Article

Sub-Block Urban Function Recognition with the Integration of Multi-Source Data

Baihua Liu 1,2,†, Yingbin Deng 2,3,†, Xin Li 1,2, Miao Li 1,* , Wenlong Jing 2,3, Ji Yang 2,3, Zehua Chen 4 and Tao Liu 1,2

1 College of Geographical Science, Harbin Normal University, Harbin 150025, China
2 Guangdong Open Laboratory of Geospatial Information Technology and Application, Laboratory of Guangdong for Utilization of Remote Sensing and Geographical Information System, Guangzhou Institute of Geography, Guangdong Academy of Sciences, Guangzhou 510070, China
3 Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), Guangzhou 511485, China
4 Guangdong Provincial Institute of Land Surveying & Planning, Guangzhou 510075, China
* Correspondence: mli@hrbnu.edu.cn
† These authors contributed equally to this work.

Abstract: The recognition of urban functional areas (UFAs) is of great significance for the understanding of urban structures and urban planning. Due to the limitation of data sources, early research was characterized by problems such as singular data, incomplete results, and inadequate consideration of the socioeconomic environment. The development of multi-source big data brings new opportunities for dynamic recognition of UFAs. In this study, a sub-block function recognition framework that integrates multi-feature information from building footprints, point-of-interest (POI) data, and Landsat images is proposed to classify UFAs at the sub-block level using a random forest model. The recognition accuracies of single- and mixed-function areas in the core urban area of Guangzhou, China, obtained by this framework are found to be significantly higher than those of other methods. The overall accuracy (OA) of single-function areas is 82%, which is 8–36% higher than that of other models. The research conclusions show that the introduction of the three-dimensional (3D) features of buildings and finer land cover features can improve the recognition accuracy of UFAs. The proposed method that uses open access data and achieves comprehensive results provides a more practical solution for the recognition of UFAs.

Keywords: multi-source data; urban functional area; sub-block; random forest

1. Introduction

Cities are composed of various functional areas that can effectively perform specific urban functions, such as residential, commercial, and industrial areas. The functional differences in these regions are the result of the combined effects of different socioeconomic activities. Today, the rapid development of urbanization has led to drastic changes in the functional layout of cities; this has disrupted the original interconnected and stable internal spatial structure of cities, and has also caused various problems in urban development [1,2]. Urban functional areas (UFAs), as the basic unit of urban planning and management, are affected not only by planning policies, but also by the way of civil life, and change with urban development. As the distribution of UFAs is related to human activities, it is of great significance to divide UFAs and understand their spatial distribution and characteristics for the coordination of human–land relationships, urban planning management, environmental change monitoring, and resource allocation optimization [3–6].

Traditionally, UFAs recognition has mainly relied on qualitative methods such as statistical investigation and expert judgment [7,8]. However, the data are difficult to obtain, and the investigation and analysis require large amounts of time and labor costs. Moreover,
the reliability of the results is also characterized by problems such as strong subjectivity and poor timeliness [9,10].

With the development of remote sensing technology, scholars have found that remote sensing images can effectively extract ground object information [11]. The dynamic monitoring of urban land-use changes can be realized by analyzing remote sensing images in different periods; thus, remote sensing images are widely used in land-use and land cover research [12,13]. Many studies have explored the possibility of using remote sensing images to classify urban land cover and functional areas [14,15]. These methods display remarkable advantages in extracting the physical properties of ground objects. However, the function of urban land is not only related to the physical properties of land blocks, but is also affected by human activity, and it is difficult to determine the socioeconomic characteristics of different regions by relying solely on remote sensing image information [16].

In recent years, social sensing data, such as point-of-interest (POI) data, travel trajectory data, and social media check-in data, have provided new opportunities for research on urban internal regional functions [17–19]. These emerging geospatial data are rich in the semantic features of human mobility and socioeconomic activities, thus contributing to a better understanding of the functional interaction patterns and social structures of urban environments [20]. Among them, POI data are favored because of the close correlation between category labels and land-use types and the relative ease of data acquisition [21]. However, the performance of methods that use only POI data is affected by the data coverage. For underdeveloped areas and privacy-conscious residential areas, the function classification of these areas is difficult to predict accurately due to the lack of POI data [22].

The integration of multi-source data can compensate for the shortcomings of single-source data in the recognition of UFAs, and can more accurately reveal the spatiotemporal characteristics of UFAs; thus, this has become an important direction in the achievement of the refined study of UFAs [23]. For example, functional area recognition has been performed based on spectral features, landscape features, and latent semantic features via the fusion of remote sensing images and social sensing data [24]. The maximum expectation algorithm has been used to extract the time-series features of trajectory data, and the Delphi method has been used to mine the semantic information of the density and category of POI data, combined with their analysis for realizing functional area identification [25]. The dominant function of each area has been identified based on the features of different functional types extracted from POIs and the distribution of the check-in intensities of different functional types estimated by the kernel density method [26]. The information mapped by different data is different, and the fusion analysis of multi-source data can more comprehensively reflect the regional features of ground objects, the characteristics of human activity, and the spatiotemporal changes in social information.

In previous studies, the fusion of remote sensing data and social sensing data, as well as the fusion of social sensing data with other social sensing data, has been common [27–30]. However, the joint use of these types of data with building information has been easily ignored in feature mining, and most researchers have used the spectral, texture, and spatial features of images without considering the relationship between different land cover features and UFAs [31]. In addition, the results of the existing studies are generally divided into several single-function areas, and a discussion of mixed-function areas is lacking [32]. As the main setting of human activities, buildings can reflect the urban form and regional functions to a certain extent, and there are obvious differences in the land cover composition within different functional areas; thus, it is important to understand the differences in the building morphology and finer land cover features of urban areas for function classification [33]. Moreover, in complex and heterogeneous urban areas, the function of areas is often not completely pure, and it would be more in line with the real situation to consider a functional mix.

To overcome these existing problems, this study proposes a functional area recognition framework that integrates multi-feature information and can successfully find single- and mixed-function areas in the study area. By deeply mining the multi-feature information of
building footprints, POIs, and remote sensing images, the fine recognition of functional areas at the sub-block level is achieved by using random forest classifiers. The contributions of this research are mainly divided into the following three points. (1) The introduction of the three-dimensional (3D) features of buildings and finer land cover features of images effectively improves the classification accuracy of UFAs. (2) By using open access multi-source data, the mapping of UFAs with coexisting single- and mixed-function areas is realized. (3) The accuracies of UFA recognition using different feature factors are compared, and the introduction of features is proven to be effective.

The remainder of this article is organized as follows. The study area and datasets used are described in Section 2. In Section 2.3, the methodology is explained. The experimental results and comparative analyses are presented in Section 3. Section 4 discusses the advantages and limitations of the research method and suggestions for future work. Finally, this study is summarized in Section 5.

2. Materials and Methods

2.1. Study Area

Guangzhou is the capital city of Guangdong Province, China. It is located at the mouth of the Pearl River and is the core of the Guangdong–Hong Kong–Macao Greater Bay Area. As of 2020, the city had 11 districts with a total area of 7434.4 km², and its latitude and longitude range from 112°57′ to 114°3′E and 22°26′ to 23°56′N, with a permanent population of 18.67 million and an urbanization rate of 86.46%. The core urban areas of Guangzhou were selected for investigation, including the four municipal districts of Liwan, Yuexiu, Tianhe, and Haizhu (Figure 1), which are the most densely populated and economically active areas in the city. In recent years, Guangzhou has been expanding rapidly. To promote the healthy and sustainable development of Guangzhou, the reasonable planning of its UFAs should be considered. However, the urban planning maps designed by the government are difficult to update in time, so it is very important to map UFAs both accurately and promptly.

Figure 1. Area of study in the core area of Guangzhou, China. The remote sensing image is a Landsat 8 image obtained in 2020 with a spatial resolution of 30 m.
2.2. Data Sources

In this study, building footprints, POI data, and remote sensing data were used to estimate and classify UFAs (Table 1). Building footprints in 2020 were collected from the Gaode Map API, which included a total of 88,922 buildings with height information, and the area and perimeter indicators of each building were calculated to construct relevant features. The POI data in 2020 were also collected from the Gaode Map API. Each record included the latitude, longitude, name, address, telephone number, POI category, and administrative district, and a total of 72,489 POIs were selected. UFAs with the blocks formed by the road network were recognized as the basic unit. The road network data were downloaded from OpenStreetMap (OSM). The Landsat 8 remote sensing images were obtained from the Geospatial Data Cloud website. The images have a spatial resolution of 30 m and were imaged on 18 February 2020. The images were pre-processed according to radiometric calibration and atmospheric correction methods recommended by the United States Geological Survey (USGS) (http://glovis.usgs.gov/ accessed on 11 November 2021). Street View imagery provided by Baidu Street View and high-resolution historical Earth imagery provided by Google Earth were used to assist in the collection and verification of the functional area samples. All data were reprojected to the Universal Transverse Mercator projection for area 49N.

Table 1. The data sources.

| Data Type          | Function                                      | Sources                                                                 |
|--------------------|-----------------------------------------------|-------------------------------------------------------------------------|
| Building footprint | Extract the building features of the block     | Gaode map (https://lbs.amap.com/api accessed on 11 October 2021)         |
| POI                | Extract the social and economic characteristics of the block | Gaode map (https://lbs.amap.com/api accessed on 25 May 2021)          |
| Road network       | The basic data of blocks                      | OpenStreetMap (http://www.openstreetmap.org accessed on 26 April 2021) |
| Landsat 8 OLI image| Extract the physical characteristics of the block | Geospatial Data Cloud (http://www.gscloud.cn/sources accessed on 11 November 2021) |
| Street View        | Assistance for sample collection and verification | Baidu Street View (https://map.baidu.com/ accessed on 10 August 2022)   |
| Google Earth       | Assistance for sample collection and verification | Google Earth (https://earth.google.com/web/ accessed on 10 August 2022) |

2.3. Methodology

This study consisted of the following three steps, as presented in Figure 2: (1) UFA segmentation; based on topology processing, the revised road vector was used to segment the basic spatial unit of urban function classification; (2) multi-feature extraction from multi-source data; (3) sub-block UFA recognition.

2.3.1. Definition of UFA Types

The definition of UFAs is based on population mobility and social service functions. Five functional types of urban areas are defined in this study: residence, commercial, industry, public service, and green space. A block with a pure function is defined as a single-function area. In addition, there are some blocks with multiple functional uses simultaneously, which are divided into the following three main types: (1) mixed-function areas with a certain function as the dominant function, but their service effect and influence are far weaker than those of single-function areas; (2) mixed-function areas with multiple dominant functions that coexist obviously and have social functions with multiple functions; (3) mixed-function areas whose functions are not dominant and that have multiple functions simultaneously, but their various functions are not prominent.
2.3.2. Blocks Segmentation

The block is the basic unit that undertakes socioeconomic functions in urban management and urban planning [34]. Thus, research on the block scale is of great significance for urban cognition. The city was divided into blocks by the urban road network, which served as the basic units for the recognition of UFAs. However, road network vectors obtained from OSM are characterized by the problem of hanging points or independent roads, which, when they are directly used to partition the study area, will lead to internal incompleteness and discontinuity. Thus, it is necessary to deal with topological errors in the road network to form polygonal closed cells bounded by road networks. Based on this, the division of UFAs can be realized.

To cull the road space without compromising the building footprint, reference was made to the method of Wang et al. [35], and the following processing was performed on the OSM data: (1) for roads with suspension points, the originally unconnected roads were connected; (2) for independent roads, the deletion process was performed; (3) after eliminating topological errors, the reclassification and hierarchical buffer processing was performed for all categories of roads in OSM (Table 2). After eliminating the road space, the blocks bounded by roads were formed. A total of 1925 blocks were obtained as the object of urban function recognition in this study.

Table 2. The buffer processing of the OSM road network.

| Road Buffer Level       | Class in OSM                                      |
|-------------------------|---------------------------------------------------|
| First buffer (20 m)     | motorway, motorway_link, tertiary, tertiary_link, track, trunk, trunk_link |
| Second buffer (10 m)    | secondary, secondary_link                         |
| Third buffer (5 m)      | living_street, path, primary, primary_link, residential, service, unclassified, unknown |
2.3.3. Extraction of Building Features

Buildings and blocks are closely related to each other. As the major sites of human activities, differences between the physical features of buildings can often reflect the functions of areas [22,36]. Many studies have confirmed that building features do have the potential to distinguish functional urban areas [37,38].

Commonly used building features include the area (A), perimeter (P), and structural ratio (SR) of a building. The area represents the floor area of the building, the perimeter represents the perimeter of the building, and the structural ratio is the ratio of the building’s perimeter to its area. There are differences in the building features of different functional areas. For commercial areas, the values of the building area, perimeter, and structural ratio within the block may vary greatly. For example, shopping centers have large footprints and complex shapes, but business offices have small footprints and regular rectangular shapes. However, buildings in residential areas tend to be highly consistent when they are compared in terms of their building area, perimeter, and structural ratio. Because of the high similarity of buildings within residential areas, the values of these features fluctuate less. The area, perimeter, and structural ratio of each building were calculated, and their sum (Sum), average (Avg), and standard deviation (SD) were determined based on the block; this helped to distinguish different types of functional areas, especially commercial and residential areas.

The floor (F) indicates the number of floors of a building, which is different from the two-dimensional (2D) features such as the area, perimeter, and structural ratio of a building. The introduction of this feature can reflect the difference in the 3D features of different buildings. Buildings of different heights often carry different social functions. For example, commercial office buildings are usually located in the city center and have higher floors; buildings in residential areas are usually relatively uniform in height; factories in industrial areas usually have lower building heights; and green space areas have fewer buildings with lower floor heights. Therefore, the total number of floors (Sum_F), the average number of floors (Avg_F), and the standard deviation of floors (SD_F) within a block were counted to distinguish different categories of functional areas.

In addition, the total number (TN) and density (D) of buildings in the block were calculated. The total number of buildings refers to the sum of all buildings in the block. The density of buildings refers to the ratio of the sum of the floor areas of buildings to the area of the block. Generally, residential areas have a dense distribution of buildings within them, and both these and commercial areas have a higher building density. In contrast, green space areas differ significantly from other areas in that both the density and the total number of buildings are relatively low. These two factors can reflect the distribution of buildings in the block to some degree.

A total of 14 related factors of these building features were obtained. First, the area, perimeter, and structure ratio of each building were calculated using ArcGIS (Figure 3). Then, using the spatial joining method to join the buildings into their blocks, the indicators, including Sum_A, Avg_A, SD_A, Sum_P, Avg_P, SD_P, Sum_SR, Avg_SR, SD_SR, Sum_F, Avg_F, SD_F, TN, and D, were obtained via statistical calculation to represent the building features of the block.
Figure 3. The thematic maps of building footprints: (a) building area; (b) building perimeter; (c) building structure ratio; (d) building floor.

2.3.4. Extraction of Socioeconomic Features

POIs are geographical points with category labels generated by human economic activities, and have been demonstrated to play an important role in functional area recognition [39]. The collected POIs contained 14 categories, including catering services, landscapes, companies and enterprises, shopping services, transportation facilities services, financial insurance services, education and cultural services, commercial residences, public facilities, life services, sports and leisure services, medical care services, government agencies and social organizations, and accommodation services. These categories are numerous and cross-repeat. Therefore, reclassification was needed for the initial POIs [40], and, before reclassification, points with low public awareness, such as bus stops, public toilets, and logistics and express delivery, were eliminated [41]. Because many road spaces were eliminated, reference was made to the POI classification criteria to reclassify them into a total of five categories: residence, commerce, industry, public service, and green space (Figure 4). It is worth noting that the company and enterprise POIs were not completely classified as industrial; the company and enterprise POIs located in buildings with more than seven floors were classified into the business office category of commercial functionality [22].

Because POIs are points abstracted from actual buildings, the area of a POI entity within a functional area has a significant impact on the function of that unit. Reference was made to the general area and scoring criteria of POIs [42], the results of street view images and online search surveys were combined, and corresponding weights were assigned to each POI category (Table 3).
Because POIs are points abstracted from actual buildings, the area of a POI entity within a functional area has a significant impact on the function of that unit. Reference was made to the general area and scoring criteria of POIs [42], the results of street view images and online search surveys were combined, and corresponding weights were assigned to each POI category (Table 3).

Table 3. The POI area weight assignments.

| Area (m²)     | POI Class                                    | Score (Points) |
|---------------|----------------------------------------------|----------------|
| <500          | C2, C3, C6, C8, C9                          | 3              |
| 500–1000      | C1, C4, C5, C7, I1, P2, P4                  | 8              |
| 1000–5000     | C10, I2, P1, P3, R1                         | 40             |
| 5000–10,000   | G1, G2                                      | 80             |
| ≥10,000       | I3                                           | 100            |

To realize the diffusion of the facility service influence of POIs and weaken the dispersion of POIs, the kernel density estimation method was used to analyze the POIs [43]. Related studies have shown that the kernel function has little effect on the kernel density estimation results, while the effect of the bandwidth is stronger [44]. By comparing the kernel density effects of different bandwidths (300 m, 500 m, 800 m, and 1000 m), 500 m was determined as the optimal bandwidth for this study. The kernel density estimation results of different categories of POIs are shown in Figure 5.
Remote sensing images can reflect the real condition of ground object distribution. Based on previous studies [31,45,46], spectral and texture features were used for recognition, and finer land cover features were introduced to enhance the recognition effect. First, radiometric calibration and atmospheric correction were performed on the Landsat 8 OLI images using ENVI software.

(1) The principal component eigenvalues obtained by principal component analysis of the images represent the spectral features of the ground objects [47,48]. The principal component analysis can be used to highlight image information, so that each band of image data has different information about the strength or weakness of the features. All the original band image data were selected for the principal component analysis transformation, and the desired information was enhanced and highlighted, while other information was suppressed. After principal component transformation, the difference of information in different bands was analyzed. Finally, the obtained first principal component (i.e., the transformed band 1) was determined to represent the spectral features of the image.

(2) The texture features of the images were described by the gray level co-occurrence matrix (GLCM) [46,49]. Eight Haralick texture measures [50], including the second moment, mean, homogeneity, entropy, dissimilarity, correlation, contrast, and variance, were selected as the texture feature factors of the images. These can reflect the image grayscale information on the direction, adjacent interval, and change magnitude. Three key parameters are involved in the processing, i.e., window size, image element spacing, and gray level. After several experiments, the sliding window size of $3 \times 3$, the window shift distance of 1, and the quantized gray level of 64 were finally selected. After principal component transformation, band 1 contained the most information. Therefore, the transformed band 1 was used to calculate the texture features of the eight statistical factors.

(3) The finer land cover features of the study area are obtained by the coverage of ground objects. The results of coverage of ground objects are obtained by image mixing decomposition at a sub-pixel scale using the random forest model on pre-processed Landsat 8 OLI imagery [51]. The spatial resolution of this image is the same as Landsat 8 OLI. The accuracy of the classification results was verified using Google Earth high-resolution image.
data. The root mean square error (RMSE) was used to evaluate the classification accuracy of each ground object. The RMSE for all ground objects in 2020 was within 0.15. This method has high classification accuracy and can be used in the study of functional area recognition. In this study, the ground objects were finally reclassified into 9 classes: bare soil, woodland, grassland, water, shade, concrete, asphalt, metal, and other impervious surfaces. The coverage of the 9 ground object classes within each block was used to represent the finer land cover characteristics of the block.

In total, one spectral feature, eight texture features, and nine land cover features of the images were obtained. Based on the spatial analysis tools of ArcGIS, the spectral features, texture features, and finer land cover features of each block were, respectively, obtained.

2.3.6. Sample Selection

Samples were selected for the five types of single-function areas. First, the function of the blocks on the high-resolution images provided by Google Earth was confirmed. Fifty blocks were selected as samples for each of the five types of single-function areas, and were randomly divided into a training set and test set with a ratio of 4:1. Then, based on Baidu Street View, the sample categories were checked for the selected blocks. An example of a finalized single-function area is shown in Figure 6.

![Figure 6](image_url)

Figure 6. The diagrams of single-function area samples: (a) residential area (R) sample; (b) commercial area (C) sample; (c) industrial area (I) sample; (d) public service area (P) sample; (e) green space area (G) sample.

2.3.7. Sub-Block-Level UFA Function Recognition

Sub-block functional recognition is based on the random forest classifier. The random forest classifier has high prediction accuracy, high tolerance for outliers and missing values, less manual intervention, and very effective performance in classification tasks [52,53]. This study adopted the Random Forest Classifier from the sklearn for classification. The parameters of Random Forest Classifier from scikit-learn are set as: n_estimators = 50, max_depth = 10, and random_state = 42. By fusing 14 building features, 5 socioeconomic features, and 18 remote sensing image features, function classification at the sub-block level was achieved using a random forest classifier with blocks as the basic units. The
classification details are exhibited in Figure 7. The model analyzes the architectural features, socioeconomic features, and remote sensing image features of the five types of single-function areas, and outputs the probability that each block is predicted as the five types of single-function areas.

![Flowchart of the sub-block-level UFA function recognition procedure.](attachment:image.png)

**Figure 7.** The flowchart of the sub-block-level UFA function recognition procedure.

The probability ratio of 50% is used as the criterion for judging the functional mixture of a block. When the probability of a block assigned to a certain function reaches or exceeds 50%, the block is determined to be a single-function area; when the probability of all functions does not reach 50%, the block is determined to be a mixed-function area. The mixed type also depends on the probability that the block is assigned to one of the five functional categories. When the probability of a block being assigned to a certain function reaches 30%, it means that the block has this type of dominant function. When the probability of a block being assigned to all five types of functional areas is less than 30%, it means that the block is a mixed-function area without dominant functions.

The accuracy of the recognition results is evaluated in terms of both single- and mixed-function areas, respectively. For single-function areas, the accuracy and reliability of the classification results are determined by using the confusion matrix. For mixed-function areas, the accuracy of the recognition results is verified by random sampling and visual comparison with the assistance of Google Earth’s high-resolution images and Baidu Street View.
2.3.8. Comparing the Recognition Effects of Different Features

The socioeconomic features extracted from POI data, the morphological features and 3D features extracted from building footprints, and the spectral, textural, and finer land cover features extracted from remote sensing images all contribute to different degrees to the recognition results of UFAs. To emphasize the advantages of data features in this study, several sets of comparison experiments were conducted based on different features, including experiments on single-source features such as POI features (POI, P), building features (Building, B), and image features (Image, I), and experiments on combined features such as POI features and building features (POI + Building, P + B), POI features and image features (POI + Image, P + I), and building features and image features (Building + Image, B + I). Moreover, to highlight the contribution of 3D building features and finer land cover features to the optimization of functional area recognition, an experiment was conducted in which 3D building features were not introduced to Building2 (B2) for building features, an experiment in which finer land cover features were not introduced to Image2 (I2) for image features, and an experiment was conducted in which 3D building features and fine land cover features were not introduced to POI+Building2 + Image2 (P + B2 + I2) for combined features.

A total of nine parallel experiments with different features were set up, and the recognition effects achieved using the combined POI features, building features, and image features (POI + Building + Image, P + B + I) used in this study were compared. The same recognition procedure as that used in this study was used to complete sub-block-level UFA function recognition.

3. Results

3.1. Recognition Results of UFAs

According to the research procedure described in Section 2.3.7, the results of urban functional zoning in the study area were obtained (Figure 8). There were 875 blocks recognized as single-function areas, accounting for 50.98% of the study area. Moreover, 1050 blocks were recognized as mixed-function areas, accounting for 49.02% of the study area.

It is evident from Figure 8 that the residential areas are mainly distributed in the Xiguan region of Liwan District and the northeast area of Haizhu District. The commercial areas are mainly located in the core of the study area such as Pearl River New Town and Canton Fair. The industrial areas are mainly concentrated in the southern area of Liwan District and the eastern area of the central part of Tianhe District, and the southern and eastern areas of Haizhu District are also slightly distributed. Public service areas are relatively evenly distributed throughout the study area, and are mainly concentrated in Tianhe District due to the number of universities located in the region. Green space areas are mainly distributed in Tianlulu Forestry Park, Guangzhou Haizhu National Wetland Park, Xiaoguwei Island Island, and other peripheral areas of the research area. The recognition results of single-function areas revealed an overall accuracy (OA) of 82% and a kappa coefficient of 0.78.

The blocks assigned as mixed-function areas in the results were selected to obtain Figure 9. It was found that the mixed-function areas including residential, public service, and commercial functions were evenly distributed in the study area, which was consistent with the land-use information presented on the electronic map. By combining Google Earth and Baidu Street View, the recognition results of mixed-function areas in this study were further verified by random sampling, and the verification procedure is presented in Table 4.
Figure 8. The recognition results of UFAs. (R: residential area; C: commercial area; I: industrial area; P: public service area; G: green space area; RF: mixed-function area dominated by residential use; CF: mixed-function area dominated by commercial use; IF: mixed-function area dominated by industrial use; PF: mixed-function area dominated by public service use; GF: mixed-function area dominated by green space use; GR: mixed-function area of green space and residential use; PG: mixed-function area of public service and green space use; PR: mixed-function area of public service and residential use; PC: mixed-function area of commercial and residential use; IC: mixed-function area of industrial and commercial use; IP: mixed-function area of commercial and residential use; IR: mixed-function area of industrial and residential use; IG: mixed-function area of industrial and green space use; MIX: mixed area without a dominant function.).

Figure 9. The recognition results of mixed-function areas.
According to two accuracy verifications, it was found that the recognition performance of residential and commercial areas was excellent; in particular, the relatively concentrated residential areas and main commercial centers in the study area were sufficiently identified. Many of the industrial areas could also be recognized, but there were industrial areas under construction with bare soil surfaces that may have been misclassified as green space areas. The recognition performance of the green space function in single-function areas was excellent, but the attributes of small parks or squares in mixed-function areas were not recognized. The results for public service areas were similar. The recognition accuracy of complete blocks formed by large schools and gymnasiuims in the study area was high, but that of small public service facilities such as hospitals and kindergartens was poor, as these facilities are mostly located in mixed-function areas.
3.2. Comparison of the Recognition Effects with Different Feature Factors

3.2.1. Comparison of the Classification Accuracy of Single-Function Areas

Parallel experiments were set up according to Section 2.3.8, and the recognition results of different features were obtained. The classification accuracy of single-function areas was verified by combining the test samples in this study, and the results are reported in Table 5. In parallel experiments using single-source features, it was found that Image (I) had the highest classification accuracy, and Building2 (B2) without 3D features had the lowest classification accuracy. In the parallel experiment of combining features, the fusion of the three types of features had the highest classification accuracy, the combination of POI + Building (P + B) had the lowest classification accuracy, and the difference in the classification accuracy between the two reached 30%.

Table 5. The classification accuracy of single-function areas with different feature factors.

| Input Feature Factors | Residential Area | Commercial Area | Industrial Area | Public Service Area | Green Space Area | Overall Accuracy |
|-----------------------|------------------|-----------------|-----------------|--------------------|-----------------|-----------------|
| P                     | 70%              | 30%             | 20%             | 50%                | 60%             | 46%             |
| B                     | 70%              | 40%             | 50%             | 0%                 | 90%             | 50%             |
| B2                    | 40%              | 30%             | 20%             | 10%                | 90%             | 38%             |
| I                     | 50%              | 60%             | 90%             | 50%                | 90%             | 64%             |
| I2                    | 50%              | 60%             | 70%             | 20%                | 90%             | 50%             |
| P + B                 | 60%              | 40%             | 70%             | 20%                | 70%             | 52%             |
| P + I                 | 80%              | 70%             | 80%             | 50%                | 90%             | 74%             |
| B + I                 | 60%              | 70%             | 80%             | 60%                | 100%            | 74%             |
| P + B2 + I2           | 60%              | 70%             | 90%             | 50%                | 100%            | 74%             |
| P + B + I             | 100%             | 70%             | 90%             | 60%                | 100%            | 82%             |

The introduction of 3D building features and finer land cover features was found to be helpful for improving the accuracy of the recognition results. In the experiment with building features, the OA of Building (B) was 50%, that of Building2 (B2) was 38%, and the OA increased by 12% after the introduction of 3D building features. In the experiment with image features, the OA of Image (I) was 64%, that of Image2 (I2) was 50%, and the OA increased by 14% after introducing the finer land cover features. In the experiment with three types of data features, the OA of POI + Building + Image (P + B + I) was 82%, that of POI + Building2 + Image2 (P + B2 + I2) was 74%, and the OA increased by 8% after introducing the 3D building features and fine land cover features.

3.2.2. Comparison of the Classification Accuracy of Mixed-Function Areas

For the classification accuracy of mixed-function areas, 30 blocks assigned as having mixed functions were randomly selected to verify their recognition results. The classification accuracy of results with different features according to the number of statistics is shown in Figure 10.

The classification accuracies of mixed-function areas were similar to those of single-function areas. The classification accuracy of Image (I) was still the highest among the single-source feature experiments, and the difference from the lowest classification accuracy was 20%. In the parallel experiment in which features were combined, the classification accuracy of the proposed method was still the highest, and reached 76.67%. Figures 11 and 12 display two specific locations for comparison.

There is a large area of factory and green space in the block (Figure 11), which should have been recognized as a mixed-function area of industrial and green space (IG). However, in the experiment using single-source features, POI (P) recognized it as a mixed-function area dominated by industry. This is because there is a lack of POI labeling of green space, and the use of only POI features could not accurately extract the green space characteristics of the block. Building (B) and Building2 (B2), respectively, classified this block as an industrial-dominated mixed-function area and a mixed-function area without a
dominant function. Building features can better classify a pure green space functional area. However, when the function of the block is mixed with other functions and green space, the recognition capability of building features for the green space function is very limited. Image (I) and Image2 (I2), respectively, recognized this block as a mixed-function area dominated by green space and a mixed-function area of public service and commercial use, and failed to correctly recognize the industrial function of the block. In the experiments using the combined features, POI + Image (P + I), POI + Building (P + B) and POI + Building2 + Image2 (P + B2 + I2) recognized the block as a mixed-function area dominated by green space and a mixed-function area dominated by industrial use, and Building + Image (B + I) recognized it as a mixed-function area dominated by green space. None of the four experiments resulted in the accurate extraction of the industrial and green space characteristics of the block, but the recognition results of the combination of POI + Building + Image (P + B + I) did achieve this. The block was designated as a mixed-function area of industrial and green space use, and the result was the most consistent with reality.

Figure 10. The classification accuracy of mixed-function areas with different features.

Figure 11. The comparison of the recognition results of different features for specific location 1. (a): the high-resolution imagery of Google Earth; (b): the result of P experiment; (c): the result of B experiment; (d): the result of B2 experiment; (e): the result of I experiment; (f): the result of I2 experiment; (g): the result of P + B experiment; (h): the result of P+I experiment; (i): the result of B + I experiment; (j): the result of P + B2 + I2 experiment; (k): the result of P + B + I experiment.
The recognition of green space and related sub-classes is of great research significance for quantitative analysis of spatial differences among different functional areas within cities.

There is a large area of residential buildings and campus in the block (Figure 12), which should be recognized as a mixed-function area of public service and residence (PR). However, in the experiments using single-source features, POI (P), Image (I), Image (I2), and Building (B2) recognized the block as a mixed-function area without a dominant function and did not extract the characteristics of public service or residence, and the recognition effects were poor. Building (B) mistakenly recognized the block as a mixed-function area dominated by commercial use, which may have been due to the large differences between the floors and morphology of the residential buildings and campus buildings in this block. In the experiment using the combined data features, Building + Image (B + I) and POI + Building2 + Image2 (P + B2 + I2) recognized the block as a mixed-function area dominated by residential use, and failed to correctly extract the public service characteristics of the block. Both POI + Building (P + B) and POI + Image (P + I) recognized the block as a mixed-function area dominated by public service use, and failed to correctly extract the residence characteristics of the block. However, POI + Building + Image (P + B + I) recognized the block as a mixed-function area of public service and residential use, which was in line with the survey results.

On the whole, the classification accuracy of single-function areas was found to be higher than that of mixed-function areas, the accuracy of combined feature experiments was better than that of single-source feature experiments, and the recognition effect of the proposed method was better than that achieved in other parallel experiments. The comparison with the classification accuracy of parallel experiments also effectively proves that the introduction of 3D building features and finer land cover features was helpful for improving the recognition effect of UFAs.

4. Discussion

In this study, open access urban multi-source data were used to classify area functions at the urban sub-block level, and a solution with easily accessible data, easy-to-implement methods, and more comprehensive classification results was proposed. Previously, some scholars integrated multi-source data and used machine learning models to classify UFAs, but did not analyze regions in urban areas with mixed functions [16,54,55]; they only classified UFAs into four or six single-function categories based on different data feature information. The present research differs from these previous studies in that the classifications of functional areas were refined, and the results achieved reliable accuracy. Fine urban functional zoning not only promotes understanding of local conditions, but also allows quantitative analysis of spatial differences among different functional areas within cities. The recognition of green space and related sub-classes has positive significance for studying ecosystem services and mitigating the urban heat island effect [56].
for urban facilities construction improvement and hot spot detection [57,58]. In addition, the distribution of fine functional areas has certain reference value for the analysis of the causes of traffic congestion [59], the evaluation of regional functional use intensity [23], and the planning calibration [60] and renewal of cities [61]. Some scholars have also finely classified UFAs, but the data sources were different [5,23,62]. Other studies have made a finer classification of UFAs based on POI data, but this method only determines the mixed functions of the area via the frequency density and category ratio of POIs, which is not sufficient for the comprehensive consideration of the urban environment [35,63,64]. Some studies integrated multi-source data for the finer classification of UFAs and achieved reliable accuracy. For example, Liu et al. [44] fused landscape features of images, human activities based on taxi trajectory data, and functional information of POIs. Shannon’s information entropy [65] was used to measure the mixed degree of functional categories for an area. However, the taxi trajectory data used in the study were not open access, thereby limiting its wide suitability.

This study achieved the finer function classification of UFAs by fusing multi-features of open access data. The main challenges and potential improvements in this study are as follows:

1. The introduced finer land cover features are reliable. Land cover features are closely related to the functional category of the block, and the finer land cover features introduced in this study were obtained based on medium-resolution data using a random forest model. The prerequisite for accurate classification is the need to ensure that the selected samples are pure. Reliable land cover classification results were obtained in this study using images with a spatial resolution of 30 m, but there was still a certain degree of error, and this error may affect the utility of finer land cover features in the recognition of UFAs. To minimize the accumulation of classification errors, the use of images with a higher spatial resolution (higher than 30 m) [51] or the construction of a bidirectional iterative self-optimizing classification model [33] will be considered in the future to optimize the accuracy of finer land cover.

2. Appropriate training samples were selected for the recognition of functional areas. To ensure the effectiveness of model training, the selected training samples must be typical. However, in fact, the number of blocks in urban areas with an almost pure function that can be used as training samples is very rare. In this work, samples were selected by the manual visual method and investigation, and it was difficult to obtain a large sample size. Additionally, because of the difficulty of directly visually determining the degree of the functional mix of blocks, there were no samples selected for the training and classification of the mixed-function areas. Instead, according to the random forest classification, each block was assigned as different functional areas based on probability ratios to determine the category of the mixed-function area. Finally, high-resolution maps and street views were used to verify the recognition results. Moreover, as a sufficient number of blocks that function as pure transportation could not be selected, to ensure the classification accuracy of the model, the categories of functional areas defined in this study did not include transportation areas. This may have led to blocks such as railway stations being recognized as having industrial or commercial functions, or as a mixture of both because of their building morphology, material, and land cover composition; however, according to the functional definition of this study, they should have been recognized as a public service area. To obtain a larger sample size and allow the model to achieve a better recognition effect, a follow-up study will consider labeling the samples and verifying the results by integrating other data sources, such as land-use data. This can not only ensure the sample size, but can also improve the scientificity of the sample selection and result verification. However, these data of government departments are not publicly available at present, and the use of land-use data obtained via image supervision classification may lead to more errors. In addition, due to the differences in the surface landscapes of different cities, it is necessary to re-select samples for land cover and functional areas during migration application.
(3) The finer and more accurate mapping of UFAs can provide a scientific basis for urban macro-decisions, public service, and “smart city” construction. At present, the recognition results of this framework have achieved reliable accuracy, but there remains room for further optimization. On the one hand, the data used in this study were all open access data, and the quality and information contained in these data are often lower than those of commercial data. In the future, commercial data with higher spatial and temporal resolution or richer socioeconomic characteristics will be used to obtain more accurate recognition results. On the other hand, the organic integration of model parameters was performed. The 37 feature factors introduced in this study were not screened, which tends to lead to the redundancy of parameters and thus affects the accuracy of the recognition results. In a more detailed study in the future, the contribution of each factor to the classification of different functional areas will be estimated. Based on this, a hierarchical recognition framework will be constructed, and the imported parameters will be screened in a hierarchical manner to remove redundancy and achieve the organic integration of multiple features.

5. Conclusions

In this study, the multi-feature information extracted from building footprints, POI data, and Landsat images was combined, and sub-block-level urban function recognition was realized by using a random forest classifier. The conclusions are shown as follows:

(1) The recognition results of UFAs in the core urban area of Guangzhou, China showed that the area of single-function areas was slightly larger than that of mixed-function areas. Moreover, the OA of single-function areas was 82%, and the recognition accuracy of mixed-function areas was 76.67%.

(2) The comparison of recognition performance of parallel experiments showed that the introduction of 3D building features and finer land cover features was helpful for improving the recognition accuracy of UFAs. The OA of single-function areas increased by 8%, and the recognition accuracy of mixed-function areas increased by 13.34%.

The strength of this study is the method of the finer classification of UFAs using an easy-to-realize random forest classifier via open access multi-source data, which has wide applicability. Furthermore, the obtained research results provide detailed functional information about urban areas, which is of great value in urban research and can be used to assist urban planning [66], population density estimation [67], and land surface temperature surveys [68].

Author Contributions: Conceptualization, B.L. and Y.D.; methodology, X.L. and Y.D.; software, X.L. and W.J.; validation, X.L. and B.L.; formal analysis, B.L. and T.L.; investigation, W.J., Z.C. and T.L.; resources, J.Y., Z.C. and W.J.; writing—original draft preparation, B.L. and X.L.; writing—review and editing, M.L., Y.D. and B.L.; data curation, Z.C. and T.L.; funding acquisition, J.Y. and M.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was fund by the Key Special Project for National Natural Science Foundation of China, grant number 41901372, 41976189 and 41901072, Science and Technology Program of Guangzhou [grant number: 202002030247], the GDAS’ Project of Science and Technology Development [grant number: 2022GDASZH-2022010202], and the Science and Technology Program of Guangdong [grant number: 2021B1212100006].

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding website.

Acknowledgments: We would like to thank the anonymous reviewers for their constructive comments; we also thank the Geographical Science Data Center of The Greater Bay Area for providing the relevant data for this study and the support of Harbin Normal University Graduate Academic Forum.
Conflicts of Interest: The authors declare that there are no real or perceived conflicts of interest involved in the submission and/or publication of this manuscript.

References

1. Zhou, G.; Li, C.; Li, M.; Zhang, J.; Liu, Y. Agglomeration and diffusion of urban functions: An approach based on urban land use conversion. Habitat Int. 2016, 56, 20–30. [CrossRef]

2. Wei, Y.D.; Ye, X. Urbanization, urban land expansion and environmental change in China. Stoch. Environ. Res. Risk Assess. 2014, 28, 757–765. [CrossRef]

3. Deng, Y.; Liu, J.; Liu, Y.; Luo, A. Detecting Urban Polycentric Structure from POI Data. Int. J. Geo Inf. 2019, 8, 283. [CrossRef]

4. Liu, B.; Deng, Y.; Li, M.; Yang, J.; Liu, T. Classification schemes and identification methods for urban functional zone: A Review of Recent Papers. Appl. Sci. 2021, 11, 9968. [CrossRef]

5. Chi, J.; Jiao, L.; Dong, T.; Gu, Y.; Ma, Y. Quantitative Identification and Visualization of Urban Functional Area Based on POI Data. J. Geomat. 2016, 41, 68–73. [CrossRef]

6. Xue, B.; Li, J.; Xiao, X.; Xie, X.; Lu, C.; Ren, W.; Jiang, L. Overview of man-land relationship research based on POI data: Theory, Method, and Application. Geogr. Geo Inf. Sci. 2019, 35, 51–60. [CrossRef]

7. Huang, N.; Zhang, L.; Liu, Q.; Wang, P. The Research on Approach and Method of Sustainable Development-Based Urban Functional Zoning and Application. Urban Dev. Stud. 2009, 16, 63–70. [CrossRef]

8. Gu, J.; Chen, X.; Yang, H. Spatial clustering algorithm on urban function oriented zone. Sci. Surv. Mapp. 2011, 36, 64–67. [CrossRef]

9. Wu, W.; Xu, Z.; Yang, X. Quantitative research of spatial development differentiation in Xi’an from the perspective of urban functional zoning. Geogr. Res. 2012, 31, 2173–2184. [CrossRef]

10. Li, Y.; Liu, C.; Li, Y. Identification of Urban Functional Areas and Their Mixing Degree Using Point of Interest Analyses. Land 2022, 11, 996. [CrossRef]

11. Huang, N.; Zhang, L.; Liu, Q.; Wang, P. The Research on Approach and Method of Sustainable Development-Based Urban Functional Zoning and Application. Urban Dev. Stud. 2009, 16, 63–70. [CrossRef]

12. Liu, Y.; Liu, X.; Gao, S.; Gong, L.; Kang, C.; Chi, G.; Shi, L. Social sensing: A new approach to understanding our socioeconomic environments. Ann. Assoc. Am. Geogr. 2015, 105, 512–530. [CrossRef]

13. Luo, J.; Zhou, C.; Yang, Y. Land-cover and land-use classification based on remote sensing intelligent Geo-interpreting model. J. Nat. Resour. 2001, 16, 179–183. [CrossRef]

14. Zhang, R.; Zhu, D. Study of land cover classification based on knowledge rules using high-resolution remote sensing images. Expert Syst. Appl. 2011, 38, 3647–3652. [CrossRef]

15. Hu, S.; Wang, L. Automated urban land-use classification with remote sensing. Int. J. Remote Sens. 2013, 34, 790–803. [CrossRef]

16. Huang, C.; Deng, L.; Fang, C. Satellite Remote Sensing and Dynamic Variation of Urban Land Use Changes—Taking Chengdu City as an Example. J. Geo Inf. Sci. 2007, 9, 118–123. [CrossRef]

17. Wang, Z.; Zhang, S.; Liu, J.; Shao, Q. Methodology for classifying and detecting intra-urban land use change: A case study of Changchun city during the last 100 years. J. Remote Sens. 2010, 14, 345–355. [CrossRef]

18. Luo, J.; Zhou, C.; Yang, Y. Land-cover and land-use classification based on remote sensing intelligent Geo-interpreting model. J. Nat. Resour. 2001, 16, 179–183. [CrossRef]

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

20. Lu, C.; Pang, M.; Zhang, Y.; Li, H.; Lu, C.; Tang, X.; Cheng, W. 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. 2020, 9, 92. [CrossRef]

21. Mao, F.; Ji, M.; Liu, T. Mining spatiotemporal patterns of urban dwellers from taxi trajectory data. Front. Earth Sci. 2016, 10, 205–221. [CrossRef]

22. Zhi, Y.; Li, H.; Wang, D.; Deng, M.; Wang, S.; Gao, J.; Duan, Z.; Liu, Y. Latent spatio-temporal activity structures: A new approach to inferring intra-urban functional regions via social media check-in data. Geo Spat. Inf. Sci. 2016, 19, 94–105. [CrossRef]

23. Liu, Y.; Liu, X.; Gao, S.; Geng, L.; Kang, C.; Zhi, Y.; Chi, G.; Shi, L. Social sensing: A new approach to understanding our socioeconomic environments. Ann. Assoc. Am. Geogr. 2015, 105, 512–530. [CrossRef]

24. Yuan, N.J.; Zheng, Y.; Xie, X.; Wang, Y.; Zheng, K.; Xiong, H. Discovering urban functional zones using latent activity trajectories. IEEE Trans. Knowl. Data Eng. 2014, 27, 712–725. [CrossRef]

25. Peng, Y.; Chen, R.; Yang, J.; Li, Y.; Jiang, H.; Liao, W.; Sun, M. Identify urban building functions with multisource data: A case study in Guangzhou, China. Int. J. Geogr. Inf. Sci. 2022, 36, 1–26. [CrossRef]

26. Yang, Z.; Gao, J.; Yang, H.; Zhao, Y. Exploring urban functional areas based on multi-source data: A case study of Beijing. Geogr. Res. 2021, 40, 477–494. [CrossRef]

27. Du, S.; Du, S.; Liu, B.; Zhang, X.; Zheng, Z. Large-scale urban functional zone mapping by integrating remote sensing images and open social data. GISci. Remote Sens. 2020, 57, 411–430. [CrossRef]

28. Yu, B.; Wang, Z.; Mu, H.; Sun, L.; Hu, F. Identification of Urban Functional Regions Based on Floating Car Track Data and POI Data. Sustainability 2019, 11, 6541. [CrossRef]

29. Xu, J.; Yang, F. A study of urban functional area identification methods based on big data of social sensing. Urban. Archit. 2017, 27, 30–34. [CrossRef]

30. Zhang, Y.; Li, Q.; Tu, W.; Mai, K.; Yao, Y.; Chen, Y. Functional urban land use recognition integrating multi-source geospatial data and cross-correlations. Comput. Environ. Urban Syst. 2019, 78, 101374. [CrossRef]
29. Yao, Y.; Li, X.; Liu, X.; Liu, P.; Liang, Z.; Zhang, J.; Mai, K. Sensing spatial distribution of urban land use by integrating points-of-interest and Google Word2Vec model. *Int. J. Geogr. Inf. Sci.* 2017, 31, 825–848. [CrossRef]
30. Chen, S.; Zhang, H.; Yang, H. Urban Functional Zone Recognition Integrating Multisource Geographic Data. *Remote Sens.* 2021, 13, 4732. [CrossRef]
31. Liu, X.; He, J.; Yao, Y.; Zhang, J.; Liang, H.; Wang, H.; Hong, Y. Classifying urban land use by integrating remote sensing and social media data. *Int. J. Geogr. Inf. Sci.* 2017, 31, 1675–1696. [CrossRef]
32. Bao, H.; Ming, D.; Guo, Y.; Zhang, K.; Zhou, K.; Du, S. DFCNN-Based Semantic Recognition of Urban Functional Zones by Integrating Remote Sensing Data and POI Data. *Remote Sens.* 2020, 12, 1088. [CrossRef]
33. Zhang, X.; Du, S.; Wang, Q. Integrating bottom-up classification and top-down feedback for improving urban land-cover and functional-zone mapping. *Remote Sens. Environ.* 2018, 212, 231–248. [CrossRef]
34. Gu, Y.; Jiao, L.; Dong, T.; Wang, Y.; Xu, G. Spatial Distribution and Interaction Analysis of Urban Functional Areas Based on Multi-source Data. *Geomat. Inf. Sci. Wuhan Univ.* 2018, 43, 1113–1121. [CrossRef]
35. Wang, J.; Ye, Y.; Fang, F. A study of urban functional zoning based on kernel density estimation and fusion data. *Geogr. Geo Inf. Sci.* 2019, 35, 66–71. [CrossRef]
36. Liu, X.; Niu, N.; Liu, X.; Jin, H.; Ou, J.; Jiao, L.; Liu, Y. Characterizing mixed-use buildings based on multi-source big data. *Int. J. Geogr. Inf. Sci.* 2018, 32, 738–756. [CrossRef]
37. Xing, H.; Meng, Y. Integrating landscape metrics and socioeconomic features for urban functional region classification. *Comput. Environ. Urban Syst.* 2018, 72, 134–145. [CrossRef]
38. Feng, Y.; Huang, Z.; Wang, Y.; Wan, L.; Liu, Y.; Zhang, Y.; Shan, X. An SOE-Based Learning Framework Using Multisource Big Data for Identifying Urban Functional Zones. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2021, 14, 7336–7348. [CrossRef]
39. Wang, Q.; Dai, Z.Y. The analysis of urban spatial structure based on POI data and principal component analysis. *Territ. Nat. Resour. Study 2018*, 6, 12–16. [CrossRef]
40. Kang, Y.; Wang, Y.; Xia, Z.; Chi, J.; Jiao, L.; Wei, Z. Identification and Classification of Wuhan Urban Districts Based on POI. *J. Geomat.* 2018, 43, 81–85. [CrossRef]
41. Zhao, W.; Li, Q.; Li, B. Extracting hierarchical landmarks from urban POI data. *Natl. Remote Sens. Bull.* 2011, 15, 973–988. [CrossRef]
42. Jia, F.; Yan, J.; Wang, T. Research on scoring evaluation model and functional regions identification constructed by big data. *Sci. Surv. Mapp.* 2021, 46, 172–178. [CrossRef]
43. Fang, Q.; Wang, S.; Yu, D.; Zhu, H.; Li, H. Analysis of resident trip characteristics based on taxi trajectory. *Geospat. Inf.* 2019, 17, 128–130. [CrossRef]
44. Luo, S.; Liu, Y.; Gao, S.; Wang, P. Quantitative identification of urban functional areas based on spatial grid. *Bull. Surv. Mapp.* 2020, 51, 214–217. [CrossRef]
45. Jia, Y.; Ge, Y.; Ling, F.; Guo, X.; Wang, J.; Wang, L.; Chen, Y.; Li, X. Urban land use mapping by combining remote sensing imagery and mobile phone positioning data. *Remote Sens.* 2018, 10, 446. [CrossRef]
46. Liu, H.; Xu, Y.; Tang, J.; Deng, M.; Huang, J.; Yang, W.; Wu, F. Recognizing urban functional zones by a hierarchical fusion method considering landscape features and human activities. *Trans. GIS 2020*, 24, 1359–1381. [CrossRef]
47. Dharani, M.; Sreenivasulu, G. Land use and land cover change detection by using principal component analysis and morphological operations in remote sensing applications. *Int. J. Comput. Appl.* 2021, 43, 462–471. [CrossRef]
48. Eklundh, L.; Singh, A. A comparative analysis of standardised and unstandardised principal components analysis in remote sensing. *Int. J. Remote Sens.* 1993, 14, 1359–1370. [CrossRef]
49. Hou, Q.; Wang, F.; Li, Y. Extraction of color image texture feature based on gray-level co-occurrence matrix. *Remote Sens. Land Resour.* 2013, 25, 26–32. [CrossRef]
50. Haralick, R.M.; Shanmugam, K.; Dinstein, I.H. Textural features for image classification. *IEEE Trans. Syst. Man Cybern.* 1973, SMC-3, 610–621. [CrossRef]
51. Liao, W.; Deng, Y.; Li, M.; Sun, M.; Yang, J.; Xu, J. Extraction and Analysis of Finer Impervious Surface Classes in Urban Area. *Remote Sens.* 2021, 13, 459. [CrossRef]
52. Breiman, L. Random forests. *Mach. Learn.* 2001, 45, 5–32. [CrossRef]
53. Guo, Y.; Chi, T.; Peng, L.; Liu, J.; Yang, L. Classification of GF-1 remote sensing image based on random forests for urban land-use. *Bull. Surv. Mapp.* 2016, 5, 73. [CrossRef]
54. Sanlang, S.; Cao, S.; Du, M.; Mo, Y.; Chen, Q.; He, W. Integrating aerial LiDAR and very-high-resolution images for urban functional zone mapping. *Remote Sens.* 2021, 13, 2573. [CrossRef]
55. Zhang, X.; Du, S.; Wang, Q.; Zhou, W. Multiscale Geosense Segmentation for Extracting Urban Functional Zones from VHR Satellite Images. *Remote Sens.* 2018, 10, 281. [CrossRef]
56. Gao, S.; Zhan, Q.; Yang, C.; Liu, H. The Diversified Impacts of Urban Morphology on Land Surface Temperature among Urban Functional Zones. *Int. J. Environ. Res. Public Health* 2020, 17, 9578. [CrossRef]
57. Yu, W.; Ai, Y.; Yang, M.; Liu, J. Detecting “Hot Spots” of facility POIs based on kernel density estimation and spatial autocorrelation technique. *Geomat. Inf. Sci. Wuhan Univ.* 2016, 41, 221–227. [CrossRef]
58. Hu, Q.; Bai, G.; Wang, S.; Ai, M. Extraction and monitoring approach of dynamic urban commercial area using check-in data from Weibo. *Sustain. Cities Soc.* 2019, 45, 508–521. [CrossRef]
59. Lai, G.; Shang, Y.; He, B.; Zhao, G.; Yang, M. Revealing Taxi Interaction Network of Urban Functional Area Units in Shenzhen, China. *ISPRS Int. J. Geo Inf.* 2022, 11, 377. [CrossRef]

60. Wei, S.; Wang, L. Community detection, road importance assessment, and urban function pattern recognition: A big data approach. *J. Spat. Sci.* 2021, 22, 1–21. [CrossRef]

61. Hao, W.; Tong, D.; Liu, Q.; Feng, C. The delimitation and classified planning and management of transformation function region: The experience and exploration of urban renewal in Shenzhen. *Urban Dev. Stud.* 2015, 42–48. [CrossRef]

62. Long, Y.; Zhang, Y.; Cui, C. Identifying Commuting Pattern of Beijing Using Bus Smart Card Data. *Acta Geogr. Sin.* 2012, 67, 1339–1352. [CrossRef]

63. Hu, Y.; Han, Y. Identification of Urban Functional Areas Based on POI Data: A Case Study of the Guangzhou Economic and Technological Development Zone. *Sustainability* 2019, 11, 1385. [CrossRef]

64. Ding, Y.; Xu, H.; Wang, C. Research on urban functional area recognition integrating OSM road network and POI data. *Geogr. Geo Inf. Sci.* 2020, 36, 57–63. [CrossRef]

65. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* 1948, 27, 379–423. [CrossRef]

66. Ding, Y.; Xie, Y.; Ye, M. Urban Planning Based on Multi-source Spatiotemporal Big Data: A Case Study of Xiacheng District of Hangzhou City. *Geomat. World* 2020, 27, 25–30. [CrossRef]

67. Song, J.; Tong, X.; Wang, L.; Zhao, C.; Prishchepov, A.V. Monitoring finer-scale population density in urban functional zones: A remote sensing data fusion approach. *Landscape Urban Plan.* 2019, 190, 103580. [CrossRef]

68. Yu, Z.; Jing, Y.; Yang, G.; Sun, R. A new urban functional zone-based climate zoning system for urban temperature study. *Remote Sens.* 2021, 13, 251. [CrossRef]