Distance Weight-Graph Attention Model-Based High-Resolution Remote Sensing Urban Functional Zone Identification

Article *in* IEEE Transactions on Geoscience and Remote Sensing · October 2021

DOI: 10.1109/TGRS.2021.3115972

Kui Zhang
China University of Geosciences (Beijing)

Dongping Ming
China University of Geosciences (Beijing)

Lu Xu
China University of Geosciences (Beijing)

Xiao Ling
China University of Geosciences (Beijing)

Distance Weight-Graph Attention Model based High Resolution Remote Sensing Urban Functional Zone Identification

Kui Zhang [a], Dongping Ming [a,*], Shigao Du [a], Lu Xu [a], Xiao Ling [a], Beichen Zeng [a], Xianwei Lv [b]

***Abstract*—The spatial arrangement of land-cover features constitutes different urban functional zone. With the same attributes of urban functional zone, the land-cover features that make up the functional zone will have similar spatial distribution characteristics. Considering the importance of understanding spatial relationships between land-cover features, the up-bottom hierarchical decomposition and semantic understanding of functional zone are achieved. Firstly, for Object Convolution Neural Network (OCNN) based land cover classification, an Equal-Area dividing algorithm is proposed to automatically generate convolution kernel position. Secondly, to extract spatial relationship features of urban land covers, a novel Distance Weight-Graph Attention Model (DW-GAM) is originally proposed for classifying urban functional zones by comparing the feature similarity of land cover relationship graph. Thirdly, considering the extreme difficulties in expressing the urban structure characteristic on single-scale, a recursive model which uses urban road network of different levels to divide multi-scale functional zones is built. Finally, taking the analysis of urban function allocation as the application objective, this paper establishes a primary framework of the spatial pattern evaluation index. Experimental results conducted on a Google Earth image of Xi'an City show that the multi-scale recursive model can accurately recognize urban functional zones by using the originally proposed DW-GAM. Then based on the outcome of urban functional zone identification, the case study of urban function allocation analysis is innovatively conducted on fine scale to give some suggestions on future urban planning, which is hence of great significance for urban function pattern analysis and urban planning.**

***Index Terms*—Urban functional zone, Distance Weight-Graph Attention Model (DW-GAM), Multi-scale recursive model, Urban function pattern**

I. Introduction

As a part of the urban complicated organism, urban functional zone is a geographical space that performs the urban function and it is embodied by the aggregation of relevant social resources. Urban functional zones with the same attributes often have the same requirements on the types and spatial allocations of land cover. However, due to the different spatial distributions of land cover, some human activities usually gather in the same one local space, thus various urban functional zones are formed. Therefore, accurately identifying urban land cover and the spatial relationship among land cover is helpful for modeling and spatial division of the urban functional zone.

VHSR remote sensing imagery can provide rich surface information of the earth, which makes it possible to extract urban land cover features accurately. The extraction accuracy of land cover features directly affects urban functional zones identification [1]. However, the spectral differences of the same land cover become larger, and the spectral characteristics of different land cover are more similar, traditional classification methods based on low and medium resolution images cannot be directly used for VHSR image classification [2]–[4]. Therefore, object-based image analysis (OBIA) method is proposed and widely used in VHSR images, which greatly improves the classification accuracy of land covers [5],[6]. Many studies applied OBIA in urban land cover classification with a set of low-level features (such as spectra, texture, shape) of the ground features, however it is difficult to express the deep-level features of land cover [7].

In order to fully express the complex features of land cover, the feature representation framework based on deep learning is widely used in VHSR images classification [8]–[10], CNN (Convolutional Neural Network) is one of the important representative models of deep learning method. Many studies classified remote sensing images by OCNN [11],[12]. The OCNN method can extract the deep features of the image including the global information, and keep the spatial detailed information [13]–[17].

The key of OCNN model is how to select convolution position in the segmented object, due to the convolution position affects classification accuracy. There are two main sampling methods: Zhang et al. [18] proposed the object convolutional position analysis (OCPA) method, which is mainly aimed at the linearly shaped objects. Lv et al. [12] proposed the random point generation algorithm, but the spatial distribution of random points is uneven. The reasonable

The work presented in this paper was supported by the National Natural Science Foundation of China (No.41671369), the National Key Research and Development Program (2017YFB0503600) and the Fundamental Research Funds for the Central Universities. (Corresponding authors: Dongping Ming)

Kui Zhang. School of Information Engineering, China University of Geosciences (Beijing), Beijing 100083, China (e-mail: zhangkui@cugb.edu.cn).

Dongping Ming. School of Information Engineering, China University of Geosciences (Beijing), Beijing 100083, China (e-mail: mingdp@cugb.edu.cn).

Xianwei Lv. State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China.

distribution of convolution position is helpful to improve urban land cover classification accuracy, and it lays a groundwork for identifying urban functional zone.

Recently, studies on extending graph based deep learning approaches have emerged in natural language understanding. This topological graph data is completely consistent with the spatial relationship between land cover features in the functional zone, which casts a light on urban functional zone identification. Graph-based deep learning models have been applied to a variety of analysis tasks, including the most common node classification [19], connection edge prediction [20]-[22], and graph classification [23]-[25]. However, graph classification is still in the exploratory stage due to the complexity of the own method and its application though it is good at extracting spatial relation features, thus applying graph classification in urban functional zone identification will be a promising way because the graph structure existing within the urban function zone can be effectively used.

Additionally, the urban functional pattern is closely related to the distribution of urban functional zone. The combination and interaction of the different functional zones determine the urban functional pattern, and the rationality of the urban functional pattern restricts urban functional zones formation [26],[27]. Reasonable planning of the functional pattern provides a strong basis for urban sustainable development [28],[29]. Current research on functional pattern mainly focuses on the allocation analysis of large-scale and multi-temporal [30]-[32], which ignores the fine scale functional pattern analysis. Therefore, it is a valuable work to extract the functional zone on fine scale and build up a fine-scale evaluation index system for urban functional allocation analysis, and then further put forward reasonable suggestions for meeting urban functional configuration requirements.

Based on the above analysis, this paper proposes an object based graph classification model, DW-GAM (Distance Weight-Graph Attention Model), which focuses on extracting spatial relation features between different objects to further extract multi-scale functional zones by jointly using multi-level road network and POI data with VHSR image. Especially, this paper establishes the systematic framework for urban functional pattern evaluation, which is further verified by a case study of urban functional zone identification and allocation analysis.

The remainder of this paper is organized as follows: Section II introduces the related work on identification of urban functional zones. Section III introduces the general workflow and the key components of the proposed methods. Section IV describes the study area and experiment results, followed by a discussion in section V. The conclusions are drawn in the last section.

II. Related Work on Identification of Urban Functional Zones

There are two main steps in urban functional zone identification: urban functional zone division and category classification [33].

A. *Urban functional zone division*

Urban functional zone division methods can be subdivided into single-scale division methods and multi-scale division methods. There are actually two approaches to single-scale division: dividing methods based on remote sensing image or vector auxiliary data. The first method uses image segmentation technology to generate urban functional zone units according to the spatial distribution features of image objects and the homogeneity of functional types [34]. Zhou et al. [35] proposed that the image was segmented into SOs (super object) by using single-scale division. The second method relies on vector auxiliary data such as road network and grid, road network is also used to divide functional zones at a single scale. However, there is no regulation on which level of road should be selected as the standard of urban functional zone division [36]-[38].

Du et al. [39] achieved the multi-scale division method through scale-adaptive approach and graph cuts algorithm. The method does not use vector auxiliary data, especially road network, which is very important for reflecting and evaluating urban functional pattern. In order to effectively express the hierarchical semantic information of urban functional zones, this paper will explore the significance of multi-scale functional zones division based on road network data.

B. *Urban functional zone classification*

For the attribute recognition of urban functional zones, the existing recognition algorithms mainly include clustering algorithm and deep learning. Clustering algorithm is mainly aimed at socially perception data with geographical location [40]-[42], such as POI. POIs reflect functional semantic information, and reveal the spatiotemporal behavior patterns of human beings within the city. Therefore, urban functional zones can be directly analyzed from social perception data. However, when there is no POI in the block unit, these methods cannot determine the attribute of functional zone.

Deep learning model, with the ability of extracting the deep features of urban functional zone, can be applied to identify functional zone. The CNN model can mine the deep image features and spatial information, which significantly improves the accuracy and calculation efficiency of the functional zone attribute recognition [43]-[45]. CNN can excavate the topological relationship characteristics between pixels within a specific window, but cannot express the spatial adjacency relation between land cover features on object level. However, the spatial attribute characteristics between adjacent land covers greatly influence the functional zone identification. Constructing a machine learning model capable of mining the spatial relationship features between land covers is the key of urban functional zone recognition.

Due to the deficiencies of currently functional zone division and classification methods, the spatial features between land cover are extracted by deep learning model, and the hidden social semantic features of POI are excavated. More details about the principle and methodology for functional zone dividing and identifying are presented in section III.

III. Methodology

Fig. 1 illustrates the workflow of urban functional zones identification which includes the following three steps：

Step 1, object-based land cover classification: VHSR image is segmented into image objects, then CNN model based on the proposed Equal-Area division algorithm is used to classify the image object to obtain the land cover classification results of the study area.

Step 2, top-down functional zone identification: the spatial distribution features of each block unit are extracted by DW-GAM, then combined with the urban semantic features represented by POI, functional zone categories of multi-scale block units can be finely identified level by level.

Step 3, urban functional allocation analysis based on the self-established assessment index system.

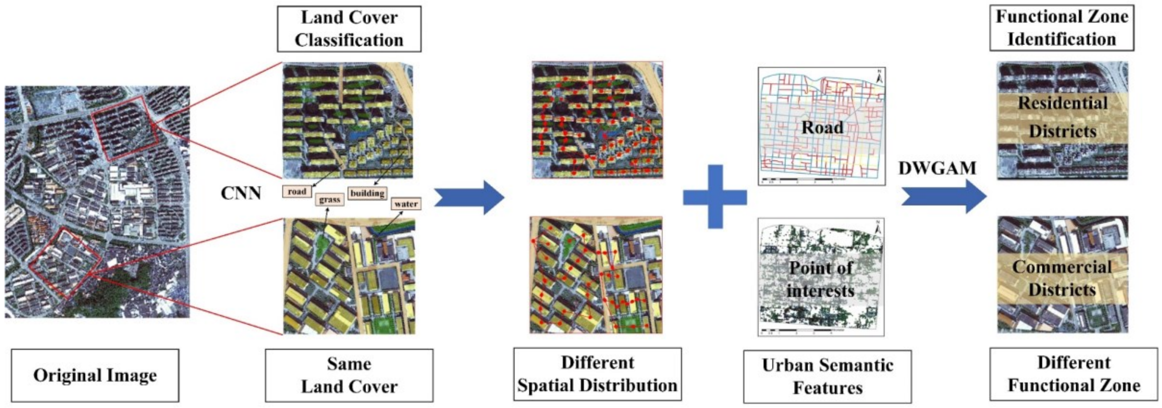


Fig. 1. Workflow of urban functional zones identification

A. Equal-Area Division Algorithm

In this paper, the OCNN method combining OBIA and CNN is introduced for urban land cover classification. As shown in Fig.2, the OCNN classification method comprises the following three steps: Firstly, the training samples are manually marked to train the parameter values of the OCNN model according to Google map and VHSR image. Secondly, VHSR image is segmented into image objects by segmentation method, and each object is processed by Equal-Area division algorithm, in order to generate convolution positions with reasonable spatial distribution inside the object. Finally, given an object, image patches are generated at each convolution position as test samples, and each patch is classified by the pre-trained model. The predicted category of each convolution position may be inconsistent, and the majority voting strategy is used to determine the attribute of the object (The final attribute of the object is determined by the largest number of predicted categories).

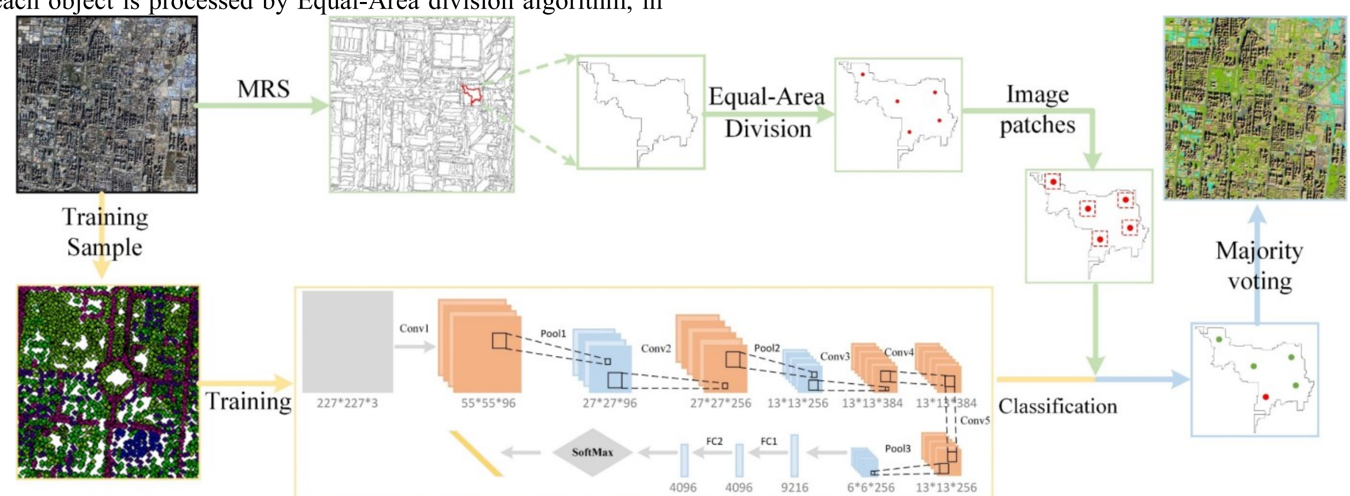


Fig.2 Flowchart of the OCNN method (The green points (p1, p2, p3, p5) represent the correctly classified category points; and one red point (p4) represents misclassified points)

The selection of convolution kernel position affects the classification accuracy of OCNN. The normally used random point generation algorithm proposed for RMV-CNN model is suitable for objects of various shapes, but the convolution kernel position may fall on the object edge, which cannot accurately extract the object features.

The OCPA was proposed based on the moment bounding (MB) box of each object to identify the position, by which the convolution position is distributed on both sides of the MB major axis. However, an idealized convolution position should meet the following requirements: first, it must be evenly and reasonably distributed among each object. Second, the algorithm for generating convolution position should adapt to image object with any shape.

Thus this paper proposes an iterative sampling strategy based on idea of equally dividing the object area. The algorithm realizes the object hierarchical division according to the object area, as shown in Fig. 3. Before dividing the secondary objects, it is necessary to judge whether the objects are distributed horizontally or vertically in order to make the subdivided object's shape relatively consistent. Each iteration will generate equal-area subdivided objects (one divides into two) until the specified even number is generated. Finally, the centroid points of all subdivided objects are generated, thus all convolution position points with odd numbers (including the centroid point of the original image object) are generated. Fig. 3 illustrates the generated convolutional kernel positions (the points with blue color) of OCNN for the case of N = 5. The Equal-Area division method is applied only to the testing phase in the study.

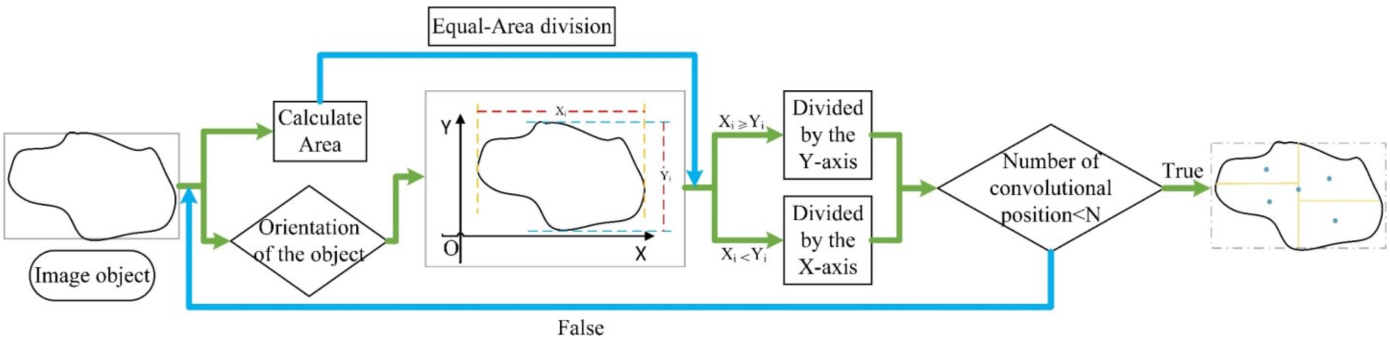


Fig. 3. Workflow of Equal-Area division algorithm

B. DW-GAM Based Urban Functional Zone Identification

1) DW-GAM

Graph-based deep learning model has made some achievements in chemistry, medicine, and other fields. Compared with other deep learning models (such as CNN), it can learn the topological relationship between nodes and non-topological attribute features (such as distance, azimuth, etc.) between adjacent nodes. Learning the deep-level spatial distribution features between land covers are necessary for classifying urban functional zone, because different urban functional zone presents specific spatial structure. Therefore, this paper attempts to introduce RNN based deep learning model for graph classification. The novel RNN model, called the Graph Attention Model (GAM) [46] predicts graph labels by traversing the nodes in part of the graph, is involved in urban functional zone identification for the first time.

For identification of urban functional zone, the two factors should be taken into account: there is a strong spatial autocorrelation between land cover features in urban functional zone; and features with spatial topological structure should satisfy the first law of geography, distance is an essential factor for the correlation between land covers, the closer the distance between land covers, the greater the correlation. Therefore, adding the distance weight to the GAM based step module, this paper proposes the DW-GAM architecture as shown in Fig. 4.

DW-GAM consists of three parts: Step Module, LSTM Module and Network Module. The input of step module is each topology data $G_i(\mathbf{A_g}, \mathbf{D_g}, \mathbf{W_g}) = \{g_i, l_i\}$($g_i$ refers to i-th graph and is comprised of an adjacency matrix **Ag**, attribute matrix **Dg**, distance weight matrix $\mathbf{W_g}$ and the rank vector r_{t-1} (a stochastic rank vector, prioritizing those whose type (i.e., node label have higher rank in r_{t-1}); l_i is graph label in the i-th graph) . Through training and learning, the output of the step model is the embedding vector(s_t) which represents information captured from the current step we took.

The information vector s_t output by step module, as the input vector of LSTM module, will update the historical vector h_t. The h_t is the latest feature vector after representation of graph features, and the model will constantly update the feature information over time. The h_t can be defined as

$$h_t = y_h(s_t, h_{t-1}; \theta_h) \tag{1}$$

where h_{t-1} is the historical vector generated at the previous time, which is equivalent to a feature vector of short-term memory, and θ_h is the training parameter.

The newly updated history vector h_t has two functions: firstly, a new rank vector r_t is defined as $r_t = y_r(h_t; \theta_r)$,r_t will be used as input vector for the next step module. Secondly, h_t is used as the input vector of the Network module. The network module is designed to predict graph label (l_t) with the help of the latest history vector h_t. The function of predicting graph label can be written as $l_t = \arg\max_i P(y = i \mid y_l(h_t; \theta_l))$.

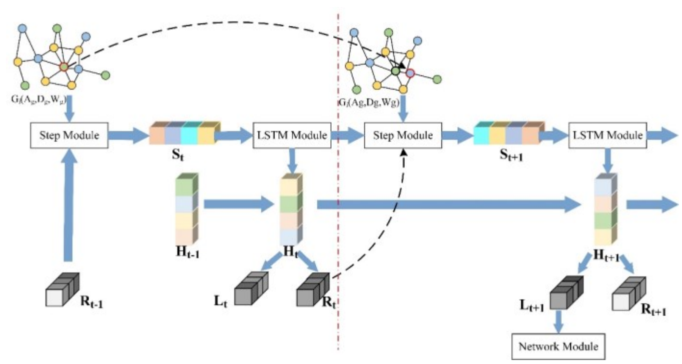


Fig. 4. The structure of DW-GAM

The main goal of the step module is to select the next node to be moved. The method of traversal nodes used in the step module of GAM is to find out the type of the node that occurs the most frequently among adjacent nodes. In this paper, the calculated distance weights are input into the step module to enable DW-GAM to find the nearest neighboring node among adjacent nodes.

Taking the case of only 3 nodes as an example, Fig. 5 illustrates the workflow of the node traversal. Firstly, the main land covers from the VHSR image classification results in each block unit are extracted. Then the land covers in each block unit are treated as nodes of the graph data, calculating the adjacency matrix (in which the number "1" represents adjacent and the number "0" represents not adjacent), attribute matrix (recording the land cover type of each node, for example, the land cover type of node 1 is 0110) and distance weight matrix (in which d1 denotes the spatial distance from node 1 to node 3, and d2 denotes the spatial distance from node 1 to node 2) for the nodes in each block unit. The distance weight matrix is to find the nearest node to the current node. Secondly, the three matrixes are input into the step module, the spatial structure features of urban land covers can be acquired by learning the spatial distribution patterns among different land covers, and then generate the new node output vectors of the module.

Compared with the computational process of GAM, the distance weight is time-consuming stage for DW-GAM. The GAM gives rise to a time complexity of *O(N)*, where N represents the number of neighbor nodes. DW-GAM needs to find the nearest node by traversing the neighbor nodes again. Each node has an acceptable number of neighbor nodes, so it does not have a significant impact on the performance of the model.

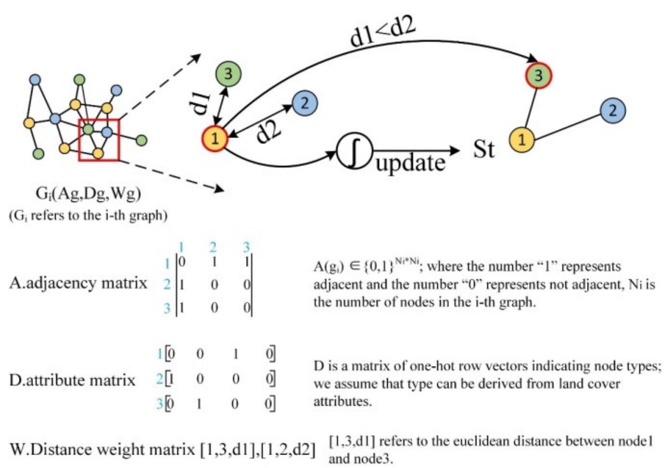


Fig. 5. Schematic diagram of the Step Module

2) *Urban Functional Zone Classification*

Previous studies of determining urban functional zone categories are mainly relying on POI, DMSP/OLS nighttime light image, GPS trajectory and other multi-source data to recognize urban functional attributes [47]-[50]. Studies have extracted urban semantic features of POI data to classify urban functional zone through clustering algorithm or mathematical formulae [51],[52]. Some research methods recognize functional zone by integrating remote sensing image and POI data [53]-[55]. Most previous studies only relied on the urban functional semantics of multi-source data [56], however failed to mine the spatial relationship features between land covers in VHSR images [57],[58].

Urban functional zone contains multiple land cover categories with different spectral and spatial features. Sometimes the same land cover may exist in different urban functional zone, however the spatial relationships or spatial arrangement of land covers is different. Given the land cover diversity in VHSR images and land cover spatial distribution variability, it is necessary to build a more robust feature learning model which can learn the spatial structure relations among land covers and thus contribute to accurate functional zone attribute identification. Therefore, DW-GAM is selected as the main model framework for the attribute identification of urban functional zones.

POIs are introduced as auxiliary data for urban functional zone classification. POI data contains a large amount of urban functional semantic information, which can be used to quantitatively identify urban functional zone. However, the number of POIs contained in each functional zone is not evenly distributed [59]. For example, the amount of POIs of commercial types is significantly more than other types, and even no distribution of POI data appears in some blocks. Moreover, it is difficult to guarantee the accuracy of the POI attributes in some blocks, so it is not feasible to only rely on POI to recognize urban functional zone because open access unstructured urban data can be biased since there is uncertainty around the data's representativeness [60]. In this paper, the quantitative identification method of POI data is another identification method of functional zones. We compare the consistency of urban functional zone identification results between by DW-GAM method and by POI data.

Considering that the number of POIs within different functional attributes is different. We identify functional zone attributes by constructing frequency density (FD) and category ratio (CR) for a block unit according to formula (2) and (3).

$$FD_{\mathrm{i}} = \frac{n_i}{N_i} \quad (i = 1,2,3,4,5) \tag{2}$$

$$CR_{\mathrm{i}} = \frac{FD_i}{\sum_{i=1}^{5} FD_i} \quad (i = 1,2,3,4,5) \tag{3}$$

where i refers to the *i-th* type of POI in the block unit; n_i is the total number of *i-th* type of POI in the block unit; N_i is the total number of all POI in the block unit; F_i refers to the frequency density of *i-th* type of POI in the block unit, and C_i is the category ratio of *i-th* type of POI in the block unit. The CR value 50% is recommended as the standard to judge the functional categories of each block unit.

Meanwhile, considering the multi-level road network provided the feasibility of fine extraction of urban functional zone, a multi-scale recursive model of functional zone identification was proposed by combining DW-GAM and POI quantitative methods. Fig. 6 presents the multi-scale recursive model of functional zone identification. We divide the block units using road network data from the large-scale to the small-scale. We use two methods to distinguish the attributes of each block unit: DW-GAM and POI quantitative method (formula 2 and formula 3). Correspondingly, DW-GAM is employed to determine the urban functional zone categories of block units on each scale. In this procedure, POI based voting is used to optimize the DW-GAM based recognition, and the voting results determine which blocks will be divided into sub-blocks on the next scale. This recursive model can recognize urban functional zones for block units on different scales.

C. Multi-Scale Urban Functional Zone Division

A variety of auxiliary data are used for the division of urban functional zones, such as Grid Data [61], Cadastral Data [62],[63] and Road Network Data [38],[64], among which block unit formed by road network is the most widely used in functional zone division because it is closest to the boundary of urban functional zone and easy to be acquired [36],[37],[65].

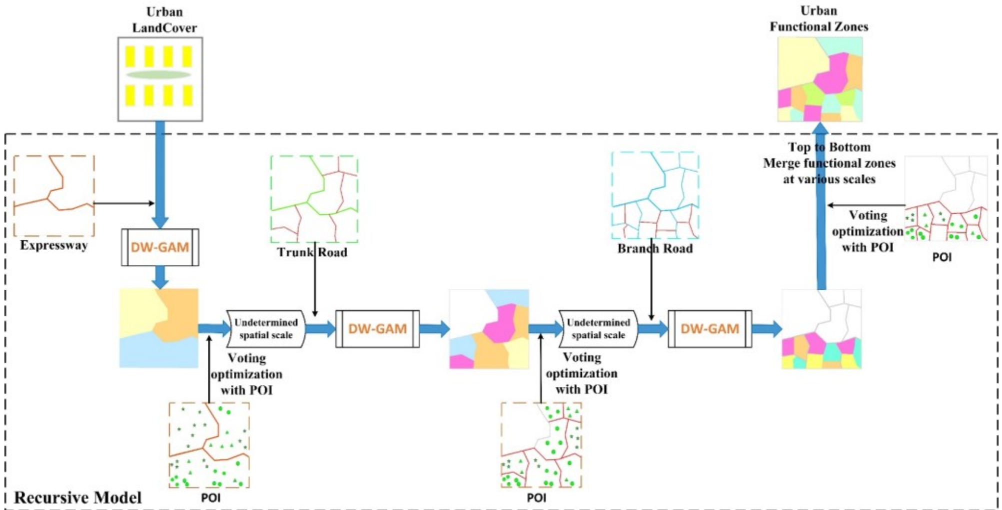


Fig. 6. Multi-Scale recursive model of functional zone identification

The urban functional zone division method based on road network takes the actual urban road distribution into consideration, however, a spatially and semantically independent functional zone may be over-divided in the road-dense area, which increases the complexity of division. Therefore, a multi-scale functional zone division method is proposed by using multi-level urban road network data, which decreases the computation of generating block units.

After calculating all CRs for all POI types within each block unit, the block units with CR value less than 50% will be further divided by the next level road network until the new CR value is more than 50%. Then the urban functional zone attribute can be determined by calculating the attribute similarity between CR based judgement and DW-GAM based recognition. As illustrated in Algorithm 1, the rules for determining multi-scale functional zone division are as follows.

Algorithm 1 Multi-scale functional zone division algorithm

1: Input: r_i, r_{i+1}, **M**, POI (r_i is road of i-th scale, **M** is DW-GAM identification result of i-th scale)
2: Initialize: correct_graph (an empty list to store geospatial unit ID without next scale division)
question_graph (an empty list to store geospatial unit ID require next scale division)
question_featureclass(a new featureclass to store next scale road)
3: Process:
4: for j=0 to n do (n is the total number of road levels)
5: for k=0 to p do (p is the block unit number of j-th scale)
6: Calculate **C[j][k]** and **M[j][k]** (**C[j][k]** is CR value of k-th block unit at j-th scale, **M[j][k]** is DW-GAM result of k-th block unit at j-th scale)
7: if (Max(**C[j][k]**)≥0.5 && $\mathbf{M_j}$ == F(Max(**C[j][k]**)) or Max(**C[j][k]**) == null) then
8: append(**M[j][k]**) to correct_graph
9: else append(**M[j][k]**) to question_graph
10: question_graph generate question_featureclass
11: question_featureclass clip r_{i+1} generate new_r, and as a new model input (new_r is the graph data generated under the new scale)
12: Output: new_r

1) For the block units with CR value more than 50%, if the functional attribute determined by CR is consistent with the functional identification result by DW-GAM, the functional zone attribute of the block unit will be easily determined, and this block unit will not be divided at the next scale.

2) For the block units with CR value more than 50%, if the functional attribute determined by CR is inconsistent with the functional identification result by DW-GAM, the block unit will be further divided by the next level road network until the achieved two functional attributes are consistent.

3) For the block unit does not contain POI (CR is a null value), the block unit is called a null value unit, the identification result of DW-GAM will be the terminal functional zone category of the block unit, and the block unit also will not be divided at the next scale.

As shown in Fig. 6, the DW-GAM identification results and the calculation results of CR are combined. The partition of the next scale of block units depends on whether the identification results of the two methods are consistent, so as to realize the top-down hierarchical division from large-scale road network to small-scale road network. On the one hand, this method completes the mutual verification of the two identification results, which ensures the extraction accuracy of the functional zone; on the other hand, it reduces the unnecessary partition of some blocks, which greatly reduces the number of DW-GAM graph data sets and improves the operation efficiency of the model.

D. Evaluation System of Urban Function Pattern Allocation

Urban functional distribution in urban space forms the urban spatial structure which also reflects the spatial manifestation of the city's society and culture in the region [66]. Therefore, the spatial pattern of urban functions determines the social and cultural development direction of the city, and a reasonable function configuration is important to promote urban sustainable development [28],[67].

Urban function configuration research is based on rough scale urban pattern analysis, which results the analysis is consequently rough. Based on the urban functional zone identification results from VHSR images, enlightened by landscape pattern analysis, this paper establishes the evaluation index of urban functional pattern as shown in Fig. 7. A series indexes of functional pattern and functional service is established from two perspectives, the spatial distribution of the same attribute functional zone and the spatial correlation between different attribute functional zones. By using the established evaluation index of urban functional pattern, it is available to realize the preliminary functional allocation analysis and provide rational urban planning suggestions for the future urban development.

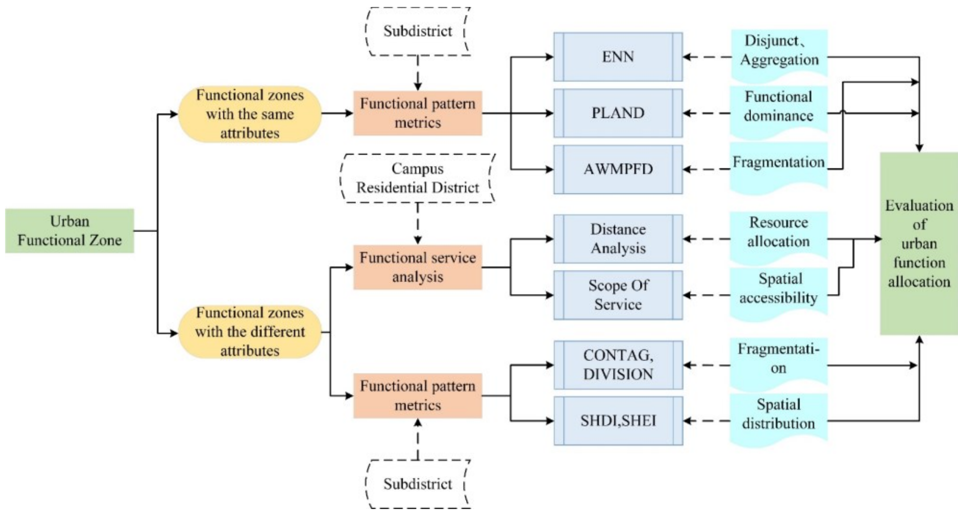


Fig. 7. Framework of urban function allocation analysis (ENN: Euclidean Nearest Neighbor Distance; PLAND: Percentage of Landscape; AWMPFD: Area-weighted mean patch fractal dimension; CONTACT: Contagion Index; DIVISION: Landscape Division Index; SHDI: Shannon's Diversity Index; SHEI: Shannon's Evenness Index)

IV. Experiment

The experiments were performed on an Ubuntu 16.04 OS with a CPU (3.4 GHz core i7-6700), RAM (8 GB), and GPU (NVIDIA GTX 1060 2 GB). PyTorch was chosen as the deep learning framework.

A. Experimental Data

As shown in Fig.8(a), the experimental data, with spatial resolution of 0.6 m and size of 12246 × 11002 pixels, was downloaded from Google Map. The image was captured on 21st July 2019 and was located in the urban area of Weiyang District, Xi'an City, China.

Two images, Fig.8(b) and (c) with spatial resolution of 0.6 m and size of 9320 × 12427 pixels, 13502 × 9527 pixels, respectively, were used to generate samples of graph dataset for DW-GAM.

POI data records the activity types of urban residents in a certain place, including living, working, dining, shopping or entertainment. We took 32840 POIs as the vector auxiliary data (Fig.8(d)), which were downloaded from Baidu map and divided into five categories: residential services, commercial services, education services, scenic spots and shanty towns.

Fig.8(e) showed the vector data of three-level road network obtained from Open Street Map (OSM), including expressways, trunk roads and branch roads.

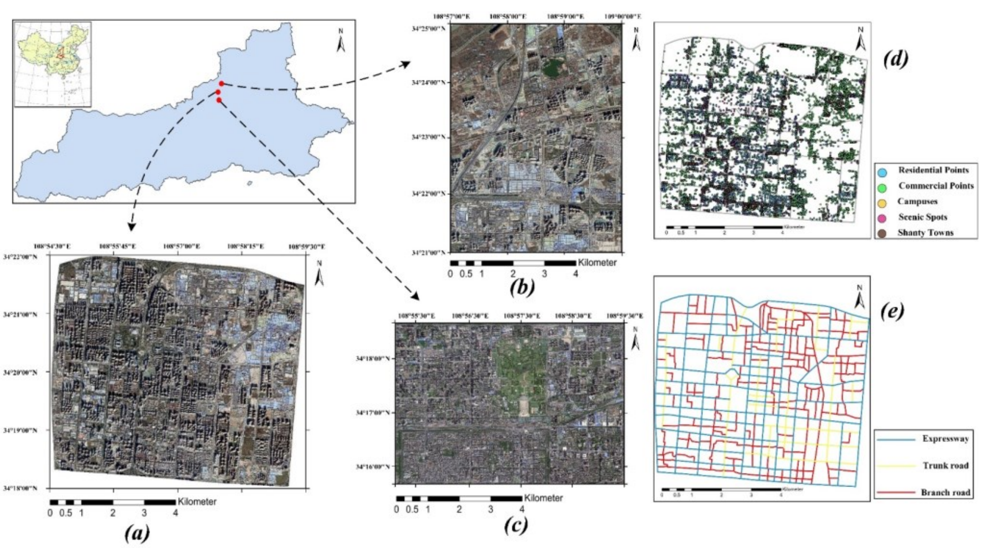


Fig. 8. Experimental data. (a) VHSR image for model predicting; (b),(c) VHSR image for model training; (d) Point Of Interests (POIs); (e) OpenStreetMap (OSM).

B. Urban Land Cover Classification based on Equal-Area division algorithm

1) Equal-Area Division Algorithm

The proposed Equal-Area division algorithm is sensitive to the segmented regions. In order to acquire a better segmented boundary, we chose the multi-resolution segmentation (MRS) algorithm. The segmentation parameters were determined by experience. For the Xi'an image, the scale parameter was set as 30, while the shape and compactness values were set as 0.2 and 0.4, respectively.

Based on the segmentation results, positions of convolution for each object were generated by Equal-Area division strategy. As the Equal-Area division results cannot be shown clearly here, this paper only demonstrated part of the results of convolutional positions. Fig. 9 showed the local results of the Equal-Area division algorithm and random points generation algorithm, respectively.

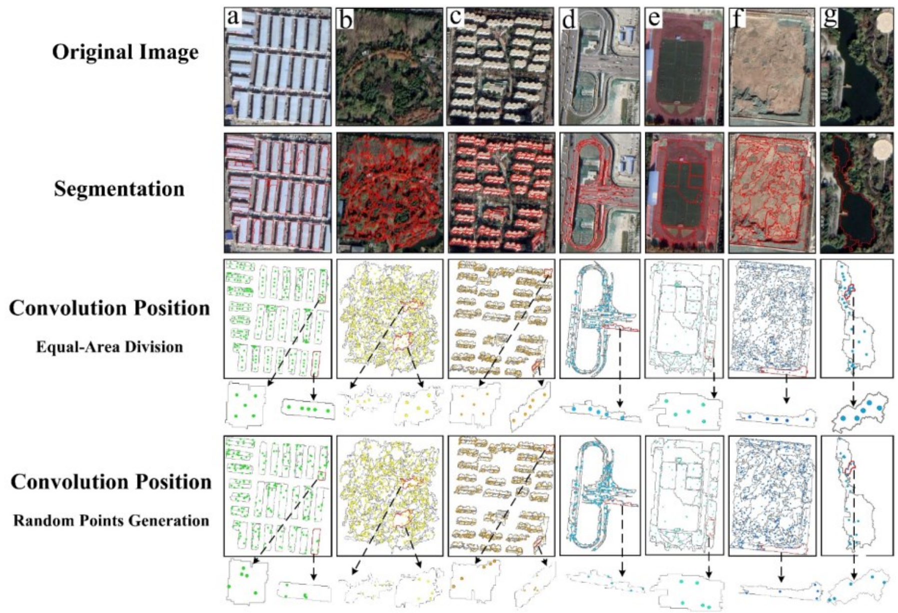


Fig. 9. Local results of convolution position. (a: building with blue roof; b: grass; c: other building; d: road; e: artificial turf field; f: bare soil; g: water)

Fig. 9 showed the local results of seven different land covers. There were some problems in the convolutional positions extracted by the random point generation algorithm. For example, some points were too close, some points were distributed on the edge of the object, so CNN was limited to mine local features of the image object. The points generated by equal-area division algorithm were suitable for objects of different shapes, and ensured that they were evenly distributed in the objects. Therefore, the features of each object can be extracted more completely by CNN, which will improve the classification accuracy to a certain extent.

2) Urban land Cover Classification

AlexNet, a fundamental CNN framework, was introduced to classify urban land cover. Some parameters need to be clearly explained. For training the CNN model, the learning rate was set as 0.01 to control the progress of learning model. The epoch is the number of the times that entire training datasets were traversed by the model. Here, the epoch of the framework this paper used was set to 100. Avoiding over-fitting of the deep learning model, the dropout rate was set as 0.5. The batch is a part of the training datasets fed into the model each time, and batch size is the number of training samples in each batch. We set the batch size to 100.

As shown in Fig. 10, after sorting out the classification results, five land cover features, namely building with blue roof, other building, grass, artificial turf field and water body, were selected to identify urban functional zones.

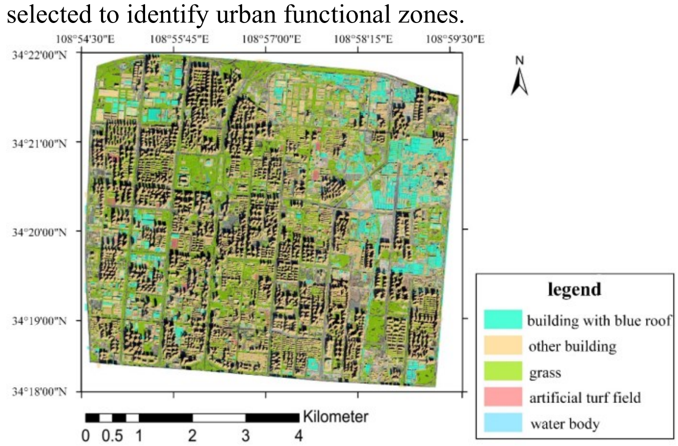


Fig. 10. Urban land cover classification result.

C. Identification of urban functional zones based on DW-GAM

1) Training Data Sampling

In this paper, the categories of urban functional zones are determined according to the code for classification of urban land use and planning standards of development land which is issued by the Ministry of Housing and Urban-Rural Development of the People's Republic of China. A label sample is composed of attribute matrix, adjacency matrix, weight matrix and its functional label. The training samples of each category were selected by means of POI data and online Baidu

Map (https://map.baidu.com). To enhance the robustness of the DW-GAM, we made training samples by random sampling from the b and c images in Fig. 8, 183 graph datasets (including 23,816 nodes) were generated for training the graph classification model, the sample sizes of commercial district, campus, residential district, scenic spot, shanty down are 45,16,87,15,20, respectively. Fig. 11 illustrates representative samples (image patch) of each functional zone category.

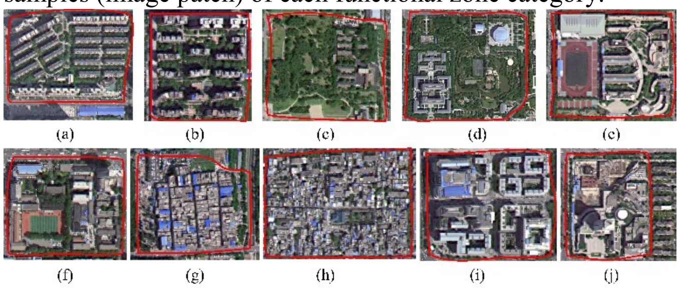
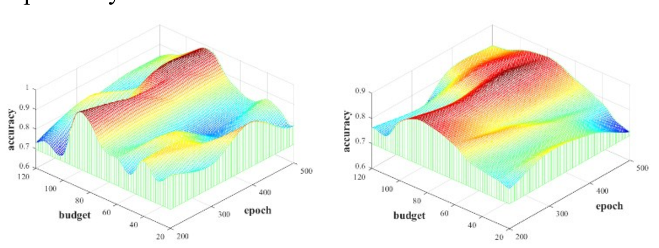


Fig. 11. The selected 10 representative samples. (a)~(b) are residential district, (c)~(d) scenic spot, (e)~(f) campus, (g)~(h) shanty down, (i) and (j) commercial district.

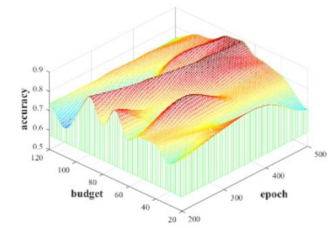
2) Parameters Setting for DW-GAM

In order to extract the deep spatial features of each graph, DW-GAM needed to set different parameters like CNN. The learning rate that controlled the progress of DW-GAM was 0.01. In addition, through a large number of experiments, it was concluded that epoch and budget have a great influence on the experimental results. The size of the epoch is to extract the deep spatial features of graphs and prevent over fitting. The budget determines the number of nodes that be traversed in each epoch. If the number of traversed nodes is small, DW-GAM cannot completely mine the deep-level features of the graph. Rather, if the number of traversed nodes is large, the model will learn the features of some fuzzy nodes, which will have an inappropriate impact on the classification accuracy.

For the multi-scale functional zones identification method, due to the inconsistency of the size of the graphs at different scales, the amount of the nodes contained in each graph is different, but the size of the graphs at the same scale is relatively consistent. We made an analysis on budget and epoch at three scales. According to the results in Fig. 12, the values of the budget and epoch parameters corresponding to the expressways, trunk roads, branch roads were set as 100, 500; 80, 400; 70, 400, respectively.



(a) Express way scale (b) Trunk road scale



(c) Branch road scale

Fig. 12. Changes of different scales with the budget values varying from 20 to 120 and the epochs varying from 200 to500.

3) Multi-Scale Urban Functional Zone Identification Results

Scale is the main issue in the urban functional zone identification. Functional zones with the same attributes may have block units of different scales. This is because of the difference in urban planning and design. For example, some residential districts are bounded by the expressways, these residential districts are relatively large in scale compared with the residential districts bounded by branch roads. Therefore, we adopted a multi-scale functional zones identification method to find the best division scale for each functional zone as far as possible, and to ensure that the functional attribute features and spatial features of one block unit are relatively homogeneous.

Fig. 13 showed the recognition results of functional zones with expressway as the first level of division scale. Expressway is mainly responsible for the long-distance and rapid transport services in the city, the division scale is relatively large. The study area was divided into 58 first-level block units, of which only 18 block units were successfully identified by the multi-scale model, and other 40 block units failed to reach the threshold of the similarity calculation. It is shown that the functional attributes of these 40 block units were highly heterogeneous. There were multiple categories of functional zones inside the block units, and these first-level block units needed to be segmented at the next scale.

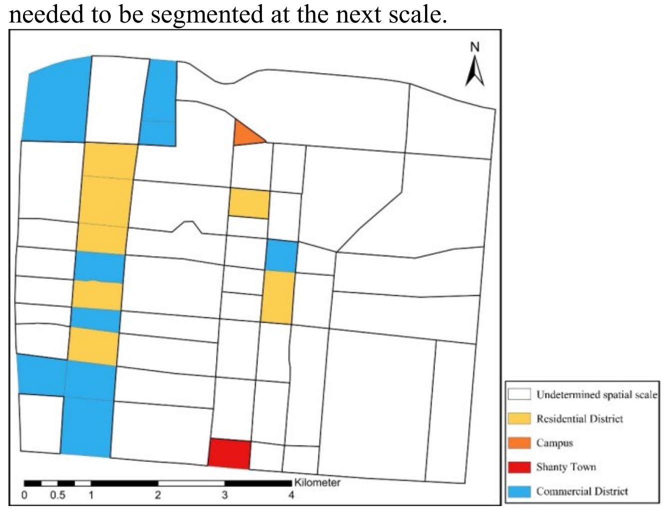


Fig. 13. Functional zone identification results at the express way scale

Taking the trunk road as the dividing boundary, a total of 88 second-level block units were divided, of which 51 second-level block units were successfully identified. The remaining 37 second-level block units needed to be divided into the next scale

(Fig. 14). Most of the residential districts in the second-level block units were identified, which also indicated that many residential districts in urban planning were designed with the trunk roads as the boundary. In addition, the trunk roads with traffic function are used to connect the main districts of the city, and residential districts are generally located near streets with convenient traffic.

Fig. 14. Functional zone identification results at the trunk road scale

Fig. 15 showed the recognition results of the last-level block unit divided by branch roads. There were 101 third-level block units identified by DW-GAM, and commercial districts and campuses were identified at the smallest scale. Commercial districts mainly contained some companies and factories. The area of the functional zones was relatively small. Therefore, the smallest scale branch roads were required to divide the functional zones unit.

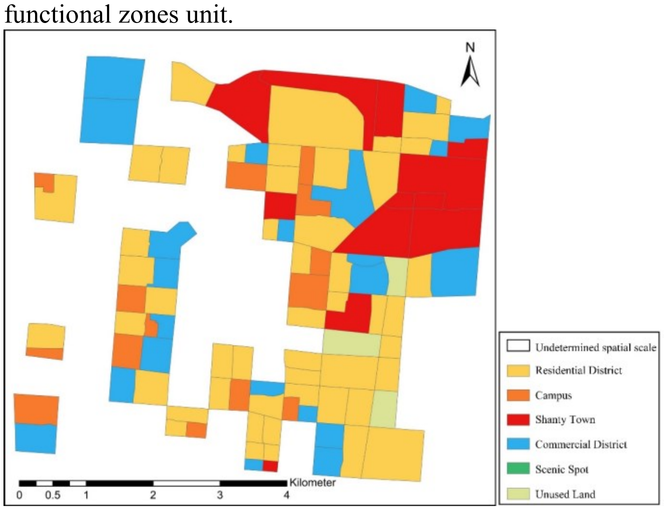


Fig. 15. Functional zone identification results under the branch road scale

Finally, we combined the classification results of the above three scales to obtain the overall urban functional zones identification results of the study area, as shown in Fig. 17. The proposed algorithm not only considered the heterogeneity of land covers in the block units, but also divided the urban functional zones from the large-scale roads network to the small-scale roads network, and realized the classification of multi-scale urban functional zones from top to bottom. A total of 170 block units in the study area were classified. For each category of urban functional zones, about 40% block units of the classification results were extracted, and the results were tested with the true value of the Google electronic vector map. The confusion matrix was calculated as shown in Fig.16, the overall classification accuracy was 85.2%.

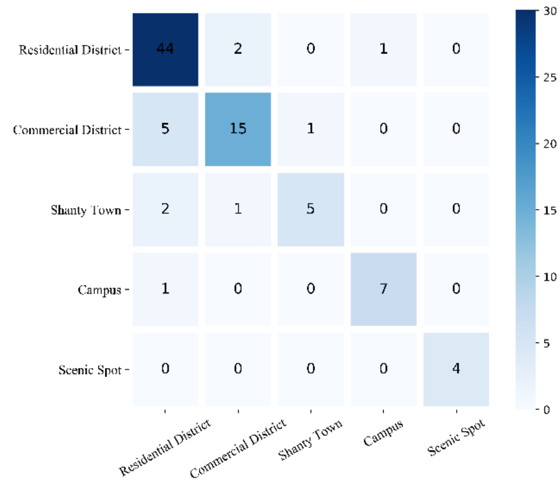


Fig.16. Confusion Matrix

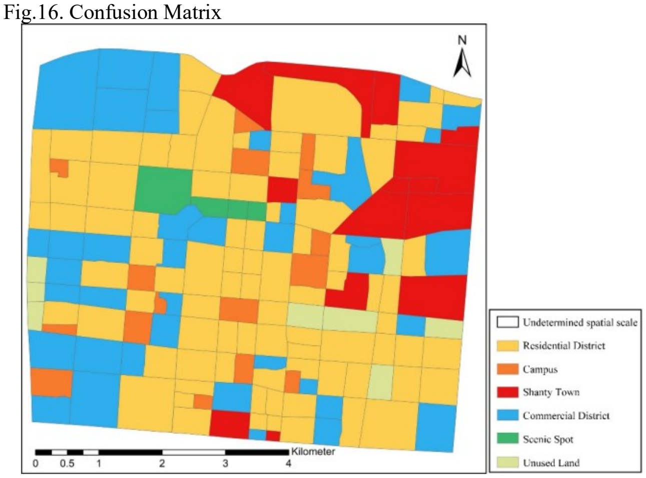


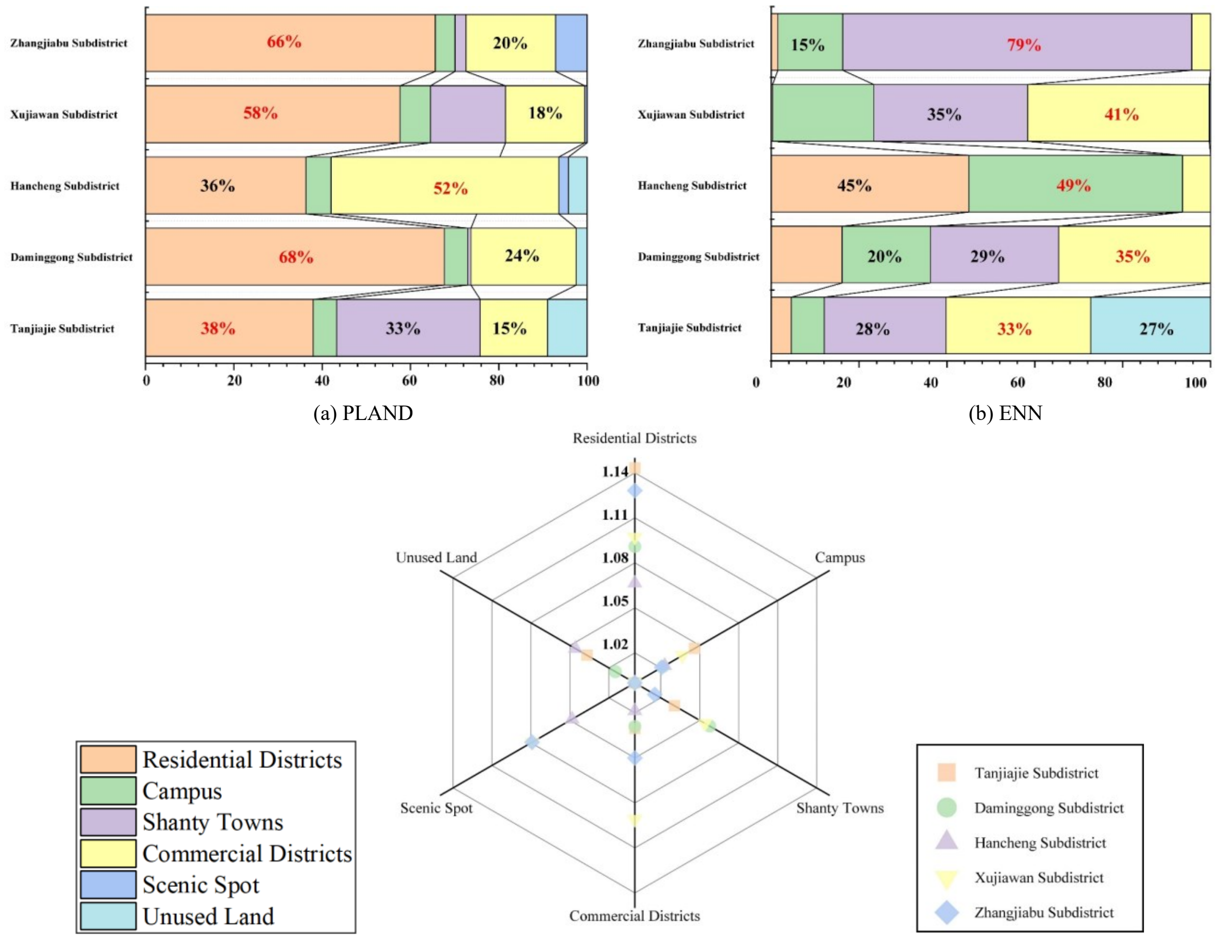
Fig. 17. Functional zones identification results

D. The analysis of Allocation for Urban Function Pattern

For the purpose of analyzing allocation of urban function reasonably and elaborately, study area was subdivided into five blocks, therefore the evaluation of spatial distribution that functional subdistrict corresponding to each block was derived.

1) Allocation analysis of Same Functional Zones

Each subdistrict owns its urban function distinctively, equipped with corresponding fundamental functional zones, meanwhile has its representative domination functional zones. As well-developed tourism zones would locate in some certain subdistricts while advanced commercial zone in another. Therefore, function patter indexes were calculated in this section, for the purpose of analyzing the aggregation degree of a total 5 subdistricts along with their domine attributes.



(a) PLAND (b) ENN

(c) The legend of Fig.16. (a) and (b) (d) AWMPFD

Fig. 18. Urban function pattern index results

The index PLAND represents for the proportion of landscape types, was calculated for analyzing regional dominant urban function, Fig.18(a) shows that except for the Hancheng subdistrict, the proportion of residential district appeared to be the highest in any other zones. Tanjiajie subdistrict locates in the east of study area where a large number of factories exist as well, shantytown proportion rises relatively. All five subdistricts contain the corresponding fundamental infrastructures such as campus, however, when analyzing the spatial rationality for such, PLAND is considered less helpful. Therefore, as shown in Fig.21, the functional service index is involved in order to solve this problem.

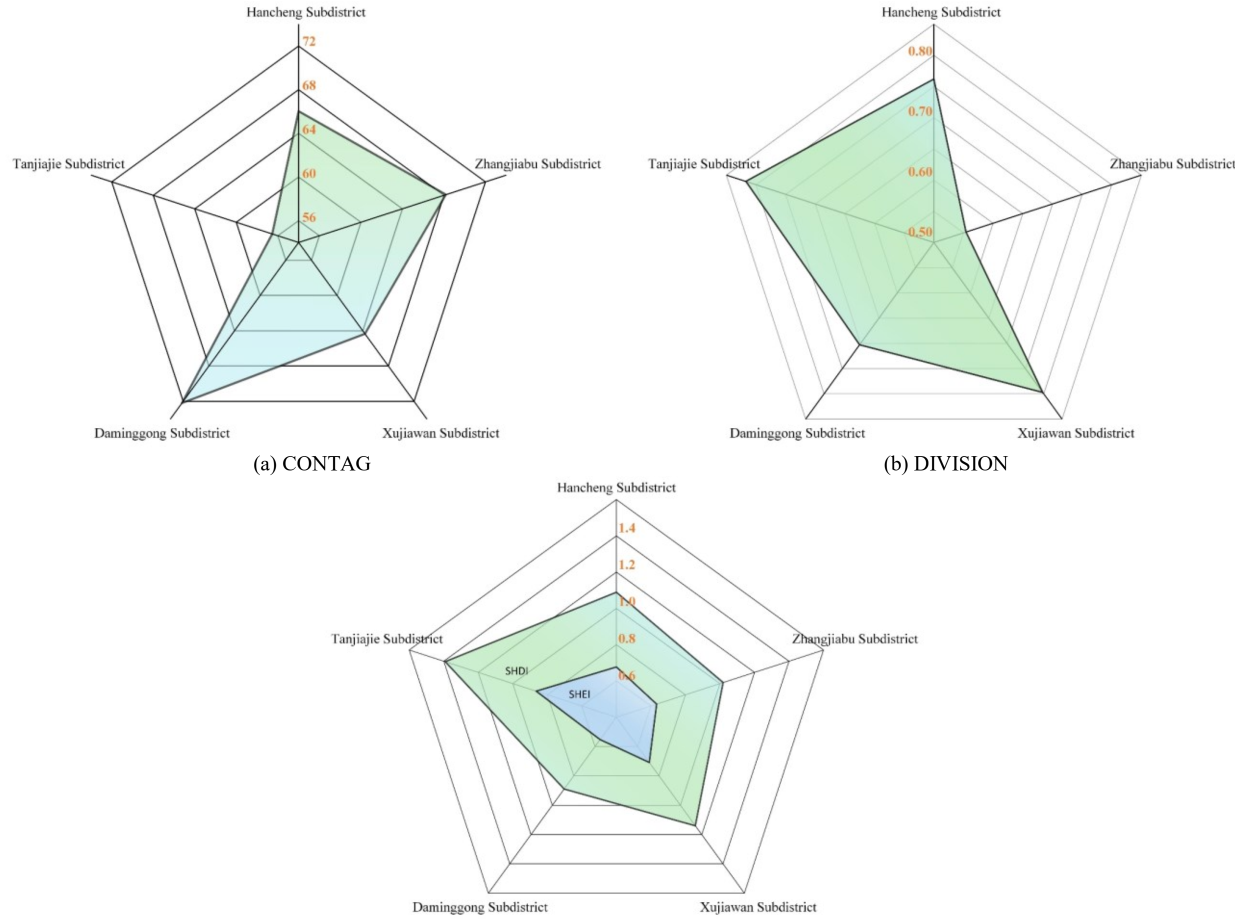
The spatial pattern of functional zones can be manifested by the MNN index. As is shown in Fig. 18(b), The commercial district shows higher MNN, indicating that there might be a distance between each district, resulting in dissociation in spatial pattern. The commercial district shows higher MNN, suggesting that there might be a distance between each zone, result in dissociation in spatial pattern. The residential districts in Xujiawan show the lowest MNN, which indicates that the residential districts are reunited. The residential district that obtains the highest MNN is located in Hancheng subdistrict, with 2 zones separately locate. In this case, proposing real estate development in nearby region is recommended.

AWMPFD is a measurement of spatial shape complexity. Fig. 18(d) shows that the residential district reaches the overall highest AWMPFD, for the more human-activity is involved in a certain functional zone, the higher AWMPFD would be. The AWNPFD of commercial districts and residential districts are both relatively high in Xujiawan and Zhangjiabu, suggesting that the urbanization progress of the study area is developed from the center to the two-way direction.

2) Allocation Analysis of Different Functional Zones

To evaluate whether an urban is well-developed or not depends on multiple aspects, such as whether it's equipped with complete fundamental facilities, and whether the functional zones have played their respective role. For example, the service scope of the campus is considered to cover each residential district while the functional zone spatial pattern of each subdistrict should be reasonable. In order to better analyze the rationality of urban functions, taking the purpose of building the spatial connection of different functional zones, the indicators for analyzing the connection between functional zones was proposed (Fig. 19).

Tanjiajie subdistrict achieves low CONTAG index and high DIVISION index, indicating that connectivity was poor and fragmentation was high among functional zones, which further suggests that Tanjiajie maintains a relatively high urbanization level in the study area. SHDI represents the diversity of functional zone categories, while SHEI illustrates the uniformity of number and area for urban functional zones. Where the SHDI and SHEI are high indicating that the subdistrict might be more balanced in functional zones distributions, for example in Fig. 19, the Tanjiajie and Xujiawan subdistricts.



(a) CONTAG

(b) DIVISION

(c) SHDI and SHEI

Fig. 19. Urban function pattern index results

Urban functional zone could play better part in human society by interrelating and interacting with each other. To evaluate the connection between each zone, the service scope evaluation index is proposed and tested by analyzing campus and residential districts as an example. Fig. 20 shows the result of nearest campus location searching within each settlement based on Euclidean Distance. As can be seen from the diagram, the school facilities in southeast study area appeared to be overloaded, yet in the southwest there's much looser, therefore the campus allocation in the direction southeast and southwest of the study area is considered less reasonable.

Fig. 21 shows the functional service scope of primary and secondary schools. In total there are 92 residential districts and 15 campuses in the study area. The service scope for campus was set respectively at the scale of 500m, 1000m, 1500m, 2000m. Within the 500m service scope locates 32 residential districts, accounting for 35% in total, when the parameter goes up to 1000, the number of residential districts rises as well, with a total of 34 and accounts for 37%, however, as the parameter further goes up to 1500, the number of residential districts decreases, with a total of 16 and accounts for 17%. Except for these, there are 10 residential districts accounting for 10%, located beyond the 1500m service scope, where the service radius is too large against the national standards, the locations of such are all to east of the study area. This suggests that the service radius of primary and secondary schools in the eastern study area is relatively large, which surpasses 1500 meters. Thus, the accessibility of schools is not so well as it's supposed to be.

Fig. 20. Search the nearest campus in each residential district (The red points represent campuses while the blue points represent residential districts)

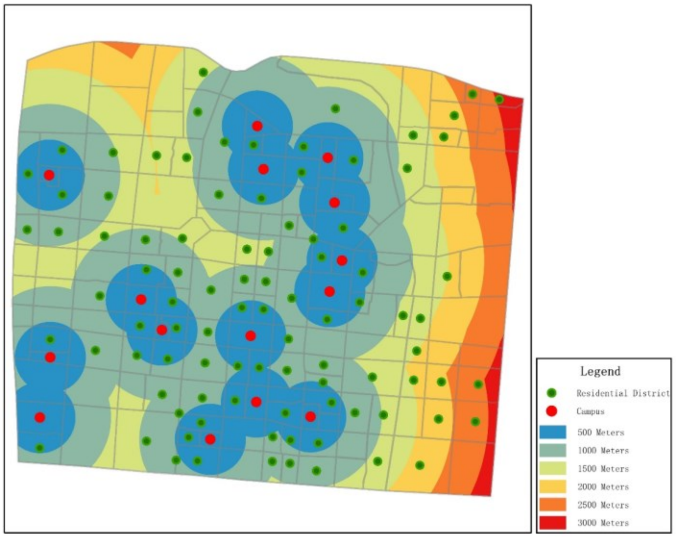


Fig. 21. The service scope of the campus

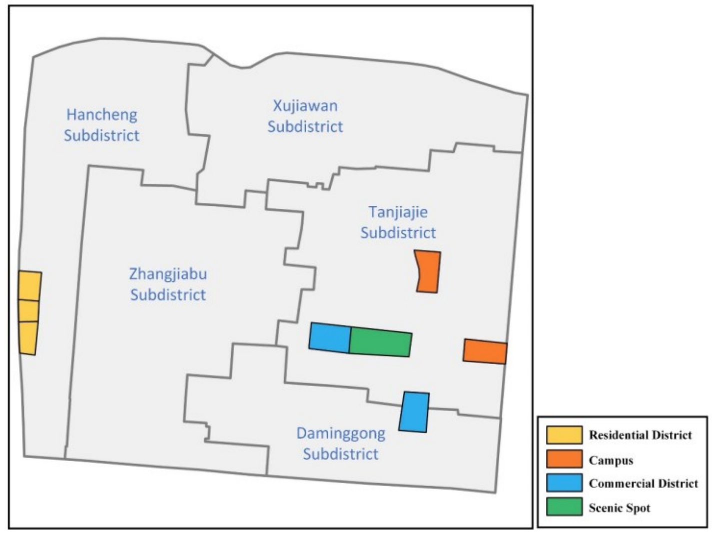


Fig. 22. Reginal planning for unused land

3) *Suggestions for Urban Function Planning*

The ultimate goal of functional allocation analysis is to propose suggestions for urban functional planning. According to the results of the allocation analysis, advices for unused land-use planning are given in Fig. 22. In the south of the Hancheng subdistrict, residential districts were discretely distributed and the school infrastructure is complete. Therefore, it is necessary to build up more residential districts on the unutilized land, gradually. The results demonstrated that Daminggong subdistrict has mainly high level of dominance by residential districts. It would be beneficial to build up a series of commercial districts, for there contains a great potential of commercial consumptions. In the east of the study area, especially in the Tanjiajie subdistrict, the spatial accessibility of school is low, for this reason, public infrastructure such as schools and scenic spots is recommended to be built in the future.

V. DISCUSSION

A. *The Effectiveness of Equal-Area Division Algorithm in OCNN Classification*

To testify the validity of the Equal-Area division algorithm, the commonly used random point generation algorithm is used as the comparison method. Combining the two algorithms with OCNN, the confusion matrix calculated is shown in Fig. 23. Fig. 24 shows the local classification results of the remaining six major land cover types except water body (because the classification accuracies of both two-convolution kernel position generation methods). It can be seen that the former has a relatively better recognition of the three types of land covers: building with blue roof, artificial turf field and bare land. Fig. 23 also shows the proposed Equal-Area division algorithm as well performs better for other three types although wrong classifications still happen. Since the convolution positions of buildings are almost generated at the boundary, CNN will irrationally extract the object features of the surrounding land covers, which causes misclassification of building.

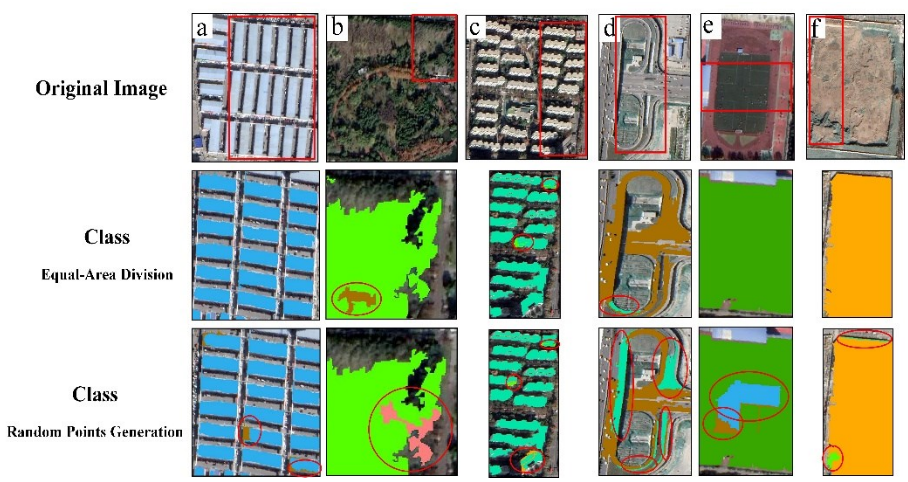
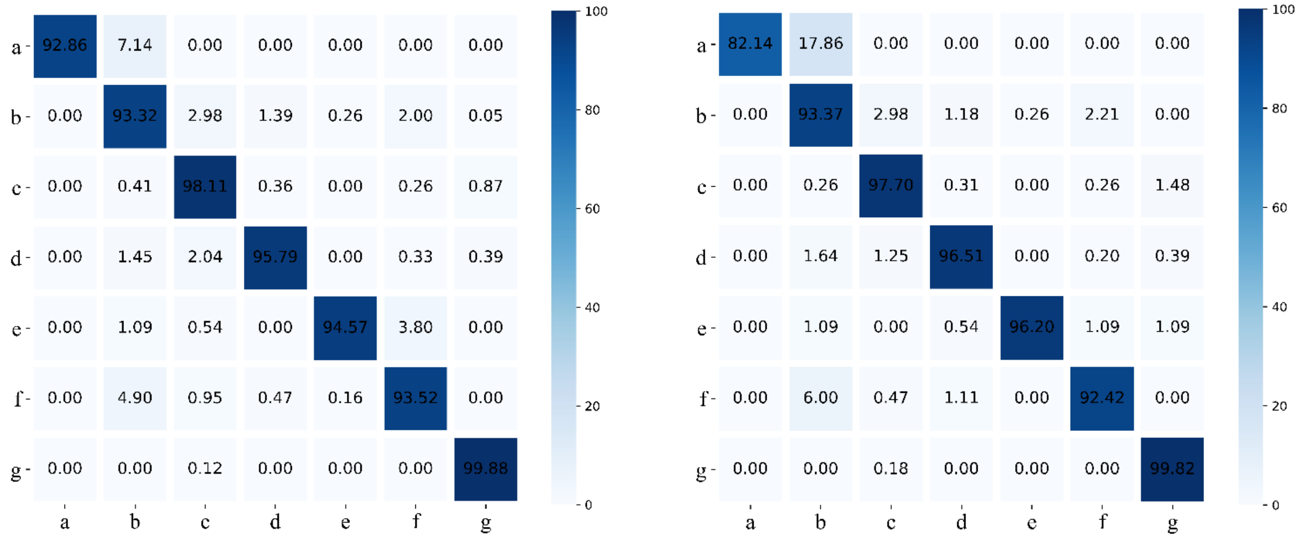


Fig. 23. Local classification results respectively by using Equal-Area division algorithm and Random Points Generation algorithm to generate convolution kernel position. (a: building with blue roof; b: grass; c: other building; d: road; e: artificial turf field; f: bare soil)



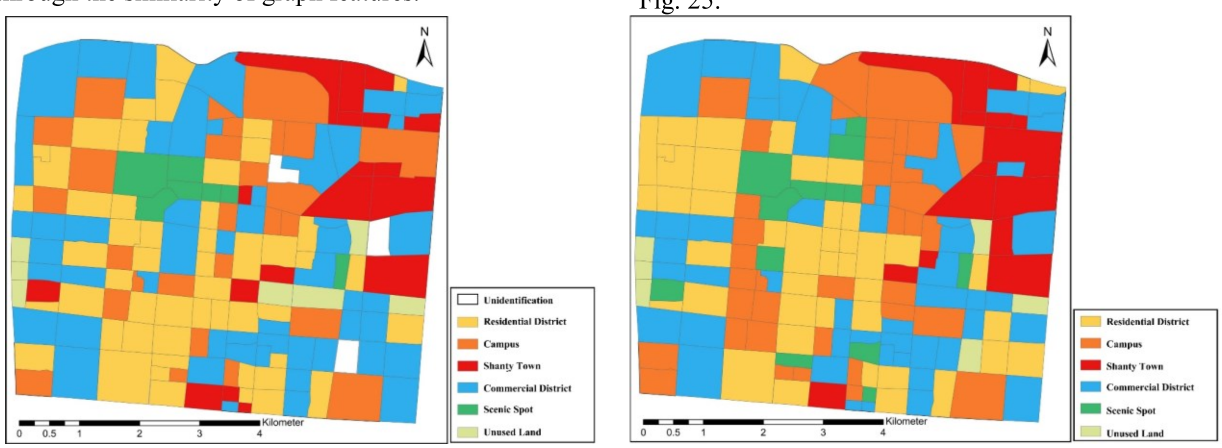
(a) Equal-Area division algorithm (b) Random Point Generation algorithm

Fig. 24. Confusion Matrix. (a: building with blue roof; b: grass; c: other building; d: road; e: artificial turf field; f: bare soil; g: water)

B. The Performance of DW-GAM for Functional Zone Attribute Identification.

The spatial distribution of urban function forms the spatial structure of the city, the spatial arrangement of various urban land use types. For example, most residential districts are formed by neatly arranged buildings and greenbelts; industrial areas consist of some office buildings, factory buildings and vegetation. Thus, this paper proposes to identify urban functional zones by exploring the different arrangement of various land covers. The DW-GAM provides the possibility to achieve this goal. DW-GAM can deeply mine the spatial relationships of different types of nodes being expressed by graphs, and then judge the functional zone attributes of each graph through the similarity of graph features.

Traditional methods for identifying functional zone attributes mostly use social perception data such as POI to give the functional semantic information, or use CNN to explore the influential features in different functional zones. Therefore, to verify the performance of DW-GAM on exploring the spatial distribution characteristics of land covers to identify urban functional zones, the urban functional zone identification results respectively by DW-GAM, traditional POI voting method and CNN classification method are compared. For the sake of keeping the variables the same, we use the block units of each level obtained by the method proposed in Fig.17 as the functional zone boundary constraints. The overall accuracy of the POI voting method and CNN method is 81.25% and 79.50%, respectively. The comparison result is illustrated in Fig. 17 and Fig. 25.



(a) By POI voting method (b) By CNN classification method

Fig. 25. Urban functional zone identification results respectively by POI voting method and CNN classification.

From Fig. 25, it can be seen that the main problems of the POI voting based method are: first, there is no POI in some blocks, which makes the blocks unclassified; second, although Shanty Town and Scenic Spot can be recognized with high accuracy, the mis-recognitions for other functional zones are a bit serious. Additionally, CNN classification method for identifying functional zones is neither ideal, especially the campuses and residential districts are severely misclassified, because CNN only focuses on deep features in fixed-size windows rather than the features of topological relations existing in the window scenes. By contrast, the proposed DW-GAM method performs best since it comprehensively considers both land cover types and their spatial topological relations.

C. The Effectiveness of multi-scale division

Actually, the boundary of each urban functional zone is not uniform. Some functional zones with large functional radiation area may be extracted by large-scale block constraints. However, some large-scale blocks containing complex spatial distribution of diverse land covers should be divided on fine scale by small-scale block constraints to further determine the attributes of urban functional zones on fine scale.

Because the types of urban functional zones are complex and the size of block units are various, we cannot determine which scale of blocks should be optimally used for dividing urban functional zone. If the scale is too large, it may lead to under-division, and the functional zones in the block cannot be fully extracted; if the scale is too small, over-division will happen, which destroys the spatial structure of land covers in the functional zone and causes the redundancy of the sample size and sick graph feature expressing within the block.

Fig. 26 demonstrates the functional zone division results respectively by single-scale and multi-scale methods. The proposed multi-scale method divides a total number of 170 urban functional zone units. If this method is not employed, 299 urban functional zone units will be divided by single scale method (using small-scale road network to constrain block division). From the perspective of data amount to be identified, the division unit is reduced by 40% and the time consuming for functional zone division and attribute identification is saved. Therefore, for the analysis of urban functional structure, the proposed method is more advantageous. Not only can the fine-scale division unit be derived from large-scale, but also the functional structure of the city can be reflected from a multi-scale perspective.

(a) By single-scale method (b) By multi-scale method

Fig. 26. Functional zone division results respectively by single-scale and multi-scale methods.

VI. CONCLUSIONS

Urban functional zone identification and functional pattern analysis are of great significance for urban planning and urban sustainable development. By combining using VHSR image, POI data and multi-scale road network data, this paper proposes a novel multi-scale recursive model that incorporates POI frequency density analysis and DW-GAM graph classification to classify urban functional zone. Especially, this paper presents a case study of urban functional zone identification and allocation analysis using self-established urban functional pattern evaluation index. The experiments and comparisons show the feasibility and superior of the proposed method. The contributions of this work can be summarized as follows.

1) An Equal-Area division algorithm is proposed to generate more reasonable convolution kernel locations for OCNN, which guarantees the accuracy of land cover classification to a certain extent and further ensures the accuracy of functional zone identification.

2) Considering multiple land covers with specific spatial relationships compose different functional zones, DW-GAM is originally proposed for urban functional zone classification. Rather than Convolutional Neural Network model, the DW-GAM is used to analyze spatial relationship features between-land cover variation in urban functional zone identification, which enhances the stringency of functional zone identification.

3) Using multi-level road network to decompose block unit step by step, a recursive model for multi-scale division of urban functional zones is proposed by combining POI and DW-GAM to realize the fine extraction of functional zones from top to bottom, which avoids the imperfection of using single road network (such as over-division or under-division of some blocks). Additionally, the identification results by jointly using DW-GAM and CR can relieve the negative influence when there is no POI in some blocks or the POI data are inaccurate.

4) As an applicable goal of this work, from the perspective of functional pattern analysis, an evaluation index system of urban functional reasonable allocation is put forward, by which suggestions for urban functional planning can be given, which is practically meaningful for enriching the research related to urban functional zone issues.

In addition, there are still some concerns as follows for future research.

1) The method proposed in this paper has some shortcomings in the functional zone attribute identification. Since the spatial structure difference between the land covers in industrial districts and commercial districts is not conspicuous enough, the model cannot divide the two functional zones accurately. In the future, more social perception data, such as GPS trajectory data, can be introduced to explore the human mobility among different functional zones for further distinguish different types of urban functional zones. Besides, aiming at better distinguishing the two types of functional zone, other network models based on graph classification can also be introduced to more accurately express the spatial relation features within the commercial districts and industrial districts.

2) The allocation of urban functional pattern may affect the utilization of resources and the social cohesion of people and society, which further determines whether the urban development is sustainable. This paper presents a primary framework for urban functional allocation analysis from the perspective of functional zone spatial pattern. However, since urban function planning involves many factors such as human, resources and environment, it is necessary to introduce more factors to analyze the urban functional allocation. For example, the change pattern of urban population can be extracted from nighttime light data, and urban housing price data can be used to reflect urban economy. In addition, the multi-scale functional pattern analysis theory and indicators system should be established in the future to more effectively express the

structural characteristics of functional zones. So that the relevant research results can better serve the geographical pattern monitoring and urban intelligent management.

Abbreviation

The following abbreviations are used in this manuscript.

DW-GAM	Distance Weight-Graph Attention Model
OCNN	Object Convolution Neural Network
VHSR	Very High Spatial Resolution
FD	Frequency Density
CR	Category Ratio
OBIA	Object-based Image Analysis
ENN	Euclidean Nearest Neighbor Distance
PLAND	Percentage of Landscape
AWMPFD	Area-weighted mean patch fractal dimension
CONTACT	Contagion Index
DIVISION	Landscape Division Index
SHDI	Shannon's Diversity Index
SHEI	Shannon's Evenness Index

References

[1] X. Huang *et al.*, "High-resolution urban land-cover mapping and landscape analysis of the 42 major cities in China using ZY-3 satellite images," *Sci. Bull.*, vol. 65, no. 12, pp. 1039–1048, 2020, doi: 10.1016/j.scib.2020.03.003.

[2] T. Blaschke *et al.*, "Geographic Object-Based Image Analysis - Towards a new paradigm," *ISPRS J. Photogramm. Remote Sens.*, vol. 87, pp. 180–191, 2014, doi: 10.1016/j.isprsjprs.2013.09.014.

[3] L. Gomez-Chova, D. Tuia, G. Moser, and G. Camps-Valls, "Multimodal Classification of Remote Sensing Images: A Review and Future Directions," *Proc. IEEE*, vol. 103, no. 9, pp. 1560–1584, 2015, doi: 10.1109/JPROC. 2015.2449668.

[4] M. Schultz, J. Voss, M. Auer, S. Carter, and A. Zipf, "Open land cover from OpenStreetMap and remote sensing," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 63, no. May, pp. 206–213, 2017, doi: 10.1016/j.jag.2017.07.014.

[5] D. Ming, J. Li, J. Wang, and M. Zhang, "Scale parameter selection by spatial statistics for GeOBIA: Using mean-shift based multi-scale segmentation as an example," *ISPRS J. Photogramm. Remote Sens.*, vol. 106, pp. 28–41, 2015, doi: 10.1016/j.isprsjprs.2015.04.010.

[6] J. Li, J. A. Benediktsson, B. Zhang, T. Yang, and A. Plaza, "Spatial Technology and Social Media in Remote Sensing: A Survey," *Proc. IEEE*, vol. 105, no. 10, pp. 1855–1864, 2017, doi: 10.1109/JPROC.2017.2729890.

[7] W. Zhou, G. Huang, A. Troy, and M. L. Cadenasso, "Object-based land cover classification of shaded areas in high spatial resolution imagery of urban areas: A comparison study," *Remote Sens. Environ.*, vol. 113, no. 8, pp. 1769–1777, 2009, doi: 10.1016/j.rse.2009.04.007.

[8] Y. Zhong *et al.*, "Open-source data-driven urban land-use mapping integrating point-line-polygon semantic objects: A case study of Chinese cities," *Remote Sens. Environ.*, vol. 247, no. March, 2020, doi: 10.1016/j.rse.2020.111838.

[9] J. Gong, C. Liu, and X. Huang, "Advances in urban information extraction from high-resolution remote sensing imagery," *Sci. China Earth Sci.*, vol. 63, no. 4, pp. 463–475, 2020, doi: 10.1007/s11430-019-9547-x.

[10] X. X. Zhu *et al.*, "Deep learning in remote sensing: a review," *arXiv*, no. december, 2017, doi: 10.1109/MGRS.2017.2762307.

[11] X. Lv, D. Ming, Y. Y. Chen, and M. Wang, "Very high resolution remote sensing image classification with SEEDS-CNN and scale effect analysis for superpixel CNN classification," *Int. J. Remote Sens.*, vol. 40, no. 2, pp. 506–531, 2019, doi: 10.1080/01431161.2018.1513666.

[12] X. Lv, D. Ming, T. Lu, K. Zhou, M. Wang, and H. Bao, "A new method for region-based majority voting CNNs for very high resolution image classification," *Remote Sens.*, vol. 10, no. 12, pp. 1–24, 2018, doi: 10.3390/rs10121946.

[13] Y. Chen, D. Ming, and X. Lv, "Superpixel based land cover classification of VHR satellite image combining multi-scale CNN and scale parameter estimation," *Earth Sci. Informatics*, vol. 12, no. 3, pp. 341–363, 2019, doi: 10.1007/s12145-019-00383-2.

[14] C. Zhang *et al.*, "A hybrid MLP-CNN classifier for very fine resolution remotely sensed image classification," *ISPRS J. Photogramm. Remote Sens.*, vol. 140, pp. 133–144, 2018, doi: 10.1016/j.isprsjprs.2017.07.014.

[15] C. Zhang *et al.*, "Joint Deep Learning for land cover and land use classification," *Remote Sens. Environ.*, vol. 221, no. May 2018, pp. 173–187, 2019, doi: 10.1016/j.rse.2018.11.014.

[16] Y. Zhai *et al.*, "Simulating urban land use change by integrating a convolutional neural network with vector-based cellular automata," *Int. J. Geogr. Inf. Sci.*, vol. 34, no. 7, pp. 1475–1499, 2020, doi: 10.1080/13658816.2020.1711915.

[17] W. Zhao, S. Du, and W. J. Emery, "Object-Based Convolutional Neural Network for High-Resolution Imagery Classification," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 10, no. 7, pp. 3386–3396, 2017, doi: 10.1109/JSTARS.2017.2680324.

[18] C. Zhang *et al.*, "An object-based convolutional neural network (OCNN) for urban land use classification," *Remote Sens. Environ.*, vol. 216, no. January, pp. 57–70, 2018, doi: 10.1016/j.rse.2018.06.034.

[19] D. Zhu *et al.*, "Understanding Place Characteristics in Geographic Contexts through Graph Convolutional Neural Networks," *Ann. Am. Assoc. Geogr.*, vol. 110, no. 2, pp. 408–420, 2020, doi: 10.1080/24694452.2019.1694403.

[20] Z. Cui, K. Henrickson, R. Ke, and Y. Wang, "Traffic Graph Convolutional Recurrent Neural Network: A Deep Learning Framework for Network-Scale Traffic Learning and Forecasting," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 11, pp. 4883–4894, 2020, doi: 10.1109/TITS.2019.2950416.

[21] K. Lei, M. Qin, B. Bai, G. Zhang, and M. Yang, "GCN-GAN: A non-linear temporal link prediction model for weighted dynamic networks," *arXiv*, pp. 388–396, 2019.

[22] T. N. Kipf and M. Welling, "Variational Graph Auto-Encoders," no. 2, pp. 1–3, 2016, [Online]. Available: http://arxiv.org/abs/1611.07308.

[23] 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. 2020-Decem, no. August, pp. 4800–4810, 2020.

[24] S. Kearnes, K. McCloskey, M. Berndl, V. Pande, and P. Riley, "Molecular graph convolutions: moving beyond fingerprints," *J. Comput. Aided. Mol. Des.*, vol. 30, no. 8, pp. 595–608, 2016, doi: 10.1007/s10822-016-9938-8.

[25] M. Li, A. Stein, and K. M. de Beurs, "A Bayesian characterization of urban land use configurations from VHR remote sensing images," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 92, no. June, p. 102175, 2020, doi: 10.1016/j.jag.2020.102175.

[26] Y. Feng, S. Du, S. W. Myint, and M. Shu, "Do urban functional zones affect land surface temperature differently? A case study of Beijing, China," *Remote Sens.*, vol. 11, no. 15, 2019, doi: 10.3390/rs11151802.

[27] M. Luck and J. Wu, "A gradient analysis of urban landscape pattern: A case study from the Phoenix metropolitan region, Arizona, USA," *Landsc. Ecol.*, vol. 17, no. 4, pp. 327–339, 2002, doi: 10.1023/A:1020512723753.

[28] X. Zhang, S. Du, S. Du, and B. Liu, "How do land-use patterns influence residential environment quality? A multiscale geographic survey in Beijing," *Remote Sens. Environ.*, vol. 249, no. August, p. 112014, 2020, doi: 10.1016/j.rse.2020.112014.

[29] S. Yuan, J., Wu, C., Du, B., Zhang, L., Wang, "Analysis of landscape pattern on urban land use based on GF-5 hyperspectral data," *J. Remote Sens.*, vol. 24, no. 4, pp. 465–478, 2020.

[30] J. Zhang *et al.*, "Exploring annual urban expansions in the Guangdong-Hong Kong-Macau Greater Bay Area: Spatiotemporal features and driving factors in 1986-2017," *Remote Sens.*, vol. 12, no. 16, 2020, doi: 10.3390/RS12162615.

[31] X. Huang, X. Han, S. Ma, T. Lin, and J. Gong, "Monitoring ecosystem service change in the City of Shenzhen by the use of high-resolution remotely sensed imagery and deep learning," *L. Degrad. Dev.*, vol. 30, no. 12, pp. 1490–1501, 2019, doi: 10.1002/ldr.3337.

[32] C. Yang *et al.*, "Spatiotemporal evolution of urban agglomerations in four major bay areas of US, China and Japan from 1987 to 2017: Evidence from remote sensing images," *Sci. Total Environ.*, vol. 671, pp. 232–247, 2019, doi: 10.1016/j.scitotenv.2019.03.154.

[33] X. Liu and J. Li, "Scientific solutions for the functional zoning of nature reserves in China," *Ecol. Modell.*, vol. 215, no. 1–3, pp. 237–246, 2008, doi: 10.1016/j.ecolmodel.2008.02.015.

[34] S. Du, S. Du, B. Liu, and X. Zhang, "Remote Sensing of Environment Mapping large-scale and fine-grained urban functional zones from VHR images using a multi-scale semantic segmentation network and object

based approach," Remote Sens. Environ., vol. 261, no. April, p. 112480, 2021, doi: 10.1016/j.rse.2021.112480.

[35] W. Zhou, D. Ming, X. Lv, K. Zhou, H. Bao, and Z. Hong, "SO–CNN based urban functional zone fine division with VHR remote sensing image," *Remote Sens. Environ.*, vol. 236, no. July 2019, p. 111458, 2020, doi: 10.1016/j.rse.2019.111458.

[36] J. Song, T. Lin, X. Li, and A. V. Prishchepov, "Mapping urban functional zones by integrating very high spatial resolution remote sensing imagery and points of interest: A case study of Xiamen, China," *Remote Sens.*, vol. 10, no. 11, 2018, doi: 10.3390/rs10111737.

[37] Y. Shen and K. Karimi, "Urban function connectivity: Characterisation of functional urban streets with social media check-in data," *Cities*, vol. 55, no. March 2019, pp. 9–21, 2016, doi: 10.1016/j.cities.2016.03.013.

[38] F. Zhang, B. Du, and L. Zhang, "Scene classification via a gradient boosting random convolutional network framework," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 3, pp. 1793–1802, 2016, doi: 10.1109/TGRS. 2015.2488681.

[39] S. Du, S. Du, B. Liu, and X. Zhang, "Context-enabled extraction of large-scale urban functional zones from very-high-resolution images: A multiscale segmentation approach," *Remote Sens.*, vol. 11, no. 16, 2019, doi: 10.3390/rs11161902.

[40] Y. Wang, T. Wang, M. H. Tsou, H. Li, W. Jiang, and F. Guo, "Mapping dynamic urban land use patterns with crowdsourced geo-tagged social media (Sina-Weibo) and commercial points of interest collections in Beijing, China," *Sustain.*, vol. 8, no. 11, pp. 1–19, 2016, doi: 10.3390/su8111202.

[41] S. Gao, K. Janowicz, and H. Couclelis, "Extracting urban functional regions from points of interest and human activities on location-based social networks," *Trans. GIS*, vol. 21, no. 3, pp. 446–467, 2017, doi: 10.1111/tgis.12289.

[42] N. Xu, J. Luo, T. Wu, W. Dong, W. Liu, and N. Zhou, "Identification and portrait of urban functional zones based on multisource heterogeneous data and ensemble learning," Remote Sens., vol. 13, no. 3, pp. 1–20, 2021, doi: 10.3390/rs13030373.

[43] H. Bao, D. Ming, Y. Guo, K. Zhang, K. Zhou, and S. Du, "DFCNN-based semantic recognition of urban functional zones by integrating remote sensing data and POI data," *Remote Sens.*, vol. 12, no. 7, 2020, doi: 10.3390/rs12071088.

[44] X. Zhang and S. Du, "A Linear Dirichlet Mixture Model for decomposing scenes: Application to analyzing urban functional zonings," *Remote Sens. Environ.*, vol. 169, pp. 37–49, 2015, doi: 10.1016/j.rse.2015.07.017.

[45] Z. Zhang, Y. Wang, Q. Liu, L. Li, and P. Wang, "A CNN based functional zone classification method for aerial images," *2016 IEEE Int. Geosci. Remote Sens. Symp.*, vol. 6, pp. 5449–5452, 2016.

[46] J. B. Lee, R. Rossi, and X. Kong, "Graph classification using structural attention," *Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, pp. 1666–1674, 2018, doi: 10.1145/3219819.3219980.

[47] W. Chen, H. Huang, J. Dong, Y. Zhang, Y. Tian, and Z. Yang, "Social functional mapping of urban green space using remote sensing and social sensing data," *ISPRS J. Photogramm. Remote Sens.*, vol. 146, no. May, pp. 436–452, 2018, doi: 10.1016/j.isprsjprs.2018.10.010.

[48] W. Tu *et al.*, "Portraying urban functional zones by coupling remote sensing imagery and human sensing data," *Remote Sens.*, vol. 10, no. 1, pp. 1–20, 2018, doi: 10.3390/rs10010141.

[49] C. Aubrecht and J. A. L. Torres, "Evaluating multi-sensor nighttime earth observation data for identification of mixed vs. residential use in urban areas," *Remote Sens.*, vol. 8, no. 2, 2016, doi: 10.3390/rs8020114.

[50] Q. Gao, J. Fu, Y. Yu, and X. Tang, "Identification of urban regions' functions in Chengdu, China, based on vehicle trajectory data," *PLoS One*, vol. 14, no. 4, pp. 1–17, 2019, doi: 10.1371/journal.pone.0215656.

[51] W. Tu *et al.*, "Coupling mobile phone and social media data: a new approach to understanding urban functions and diurnal patterns," *Int. J. Geogr. Inf. Sci.*, vol. 31, no. 12, pp. 2331–2358, 2017, doi: 10.1080/13658816.2017.1356464.

[52] Y. Zhang, Q. Li, W. Tu, K. Mai, Y. Yao, and Y. Chen, "Functional urban land use recognition integrating multi-source geospatial data and cross-correlations," *Comput. Environ. Urban Syst.*, vol. 78, no. 3688, p. 101374, 2019, doi: 10.1016/j.compenvurbsys.2019.101374.

[53] R. Cao *et al.*, "Deep learning-based remote and social sensing data fusion for urban region function recognition," *ISPRS J. Photogramm. Remote Sens.*, vol. 163, no. February, pp. 82–97, 2020, doi: 10.1016/j.isprsjprs.2020.02.014.

[54] Y. Yao *et al.*, "Sensing spatial distribution of urban land use by integrating points-of-interest and Google Word2Vec model," *Int. J. Geogr. Inf. Sci.*, vol. 31, no. 4, pp. 825–848, 2017, doi: 10.1080/13658816.2016.1244608.

[55] X. Zhang, S. Du, and Z. Zheng, "Heuristic sample learning for complex urban scenes: Application to urban functional-zone mapping with VHR images and POI data," *ISPRS J. Photogramm. Remote Sens.*, vol. 161, no. December 2019, pp. 1–12, 2020, doi: 10.1016/j.isprsjprs.2020.01.005.

[56] Z. Huang, H. Qi, C. Kang, Y. Su, and Y. Liu, "An ensemble learning approach for urban land use mapping based on remote sensing imagery and social sensing data," *Remote Sens.*, vol. 12, no. 19, pp. 1–18, 2020, doi: 10.3390/rs12193254.

[57] S. Hu, Z. He, L. Wu, L. Yin, Y. Xu, and H. Cui, "A framework for extracting urban functional regions based on multiprototype word embeddings using points-of-interest data," *Comput. Environ. Urban Syst.*, vol. 80, no. May 2019, p. 101442, 2020, doi: 10.1016/j.compenvurbsys.2019.101442.

[58] S. Jiang, A. Alves, F. Rodrigues, J. Ferreira, and F. C. Pereira, "Mining point-of-interest data from social networks for urban land use classification and disaggregation," *Comput. Environ. Urban Syst.*, vol. 53, pp. 36–46, 2015, doi: 10.1016/j.compenvurbsys.2014.12.001.

[59] H. Cui, L. Wu, S. Hu, R. Lu, and S. Wang, "Recognition of urban functions and mixed use based on residents' movement and topic generation model: The case of Wuhan, China," *Remote Sens.*, vol. 12, no. 18, 2020, doi: 10.3390/RS12182889.

[60] Z. Zhu *et al.*, "Understanding an urbanizing planet: Strategic directions for remote sensing," *Remote Sens. Environ.*, vol. 228, no. May, pp. 164–182, 2019, doi: 10.1016/j.rse.2019.04.020.

[61] Y. Chi, J., Jiao, L., Dong, T., Gu, Y., Ma, "Quantitative identification and visualization of urban functional area based on POI data," *J. Geomatics*, vol. 41, no. 2, pp. 68–73, 2016, doi: 10.14188/j.2095-6045.2016.02.017.

[62] Y. Chen *et al.*, "Delineating urban functional areas with building-level social media data: A dynamic time warping (DTW) distance based k-medoids method," *Landsc. Urban Plan.*, vol. 160, pp. 48–60, 2017, doi: 10.1016/j.landurbplan.2016.12.001.

[63] R. E. Plotnick, R. H. Gardner, and R. V. O'Neill, "Lacunarity indices as measures of landscape texture," *Landsc. Ecol.*, vol. 8, no. 3, pp. 201–211, 1993, doi: 10.1007/BF00125351.

[64] W. Tu, Y. Zhang, Q. Li, K. Mai, and J. Cao, "Scale Effect on Fusing Remote Sensing and Human Sensing to Portray Urban Functions," *IEEE Geosci. Remote Sens. Lett.*, pp. 1–5, 2020, doi: 10.1109/lgrs.2020.2965247.

[65] W. Zhao, Y. Bo, J. Chen, D. Tiede, B. Thomas, and W. J. Emery, "Exploring semantic elements for urban scene recognition: Deep integration of high-resolution imagery and OpenStreetMap (OSM)," *ISPRS J. Photogramm. Remote Sens.*, vol. 151, no. March, pp. 237–250, 2019, doi: 10.1016/j.isprsjprs.2019.03.019.

[66] X. Huang and Y. Wang, "Investigating the effects of 3D urban morphology on the surface urban heat island effect in urban functional zones by using high-resolution remote sensing data: A case study of Wuhan, Central China," *ISPRS J. Photogramm. Remote Sens.*, vol. 152, no. April, pp. 119–131, 2019, doi: 10.1016/j.isprsjprs.2019.04.010.

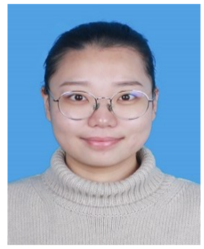
[67] L. Dong, Z. Huang, J. Zhang, and Y. Liu, "Understanding the mesoscopic scaling patterns within cities," *arXiv*, no. 0123456789, pp. 1–12, 2020, doi: 10.1038/s41598-020-78135-2.

Kui Zhang is currently pursuing the master's degree from the School of Information Engineering, China University of Geosciences, Beijing. His researches focus on recognizing urban functional zones by integrating with social perception data and mining the spatial relationships of geographic objects in high spatial resolution remote sensing image.

Dongping Ming received the B.E. degree in land administration and cadastral surveying from the Wuhan Technical University of Surveying and Mapping, Wuhan, China, in 1999, the M.E. degree in cartography and geographic information engineering from Wuhan University, Wuhan, in 2002, and the Ph.D. degree in cartography and geographic information system from the Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China, in 2006. She is currently a full professor with the School of Information Engineering, China University of Geosciences, Beijing. Her research interests include remote-sensing image processing and analysis, and information extraction from high spatial resolution satellite remote sensing image**.**

Shigao Du is currently pursuing the master's degree from the School of Information Engineering, China University of Geosciences, Beijing. His research interests are very related to image cognition, comprehending of images (spatio-temporal analysis based on geographic object) and deep learning.

Lu Xu is currently pursuing the PH.D. degree from the School of Information Engineering, China University of Geosciences, Beijing. He focuses on comprehending very high spatial resolution images. The specific research interests include Object-Based Image Analysis (OBIA) theory and image information extraction by deep learning methods.



Xiao Ling is currently studying in China University of Geoscience as a postgraduate student. Her research interests include Geo-hazardous information extractions and quantitative analysis, specifically, landslide image classification and susceptibility assessment.

Beichen Zeng is currently pursuing the master's degree from the School of Information Engineering, China University of Geosciences, Beijing. His research interests include deep learning and high spatial resolution remote sensing imagery analysis.

Xianwei Lv received the M.S. degree in the School of Information Engineering China University of Geosciences (Beijing) in 2019. Now, he is currently pursuing the Ph.D. degree in photogrammetry and remote sensing with the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan Uiversity. He does research in deep learning for very high-resolution image processing and applications.