# Identifying Urban Building Function by Integrating Remote Sensing Imagery and POI Data

Anqi Lin, Xiaomeng Sun, Hao Wu ⬤ , Wenting Luo, Danyang Wang, Dantong Zhong, Zhongming Wang, Lanting Zhao, and Jiang Zhu

*Abstract*—Identifying urban building function plays a critical role in understanding the complexness of urban construction and improving the effectiveness of urban planning. The emergence of user generated contents has brought access to massive semantic information which complements the traditional remote sensing data for identifying urban building functions and exploring the spatial structure in urban environment. This article proposes a stepwise identification framework for urban building functions based on remote sensing imagery and point of interests (POIs) data, which merges the spatial similarity of buildings and kernel density to improve the identification accuracy and completeness. Taking Wuhan as an example, Google earth images and POI data were obtained to identify the seven primary categories for the individual buildings in the core urban area. The results suggest that the proposed stepwise framework is feasible to identify the urban building functions as the identification results exhibit the superiority in terms of accuracy and completeness. Our results suggest that the identification of urban building function is sensitive to the bandwidth of kernel density estimation and 200 meter is the optimal size. The findings also indicate that significant spatial agglomeration exists in residential and commercial buildings at both macro and microlevels.

*Index Terms*—Google earth image, kernel density estimation (KDE), point of interest (POI) data, spatial similarity, urban buildings, user generate contents (UGCs).

## I. INTRODUCTION

URBAN LAND use information is essential for urban and environmental studies [1]. As the fundamental units of cities, buildings serve as the functional spaces for living, working, entertaining and other socio-economic activities, and

Anqi Lin, Xiaomeng Sun, Hao Wu, Wenting Luo, Danyang Wang, Dantong Zhong, Zhongming Wang, and Lanting Zhao are with the College of Urban and Environmental Sciences, Central China Normal University, Wuhan 430079, China, and also with the Hubei Provincial Key Laboratory for Geographical Process Analysis and Simulation, Central China Normal University, Wuhan 430079, China. (e-mail: linanqi@mails.ccnu.edu.cn; villion@mails.ccnu.edu.cn; haowu@mail.ccnu.edu.cn; jinyinhua@mails.ccnu.edu.cn; wangdy27@mails.ccnu.edu.cn; winter@mails.ccnu.edu.cn; wangzhongming@mails.ccnu.edu.cn; 2017212606zlt@mails.ccnu.edu.cn).

Jiang Zhu is with the KQ GEO Technologies Co., Ltd, Beijing 100176, China. (e-mail: zhujiang@kqgeo.com).

Digital Object Identifier 10.1109/JSTARS.2021.3107543

reflect detailed land use information [2]. That is, building function refers to its actual use, such as commercial, residential, or industrial, rather than the initial designed purpose. Identifying the building functions could not only facilitate a better understanding of urban morphology but also contribute to various applications in emergency response [3], policy making [4], resources management [5], and other fields. However, information of urban building functions was mostly collected from field investigations by government agency, which could be time-consuming and rarely open to public. Therefore, an automated and efficient way to identify urban building function is absolutely important for refining urban land use mapping, and thereby allows the further applications to the benefit of urban development and urban planning.

Urban land use mapping largely depends on the interpretation of remote sensing imageries as they reflect land use information in a large extent with high temporal frequency [6], [7]. The history of land use mapping by remote sensing technology dates back to 1970s when the first satellite remote sensing for Earth observation has provided the unique opportunity for quantitative analysis and dynamic monitoring in urban land use [8], [9]. Not until around 2000 has a variety of very-high-resolution (VHR) satellite sensors been available and shown a clearer representation of the ground truth [10]–[13]. More recently, plenty of virtual global platforms, such as Google Earth, NASA World Wind, and Bing Maps, have opened a new era of digital earth [14], [15]. For instance, the high resolution Google earth images, as free and open access data source, have become one of the primary supplementary date source of the traditional land use mapping [16], [17]. In specific, the richness in texture, tone and geometric features of Google earth images provide detailed information for land use classification at object level, including buildings and other artificial structures [18], [19]. However, remote sensing imageries are typically more useful for delineating the physical layout of ground objects rather than identifying their specific functional usage for human activities due to the deficiency in semantic information [20]. Fortunately, the emergence of user generate content (UGC) has demonstrated highly successful in allowing for active users to participate in collecting, updating and sharing the massive data that reflect human activities and social attributes [21]–[23]. Point of interests (POIs), a typical UGC data, has greatly addressed the semantic gap in remote sensing images [24]–[26]. Preliminary studies have introduced POI data as the reference data or training data for land use classification [27]–[29]. With the explosive growth in recent

years, POI data have been playing a more important role in land use study. Zhang *et al.* [30] made use of POI data to automatically delineate the urban functional zones in Hangzhou, China. Hong and Yao [31] took advantage of POI data to identify the urban functional zones in Guangzhou, China. Andrade *et al.* [32] conducted land use identification based on the numeric features extracted from POI data in Lisbon city, Portugal. Besides the single data source, there are studies fusing multiple sources of data for identifying urban functional zones. Zhang *et al.* [33] developed hierarchical semantic cognition for urban functional zones using VHR images and POI data. Zhong *et al.* [34] integrated POI, OpenStreetMap data, and VHR images for urban land use mapping. Liu *et al.* [35] recognized urban functional zones combined landscape features from Landsat image and human activities from POI and taxi GPS trajectories. Despite these significant contributions for classifying urban functional zone, few research has yet been attempted to combine remote sensing image and POIs for a better interpretation of urban land use at the finer scale, that is urban building unit rather than the street block or parcel. After all, not only has the variety of building functions increased in metropolitan areas, but also many sophisticated urban analyses regarding fine urban population mapping, emergency management, environmental monitoring and so on deeply depend on the detailed classification of urban buildings.

Besides the data sources, identification approach is another core part of land use study. Various models have been developed and applied to identify the urban functional areas, which taking advantage of semantic information from UGC data to complement the remote sensing imagery. The essential and preliminary way is to extract semantic information from POI or other UGC data spatially associated with buildings or parcels to classify the functions [36]–[38]. Later researches introduced the frequency index of POI data to distinguish the functions of urban areas [39], [40]. Furthermore, a topic-based model inferring on identify the functions of each region was proposed [41]. K-means algorithm based on object distance is suitable for processing high-dimensional objects [42] and it has been applied to identify the categories of land use clusters in many studies [43], [44]. However, few studies have yet taken the spatial distribution of POI data into account so that they failed to identify built-up areas where missing POI data, or resulted in low accuracy. Fortunately, the kernel density estimation (KDE) method, which exploits the strength of the estimation nonparametric random variables of point distribution, has significantly enabled us to transform the POI data into continuous surface features [45]. That is to say, KDE can be utilized to extract distribution characteristics of different land use from POI semantic information, enabling us to infer the building functions particularly for functional buildings with significant spatial aggregations. Not only the distribution characteristics of POI data deserve further exploitation, but also are the spatial features of buildings extracted from VHR images worthy further interpretation. For example, in a residential community or business district, buildings show certain textural or geometric similarities. That is to say those spatial features can be used to infer the building functions under the assumption that buildings with the similar spatial features always have
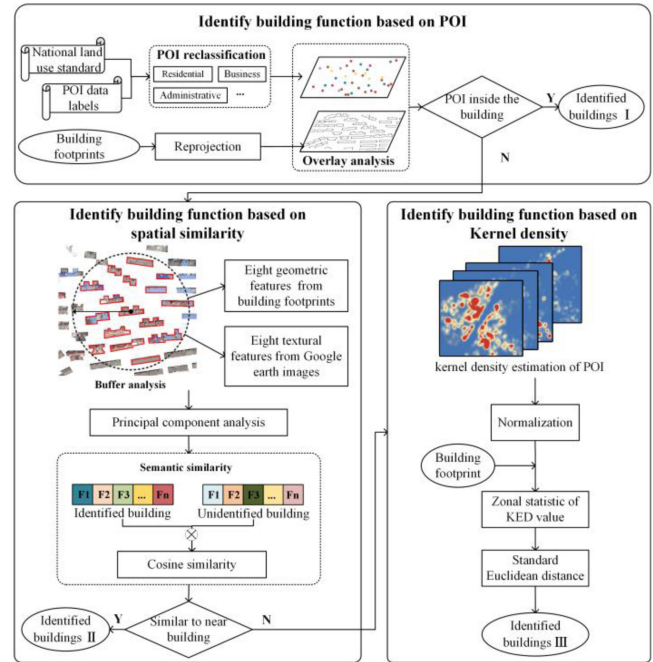


Fig. 1. Identification framework of urban building functions.

the similar functions within a certain region. Therefore, it is critically important to develop a hybrid identification method taking advantage of spatial similarity and POI distribution to achieve 100% identification rate of urban buildings with high accuracy.

The objective of this article is to propose a stepwise identification framework for urban building functions based on Google earth image and POI data, which merges the spatial similarity of buildings and KDE to improve the identification accuracy and completeness, and further examines the spatial aggregation characteristics of different type of buildings. The thorough understanding of urban building function will help policy makers take a deep insight into urban structures and improve the effectiveness of urban planning.

## II. METHODS

### A. Identification Framework of Urban Building Functions Integrating Remote Sensing Imagery and POI Data

To reveal the urban building functions, we proposed a stepwise identification framework of urban building functions, which integrates remote sensing imagery and POI data. As shown in Fig. 1, building functions were identified by three steps sequentially, including the direct usage of POI data, measurement of spatial similarity using VHR images and building footprint geometries, and KDE of POI data in different categories.

In the first step, we reclassified POI data according to the classification standards for urban development land and coordinate, and then decide building function according to POI inside buildings or within the distance of 5 m. Notably, POI data with low importance, but high density were used to decide building function by data frequency, and POI with high importance but low density were used by the priority order. Second, we
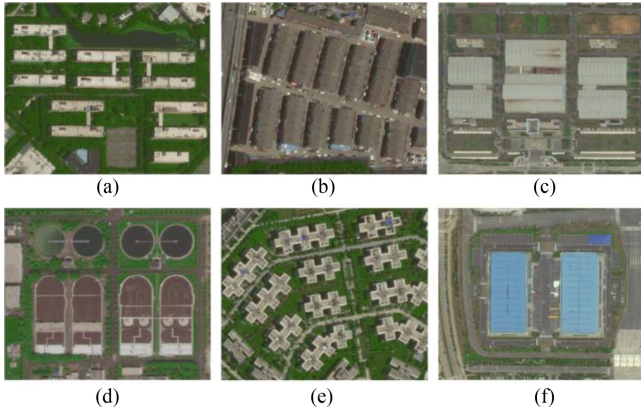
Fig. 2. Typical examples of different functional buildings with similar geometric and textural features.



Fig. 3. Graph illustration of eight building geometrical features.

applied spatial similarity to infer the building functions for those have no POI data inside the boundary. In specific, eight geometric features and eight textural features were extracted from building footprint geometries and VHR images, principle components analysis were employed to reduce dimensions for the spatial features. So that cosine similarity between unidentified and identified buildings in a buffer zone could be calculated to infer the building function. Third, in order to improve recognition rate of building functions, kernel densities from each POI category were calculated and normalized, and zone statistic of them were conducted for all buildings. Normalized Euclidean distance (NED) was performed to calculate the similarity of kernel densities between identified and unidentified buildings to complete the function identification.

### B. Identifying Building Functions Based on Spatial Similarity

Although associating the POI type to the overlapped building is the most efficient way to decide building functions, there is not always POI data inside or near the building. Under the circumstances, additional methods are needed which could infer the building function by extra information. Actually, it is observed from Google earth images that buildings with same functions in a small extent always show the similar geometric and textural features as shown in Fig. 2. So that we could take advantage of those features to infer the building function when there is identified building with similar spatial patterns.

To fully capture the spatial similarity of buildings, multiple geometric features and textural features were deriving from building footprint geometry and Google earth images. Specifically, eight classical geometric features, including area, perimeter, compactness, number of nodes, radius shape index [46], regularity, and aspect ratio orientation, were selected and calculated for all buildings. Detailed illustrations and equations of the eight geometric features are shown in Fig. 3. Meanwhile, texture is an important spatial feature measuring the spatial distribution of tones across the pixels of remote sensing image, which has been widely used for object identification and image segmentation [47]. The textural features were extracted from Google earth images by grey-level co-occurrence matrix (GLCM). Despite
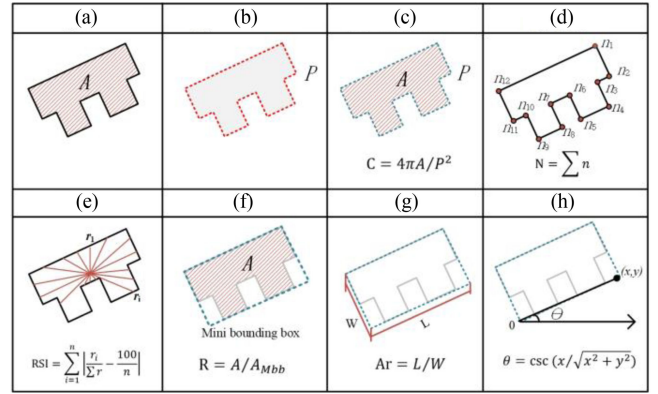
dozens of textural features can be derived from GLCM, we selected the most important and popular eight features including mean, variance, homogeneity, contrast, dissimilarity, entropy, second moment and correlation [48]. Detailed explanations and equations of each measure can refer to the literature [49].

Since all features are derived from either building geometry or Google earth images, it is inevitable that information redundancy exists between them. In order to reduce the dimensionality of selected features, principal component analysis, one of the most popular multivariate statistical methods, was applied to keep the variance of raw data as well as minimize information loss [50]. With the results of $n$ principal components, we could form the spatial feature vector of identified building, $F_i = [F1, F2, \ldots Fn]$, and unidentified building, $F_u = [F1', F2', \ldots Fn']$. So that spatial similarity between unidentified building and surrounding identified building can be calculated by cosine similarity

$$\text{Cosine } (F_i, \ F_u) = \frac{\sum_{i=1}^{n} F_i \cdot F_u}{\sqrt{\sum_{i=1}^{n} F_i^2} \cdot \sqrt{\sum_{i=1}^{n} F_u^2}}. \quad (1)$$

The original cosine value ranges from $-1$ to $+1$, but the spatial similarity was decided by the magnitude of cosine value regardless of the direction between two spatial feature vectors. That is, the greater the absolute value of cosine value, the higher the spatial similarity. In the buffer zone with one unidentified building in center, we iteratively compared its spatial features with identified buildings, the function of most similar identified building was assigned to the unidentified one.

### C. Identifying Building Functions Based on KDE

For those buildings could not be matched with nearby buildings by spatial similarity, KDE of POI data was applied to identify the function. Differences in the spatial distribution of POI density reflect the distribution characteristic of functional area [51]. Therefore, POIs were converted to continuous density surface representing the distribution characteristic by KDE method, and applied to infer the building function.

KDE is a useful method to estimate unknown density function in probability theory. It computes the density contribution of each sample point in the specified range to the centre point of

each grid through the kernel function, and finally generates a smooth surface indicating the density map [52]. The amount of POI data in different categories varies greatly, for example, the number of POI data in commercial and business category is greater than the number in other categories. To avoid the unbalanced distribution POI data in distinct categories and further improve the identification rate, the normalized of KDE is applied in the third step, which can be expressed as [53]

$$f(x) = \frac{1}{Nh^d} \sum_{i=1}^{N} K\left(\frac{x - x_i}{h}\right) \qquad (2)$$

where $f(x)$ is the KDE calculation function at the location $x$; $d$ is the spatial dimension; $h$ is the bandwidth; $N$ is the number of points which distance $\leq h$ from the location $x_i$; and $K$ is the spatial weight function.

NED is a similarity measure applied for calculating the similarity of kernel density between unidentified buildings and each building type, which has been previously used for land use mapping [54], [55]. With the kernel densities of POIs in each category, the average KDE values inside the building geometry were calculated as the kernel density of each building. The identified buildings based on POI and spatial similarity were selected as training samples, and their kernel density values were used for calculating the average value and standard deviation of different building types. The results were then used for computing NED as follows:

$$S_i = \sqrt{\frac{1}{m} \sum_{j=1}^{m} \left(\frac{x_j - \overline{x_i}}{\sigma_i}\right)^2} \qquad (3)$$

where $\overline{x_i}$ and $\sigma_i$ represent the average and standard deviation of the $i$th type of building kernel densities, respectively, $m$ represents the total amounts of identification types, $x_j$ represents the kernel density values of the $j$th unidentified building. Among all NED values of different types, the type with minimum NED (highest similarity) was assigned to the unidentified building.

## III. STUDY AREA AND DATA PROCESSING

### A. Study Area

Wuhan is the capital city of Hubei province in the central China covering a region of longitudes from 29°58′N to 31°22′N and latitudes from 113°41′E to 115°05′E, which is one of most important economic development centers in China. The administrative area of Wuhan city covers an area of approximately 8569.15 Km$^2$ and has a permanent residential population of nearly 11.21 million. With the rapid growth of regional economy and government management of urban construction, the functions of building have become more diverse and more complex [56]. Given that the diversity of building functions and rapid urbanization in Wuhan, it provides an excellent case for identifying the urban building functions. For this article, the most urbanized area inside third ring road of Wuhan was selected as the study area. The geographical location of the study area is shown in Fig. 4.
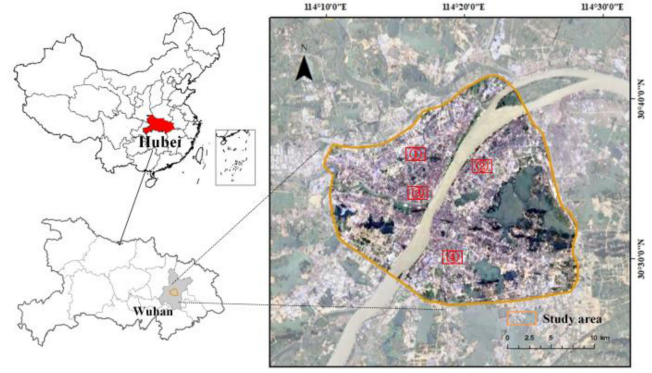


Fig. 4. Study area, Wuhan, China.

### B. Data Processing

Three different datasets were used in this article, including google earth images, POI data, and building footprint geometries in the study area. Google earth images were downloaded through GGGIS,[1] an open-access platform specialized for Google Earth. These Google earth images, acquired on March 2019, have red, green and blue bands with a spatial resolution of 1.2 m. Building footprint geometries in the format of shapefile were collected from National Platform for Common Geospatial Information Services.[2] The geometries were re-projected and modified to keep consistency with Google earth images.

POI data not only contain the spatial location but also provide a great number of attribute information, including name, address, and type, which could contribute to identify the urban building function. POI data are usually the locations representing real geographical entities generated by commercial or crowed-sourced platforms. In this article, 120 589 POI records were collected through application programming interfaces of Amap, one of the most popular map applications in China. POI data from Amap have been proven reliable source for POI data since they were widely used in urban research due to the high precision, wide coverage and frequent update [57], [58]. The original POI data are classified into dozens of types, which are inconsistent with the national standard of land use category. Thus, POI data were reclassified according to the National Urban Land Use and Planning Standard (GB50137-2011). So that POI data were reclassified into seven primary categories as shown in Table I: residential (R); administration and public services (A); commercial and business facilities (B); municipal utilities (U); logistics and warehouse (W); road, street, and transportation (S); and industrial, manufacturing (M). It was noticed that the public awareness and importance of different types of POI did not correspond to the density. In particular, the commercial type of POI occupied the largest proportion of the entire POI data, but most of them are small business that could not decide the primary use of the building. So that, POIs were divided to two types, type I: low importance but high density (B and W), type II: high importance but low density (R, A, S, T, M). If

---

[1]http://www.gggis.com/
[2]https://www.tianditu.gov.cn/

TABLE I
CATEGORY MAPPING DICTIONARY BETWEEN NATIONAL STANDARD
AND POI LABEL

| Primary Categories | Quantity | POI labels |
|---|---|---|
| Residential (R) | 26452 | Residential, Service facilities |
| Administration and public services (A) | 23029 | Administrative office, Book and exhibition facilities, Cultural activities, Universities and colleges, Secondary professional school, Primary and secondary schools, Special education, Scientific research, Stadiums and gymnasiums, Physical training, Hospital, Health and epidemic prevention, Special medical treatment, Other medical and health care, Social welfare facilities, Cultural relics and monuments, Foreign affairs, Religious facilities |
| Commercial and business facilities (B) | 66107 | Retail business, Wholesale Market, Catering, Hotel, Finance and insurance, Art media, Other business facilities, Entertainment, Recreation and sports, Refueling station, Other public facilities business outlets, Other service facilities |
| Industrial, manufacturing (M) | 663 | Factory |
| Logistics and warehouse (W) | 3018 | Logistics warehousing |
| Road, street and transportation (S) | 120 | Urban Road, Urban rail transit, Transportation hub, Public transport station, Social parking lot, Other transportation facilities |
| Municipal utilities (U) | 1200 | Water supply, Power supply, Supply gas, Heat supply, Communication facilities, Radio and television facilities, Drainage facilities, Sanitation facilities, Environmental protection facilities, Firefighting facilities, Flood control facilities, Other public facilities |

there are only type I of POIs in the building, the function of building will be decided by the frequency of POI. Otherwise, the building function was decided by the priority order in type II as listed above. Specifically, residence is recognized as the primary usage regardless any auxiliary usage. Administration and public services, mainly including governmental office, school and hospital, were of greater public awareness over the rest types.

## IV. RESULTS

### A. Results of Identified Buildings Based on POI Data

The first and most efficient identification way is to associate POI data and building data by overlay analysis which decides building function depending on whether POI data are inside the any of the buildings. By this method, 21 701 buildings were assigned proper functions in the study area, including 8069 residential, 3780 administration and public services, 9232 Commercial and business facilities, 89 Industrial or manufacturing, 289 logistics and warehouses, 231 municipal utilities, and 11 road, street and transportation. Business, residential, and administrative buildings have been largely identified as there is a great number of POI data with these three categories in the study
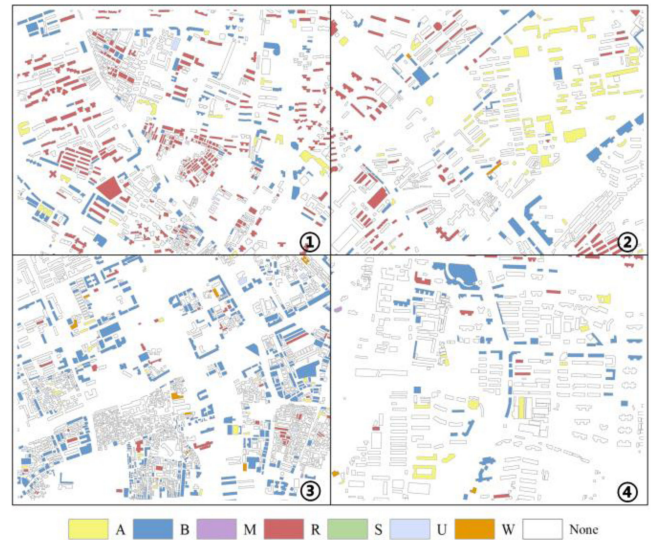


Fig. 5. Identification results of urban building functions based on POI data.

area. It was also found that only 35.15% POI data were inside the building or within the distance of 5 m, and 27.98% building were identified. That is to say large proportion of buildings could not be directly identified by POI data.

In detail, four typical areas, located in Jiang'an district, Qingshan district, Qiaokou district and Hongshan district, were selected to exhibit the identification results (see Fig. 5). Locations and extents of them have been labelled in Fig. 4. Obviously, four areas show different identification rate, and identified buildings in area ④ is far less than the other three areas. This is because area ④ is at the edge of core urbanized area in Wuhan, and POI data were less dense here. The results indicate that the spatial heterogeneity of POI data lead to the imbalance identification rate in different areas in the city. Area with higher POI density shows more identified buildings. Meanwhile, semantic information of 64.85% POI data remains unused in this step, which suggests that additional method is necessary if we would like to make full use of POI and improve the identification rate.

### B. Results of Identified Buildings Based on Spatial Similarity

Taking advantage of Google earth images and building footprint, eight textural features and eight geometric features, were calculated for 77 545 buildings in the study area. Notably, although textural analysis was performed for red, green, and blue band of Google earth images, only the textural features derived from red band were used. This is due to the high consistency among the results from different bands, which has been proven by the high values of correlation coefficients. Moreover, there are similar patterns between different textural features as shown in Fig. 6. For instance, results of variance [see Fig. 6 (b)] and contrast [see Fig. 6 (d)] look like each other. Also, the similar trend among different geometric features could be observed. Hence, PCA was conducted to reduce the dimensionality and keep the most important information. It can be seen from component matrix as given in Table II that there are two textural
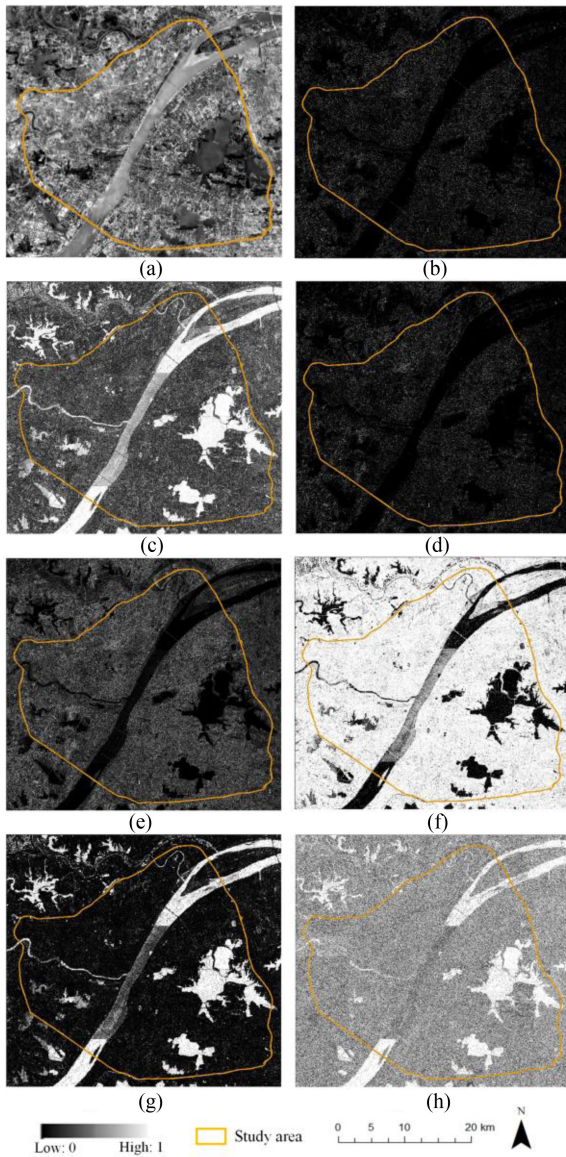
Fig. 6. Eight textural features of Google Earth image in the study area.

TABLE II
COMPONENT MATRIX OF SPATIAL FEATURES

| Features | F1 | F2 | F3 | F4 | F5 |
|---|---|---|---|---|---|
| Mean | 0.242 | -0.013 | 0.152 | 0.360 | 0.573* |
| Variance | 0.892* | 0.039 | -0.008 | 0.119 | 0.043 |
| Homogeneity | -0.908* | -0.029 | 0.025 | 0.171 | 0.071 |
| Contrast | 0.891* | 0.016 | -0.011 | 0.118 | -0.112 |
| Dissimilarity | 0.9608* | 0.046 | -0.021 | 0.020 | -0.121 |
| Entropy | 0.731* | -0.114 | 0.012 | -0.214 | 0.100 |
| Second Moment | -0.567* | -0.328 | 0.074 | 0.337 | 0.224 |
| Correlation | -0.118 | -0.076 | 0.092 | 0.145 | 0.791* |
| Area | -0.073 | -0.003 | -0.218 | 0.850* | -0.072 |
| Perimeter | -0.034 | 0.393 | -0.539 | 0.630* | -0.116 |
| Compactness | -0.010 | -0.785* | 0.527 | -0.103 | -0.018 |
| Nodes | 0.012 | 0.015 | -0.867* | 0.205 | -0.012 |
| Radius shape index | 0.002 | 0.890* | -0.217 | 0.059 | 0.035 |
| regularity | -0.016 | -0.135 | 0.915* | -0.024 | -0.005 |
| Aspect Ratio | 0.049 | 0.898* | 0.193 | -0.032 | -0.142 |
| Orientation | 0.015 | -0.004 | 0.588* | -0.167 | -0.063 |

∗Indicates the largest loading of each feature on a certain component among the five.
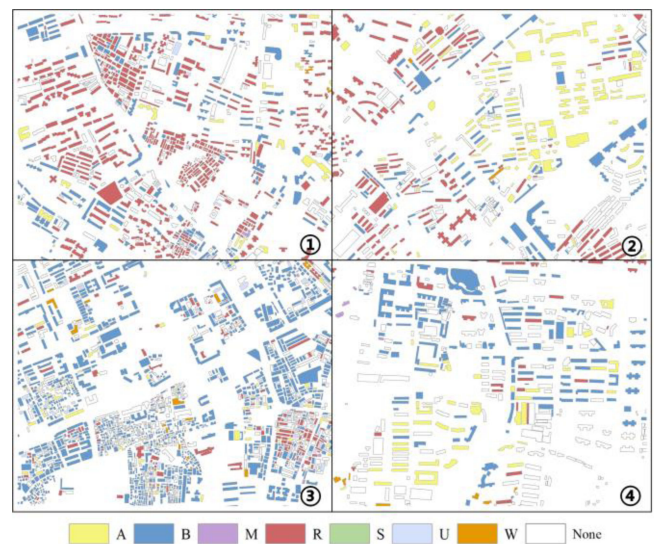


Fig. 7. Identification results of urban building functions based on spatial similarity method.

components and three geometric components extracted from the raw sixteen spatial features.

The spatial feature vectors consisting of the five components were calculated and used for measuring the cosine similarity between the unidentified and identified buildings in the buffer zone. Under the assumption that nearby buildings with similar spatial features are of similar function, spatial similarity contributes to assign functions for 33.67% buildings, including 7496 residential, 4213 administration, 13204 commercial, 278 industrial, 475 logistics and warehouses, 413 municipal utilities, and 7 transportation. The identification results for the selected four areas by spatial similarity are shown in Fig. 7.

Comparing with the identification results based on POI, a large amount of buildings was labeled, especially for commercial (area ③) and residential (area ①) usage. Areas ② and ④ have less identified buildings because POI-based method did not provide sufficient examples that can be used to infer the function

of nearby buildings. That is to say, the identification performance of spatial similarity method heavily depends on the sample sets, namely identified buildings based on POI. This method greatly improved the identification rate with the assistant of remote sensing imagery, but still could not finish the identification left by the first step.

## C. Results of Identified Buildings Based on KDE

To fully make use of POI data, we introduced KDE to complete the identification. Based on the kernel density of POI in different categories, the rest 38.38% buildings were all assigned proper functions, including 2897 residences and 26 864 commercial and business facilities. The number of Commercial and business facilities is much higher than that of residential,
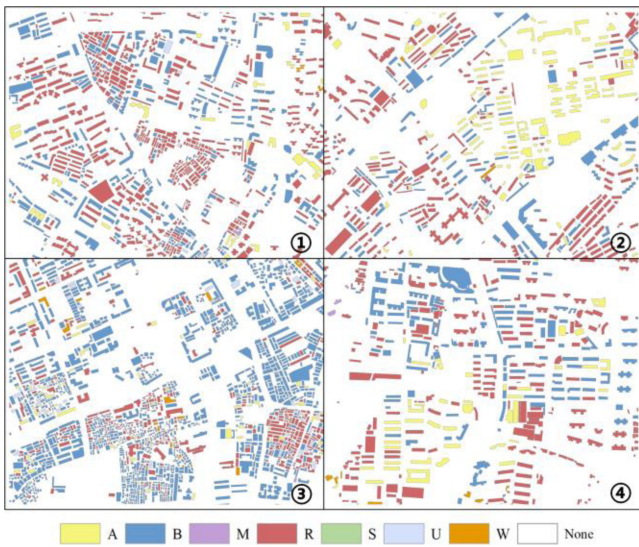
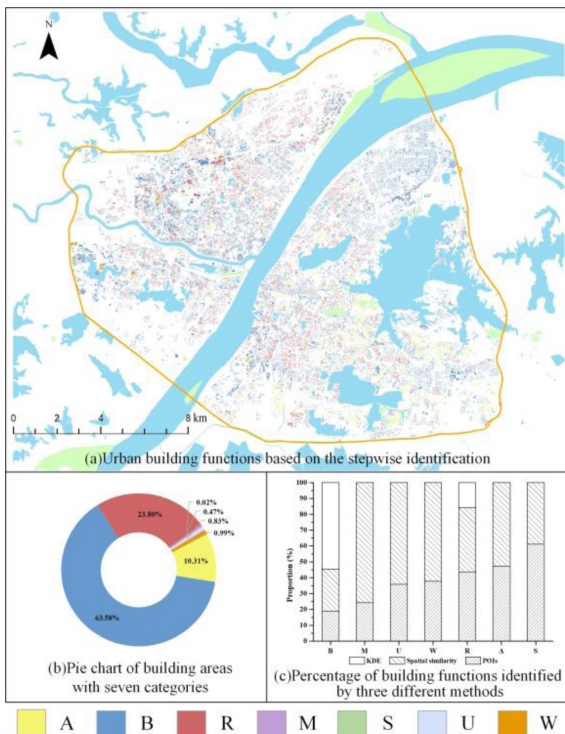Fig. 8. Identification results of urban building functions based on KDE.



Fig. 9. Identification results of urban building functions based on the POI data, spatial similarity and KDE method.

which is related to the number of POIs and the spatial patterns of different categories of buildings. As shown in Fig. 8, the significant improvement can be seen in areas ② and ④ that KDE mainly complements the identification of residential and commercial buildings.

In the study area, the identification results of building functions based on three different methods are shown in Fig. 9. Besides 27.98% building functions directly identified by POI

data, spatial similarity and KDE method contributed to the identification of 33.64% residential building and 38.38% business buildings, respectively. As shown in Fig. 9(b), the primary types are business, residential and administrative buildings which occupy 63.58%, 23.80%, and 10.31% of the entire building areas in the study area. Moreover, the proportion of each category of identified buildings based on the three methods were calculated [see Fig. 9(c)], and the results suggest that spatial similarity and KDE made great improve of building function identification on the basis of POI-based results.

## V. DISCUSSION

### A. Performance Assessment of Identification Results

As the proposed step-wise identification integrates spatial similarity and KDE, it is necessary to compare its performance with the single method. This could be done through comparing the identification accuracy between different methods. Hence, we separated buildings identified based on POI data as training sample and test sample in order to test the performance of spatial similarity and KDE. That is to say, the identification results of test samples were compared with the results based on POI to calculate the producer's accuracy, user's accuracy, and overall accuracy (OA) as given in Table III. Notably, most of Road, street and transportation can be directly identified based on the POI data, it is unnecessary to include this type when using spatial similarity or KDE. For spatial similarity method, the OA is 0.68, and the highest accuracy is 0.78 for residential buildings. The accuracy of administration, commercial and residential buildings are much higher than the other three types. Notably, there is a significant difference between producer's accuracy and user's accuracy of Municipal utilities as many municipal utilities have been incorrectly identified as commercial and business facilities. This is mainly due to retail stores are always mixed with factories.

Because majority of POI belongs to residential and commercial types (see Table I), and they are the main building functions in urban area, KDE method was only employed to identify these two categories. Although the accuracy rates based on KDE for residential (0.65), and commercial and business (0.67) are lower than that of spatial similarity, the identification rate greatly increase from 61.62% to 100% for whole buildings in study area. After identification based on POI recognition method, spatial similarity and KDE, the accuracy rates of commercial and business and residential are 0.66 and 0.78, respectively. Therefore, stepwise identification of urban building functions, which merges the KDE of POI data and the spatial similarities of building data can improve the identification accuracy and efficiency effectively.

To further examining the accuracy of identified building function by proposed method, the results were compared with urban planning map of Wuhan from Wuhan planning and design institute.[3] Since the urban planning map only exhibits the function of each street blocks, the building function identified in this article could not compare with it directly. Instead, the use of each

[3]http://www.wpdi.cn/project-1-i_11296.htm

TABLE III
ACCURACY OF IDENTIFYING BUILDINGS BASED ON THE SPATIAL SIMILARITY AND KDE

| Method | Primary Categories | Producer's Accuracy | User's Accuracy | Overall Accuracy |
|---|---|---|---|---|
| Spatial similarity | Residential | 0.78 | 0.75 | |
| | Administration and public services | 0.45 | 0.53 | |
| | Commercial and business facilities | 0.66 | 0.65 | 0.68 |
| | Industrial, manufacturing | 0.20 | 0.33 | |
| | Logistics and warehouse | 0.05 | 0.04 | |
| | Municipal utilities | 0.24 | 0.07 | |
| KDE | Residential | 0.65 | 0.63 | 0.66 |
| | Commercial and business facilities | 0.67 | 0.69 | |
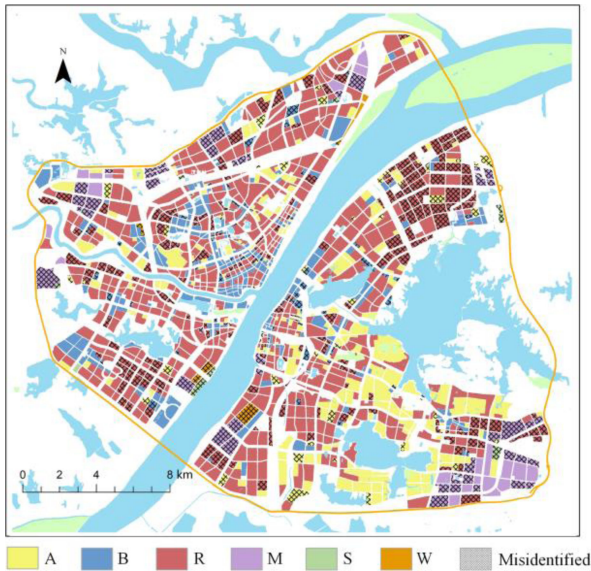


Fig. 10. Results of comparing identified urban functional zones between proposed method and urban planning map.

TABLE IV
SPATIAL SIMILARITY BETWEEN SAME/ DIFFERENT BUILDING TYPES

| Building type | with same type of building | | with other types of building | |
|---|---|---|---|---|
| | Average spatial similarity | Spatial similarity >0.5 | Average spatial similarity | Spatial similarity > 0.5 |
| Administration and public services | 0.54 | 58.77% | 0.31 | 27.64% |
| Commercial and business facilities | 0.58 | 63.87% | 0.29 | 26.66% |
| Residential | 0.65 | 64.52% | 0.25 | 18.51% |

street blockwas decided by the primary building function (largest amount) inside the block, and the results were compared with planning map. As shown in Fig. 10, the urban functional zones on planning map were reclassified to accord with the land use categories used in this article, and the misidentified blocks were shadowed. Statistically, there are 1372 blocks in the study area, and 71.33% of them were correctly identified, including 78.89% adminstrative blocks, 78.80% business blocks, 70% residential blocks, 37.50% warehouse blocks, and 32.86% manufactory blocks. The results suggest that the inferred function of buildings are quite reliable when identifying the urban functional zone especially for adminstrative, business and residential types.

In order to prove the hypnosis of spatial similarity method, the cosine similarities were measured for the buildings (identified by POIs) in the same type as well as in different types within the small distance (100 m). Notably, only three building types, administration, commercial and residential, were examined because the identification accuracy of them by spatial similarity is much higher than the other types (see Table III), and the results were given in Table IV. It can be observed that the average

spatial similarities of the same type were greatly higher than that of different types. Also, buildings in the same types have larger proportion of high spatial similarity, greater than 0.5. The results provide the evidence supporting the assumption behind building function identification based on spatial similarity.

## B. Impact of KDE Bandwidth Sensitivity on Building Identification Accuracy

It is generally believed that in KDE, the most critical issue is the selection of bandwidth [59], [60]. Improper bandwidth will seriously affect the accuracy of the KDE [61]. However, this is a parameter that usually decided by personal judgement without verification. Since identification results using KDE method are sensitive to the bandwidth, this article determines the bandwidth selection of KDE through a large number of experiments to find the optimal one that brings out the highest overall classification accuracy. Also, considering the effect of the proportions of training samples, 90%, 75%, and 50% of the identified buildings were selected as the training samples to exam the OA under different bandwidths.

As shown in Fig. 11, the similar curve pattern under the 90%, 5%, and 50% of training samples illustrates that the proportions of training samples have few influences on the identification results. Most surprisingly, we found interesting results for different and widths of KDE. It can be seen that OA of
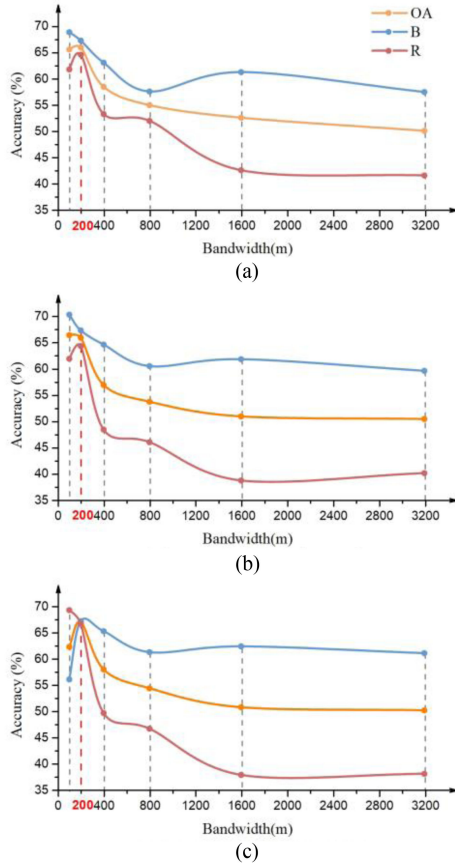
Fig. 11. Accuracy comparison of the identification results with different bandwidth.

TABLE V
MORAN'S I INDEX OF BUILDINGS IN DIFFERENT CATEGORIES

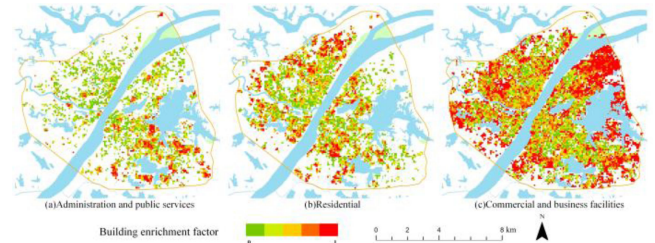| Primary Categories | Moran's I | z-value |
|---|---|---|
| Commercial and business facilities | 0.57** | 93.12 |
| Residential | 0.54** | 87.93 |
| Administration and public services | 0.5** | 81.13 |
| Industrial, manufacturing | 0.39** | 16.11 |
| Logistics and warehouse | 0.3** | 15.96 |
| Municipal utilities | 0.14** | 8.15 |
| Road, street and transportation | 0.12 | 1.4 |

∗∗Moran's I index is significant at the 0.01 level.



Fig. 12. Enrichment factors of buildings in different categories.

urban building functions using KDE method shows an increasing trend from 100 to 200 meter and a decreasing trend after 200 meter. Although the trends for residential buildings (R) and the commercial buildings (B) are slightly different from that of the OA, we can still conclude that 200 m is the optimal bandwidth in identification of urban building functions based on the KDE method. This is perhaps because that the average size of the street blocks in study area is about 200 m, and buildings in the street blocks are more likely of homogeneous functions. So that it is more appropriate to interpret building functions according to the spatial characteristics extracted from KDE at 200 m scale.

## C. Examining the Spatial Clustering of Buildings With Different Functions

Different types of buildings exhibit different clustering in space, and this aggregation feature can be taken advantage of to identify the type of building [62]. Through the above experiments results, we observed that different types of buildings have different aggregation patterns, but this is only the inference from observation. In this section, we will use two indicators to quantify the spatial clustering of different types of buildings from the macro and microperspectives.

In order to measure the spatial auto-correlation of the buildings with different functions, we calculated global Moran's I index for each category (see Table V), the Moran's I index and z-value of three building functions, including commercial and business facilities, residential, and administration and public services are much higher than the other building functions, which demonstrates that these three types of buildings have significant spatial clustering characteristics. While the other types of buildings show few or none clustering pattern since they consist of extremely low proportions of urban buildings in the entire study area.

To furtherly explore the spatial clustering characteristics of buildings at microlevel, we introduced the building enrichment factor (BEF) [31]. A large BEF value indicates high spatial aggregation degree. As shown in Fig. 12, we calculate the BEF of three building types with high spatial autocorrelation. Compared with other two types, most BEFs for Commercial and business facilities [see Fig. 11(c)] are close to one, which suggests that this type of buildings has a much higher degree of enrichment. This is because that Wuhan is the most important economical center in central China and one amongst the new first tier cities. Moreover, residential buildings are commonly surrounded with commercial buildings as we can see that where the red areas are clustered in Fig. 11 (b) mostly fill up the green areas in Fig. 11 (c). On the one hand, areas where residence clustered are attractive for those who want to start their business. On the other hand, places where full of business facilities providing convenient services are the optimal location for living in the cities. As to the Administration and public buildings, they present the lowest enrichment with few red areas [see Fig. 11 (a)]. These buildings

are the urban infrastructure for public service, such as administrative offices, universities, primary and secondary schools, nursing homes, hospitals, etc. It makes sense that this type of facilities should be scattered among the city to serve citizens as much as possible. In general, residential and business buildings have obvious spatial clustering characteristics all over the study area, while Administration and public services buildings show clustering characteristics in local areas and relatively dispersed at macrolevel.

## VI. CONCLUSION

Along with the rapid urbanization, the diversity and complexity of building functions have increased in metropolitan areas. Identifying land use at a finer scale contributes to a new perspective for urban design, environment monitoring, and emergency management, etc. To address this issue, this article proposes a stepwise identification framework for urban building functions based one remote sensing imagery and POI data, which merges the spatial similarity of buildings and KDE of POI data. The main conclusions can be drawn as follows.

1) Two proposed methods, spatial similarities based on Google earth images and building geometry, and KDE based on POI data, have proven efficient ways of identifying urban building functions as they merge the strength of geometrical characteristic of building and spatial distribution of POI data. This has improved identification rate up to 100% with high OA.

2) The identification results are sensitive to the bandwidth of KDE method. In this article, the iterative methods with a series of bandwidth indicate that 200 m is the optimal size in identification of urban building functions.

3) Different functional buildings exhibit various spatial correlation and spatial agglomeration. Notably, residential and commercial buildings are significantly more clustered than other types of urban buildings at both macro and microlevels. A deep insight into spatial characterizes of different functional buildings assists policy makers in optimizing the effectiveness of urban management.

Although the achievements in this article, there are still some aspects deserving being investigated in future study. Since only POI and Google earth images were used in this article, it is necessary to improve the proposed method by incorporating with multiple UGC and remote sensing data sources, such as taxi GPS trajectories data, social media data, remote sensing of night-time light, etc., to further enrich the semantic information regarding the urban building function. Also, we could promote this methodology to identify building functions in other cities around the world to enhance the applicability of the proposed identification framework.

## REFERENCES

[1] X. Liu *et al.*, "Classifying urban land use by integrating remote sensing and social media data," *Int. J. Geographical Inf. Sci.*, vol. 31, no. 8, pp. 1675–1696, 2017.

[2] L. Zhuo, Q. Shi, C. Zhang, Q. Li, and H. Tao, "Identifying building functions from the spatiotemporal population density and the interactions of people among buildings," *ISPRS Int. J. Geo-Inf.*, vol. 8, no. 6, 2019, Art. no. 247.

[3] H. R. Pourghasemi *et al.*, "Assessment of urban infrastructures exposed to flood using susceptibility map and Google Earth Engine," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 1923–1937, Dec. 2021.

[4] C. Zhong, X. Huang, S. Müller Arisona, G. Schmitt, and M. Batty, "Inferring building functions from a probabilistic model using public transportation data," *Comput., Environ. Urban Syst.*, vol. 48, pp. 124–137, 2014.

[5] S. Srivastava, J. E. Vargas-Muñoz, D. Swinkels, and D. Tuia, "Multilabel building functions classification from ground pictures using convolutional neural networks," in *Proc. 2nd ACM SIGSPATIAL Int. Workshop AI Geographic Knowl. Discov.*, pp. 43–46, 2018.

[6] M. Li, S. Zang, B. Zhang, S. Li, and C. Wu, "A review of remote sensing image classification techniques: The role of spatio-contextual information," *Eur. J. Remote Sens.*, vol. 47, no. 1, pp. 389–411, 2014.

[7] J. Arndt and D. Lunga, "large-scale classification of urban structural units from remote sensing imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 2634–2648, Jan. 2021.

[8] Y. Chang, K. Hou, X. Li, Y. Zhang, and P. Chen, "Review of land use and land cover change research progress," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 113, 2018, Art. no. 012087.

[9] P. Treitz, "Remote sensing for mapping and monitoring land-cover and land-use change," *Prog. Plan.*, vol. 64, no. 4, pp. 269–279, 2004.

[10] X. Y. Bian, C. Chen, L. Tian, and Q. Du, "Fusing local and global features for high-resolution scene classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 6, pp. 2889–2901, Jun. 2017.

[11] W. Zhou, D. P. Ming, X. W. Lv, K. Q. Zhou, H. Q. Bao, and Z. L. Hong, "SO-CNN based urban functional zone fine division with VHR remote sensing image," *Remote Sens. Environ.*, vol. 236, Jan 2020, Art. no. 111458.

[12] Z. Y. Lv, T. F. Liu, and J. A. Benediktsson, "Object-oriented key point vector distance for binary land cover change detection using VHR remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 9, pp. 6524–6533, Sep. 2020.

[13] X. Z. Shi, S. L. Fu, J. Chen, F. Wang, and F. Xu, "Object-level semantic segmentation on the high-resolution gaofen-3 FUSAR-Map dataset," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 3107–3119, 2021.

[14] G. Pulighe, V. Baiocchi, and F. Lupia, "Horizontal accuracy assessment of very high resolution Google Earth images in the city of Rome, Italy," *Int. J. Digit. Earth*, vol. 9, no. 4, pp. 342–362, 2015.

[15] M. F. Goodchild *et al.*, "Next-generation digital Earth," *Proc. Nat. Acad. Sci. USA*, vol. 109, no. 28, pp. 11088–11094, Jul. 2012.

[16] Q. Hu *et al.*, "Exploring the use of Google Earth imagery and object-based methods in land use/cover mapping," *Remote Sens.*, vol. 5, no. 11, pp. 6026–6042, 2013.

[17] R. K. Horota *et al.*, "Printgrammetry-3-D model acquisition methodology from Google Earth imagery data," (in English), *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 2819–2830, May 2020.

[18] J. L. San Emeterio and C. Mering, "Mapping of African urban settlements using Google Earth images," (in English), *Int. J. Remote Sens.*, vol. 42, no. 13, pp. 4886–4901, 2021.

[19] Y. Zhong, Y. Su, S. Wu, Z. Zheng, and L. Zhang, "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, 2020, Art. no. 111838.

[20] D. Yang, C.-S. Fu, A. C. Smith, and Q. Yu, "Open land-use map: A regional land-use mapping strategy for incorporating OpenStreetMap with earth observations," *Geo-Spatial Inf. Sci.*, vol. 20, no. 3, pp. 269–281, 2017.

[21] M. F. Goodchild, "Citizens as voluntary sensors: Spatial data infrastructure in the world of Web 2.0," *Int. J. Spatial Data Infrastructures Res.*, vol. 2, pp. 24–32, 2007.

[22] S. Elwood, "Volunteered geographic information: Future research directions motivated by critical, participatory, and feminist GIS," *GeoJ.*, vol. 72, no. 3, pp. 173–183, 2008.

[23] H. Wu, A. Lin, K. C. Clarke, W. Shi, A. Cardenas-Tristan, and Z. Tu, "A comprehensive quality assessment framework for linear features from volunteered geographic information," *Int. J. Geographical Inf. Sci.*, vol. 35, no. 9, pp. 1826–1847, 2021.

[24] A. Lin *et al.*, "A big data-driven dynamic estimation model of relief supplies demand in urban flood disaster," *Int. J. Disaster Risk Reduction*, vol. 49, 2019, Art. no. 101682.

[25] X. Zhu and C. Zhou, "POI inquiries and data update based on LBS," in *Proc. Int. Symp. Inf. Eng. Electron. Commerce*, pp. 730–734, 2009.

[26] K. Liu, P. Qiu, S. Qiu, F. Lu, and L. Yin, "Investigating urban metro stations as cognitive places in cities using points of interest," *Cities,* vol. 97, 2020, Art. no. 102561.

[27] N. Luo, T. Wan, H. Hao, and Q. Lu, "Fusing high-spatial-resolution remotely sensed imagery and openstreetmap data for land cover classification over urban areas," *Remote Sens.,* vol. 11, no. 1, 2019, Art. no. 88.

[28] C. C. Fonte and N. Martinho, "Assessing the applicability of openstreetmap data to assist the validation of land use/land cover maps," *Int. J. Geographical Inf. Sci.,* vol. 31, no. 12, pp. 2382–2400, 2017.

[29] H. Wu, A. Lin, X. Xing, D. Song, and Y. Li, "Identifying core driving factors of urban land use change from global land cover products and POI data using the random forest method," *Int. J. Appl. Earth Observ. Geoinf.,* vol. 103, 2021, Art. no. 102475.

[30] X. Zhang, W. Li, F. Zhang, R. Liu, and Z. Du, "Identifying urban functional zones using public bicycle rental records and point-of-interest data," *ISPRS Int. J. Geo-Inf.,* vol. 7, no. 12, 2018, Art. no. 459.

[31] Y. Hong and Y. Yao, "Hierarchical community detection and functional area identification with OSM roads and complex graph theory," *Int. J. Geographical Inf. Sci.,* vol. 33, no. 8, pp. 1569–1587, 2019.

[32] R. Andrade, A. Alves, and C. Bento, "POI mining for land use classification: A case study," *ISPRS Int. J. Geo-Inf.,* vol. 9, no. 9, 2020, Art. no. 493.

[33] X. Zhang, S. Du, and Q. Wang, "Hierarchical semantic cognition for urban functional zones with VHR satellite images and POI data," *ISPRS J. Photogramm. Remote Sens.,* vol. 132, pp. 170–184, 2017.

[34] 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, 2020, Art. no. 111838.

[35] H. Liu et al., "Recognizing urban functional zones by a hierarchical fusion method considering landscape features and human activities," *Trans. GIS,* vol. 24, no. 5, pp. 1359–1381, 2020.

[36] C. C. Fonte, M. Minghini, V. Antoniou, J. Patriarca, and L. M. See, "Classification of building function using available sources of VGI," *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.,* vol. XLII-4, pp. 209–215, 2018.

[37] Y. Song, Y. Long, P. Wu, and X. Wang, "Are all cities with similar urban form or not? Redefining cities with ubiquitous points of interest and evaluating them with indicators at city and block levels in China," *Int. J. Geographical Inf. Sci.,* vol. 32, no. 12, pp. 2447–2476, 2018.

[38] J. Jokar Arsanjani, M. Helbich, M. Bakillah, J. Hagenauer, and A. Zipf, "Toward mapping land-use patterns from volunteered geographic information," *Int. J. Geographical Inf. Sci.,* vol. 27, no. 12, pp. 2264–2278, 2013.

[39] X. Liu and Y. Long, "Automated identification and characterization of parcels with openstreetmap and points of interest," *Environ. Planning B, Planning Des.,* vol. 43, no. 2, pp. 341–360, 2016.

[40] Y. Hu and Y. Han, "Identification of urban functional areas based on POI data: A case study of the Guangzhou economic and technological development zone," *Sustainability,* vol. 11, no. 5, 2019, Art. no. 1385.

[41] Y. Jing, Z. Yu, and X. Xing, "Discovering regions of different functions in a city using human mobility and POIs," in *Proc., 18th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, pp. 186–194, 2012.

[42] B. Dash, D. Mishra, A. K. Rath, and M. Acharya "A hybridized K-means clustering approach for high dimensional dataset," *Int. J. Eng., Sci. Technol.,* vol. 2, pp. 59–66, 2010.

[43] 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," *Sustainability,* vol. 8, no. 11, 2016, Art. no. 1202.

[44] X. Liu, Y. Tian, X. Zhang, and Z. Wan, "Identification of urban functional regions in Chengdu based on taxi trajectory time series data," *ISPRS Int. J. Geo-Inf.,* vol. 9, no. 3, 2020, Art. no. 158.

[45] W. Yu and T. Ai, "The visualization and analysis of urban facility pois using network kernel density estimation constrained by multi-factors," *Boletim de Ciências Geodésicas,* vol. 20, no. 4, pp. 902–926, 2014.

[46] R. R. Boyce and W. A. Clark, "The concept of shape in geography," *Geographical Rev.,* vol. 54, no. 4, pp. 561–572, 1964.

[47] R. Trias-Sanz, G. Stamon, and J. Louchet, "Using colour, texture, and hierarchial segmentation for high-resolution remote sensing," *ISPRS J. Photogramm. Remote Sens.,* vol. 63, no. 2, pp. 156–168, 2008.

[48] A. Gebejes and R. Huertas, "Texture characterization based on grey-level co-occurrence matrix," *Databases,* vol. 9, no. 10, pp. 375–378, 2013.

[49] M. Hall-Beyer, "Practical guidelines for choosing GLCM textures to use in landscape classification tasks over a range of moderate spatial scales," *Int. J. Remote Sens.,* vol. 38, no. 5, pp. 1312–1338, 2017.

[50] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley Interdisciplinary Rev., Comput. Statist.,* vol. 2, no. 4, pp. 433–459, 2010.

[51] Y. Wang, Y. Gu, M. Dou, and M. Qiao, "Using spatial semantics and interactions to identify urban functional regions," *ISPRS Int. J. Geo-Inf.,* vol. 7, no. 4, 2018, Art. no. 130.

[52] C. Lu et al., "Mapping urban spatial structure based on POI (point of interest) data: A case study of the central city of Lanzhou, China," *ISPRS Int. J. Geo-Inf.,* vol. 9, no. 2, 2020, Art. no. 92.

[53] D. Khosrow, "Density estimation for statistics and data analysis," *Technometrics,* vol. 29, no. 4, pp. 495–495, 2012.

[54] T. Hu, J. Yang, X. Li, and P. Gong, "Mapping urban land use by using Landsat images and open social data," *Remote Sens.,* vol. 8, no. 2, 2016, Art. no. 151.

[55] F. Gao et al., "Mapping impervious surface expansion using medium-resolution satellite image time series: A case study in the Yangtze River Delta, China," *Int. J. Remote Sens.,* vol. 33, no. 24, pp. 7609–7628, 2012.

[56] H. Wu et al., "Examining the sensitivity of spatial scale in cellular automata Markov chain simulation of land use change," *Int. J. Geographical Inf. Sci.,* vol. 33, no. 5, pp. 1040–1061, 2019.

[57] Y. Deng, J. Liu, Y. Liu, and A. Luo, "Detecting urban polycentric structure from POI data," *ISPRS Int. J. Geo-Inf.,* vol. 8, no. 6, 2019, Art. no. 283.

[58] S. Wang, G. Xu, and Q. Guo, "Street centralities and land use intensities based on points of interest (POI) in Shenzhen, China," *ISPRS Int. J. Geo-Inf.,* vol. 7, no. 11, 2018, Art. no. 425.

[59] J. Chen, Y. Zhang, and Y. Yu, "Effect of MAUP in spatial autocorrelation," *Acta Geographica Sinica,* vol. 66, no. 12, pp. 1597–1606, 2011.

[60] M. S. Gerber, "Predicting crime using Twitter and Kernel density estimation," *Decis. Support Syst.,* vol. 61, pp. 115–125, May 2014.

[61] N.-B. Heidenreich, A. Schindler, and S. Sperlich, "Bandwidth selection for kernel density estimation: A review of fully automatic selectors," *AStA Adv. Stat. Anal.,* vol. 97, no. 4, pp. 403–433, 2013.

[62] W. Yue, "A literature review of spatial clustering applications," *Guide Sci. Educ.,* no. 9, pp. 30–32, 2017.
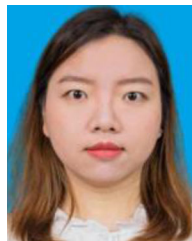
**Anqi Lin** received the Master degree in geographic information science from the University of Pittsburgh, Pittsburgh, PA, USA, in 2016. She is currently working toward the Ph.D. degree in geographic information science with the College of Urban and Environmental Sciences, Central China Normal University, Wuhan, China.

Her research interests include spatial big data analysis and mining, land use mapping, and urban environment modeling.



**Xiaomeng Sun** received the B.S. degree in geographic information science in 2021 from Central China Normal University, Wuhan, China, where he is working toward the Master degree in geographic information science.

His research interests include geospatial data mining and machine learning.



**Hao Wu** received the B.S degree in surveying engineering in 2000 from Wuhan University, where he received Ph.D. degree in geographic information science in 2007.

He is currently a Professor with the College of Urban and Environmental Sciences, Central China Normal University, Wuhan, China. His research interests focus on Earth observation science, urban environmental simulation, and volunteer geographic information.

**Wenting Luo** received the B.S. degree in geographic information science from Jiangxi Normal University, Nanchang, China, in 2021. She is currently working toward the Master's degree in geographic information science with the College of Urban and Environmental Sciences, Central China Normal University, Wuhan, China.

Her research interests are big data analysis and land use simulation.

**Zhongming Wang** received the B.S. degree in geographic information science from Central China Normal University, Wuhan, China, in 2021.

His research interests include spatial data analysis and machine learning.

**Danyang Wang** received the B.S. degree in geographic information science from Central China Normal University, Wuhan, China, in 2021.

Her research interests are big data analysis and land use simulation.

**Lanting Zhao** received the B.S. degree in human geography and urban and rural planning, in 2021, from Central China Normal University, Wuhan, China, where she is currently working toward the Master's degree in human geography with the College of Urban and Environmental Sciences.

Her research interests include urban land use mapping and user generate content.

**Dantong Zhong** received the B.S. degree in geographic information science, in 2021, from Central China Normal University, Wuhan, China, where she is currently working toward the Master's degree in Geographic information science with the College of Urban and Environmental Sciences.

Her research interests include environmental remote sensing and urban land use mapping.

**Jiang Zhu** received the B.S. degree in computer software from Jiangxi Normal University, in 2006. He is currently the general manager of software department, KQ GEO Technologies Co., Ltd, Beijing, China. He is dedicated to land use monitoring and software development.