*remote sensing*

MDPI

*Article*

# Multi-Source Data-Driven Extraction of Urban Residential Space: A Case Study of the Guangdong–Hong Kong–Macao Greater Bay Area Urban Agglomeration

**Xiaodie Yuan, Xiangjun Dai, Zeduo Zou, Xiong He, Yucong Sun and Chunshan Zhou ***

School of Geography and Planning, Sun Yat-sen University, Guangzhou 510275, China; yuanxd8@mail2.sysu.edu.cn (X.Y.); daixj8@mail2.sysu.edu.cn (X.D.); zouzd@mail2.sysu.edu.cn (Z.Z.); hexiong6@mail2.sysu.edu.cn (X.H.); sunyc7@mail2.sysu.edu.cn (Y.S.)
* Correspondence: zhoucs@mail.sysu.edu.cn

**Abstract:** The accurate extraction of urban residential space (URS) is of great significance for recognizing the spatial structure of urban function, understanding the complex urban operating system, and scientific allocation and management of urban resources. The traditional URS identification process is generally conducted through statistical analysis or a manual field survey. Currently, there are also superpixel segmentation and wavelet transform (WT) processes to extract urban spatial information, but these methods have shortcomings in extraction efficiency and accuracy. The superpixel wavelet fusion (SWF) method proposed in this paper is a convenient method to extract URS by integrating multi-source data such as Point of Interest (POI) data, Nighttime Light (NTL) data, LandScan (LDS) data, and High-resolution Image (HRI) data. This method fully considers the distribution law of image information in HRI and imparts the spatial information of URS into the WT so as to obtain the recognition results of URS based on multi-source data fusion under the perception of spatial structure. The steps of this study are as follows: Firstly, the SLIC algorithm is used to segment HRI in the Guangdong–Hong Kong–Macao Greater Bay Area (GBA) urban agglomeration. Then, the discrete cosine wavelet transform (DCWT) is applied to POI–NTL, POI–LDS, and POI–NTL–LDS data sets, and the SWF is carried out based on different superpixel scale perspectives. Finally, the OSTU adaptive threshold algorithm is used to extract URS. The results show that the extraction accuracy of the NLT–POI data set is 81.52%, that of the LDS–POI data set is 77.70%, and that of the NLT–LDS–POI data set is 90.40%. The method proposed in this paper not only improves the accuracy of the extraction of URS, but also has good practical value for the optimal layout of residential space and regional planning of urban agglomerations.

**Keywords:** residential functional area; superpixel wavelet fusion; image fusion; the Guangdong–Hong Kong–Macao Greater Bay Area (GBA); urban agglomeration

check for updates

**Citation:** Yuan, X.; Dai, X.; Zou, Z.; He, X.; Sun, Y.; Zhou, C. Multi-Source Data-Driven Extraction of Urban Residential Space: A Case Study of the Guangdong–Hong Kong–Macao Greater Bay Area Urban Agglomeration. *Remote Sens.* **2024**, *16*, 3631. https://doi.org/10.3390/rs16193631

Academic Editors: Zihao Zheng and Qiming Zheng

Received: 21 August 2024
Revised: 22 September 2024
Accepted: 25 September 2024
Published: 29 September 2024

**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

## 1. Introduction

Dwelling is one of the basic functions of urban space [1]. As an essential part of urban functional space (UFS), urban residential space (URS) can be regarded as the concrete expression of structural factors in a specific mode of production and is the "refraction" of social structure at the spatial level with both physical and social attributes [2–4]. From the physical perspective, URS is manifested as the house, the community, the street, and the settlement area [5,6]. From the social perspective, URS is the embodiment of established social relations and social space order [7], reflecting the specific social connotations of the area. Especially under the condition of a pure market economy, the production, use, consumption, and transition of URSs are closely related to the differentiation and reorganization of social structure [8]. That is, different social groups and classes occupy corresponding housing resources according to their own purchasing ability, thus producing different living patterns and forms [9,10]. Along with China's urbanization processes from rapid progress to

high-quality development, as well as the transformation of the economic system and social structure, the urban spatial structure has entered a transformation period of reorganization and adjustment, and the phenomenon of reconstruction and differentiation of URS has appeared [11,12]. At the same time, it also gives rise to a series of social problems, which need more attention in the management of URS in the future. In this context, the extraction of URS is conducive to obtaining its spatial status more accurately in the reconstruction of urban space and also helping lay the foundation for scientific urban spatial planning. These are critical for the promotion of healthy urban and social development.

Research on the identification and extraction of urban space mostly focuses on the extraction of urban built-up areas and the identification of urban spatial structure, while the focus on URS is mostly on the classification and identification of UFS. Currently, specific research on URS is insufficient [13,14]. In the aspect of UFS identification, traditional methods mostly rely on a large number of manual collections and a field exploration of the urban land survey area, which not only requires a huge workload, but also, the extraction results are greatly influenced by subjective investigators [15,16]. In addition, traditional research relies heavily on traditional data such as urban planning and socioeconomic statistical yearbooks. These traditional data have certain value, but their update cycle is long, and the data are greatly affected by administrative boundaries, statistical caliber, and other factors [17,18]. At the same time, local government officials may inflate statistics to boost their performance [19]. The spatial distribution of URS is very important for the government to conduct more scientific urban planning. Therefore, under the defects of traditional data and methods, how to identify and extract URS more accurately is an important issue that needs to be solved urgently. In this context, an efficient method based on satellite remote sensing imagery has emerged to identify UFS or land use types [20]. However, a remote sensing survey mainly depends on spectral features of topographies and cannot directly detect the social and economic functions of UFS [21–23]. With the great development of computer technology, the data that can reflect urban social and economic functions are more and more abundant and more convenient to obtain. Research focused on identifying UFS based on geospatial big data has gradually become one of the hot spots in urban studies.

At present, research data on the automatic extraction of UFS mainly rely on POI data and NTL data. POI data have geospatial identifiers, including name, zoning type, geographic location, and other information, and they are widely used in the study of UFS demarcation [24], urban spatial structure [25,26], land use identification [27,28], and so on. Although the classification system of POI data does not exactly correspond to the urban land use, this study found that 76.7% of POI data can be preliminarily matched with the urban land use type, so it is feasible to infer the function of urban land use with POI data [29]. In specific research, relevant scholars mostly judge the type of land use by the frequency of occurrence of various types of POIs in each research unit and identify a single functional space or mixed functional space according to whether the percentage of certain types of POIs account for more than 50% of the total [14,24]. In addition, relevant studies have obtained more accurate extraction results through machine learning [30]. UFS recognition based on POIs can achieve higher classification accuracy, but, in essence, such research only uses basic information such as the category and spatial distribution of POIs. This tends to make the results of UFS recognition more inclined to select the use type with the higher number of POIs [31], rather than necessarily the dominant socioeconomic function of the functional area. This may cause the real leading function of the research unit to be masked by the surrounding noise, causing the research results to fail to meet the specific requirements of urban and rural planning practices [32,33].

NTL data are commonly used for urban spatial identification and extraction due to their significant difference in the spatial distribution of Digital Number (DN) values [34–36]. Common NTL data are mainly from the Defense Meteorological Satellite Program/Operational Linescan System (DMSP/OLS), the National Purchasing Partners Visible Infrared Imager Radiometer Suite on Suomi National Polar Orbit Partnership satel-

lite (NPP-VIIRS), and Luojia-1 night light data. However, the resolution of these data is relatively low, and there are pixel saturation problems. The extraction threshold of urban space varies with the study area. In addition, the commonly used NTL data are inconsistent and comparable in sensor parameters and acquisition times, which reduced the extraction accuracy of URS to some extent.

At present, remote sensing image (RSI) data have become an advantageous data source for the dynamic monitoring of urban space because of its strong continuity, large coverage, and real-time collection. Among them, High-resolution Image (HRI) data with abundant texture information and multiple imaging spectral bands are widely used in the study of UFS extraction [32,37]. For example, Google Earth images with high resolution are freely available and are considered to be one of the main complementary data sources for urban spatial research [38,39]. However, due to the lack of semantic information, HRI is usually more helpful to describe the spatial layout of ground objects than to identify their functional uses [40]. At present, the use of RSI data to extract UFS is mainly based on the difference of pixels or research objects. The research methods include supervised classification, unsupervised classification, Artificial Neural Network (ANN), Normalized Difference Built-up Index (NDBI), Principal Component Analysis (PCA), information extraction model method based on spectral knowledge, and so on [36,41,42]. However, these methods also have some problems, such as low extraction accuracy, data need to be purchased, a lack of historical archive data, large data processing workload, and long model operation time [43,44].

Due to the characteristics of each type of data, it is difficult to accurately identify and extract URS with a single data set. In recent years, some new data types have become more and more widely used. These data have the advantages of fast collection speed and wide sample range, which is very suitable for urban research. Related studies have also found significant spatial correlations in these big data and have begun to fuse multiple data sets to improve the precision of urban spatial estimates. At present, POI data, NTL data, LandScan (LDS) data, and HRI data are widely used in data fusion. For example, POI data greatly addresses the semantic gap in RSI by utilizing large amounts of user-generated content [45–47]. Currently, studies have been conducted to use POI data as training data for urban spatial recognition [48] and fused it with RSI data to identify urban buildings [49] and UFS [50]. POI data also correlates well with NTL data. Related studies integrate POI data and NTL data to identify urban built-up areas, and the results show that the recognition results after data fusion are more accurate [34,51]. The above studies show that exploring the laws of urban space more precisely through data fusion will become a new hotspot in the field of urban research.

In addition, LDS data, as an early attempt to draw a spatiotemporal population map with a global resolution of 30 × 30 arcseconds [52], is the best census data in the existing population database, which can better reflect the spatial pattern of population and is widely applied in urban research [53,54]. At the same time, it is also gradually used for multi-source data fusion to identify urban space more accurately [55]. However, in the current research on the extraction process of URS, NTL data, and POI data are mainly used, and RSI data and LDS data are relatively insufficient.

In recent years, machine learning (ML) has been gradually applied to the research of UFS extraction, and the accuracy of the research results is far superior to other methods. Among them, superpixel segmentation in statistical learning is a typical image segmentation technology for processing HRI. It segments neighboring pixels with similar features, which greatly reduces the complexity of image post-processing, and is often used in the extraction research of urban buildings, roads, and other elements (Figure 1), but at the same time, there are some disadvantages such as excessive segmentation and poor boundary handling. In the fusion method of multi-source data, pixel-level fusion has a significant precision advantage in the current image fusion framework. At present, frequency domain transform is commonly used for image fusion, and the specific fusion methods mainly include Fourier transform (FT), wavelet transform (WT), and discrete cosine wavelet transform (DCWT).

Compared with FT, WT has a stronger adaptive ability, while DCWT has the advantages of energy concentration, lossless compression, high coding efficiency, color consistency, etc., which can be better applied to the fusion of urban multi-source data. But DCWT also has some disadvantages such as information loss and slow computation speed. Based on this, this study combined superpixel segmentation and DCWT, fully considered the distribution law of spatial information in HRI, embedded the spatial information of urban residential areas into the wavelet transform, creatively proposed the superpixel wavelet fusion (SWF) method, and provided a scientific and efficient way for URS extraction.

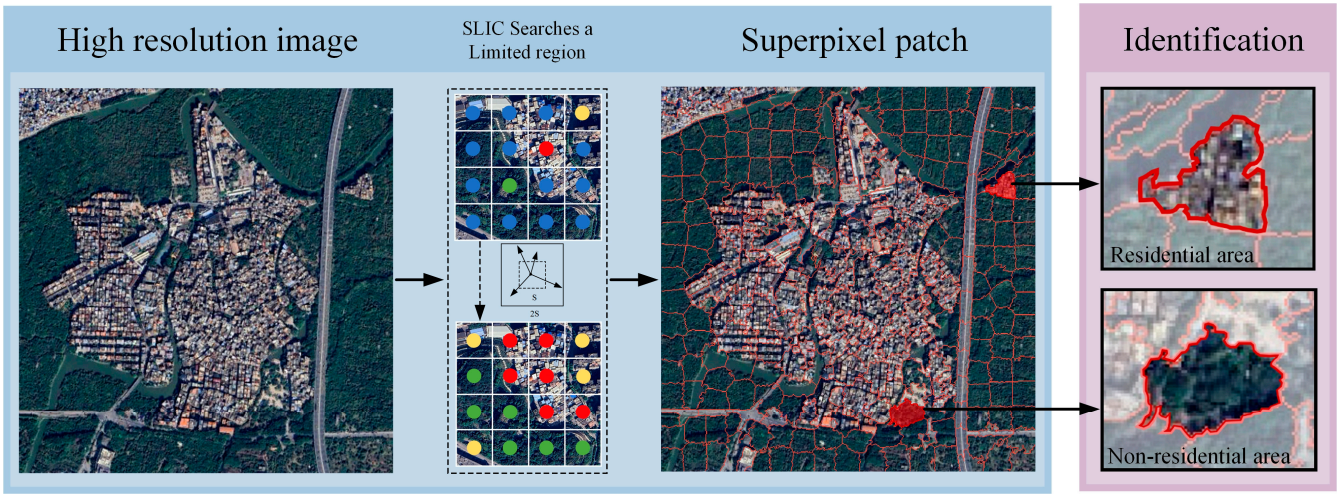


**Figure 1.** Residential space localization and non-residential space recognition based on superpixel segmentation. (Note: the blue circles represent the initial seed points, while the red, yellow, and green circles represent the sampling points of different feature types.)

Based on the theory of DCWT and superpixel segmentation, this study creatively proposed a SWF method to identify URS based on the fusion of multi-source data (POI data, NTL data, HRI data, and LDS data) and extracted URS using OSTU threshold value. The distribution of URS is studied quantitatively. In this study, a new idea and method of URS extraction is proposed to serve the subsequent urban planning and urban development more scientifically.

The rest of the paper is organized as follows. The second part is the research area and data, including the overview of the research area, the source of the research data, and the pretreatment. The third part is the research methods, including the principle and application of the main research methods, accuracy verification methods, and indicators. The fourth part is the research results, including the differences between URS extracted from different data. The fifth part is the discussion. The sixth part is the conclusion.

## 2. Research Area and Data

### *2.1. Research Area*

The Guangdong–Hong Kong–Macao Greater Bay Area (GBA) urban agglomeration is composed of two special administrative regions (Hong Kong and Macao) and nine cities in the Pearl River Delta urban agglomeration (Guangzhou, Shenzhen, Zhuhai, Zhongshan, Jiangmen, Foshan, Dongguan, Huizhou, and Zhaoqing) (Figure 2). It is located in the southeast coastal region of China, with a total area of about 56,000 km$^2$. It is recognized as one of the most internationalized and economically dynamic urban agglomerations in China [56].

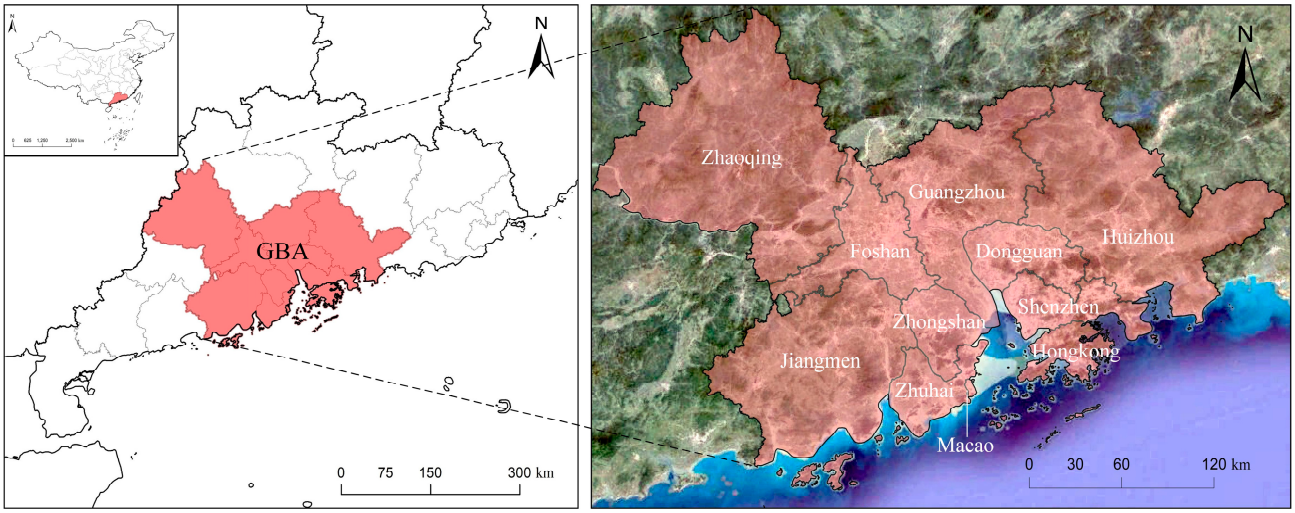


**Figure 2.** Research area of this work—the GBA urban agglomeration.

As one of the most important UFSs, the URS of the GBA urban agglomeration carries the daily life of a large population. With the continuous agglomeration of the population, the URS will face more and more pressure, and it is urgent to optimize the spatial layout and accurately identify that the URS has become one of the main challenges at present.

### *2.2. Research Data and Preprocessing*

The data used in this study are shown in Figure 3, including POI data, NTL data, LDS data, and HRI data. Details of the data are shown in Table 1.

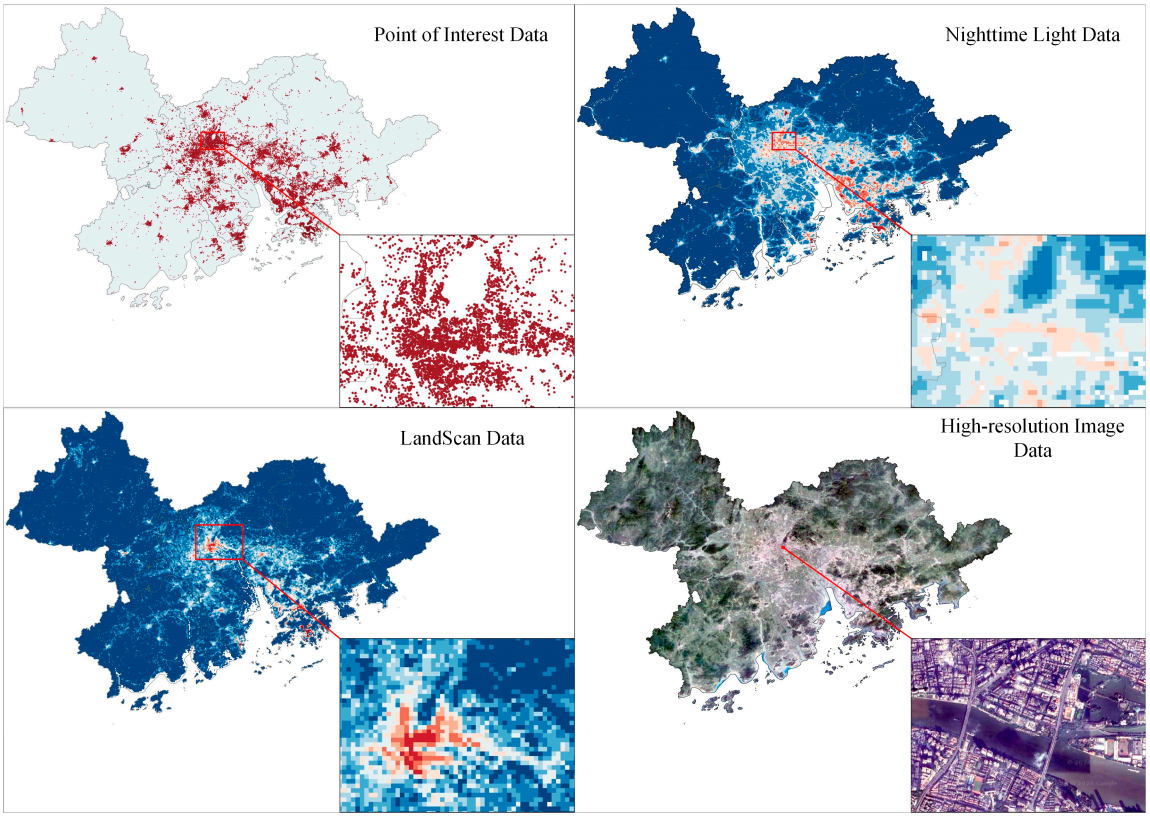


**Figure 3.** Data presentation.

**Table 1.** Basic information on various types of research data.

| Type | Source | Resolution | Release Time |
|---|---|---|---|
| POI | https://www.amap.com/ accessed on 30 January 2023 | — | 2022 |
| NTL | http://geodata.nnu.edu.cn/ accessed on 9 March 2024 | 500 m × 500 m | 2022 |
| LDS | https://landscan.ornl.gov/ accessed on 16 June 2024 | 1000 m × 1000 m | 2022 |
| HRI | https://earth.google.com/ accessed on 16 June 2024 | 500 m × 500 m | 2022 |

#### 2.2.1. POI Data

This study uses POI data collected from the AMAP location service interface [57]. These data are defined by geographic coordinates and some additional attributes, usually including latitude and longitude, name, category, address, opening hours, and contact information. POI data provide an accurate representation of land use by providing a distribution of urban functions. Amap classifies POI data into 23 categories, 267 medium categories, and 869 subcategories. In this study, the residential area (middle class) type was selected and the data were further cleaned, including screening and merging of duplicate parts, rapid removal of missing information, and other preprocessing operations. On this basis, a total of 63,650 POI data of residential areas in the GBA urban agglomeration were obtained and adopted.

#### 2.2.2. NTL Data

The NPP-VIIRS data are the second generation of night light remote sensing sensors after the DMSP/0LS data, with a Day/Night Band (DNB) that can detect night light information in the wavelength range of 0.5 to 0.9 μm. This study selected a global class NPP-VIIRS annual data product calibrated across sensors with a spatial resolution of 500 m [58]. In this study, ArcGIS 10.8 was used to clip and turn the NTL raster data according to the study area, and then the light value was connected to the grid according to the spatial position of the pixel, which was applied to reflect the spatial scope of the URS in the study area.

#### 2.2.3. LDS Data

The LDS data were derived from the 2022 LandScan Global data of the Oak Ridge National Laboratory [59]. A weighted model based on GIS and a partition density model was established. In this study, ArcGIS 10.8 was used to clip according to the study area and turn the grid to a point, interpolated by the inverse distance weight (IDW) method, and then connected the population density value to the grid according to the spatial position of the pixel to reflect the urban population density.

#### 2.2.4. HRI Data

HRI data have many advantages in urban spatial research, such as fast acquisition time, large distance, long time series, and high resolution, and have become one of the indispensable data sources in related research [60–62]. In this study, HRI data from Google Earth in 2022 were used, and ArcGIS 10.8 was used to reduce and unify the row and column number of raster data pixels according to the study area for subsequent image fusion.

## 3. Research Methods

Based on the theory of SLIC image superpixel segmentation and DCWT, this paper proposes SWF and uses OSTU adaptive threshold calculation to extract URS. The structure of the main analysis framework of this study is shown in Figure 4.

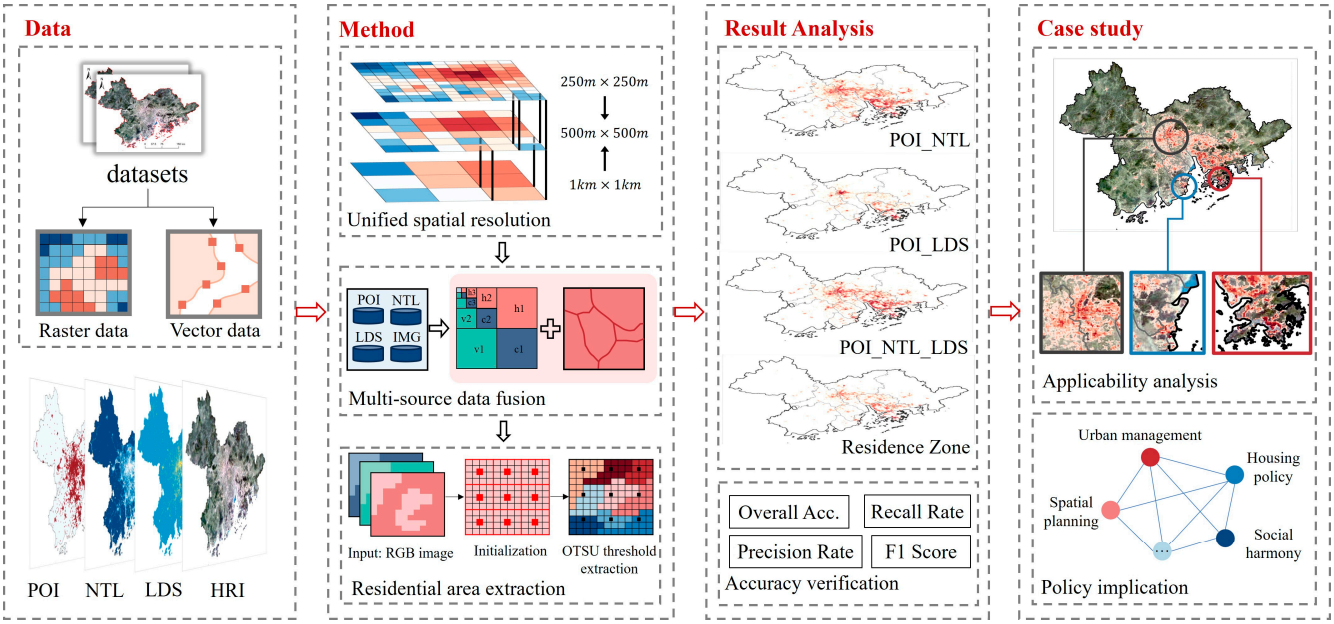


**Figure 4.** Analysis frame diagram.

### *3.1. Theoretical Basis*

#### 3.1.1. SLIC Image Superpixel Segmentation

A superpixel is considered to be a pixel region consisting of neighboring pixels with similar characteristics. After processing by the superpixel algorithm, a large number of image pixels are replaced by a small amount of information, and its processing is simplified while the processing efficiency is largely promoted. The commonly used superpixel algorithms include graph-based algorithms, gradient-ascending methods, and simple linear iterative clustering (SLIC) algorithms. Compared with other superpixel segmentation methods, the SLIC algorithm has the advantages of fast running speed, higher memory efficiency, better contours of generated superpixels, and fewer parameters to be set. In this paper, the SLIC algorithm is used for superpixel segmentation.

SLIC, also known as simple linear iterative clustering, is improved based on the K-means algorithm [63]. The calculation steps of the superpixel center point are as follows:

(1) Assuming that the image size is $M \times N$ and the number of superpixels is k, this image is evenly divided into k superpixel blocks, so that the length and width of each superpixel block are equal to $SL = \sqrt{(M \times N)/k}$, and the center point is equal $(SL/2, SL/2)$;

(2) The center point of each superpixel block may be at the noise point or the pixel mutation. To reduce this probability, gradient calculation is performed by using the differential method, and the point with the smallest gradient value is the new center point.

The difference gradient is calculated as follows:

$$d_x(g,h) = I(g+1,h) - I(g,h) \quad (1)$$

$$d_y(g,h) = I(g,h+1) - I(g,h) \quad (2)$$

$$G(x,y) = d_x(g,h) + d_y(g,h) \quad (3)$$

where $I(g,h)$ is the value of the pixel at position $(g,h)$, and $G(x,y)$ is the gradient value of the pixel at position $(g,h)$.

The steps of pixel clustering are as follows:

(1) Assign a cluster label to each pixel in a neighborhood around each center point that is $2SL \times 2SL$, that is, twice the size of the expected superpixel;

(2) Calculate the color distance and spatial distance between the pixel point and the center point in the search range, and the formula is as follows:

$$d_c = \sqrt{(L_h - L_g)^2 + (A_h - A_g)^2 + (B_h - B_g)^2} \quad (4)$$

$$d_s = \sqrt{(X_h - X_g)^2 + (Y_h - Y_g)^2} \quad (5)$$

$$D = \sqrt{\left(\frac{d_c}{m}\right)^2 + \left(\frac{d_s}{SL}\right)^2} \quad (6)$$

where $d_c$ is the color distance between the center point of the g-th superpixel block and the h-th pixel point in the corresponding search range; $d_s$ is the corresponding spatial distance; D is the final distance; and m is the maximum possible distance of LAB space, the value range is generally 10 [1,42];

(3) Each pixel corresponds to multiple superpixel block centers; of course, there are multiple distances and the center point corresponding to the minimum value is the center of the pixel;

(4) After completing 1 iteration, the center point coordinates of each superpixel block are recalculated and 10 iterations are performed again. In the process of segmentation of residential area image, it can be seen by many experiments that k is the most appropriate value of 300.

3.1.2. Discrete Cosine Wavelet Transform (DCWT)

Compared with FT, WT is considered to have a more adaptive capability to discretize images with various resolutions, which facilitates image post-processing [64,65].

The WT generates the corresponding scale and displacement function through a parent wavelet hill $\psi(t)$, which is defined as follows:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right) \quad (7)$$

where a and b represent the scale factor and displacement factor, respectively.

The WT of the signal $f(t)$ can be defined as follows:

$$W(a,b) = \int_R \psi_{a,b}(t) f(t) dt \quad (8)$$

The inverse WT of $f(t)$ can be defined as follows:

$$f(t) = \int_{a=0}^{+\infty} \int_{b=-\infty}^{+\infty} \frac{1}{C_\psi} W(a,b)\psi_{a,b}(t) da db \quad (9)$$

where $C_\psi = \int_R \frac{|\psi(\omega)|}{|\omega|} d\omega$, $\psi(\omega)$ is the FT of $\psi(t)$.

The continuous wavelet transform (CWT) requires that the variables and functions be continuous variables and continuous functions, respectively. However, as a digital signal, an image is considered to be hardly able to fulfill this condition [66]. To solve the shortcomings of CWT, DCWT was presented. A key advantage of the DCWT over the Ft is its ability to capture both frequency and position information (time position). a and b are restricted to a certain range, and they are discretized, specifically, as follows:

(1) The discretization of a: power series processing of a; that is, $a > 0$, and the corresponding wavelet function is $a_0^{\frac{-j}{2}} \psi\left[a_0{}^{-j}(t-b)\right], j = 0, 1, 2, \cdots$;

(2) The uniform dispersion of b, $\psi_{a,b}(t)$ becomes the following:

$$a_0^{\frac{-j}{2}} \psi\left[a_0{}^{-j}\left(t - ka_0^j b_0\right)\right] = a_0^{\frac{-j}{2}} \psi\left[a_0{}^{-j} t - kb_0\right] \quad (10)$$

DCWT was defined as follows:

$$W\left(a_0^j, kb_0\right) = \int \psi_{a_0^j kb_0}(t) f(t) dt, j = 0, 1, 2, \cdots, k \tag{11}$$

### *3.2. Superpixel Wavelet Fusion*

Based on SLIC superpixel segmentation and DCWT theory, a multi-source data fusion method based on superpixel wavelet fusion (SWF) is designed in this paper. Specifically, firstly, the SLIC algorithm is used to divide the HRI of GBA into superpixels, and 800,400,200 superpixel blocks are obtained successively. Then, the data from two different sources are obtained by DCWT to obtain the high- and low-frequency characteristics at three scales. Next, high- and low-frequency feature fusion of data from different sources is carried out based on different scale superpixel perspectives. Among them, the absolute maximum method is used for high-frequency features, and the coefficient weighting strategy is used for low-frequency features.

In DCWT, the Mallat algorithm is one of the most commonly used decomposition algorithms [67], and it can be expressed as follows:

$$\begin{cases} C_{j+1}(m,n) = \sum\limits_{r\in Z}\sum\limits_{c\in z} H_{r-2m} H_{c-2n} C_j \\ D^H_{j+1}(m,n) = \sum\limits_{r\in Z}\sum\limits_{c\in z} G_{r-2m} H_{c-2n} C_j \\ D^V_{j+1}(m,n) = \sum\limits_{r\in Z}\sum\limits_{c\in z} H_{r-2m} G_{c-2n} C_j \\ D^D_{j+1}(m,n) = \sum\limits_{r\in Z}\sum\limits_{c\in z} G_{r-2m} G_{c-2n} C_j \end{cases} \tag{12}$$

In the formula, $H_r$ and $H_c$ represent high-pass filters; $G_r$ and $G_c$ represent low-pass filters; r and c represent the rows and columns of the image; $C_{j+1}$ represents the low-frequency part, and can also be represented by LL; $D^H_{j+1}$, $D^V_{j+1}$, $D^D_{j+1}$ represent the edge details of the image in the x, y, xy directions, that is, the high-frequency part, which can be represented by LH, HL, and HH, respectively.

The three-layer decomposition principle of WT can be shown in Figure 5. In the first decomposition, four sub-bands can be obtained, among which three are high-frequency sub-bands and one is a low-frequency sub-band. Only low-frequency sub-bands are decomposed in the next decomposition.

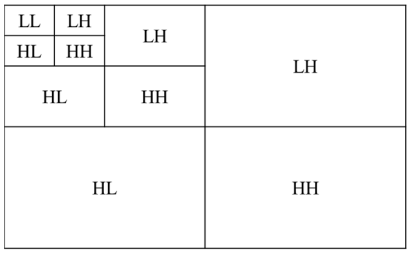


**Figure 5.** Schematic diagram of wavelet decomposition.

The two-dimensional Mallat algorithm for image reconstruction [68] is as follows:

$$\begin{aligned} &C_{j+1}(m,n) \\ &= \sum_{r\in Z}\sum_{c\in z} H^*_{m-2r} H^*_{n-2c} C_{j+1}(r,c) \\ &+ \sum_{r\in Z}\sum_{c\in z} G^*_{m-2r} H^*_{n-2c} D^H_{j+1}(r,c) \\ &+ \sum_{r\in Z}\sum_{c\in z} H^*_{m-2r} G^*_{n-2c} D^V_{j+1}(r,c) \\ &+ \sum_{r\in Z}\sum_{c\in z} G^*_{m-2r} G^*_{n-2c} D^D_{j+1}(r,c) \end{aligned} \tag{13}$$

Here, the two original images to be fused are set as A and B, and the fused images are set as F. The steps of image fusion are as follows: First, images A and B are decomposed using WT's two-dimensional Mallat algorithm to obtain the low-frequency sub-band and high-frequency sub-band. Then, the sub-band coefficients at different levels of the image are fused using various rules. Finally, the fused wavelet coefficients are reconstructed using the two-dimensional Mallat algorithm to obtain F.

The selection of fusion rules is very important in the process of image fusion by WT. Currently, pixel fusion rules [69] and window fusion rules [70] are widely used. Among them, the correlation of adjacent pixels is taken into account in the window fusion rule, and richer details can be obtained, resulting in better image vision. Therefore, the window fusion rule is selected in this paper to achieve image fusion.

$$L_N^F(p,q) = \alpha_1 L_N^A(p,q) + \alpha_2 L_N^A(p,q) \tag{14}$$

Images A and B are decomposed by N layer WT, and their lowest frequency sub-band numbers are $L_N^F(p,q)$ and $H_i^F(p,q)$, $H_N^A(p,q)$, and $H_N^B(p,q)$, respectively, representing their highest frequency sub-band coefficients. $(p,q)$ is the coordinate of the coefficient, and $L_N^F(p,q)$ and $H_i^F(p,q)$, respectively, represent the coefficient of image fusion. For the lowest frequency sub-band coefficients $L_N^A(p,q)$ and $L_N^B(p,q)$, the fusion rules are as follows:

In the formula, $\alpha_1 + \alpha_2 = 1$.

For the highest frequency sub-band coefficients $H_N^A(p,q)$ and $H_N^B(p,q)$, the following can be obtained:

$$H_i^F(p,q) = \begin{cases} H_i^A(p,q), & \left|H_i^A(p,q)\right| \geq \left|H_i^B(p,q)\right| \\ H_i^B(p,q), & \left|H_i^A(p,q)\right| < \left|H_i^B(p,q)\right| \end{cases} \tag{15}$$

### *3.3. Precision Validation*

Random points were sampled in the study area, and the accuracy, accuracy, recall rate, and F1 score of statistical classification indicators were used to quantitatively evaluate the accuracy of extraction results [71].

$$\text{accurary} = \frac{T_P + T_N}{T_P + F_P + T_N F_N} \tag{16}$$

$$\text{precision} = \frac{T_P}{T_P + F_P} \tag{17}$$

$$\text{recall} = \frac{T_P}{T_P + F_N} \tag{18}$$

$$\text{F1-score} = 2 \times \frac{\text{prcision} \times \text{recall}}{\text{prcision} + \text{recall}} \tag{19}$$

where $T_P$ is the sample that is predicted to be URS and actually is; $F_P$ is the sample predicted as URS, but actually is not; $T_N$ is the sample that is predicted to be non-URS and actually is not; $F_N$ indicates the sample that is predicted to be a non-URS but actually is.

## 4. Research Results

### *4.1. Multi-Source Data Fusion*

Since single-source data have errors in the extraction of URS, while POI data, NTL data, LDS data, and HRI data have relatively strong spatial interrelationships, these types of data can be integrated to increase the precision of URS extraction. In this study, the input RSI is transformed by wavelet, the low-frequency information is segmented by the weighted average method, the high-frequency information is segmented by the larger coefficient absolute value method, and then the fused data set is obtained by inverse wavelet transform (IWT). The specific fusion process is shown in Figure 6.

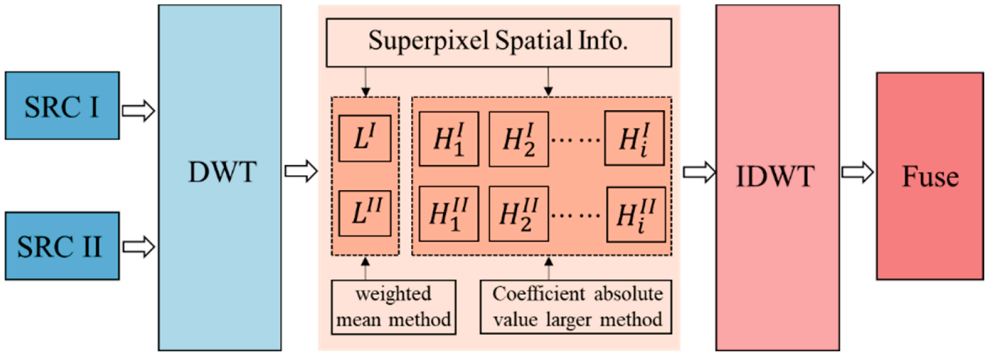


**Figure 6.** Fusion process of SWT.

### *4.2. Extraction of URS in GBA Urban Agglomeration*

#### 4.2.1. Remote Sensing Image Processing Based on Superpixel Segmentation

Superpixel segmentation before extracting image features can effectively reduce the amount of computation and image noise, so this study adopts superpixel segmentation technology to control the number and size of each superpixel. The larger the segmentation coefficient, the smaller the number of superpixels, the larger the size of the obtained segmentation map, and the larger the retained ground object information. On the contrary, the smaller the segmentation coefficient, the more the number of superpixels, the smaller the size of the obtained segmentation map, the larger the retained ground object information, and the more noise. To analyze the effect of the number of superpixel blocks on the URS recognition results, the segmentation coefficients were set as 200, 400, and 800, respectively, and these three scales were used for image fusion of different data sets, respectively (Figure 7).




**Figure 7.** Three scales of superpixel segmentation.

#### 4.2.2. URS Feature Extraction Based on Different Fusion Data

Because of the shortcomings of using POI data, NTL data, or LDS data alone to identify URS [51], image fusion of these three kinds of data is carried out. The DB1 wavelet base (Haar wavelet) was used for frequency decomposition on three-pixel scales, and image fusion was performed on three data sets, respectively, to obtain three kinds of fusion data,

POI–NTL, POI–LDS, and POI–NTL–LDS, as shown in Figure 8. The areas identified as possible residential areas are shown in red and divided into five levels according to their probability of being residential areas, namely, very low, low, medium, high, and very high.

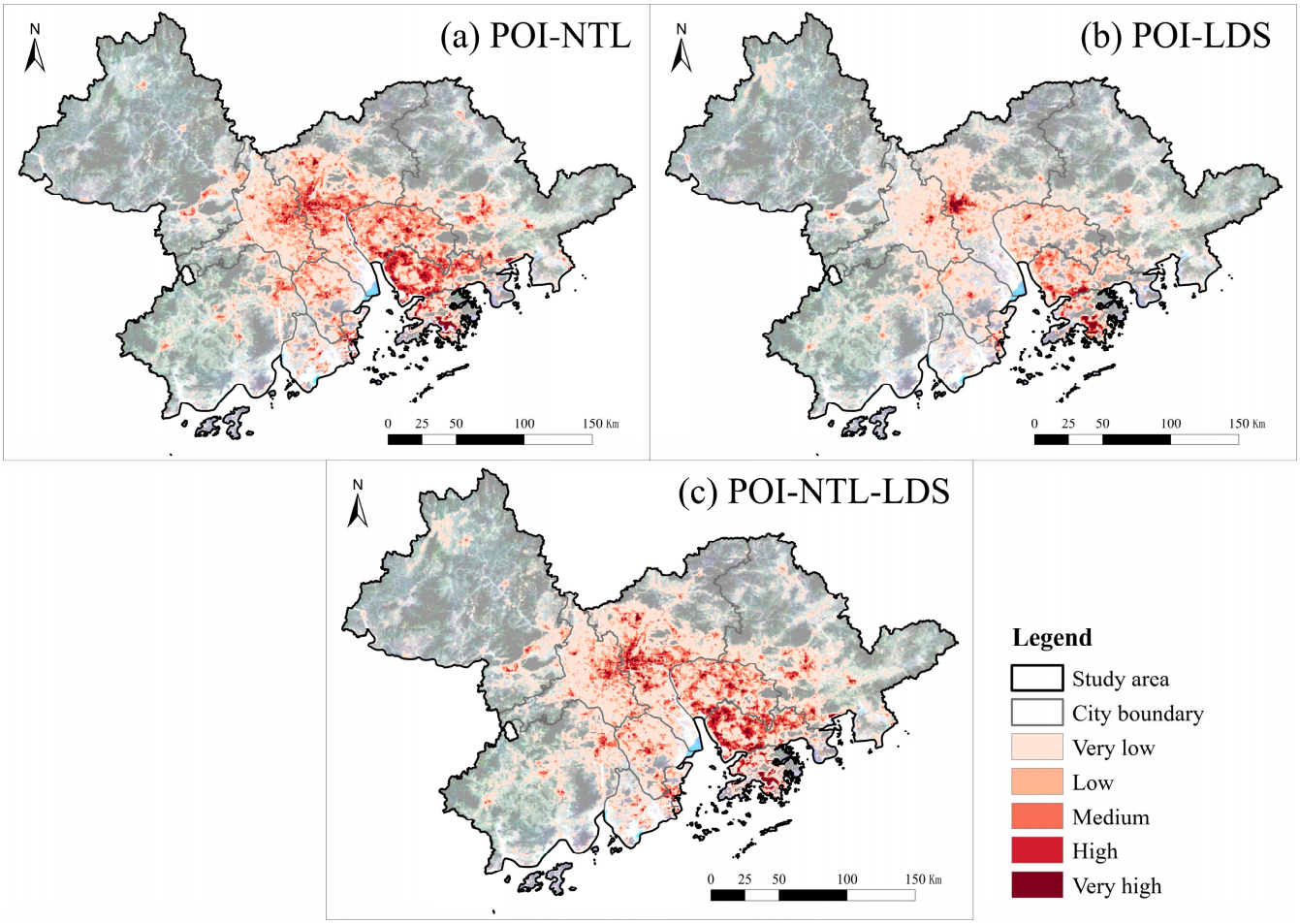


**Figure 8.** Image after the fusion of POI data, NTL data, and LDS data by SWT.

In terms of spatial distribution, the high value of URS identified by POI–NTL data is mainly located in the central city of Guangzhou and Foshan, the Zhuhai–Macao cluster area on the west bank of the Pearl River Estuary, the Dongguan–Shenzhen cluster area on the east bank of the Pearl River Estuary, and the Kowloon City District of Hong Kong. However, the urban living space obtained by POI–LDS is quite different from that obtained by the former. Compared with the more continuous URS obtained by POI–NTL data, the URS obtained by POI–LDS is mainly patched, and its distribution is mainly concentrated in Guangzhou, Foshan, Zhongshan, Zhuhai–Macao, Dongguan, Nanshan, and Futian in Shenzhen, the Kowloon City District of Hong Kong, and the north coast of Hong Kong Island. POI–NTL–LDS shows better authenticity because it combines the advantages of three kinds of data. Its high values are mainly clustered in the central urban areas of the core cities in GBA urban agglomeration and show finer image texture features.

The results of this study show that there are large differences in the results of URS estimated by POI–NTL and POI–LDS data sets. The POI–NTL data set integrates the point information of URS and the light information generated by urban human activities, but the obtained URS is too large. The POI–LDS data set integrates the point information and population spatial distribution information of URS. Compared with POI–NTL, POI–LDS cannot be affected by the lighting of non-URS, but the obtained URS are smaller. However, a significant advantage is shown in the data set obtained by merging the three types of

data. The URS obtained by POI–NTL–LDS contains information about the spatial location of residential areas and the daily activity range of urban residents, making it possible to present more details.

By comparing the recognition results of the three kinds of fusion data, it can be found that the URS identified by the three kinds of fusion data have certain similarities in spatial structure, and their probability of being recognized as URSs shows a tendency of spreading outward and decreasing from the central urban area of the core city. Furthermore, the distribution of the high values of the three data types in the city clusters also has good consistency, mainly concentrated in the main urban areas of the central cities of the GBA urban agglomeration. All these features indicate that POI–NTL, POI–LDS, and POI–NTL–LDS fully demonstrate the strengths of POI data, reflecting the structure by mirroring the point information of URS.

Because POIs only have spatial location and category attributes, it is difficult to reflect the spatial dimension of the research object, which to some extent leads to certain errors in URS extraction. Due to its spatial brightness attribute, NTL data can reveal the spatial dispersion of human activities through brightness values. Still, it will also have high brightness values in the road, airport, and other spaces, and different residential areas have large differences in light brightness values, thus reducing the accuracy of URS extraction. POI–NTL data retain the benefits of the respective data, that is, taking into account the light value of the URS and the spatial location information of the POIs to obtain a more reasonable URS. In addition, since the URS is closely related to the distribution of population, the LDS data can compensate for the incomplete collection of POI data and residential lighting thresholds of the NTL data to increase the precision of URS extraction. All in all, POI–NTL–LDS data can more fully express the spatial distribution of URS and more clearly express the details of the space structure.

### *4.3. Accuracy Verification Based on OTSU*

OTSU algorithm is applied to automatically compute the classification threshold of gray level image for POI–NTL–LDS data with better feature recognition effect. After standardized image processing, the optimal POI–NTL–LDS threshold was 44 according to the calculation results (Figure 9). Finally, according to the optimal threshold, the results of URS in GBA urban agglomeration were extracted (Figure 10).

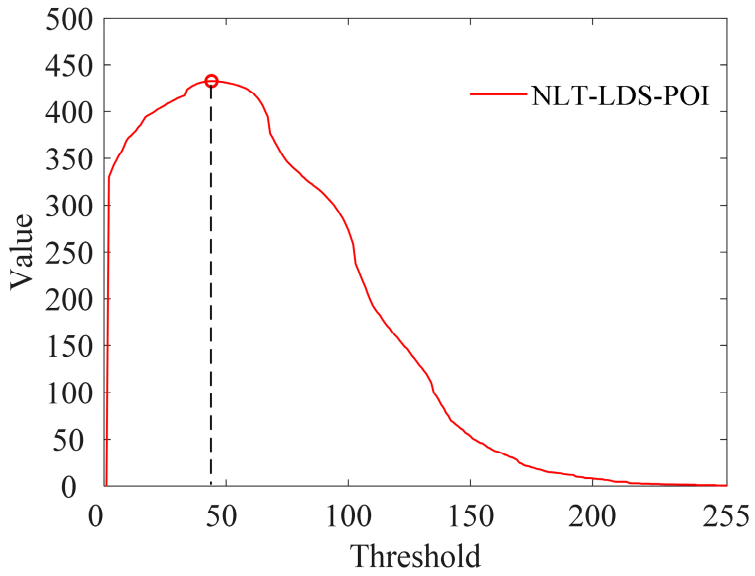


**Figure 9.** Threshold extraction of POI–NTL–LDS data set.

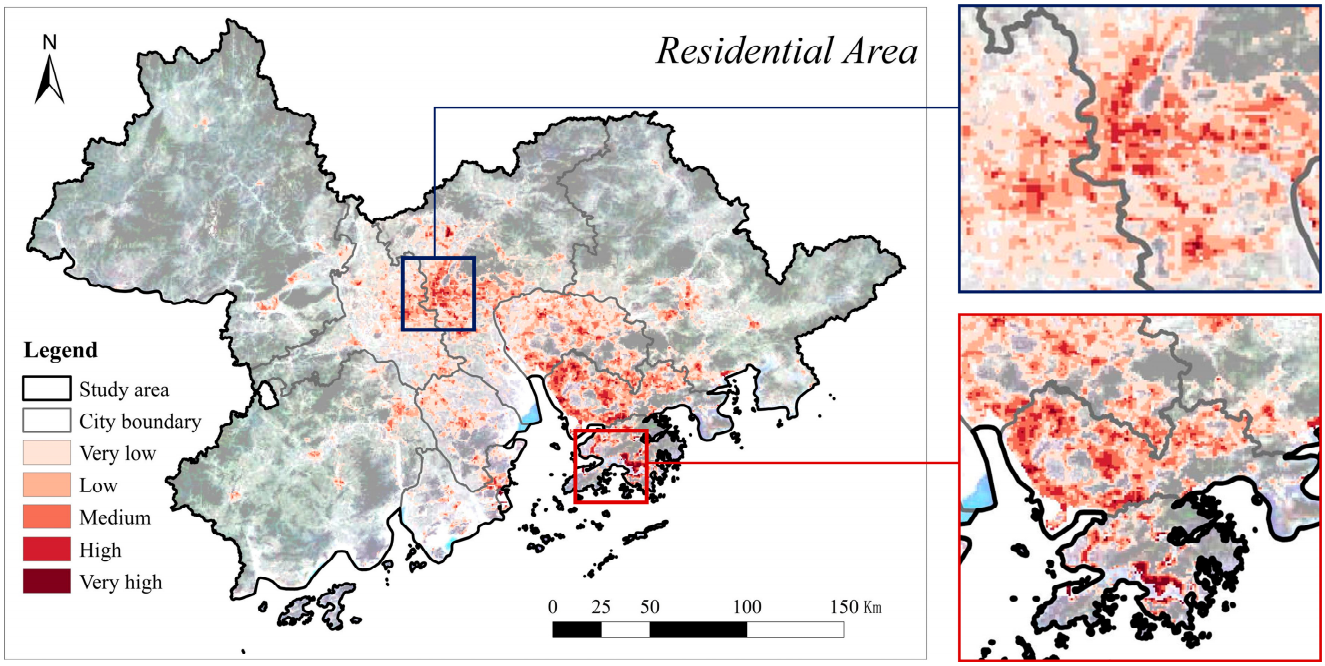


**Figure 10.** Residential area results extracted by the OSTU adaptive threshold calculation.

To have an accurate evaluation of the research results, this study also used the OSTU adaptive threshold calculation to extract the POI–NTL data set and POI–LDS data set, respectively (Figure 11). Then, referencing the Google Earth remote sensing image of the same phase, 3000 verification points are manually marked within the residential area (Figure 12), and whether the verification points are URS is manually determined. Finally, three different fusion data extraction results were validated using total accuracy, accuracy, recall rate, and F1 score in the quantitative accuracy evaluation (Table 2). The overall accuracy of the URS extracted by the NLT–POI data set is 0.8152, the precision rate is 0.9228, the recall rate is 0.6973, and the F1 score is 0.7943. The overall accuracy of the URS extracted by the LDS–POI data set was 0.7770, the accuracy rate was 0.8188, the recall rate was 0.7245, and the F1 score was 0.7688. The overall accuracy of the URS extracted by the POI–NTL–LDS data set is 0.9040, the accuracy rate is 0.9563, the recall rate is 0.8514, and the F1 score is 0.9008. The results show that the URS extracted by the POI–NTL–LDS data set has high precision.

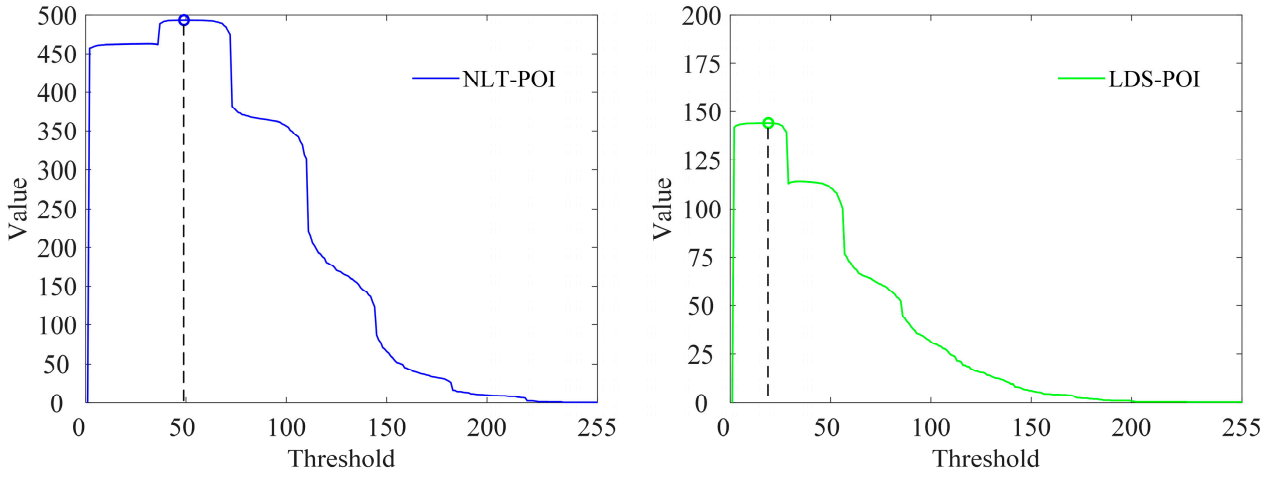


**Figure 11.** Threshold extraction of POI–NTL and POI–LDS data set.

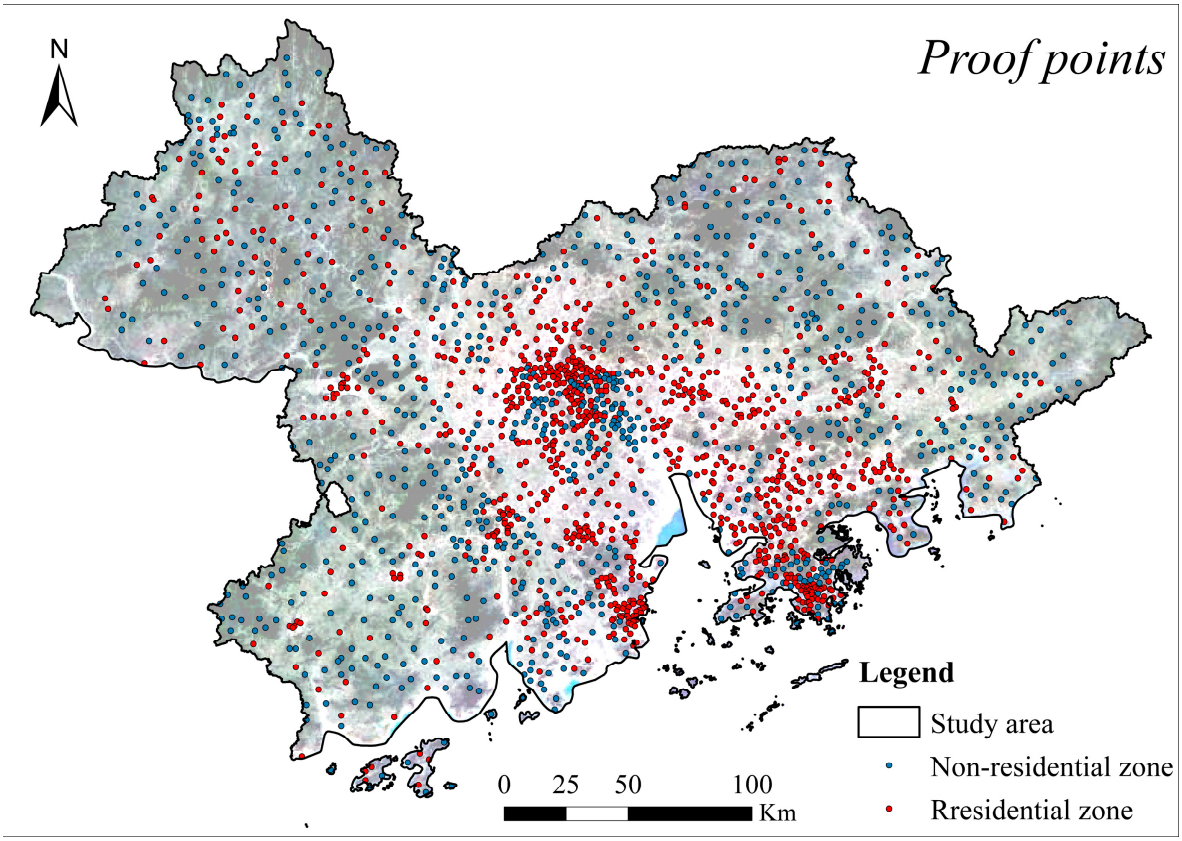


**Figure 12.** Random verification points.

**Table 2.** Accuracy verification results of different fusion data.

| Fused Data | Accuracy | Precision | Recall | F1-Score |
|---|---|---|---|---|
| NLT–POI | 0.8152 | 0.9228 | 0.6973 | 0.7943 |
| LDS–POI | 0.7770 | 0.8188 | 0.7245 | 0.7688 |
| NLT–LDS–POI | 0.9040 | 0.9563 | 0.8514 | 0.9008 |

From the extracted URS, the recognition results of different data fusions also have certain differences. Three levels of medium, high, and very high are taken as URSs with high confidence, and their area and proportion are calculated, respectively. It is calculated that the area of high confidence URS estimated by POI–NTL is 3106.80 km$^2$, representing 6.18% of the GBA urban agglomeration. The high confidence URS identified by POI–LDS is 784.24 km$^2$, representing 1.56% of the total area of the GBA urban agglomeration. The high confidence URS identified by POI–NTL–LDS is 2980.97 km$^2$, representing 5.93% of the total area of the GBA urban agglomeration. It can be seen that the URS extracted by NLT–POI data is larger than that of POI–LDS data, but the accuracy is only slightly higher than that of POI–LDS data. The URS extracted from NLT–LDS–POI data is similar to that from NLT–POI data, but it has the highest accuracy. This shows that three data fusions can significantly increase the precision of URS extraction after two data fusions.

## 5. Discussion

This paper analyzes the main characteristics of POI data, NTL data, LDS data, and HRI data. The performance of different data fusions in URS estimates is analyzed while fully considering the characteristics and advantages of various data. This study believes that the URS extracted by the SWF method and multi-source data fusion have significant advantages.

POI data and NTL data are now more and more applied to the identification and extraction of URS, but they both have obvious defects. POI data do not contain some information, such as building area and building scale, and different types of POIs are unbalanced, resulting in the redundancy of POIs in the same cell. The phenomenon of light overflow and supersaturation exists in NTL data in urban spatial research, which reduces the research accuracy. Therefore, relevant studies began to try to obtain more accurate research results by integrating POI data and NTL data. Most studies show that the fusion data can successfully compensate for the shortcomings of single-source data [32,34], especially in the extraction of UFS, and its extraction accuracy has been significantly improved [43]. Although single-source data or the fusion of POI data and NTL data have been used to identify URS, taking into account residential location information and human activity characteristics, most studies have ignored the key impact of population spatial distribution on URS. However, the traditional population distribution data rely on the statistical data of administrative divisions, and it is difficult to express the spatial differences of population distribution within administrative divisions. LDS, as the most accurate global population distribution data, has a higher granularity in reflecting the spatial pattern of the population.

Based on the characteristics of URS, the effect of population agglomeration on the estimate of URS is more fully considered in this study. Based on the POI–NTL data set and POI–LDS data set, the pixel-level fusion of the three kinds of data is further carried out. The accuracy of the URS extraction by the fused NLT–LDS–POI data set reaches 90.40%, which is higher than the previous results, and also higher than other studies on URS. In addition, the NLT–LDS–POI data set further integrates the population distribution data on the premise of highlighting the spatial function, making the extraction of URS more reasonable and accurate.

At present, in the relevant research on urban space recognition, dichotomy and threshold methods are mainly used to process the image pixels, which take a long time and have insufficient accuracy [72,73]. URS recognition is essentially the extraction of image features. Therefore, in this study, SLIC image superpixel segmentation is used to preprocess HRI, and DCWT is used to segment low-frequency information and high-frequency information of images by different methods, and feature fusion of different data is realized by superpixel wavelet fusion. Finally, the optimal URS extraction range is obtained by OTSU adaptive threshold calculation. This method can obtain the estimated results of URS faster and more precisely, which is meaningful for the application of multi-source data in urban spatial investigation.

The verification results show that the three data fusion data sets have significant advantages in the extraction of URS. In other types of data sets, the URS extracted by the LDS–POI data set has low precision, which also reflects the spatial mismatch between population and housing; that is, the population is more concentrated in some URS, which leads to a large number of URS with few people living, and obvious housing vacancy phenomenon. This also provides a way for us to further study the problem of UFS mismatch.

However, this study did have some limitations. First, the role of LDS data in improving the extraction accuracy of URS is limited. LDS data are based on demographic data and obtained by image analysis technology and a multiple partition density model, which is easily affected by statistical rules and the caliber of official statistical data. Secondly, the spatial scale of the study needs to be refined. As one of the UFS, URS has different types, spatial scales, spatial distribution patterns, etc. The spatial scale of the current research is not enough to support more in-depth research on URS. Finally, URS is dynamic and it will change with the actual development of the city.

These limitations point the way for future research. For example, open access urban land survey data or cadastral maps have the potential to provide more detailed and authoritative data for URS identification and extraction. The fusion application of these data is conducive to improving the accuracy of URS extraction. In addition, focusing the research area on the central area of the city, identifying smaller spatial scales, further

identifying the type of URS, and exploring its characteristics, rules, and influencing factors may be more valuable for the study of urban space [74]. Finally, to better promote the formulation of regional policies for urban agglomerations, it is also necessary to study the dynamics of URS in a long time series and simulate the future URS, so as to better serve urban planning and practice [75].

## 6. Conclusions

Identifying and clarifying the layout of URS is conducive to further research on the differentiation of URS and the geographical matching relationship between URS and other UFS, laying a foundation for optimizing and adjusting the structure of UFS. This has positive implications for promoting more scientific urban spatial planning and more harmonious social development. In this study, the SLIC image superpixel segmentation is firstly carried out on HRI, and on the basis of data fusion of POI data, NTL data, and LDS data, different data sets are fused at pixel level through SWT. Finally, the optimal range of URS is obtained through the OSTU adaptive threshold calculation. By verifying the accuracy of URS extracted from different data, it is found that the precision of estimates of NLT–POI, LDS–POI, and NLT–LDS–POI data sets were 81.52%, 77.70%, and 90.40%, respectively. The main conclusions are as follows: (1) The automatic identification and extraction process of URS is optimized by using the SWF method. This study introduces a SWF method to address multi-source data integration in URS extraction for the first time. This method improves the accuracy of extraction and provides a new perspective and method for detecting complex change characteristics of modern URS. (2) Compared with the previous method, SWF fully considers the spatial distribution law of HRI, and the spatial information of URS is embedded into the WT to obtain the identification results of URS with multi-source data fusion under the perception of spatial structure. (3) The data set after the integration of population distribution data strengthens the spatial function of URS by reflecting the spatial agglomeration of the population and helps to improve the accuracy of URS extraction.

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

**Funding:** This research was funded by the National Natural Science Foundation of China under grant no. 42371208.

**Data Availability Statement:** The data for this research were retrieved from: https://www.amap.com/ (accessed on 30 January 2023); http://geodata.nnu.edu.cn/ (accessed on 9 March 2024); https://landscan.ornl.gov/ (accessed on 16 June 2024); https://earth.google.com/ (accessed on 16 June 2024).

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Corbusier, L.; Eardley, A. *The Athens Charter*; Grossman Publishers: New York, NY, USA, 1978.
2. Bourdieu, P. Social Space and the Genesis of Appropriated Physical Space. *Int. J. Urban Reg. Res.* **2018**, *42*, 106–114. [CrossRef]
3. Kang, Y.; Zhang, F.; Gao, S.; Peng, W.; Ratti, C. Human settlement value assessment from a place perspective: Considering human dynamics and perceptions in house price modeling. *Cities* **2021**, *118*, 103333. [CrossRef]
4. Jacobs, A.; Appleyard, D. Toward an Urban Design Manifesto. In *The City Reader*; Routledge: London, UK, 2015; pp. 640–651.
5. Babalola, O.D.; Ibem, E.O.; Olotuah, A.O.; Opoko, A.P.; Adewale, B.A.; Fulani, O.A. Housing quality and its predictors in public residential estates in Lagos, Nigeria. *Environ. Dev. Sustain.* **2020**, *22*, 3973–4005. [CrossRef]
6. Sheikh, W.T.; van Ameijde, J. Promoting livability through urban planning: A comprehensive framework based on the "theory of human needs". *Cities* **2022**, *131*, 103972. [CrossRef]
7. Blomley, N. Precarious Territory: Property Law, Housing, and the Socio-Spatial Order. *Antipode* **2020**, *52*, 36–57. [CrossRef]
8. Jabareen, Y.; Eizenberg, E. Theorizing urban social spaces and their interrelations: New perspectives on urban sociology, politics, and planning. *Plan. Theory* **2021**, *20*, 211–230. [CrossRef]

9. Arundel, R.; Hochstenbach, C. Divided access and the spatial polarization of housing wealth. *Urban Geogr.* **2020**, *41*, 497–523. [CrossRef]
10. Tong, D.; Zhang, Y.; MacLachlan, I.; Li, G. Migrant housing choices from a social capital perspective: The case of Shenzhen, China. *Habitat Int.* **2020**, *96*, 102082. [CrossRef]
11. Lai, Y.; Tang, B.; Chen, X.; Zheng, X. Spatial determinants of land redevelopment in the urban renewal processes in Shenzhen, China. *Land Use Policy* **2021**, *103*, 105330. [CrossRef]
12. Yue, W.; Chen, Y.; Thy, P.T.M.; Fan, P.; Liu, Y.; Zhang, W. Identifying urban vitality in metropolitan areas of developing countries from a comparative perspective: Ho Chi Minh City versus Shanghai. *Sust. Cities Soc.* **2021**, *65*, 102609. [CrossRef]
13. Cao, R.; Tu, W.; Yang, C.; Li, Q.; Liu, J.; Zhu, J.; Zhang, Q.; Li, Q.; Qiu, G. Deep learning-based remote and social sensing data fusion for urban region function recognition. *ISPRS J. Photogramm. Remote Sens.* **2020**, *163*, 82–97. [CrossRef]
14. Xue, B.; Xiao, X.; Li, J.; Zhao, B.; Fu, B. Multi-source Data-driven Identification of Urban Functional Areas: A Case of Shenyang, China. *Chin. Geogr. Sci.* **2023**, *33*, 21–35. [CrossRef]
15. Wang, Y.; Huang, H.; Yang, G.; Chen, W. Ecosystem Service Function Supply-Demand Evaluation of Urban Functional Green Space Based on Multi-Source Data Fusion. *Remote Sens.* **2023**, *15*, 118. [CrossRef]
16. Deng, Y.; He, R. Refined Urban Functional Zone Mapping by Integrating Open-Source Data. *ISPRS Int. J. Geo Inf.* **2022**, *11*, 421. [CrossRef]
17. Zhou, L.; Wei, L.; Lopez-Carr, D.; Dang, X.; Yuan, B.; Yuan, Z. Identification of irregular extension features and fragmented spatial governance within urban fringe areas. *Appl. Geogr.* **2024**, *162*, 103172. [CrossRef]
18. Tian, Y.; Mao, Q. The effect of regional integration on urban sprawl in urban agglomeration areas: A case study of the Yangtze River Delta, China. *Habitat Int.* **2022**, *130*, 102695. [CrossRef]
19. Gao, J. Mitigating Pernicious Gaming in Performance Management in China: Dilemmas, Strategies and Challenges. *Public Perform. Manag. Rev.* **2021**, *44*, 321–351. [CrossRef]
20. Chen, C.; Yan, J.; Wang, L.; Liang, D.; Zhang, W. Classification of Urban Functional Areas from Remote Sensing Images and Time-Series User Behavior Data. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2021**, *14*, 1207–1221. [CrossRef]
21. Zhang, K.; Ming, D.; Du, S.; Xu, L.; Ling, X.; Zeng, B.; Lv, X. Distance Weight-Graph Attention Model-Based High-Resolution Remote Sensing Urban Functional Zone Identification. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–16. [CrossRef]
22. Zhou, C.; He, Z.; Lou, A.; Plaza, A. RGB-to-HSV: A Frequency-Spectrum Unfolding Network for Spectral Super-Resolution of RGB Videos. *IEEE Trans. Geosci. Remote Sens.* **2024**, *62*, 5609318. [CrossRef]
23. Zhou, C.; Shi, Q.; He, D.; Tu, B.; Li, H.; Plaza, A. Spectral-Spatial Sequence Characteristics-Based Convolutional Transformer for Hyperspectral Change Detection. *CAAI Trans. Intell. Technol.* **2023**, *8*, 1237–1257. [CrossRef]
24. Huang, C.; Xiao, C.; Rong, L. Integrating Point-of-Interest Density and Spatial Heterogeneity to Identify Urban Functional Areas. *Remote Sens.* **2022**, *14*, 4201. [CrossRef]
25. Yeow, L.W.; Low, R.; Tan, Y.X.; Cheah, L. Point-of-Interest (POI) Data Validation Methods: An Urban Case Study. *ISPRS Int. J. Geo Inf.* **2021**, *10*, 735. [CrossRef]
26. Zhou, N. Research on urban spatial structure based on the dual constraints of geographic environment and POI big data. *J. King Saud Univ. Sci.* **2022**, *34*, 101887. [CrossRef]
27. Wu, R.; Wang, J.; Zhang, D.; Wang, S. Identifying different types of urban land use dynamics using Point-of-interest (POI) and Random Forest algorithm: The case of Huizhou, China. *Cities* **2021**, *114*, 103202. [CrossRef]
28. Andrade, R.; Alves, A.; Bento, C. POI Mining for Land Use Classification: A Case Study. *ISPRS Int. J. Geo Inf.* **2020**, *9*, 493. [CrossRef]
29. Estima, J.; Painho, M. Investigating the potential of open street map for land use/land cover production: A case study for continental Portugal. In *OpenStreetMap in GIScience*; Springer: Cham, Switzerland, 2015; pp. 273–293.
30. Casali, Y.; Aydin, N.Y.; Comes, T. Machine learning for spatial analyses in urban areas: A scoping review. *Sust. Cities Soc.* **2022**, *85*, 104050. [CrossRef]
31. Yu, M.; Li, J.; Lv, Y.; Xing, H.; Wang, H. Functional Area Recognition and Use-Intensity Analysis Based on Multi-Source Data: A Case Study of Jinan, China. *ISPRS Int. J. Geo Inf.* **2021**, *10*, 640. [CrossRef]
32. Lu, W.; Tao, C.; Li, H.; Qi, J.; Li, Y. A unified deep learning framework for urban functional zone extraction based on multi-source heterogeneous data. *Remote Sens. Environ.* **2022**, *270*, 112830. [CrossRef]
33. Tian, Y.; Qian, J. Suburban identification based on multi-source data and landscape analysis of its construction land: A case study of Jiangsu Province, China. *Habitat Int.* **2021**, *118*, 102459. [CrossRef]
34. He, X.; Zhou, C.; Zhang, J.; Yuan, X. Using Wavelet Transforms to Fuse Nighttime Light Data and POI Big Data to Extract Urban Built-Up Areas. *Remote Sens.* **2020**, *12*, 3887. [CrossRef]
35. Zhang, J.; Zhang, X.; Tan, X.; Yuan, X. Extraction of Urban Built-Up Area Based on Deep Learning and Multi-Sources Data Fusion-The Application of an Emerging Technology in Urban Planning. *Land* **2022**, *11*, 1212. [CrossRef]
36. Hu, S.; Huang, S.; Hu, Q.; Wang, S.; Chen, Q. An commercial area extraction approach using time series nighttime light remote sensing data--Take Wuhan city as a case. *Sust. Cities Soc.* **2024**, *100*, 105032. [CrossRef]
37. Huang, X.; Yang, J.; Li, J.; Wen, D. Urban functional zone mapping by integrating high spatial resolution nighttime light and daytime multi-view imagery. *ISPRSJ. Photogramm. Remote Sens.* **2021**, *175*, 403–415. [CrossRef]

38. Hu, Q.; Wu, W.; Xia, T.; Yu, Q.; Yang, P.; Li, Z.; Song, Q. Exploring the Use of Google Earth Imagery and Object-Based Methods in Land Use/Cover Mapping. *Remote Sens.* **2013**, *5*, 6026–6042. [CrossRef]
39. Horota, R.K.; Aires, A.S.; Marques, A.; Rossa, P.; de Souza, E.M.; Gonzaga, L.; Veronez, M.R. Printgrammetry-3-D Model Acquisition Methodology from Google Earth Imagery Data. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2020**, *13*, 2819–2830. [CrossRef]
40. Yang, D.; Fu, C.; Smith, A.; Yu, Q. Open land-use map: A regional land-use mapping strategy for incorporating OpenStreetMap with earth observations. *Geo Spat. Inf. Sci.* **2017**, *20*, 269–281. [CrossRef]
41. Wu, H.; Luo, W.; Lin, A.; Hao, F.; Olteanu-Raimond, A.; Liu, L.; Li, Y. SALT: A multifeature ensemble learning framework for mapping urban functional zones from VGI data and VHR images. *Comput. Environ. Urban Syst.* **2023**, *100*, 101921. [CrossRef]
42. Zang, N.; Cao, Y.; Wang, Y.; Huang, B.; Zhang, L.; Mathiopoulos, P.T. Land-Use Mapping for High-Spatial Resolution Remote Sensing Image Via Deep Learning: A Review. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2021**, *14*, 5372–5391. [CrossRef]
43. Chen, S.; Ogawa, Y.; Zhao, C.; Sekimoto, Y. Large-scale individual building extraction from open-source satellite imagery via super-resolution-based instance segmentation approach. *ISPRS J. Photogramm. Remote Sens.* **2023**, *195*, 129–152. [CrossRef]
44. Chen, J.; Chen, L.; Li, Y.; Zhang, W.; Long, Y. Measuring Physical Disorder in Urban Street Spaces: A Large-Scale Analysis Using Street View Images and Deep Learning. *Ann. Am. Assoc. Geogr.* **2023**, *113*, 469–487. [CrossRef]
45. Lin, A.; Wu, H.; Liang, G.; Cardenas-Tristan, A.; Wu, X.; Zhao, C.; Li, D. A big data-driven dynamic estimation model of relief supplies demand in urban flood disaster. *Int. J. Disaster Risk Reduct.* **2020**, *49*, 101682. [CrossRef]
46. Zhu, X.; Zhou, C. POI Inquiries and data update based on LBS. In Proceedings of the IEEC 2009: First International Symposium on Information Engineering and Electronic Commerce, Ternopil, Ukraine, 6–17 May 2009; pp. 730–734.
47. Liu, K.; Qiu, P.; Gao, S.; Lu, F.; Jiang, J.; Yin, L. Investigating urban metro stations as cognitive places in cities using points of interest. *Cities* **2020**, *97*, 102561. [CrossRef]
48. Fonte, C.C.; Martinho, N. Assessing the applicability of OpenStreetMap data to assist the validation of land use/land cover maps. *Geogr. Inf. Syst.* **2017**, *31*, 2382–2400. [CrossRef]
49. Lin, A.; Sun, X.; Wu, H.; Luo, W.; Wang, D.; Zhong, D.; Wang, Z.; Zhao, L.; Zhu, J. Identifying Urban Building Function by Integrating Remote Sensing Imagery and POI Data. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2021**, *14*, 8864–8875. [CrossRef]
50. Chen, Y.; He, C.; Guo, W.; Zheng, S.; Wu, B. Mapping urban functional areas using multi-source remote sensing images and open big data. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2023**, *16*, 7919–7931. [CrossRef]
51. He, X.; Yuan, X.; Zhang, D.; Zhang, R.; Li, M.; Zhou, C. Delineation of Urban Agglomeration Boundary Based on Multisource Big Data Fusion—A Case Study of Guangdong–Hong Kong–Macao Greater Bay Area (GBA). *Remote Sens.* **2021**, *13*, 1801. [CrossRef]
52. Dobson, J.E.; Bright, E.A.; Coleman, P.R.; Durfee, R.C.; Worley, B.A. LandScan: A global population database for estimating populations at risk. *Photogramm. Eng. Remote Sens.* **2000**, *66*, 849–857.
53. Yuan, X.; Chen, B.; He, X.; Zhang, G.; Zhou, C. Spatial Differentiation and Influencing Factors of Tertiary Industry in the Pearl River Delta Urban Agglomeration. *Land* **2024**, *13*, 172. [CrossRef]
54. Meng, M.; Shang, Y.; Yang, Y. Did highways cause the urban polycentric spatial structure in the Shanghai metropolitan area? *J. Transp. Geogr.* **2021**, *92*, 103022. [CrossRef]
55. He, X.; Cao, Y.; Zhou, C. Evaluation of Polycentric Spatial Structure in the Urban Agglomeration of the Pearl River Delta (PRD) Based on Multi-Source Big Data Fusion. *Remote Sens.* **2021**, *13*, 3639. [CrossRef]
56. He, X.; Zhang, R.; Yuan, X.; Cao, Y.; Zhou, C. The role of planning policy in the evolution of the spatial structure of the Guangzhou metropolitan area in China. *Cities* **2023**, *137*, 104284. [CrossRef]
57. AutoNavi. Amap API. 2021. Available online: https://lbs.amap.com/api/webservice/guide/api/search/ (accessed on 30 January 2023).
58. Chen, Z.; Yu, B.; Yang, C.; Zhou, Y.; Yao, S.; Qian, X.; Wang, C.; Wu, B.; Wu, J. An extended time series (2000–2018) of global NPP-VIIRS-like nighttime light data from a cross-sensor calibration. *Earth Syst. Sci. Data.* **2021**, *13*, 889–906. [CrossRef]
59. Sims, K.; Reith, A.; Bright, E.; Kaufman, J.; Pyle, J.; Epting, J.; Gonzales, J.; Adams, D.; Powell, E.; Urban, M.; et al. *Land Scan Global 2022 [Data Set]*; Oak Ridge National Laboratory: Oak Ridge, TN, USA, 2023.
60. Chen, W.; Huang, H.; Dong, J.; Zhang, Y.; Tian, Y.; Yang, Z. Social functional mapping of urban green space using remote sensing and social sensing data. *ISPRS J. Photogramm. Remote Sens.* **2018**, *146*, 436–452. [CrossRef]
61. Zhang, C.; Sargent, I.; Pan, X.; Li, H.; Gardiner, A.; Hare, J.; Atitinson, P.M. An object-based convolutional neural network (OCNN) for urban land use classification. *Remote Sens. Environ.* **2018**, *216*, 57–70. [CrossRef]
62. Van de Voorde, T.; Jacquet, W.; Canters, F. Mapping form and function in urban areas: An approach based on urban metrics and continuous impervious surface data. *Landsc. Urban Plan.* **2011**, *102*, 143–155. [CrossRef]
63. Achanta, R.; Shaji, A.; Smith, K.; Lucchi, A.; Fua, P.; Suesstrunk, S. SLIC Superpixels Compared to State-of-the-Art Super-pixel Methods. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *34*, 2274–2281. [CrossRef]
64. Shensa, M.J. The discrete wavelet transform: Wedding the a trous and Mallat algorithms. *IEEE Trans. Signal Process.* **1992**, *40*, 2464–2482. [CrossRef]
65. Sifuzzaman, M.; Islam, M.R.; Ali, M.Z. Application of wavelet transform and its advantages compared to Fourier transform. *J. Phys. Sci.* **2009**, *13*, 121–134.
66. Farge, M. Wavelet transforms and their applications to turbulence. *Annu. Rev. Fluid Mech.* **1992**, *24*, 395–458. [CrossRef]

67. Mallat, S.G. Multifrequency channel decompositions of images and wavelet models. *IEEE Trans. Acoust. Speech Signal Process.* **1989**, *37*, 2091–2110. [CrossRef]
68. Mallat, S.G. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Trans. Pattern Anal. Mach. Intell.* **1989**, *11*, 674–693. [CrossRef]
69. Lewis, J.J.; O'Callaghan, R.J.; Nikolov, S.G.; Bull, D.R.; Canagarajah, N. Pixel- and region-based image fusion with complex wavelets. *Inf. Fusion* **2007**, *8*, 119–130. [CrossRef]
70. Pajares, G.; de la Cruz, J.M. A wavelet-based image fusion tutorial. *Pattern Recognit.* **2004**, *37*, 1855–1872. [CrossRef]
71. Luescher, P.; Weibel, R. Exploiting empirical knowledge for automatic delineation of city centres from large-scale topographic databases. *Comput. Environ. Urban Syst.* **2013**, *37*, 18–34. [CrossRef]
72. Du, S.; Du, S.; Liu, B.; Zhang, X. Mapping large-scale and fine-grained urban functional zones from VHR images using a multi-scale semantic segmentation network and object based approach. *Remote Sens. Environ.* **2021**, *261*, 112480. [CrossRef]
73. Shang, R.; Peng, P.; Shang, F.; Jiao, L.; Shen, Y.; Stolkin, R. Semantic Segmentation for SAR Image Based on Texture Complexity Analysis and Key Superpixels. *Remote Sens.* **2020**, *12*, 2141. [CrossRef]
74. He, X.; Zhou, Y. Urban spatial growth and driving mechanisms under different urban morphologies: An empirical analysis of 287 Chinese cities. *Landsc. Urban Plan.* **2024**, *248*, 105096. [CrossRef]
75. He, X.; Zhou, Y.; Yuan, X.; Zhu, M. The coordination relationship between urban development and urban life satisfaction in Chinese cities-An empirical analysis based on multi-source data. *Cities* **2024**, *150*, 105016. [CrossRef]


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