Graph Convolutional Networks-Based Super-Resolution Land Cover Mapping

Xining Zhang, Yong Ge, *Senior Member, IEEE*, Feng Ling, Jin Chen, Yuehong Chen, and Yuanxin Jia

***Abstract*—Super-resolution mapping (SRM) is an effective technology to solve the problem of mixed pixels because it can be used to generate fine-resolution land cover maps from coarse-resolution remote sensing images. Current methods based on deep neural networks have been successfully applied to SRM, as they can learn complex spatial patterns from training data. However, they lack the ability to learn structural information between adjacent land cover classes, which is vital in the reconstruction of spatial distribution. In this article, an SRM method based on graph convolutional networks (GCNs), named $\mathrm{SRM_{GCN}}$, is proposed to improve SRM results by capturing structure information on the graph. In $\mathrm{SRM_{GCN}}$, a supervised inductive learning strategy with mini-graphs as input is considered, which is an extension of the GCN framework. Furthermore, two operations are designed in terms of adjacency matrix construction and an information propagation rule to help reconstruct detailed information of geographical objects. Experiments on three datasets with different spatial resolutions demonstrate the qualitative and quantitative superiority of $\mathrm{SRM_{GCN}}$ over three other popular SRM methods.**

***Index Terms*—Deep neural networks (DNNs), graph convolutional networks (GCNs), land cover, subpixel, super-resolution mapping (SRM).**

I. Introduction

REMOTE sensing images are considered the most important input to produce land cover maps. However, mixed pixels, which refer to pixels that contain more than one land cover class, constitute a significant obstacle to accurate classification [1]–[3]. Due to the complex distribution patterns of geographical objects and the limitation of the sensor's spatial resolution [1], [4], mixed pixels are common in low-, medium-, and even high-resolution remote sensing images. Hard classification simply assigns the most likely land cover class to each mixed pixel; however, this process may inevitably bring tremendous information loss [5]. Soft classification provides membership grades for each class which can also be considered as the area proportions of different land cover classes within coarse-resolution pixels in the form of fraction images [5]–[7]; however, it is unable to determine the correct position of land cover classes inside the mixed pixel. Therefore, super-resolution mapping (SRM) technology was developed to address this problem [6]–[9].

In the past 20 years, various SRM technologies have been developed. These technologies can be roughly classified into two groups, according to the model used to describe the spatial pattern of land cover classes. One group uses explicit prior models to analyze the spatial distribution of land cover classes through clear formulations. The maximal spatial dependence model based on the first law of geography [10] is used most fundamentally and commonly [8]. The spatial position of each class in mixed pixels is determined by maximizing the spatial dependence at the subpixel scale, the subpixel/pixel scale, the object scale [11], or multiple scales [12]. Typical algorithms include the spatial attraction model (SAM) [13], the vectorial boundary model (VBSPM) [14], the maximum *a posteriori* (MAP) based methods [15], etc. In recent years, some optimization algorithms have been developed to acquire optimal SRM results. Ma *et al*. [16] proposed multiobjective subpixel land-cover mapping framework to resolve the regularization parameter determination problem in the MAP-based SPM methods; based on this, Song *et al*. [17] introduced multiple shifted images to the multiobjective model to produce more accurate mapping results. However, the aforementioned methods need to take the fraction images generated by soft classification as input, and cannot fully exploit the spectral information of the original image. Therefore, some SRM approaches incorporating the original spectral information and the concept of spatial dependence have been proposed [18]. Although maximal spatial dependence is a widely used basic assumption of SRM, the spatial heterogeneity distribution characteristics of geographical objects are also universal [19]; this makes the maximal spatial dependence model unsuitable for some situations, especially for intricate and fragmented patterns in an extreme situation [20].

The other group uses learning-based methods to address SRM issues without explicit prior models, but with implicit relationships between coarse-resolution images and fine-resolution land cover maps learned from sample pairs. Support vector regression

Manuscript received April 3, 2021; revised June 20, 2021; accepted July 22, 2021. Date of publication July 27, 2021; date of current version August 12, 2021. This work was supported by the National Natural Science Foundation for Distinguished Young Scholars of China under Grant 41725006. *(Corresponding authors: Yong Ge; Jin Chen.)*

Xining Zhang and Yong Ge are with the State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China, and also with the University of Chinese Academy of Sciences, Beijing 100049, China (e-mail: zhangxn@lreis.ac.cn; gey@lreis.ac.cn).

Feng Ling is with the Key Laboratory of Monitoring and Estimate for Environment and Disaster of Hubei Province, Innovation Academy for Precision Measurement Science and Technology, Chinese Academy of Sciences, Wuhan 430077, China (e-mail: lingf@whigg.ac.cn).

Jin Chen is with the State Key Laboratory of Earth Surface Processes and Resource Ecology, Beijing Normal University, Beijing 100875, China (e-mail: chenjin@bnu.edu.cn).

Yuehong Chen is with the College of Hydrology and Water Resources, Hohai University, Nanjing 210098, China (e-mail: chenyh@lreis.ac.cn).

Yuanxin Jia is with the Academy of Forest Inventory and Planning, National Forestry and Grassland Administration, Beijing 100714, China (e-mail: jiayx@lreis.ac.cn).

Digital Object Identifier 10.1109/JSTARS.2021.3100400

[21] and back-propagation neural networks [22], [23] have been proposed to achieve this goal. Both of them are dedicated to describing the relationship between fractions in the local window and the spatial distribution of subpixels in the central coarse pixel, but they have difficulty learning elusive non-linear hidden representation. Recently, methods based on deep neural networks (DNNs), such as convolution neural networks (CNNs) [24]–[26] and generative adversarial networks [27] for SRM have gradually emerged. Some SRM methods based on DNNs comprise two steps: fraction image super-resolution and land cover allocation for subpixels [28], [29], and the successful experience of DNNs in image super-resolution can be used as references for the first step of SRM. He *et al*. [30] summarized a unified deep learning framework for SRM, into which the majority of the state-of-the-art networks could be easily embedded. DNNs can extract complicated feature representations from low to high levels by composing multiple simple but nonlinear modules [31], which have considerable advantages in modeling the complex relationships between coarse-resolution images and fine-resolution land cover categorical maps [24], [32]. Although compared with other approaches, DNNs-based methods have shown superiority in SRM, it is still a challenge to recover the details of SRM results with existing learning modes. Moreover, most current SRM methods based on DNNs choose to design deeper and more complex network structures with many parameters, ignoring the potential structural relationship between adjacent land cover classes; this hinders the representation capability of networks.

In recent years, owing to good performance in flexibly extracting structural feature representations from complicated graph data, graph convolutional networks (GCNs) [33] have attracted more attention. Although GCNs initially focused on applications in which the data were represented typically in the form of graphs, including social networks, traffic networks, citation networks, and so on, the important roles of GCNs in the field of image processing have been explored and proven lately. The research on image processing revealed the strong performance of GCNs in image recognition [34], [35], image semantic segmentation [36], and hyperspectral image classification [37]–[40], even if images have long been treated as regular grids in Euclidean space rather than graph structure data. It was found that in remote sensing image classification, local spatial structure information between adjacent land covers can be encoded by GCNs to flexibly preserve the class boundaries [37], which can also provide new perspectives for better addressing the SRM problem.

In this article, an SRM approach based on GCN is proposed, called $\mathrm{SRM_{GCN}}$. This approach considers an enhanced spatial neighborhood relation by building graph structures based on spatial distance and class-fractional similarity; hence, the inherent relationship between adjacent land covers can be captured by utilizing the learned model. To the best of our knowledge, this is the first time that GCN has been applied to solving SRM problems. To exhaustively exploit the input features and reconstruct detailed information of geographical objects, we have improved the GCN [33] in two aspects. First, in the construction of the adjacency matrix, dynamic edge weight adjustment is realized by adding a layer-wise trainable parameter. Second, in the rule of information propagation, we introduce an input feature transformation matrix into each layer to fuse the relatively shallow features. On top of this, we extend the original GCN proposed for semisupervised node classification to a generalizable classifier with a supervised inductive learning strategy in which the mini-graphs are constructed to train the GCN. In this way, GCN can be used to solve node classification where unlabeled data are inaccessible during the training stage.

The remainder of this article is organized as follows. Section II elaborates on the proposed $\mathrm{SRM_{GCN}}$. Section III presents experimental results of three datasets with different spatial resolutions using the proposed method in comparison with three popular SRM methods. Sections IV and V state discussions and conclusions, respectively.

II. Methodology

A. Basic Theory of SRM

The aim of SRM is to determine the appropriate spatial distribution of land cover classes within coarse mixed pixels by taking the fraction images as input [7], [8]. In the real situation, the fraction images are obtained through soft classification. However, the process of soft classification inevitably contains some errors and uncertainties. Therefore, the experimental part of this article uses synthetic fraction images to verify the effect of the proposed SRM method without soft classification errors. For a given fine-resolution land cover map, synthetic coarse fraction images can be simulated by averaging the pixel values of the original fine-resolution map contained within each coarse pixel with a certain scale factor [24], [27]. The averaging filter in this process can be regarded as an error free soft classifier that does not require training. During the SRM process, each coarse pixel is divided into small subpixels and each subpixel is regarded as pure so that only one of the classes is assigned to it. As a result, a finer resolution representation of the land cover map can be attained.

To be specific, let C be the number of land cover classes and suppose that fraction images with spatial resolution r have been estimated by a soft classifier. Assuming fraction images cover n pixels, then the coarse fraction images can be represented as $X = \{x_{i,c} \,|\, i = 1, 2, \ldots, n; c = 1, 2, \ldots, C\}$. Given a scale factor S, a fine spatial-resolution land cover map with spatial resolution R $(R = r/S)$ can be obtained through SRM by decomposing each pixel into $S \times S$ subpixels.

The fine land cover map can be represented as $Y = \{y_{j,c} \,|\, j = 1, 2, \ldots, N; c = 1, 2, \ldots, C\}$ and $N = n \times S^2$, where $y_{j,c} \in \{0,\ 1\}$ is defined in (1); this means that each subpixel is allocated a value 1 or 0 for each class. Meanwhile, each subpixel should be allocated only one land cover class, meaning that the condition $\sum_{c=1}^{C} y_{j,c} = 1$ should be satisfied for all $j = 1, 2, \ldots, N$ to ensure the uniqueness of subpixel class

$$y_{j,c} = \begin{cases} 1, & \text{subpixel } j \text{ belongs to class } c \\ 0, & \text{otherwise} \end{cases}. \tag{1}$$

B. Definition of GCN

Based on spectral theory, GCN [33] is a multilayer neural network architecture that directly operates on arbitrary graph structure data and aims to generate node representations via

a layer-wise propagation rule. The aggregation of feature information is conducted in each graph convolutional layer to extract high-level features from neighbors to their center nodes. Formally, let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be an undirected graph with N nodes $v_i \in \mathcal{V}$, edges $(v_i, v_j) \in \mathcal{E}$. An adjacency matrix $\boldsymbol{A}$ (weighted or unweighted) of $\mathcal{G}$ indicates whether each pair of nodes is connected and contains the structure information of the graph.

To acquire the hidden representation of nodes, spectral convolutions on a graph are defined by multiplying a graph signal with C input channels $\boldsymbol{x} \in \mathbb{R}^{N \times C}$ with a spectral filter $g_\theta(\boldsymbol{\Lambda}) = \mathrm{diag}(\boldsymbol{\theta})$ in the Fourier domain, as (2) expresses

$$g_\theta \star \boldsymbol{x} = \boldsymbol{U} g_\theta(\boldsymbol{\Lambda}) \boldsymbol{U}^T \boldsymbol{x} \tag{2}$$

where $\boldsymbol{U}$ is the matrix of eigenvectors of the symmetric normalized Laplacian matrix $\mathcal{L} = \boldsymbol{I}_N - \boldsymbol{D}^{-1/2} \boldsymbol{A} \boldsymbol{D}^{-1/2} = \boldsymbol{U} \boldsymbol{\Lambda} \boldsymbol{U}^T$. Here, $\boldsymbol{I}_N$ is an identity matrix. $\boldsymbol{D}$ is a diagonal degree matrix of the graph with element $\boldsymbol{D}_{ii} = \sum_j \boldsymbol{A}_{ij}$. $\boldsymbol{\Lambda}$ is a diagonal matrix whose elements are the eigenvalues of $\mathcal{L}$. In (2), $\boldsymbol{U}^T \boldsymbol{x}$ represents the graph Fourier transform of $\boldsymbol{x}$. $g_\theta(\boldsymbol{\Lambda})$ is considered a function of the eigenvalues of $\mathcal{L}$.

To reduce the computational costs and realize a local connection [33], [41], [42], Hammond *et al*. [43] established an approximating formula about the spectral filter $g_\theta(\boldsymbol{\Lambda})$ with Chebyshev polynomials $T_k(\cdot)$ up to the order K^{th} to circumvent computing the eigenvectors of the Laplacian matrix $\mathcal{L}$. Then the convolution of a graph signal $\boldsymbol{x}$ with a spectral filter $g_{\theta'}$ can be defined as

$$g_{\theta'} \star \boldsymbol{x} \approx \sum_{k=0}^{K} \boldsymbol{\theta}_k{}' T_k\left(\tilde{\mathcal{L}}\right) \boldsymbol{x} \tag{3}$$

where $\tilde{\mathcal{L}} = 2/(\lambda_{\max})\mathcal{L} - \boldsymbol{I}_N$ with $\lambda_{\max}$ being the largest eigenvalue of $\mathcal{L}$. $\boldsymbol{\theta}_k{}' \in \mathbb{R}^K$ is a vector of Chebyshev coefficients. This formulation is K-localized since it only depends on the nodes in the K^{th}-order neighborhoods of the central node.

Considering that a linear function of spectral convolutions on a graph can improve modeling capacity on a number of domains, Kipf and Welling [33] limited $K = 1$ to simplify this model and further approximated $\lambda_{\max}$ of $\mathcal{L}$ by 2. By doing so, (3) can be rewritten as

$$g_{\theta'} * \boldsymbol{x} \approx \theta\left(\boldsymbol{I}_N + \boldsymbol{D}^{-\frac{1}{2}} \boldsymbol{A} \boldsymbol{D}^{-\frac{1}{2}}\right) \boldsymbol{x} \tag{4}$$

by allowing the free parameters satisfy $\theta = \theta_0{}' = -\theta_1{}'$ to address overfitting and minimize the number of operations per layer. Kipf and Welling further introduced a renormalization trick and applied it to the convolution matrix to enhance stability during network training: $\boldsymbol{I}_N + \boldsymbol{D}^{-1/2} \boldsymbol{A} \boldsymbol{D}^{-1/2} \rightarrow \tilde{\boldsymbol{D}}^{-1/2} \tilde{\boldsymbol{A}} \tilde{\boldsymbol{D}}^{-1/2}$, with $\tilde{\boldsymbol{A}} = \boldsymbol{A} + \boldsymbol{I}_N$ and $\tilde{\boldsymbol{D}}_{ij} = \sum_j \tilde{\boldsymbol{A}}_{ij}$. The $\tilde{\boldsymbol{A}}$ means an adjacency matrix with added self-loops. Therefore, a layer-wise propagation rule in GCN is as follows:

$$\boldsymbol{H}^{(l)} = \sigma\left(\tilde{\boldsymbol{D}}^{-\frac{1}{2}} \tilde{\boldsymbol{A}} \tilde{\boldsymbol{D}}^{-\frac{1}{2}} \boldsymbol{H}^{(l-1)} \boldsymbol{W}^{(l)} + \boldsymbol{b}^{(l)}\right) \tag{5}$$

where $\boldsymbol{H}^{(l)} (l \geq 1)$ denotes the output in the l^{th} graph convolutional layer; $\sigma(\cdot)$ represents a nonlinear activation function, such as the rectified linear unit (ReLU) and Softmax used in this article; $\boldsymbol{W}^{(l)}$ and $\boldsymbol{b}^{(l)}$ denote the trainable weight matrix of filter parameters and the bias vector, respectively, included by the l^{th} layer.

C. The Proposed SRM Method

The framework is presented in Fig. 1. When input fraction images are given, they are preprocessed by interpolating and splitting to generate mini-graphs. In the training stage, the proposed GCN is designed to train labeled data regarding subpixels as nodes by minimizing a certain loss function. In the testing stage, unlabeled data are fed into the trained model to produce final fine-resolution land cover maps through two class-allocation algorithms. The three critical aspects are detailed below.

1) Mini-graphs Generation: In studies to date, the applications of GCN in remote sensing image classification are mostly based on feeding all samples in one big graph into the network [37]–[39]; two main problems arise from this. First, the massive nodes and edges of all samples will slow calculation and run out of memory. Second, it is impossible to build a generalizable classifier to predict unknown input through inductive learning.

To solve these problems, a training strategy based on mini-graphs is adopted in this article. As shown in Fig. 2, on the one hand, input coarse fraction images are first interpolated to nominal fine resolution images with a scale factor of S. Here, this operation is realized by the nearest neighbor interpolation. Then the nominal fine resolution images are used to extract a series of patches by a certain size of sliding window with a suitable sampling stride. For each patch, a mini-graph is constructed with each subpixel in the patch taken as a node; the graph node features are initialized with class fractions of subpixels after interpolation, which is locally expressed in Fig. 2 (a). Each mini-graph is a spatial distance-based structure graph in which each node as center node connects to its first-order neighborhoods and Fig. 2 (b) shows additional details. On the other hand, the reference map is extracted from the original labeled classification map using the same size of sliding window and sampling stride; it then forms a sample pair with the corresponding mini-graph. In this way, the SRM is transformed into a graph node classification problem at subpixel scale. Subsequently, these sample pairs based on mini-graphs are used for training and predicting in batches.

2) Proposed SRM Network: The proposed SRM model ($\mathrm{SRM_{GCN}}$) adopts the GCN as the backbone. The proposed model improves the construction of the adjacency matrix and the rule of information propagation in each graph convolutional layer to obtain node attribute information in a more detailed and comprehensive manner.

In the first place, the GCN is dependent on the adjacency matrix $\boldsymbol{A}$ to propagate information between nodes; hence, how to build $\boldsymbol{A}$ is critical for the GCN [44]. Here, we select the similarity distance-based weighted $\boldsymbol{A}$. Each element of it, also known as the weight of the edge between each pair of nodes, can be computed by using the heat kernel function [40]. However, in a graph-based SRM problem with subpixels as nodes, the similarity between nodes will change after aggregation of layer-wise information. There is a high computational cost associated with assigning each edge a dynamic adjustment parameter or recalculating the similarity distances of each layer. Thus, based

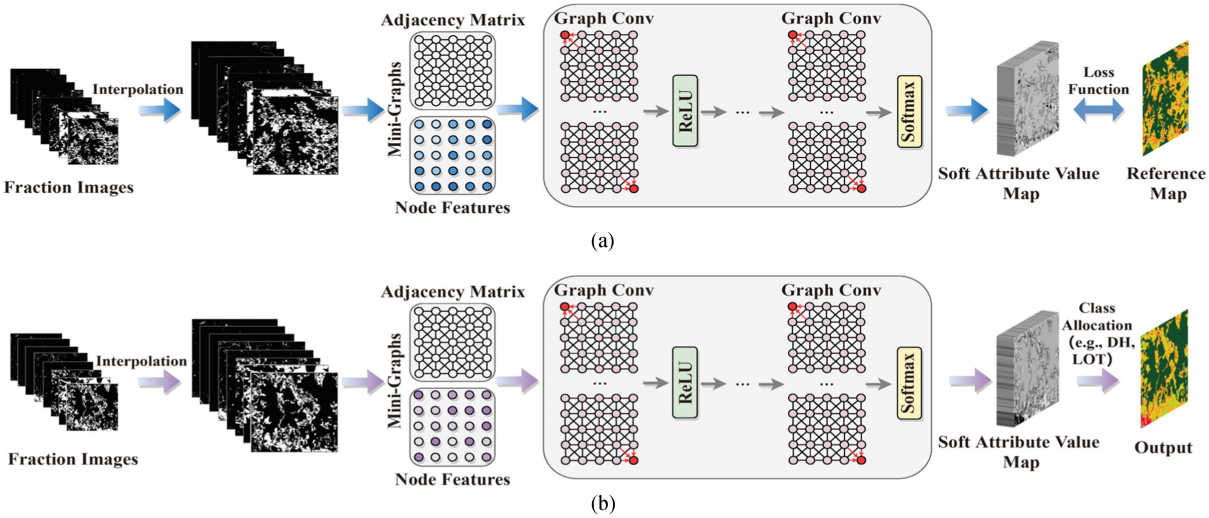


Fig. 1. Framework of our method.

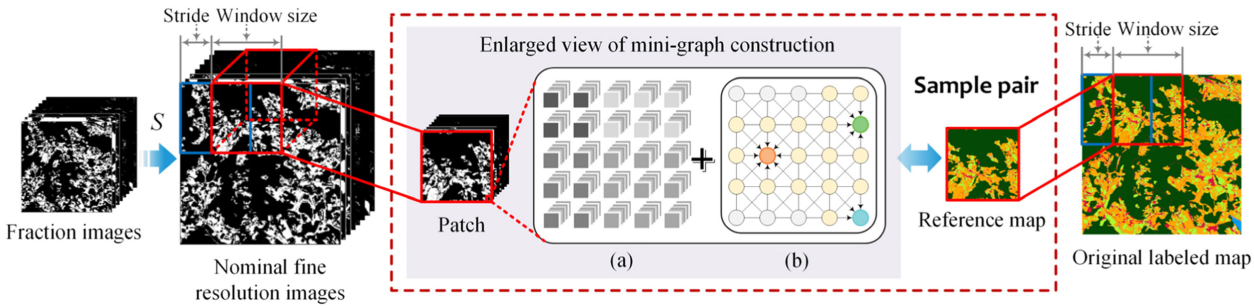


Fig. 2. Generating progress of the mini-graph and the sample pair. In the middle is an enlarged schematic view of the lower right corner of the mini-graph constructed based on the patch. (a) Graph node features derived from class fractions of subpixels. (b) First-order neighborhood diagram in the mini-graph and every node denotes a subpixel node. The orange node is an inner node on the graph with eight yellow first-order neighborhoods. The blue node is a corner node on the graph with three yellow first-order neighborhoods. The green node is a border node on the graph with five yellow first-order neighborhoods. The arrows point to the direction of information aggregation, that is, to their center nodes.

on the weighted $\boldsymbol{A}$, an additional trainable edge weighted parameter is added in each layer so that the correlation between land cover information represented by adjacent subpixels defined by elements of $\boldsymbol{A}$ can be dynamically adjusted via layer-wise training. Each element of $\boldsymbol{A}^{(l)}$ in the l^{th} layer be measured as

$$\boldsymbol{A}_{ij}^{(l)} = \begin{cases} \theta^{(l)} e^{-\gamma \|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2}, & \text{if } v_i \in N(v_j) \text{ or } v_i \in N(v_j) \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

where $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ denote the feature vectors associated with the nodes v_i and v_j. $N(v_i)$ (or $N(v_j)$) is the set of neighbors of v_i (or v_j). $\theta^{(l)}$ is the trainable parameter in the l^{th} layer with an initial value of 1, and each element of $\boldsymbol{A}$ is multiplied by this parameter to get the $\boldsymbol{A}^{(l)}$. As for the parameter γ, we empirically set 0.2 referring to [37] and [39].

In the second place, the propagation rule expressed in (5) can only propagate the aggregated features, but cannot directly fuse the input features (in other words, the relatively shallow information) of each layer. Actually, the input features are expected to help supplement the boundary and contour information of geographical objects, so we add a trainable input feature transformation matrix $\boldsymbol{W}_2^{(l)}$ of transformation parameters to the above propagation rule to further improve the process of feature representation. Now, we have the following layer-wise propagation rule in this article:

$$\boldsymbol{H}^{(l)} = \sigma\left(\tilde{\boldsymbol{D}}^{(l)-\frac{1}{2}} \tilde{\boldsymbol{A}}^{(l)} \tilde{\boldsymbol{D}}^{(l)-\frac{1}{2}} \boldsymbol{H}^{(l-1)} \boldsymbol{W}_1^{(l)} + \boldsymbol{H}^{(l-1)} \boldsymbol{W}_2^{(l)} + \boldsymbol{b}^{(l)}\right) \tag{7}$$

where $\boldsymbol{W}_1^{(l)}$ is equivalent to $\boldsymbol{W}^{(l)}$ in (5) and $\boldsymbol{W}_2^{(l)}$ has the same dimensions with $\boldsymbol{W}_1^{(l)}$. $\tilde{\boldsymbol{A}}^{(l)}$ and $\tilde{\boldsymbol{D}}^{(l)}$ vary with $\boldsymbol{A}^{(l)}$ in each layer. $\tilde{\boldsymbol{D}}^{(l)-\frac{1}{2}} \tilde{\boldsymbol{A}}^{(l)} \tilde{\boldsymbol{D}}^{(l)-\frac{1}{2}} \boldsymbol{H}^{(l-1)} \boldsymbol{W}_1^{(l)}$ represents the features after information aggregation; $\boldsymbol{H}^{(l-1)} \boldsymbol{W}_2^{(l)}$ is the transformation of input features which represents relatively shallow features.

In the $\mathrm{SRM_{GCN}}$ network, a shallow GCN with four convolutional layers is used to train and predict the input datasets, where the first three graph convolutional layers are followed by the ReLU function and the last graph convolutional layer

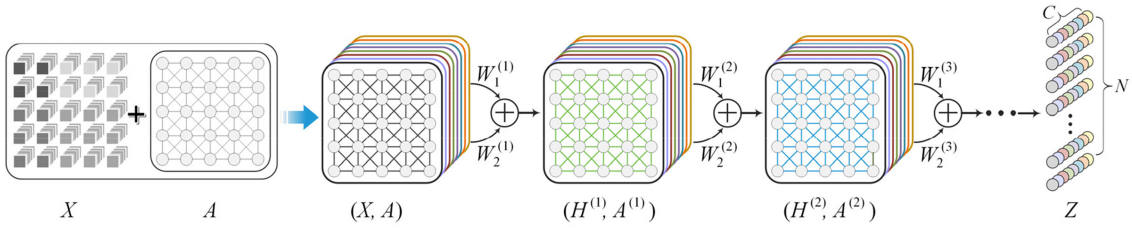


Fig. 3. Diagram of propagation of feature information. The superscripts l ($l = 1, 2, \ldots$) of H denote the output of the l^{th} layer. The superscripts l ($l = 1, 2, \ldots$) of W_1 and W_2 denote trained variables of the l^{th} layer. The superscripts l ($l = 1, 2, \ldots$) of $\boldsymbol{A}$ are calculated from the trained variables $\theta^{(l)}$ in the l^{th} layer.

is followed by the Softmax function. In general, the proposed network takes an initial node feature matrix $\boldsymbol{X}$ and an adjacency matrix $\boldsymbol{A}$ as inputs and produces node features $\boldsymbol{Z} \in \mathbb{R}^{N \times C}$ by stacking several graph convolutional layers of the form of (7), where C is the output dimension (i.e., the number of land cover classes in our cases). Finally, a C-dimensional feature vector for every node is produced, which represents a soft attribute estimation value of each subpixel, that is, the probability of each subpixel belonging to each class. For clarity, Fig. 3 demonstrates the propagation of feature information when stacking the first two graph convolutional layers.

For the SRM task, we adopt the cross-entropy loss function to evaluate the difference between the network output and the labels in the training set, which is

$$\text{loss} = -\sum_{g \in y_G} \sum_{c=1}^{C} \boldsymbol{Y}_{gc} \ln \boldsymbol{Z}_{gc} \tag{8}$$

where y_G is the node indices in the training set and $\boldsymbol{Y}$ denotes the label matrix.

3) Class-Allocation Algorithm: The allocation of classes for subpixels after the soft attribute estimation values are obtained at the subpixel scale is also a significant issue that directly affects the result of SRM [45]. In this article, we adopt two algorithms, direct hardening (DH) [46], [47] and linear optimization techniques (LOTs) [48], to complete class allocation.

Like the traditional hard classification, DH assigns each subpixel to the class with the maximum soft attribute value [47]. Its formula is shown in (9), where $p_{j,c}$ and $p_{j,c'}$ denote the soft attribute value of class c and c' respectively. This approach is simple and efficient, but it easily creates overly smooth mapping results because it does not take constraints from class fractions into account

$$y_{j,c} = \begin{cases} 1, \text{ if } p_{j,c} = max\left(p_{j,c'}\right) \text{ and } 1 \le c, c' \le C \\ 0, \text{ otherwise.} \end{cases} \tag{9}$$

LOT introduces an objective function for maximizing the sum of soft attribute values of all subpixels in each coarse resolution pixel. This objective function expressed in (10) needs to be subjected to constraints from class uniqueness and class fractions given in (11) [45], [48], where $x_{i,c|j}$ denotes the class fraction value of class C in pixel i corresponding to the given subpixel j. Both the function to maximize and the constraints are linear equations, so the theoretically optimal solution (i.e., optimal land cover classes of subpixels) of the above mathematical model can be find by LOT. Several algorithms can be applied, and the branch-and-bound algorithm [49] is employed for this purpose in this article. Although LOT is a little time-consuming compared with DH due to many iterations, the experiments in related studies have shown that it can usually produce pleasing mapping results [45], [50]. Note that this method is only applied to mixed pixels, whereas subpixels of pure pixels are still allocated land cover classes by the DH method to save time

$$\text{Maximize } F = \sum_{j=1}^{S \times S} \sum_{c=1}^{C} y_{j,c} \times p_{j,c} \tag{10}$$

$$\text{subject to } \begin{cases} \sum\limits_{c=1}^{C} y_{j,c} = 1 \\ \sum\limits_{j=1}^{S \times S} y_{j,c} = \text{round}\left(x_{i,c|j} \times S^2\right) \end{cases}. \tag{11}$$

III. Experiment and Analysis

To eliminate the uncertainty of soft classification, experiments were performed on three different resolution datasets with synthetic coarse fraction images to evaluate the proposed SRM_{GCN} method. The synthetic images were created by degrading the actual fine land cover map to coarse fraction images by applying an averaging filter with a scale factor of 3. The three original land cover maps all represented the real distribution of geographical objects on the earth's surface and were, respectively, used as reference images to assess the effectiveness of mapping results. A traditional hard classifier and three existing SRM methods with relatively remarkable performances, VBSPM [14], SAM [13], and SRM_{CNN} [25], were compared with the proposed SRM_{GCN} method.

A. Datasets

The first dataset was extracted from GlobeLand30 V2020, the 30-m resolution global land cover data product containing 966 tiles developed by China. This data could be obtained from the Global Land Cover Data Product Service website of the National Geomatics Center of China (www.globeland30.org; DOI: 10.11769). We selected one of the tiles marked N49_30_2020LC030 for experiments. In this tile comprising 19796 × 18988 pixels, we randomly extracted 80 chips with 540 × 540 pixels, 65 of which were used to generate mini-graphs for training, and the rest were used to generate mini-graphs for testing. The land cover map of the study area, which is regarded as the reference image, includes eight primary land cover classes:

TABLE I
KEY INFORMATION ABOUT DATASETS

Dataset	Spatial Resolution	Number of Classes	Proportion of Pure Pixels
GlobeLand30	30m	8	77.00%
Slovenia	10m	8	72.30%
Zurich Summer	0.61m	9	84.57%

Cultivated Land, Forest, Grass Land, Shrubland, Wetland, Water Body, Artificial Surfaces, and Bareland.

The second dataset is the example dataset of EOPatches for Slovenia 2019 (http://eo-learn.sentinel-hub.com). This dataset contains 25 groups of Sentinel-2 time series images and the corresponding 10-m resolution land use and land cover maps with 500×500 pixels. In our experiments, 20 land cover maps were used to generate training set and 5 were used to generate testing set. The land cover classes of this dataset are summarized into eight classes: Cultivated Land, Forest, Grassland, Shrubland, Water, Wetland, Artificial Surface, and Background.

The third dataset is a set of 20 chips very-high-resolution images named Zurich Summer (https://sites.google.com/site/michelevolpiresearch/data/zurich-dataset), which were captured by the QuickBird satellite in August 2002 over the city of Zurich, Switzerland [51]. The spatial resolution of the corresponding ground truth masks considered as reference images is 0.61 m as pansharpened panchromatic images, and the average size of the masks is 1182×1051 pixels. We used four chips (No. 1, 5, 6 and 13) to generate mini-graphs for testing and the rest to generate mini-graphs for training. Nine different urban and periurban classes were manually annotated in ground truth masks: roads; buildings; trees; grass; bare soil; water; railways; swimming pools; and others (Background).

Some key information about these three datasets is given in Table I. Note that the proportion of pure pixels in the last column refers to the ratio of the number of pixels containing only one land cover class to the total number of coarse resolution pixels in fraction images.

B. Experimental Settings

According to Section II-C, a series of mini-graphs of each dataset were generated and split into a training set, a validation set, and a test set to test our model. The training sets of the three datasets involve samples of all classes. The size of the patch used to construct the mini-graph and the number of mini-graphs of the three datasets are given in Table II.

In our experiments, the numbers of input and output channels of other layers are set to 64; the exception is that the number of input channels of the first layer and output channels of the last layer are equal to the number of land cover classes. The proposed network was completed via the Python-3.6 platform [52] and PyTorch-1.4.0 framework [53]. Training was run on an NVIDIA Quadro RTX 5000 GPU with 16 GB GPU memory. We selected the Adam algorithm [54] to optimize the network with a learning rate equal to 0.005. In the process of network training, the number of epochs and batch size were set to 200 and 8, respectively.

TABLE II
SIZE OF PATCHES AND NUMBER OF MINI-GRAPHS

Dataset		Patch Size	Number of Mini-graphs
GlobeLand30	Training	180×180	1200
	Validation		300
	Test		180
Slovenia	Training	166×166	450
	Validation		50
	Test		45
Zurich Summer	Training	166×166	848
	Validation		212
	Test		200

C. Results

From the perspective of visual evaluation and quantitative accuracy assessment, the effectiveness of the proposed $\mathrm{SRM_{GCN}}$ and aforementioned baseline methods on the test set of the three datasets are evaluated. The per-class accuracy, overall accuracy (OA), average accuracy (AA), and kappa coefficient (Kappa) are employed as the metrics for the latter.

1) *Results on the GlobeLand30 Dataset*: A visual comparison of SRM results generated by different methods on the GlobeLand30 dataset is shown in Fig. 4(c)–(g). The reference map and hard classification map are also provided in Fig. 4(a) and (b). The classification maps of VBSPM and SAM result in salt-and-pepper noise on the edge of the objects, especially linear objects. Contrarily, $\mathrm{SRM_{CNN}}$, $\mathrm{SRM_{GCN(DH)}}$, and $\mathrm{SRM_{GCN(LOT)}}$ are able to yield smoother visual effects. Additionally, it is also notable that the linear water body is discontinuous in the $\mathrm{SRM_{CNN}}$ result. In contrast, since $\mathrm{SRM_{GCN(DH)}}$ and $\mathrm{SRM_{GCN(LOT)}}$ take effort to construct and learn the spatial relationship between nodes, the continuity and integrity of these elongated features are well preserved. Note that, in the $\mathrm{SRM_{GCN(DH)}}$ result, some subpixels of "Shrubland" dotted on the "grass land" are misclassified into "grass land" because the DH class-allocation algorithm brings overly smooth results to some extent.

Table III quantitatively reports the assessment results obtained by different methods on the GlobeLand30 dataset. Since the input fraction images produced by degrading reference images are error free, the values of the assessment indexes of the five SRM results are all relatively high. Consistent with the above visual evaluation, the OA, AA, and Kappa values of the DNNs-based SRM methods, $\mathrm{SRM_{CNN}}$, $\mathrm{SRM_{GCN(DH)}}$, and $\mathrm{SRM_{GCN(LOT)}}$, are higher than those of VBSPM and SAM. Our GCN-based methods perform better than the $\mathrm{SRM_{CNN}}$ because elongated features are fragmentary using $\mathrm{SRM_{CNN}}$. In general, $\mathrm{SRM_{GCN(LOT)}}$ with LOT class-allocation algorithm achieves the optimal performances in AA and most specific class accuracies, especially "Shrubland," whose geo-objects are spotted, and $\mathrm{SRM_{GCN(DH)}}$ provides the highest values in OA and Kappa.

2) *Results on the Slovenia Dataset*: The reference map, hard classification result, and five SRM results on the Slovenia dataset are displayed in Fig. 5. It is obvious

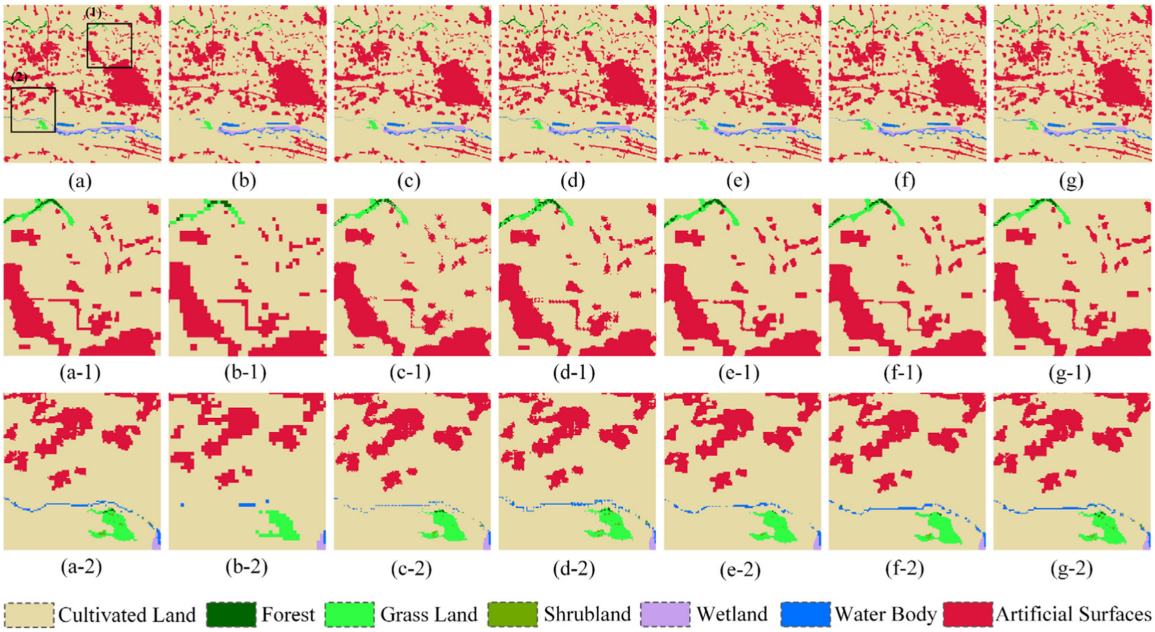


Fig. 4. Example of classification maps with 496 × 496 pixels from the test set of the GlobeLand30 dataset. (a) Reference map. (b) Hard classification map. (c)–(g) SRM results of VBSPM, SAM, $\mathrm{SRM_{CNN}}$, $\mathrm{SRM_{GCN(DH)}}$, and $\mathrm{SRM_{GCN(LOT)}}$, respectively. The second and third rows are zoomed-in area 1 and zoomed-in area 2 of the first row, respectively.

TABLE III
PER-CLASS ACCURACY (%), OA (%), AA (%), AND KAPPA COEFFICIENT ON THE GLOBELAND30 DATASET

Class Name	VBSPM	SAM	$\mathrm{SRM_{CNN}}$	$\mathrm{SRM_{GCN(DH)}}$	$\mathrm{SRM_{GCN(LOT)}}$
Cultivated Land	96.14	96.84	97.88	97.80	**97.89**
Forest	96.54	95.88	97.17	**97.25**	96.81
Grass Land	71.44	77.44	79.60	80.61	**81.67**
Shrubland	30.09	46.02	29.20	35.40	**51.33**
Wetland	93.10	95.25	94.43	95.55	**96.53**
Water Body	96.42	97.10	97.99	98.42	**98.53**
Artificial Surfaces	94.40	96.46	97.83	**98.23**	97.92
Bareland	97.57	**99.05**	98.58	98.69	98.58
OA	94.71	95.18	96.41	**96.51**	96.39
AA	84.46	88.08	86.58	87.74	**89.91**
Kappa	0.9143	0.9223	0.9420	**0.9437**	0.9418

Note: The best one is shown in bold.

that the traditional hard classification can only provide an extremely rough description of the distribution pattern of geographical objects, whereas the other five SRM methods can portray this spatial distribution in more detail. In addition, it can be seen clearly from the comparison of two critical zoomed-in areas that the slender linear artificial surface objects are disconnected and even almost vanish in the fine classification maps produced by VBSPM and $\mathrm{SRM_{CNN}}$. Although SAM retains more details, it is unable to determine the correct spatial position of each class but creates discrete messy small patches. Due to building the spatial contextual structure via graph, $\mathrm{SRM_{GCN(DH)}}$ and $\mathrm{SRM_{GCN(LOT)}}$ generate results that are more similar to the reference maps than other methods with the correct and complete linear information retained in the regions that are difficult to recover. Also, as expected, details produced by $\mathrm{SRM_{GCN(LOT)}}$ are better reconstructed because of the increased constraints from class fractions by utilizing the LOT algorithm.

The accuracy assessments of the five methods on the Slovenia dataset bear out the aforementioned findings. Similar to the GlobeLand30 dataset, the results in Table IV indicate that the OA, AA, and Kappa of VBSPM and SAM are at least 2% lower than those gained by the SRM methods based on deep learning, i.e., $\mathrm{SRM_{CNN}}$, $\mathrm{SRM_{GCN(DH)}}$, and $\mathrm{SRM_{GCN(LOT)}}$. Moreover, the highest values of OA, AA, and Kappa are captured by our proposed two GCN-based methods, i.e., $\mathrm{SRM_{GCN(DH)}}$ and $\mathrm{SRM_{GCN(LOT)}}$, which validates the strength of the graph structural features in SRM technology. As for the way of class allocation, the LOT results are not that different from the DH results, whereas the visual effect of the former is superior to that of the latter, as mentioned above.

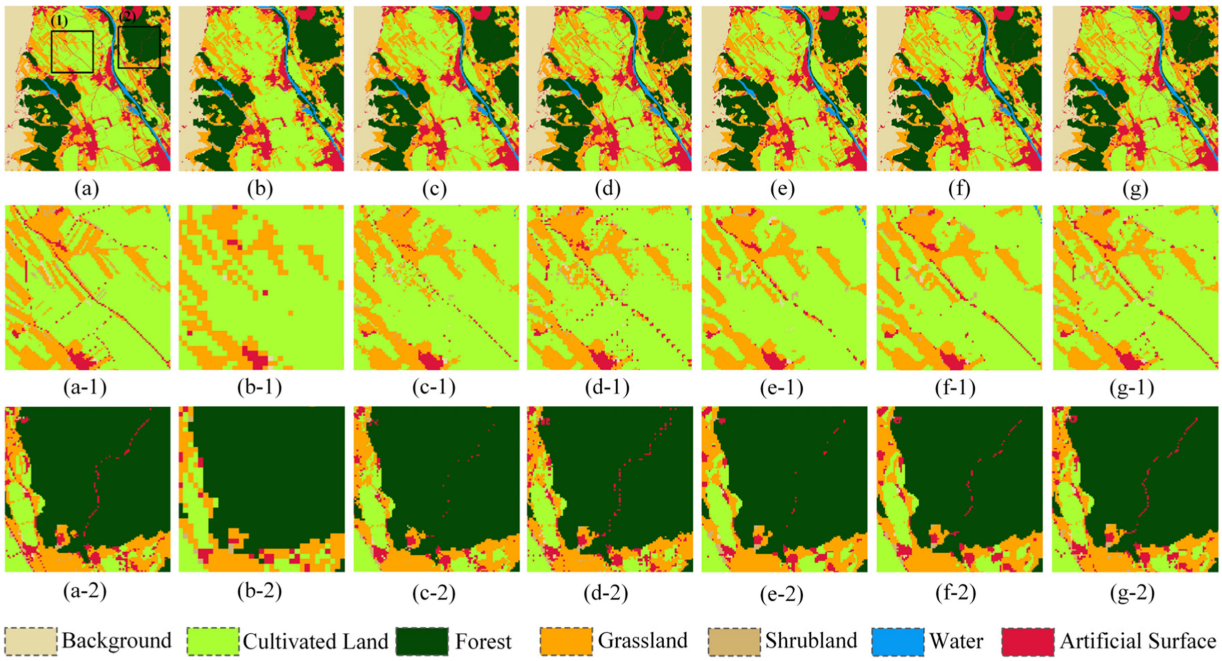


Fig. 5. Example of classification maps with 390 × 390 pixels from the test set of the Slovenia dataset. (a) Reference map. (b) Hard classification map. (c)–(g) SRM results of VBSPM, SAM, $\mathrm{SRM}_{\mathrm{CNN}}$, $\mathrm{SRM}_{\mathrm{GCN(DH)}}$, and $\mathrm{SRM}_{\mathrm{GCN(LOT)}}$, respectively. The second and third rows are zoomed-in area 1 and zoomed-in area 2 of the first row, respectively.

TABLE IV
PER-CLASS ACCURACY (%), OA (%), AA (%), AND KAPPA COEFFICIENT ON THE SLOVENIA DATASET

Class Name	VBSPM	SAM	$\mathrm{SRM}_{\mathrm{CNN}}$	$\mathrm{SRM}_{\mathrm{GCN(DH)}}$	$\mathrm{SRM}_{\mathrm{GCN(LOT)}}$
Background	97.79	97.82	**98.44**	98.34	98.39
Cultivated Land	87.33	87.52	**91.54**	90.63	91.34
Forest	98.35	98.18	98.89	98.79	**98.91**
Grassland	89.63	87.81	92.39	**93.28**	91.32
Shrubland	57.26	65.88	72.20	71.37	**76.71**
Water	80.64	84.17	87.07	87.01	**88.99**
Wetland	78.75	83.75	81.88	**88.12**	85.62
Artificial Surface	71.49	76.08	77.60	79.92	**81.59**
OA	92.09	92.14	94.31	**94.49**	94.42
AA	82.65	85.15	87.50	88.43	**89.11**
Kappa	0.8920	0.8930	0.9225	**0.9249**	0.9241

Note: The best one is shown in bold.

3) *Results on the Zurich Summer Dataset*: Fig. 6 illustrates the SRM results of five different methods on the Zurich summer dataset, where some enlarged areas of the first row are further provided in the last two rows for better visual inspection. Although the results of the five methods show no significant differences in this dataset from the overall perspective, the local views still indicate that $\mathrm{SRM}_{\mathrm{CNN}}$, $\mathrm{SRM}_{\mathrm{GCN(DH)}}$, and $\mathrm{SRM}_{\mathrm{GCN(LOT)}}$ are excellent at suppressing speckle noise and blurred boundaries compared with VBSPM and SAM. For instance, "buildings," "roads," and "trees" all have clearer boundaries in the results generated by these three SRM methods based on DNNs, particularly when there are small protrusions in the boundaries.

The quantitative evaluation results of the OA, AA, and Kappa on the Zurich Summer dataset are given in Table V. It is apparent that the performance of all methods is better than that of the GlobeLand30 and Slovenia datasets. This can be attributed mainly to the higher proportion of pure pixels (see the last column of Table I) and simpler spatial patterns of the Zurich Summer dataset. Consistent with visual perception, the statistical results of the evaluation indexes of all methods are similar; however, slight superiorities can still be observed when $\mathrm{SRM}_{\mathrm{GCN(DH)}}$ and $\mathrm{SRM}_{\mathrm{GCN(LOT)}}$ are compared with $\mathrm{SRM}_{\mathrm{CNN}}$ and other competitors in terms of all experimental assessment metrics, which highlights the potential of the proposed GCN-based method in SRM analysis. It is also worth noting that $\mathrm{SRM}_{\mathrm{GCN(LOT)}}$ secures maximum values with respect to almost all statistics

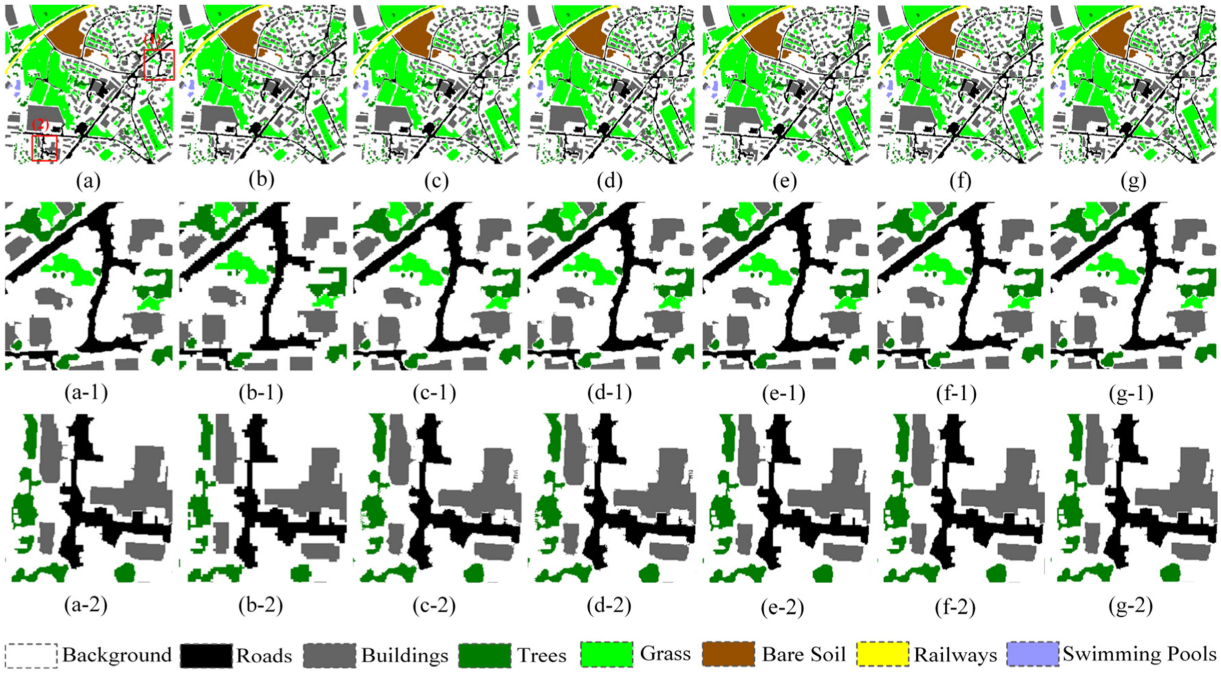


Fig. 6. Example of classification maps with 1362 × 1293 pixels from the test set of the Zurich Summer dataset. (a) Reference map. (b) Hard classification map. (c)–(g) SRM results of VBSPM, SAM, $\mathrm{SRM_{CNN}}$, $\mathrm{SRM_{GCN(DH)}}$, and $\mathrm{SRM_{GCN(LOT)}}$, respectively. The second and third rows are zoomed-in area 1 and zoomed-in area 2 of the first row, respectively.

TABLE V
PER-CLASS ACCURACY (%), OA (%), AA (%), AND KAPPA COEFFICIENT ON THE ZURICH SUMMER DATASET

Class Name	VBSPM	SAM	$\mathrm{SRM_{CNN}}$	$\mathrm{SRM_{GCN(DH)}}$	$\mathrm{SRM_{GCN(LOT)}}$
Background	98.37	99.04	99.37	99.39	**99.51**
Roads	98.16	99.04	99.29	**99.59**	99.51
Buildings	98.29	99.20	99.48	**99.59**	99.58
Trees	96.76	98.32	98.97	98.98	**99.15**
Grass	98.92	99.42	99.57	99.67	**99.70**
Bare Soil	99.61	99.83	99.86	99.89	**99.89**
Water	97.36	94.96	98.68	**99.28**	99.16
Railways	98.92	99.65	99.69	99.71	**99.78**
Swimming pools	98.50	99.35	99.32	99.25	**99.54**
OA	98.28	99.07	99.38	99.46	**99.52**
AA	98.32	98.76	99.36	99.48	**99.54**
Kappa	0.9755	0.9868	0.9911	0.9923	**0.9932**

Note: The best one is shown in bold.

except that the per-class accuracies of "roads," "buildings," and "water" are modestly lower than the results of $\mathrm{SRM_{GCN(DH)}}$.

IV. DISCUSSION

A. *Usefulness of the Proposed Graph Convolutional Layer*

As mentioned in the introduction and Section II-C, we are committed to improving the original graph convolution operation in the construction of the adjacency matrix and the information propagation rule. Here, we conducted an ablation study on the Slovenia dataset to illustrate the usefulness of these two operations. The DH algorithm was employed for all approaches to determine the optimal land-cover labels of subpixels, and the network structures were kept identical to the abovementioned experiments except for graph convolution operations of convolutional layers. At a scale factor of 3, the visual comparison of the SRM results using four different graph convolution approaches on another test area of the Slovenia dataset are exhibited in Fig. 7(c)–(f), and their varied accuracy assessment results are given in Table VI. Approach 1 uses the original convolutional layers with the similarity distance-based weighted adjacency matrix, but without the trainable edge weighted parameter based on it and the trainable input feature transformation matrix. On this foundation, approaches 2 and 3, respectively, only use the proposed adjacency matrix construction operation and information aggregation operation. The last graph convolution approach is our proposed $\mathrm{SRM_{GCN(DH)}}$, which contains both of the above operations.

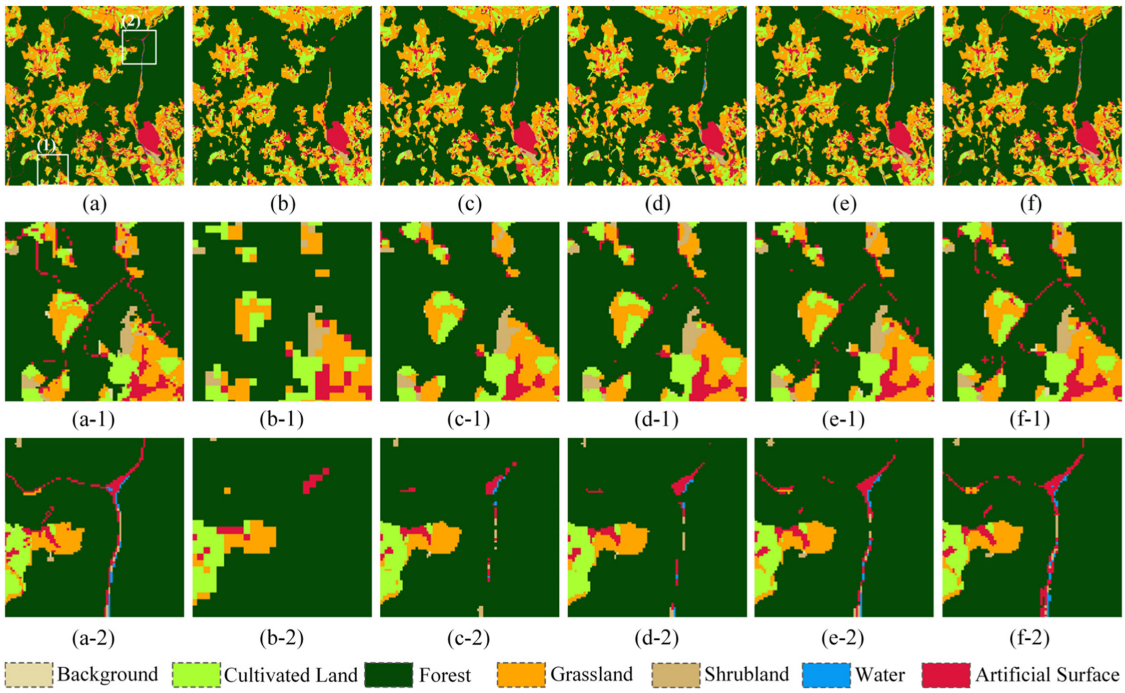


Fig. 7. Example of classification maps with 440 × 440 pixels of the Slovenia dataset achieved by different graph convolution approaches. (a) Reference map. (b) Hard classification map. (c)–(e) SRM results using the graph convolution of approaches 1 to 3, respectively. (f) SRM result of $\mathrm{SRM}_{\mathrm{GCN(DH)}}$. The second and third rows are zoomed-in area 1 and zoomed-in area 2 of the first row, respectively.

TABLE VI
PER-CLASS ACCURACY (%), OA (%), AA (%), AND KAPPA COEFFICIENT ACHIEVED BY DIFFERENT GRAPH CONVOLUTION APPROACHES ON THE SLOVENIA DATASET

Class Name	Approach 1	Approach 2	Approach 3	$\mathrm{SRM}_{\mathrm{GCN(DH)}}$
Background	98.13	98.14	98.01	**98.34**
Cultivated Land	89.99	90.10	**91.61**	90.63
Forest	98.72	98.70	**99.05**	98.79
Grassland	90.93	90.51	92.51	**93.28**
Shrubland	64.41	63.52	**74.00**	71.37
Water	79.60	85.27	86.16	**87.01**
Wetland	76.88	80.62	81.88	**88.12**
Artificial Surface	71.63	75.86	79.02	**79.92**
OA	93.05	93.32	94.49	**94.49**
AA	83.79	85.34	87.78	**88.43**
Kappa	0.9050	0.9089	0.9248	**0.9249**

Note: The best one is shown in bold.

It can be observed from Fig. 7 that the SRM results obtained by different graph convolution approaches have significant divergences in the recovery of small-sized patches and elongated features. By and large, due to the lack of layer-wise input features, the resultant classification maps of approaches 1 and 2 are both visually smooth, such that the details fail to be restructured, even though the performance of approach 2 is a little better than that of Approach 1 because of an extra parameter for each layer to regulate the edge weights (e.g., the partial linear features of "artificial surface" are captured in Fig. 7(d–1), but not in (c–1)). Contrarily, the results of approach 3 and $\mathrm{SRM}_{\mathrm{GCN(DH)}}$ retain more details, especially the narrow linear information of "water," "artificial surface," and "shrubland." This further demonstrates the significance of layer-wise input features in the SRM process based on GCN. Most notably, $\mathrm{SRM}_{\mathrm{GCN(DH)}}$ adopts the proposed approaches of adjacency matrix construction and information propagation and is capable of restoring the details to the greatest extent, including boundaries of areal objects, linear structural features, and fragmented patches. In Table VI, consistent with visual results, there is a noticeable performance drop in the quantitative results of approaches 1 and 2 compared with other graph convolution approaches that directly integrate the input features' transformation information of each layer. Although the numerical value results of $\mathrm{SRM}_{\mathrm{GCN(DH)}}$

TABLE VII
OA (%), AA (%), AND KAPPA COEFFICIENT ON THREE DATASETS AFTER EXCLUDING PURE PIXELS

Dataset	Assessment Index	VBSPM	SAM	$\mathrm{SRM}_{\mathrm{CNN}}$	$\mathrm{SRM}_{\mathrm{GCN(DH)}}$	$\mathrm{SRM}_{\mathrm{GCN(LOT)}}$
GlobeLand30	Adjusted OA	79.03	80.86	85.76	**86.17**	85.66
	Adjusted AA	74.33	80.33	80.21	81.96	**84.12**
	Adjusted Kappa	0.6928	0.7233	0.7932	**0.7994**	0.7928
Slovenia	Adjusted OA	76.47	76.60	83.09	**83.61**	83.40
	Adjusted AA	70.54	73.82	78.94	79.92	**81.03**
	Adjusted Kappa	0.6846	0.6896	0.7742	**0.7807**	0.7798
Zurich Summer	Adjusted OA	89.59	94.39	96.29	96.74	**97.12**
	Adjusted AA	89.70	93.63	96.36	96.79	**97.20**
	Adjusted Kappa	0.8481	0.9184	0.9461	0.9527	**0.9581**

Note: The best one is shown in bold.

and approach 3 are relatively close except for several per-class accuracies, $\mathrm{SRM}_{\mathrm{GCN(DH)}}$ contributes to the more satisfactory visual effect, as seen in Fig. 7. The better performance of our method is attributed to fully considering the spatial relationship between adjacent nodes representing land covers, while making full use of relatively shallow shape information.

B. Comparison of SRM Results on Three Datasets

Our experiments have obtained the SRM results using different methods on three datasets with different spatial resolutions. Since the overall assessment results of OA, AA, and Kappa in Tables III–V are affected by the proportion of pure pixels in the fraction images with different spatial resolutions, we summarized the statistical results of removing pure pixels on the three datasets, as given in Table VII. In Table VII, adjusted OA, adjusted AA, and adjusted kappa denote the OA, AA, and Kappa coefficients after excluding pure pixels, respectively. From the resultant table, we find that the SRM results of all methods on the three datasets have a certain degree of reduction after removing the pure pixels; the Zurich Summer dataset has the smallest reduction. Meanwhile, the gap in the quantitative evaluation results of various methods is further increased, especially between DNNs-based methods and the traditional methods; the proposed two methods can result in more accurate spatial location results of subpixels within mixed pixels.

Moreover, the accuracy results of the five methods on the GlobeLand30 dataset after pure pixels are removed are slightly higher than those of the Slovenia dataset. On the Zurich summer dataset, the results of all methods increase significantly, which is similar to the conclusions obtained from comparing Tables III–V. Under the same scale factor, the reason for these results may be related to the spatial distribution characteristics of the geographical objects in different datasets. To clarify this further, we calculated the values of two landscape indices on the test set of the three datasets to represent the spatial pattern characteristics; the results are illustrated in Fig. 8. The number of patches (NP) per 10 000 grids can be used to describe the heterogeneity of the entire landscape, and its value has a good positive correlation with the degree of fragmentation of the landscape. Generally speaking, the larger the NP, the higher the degree of fragmentation. The aggregation index (AI) [55] is a ratio of actual shared edges to maximal possible shared edges,

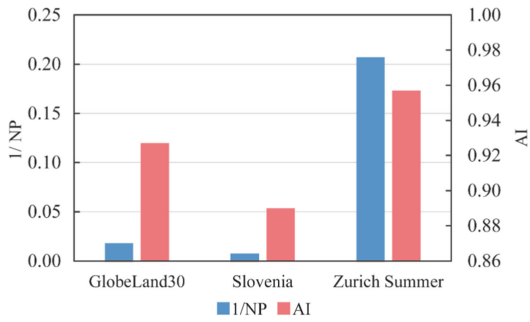


Fig. 8. NP and AI of three datasets. Note that we use the reciprocal of NP, i.e., 1/NP, to directly reflect the positive correlation with accuracy.

which is designed to measure aggregation of spatial patterns. The AI of the highest level of aggregation is equal to 1, and the AI of the lowest level of aggregation is equal to 0.

Fig. 8 indicates that compared with other datasets, the spatial pattern of the Slovenia dataset is the most fragmented, which makes the SRM process more difficult because more complicated land covers need to be located; hence, it has the lowest accuracy. In contrast, the spatial structure of the Zurich summer dataset is highly aggregated, which can bring more accurate results with a simpler SRM process.

C. Impact of Different Scale Factors on the SRM Results

The influence of different scale factors on the SRM results is also worthy of investigation. Therefore, the SRM performances with different scale factors were analyzed by studying on the GlobeLand30 dataset. Based on the experiments with a scale factor of 3 in Section III, two scale factors $S = 4$ and $S = 5$ were added and tested on the GlobeLand30 dataset for evaluating the accuracy of the five SRM methods. Fig. 9 and Table VIII compare the visual performance and quantitative assessments, respectively.

The results shown in Fig. 9(b)–(f) indicate that speckle noise and blurred boundary are significantly improved by three DNNs-based methods, especially at the scale factor of 5. In addition, the results of $\mathrm{SRM}_{\mathrm{GCN}}$ are closer to the reference image, because the mapping results of $\mathrm{SRM}_{\mathrm{CNN}}$ look slightly smoother, which makes some details disappear. Moreover, the mapping accuracy gradually decreased with an increased scale factor, which can

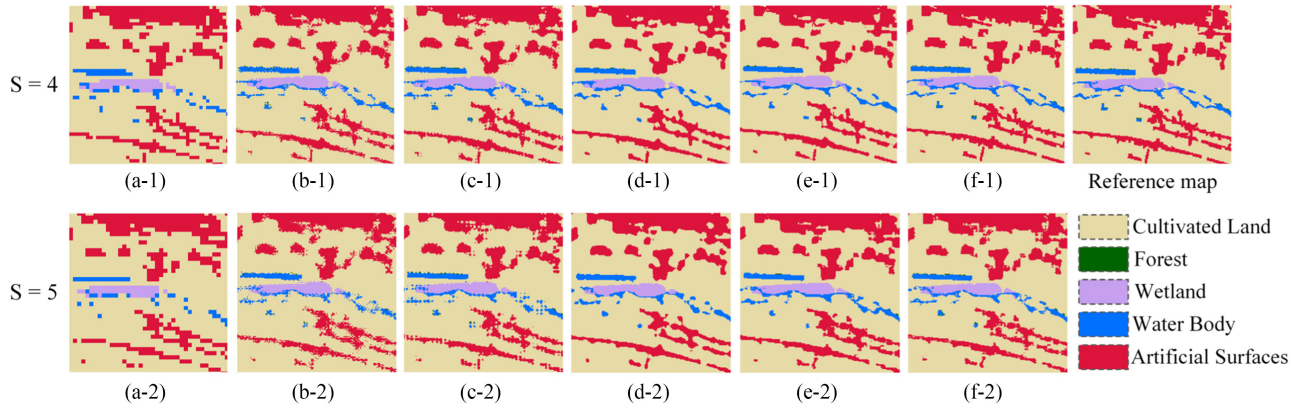


Fig. 9. Classification maps with 190 × 190 pixels of different scale factors of the GlobeLand30 dataset. (a) Hard classification map. (b)–(f) SRM results of VBSPM, SAM, $\mathrm{SRM_{CNN}}$, $\mathrm{SRM_{GCN(DH)}}$, and $\mathrm{SRM_{GCN(LOT)}}$, respectively.

TABLE VIII
OA (%), AA (%), AND KAPPA COEFFICIENT OF DIFFERENT SCALE FACTORS ON THE GLOBELAND30 DATASET

Scale Factor	Assessment Index	VBSPM	SAM	$\mathrm{SRM_{CNN}}$	$\mathrm{SRM_{GCN(DH)}}$	$\mathrm{SRM_{GCN(LOT)}}$
S=4	OA	92.78	92.95	94.77	**94.82**	94.47
	AA	81.29	82.86	**84.70**	82.39	84.53
	Kappa	0.8833	0.8865	0.9153	**0.9161**	0.9109
S=5	OA	90.38	91.07	**93.06**	92.92	92.45
	AA	78.72	81.03	78.61	78.99	**82.85**
	Kappa	0.8445	0.8561	**0.8874**	0.8852	0.8784

Note: The best one is shown in bold.

TABLE IX
PARAMETERS OF $\mathrm{SRM_{GCN}}$ AND $\mathrm{SRM_{CNN}}$ MODELS ON THE SLOVENIA DATASET

Model	Layer	W_1	W_2	b	θ	Total Number of Parameters
$\mathrm{SRM_{GCN}}$	Graph_Conv1	8×64	8×64	64	1	18636
	Graph_Conv2	64×64	64×64	64	1	
	Graph_Conv3	64×64	64×64	64	1	
	Graph_Conv4	64×8	64×8	8	1	
$\mathrm{SRM_{CNN}}$			–			332296

be more easily seen by observing the quantitative results from Tables III and VIII. The possible reason is that the spatial locations of subpixels within mixed pixels increased with the increase of scale factor which makes the SRM process more complicated [45].

D. Further Comparison between $\mathrm{SRM_{CNN}}$ and $\mathrm{SRM_{GCN}}$

To further understand the CNN-based and GCN-based SRM methods used in this article, a further comparison is made with respect to model structure and number of parameters. Taking the Slovenia dataset as an example, where the number of output channels in this dataset is 8, the statistical results of the model structure and parameters of each trainable variable about $\mathrm{SRM_{GCN}}$ are given in Table IX, and the total number of parameters of $\mathrm{SRM_{CNN}}$ at a scale factor of 3 and $\mathrm{SRM_{GCN}}$ are also reported in it.

Compared with the model structure of $\mathrm{SRM_{CNN}}$ in [25], the proposed $\mathrm{SRM_{GCN}}$ has a fairly simple architecture, with only four convolutional layers operating in the graph. In addition, $\mathrm{SRM_{CNN}}$ requires approximately 18 times the model parameters of our methods. This indicates that the proposed method only needs significantly fewer parameters to better learn the complex nonlinear relationship between the coarse fraction images and fine resolution land cover maps; this also reveals the effectiveness of the structure information defined on the graph in SRM. In the meantime, the smaller number of parameters also makes it possible to learn with smaller samples.

E. Limitations and Future Work

Although this study has shown the effectiveness of structure information constructed by $\mathrm{SRM_{GCN}}$, the study has some limitations. First, in terms of encoding an image into a graph, our experiments use subpixels as nodes and the first-order neighborhoods of each node as edges. As a result, there are many nodes and edges in the graph; this means a large amount of calculation is required, and it is inconvenient to extract the long-range spatial

relations that help improve the effect of SRM. Therefore, other graph construction methods based on remote sensing images are worth exploring. Inspired by the advantages of object-based image analysis in processing remote sensing images [11], [56], using objects as graph nodes is a promising way to address the above problems and is more conducive to maximizing the advantages of GCN.

Second, to simply compare the performances of the SRM algorithms, the synthetic coarse fraction images are used as input in this article, and the results produced by our experiments are the most ideal. However, the errors caused by soft classification algorithms are inescapable. For the sake of avoiding uncertainty propagation from soft classification to SPM, several DNNs-based SRM methods taking coarse remote sensing multiband images as input rather than fraction images have achieved the joint optimization of soft classification and SRM [30], [32]. So, similar approaches are expected to be further examined for the GCN-based methods.

Finally, in our experiments, $\mathrm{SRM_{GCN}}$ performs better than other baseline methods, but the results still have some shortcomings compared with reference maps. Some fragmented plots still disappear, and the boundaries of the geographical objects are still somewhat smooth. One reason is that only applying fraction images to extract features at the subpixel scale is insufficient to tackle the SRM problem [25], [50]. Thus, auxiliary data [50] and information on other scales [12], [57], such as the object-based features from coarse remote sensing images (e.g., shape and texture features), could be considered in the GCN-based SRM process.

V. Conclusion

Inspired by the successful applications of GCN in graph structure data, a GCN-based SRM method with simple architecture, named $\mathrm{SRM_{GCN}}$, is proposed in this article. As opposed to the original graph convolution operation, we designed two operations in terms of adjacency matrix construction and the information propagation rule to improve the performance of GCN in SRM. Meanwhile, to build a generalizable classifier, we propose a supervised inductive learning strategy of GCN with the mini-graphs as input. The experimental results of three datasets with different spatial resolutions demonstrate that the $\mathrm{SRM_{GCN}}$ is capable of yielding better performance in both visual and quantitative assessments. It is particularly conducive to restructuring small patches, and linear features draw support from the graph structure information. In future studies, the integration of other available information, such as object-based features and auxiliary data, with $\mathrm{SRM_{GCN}}$ is recommended for more accurate SRM analysis. Simultaneously, other means of graph construction for remote sensing images are worth exploring to further improve the restorability of GCN-based SRM methods.

References

[1] P. Fisher, "The pixel: A snare and a delusion," *Int. J. Remote Sens.*, vol. 18, no. 3, pp. 679–685, Feb. 1997.

[2] P. M. Atkinson, "Downscaling in remote sensing," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 22, no. 1, pp. 106–114, Jun. 2013.

[3] Y. Ge, "Sub-pixel land-cover mapping with improved fraction images upon multiple-point simulation," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 22, no. 1, pp. 115–126, Jun. 2013.

[4] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Proc. Mag.*, vol. 19, no. 1, pp. 44–57, Feb. 2002.

[5] D. Lu and Q. Weng, "A survey of image classification methods and techniques for improving classification performance," *Int. J. Remote Sens.*, vol. 28, no. 5, pp. 823–870, Mar. 2007.

[6] Y. Ge, S. Li, and V. C. Lakhan, "Development and testing of a subpixel mapping algorithm," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 7, pp. 2155–2164, Jul. 2009.

[7] P. M. Atkinson, "Issues of uncertainty in super-resolution mapping and their implications for the design of an inter-comparison study," *Int. J. Remote Sens.*, vol. 30, no. 20, pp. 5293–5308, Oct. 2009.

[8] P. M. Atkinson, "Mapping sub-pixel boundaries from remotely sensed images," in *Innovations in GIS*, vol. 4, Z. Kemp, Ed., London, U.K.: Taylor and Francis, 1997, pp. 166–180.

[9] Q. Wang, C. Zhang, X. Tong, and P. M. Atkinson, "General solution to reduce the point spread function effect in subpixel mapping," *Remote Sens. Environ.*, vol. 251, pp. 112054, Dec. 2020.

[10] W. R. Tobler, "A computer movie simulating urban growth in the detroit region," *Econ. Geography*, vol. 46, no. 2, pp. 234–240, Jun. 1970.

[11] Y. Chen, Y. Ge, G. B. Heuvelink, R. An, and Y. Chen, "Object-based superresolution land-cover mapping from remotely sensed imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 1, pp. 328–340, Sep. 2017.

[12] Y. Chen, Y. Ge, Y. Chen, Y. Jin, and R. An, "Subpixel land cover mapping using multiscale spatial dependence," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 9, pp. 5097–5106, Mar. 2018.

[13] K. C. Mertens, B. D. Baets, L. P. C. Verbeke, and R. R. D. Wulf, "A sub-pixel mapping algorithm based on sub-pixel/pixel spatial attraction models," *Int. J. Remote Sens.*, vol. 27, no. 15, pp. 3293–3310, Aug. 2006.

[14] Y. Ge, Y. Chen, S. Li, and Y. Jiang, "Vectorial boundary-based sub-pixel mapping method for remote-sensing imagery," *Int. J. Remote Sens.*, vol. 35, no. 5, pp. 1756–1768, Mar. 2014.

[15] K. Wu, P. Li, and L. Zhang, "Sub-pixel mapping of remote sensing image based on MAP model," in *Proc. 4th Int. Conf. Image Graph*, 2007, pp. 742–746.

[16] A. Ma, Y. Zhong, D. He, and L. Zhang, "Multiobjective subpixel land-cover mapping," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 1, pp. 422–435, Jan. 2017.

[17] M. Song, Y. Zhong, A. Ma, X. Xu, and L. Zhang, "Multiobjective subpixel mapping with multiple shifted hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 11, pp. 8176–8191, May. 2020.

[18] X. Xu *et al.*, "A new spectral-spatial sub-pixel mapping model for remotely sensed hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 11, pp. 6763–6778, Nov. 2018.

[19] Y. Ge *et al.*, "Principles and methods of scaling geospatial earth science data," *Earth Sci. Rev.*, vol. 197, Jul. 2019, Art. no. 102897.

[20] F. Ling *et al.*, "Learning-Based superresolution land cover mapping," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 7, pp. 3794–3810, Mar. 2016.

[21] Y. Zhang, Y. Du, F. Ling, S. Fang, and X. Li, "Example-based super-resolution land cover mapping using support vector regression," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1271–1283, Apr. 2014.

[22] L. Wang, Y. Zhang, and J. Li. "BP neural network based subpixel mapping method," in *Proc. Int. Conf. Intell. Comput.*, 2006, pp. 755–760.

[23] L. Zhang, K. Wu, Y. Zhong, and P. Li, "A new sub-pixel mapping algorithm based on a BP neural network with an observation model," *Neurocomputing*, vol. 71, no. 10–12, pp. 2046–2054. Jun. 2008.

[24] F. Ling and G. M. Foody, "Super-resolution land cover mapping by deep learning," *Remote Sens. Lett.*, vol. 10, no. 6, pp. 598–606, Jun. 2019.

[25] Y. Jia, Y. Ge, Y. Chen, S. Li, G. Heuvelink, and F. Ling, "Super-resolution land cover mapping based on the convolutional neural network," *Remote Sens.,* vol. 11, no. 15, pp. 1815, Aug. 2019.

[26] P. V. Arun, I. Herrmann, K. M. Budhiraju, and A. Karnieli, "Convolutional network architectures for super-resolution/sub-pixel mapping of drone-derived images," *Pattern Recognit.*, vol. 88, pp. 431–446, Apr. 2019.

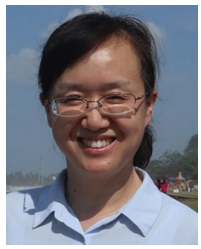
[27] C. Shang, X. Li, G. M. Foody, Y. Du, and F. Ling, "Superresolution land cover mapping using a generative adversarial network," *IEEE Geosci. Remote Sens. Lett.*, to be published.

[28] X. Ma, Y. Hong, Y. Song, and Y. Chen, "A super-resolution convolutional-neural-network-based approach for subpixel mapping of hyperspectral images," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, vol. 12, no. 12, pp. 4930–4939, Sep. 2019.

[29] X. Ma, Y. Hong, and Y. Song, "Super resolution land cover mapping of hyperspectral images using the deep image prior-based approach," *Int. J. Remote Sens.*, vol. 41, no. 7, pp. 2818–2834, Dec. 2019.
[30] D. He, Y. Zhong, X. Wang, and L. Zhang, "Deep convolutional neural network framework for subpixel mapping," *IEEE Trans. Geosci. Remote Sens.*, to be published, doi: 10.1109/TGRS.2020.3032475.
[31] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May. 2015.
[32] P. V. Arun, K. M. Buddhiraju, and A. Porwal, "CNN based sub-pixel mapping for hyperspectral images," *Neurocomputing*, vol. 311, pp. 51–64, May. 2018.
[33] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in *Proc. Int. Conf. Learn. Represent*. 2017, pp. 1–14.
[34] M. Fey, J. E. Lenssen, F. Weichert, and H. Müller, "Splinecnn: Fast geometric deep learning with continuous b-spline kernels," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 869–877.
[35] F. Monti, D. Boscaini, J. Masci, E. Rodola, J. Svoboda, and M. M. Bronstein, "Geometric deep learning on graphs and manifolds using mixture model cnns," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 5115–5124.
[36] Y. Lu, Y. Chen, D. Zhao, B. Liu, Z. Lai, and J. Chen, "CNN-G: Convolutional neural network combined with graph for image segmentation with theoretical analysis," *IEEE Trans. Cogn. Develop. Syst.*, to be published. doi: 10.1109/TCDS.2020.2998497
[37] 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. 2019.
[38] Q. Liu, L. Xiao, J. Yang, and Z. Wei, "CNN-Enhanced graph convolutional network with Pixel-and superpixel-level feature fusion for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, to be published, doi: 10.1109/TGRS.2020.3037361.
[39] S. Wan, C. Gong, P. Zhong, S. Pan, G. Li, and J. Yang, "Hyperspectral image classification with context-aware dynamic graph convolutional network," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 1, pp. 597–612, May. 2020.
[40] D. Hong, L. Gao, J. Yao, B. Zhang, A. Plaza, and J. Chanussot, "Graph convolutional networks for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens*, vol. 59, no. 7, pp. 5966–5978, Jul. 2021, doi: 10.1109/TGRS.2020.3015157.
[41] M. Defferrard, X. Bresson, and P. Vandergheynst, "Convolutional neural networks on graphs with fast localized spectral filtering," in *Proc. Adv. Neural Inf. Process. Syst.*, 2016, pp. 3844–3852.
[42] X. Yan, T. Ai, M. Yang, and H. Yin, "A graph convolutional neural network for classification of building patterns using spatial vector data," *ISPRS J. Photogramm. Remote Sens.*, vol. 150, pp. 259–273, Mar. 2019.
[43] D. K. Hammond, P. Vandergheynst, and R. Gribonval, "Wavelets on graphs via spectral graph theory," *Appl. Comput. Harmon. Anal.*, vol. 30, no. 2, pp. 129–150, Mar. 2011.
[44] Y. Yang and Y. Qi, "Image super-resolution via channel attention and spatial graph convolutional network," *Pattern Recognit,* vol. 112, , Dec. 2021, Art. no. 107798.
[45] Q. Wang, W. Shi, and L. Wang, "Allocating classes for soft-then-hard subpixel mapping algorithms in units of class," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 5, pp. 2940–2959, Jul. 2013.
[46] X. Xu, Y. Zhong, L. Zhang, and H. Zhang, "Sub-pixel mapping based on a MAP model with multiple shifted hyperspectral imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 580–593, Apr. 2013.
[47] A. J. Tatem, H. G. Lewis, P. M. Atkinson, and M. S. Nixon, "Increasing the spatial resolution of agricultural land cover maps using a Hopfield neural network," *Int. J. Geographical Inf. Sci.*, vol. 17, no. 7, pp. 647–672, Oct. 2003.
[48] J. Verhoeye and R. D. Wulf, "Land cover mapping at sub-pixel scales using linear optimization techniques," *Remote Sens. Environ.*, vol. 79, no. 1, pp. 96–104, Jan. 2002.
[49] E. L. Lawler and D. E. Wood, "Branch-and-bound methods: A survey," *Oper. Res.*, vol. 14, pp. 699–719, 1966.
[50] Y. Chen, Y. Ge, G. B. Heuvelink, J. Hu, and Y. Jiang, "Hybrid constraints of pure and mixed pixels for soft-then-hard super-resolution mapping with multiple shifted images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 5, pp. 2040–2052, Apr. 2015.
[51] M. Volpi and V. Ferrari, "Semantic segmentation of urban scenes by learning local class interactions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, 2015, pp. 1–9.
[52] T. E. Oliphant, "Python for scientific computing," *Comput. Sci. Eng.*, vol. 9, no. 3, pp. 10–20, Jun. 2007.
[53] A. Paszke *et al.*, "Pytorch: An imperative style, high-performance deep learning library," in *Proc. Adv. Neural Inf. Process. Syst.*, 2019, pp. 8026–8037.
[54] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Represent.*, 2014, pp. 1–41.
[55] H. He, B. DeZonia, and D. Mladenoff, "An aggregation index (AI) to quantify spatial patterns of landscapes," *Landscape Ecol.*, vol. 15, no. 7, pp. 591–601, Jan. 2000.
[56] T. Blaschke *et al.*, "Geographic object-based image analysis–towards a new paradigm," *ISPRS J. Photogramm. Remote Sens.*, vol. 87, pp. 180–191, Dec. 2014.
[57] P. Wang, G. Zhang, L. Wang, H. Leung, and H. Bi, "Subpixel land cover mapping based on dual processing paths for hyperspectral image," *IEEE J. Sel. Topics. Appl. Earth Observ. Remote Sens.*, vol. 12, no. 6, pp. 1835–1848, Apr. 2019.

Xining Zhang received the B.Eng. degree in geomatics engineering from the China University of Geosciences, Beijing, China, in 2016, and M.S.E. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2019. She is currently working toward the Ph.D. degree in cartography and geographic information system with the State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China.

Her current research interests focus on remote sensing image processing and intelligent super-resolution mapping.



Yong Ge (Member, IEEE) received the Ph.D. degree in cartography and geographical information system from the Chinese Academy of Sciences (CAS), Beijing, China, in 2001.

She is currently a Professor with the State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, CAS. Her research interests broadly focus on the statistical aspects of spatial and spatio-temporal data. She has authored more than 200 papers in the field of spatial statistics and related fields. She has been involved in the organization of several international conferences and workshops. She is currently an Associate Editor of computers and geosciences and the Editor board member of Spatial Statistics. She is also the Guest Editor of various special issues of journals.

Feng Ling received the B.S. degree in geology and the M.S. degree in geophysical prospecting and information technology from the China University of Geosciences, Wuhan, China, in 1999 and 2002, respectively, and the Ph.D. degree in system analysis and integration from the Huazhong University of Science and Technology, Wuhan, China, in 2006.

He is currently a Professor with the Innovation Academy for Precision Measurement Science and Technology, Chinese Academy of Sciences, Wuhan, China. He has authored more than 80 refereed journal articles. His research interests include super-resolution mapping and spatial–temporal fusion of remotely sensed imagery.

Dr. Ling was the recipient of the 2016 Boeing Award for Best Scientific Paper in Image Analysis and Interpretation from the American Society for Photogrammetry and Remote Sensing as the corresponding author. He has been supported by several talent programs, including the Hubei Province Natural Science Fund for Distinguished Young Scholars in 2013 and the Distinguished Young Scientist Grant of the Chinese Academy of Sciences in 2014.

Jin Chen received the B.A. and M.S. degrees in geography from Beijing Normal University, Beijing, China, in 1989 and 1992, respectively, and the Ph.D. degree in civil engineering from Kyushu University, Fukuoka, Japan, in 2000.

He was a Postdoctoral Researcher with the University of California, Berkeley, CA, USA, from 2000 to 2001, and with the National Institute of Environmental Studies, Tsukuba, Japan, from 2001 to 2004. He is currently a Professor with the Institute of Remote Sensing Science and Engineering, Faculty of Geographical Science, Beijing Normal University, Beijing, China. His research interests include remote sensing modeling and vegetation parameter retrieval through the inversion of remote sensing models.

Yuehong Chen received the B.S. degree from Hohai University, Nanjing, China, in 2010, and the M.Sc. and Ph.D. degrees from the State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, University of Chinese Academy of Sciences, Beijing, China, in 2013 and 2016, respectively.

He is currently an Associate Professor with the College of Hydrology and Water Resources, Hohai University, China. His current research interest includes remote sensing image processing.

Yuanxin Jia received the Ph.D. degree in cartography and geographical information system from the State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, University of Chinese Academy of Sciences, Beijing, China, in 2020.

He is currently with the Academy of Forest Inventory and Planning, National Forestry and Grassland Administration, Beijing, China. His research interests include natural resource inventorying and monitoring, remote-sensing image processing, and super-resolution mapping.