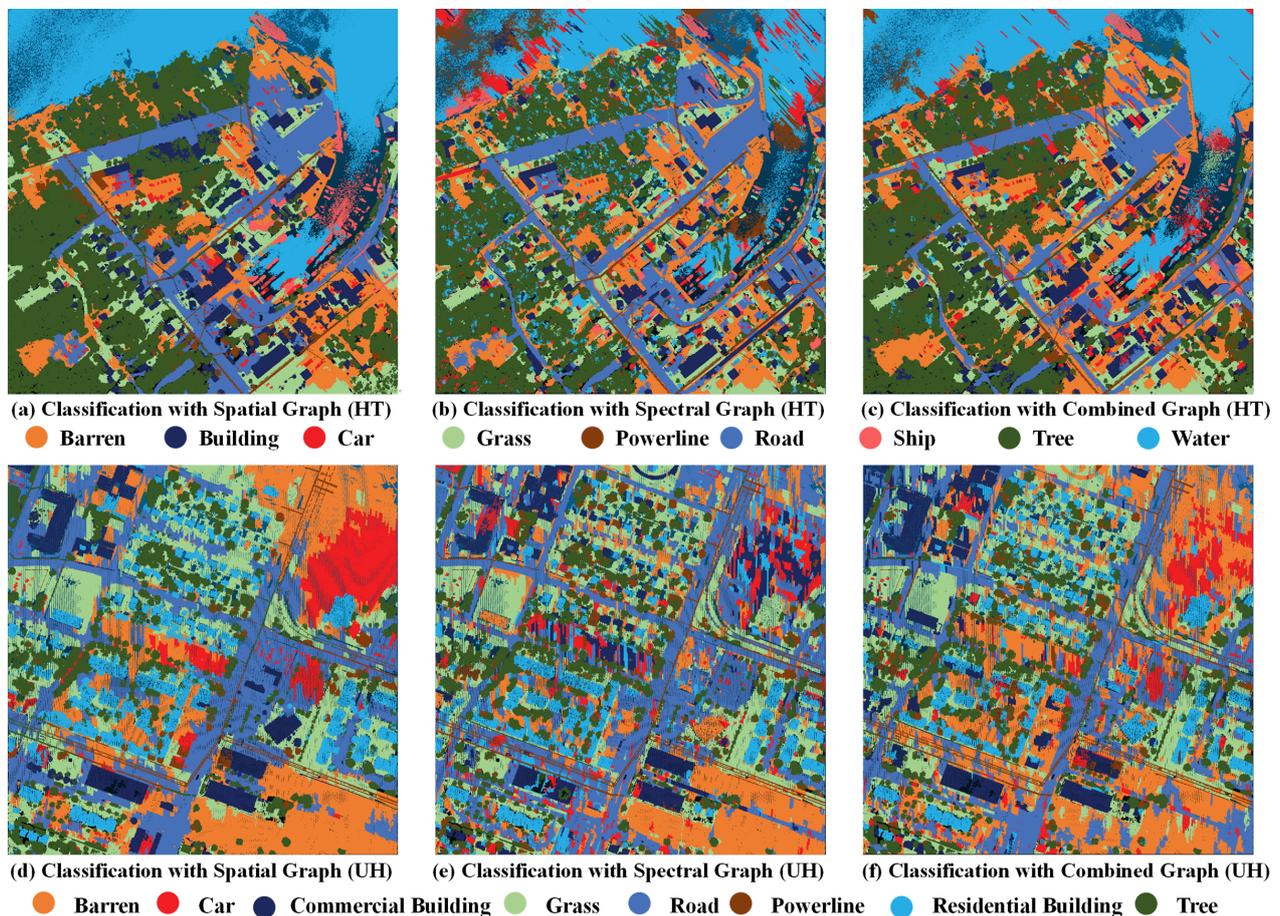


Figure 2. Visualization of differences between three technical routes for constructing graphs from two datasets: (a) classification with a spatial graph for the HT dataset; (b) classification with a spectral graph for the HT dataset; (c) classification with a combined graph on the HT dataset; (d) classification with a spatial graph for the UH dataset. (e) classification with a spectral graph for the UH dataset. (f) classification with a combined graph for the UH dataset.

To address this problem, we developed a deep spatial graph convolution network with adaptive spectral aggregated residuals (DSGCN-ASR), which inputs both spatial and spectral graphs into the network. The proposed DSGCN-ASR uses a spatial graph to perform multiple layers of graph convolutions on multispectral point clouds and uses the information aggregated by the spectral graph as residuals, which are adaptively added during each convolution. Specifically, the main contributions can be summarized as follows:

1. A novel framework was developed for the simultaneous use of spatial and spectral information in multispectral point clouds. The spatial and spectral graphs are treated differently to preserve the robustness of the spatial graph in capturing the nearby land cover relationships while harnessing the discriminative power of the spectral graph in distinguishing between various features in proximity.
2. A deep graph neural network, DSGCN-ASR, was developed to learn the implicit relationships between points in a multispectral point cloud to overcome the insufficient capability of shallow graph neural networks in fitting the nonlinearity of multispectral

point clouds in complex remote sensing scenes. Additionally, the spectral aggregated residuals were adaptively added to learn the spectral relationship between points, simultaneously addressing the oversmoothing problem of deep features.

The remainder of this paper is organized as follows. Section 2 describes the methodology and specific algorithms for the proposed DSGCN-ASR. Section 3 outlines the performance of the proposed DSGCN-ASR through experiments, and Section 4 provides the conclusions.

2. Methodology

In this section, we describe, in detail, the principles and implementation of the proposed DSGCN-ASR and provide the corresponding algorithm. The overall network structure is shown in Figure 3.

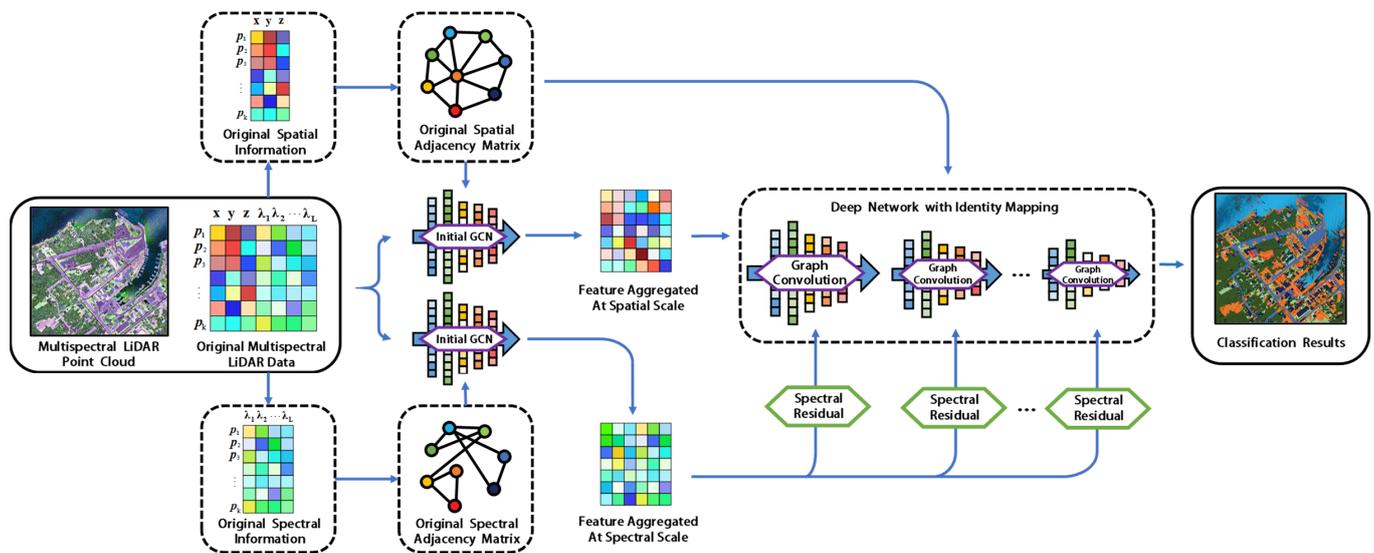


Figure 3. Overall structure of the proposed DSGCN-ASR.

2.1. Construction of Spatial and Spectral Graphs

The data form of a multispectral point cloud can be viewed as a set of point clouds collected by multiple lasers of different wavelengths in the same scene. However, in practice, multiple bands of data are commonly integrated into a single point cloud. The integrated multispectral point cloud can be denoted as $\mathbf{P} = [p_1, p_2, p_3, \dots, p_k] \in \mathbb{R}^{(L+3) \times k}$, where L is the number of bands, and k is the number of points in the multispectral point cloud. A single point can be represented as $p_i = [x, y, z, \lambda_1, \lambda_2, \dots, \lambda_L]$, where $i \in (1, k)$ is the index of the point.

For a graph $(G = (V, E))$, V is the set of nodes, and E is the set of edges. For each node (i) , its corresponding feature (x_i) can be represented by matrix $X_{N \times D}$, where N denotes the number of nodes, and D denotes the feature dimension of each node. For a multispectral point cloud, matrix $X_{N \times D}$ corresponds to point set $\mathbf{P} = [p_1, p_2, p_3, \dots, p_k] \in \mathbb{R}^{(L+3) \times k}$, which can be obtained by transposing \mathbf{P} . Thus, the number of nodes (N) is the number of points (k), and the number of features (D) is equal to $(L + 3)$.

Regarding the set of edges (E), we separately measure the similarity between each point for the spatial and spectral information and compute two adjacency matrices, i.e., a spatial adjacency matrix and a spectral adjacency matrix. Specifically, we separately compute the Euclidean distance between points for the spatial and spectral information to obtain the distance matrix. Because the larger the Euclidean distance, the weaker the correlation between the points, each value in the matrix is subtracted from the maximum

value in the distance matrix. Finally, the overall matrix is max–min-normalized to obtain adjacency matrices $\mathbf{A}_{spatial}$ and $\mathbf{A}_{spectral}$.

$$\mathbf{A}_{spatial} = Normalized(DisSpatial.max - DisSpatial) \quad (1)$$

$$DisSpatial = Dis\mathbf{X} + Dis\mathbf{Y} + Dis\mathbf{Z} \quad (2)$$

$$\mathbf{A}_{spectral} = Normalized(DisSpectral.max - DisSpectral) \quad (3)$$

$$DisSpectral = Dis\lambda_1 + Dis\lambda_1 + \dots + Dis\lambda_L \quad (4)$$

where $DisSpatial$ and $DisSpectral$ are the spatial and spectral distance matrices, respectively. $DisSpatial.max$ is a matrix of the same dimension as $DisSpatial$, with each element corresponding to the max value in $DisSpatial$, and $DisSpectral.max$ is the same as for $DisSpectral$. $Dis\mathbf{X}$, $Dis\mathbf{Y}$, $Dis\mathbf{Z}$, $Dis\lambda_1$, $Dis\lambda_2$, \dots , $Dis\lambda_L$ are the distance matrices of the spatial and spectral information. The calculation process is shown in Algorithm 1.

Algorithm 1: Construction of graphs for multispectral point cloud.

Input: Multispectral point cloud, $\mathbf{P} = [p_1, p_2, p_3, \dots, p_k] \in \mathbb{R}^{(L+3) \times k}$

Output: Feature matrix, $X_{k \times (L+3)}$, Spatial adjacency matrix, $\mathbf{A}_{spatial}$, Spectral adjacency matrix, $\mathbf{A}_{spectral}$

1 Expand each point p_i in multispectral point cloud

$\mathbf{P} = [p_1, p_2, p_3, \dots, p_k] \in \mathbb{R}^{(L+3) \times k}$ into its feature vector $p_i = [x, y, z, \lambda_1, \lambda_2, \dots, \lambda_L]$,

where $(i = 1, 2, 3, \dots, k)$. \mathbf{P} can be represent as

$$\begin{bmatrix} x_1 & y_1 & z_1 & \lambda_{11} & \dots & \lambda_{L1} \\ x_2 & y_2 & z_2 & \lambda_{12} & \dots & \lambda_{L2} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ x_k & y_k & z_k & \lambda_{1k} & \dots & \lambda_{Lk} \end{bmatrix}$$

2 Split each column in multispectral point cloud \mathbf{P} as separate vectors \mathbf{X} , \mathbf{Y} , \mathbf{Z} , λ_1 , $\lambda_2, \dots, \lambda_L$, then max-min normalize each vector.

3 For each vector, perform the following calculation, taking \mathbf{X} as an example.

$Dis\mathbf{X} = \mathbf{X}^2repeat - 2\mathbf{X}\mathbf{X}^T + \mathbf{X}^2repeat^T =$

$$\begin{bmatrix} (x_1 - x_1)^2 & (x_1 - x_2)^2 & \dots & (x_1 - x_k)^2 \\ (x_2 - x_1)^2 & (x_2 - x_2)^2 & \dots & (x_2 - x_k)^2 \\ \vdots & \vdots & \ddots & \vdots \\ (x_k - x_1)^2 & (x_k - x_2)^2 & \dots & (x_k - x_k)^2 \end{bmatrix}, \mathbf{X}^2repeat = \begin{bmatrix} x_1^2 & \dots & x_1^2 \\ x_2^2 & \dots & x_2^2 \\ \vdots & \ddots & \vdots \\ x_k^2 & \dots & x_k^2 \end{bmatrix}.$$

Obtain the following matrix $Dis\mathbf{X}$, $Dis\mathbf{Y}$, $Dis\mathbf{Z}$, $Dis\lambda_1$, $Dis\lambda_2, \dots, Dis\lambda_L$.

4 Feature matrix can be obtain as $X_{k \times (L+3)} = [\mathbf{X}, \mathbf{Y}, \mathbf{Z}, \lambda_1, \lambda_2, \dots, \lambda_L]$

5 Calculate the spatial and spectral distance matrices

$$DisSpatial = Dis\mathbf{X} + Dis\mathbf{Y} + Dis\mathbf{Z} \quad DisSpectral = Dis\lambda_1 + Dis\lambda_1 + \dots + Dis\lambda_L$$

6 Calculate the spatial and spectral adjacency matrices

$$\mathbf{A}_{spatial} = Normalized(DisSpatial.max - DisSpatial),$$

$$\mathbf{A}_{spectral} = Normalized(DisSpectral.max - DisSpectral)$$

7 **Return:** $X_{k \times (L+3)}$, $\mathbf{A}_{spatial}$, $\mathbf{A}_{spectral}$

2.2. Deep Spatial Graph Convolution Network with Adaptive Spectral Aggregated Residuals

DSGCN-ASR effectively addresses the limitations of shallow graph neural networks in capturing the nonlinearity of multispectral point clouds in complex remote sensing scenes. In addition, it tackles the problem of previous methods lacking the ability to fully exploit joint spatial–spectral information. By incorporating several key techniques, DSGCN-ASR provides enhanced modeling and classification capabilities, ensuring the optimal use of spatial–spectral information.

Convolutional neural networks (CNNs) are important in the field of computer vision. The central CNN technique involves extracting features using a convolutional kernel by weighting the pixel values in the neighborhood of the pixel. Similarly, graph convolution aggregates information from related nodes according to the set of edges (E) to achieve feature extraction. The generalization of the graph convolutional network can be denoted as

$$\mathbf{H}^{l+1} = f(\mathbf{H}^l, \mathbf{A}) = \sigma(\mathbf{A}\mathbf{H}^l\mathbf{W}^l) \quad (5)$$

where \mathbf{H} denotes the hidden layer of the network, \mathbf{A} denotes the adjacency matrix, l is the index of the hidden layer, and \mathbf{W}^l is the weight parameter matrix of the l th layer.

As the convolutional neural network aggregates information based on the pixel neighborhoods of the images and, according to Tobler's first law of geography, a stronger correlation exists between neighboring land covers, we use the spatial graph for graph convolution operations in the deep backbone network. This also allows the model to extract more complex and abstract features, strengthening its capacity to capture the intricate non-linearity present in the data. Additionally, DSGCN-ASR incorporates the adaptive spectral aggregated residuals (ASR) technique. ASR adaptively adjusts the weights of spectral features from the multiple channels in multispectral data. The feature obtained after graph convolution of the spectral graph is used as the residuals to be added to the hidden layer. Specifically, two initial graph convolutions of the same dimension are introduced before the backbone. After initial convolution of the spatial graph, the initial input feature (\mathbf{H}^0) is obtained; after initial convolution of the spectral graph, the spectral residual feature (\mathbf{R}) is obtained. The initial convolution can be represented as

$$\mathbf{H}^0 = \sigma(\mathbf{A}_{spatial} X_{k \times (L+3)} \mathbf{W}^{spatial}) \quad (6)$$

$$\mathbf{R} = \sigma(\mathbf{A}_{spectral} X_{k \times (L+3)} \mathbf{W}^{spectral}) \quad (7)$$

where \mathbf{H}^0 serves as the input to the backbone, and \mathbf{R} serves as the residual.

The pattern of combining spectral residuals is shown in Figure 4. In the aggregation of spectral residuals, we use both concatenation and summation. The spectral residuals (\mathbf{R}) are concatenated to the right of the hidden layer (\mathbf{H}^l) before convolution in each layer. Given the use of spatial graphs in the graph convolution operation, we classify the hidden layer features of the network as spatial. As such, the residuals aggregated using spectral graphs are classified as spectral. To balance the contributions of spatial–spectral information, we introduce a trainable adaptive parameter (α). After convolution, the spectral residuals (\mathbf{R}) are summed with the hidden layer using an adaptive weight (α) and added as residuals to the new hidden layer. This adaptive weighting mechanism enables the model to focus on the most informative channels, enhancing its ability to capture nonlinearity and improving classification accuracy. Thus, hidden layer propagation with adaptive spectral residuals can be represented as

$$\mathbf{H}^{l+1} = \sigma\left(\left\langle \mathbf{A}_{spatial} \mathbf{H}^l \mid \mathbf{R} \right\rangle \mathbf{W}^l + \left((1 - \alpha) \mathbf{H}^l + \alpha \mathbf{R} \right)\right) \quad (8)$$

where the adaptive weight (α) is a trainable parameter defined in the network. To ensure that the weight of the spectral residuals does not exceed 0.5 for each addition, we apply a sigmoid function to α and divide it by 2 before each use.

$$\alpha = \text{sigmoid}(\alpha) / 2 \quad (9)$$

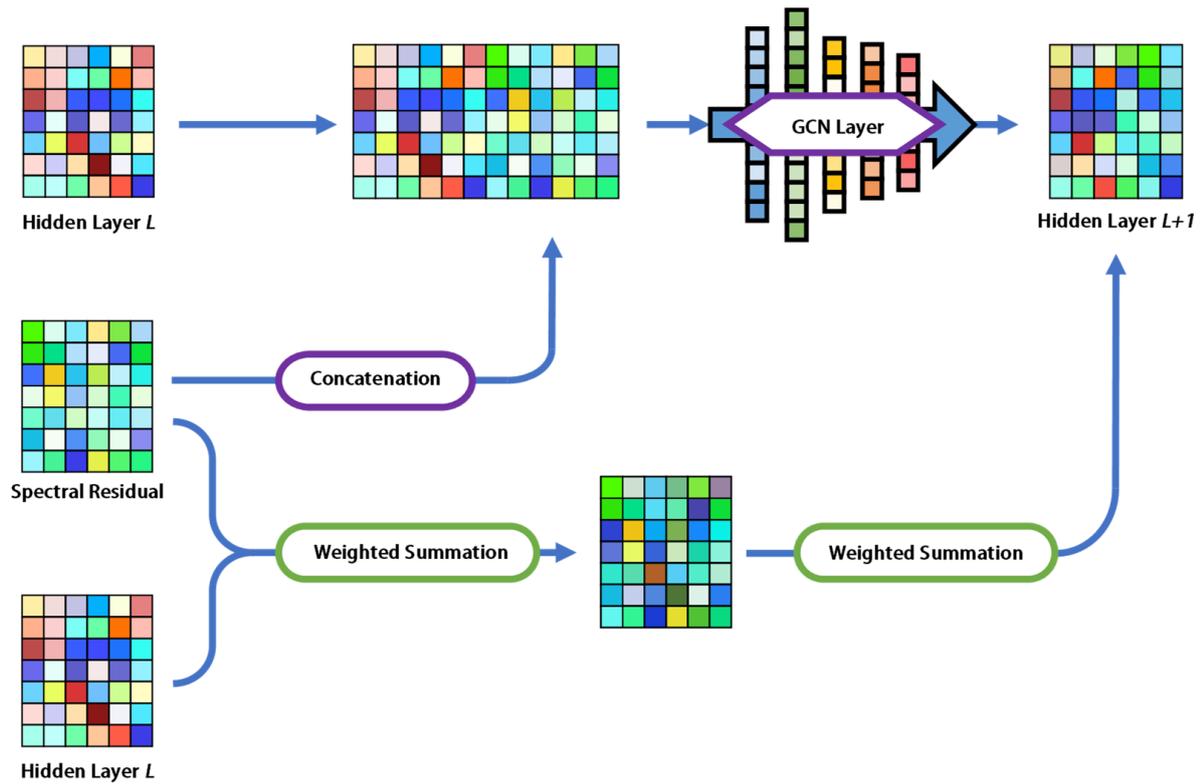


Figure 4. Combining patterns of adaptive spectral residuals in each layer.

To address the issue of deep-feature oversmoothing, we introduce a weight parameter, denoted as β , which decreases as the depth of the network increases. This weight parameter balances the contribution of the deep network weights, inspired by the concept of identity mapping in previous research [33]. As the network layers deepen, the contribution of convolution to forward propagation diminishes, while the adaptive combination of spectral residuals with the previous hidden layer gradually takes precedence. This approach effectively alleviates the problem of the oversmoothing of deep features caused by excessive feature aggregation. The idea of identity mapping has been demonstrated to be effective in mitigating this issue in previous studies, and we incorporate this idea by introducing β . The final hidden-layer propagation pattern can be denoted using the following equation:

$$\mathbf{H}^{l+1} = \sigma\left(\beta\left\langle \mathbf{A}_{spatial} \mathbf{H}^l \mid \mathbf{R} \right\rangle \mathbf{W}^l + (1 - \beta)\left((1 - \alpha)\mathbf{H}^l + \alpha\mathbf{R}\right)\right) \quad (10)$$

where β decays with the layer number (l) of the network as follows:

$$\beta = \ln(1/2l + 1) \quad (11)$$

In the proposed DSGCN-ASR, we use negative log-likelihood loss to train the model. The loss function can be denoted as

$$\mathcal{L}_{loss} = -\sum_{i=1}^k \sum_{j=1}^C y_{ij} \log(p_{ij}) \quad (12)$$

where k is the number of points in the dataset, C is the number of classes, y_{ij} is the ground truth of the i th point belonging to the j th class, and p_{ij} is the predicted probability of the i th point belonging to the j th class.

By integrating these techniques, DSGCN-ASR effectively enhances the collaborative use of spatial-spectral information and strengthens the nonlinear fitting capability.

Consequently, it provides notable advancements compared with prior methods in the classification performance of multispectral point clouds in complex remote sensing scenes.

3. Experiments

We conducted a series of comparative experiments, ablation studies, and parametric analyses using the proposed DSGCN-ASR. Two multispectral point cloud datasets of real scenes were used to conduct the experiments, i.e., Harbor of Tobermory (HT) and University of Houston (UH), as shown in Figure 1.

The HT dataset was further subjected to manual labeling, incorporating nine distinct classes, namely barren, building, car, grass, powerline, road, ship, and tree, following the labeling scheme established in a previous study [21]. The UH dataset underwent manual classification into eight classes, encompassing barren, car, commercial buildings, grass, road, powerline, residential buildings, and tree, as shown in Figure 5. All experiments were performed on a device with an Intel (R) Core (TM) CPU i5-12600KF @3.70 GHz and one NVIDIA GeForce RTX 3060 GPU with 12 GB of memory. However, because HT contains 7,181,982 points and UH contains 4,436,470 points, which cannot be directly processed by the current device, we used a previously reported method [34] to segment the multispectral point cloud into superpoints. The HT dataset was segmented into 9606 superpoints, and UH was segmented into 9350 superpoints. We used 10% of the superpoints as the training set.

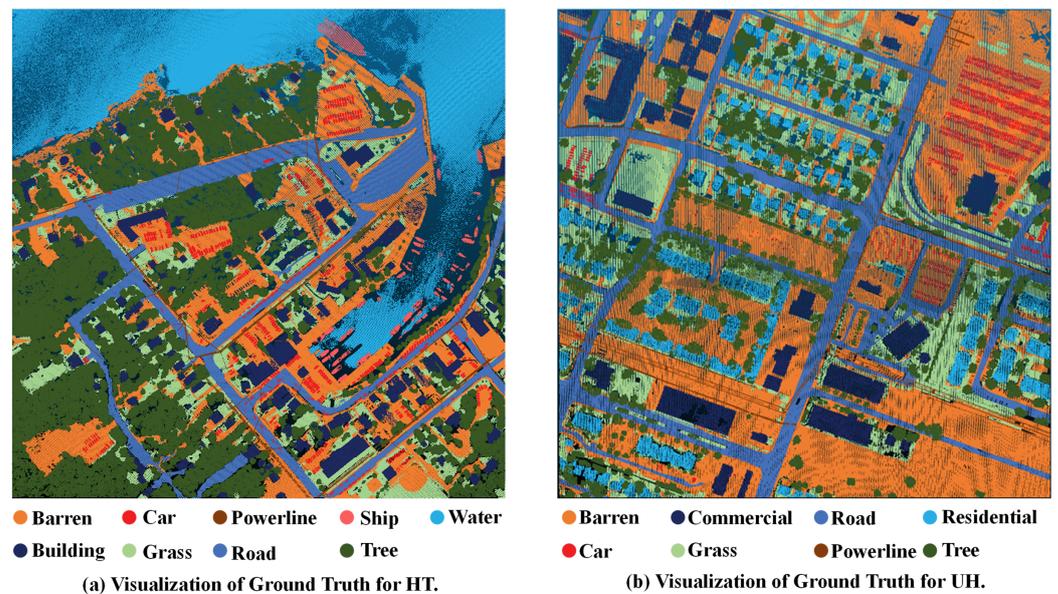


Figure 5. Visualization of ground truth for two datasets: (a) HT; (b) UH.

To numerically measure the multispectral point cloud classification performance, we used precision, recall, F score, and IoU to evaluate each set of experiments. The above metrics were used for each class. To evaluate the overall performance of the whole scene, we used macro averaging to calculate the above metrics, in addition to overall accuracy (OA).

3.1. Comparative Experiments

To validate the performance of the proposed DSGCN-ASR, several state-of-the-art graph neural networks (GCN [30], GAT [31], GCBNet [32], GCNII [33], and MaSGCN [29]), were selected to classify multispectral point clouds and were comparatively analyzed. For this comparison, we constructed the graphs following the method outlined in Algorithm 1 and combined the spatial–spectral graphs on the same scale.

$$\mathbf{A} = \text{Normalized} \left(\frac{\mathbf{A}_{\text{spatial}} + \mathbf{A}_{\text{spectral}}}{2} \right) \quad (13)$$

The input points were all $X_{k \times (L+3)}$, and the same 10% were used as training samples. The final classification results of all methods were remapped back to the original data according to the index of the superpoints and evaluated on the original data. The labels of the superpoints were generated based on the voting of the labels of the points within the same superpoint. Therefore, the segmentation of superpoints inevitably caused a loss of classification performance, which we later specifically analyzed.

3.1.1. HT Classification Results

The overall evaluation metrics for the HT classification results are shown in Table 1. The proposed DSGCN-ASR outperforms the other methods overall in the classification of HT, with an OA of 87.57%, macro precision of 74.23%, macro recall of 69.45%, macro F score of 71.76%, and MIoU of 59.51%. The OA of DSGCN-ASR is higher than that of the second-best method by 2.73%, outperforming the next-best method by 1.74% and 2.14% for macro recall and MIoU, respectively.

Table 1. Overall evaluation metrics (%) for classification results on the HT dataset.

Method	GCN [30]	GCNII [33]	GAT [31]	GCBNet [32]	MaSGCN [29]	DSGCN-ASR (Ours)
OA	81.36	84.83	84.77	84.70	82.81	87.57
Macro precision	72.30	79.17	71.71	77.84	69.55	74.23
Macro recall	61.32	64.68	62.04	66.58	67.71	69.45
Macro F score	66.36	71.19	66.53	71.77	68.62	71.76
MIoU	51.84	55.96	53.08	57.36	54.94	59.51

Maximum values in the same metrics are marked in bold.

The classification results on the HT dataset are visualized in Figure 6. The visualized results show that the proposed DSGCN-ASR learns the information of the spatial distribution of the land cover as expected and achieves a fine delineation of the boundary with the help of the spectral information.

The other methods produced different apparent misclassifications; in contrast, the classification result of the proposed DSGCN-ASR is more in line with the ground truth. GCN and GAT produced cluttered distributions of misclassified points, GCNII and GCBNet classified large areas of water as car, and MaSGCN classified an area of water as barren.

The evaluation metrics for the classification results of each class are shown in Table 2. The classification results of the proposed DSGCN-ASR are more balanced, performing relatively well in all classes. In particular, the proposed method outperforms the other methods in classifying building, grass, tree, and water classes. Combining the visualizations revealed that the DSGCN-ASR confused the barren and road; tree and powerline; and car, ship, and building classes because the spectral information of these classes is similar, and they are relatively close in spatial distribution, which makes distinguishing them challenging. In addition, the small number of powerline points in the dataset may have hindered the ability to learn these features, leading to poor performance.

Car is a small land cover target, and the spectral information associated with the car class is relatively complex. This results in a low number of points for car in the original point cloud data, so car is easily confused with barren. As such, when segmenting the multispectral point cloud, we were unable to oversegment it due to the memory limitations of the experimental equipment, which led to a reduction in the number of effective samples. As shown in Figure 6b, superpoint segmentation strongly impacted car classifications, which is the reason for the poor performance of all models on this class. A similar situation occurred for other classes.

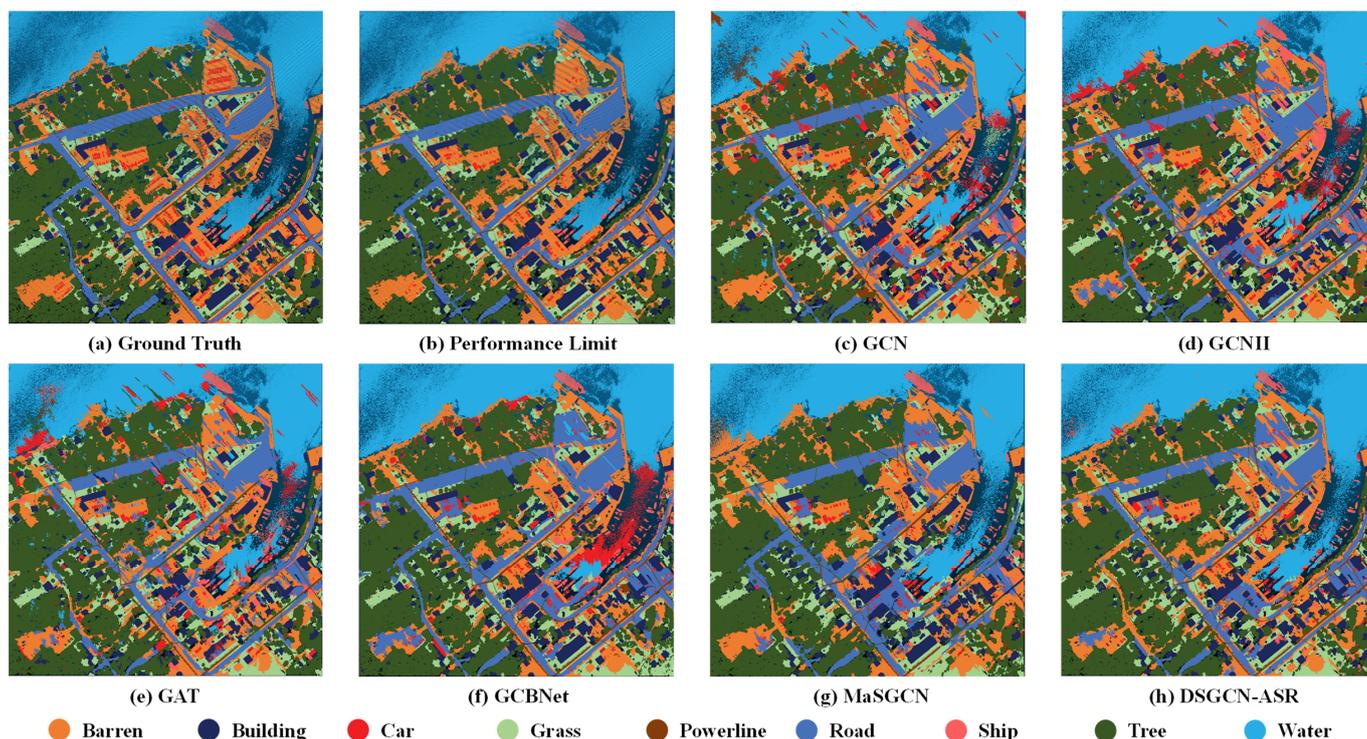


Figure 6. Visualization of classification results on HT dataset. (a) Visualization of the ground truth; (b) performance limit due to superpoint segmentation; visualization of classification results of (c) GCN, (d) GCNII, (e) GAT, (f) GCBNet, (g) MaSGCN, and (h) DSGCN-ASR (ours).

Table 2. Evaluation metrics (%) for classification results in each class on the HT dataset.

Method	Class	Barren	Building	Car	Grass	Powerline	Road	Ship	Tree	Water
GCN [30]	Precision	75.50	72.88	33.42	87.70	66.37	82.38	57.34	83.85	91.24
	Recall	82.45	69.97	20.24	81.09	4.89	70.57	34.61	99.20	88.85
	F-score	78.82	71.39	25.21	84.27	9.11	76.02	43.16	90.88	90.03
	IoU	65.04	55.51	14.42	72.81	4.77	61.31	27.52	83.29	81.86
GCNII [33]	Precision	71.42	79.15	38.97	88.60	82.17	88.57	82.58	89.35	91.71
	Recall	85.88	77.71	28.40	88.44	9.12	67.80	31.06	99.40	94.29
	F-score	77.99	78.42	32.86	88.52	16.42	76.80	45.14	94.11	92.98
	IoU	63.92	64.50	19.66	79.41	8.94	62.34	29.15	88.87	86.89
GAT [31]	Precision	72.37	71.53	29.14	88.46	51.50	79.46	70.75	92.60	89.59
	Recall	79.67	69.64	18.86	82.95	21.48	67.78	35.53	98.98	83.46
	F-score	75.85	70.57	22.90	85.62	30.32	73.16	47.31	95.68	86.42
	IoU	61.09	54.53	12.93	74.85	17.87	57.67	30.98	91.72	76.08
GCBNet [32]	Precision	62.48	90.04	42.23	88.37	69.87	91.39	75.02	90.54	90.63
	Recall	85.51	69.97	32.96	81.76	41.27	58.92	32.42	99.33	97.02
	F-score	72.21	78.75	37.02	84.94	51.89	71.65	45.27	94.73	93.71
	IoU	56.50	64.95	22.72	73.82	35.04	55.83	29.26	89.99	88.17
MaSGCN [29]	Precision	59.29	81.43	40.32	71.74	41.06	73.88	67.51	94.90	95.82
	Recall	71.21	73.56	57.01	67.02	33.76	55.67	56.19	97.41	97.60
	F-score	64.70	77.29	47.23	69.30	37.05	63.50	61.33	96.13	96.70
	IoU	47.82	62.99	30.92	53.02	22.74	46.52	44.23	92.56	93.62
DSGCN-ASR (ours)	Precision	72.18	88.36	29.28	90.95	69.95	77.08	47.02	95.41	97.86
	Recall	78.40	78.74	38.73	86.34	33.12	64.81	48.64	99.05	97.24
	F-score	75.16	83.27	33.35	88.59	44.96	70.41	47.81	97.19	97.55
	IoU	60.21	71.34	20.01	79.51	29.00	54.34	31.42	94.54	95.21

Maximum values in the same metrics are marked in bold.

3.1.2. UH Classification Results

The overall evaluation metrics for the classification results on the UH dataset are shown in Table 3. In the UH classification, DSGCN-ASR performs better than the other methods, with an OA of 78.20%, macro precision of 73.03%, macro recall of 65.41%, macro F score of 69.21%, and MIoU of 54.02%. The OA of DSGCN-ASR is higher than that of the second-best method by 8.73%, outperforming the second-best method by 3.04%, 0.80%, and 3.07% for macro recall, macro F score, and MIoU, respectively.

Table 3. Overall evaluation metrics (%) for classification results on the UH dataset.

Method	GCN [30]	GCNII [33]	GAT [31]	GCBNet [32]	MaSGCN [29]	DSGCN-ASR (Ours)
OA	67.39	66.80	61.31	69.47	64.53	78.20
Macro precision	67.76	72.75	66.31	75.26	68.48	73.03
Macro recall	54.30	58.29	50.87	62.37	57.86	65.41
Macro F score	60.29	64.72	57.57	68.21	62.72	69.01
MIoU	42.89	46.52	38.81	50.95	45.03	54.02

Maximum values in the same metrics are marked in bold.

The classification results on the UH dataset are visualized in Figure 7. The figure demonstrates that the classification results produced by the proposed DSGCN-ASR are close to the ground truth and performance limit. This is especially evident in the parking lot area in the upper-right corner of the scene. In addition, the rectangular area in the middle of the scene demonstrates the contrast among the methods. The ground truth for this area is regular rectangular barren land; however, GCNII, GCBNet, and MasGCN all misclassify this area as road or car. GCN and GAT perform relatively better in this area but are more disturbed than the proposed DSGCN-ASR. However, DSGCN-ASR, as with MaSGCN, incorrectly classifies road in the upper-right corner of the scene as barren. Overall, the proposed DSGCN-ASR retains the robustness of the spatial graph regarding land cover distribution with less interference on the UH dataset when using the spectral graph to enhance the accuracy of boundary classification.

The evaluation metrics for the classification results for each class are shown in Table 4. The proposed DSGCN-ASR achieved relatively high-quality performance for all classes; however, the metrics for car are poor. Combining the visualizations, we concluded that the poor classification of car was due to the effect of superpoint segmentation, as shown in Figure 7b. The proposed DSGCN-ASR provides substantial advantages over the other methods in the classes of barren, road, powerline, and tree. This conclusion is consistent with the visualization results. Commercial and residential buildings are difficult to distinguish because they are both buildings that have similar spatial and spectral information. However, the proposed DSGCN-ASR outperforms the other methods on the UH dataset in general.

Table 4. Evaluation metrics (%) for classification results in each class on the UH dataset.

Method	Class	Barren	Car	Commercial	Grass	Road	Powerline	Residential	Tree
GCN [30]	Precision	54.13	36.22	70.84	81.62	76.47	70.78	74.60	77.39
	Recall	80.05	15.41	49.31	74.47	51.76	16.41	51.94	95.05
	F score	64.59	21.63	58.14	77.88	61.73	26.64	61.24	85.32
	IoU	47.70	12.12	40.99	63.77	44.65	15.37	44.13	74.39
GCNII [33]	Precision	44.95	45.80	81.14	84.53	85.18	81.93	77.36	81.10
	Recall	83.79	13.00	59.87	73.33	49.13	22.34	68.81	96.07
	F score	58.51	20.26	68.90	78.53	62.31	35.10	72.83	87.95
	IoU	41.35	11.27	52.56	64.65	45.26	21.29	57.27	78.50

Table 4. Cont.

Method	Class	Barren	Car	Commercial	Grass	Road	Powerline	Residential	Tree
GAT [31]	Precision	38.88	55.28	59.79	80.38	77.53	62.68	79.12	76.85
	Recall	80.00	12.65	31.70	74.45	47.13	20.40	46.77	93.85
	F score	52.33	20.59	41.43	77.30	58.62	30.78	58.79	84.50
	IoU	35.44	11.48	26.13	63.00	41.47	18.19	41.63	73.16
GCBNet [32]	Precision	46.94	60.35	80.20	87.11	85.25	72.86	83.07	86.28
	Recall	85.38	20.16	71.41	78.00	46.53	30.04	71.46	95.97
	F score	60.58	30.22	75.55	82.30	60.20	42.54	76.83	90.87
	IoU	43.45	17.80	60.71	69.93	43.06	27.02	62.37	83.26
MaSGCN [29]	Precision	43.20	39.28	72.71	85.18	78.14	64.96	86.02	78.32
	Recall	80.52	20.97	73.93	66.49	42.97	33.89	49.76	94.35
	F score	56.23	27.35	73.31	74.68	55.45	44.54	63.05	85.59
	IoU	39.11	15.84	57.87	59.59	38.36	28.65	46.04	74.81
DSGCN-ASR (ours)	Precision	75.78	33.89	67.28	83.02	72.70	79.49	83.82	88.28
	Recall	80.72	28.60	71.00	80.46	68.65	32.39	67.05	94.42
	F score	78.18	31.02	69.09	81.72	70.62	46.02	74.51	91.24
	IoU	64.17	18.36	52.77	69.09	54.58	29.89	59.37	83.90

Maximum values in the same metrics are marked in bold.

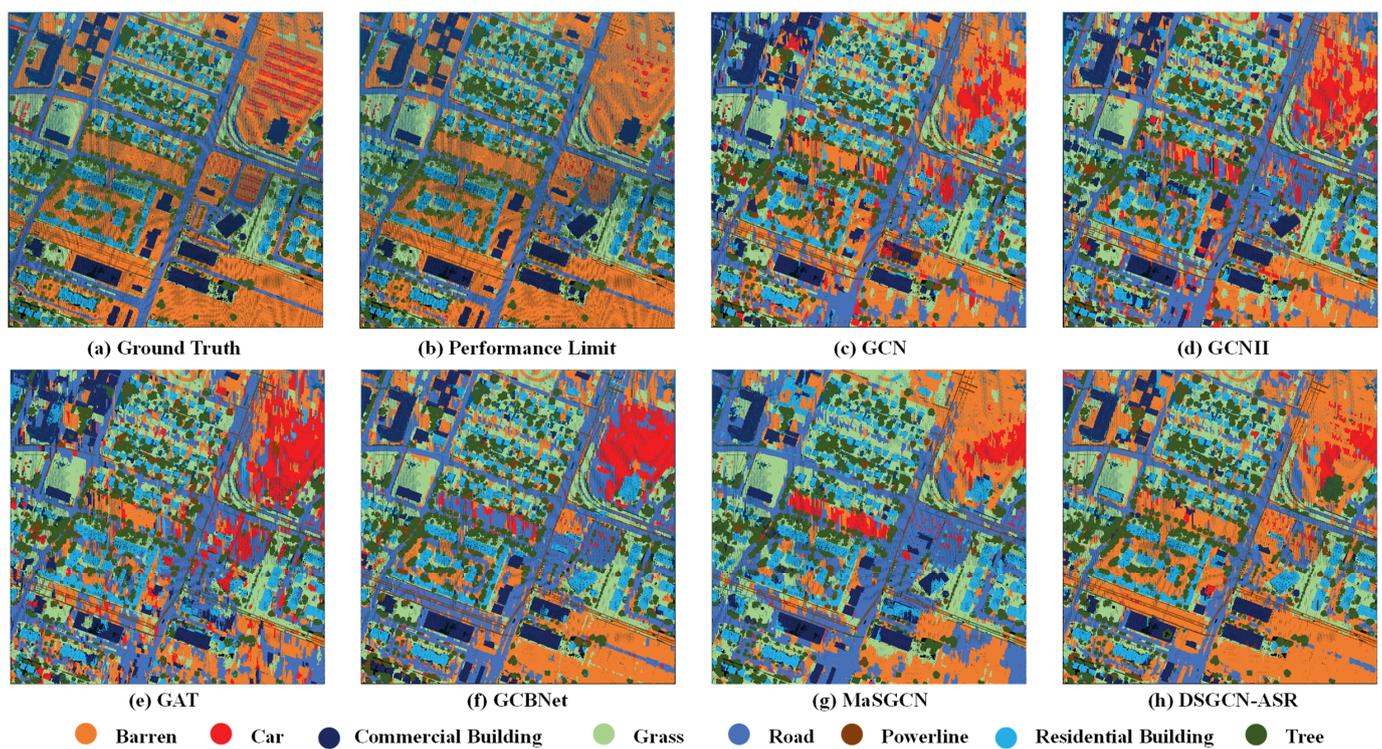


Figure 7. Visualization of classification results on the UH dataset: (a) visualization of the ground truth; (b) performance limit due to superpoint segmentation. Visualization of classification results of (c) GCN, (d) GCNII, (e) GAT, (f) GCBNet, (g) MaSGCN, and (h) DSGCN-ASR (ours).

3.2. Ablation Studies

Ablation studies were conducted to validate the effectiveness of the proposed joint-use scheme of spatial–spectral graphs. Different experimental groups were set up by controlling the graphs used in the backbone and residuals, which were used to analyze the respective contributions of the spatial and spectral graphs in the network and to validate the proposed DSGCN-ASR.

For each dataset, we conducted the following sets of experiments: (a) using the spatial graph in the backbone and the residuals; (b) using the spatial graph in the backbone and the spectral graph in the residuals; (c) using an equal-scale combined spatial–spectral graph (Equation (13)) in the backbone and residuals; (d) using the spectral graph in the backbone and the spatial graph in the residuals; and (e) using the spectral graph in the backbone and the residuals. The setup of experiments is shown in Table 5.

Table 5. Experimental setup for ablation studies.

Group	Backbone	Residuals
I	Spatial Graph	Spatial Graph
II *	Spatial Graph	Spectral Graph
III	Combined Graph	Combined Graph
IV	Spectral Graph	Spatial Graph
V	Spectral Graph	Spectral Graph

* In the proposed DSGCN-ASR, we use the spatial graph in the backbone and the spectral graph in the residual, as in II.

The overall evaluation metrics for the ablation results on the HT and UH datasets are shown in Table 6; the evaluation metrics for each class on the HT and UH datasets are shown in Table 7 and Table 8, respectively. The ablation results are visualized in Figure 8. In the experimental setup, the spatial graph contributed progressively less, and the spectral graph progressively dominated from groups I to V. The difficulty in achieving accurate classification performance using only one of the spatial and spectral graphs was noted based on the results for groups I and V. The results for groups I and V are consistent with our analysis in the Introduction, with spatial graphs tending to classify spatially adjacent points into the same class and spectral graphs being better at distinguishing spatially neighboring land cover.

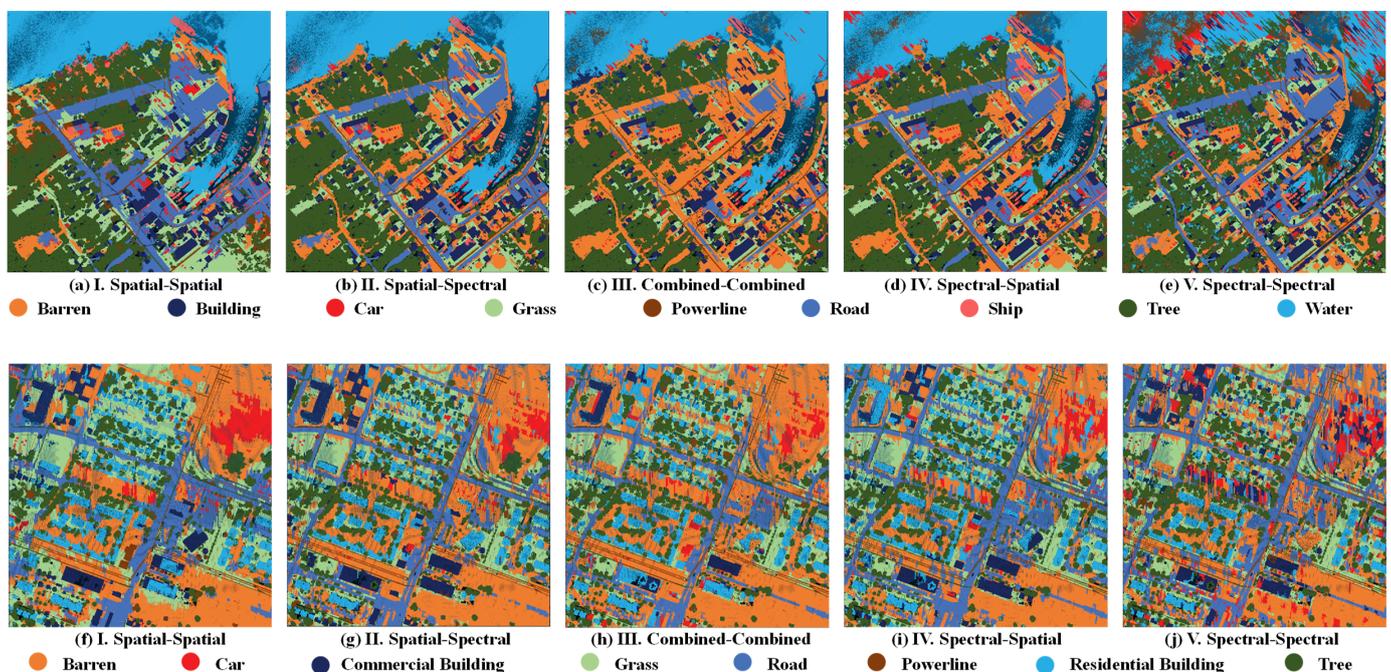


Figure 8. Visualization of ablation results: (a) group I, (b) group II, (c) group III, (d) group IV, and (e) group V on the HT dataset ; (f) group I, (g) group II, (h) group III, (i) group IV, and (j) group V on the UH dataset.

Table 6. Evaluation metrics (%) for ablation studies.

Dataset	Group	Setup	OA	Macro Precision	Macro Recall	Macro F Score	MIoU
HT	I	Spatial–Spatial	70.85	67.33	55.73	60.99	42.79
	II *	Spatial–Spectral	87.57	74.23	69.45	71.76	59.51
	III	Combined–Combined	83.04	67.97	67.35	67.66	52.14
	IV	Spectral–Spatial	87.09	76.70	66.81	71.41	57.67
	V	Spectral–Spectral	77.87	63.19	52.74	57.49	43.02
UH	I	Spatial–Spatial	68.50	67.23	55.90	61.04	43.50
	II *	Spatial–Spectral	78.20	73.03	65.41	69.01	54.02
	III	Combined–Combined	75.59	63.11	65.51	64.29	47.60
	IV	Spectral–Spatial	74.81	66.41	63.18	64.76	49.46
	V	Spectral–Spectral	68.94	63.83	55.87	59.58	44.61

* In the proposed DSGCN-ASR, we use the spatial graph in the backbone and the spectral graph in the residual, as in group II. Maximum values in the same metrics are marked in bold.

Table 7. Evaluation metrics (%) for ablation studies in each class on the HT dataset.

Group	Class	Barren	Building	Car	Grass	Powerline	Road	Ship	Tree	Water
I	Precision	18.05	68.02	46.27	79.71	53.95	76.38	80.08	84.60	98.93
	Recall	72.97	63.95	48.93	44.93	5.21	42.80	34.07	96.96	91.77
	F score	28.94	65.93	47.56	57.47	9.50	54.86	47.80	90.36	95.22
	IoU	16.92	49.17	31.20	40.32	4.98	37.80	31.40	82.42	90.87
II *	Precision	72.18	88.36	29.28	90.95	69.95	77.08	47.02	95.41	97.86
	Recall	78.40	78.74	38.73	86.34	33.12	64.81	48.64	99.05	97.24
	F score	75.16	83.27	33.35	88.59	44.96	70.41	47.81	97.19	97.55
	IoU	60.21	71.34	20.01	79.51	29.00	54.34	31.42	94.54	95.21
III	Precision	91.12	87.90	24.26	74.48	73.92	25.14	48.63	93.19	93.04
	Recall	61.04	74.38	27.80	96.36	14.50	68.74	70.33	99.23	93.75
	F score	73.11	80.58	25.91	84.02	24.24	36.82	57.50	96.12	93.39
	IoU	57.62	67.48	14.88	72.44	13.79	22.56	40.35	92.52	87.61
IV	Precision	78.32	84.55	39.76	88.41	82.76	80.76	54.18	93.08	88.48
	Recall	80.99	80.07	34.71	94.18	16.80	69.97	34.58	98.72	91.23
	F score	79.63	82.25	37.06	91.20	27.94	74.98	42.21	95.82	89.83
	IoU	66.16	69.85	22.75	83.83	16.24	59.97	26.75	91.98	81.54
V	Precision	61.76	75.02	14.23	91.46	68.60	78.73	39.44	86.66	52.80
	Recall	86.48	52.65	6.70	84.11	19.96	59.22	23.60	97.82	44.07
	F score	72.06	61.88	9.11	87.63	30.92	67.60	29.53	91.90	48.04
	IoU	56.32	44.80	4.77	77.98	18.29	51.06	17.32	85.02	31.62

* In the proposed DSGCN-ASR, we use the spatial graph in the backbone and the spectral graph in the residual, as in group II. Maximum values in the same metrics are marked in bold.

The experiments in group III showed that the equiscale combination of spatial and spectral graphs, to some extent, could increase the accuracy of classification and achieve relatively good metrics. However, this combination also inherits the drawbacks of both graphs, with a simultaneous lack of clear distinction at the boundaries and interference from the chaotic spectral information.

Numerically, group II, which is also used in the proposed DSGCN-ASR, achieved the best performance among the five groups of ablation investigations. Compared with the second-best group, group II was 0.48% ahead in OA, 2.1% ahead in macro recall, 0.35% ahead in macro F score, and 1.83% ahead in MIoU on the HT dataset. On the UH dataset, group II achieved a 2.60% OA lead, a 5.81% macro precision lead, a 4.25% macro F-score lead, and a 4.55% MIoU lead. Group IV, using the spectral graphs in the backbone and the spatial graphs in the residuals, also achieved good results, slightly outperforming group III

overall. This tangentially corroborated the superiority of distinguishing the use of spatial and spectral graphs as proposed in this study.

Table 8. Evaluation metrics (%) for ablation studies in each class on the UH dataset.

Group	Class	Barren	Car	Commercial	Grass	Road	Powerline	Residential	Tree
I	Precision	64.18	55.39	58.51	80.67	59.57	63.56	81.98	73.95
	Recall	77.81	28.15	52.84	62.45	60.15	24.33	48.89	92.55
	F score	70.34	37.33	55.53	70.40	59.86	35.19	61.25	82.21
	IoU	54.25	22.95	38.44	54.32	42.71	21.35	44.15	69.80
II *	Precision	75.78	33.89	67.28	83.02	72.70	79.49	83.82	88.28
	Recall	80.72	28.60	71.00	80.46	68.65	32.39	67.05	94.42
	F score	78.18	31.02	69.09	81.72	70.62	46.02	74.51	91.24
	IoU	64.17	18.36	52.77	69.09	54.58	29.89	59.37	83.90
III	Precision	80.37	26.77	24.25	73.37	58.11	65.09	85.75	91.21
	Recall	73.47	22.67	82.86	84.46	69.83	44.18	51.48	95.16
	F score	76.76	24.55	37.52	78.53	63.43	52.63	64.33	93.14
	IoU	62.29	13.99	23.09	64.64	46.45	35.71	47.42	87.16
IV	Precision	65.61	34.63	52.98	83.17	79.19	42.12	83.21	90.40
	Recall	82.20	25.31	70.11	78.63	58.59	46.62	50.57	93.45
	F score	72.97	29.25	60.35	80.84	67.35	44.25	62.90	91.90
	IoU	57.45	17.13	43.22	67.84	50.78	28.41	45.88	85.02
V	Precision	56.16	31.29	48.67	82.89	69.42	67.43	63.09	91.66
	Recall	77.00	10.61	28.64	80.31	53.34	41.64	62.49	92.90
	F score	64.95	15.85	36.06	81.58	60.33	51.48	62.79	92.27
	IoU	48.09	8.60	22.00	68.89	43.19	34.67	45.76	85.66

* In the proposed DSGCN-ASR, we use the spatial graph in the backbone and the spectral graph in the residual, as in group II. Maximum values in the same metrics are marked in bold.

Within the network architecture, the integration of information between spatial- and spectral-graph-based aggregation is primarily governed by a trainable adaptive weight. This allows groups II and IV to achieve an approximate integration of information. Group II employs the spatial graph in the backbone for convolution, effectively accessing the spatial distribution of relationships among land cover classes. In comparison, group IV uses the spectral graph for convolution, prioritizing the spectral similarities between land cover classes. However, this approach results in the inclusion of some irrelevant connections. This finding is indirectly supported by the observation of Figure 8d,i, where the visualization of the results reveals numerous scattered, misclassified points. This performance aligns with that of group V. The outcomes of the ablation studies provide evidence for the practicality and effectiveness of the proposed joint spatial–spectral use strategy employed by DSGCN-ASR.

3.3. Parametric Analysis

We then conducted experiments to analyze the impact of the α and β parameters on the classification performance. We specifically focused on these parameters while keeping all other settings constant. We set α to a fixed value of 0, 0.25, 0.5, 0.75, or 1 for comparison with the case of an adaptive α . We followed the same approach for β and performed five sets of experiments with values of 0, 0.25, 0.5, 0.75, and 1 for comparison with a decreasing β . The results of the parametric analysis experiment for α are visualized in Figure 9 and in Figure 10 for β . The evaluation metrics for the parametric analysis of α and β are shown in Table 9 and Table 10, respectively. To more intuitively show the impact of the parameters on the classification results, we also plotted histograms and line graphs, as shown in Figure 11.

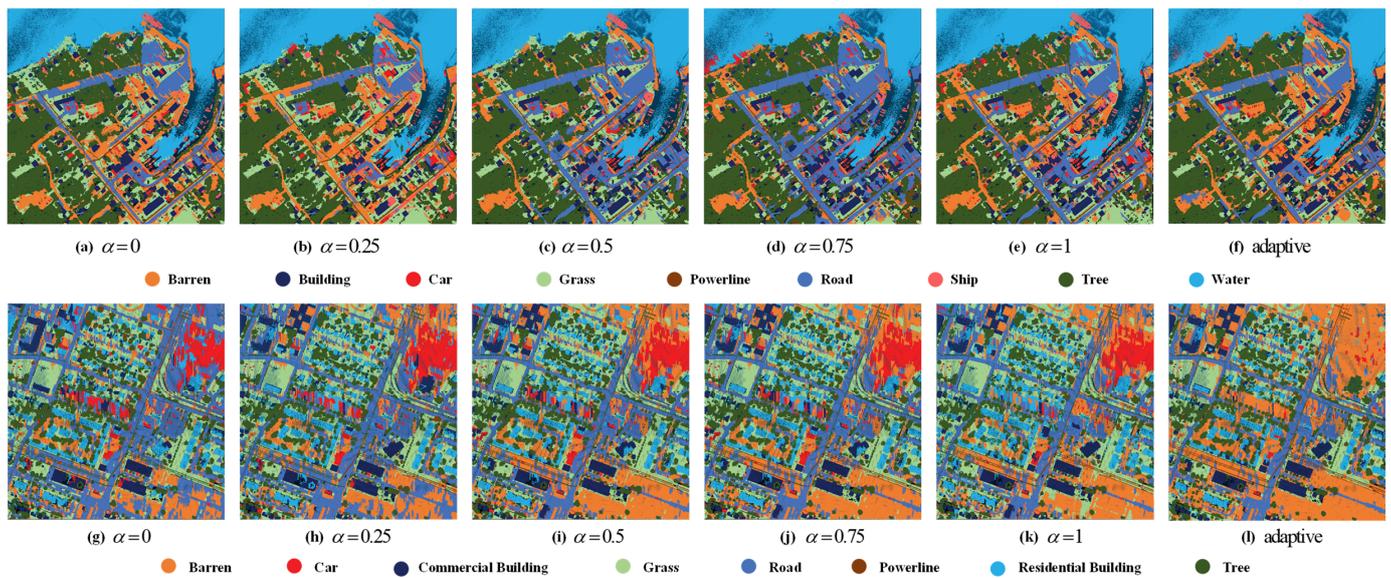


Figure 9. Visualization of the parametric analysis experiment for α when set to (a) 0, (b) 0.25, (c) 0.5, (d) 0.75, (e) 1, and (f) an adaptive value on the HT dataset and (g) 0, (h) 0.25, (i) 0.5, (j) 0.75, (k) 1, and (l) an adaptive value on the UH dataset.

Table 9. The evaluation metrics (%) for parametric analysis of α .

Dataset	α	OA	Macro Precision	Macro-Recall	Macro F Score	MIoU
HT	0	84.38	72.00	66.29	69.03	55.96
	0.25	79.68	65.89	59.85	62.73	48.24
	0.5	82.13	71.49	64.13	67.61	53.16
	0.75	78.59	71.19	61.48	65.98	50.42
	1	80.51	71.60	58.87	64.61	49.66
	Adaptive *	87.57	74.23	69.45	71.76	59.51
UH	0	61.68	62.52	53.34	57.56	40.42
	0.25	66.26	68.35	59.41	63.57	46.04
	0.5	70.86	70.47	59.05	64.26	47.89
	0.75	68.69	71.27	57.68	63.76	45.56
	1	67.92	69.20	56.81	62.39	44.94
	Adaptive *	78.20	73.03	65.41	69.01	54.02

* In the proposed DSGCN-ASR, we use an adaptive α value. Maximum values in the same metrics are marked in bold.

The α parameter plays a crucial role in controlling the weight of the spectral residuals in each layer of the network. As α increases, the model incorporates more spectral information, enhancing its ability to differentiate between different land cover classes in the immediate neighborhood. However, excessively large values of α can compromise the robustness of the model, leading to patchier misclassifications. These findings align with the conclusions drawn in the Introduction. The adaptive spectral residual strategy employed in our approach allows the model to autonomously adjust the acquisition weights of spectral information in each layer. As a result, the final classification performance is substantially superior to that achieved by other groups using a fixed α .

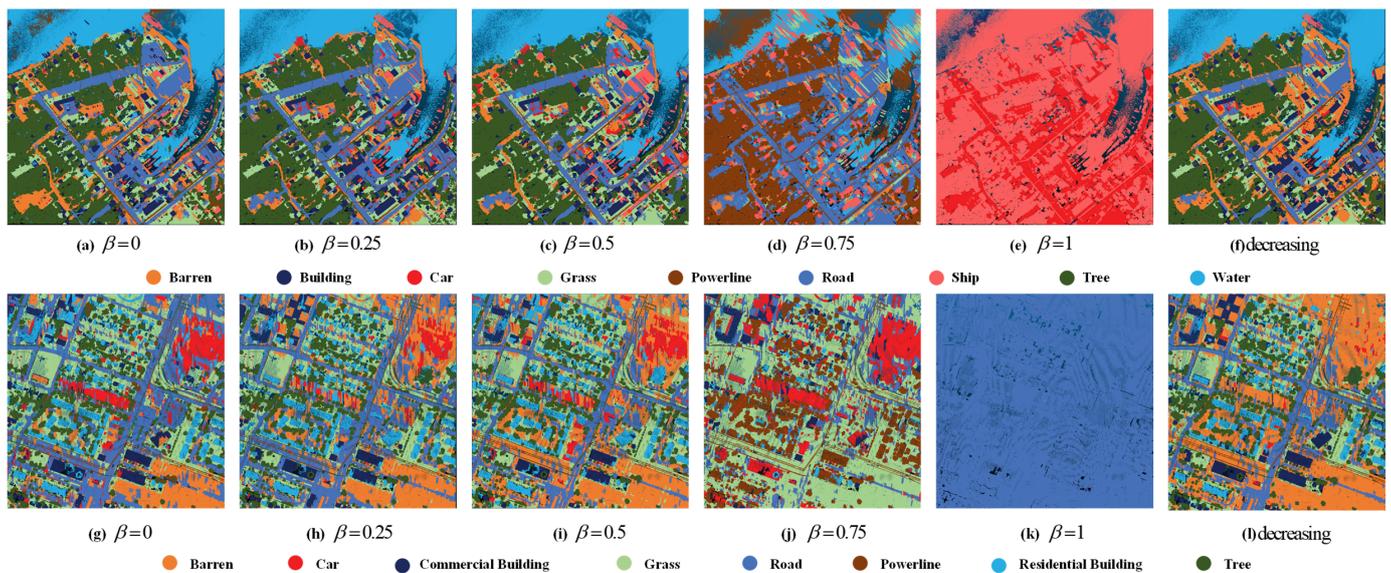


Figure 10. Visualization of the parametric analysis experiment for β when set to (a) 0, (b) 0.25, (c) 0.5, (d) 0.75, (e) 1, and (f) a decreasing value on the HT dataset and (g) 0 (h) 0.25, (i) 0.5, (j) 0.75, (k) 1, and (l) a decreasing value on the UH dataset.

Table 10. The evaluation metrics (%) for parametric analysis of β .

Dataset	β	OA	Macro Precision	Macro Recall	Macro F Score	MIoU
HT	0	80.94	68.58	59.37	63.64	47.76
	0.25	81.36	70.31	67.17	68.70	54.27
	0.5	79.68	70.64	59.71	64.72	49.12
	0.75	14.88	33.62	NAN	NAN	10.03
	1	1.05	13.18	NAN	NAN	0.25
	Decreasing *	87.57	74.23	69.45	71.76	59.51
UH	0	61.14	67.58	52.92	59.36	40.03
	0.25	70.37	67.70	62.01	64.73	48.71
	0.5	73.06	67.74	62.25	64.88	48.77
	0.75	23.34	41.63	NAN	NAN	12.10
	1	14.54	12.50	NAN	NAN	1.82
	Decreasing *	78.20	73.03	65.41	69.01	54.02

* In the proposed DSGCN-ASR, we use a decreasing β . Maximum values in the same metrics are marked in bold.

Deep graph convolutional networks often oversmooth deep features. To tackle this issue, we introduced the β parameter based on the concept of identity mapping [33]. The value of β determines the proportion of the hidden layer features obtained from the previous convolution in the model. The experimental results indicate that as β increases, the model becomes more susceptible to the oversmoothing of deep features. The model fails in the two groups where β exceeds 0.5. Our approach employs the strategy of a decreasing β with an increasing number of model layers, which was validated in a previous study [33]. Once again, this strategy proves effective in mitigating the problem of oversmoothing in our approach.

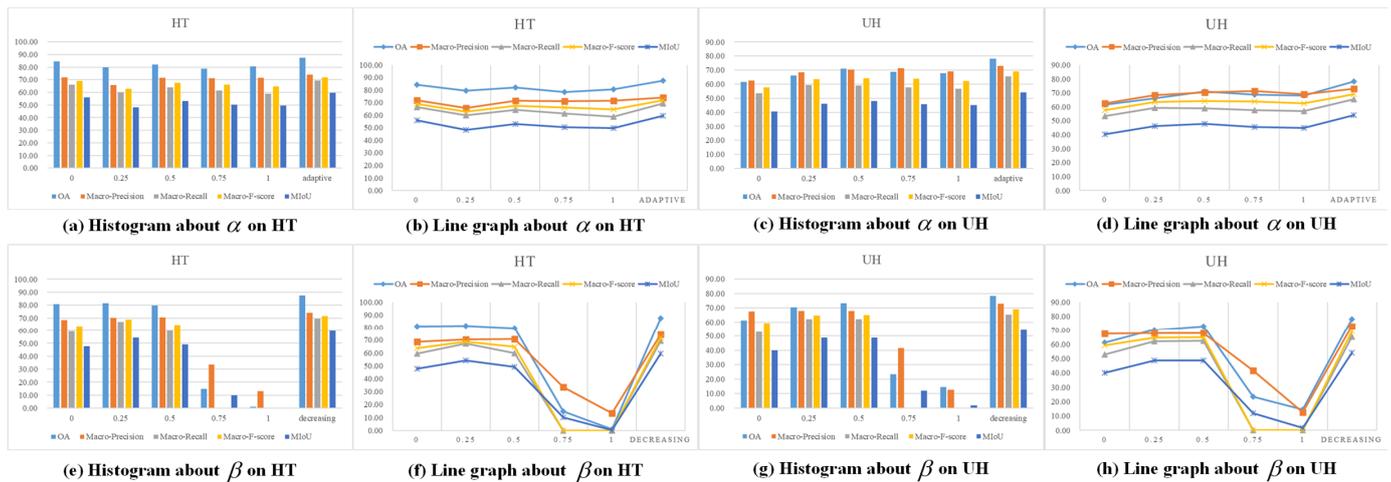


Figure 11. Histograms and line graphs of results of parametric analysis experiment. (a) Histogram of α on the HT dataset. (b) Line graph of α on the HT dataset. (c) Histogram of α on the UH dataset. (d) Line graph of α on the UH dataset. (e) Histogram of β on the HT dataset. (f) Line graph of β on the HT dataset. (g) Histogram of β on the UH dataset. (h) Line graph of β on the UH dataset.

The histograms and line graphs provide a clearer visualization of the impact of the α and β parameters on the classification performance of the model. These visual representations highlight the advantage of our developed strategy in the parametric analysis experiments. The results demonstrate the effectiveness of our approach in improving classification performance.

4. Conclusions

This study focused on the classification of multispectral point cloud data, and we developed a novel method called DSGCN-ASR. In contrast to existing methods, DSGCN-ASR adopts a differentiated treatment of spatial and spectral graphs, effectively leveraging their respective advantages to enhance classification performance. By preserving the robustness of the spatial graph for extraction of land cover relationships and applying the discriminatory ability of the spectral graph to distinguish neighboring land cover classes, DSGCN-ASR achieves superior classification performance. Experimental validation using real-world multispectral point cloud datasets and comparisons with state-of-the-art graph-based methods demonstrated the efficacy of DSGCN-ASR in effectively leveraging spatial–spectral information. This study provides valuable insights into the joint use of spatial–spectral information in multispectral point clouds, contributing to accurate mapping and fine-grained land cover classification. Further exploration of this method holds promise for the advancement of the field of elaborate mapping in navigation systems.

Author Contributions: Conceptualization, Q.W., Z.Z. and T.S.; methodology, Q.W. and Z.Z.; software, Z.Z.; validation, Z.Z.; formal analysis, Z.Z., X.C. and J.S.; resources, Q.W., J.S. and T.S.; data curation, Z.Z. and X.C.; writing—original draft preparation, Z.Z.; writing—review and editing, Q.W. and J.S.; visualization, Z.Z.; supervision, T.S.; project administration, Z.W.; funding acquisition, Z.W., Q.W., and T.S. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded in part by the Youth Project of the National Natural Science Foundation of China under grant 62201237, the Yunnan Fundamental Research Projects under grants 202101BE070001-008 and 202301AV070003, the Youth Project of the Xingdian Talent Support Plan of Yunnan Province under grant KKRD202203068, and the Major Science and Technology Projects in Yunnan Province under grants 202202AD080013 and 202302AG050009.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The HT dataset used in this study was obtained through an online application at <https://www.isprs.org/news/newsletter/2017-03/index.html> (accessed on March 2017). The UH dataset was obtained from <http://www.classic.grss-ieee.org/community/technical-committees/data-fusion/2018-ieee-grss-data-fusion-contest/> (accessed on 16 February 2017), and the data were manually labeled.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Qi, C.R.; Su, H.; Mo, K.; Guibas, L.J. Pointnet: Deep learning on point sets for 3d classification and segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 652–660.
2. Jing, Z.; Guan, H.; Zhao, P.; Li, D.; Yu, Y.; Zang, Y.; Wang, H.; Li, J. Multispectral LiDAR point cloud classification using SE-PointNet++. *Remote Sens.* **2021**, *13*, 2516. [[CrossRef](#)]
3. Chen, Y.; Liu, G.; Xu, Y.; Pan, P.; Xing, Y. PointNet++ network architecture with individual point level and global features on centroid for ALS point cloud classification. *Remote Sens.* **2021**, *13*, 472. [[CrossRef](#)]
4. Zhang, J.; Lin, X.; Ning, X. SVM-based classification of segmented airborne LiDAR point clouds in urban areas. *Remote Sens.* **2013**, *5*, 3749–3775. [[CrossRef](#)]
5. Lodha, S.K.; Fitzpatrick, D.M.; Helmbold, D.P. Aerial lidar data classification using adaboost. In Proceedings of the Sixth International Conference on 3-D Digital Imaging and Modeling (3DIM 2007), Montreal, QC, Canada, 21–23 August 2007; pp. 435–442.
6. Chehata, N.; Guo, L.; Mallet, C. Airborne lidar feature selection for urban classification using random forests. In Proceedings of the Laserscanning, Paris, France, 1–2 September 2009.
7. Munoz, D.; Bagnell, J.A.; Vandapel, N.; Hebert, M. Contextual classification with functional max-margin markov networks. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 975–982.
8. Niemeyer, J.; Wegner, J.D.; Mallet, C.; Rottensteiner, F.; Soergel, U. Conditional random fields for urban scene classification with full waveform LiDAR data. In Proceedings of the Photogrammetric Image Analysis: ISPRS Conference, PIA 2011, Munich, Germany, 5–7 October 2011; Springer: Berlin/Heidelberg, Germany, 2011; pp. 233–244.
9. Yu, Y.; Jiang, T.; Gao, J.; Guan, H.; Li, D.; Gao, S.; Tang, E.; Wang, W.; Tang, P.; Li, J. CapViT: Cross-context capsule vision transformers for land cover classification with airborne multispectral LiDAR data. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *111*, 102837. [[CrossRef](#)]
10. Yu, Y.; Liu, C.; Guan, H.; Wang, L.; Gao, S.; Zhang, H.; Zhang, Y.; Li, J. Land cover classification of multispectral lidar data with an efficient self-attention capsule network. *IEEE Geosci. Remote Sens. Lett.* **2021**, *19*, 6501505. [[CrossRef](#)]
11. Wang, Y.; Sun, Y.; Liu, Z.; Sarma, S.E.; Bronstein, M.M.; Solomon, J.M. Dynamic graph cnn for learning on point clouds. *Acm Trans. Graph.* **2019**, *38*, 146. [[CrossRef](#)]
12. Liu, Y.; Fan, B.; Xiang, S.; Pan, C. Relation-shape convolutional neural network for point cloud analysis. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 8895–8904.
13. Bakula, K.; Kupidura, P.; Jełowicki, Ł. Testing of land cover classification from multispectral airborne laser scanning data. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *41*, 161–169. [[CrossRef](#)]
14. Morsy, S.; Shaker, A.; El-Rabbany, A. Multispectral LiDAR data for land cover classification of urban areas. *Sensors* **2017**, *17*, 958. [[CrossRef](#)]
15. Sun, J.; Shi, S.; Chen, B.; Du, L.; Yang, J.; Gong, W. Combined application of 3D spectral features from multispectral LiDAR for classification. In Proceedings of the 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA, 23–28 July 2017; pp. 5264–5267.
16. Teo, T.A.; Wu, H.M. Analysis of land cover classification using multi-wavelength LiDAR system. *Appl. Sci.* **2017**, *7*, 663. [[CrossRef](#)]
17. Matikainen, L.; Hyyppä, J.; Litkey, P. Multispectral airborne laser scanning for automated map updating. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *41*, 323–330. [[CrossRef](#)]
18. Pan, S.; Guan, H.; Chen, Y.; Yu, Y.; Gonçalves, W.N.; Junior, J.M.; Li, J. Land-cover classification of multispectral LiDAR data using CNN with optimized hyper-parameters. *Isprs J. Photogramm. Remote Sens.* **2020**, *166*, 241–254. [[CrossRef](#)]
19. Yu, Y.; Guan, H.; Li, D.; Gu, T.; Wang, L.; Ma, L.; Li, J. A hybrid capsule network for land cover classification using multispectral LiDAR data. *IEEE Geosci. Remote Sens. Lett.* **2019**, *17*, 1263–1267. [[CrossRef](#)]
20. Hu, Q.; Yang, B.; Xie, L.; Rosa, S.; Guo, Y.; Wang, Z.; Trigoni, N.; Markham, A. Randa-net: Efficient semantic segmentation of large-scale point clouds. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 11108–11117.
21. Wang, Q.; Gu, Y. A discriminative tensor representation model for feature extraction and classification of multispectral LiDAR data. *IEEE Trans. Geosci. Remote Sens.* **2019**, *58*, 1568–1586. [[CrossRef](#)]
22. Niemeyer, J.; Rottensteiner, F.; Soergel, U.; Heipke, C. Hierarchical higher order crf for the classification of airborne lidar point clouds in urban areas. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *41*, 655–662. [[CrossRef](#)]

23. Wichmann, V.; Bremer, M.; Lindenberger, J.; Rutzinger, M.; Georges, C.; Petrini-Monteferri, F. Evaluating the potential of multispectral airborne lidar for topographic mapping and land cover classification. *Isprs Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2015**, *2*, 113–119. [[CrossRef](#)]
24. Anguelov, D.; Taskarf, B.; Chatalbashev, V.; Koller, D.; Gupta, D.; Heitz, G.; Ng, A. Discriminative learning of markov random fields for segmentation of 3d scan data. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 20–25 June 2005; Volume 2, pp. 169–176.
25. Wu, Z.; Pan, S.; Chen, F.; Long, G.; Zhang, C.; Philip, S.Y. A comprehensive survey on graph neural networks. *IEEE Trans. Neural Netw. Learn. Syst.* **2020**, *32*, 4–24. [[CrossRef](#)]
26. Wang, L.; Huang, Y.; Hou, Y.; Zhang, S.; Shan, J. Graph attention convolution for point cloud semantic segmentation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 10296–10305.
27. Zhao, P.; Guan, H.; Li, D.; Yu, Y.; Wang, H.; Gao, K.; Junior, J.M.; Li, J. Airborne multispectral LiDAR point cloud classification with a feature reasoning-based graph convolution network. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *105*, 102634. [[CrossRef](#)]
28. Wen, C.; Li, X.; Yao, X.; Peng, L.; Chi, T. Airborne LiDAR point cloud classification with global-local graph attention convolution neural network. *Isprs J. Photogramm. Remote Sens.* **2021**, *173*, 181–194. [[CrossRef](#)]
29. Wang, Q.; Gu, Y.; Yang, M.; Wang, C. Multi-attribute smooth graph convolutional network for multispectral points classification. *Sci. China Technol. Sci.* **2021**, *64*, 2509–2522. [[CrossRef](#)]
30. Kipf, T.N.; Welling, M. Semi-supervised classification with graph convolutional networks. In Proceedings of the International Conference on Learning Representations, Toulon, France, 24–26 April 2017; pp. 1–14.
31. Velickovic, P.; Cucurull, G.; Casanova, A.; Romero, A.; Lio, P.; Bengio, Y. Graph attention networks. In Proceedings of the International Conference on Learning Representations, Vancouver, BC, Canada, 30 April–3 May 2018.
32. Zhang, T.; Wang, X.; Xu, X.; Chen, C.P. GCB-Net: Graph convolutional broad network and its application in emotion recognition. *IEEE Trans. Affect. Comput.* **2019**, *13*, 379–388. [[CrossRef](#)]
33. Chen, M.; Wei, Z.; Huang, Z.; Ding, B.; Li, Y. Simple and deep graph convolutional networks. In Proceedings of the International Conference on Machine Learning, Virtual, 13–18 July 2020; pp. 1725–1735.
34. Lin, Y.; Wang, C.; Zhai, D.; Li, W.; Li, J. Toward better boundary preserved supervoxel segmentation for 3D point clouds. *Isprs J. Photogramm. Remote Sens.* **2018**, *143*, 39–47. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.