International Journal of Geo-Information

MDPI

Article

Identifying the Spatial Range of the Pearl River Delta Urban Agglomeration by Fusing Nighttime Light Data with Weibo Sign-In Data

Yongwang Cao [1,2], Song Liu [2] and Zaigao Yang [2,*]

[1] School of Economics and Finance, South China University of Technology, Guangzhou 510006, China; caoyw6@mail2.sysu.edu.cn
[2] Guangzhou Academy of Social Sciences, Guangzhou 510410, China; liusong6@mail2.sysu.edu.cn
* Correspondence: yzgao@gz.gov.cn; Tel.: +86-020-86464381

Abstract: Accurately identifying the spatial range of urban agglomerations holds significant practical importance for the precise allocation of various elements and coordinated development within urban agglomerations. However, current research predominantly focuses on the physical spaces of urban agglomerations, overlooking their sphere of influence. This study begins with the spatial interactions of population elements within urban agglomerations and fuses Weibo sign-in data with NTL data to identify the spatial range of urban agglomerations. It further compares and validates the results before and after the fusion of data. The results reveal that the accuracy of identifying the spatial range of urban agglomerations with the fusion of NTL data and Weibo sign-in data has improved by 7%, with a Kappa increase of 0.1766 compared to using NTL data alone, which indicates that fusing social media data can significantly enhance the accuracy of identifying the spatial range of urban agglomerations. This study proposes a novel approach for identifying the spatial range of urban agglomerations through the fusion of NTL data and social media data from a data fusion perspective. On one hand, it supplements the application of data fusion in the study of urban agglomeration spaces; on the other hand, it accurately identifies the spatial range of urban agglomerations, which holds great practical value for the sustainable development of urban agglomerations.

Keywords: nighttime light (NTL) data; Pearl River Delta (PRD) urban agglomeration; Weibo sign-in data; spatial characteristics; spatial range

check for updates

Citation: Cao, Y.; Liu, S.; Yang, Z. Identifying the Spatial Range of the Pearl River Delta Urban Agglomeration by Fusing Nighttime Light Data with Weibo Sign-In Data. *ISPRS Int. J. Geo-Inf.* **2024**, *13*, 214. https://doi.org/10.3390/ijgi13060214

Academic Editors: Wolfgang Kainz and Jamal Jokar Arsanjani

Received: 20 February 2024
Revised: 8 June 2024
Accepted: 18 June 2024
Published: 19 June 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

Urban agglomerations represent the pinnacle spatial organizational form of urban development in its mature phase, denoting regions within specific territorial boundaries that are generally comprised of multiple large cities [1,2]. These urban clusters emerge over areas characterized by tight spatial organization and closely-knit economic linkages, which are facilitated by an advanced infrastructure network of transport and communication, ultimately achieving a high degree of urban coalescence and integration [3]. As urban agglomerations experience rapid development, the spatial influence of urban agglomerations extends far beyond these administrative borders and built-up areas [4]. While the existing administrative boundaries and built-up areas align more closely with the actual form and developmental status of urban agglomerations, they are also limited in their ability to reflect the flow of factors and connections within the agglomerations [5]. Consequently, they may fail to comprehensively capture the dynamics and complexities of urban agglomerations [6]. Therefore, for the rational allocation of resources and factors within urban agglomerations, it is imperative to accurately identify their spatial ranges. Such identification can facilitate a more effective resolution of the developmental challenges within urban agglomerations, including optimizing resource allocation and promoting rational interaction and coordinated development among cities within the region [7,8]. To address these challenges, it is

necessary to go beyond the conventional methods of identifying urban agglomerations and further analyze the inherent spatial influences in the contemporary era. Subsequently, a more precise method for identifying urban agglomerations can be proposed.

In past studies on identifying the spatial ranges of urban agglomerations, commonly used data types generally include statistical survey data and remote sensing data [9]. NTL data, as an important branch of remote sensing data, can capture the distribution of nighttime light sources on the ground, providing intuitive indicators of urbanization levels and economic activities [10]. NTL data not only vividly reflect the differences between urban agglomeration areas and surrounding regions but also excel in unveiling areas of rapid urbanization [11]. Furthermore, NTL data can be used to monitor the expansion trends of urban spaces [12]. This leads to the widespread application of NTL data in research related to urban interior spaces, including the delineation of urban boundaries, identification of urban centers, and analysis of factors influencing the spatial structure of urban spaces [13,14]. Currently, commonly used NTL data include NPP/VIIRS, DMSP/OLS, and Luojia-01. The DMSP/OLS NTL data have a spatial resolution of 1 km and cover the time period from 1992 to 2012. The NPP/VIIRS NTL data have a spatial resolution of 500 m and span from 2013 to 2024. In contrast, the Luojia-01 NTL data boast a higher spatial resolution of 130 m but are mainly concentrated in the years 2018–2019. Additionally, in the process of identifying the spatial range of an urban area, NTL data face limitations due to their inherent spillover effect and the inability to fully reflect the internal factor flows within an urban interior space, resulting in certain inaccuracies in the spatial range identification [15]. Therefore, when using NTL data to identify the spatial range of an urban area, it is necessary to fuse other types of data and methods to mitigate the impacts of the nighttime light spillover effect and comprehensively reflect the internal factor flows, aiming to achieve more comprehensive and precise results in identifying the spatial range of an urban interior space.

In the contemporary information society, the development of big data technology provides new avenues for identifying and understanding the spatial range of an urban interior space [16]. Particularly, big data from social media, originating from residents' daily activities, offers a unique perspective for revealing the interactions and connections within urban interior spaces [17]. Weibo sign-in data, as a form of social media data, can directly map the social and economic connections between cities through online interactions among crowds [18]. For urban spaces, the flow of internal factors is a crucial prerequisite for identifying their spatial range, and analyzing the spatial interactions of different users within Weibo sign-in data can easily determine the range of the internal population element flow, thereby defining the spatial range of an urban space [19,20]. A considerable amount of research utilizes the characteristics of Weibo sign-in data to analyze content related to urban agglomerations, including the identification of urban agglomeration boundaries and the spatial connections between different cities within urban agglomerations [21]. These studies reflect the promising prospects of applying Weibo sign-in data to this field. However, there are also some limitations to the use of social media data in urban spatial analysis [22]. Firstly, social media users do not represent the entire population, which may lead to sampling bias. Secondly, social media data may be subject to data bias, such as the uneven distribution of user activity and the presence of false information [23]. Therefore, like other big data, social media data represent a simulation effect in a geo-virtual space and need to be used in conjunction with a physical space.

In the context of identifying the spatial range of an urban area, both NTL data and social media data serve as significant sources of information. However, the distinct features of these data types vary in their application. NTL data's advantage lies in its provision of intuitive, continuous geospatial information, which can clearly depict the physical form of cities and urban agglomerations [24]. Nevertheless, this method also has its limitations, as it cannot offer insights into the socio-economic activities and population movements within urban agglomerations [25]. On the other hand, the advantage of Weibo sign-in data is its reflection of the dynamism and immediacy of human interactions, offering a more

authentic representation of the urban agglomeration's influence [26]. Moreover, social media data can provide in-depth insights into the lifestyles, consumption habits, and cultural characteristics of urban agglomeration residents, thereby offering more comprehensive information support for urban planning and regional development [27]. However, the inherent characteristics of the data can lead to sampling biases and data skewness issues, making the use of Weibo sign-in data in urban agglomeration-related research somewhat inadequate [28]. Consequently, an increasing number of studies are engaging in research from the perspective of fusing NTL data with urban big data, considering the combination of both data types to identify and understand the spatial ranges of urban agglomerations more comprehensively and accurately [29].

In recent years, data fusion has been widely applied in research related to urban interior spaces. By fusing and analyzing information from diverse data sources, the precision of observations can be enhanced. Currently, the more commonly used data fusion techniques include algebraic methods, Intensity–Hue–Saturation (IHS) transformation, wavelet transformation, Principal Component Transformation (PCT), and K-T transformation, among others [17,30]. In terms of data fusion, the variability in the outcomes is largely dependent on the method of fusion employed, with different techniques yielding varying effects on the fused data. Specifically, wavelet transformation is noted for its ability to retain the information characteristics of the original data images as much as possible during the image fusion process, and it is increasingly applied in urban spatial studies [31,32]. Existing research, through the fusion of various data types, has substantiated this point, for instance, by integrating NTL data with Point-of-Interest (POI) data to extract urban built-up areas [33], delineate urban boundaries [34], etc. The findings from these studies indicate that data fusion generally provides stronger results in urban spatial applications than the use of single data sources [35]. Recent studies have also explored the fusion of NTL data with Weibo data to analyze intra-urban spaces. For example, researchers have found that NTL data can reflect the economic development level of a city, while Weibo sign-in data capture residents' emotional needs regarding urban spaces. By fusing NTL data with Weibo sign-in data, a comprehensive assessment of urban life satisfaction in China has been conducted [36]. Therefore, building upon this foundation, we aim to effectively fuse social media data with NTL data through data fusion techniques, thereby more accurately identifying the spatial range of urban agglomerations.

This study is conducted from several perspectives. Firstly, we identify the spatial range of urban agglomerations based on NTL data. Secondly, this study fuses NTL data with social media data to identify the spatial range of urban agglomerations. Thirdly, this study compares and validates the results obtained from these two identification methods. The primary contribution of this study lies in proposing a novel method for identifying urban spatial boundaries by fusing NTL data with Weibo sign-in big data through neural networks. This approach offers a new perspective for researching intra-urban spaces. Additionally, identifying more accurate urban spatial boundaries aids in optimizing urban spatial structures and formulating development policies, thereby promoting high-quality and sustainable urban development.

2. Materials and Methods

2.1. Study Area

The PRD urban agglomeration encompasses nine cities in Guangdong Province: Guangzhou, Shenzhen, Zhuhai, Foshan, Huizhou, Dongguan, Zhongshan, Jiangmen, and Zhaoqing. The PRD urban agglomeration occupies 30.75% of the land area of Guangdong Province, yet it aggregates 70% of its population and generates 85% of the province's GDP, making it one of the most dynamic urban agglomerations in China. As of 2020, the urban population of the PRD reached 78.6 million, with an urbanization rate of 87.5%, which is among the highest in China [37]. With ongoing urbanization, the spatial range of the PRD urban agglomeration has undergone significant changes. This region has not only expanded in physical space but also seen an increasing influence in economic, social, and

cultural aspects, making the actual range of influence of the PRD urban agglomerations far beyond its traditional administrative boundaries. Therefore, identifying the spatial range of the PRD urban agglomeration aids in better understanding the synergistic development among different cities and elements within the area, providing a scientific basis for urban planning and development. Moreover, the experiences and methods in delineating the spatial range of the PRD urban agglomeration can offer insights and references for the development of other urban areas. Figure 1 presents the study area and the results of NTL data processing.

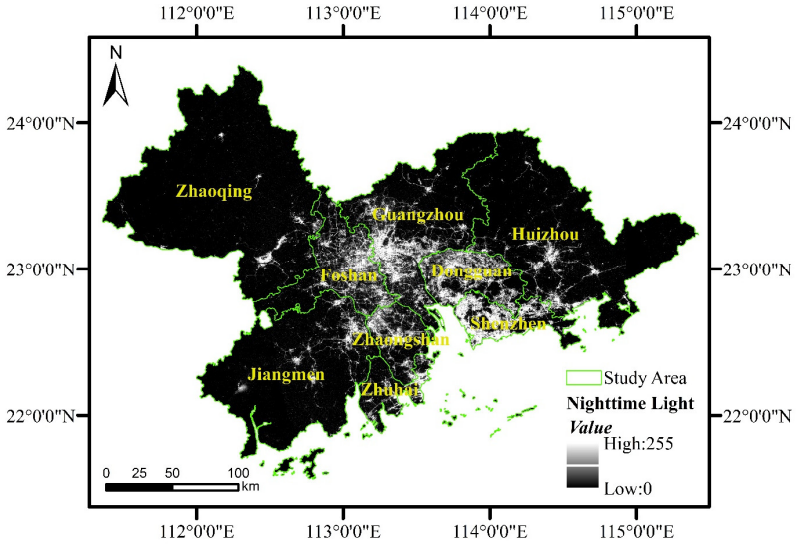

Figure 1. Study area and processing results of the NTL data.

2.2. Study Data

The data utilized in this study include NPP/VIIRS NTL data and social media data. The specific acquisition methods and processing procedures for different data are as follows:

2.2.1. NPP/VIIRS NTL Data

The NPP/VIIRS NTL data are obtained through observations by the Visible Infrared Imaging Radiometer Suite (VIIRS) on the Suomi National Polar-orbiting Partnership (NPP) satellite, operated by the National Aeronautics and Space Administration (NASA). First, the EOG website is accessed, and the appropriate product (e.g., monthly or annual composite data) is selected to download the required data files in HDF5 or GeoTIFF format. During the processing phase, geographic information system (GIS) software or programming languages (such as the GDAL and Rasterio libraries in Python) are used to unzip and read the data. Next, radiometric correction is performed, and clouds, fires, and other anomalously bright spots are removed. Subsequently, projection transformation, resampling, and standardization is conducted to ensure the data align with other geographic data sources. Finally, brightness thresholding and spatial analysis is applied to identify the boundaries of urban agglomerations. As a widely used source of NTL data, NPP/VIIRS offers superior spatial resolution and temporal series, presenting distinct advantages over other NTL datasets [38]. In this study, we access the 2022 NPP/VIIRS NTL data for the PRD urban agglomeration from the NASA Earth Observing System Data and Information System (EOSDIS) website. The acquired data undergo preprocessing, including cloud removal, radiance calibration, and image contrast adjustment, resulting in the preprocessed NTL data visualization of the PRD urban agglomeration, as shown in Figure 1.

2.2.2. Weibo Sign-In Data

Social media data are sourced from the Sina Weibo platform, which, along with Zhihu, Toutiao, and Baidu Tieba, constitutes the landscape of social media in China. Sina Weibo,

launched in 2009, stands as China's first mainstream social media platform. By the end of 2022, it maintained a stable active user base of 586 million, ranking it at the forefront of mainstream social media in China. Weibo sign-in data refer to the information generated by users' sign-in activities on Sina Weibo, documenting the specific locations and related details of user sign-ins. Compared to data from other social platforms, Sina Weibo sign-in data are characterized by their real-time nature and spatiotemporal attributes, capturing the timing and geographical location information of user sign-ins at various places [39]. In this study, we access the Sina Weibo open platform to acquire Weibo sign-in data for the year 2022, totaling approximately 10.34 million entries. The data attributes include geographical coordinates (latitude and longitude), sign-in locations, Weibo links, blogger homepage links, text content, links to images and videos, posting time, number of shares, comments, likes, and follower counts. Given the presence of erroneous or invalid sign-in entries within the acquired data, we process the data through cleaning, geocoding, and spatial clustering, among other methods. First, the Weibo API interface (e.g., the location service API) is used to collect sign-in data, including user IDs, sign-in times, sign-in locations, and other relevant information. After obtaining the data, data cleaning is performed, which includes removing duplicate sign-in records to ensure each record is unique; handling missing values by either deleting records with significant missing data or using interpolation methods to fill gaps; and standardizing the sign-in time format, such as converting it to a standard time format. The cleaned data can then undergo further processing, including converting geographic coordinates to the required coordinate system to ensure alignment with other geographic data sources. Consequently, we obtain the spatial distribution of Sina Weibo sign-in data for the PRD urban agglomeration as shown in Figure 2, with sign-in frequencies aggregated into 1 km grid squares.

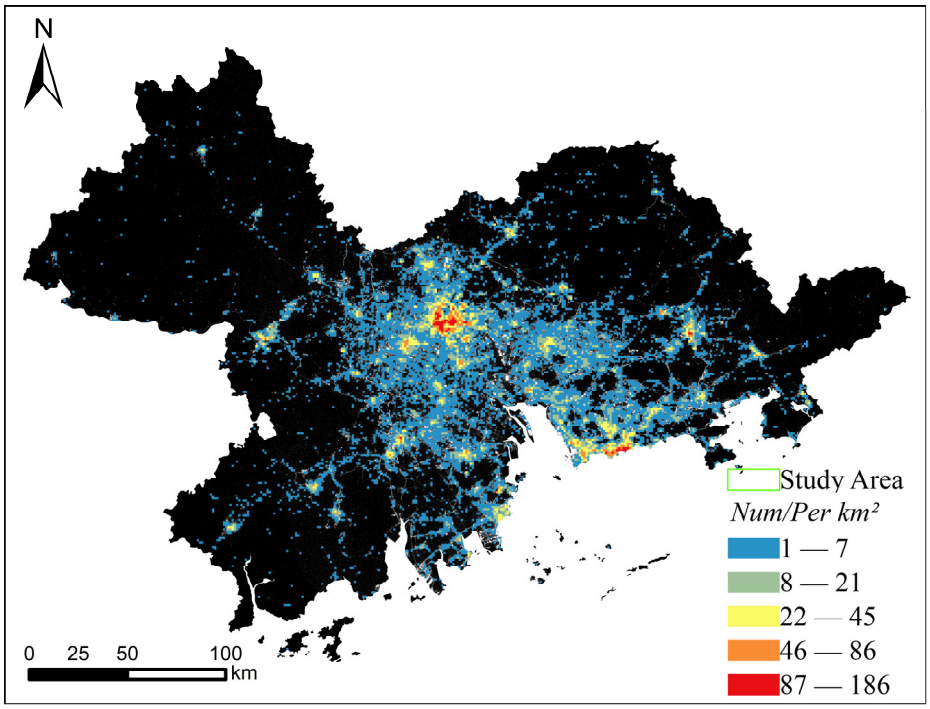

Figure 2. Processing results of Weibo sign-in data of PRD urban agglomeration.

In addition to NTL data and Weibo data, we also utilize 5000 randomly selected pixel validation points. These points are randomly chosen within the administrative boundaries of cities in the PRD urban agglomeration. Using high-resolution imagery from Google Earth, we verify that 1551 of these random pixels are within the urban agglomeration spatial boundaries, while 3449 are within the administrative boundaries but outside the

urban agglomeration spatial boundaries. These 5000 random pixel points are used for the subsequent verification of urban spatial identification accuracy.

2.3. Methods

The study workflow is presented in Figure 3.

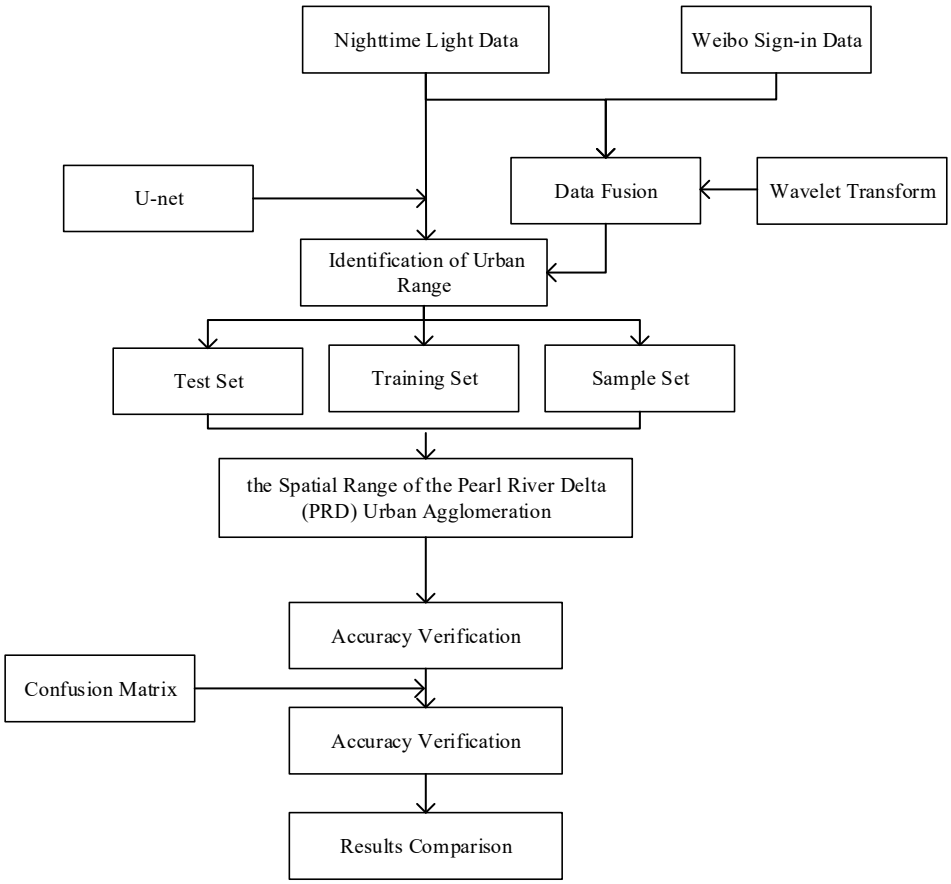

Figure 3. Workflow.

2.3.1. U-Net Automated Extraction of Urban Spaces

The essence of extracting urban agglomeration spaces lies in analyzing images generated from NTL and Weibo sign-in data to identify and extract areas with urban spatial characteristics. In the process of image feature extraction, the methods currently widely used include traditional techniques based on statistical and geometric features as well as methods based on deep learning [40]. Traditional approaches typically employ image segmentation and classification algorithms based on characteristics such as thresholds, textures, and shapes [41]. Meanwhile, deep learning-based methods utilize neural network architectures to automatically learn features within images for accurate classification and segmentation. Among these, U-net is a deep learning network architecture specifically designed for image semantic segmentation tasks and can be applied to the automated extraction of urban spaces [41], which is particularly important for understanding rapidly changing and highly complex urban agglomerations like the PRD urban agglomeration. The uniqueness of U-net lies in its encoder–decoder structure, which achieves rapid information transmission and feature extraction through stacked convolutional neural networks.

Specifically, the U-net neural network extracts urban spatial features through its unique encoder–decoder architecture. The encoder part consists of multiple convolutional layers and max-pooling layers, which extract low-level features of the image by progressively reducing the spatial dimensions. Each convolutional layer performs convolution operations using convolutional kernels, followed by pooling layers that reduce the size of the feature

maps to retain the main information. The decoder part gradually restores the spatial dimensions of the image through up-sampling layers (such as deconvolution layers) and convolutional layers, while using skip connections to transfer high-resolution features from the encoder to the decoder, preserving detailed feature information. Ultimately, U-net can accurately segment and identify detailed features within urban spaces, making it particularly suitable for processing and fusing NTL data and Weibo sign-in data to identify the spatial extent of urban agglomerations and human activity patterns.

The equations of the component layers of U-net are as follows:

Layer convolution:

$$C_{out\,j} = bias(C_{out}) + \sum_{k=0}^{C_{in}-1} weight(C_{out\,j}, k) * input(k) \tag{1}$$

Layer max-pooling:

$$out(C_j, h, w) = max_{m=0,\ldots\ldots,kH-1} max_{n=0,\ldots\ldots,kW-1} input(C_j, stride[0] \times h + m, stride[1] \times w + n) \tag{2}$$

Layer ReLU:

$$ReLU(x) = max(0, x) \tag{3}$$

Layer softmax:

$$Softmax(x_i) = \frac{exp(x_i)}{\sum_{j=1}^{k} exp(x_i)} \tag{4}$$

Layer cross-entropy:

$$loss(x, y) = -log(\frac{exp(x[y])}{\sum_{j=1}^{k} exp(x[y])}) \tag{5}$$

In Equation (1), the sizes of the input and output images are (C_{in}, H, W) and (C_{out} t, H_{out}, W_{out}), C denotes the number of channels, H is the height of the input planes in pixels, W is the width in pixels, $*$ is the valid cross-correlation operator, and j is the j-th channel of the output feature map. In Equation (2), (kH, kW) denotes the kernel size of the pooling, and h and w refer to the height and width of the output image, respectively. In Equation (3), x denotes the pixel values of the input feature map. In Equation (4), x_i is the i-th pixel value of the input feature map, and K is the number of classes. In Equation (5), x and y refer to the predicted and reference pixel values, respectively, and K is the number of classes.

2.3.2. Wavelet Transform (WT)

Image fusion refers to the process of fusing image data about the same target, collected from multiple sources, through image processing to maximally extract beneficial information from each channel, thereby enhancing the spatial resolution of the original image and facilitating detection. As an excellent algorithm capable of fusing different data images [42,43], wavelet transform is a pixel-scale image fusion algorithm. The principle of wavelet transform is to magnify the local features of images based on the interrelation of time and frequency in different images. During the transformation process, through a dynamic "time–frequency" observation window, the characteristic parts of images are focused on and analyzed, achieving a unification of different images in time and frequency from decomposition to fusion [44]. The wavelet transform process for fusing NTL data with Weibo sign-in data includes two main steps: data preparation and wavelet transformation. First, the NTL data and Weibo sign-in data need to be preprocessed, which involves radiometric correction, the removal of outliers, deduplication, and the handling of missing values. Then, wavelet transformation is applied separately to both datasets, converting them to the frequency domain and performing multi-scale decomposition to obtain spatial information at different resolutions. Finally, the wavelet coefficients of the two datasets are fused using methods such as weighted averaging or rule-based fusion, resulting in fused

urban spatial data. This approach enables the more accurate and comprehensive identification of urban agglomeration spatial extents and human activity patterns. Therefore, with its perfect reconstruction ability, wavelet transform ensures that there is no information loss or redundant information during the fusion process. Moreover, by decomposing the image into relatively independent parts in time and frequency domains while retaining the original image's detail information, wavelet transform allows for the optimal observation effect after image fusion. The formula for wavelet transform is given as follows:

$$WT(\alpha, \tau) = f(t)\varphi(t) = \frac{1}{\sqrt{\alpha}} f(t) \int_{-\infty}^{+\infty} \varphi(\frac{t-b}{\alpha}) dt \quad (6)$$

where $f(t)$ is the image signal vector, $\varphi(t)$ is the wavelet transform function, α is the scale of wavelet transform, τ is the translation of the image signal, and b is a parameter.

2.3.3. Accuracy Verification

To verify the accuracy of urban agglomeration spatial ranges identified using NTL data and after fusion with Weibo Sign-in data, this study employs a method of verification using a confusion matrix of random pixel points [45]. The study validates the identified results by creating 5000 random pixel validation points within the spatial range of the PRD urban agglomeration, with a training-to-validation data ratio of 1:1. The formula for the confusion matrix is as follows:

$$k = \frac{p_o - p_e}{1 - p_e} \quad (7)$$

$$p_e = \frac{a_1 \times b_1 + a_2 \times b_2 + \ldots\ldots a_i \times b_i}{n \times n} \quad (8)$$

where p_o is the overall accuracy, a is the real sample number of each category, b is the predicted sample number of each category, and n is the total sample number. The Kappa coefficient ranges from -1 to 1, where -1 indicates total disagreement, 0 indicates random agreement, and 1 indicates perfect agreement. Typically, the closer the Kappa coefficient is to 1, the higher the consistency between the classification results and the actual situation and the greater the accuracy.

3. Results

3.1. Urban Agglomeration Spatial Identification Based on NTL Data

Typically, the distribution of NTL data within the spatial range of urban agglomerations exhibits certain characteristics. Firstly, high-value areas of NTL data are primarily concentrated in the central areas of urban agglomerations, such as core urban areas, economically developed zones, commercial districts, and transportation hubs. Secondly, as the distance from the center of the urban agglomeration increases, the distribution of NTL data becomes relatively dispersed and diffused. Therefore, NTL data can be used to identify the spatial range of urban agglomerations based on the variations in their luminance characteristics. In automated urban spatial extraction, U-net performs supervised learning by training on annotated urban spatial images. The model learns features related to urban agglomeration spaces, enabling it to segment image pixels into regions with urban agglomeration attributes or non-urban agglomeration attributes. Thus, the U-net-based automated method efficiently and accurately extracts urban agglomeration spaces, aiding researchers in analyzing and understanding the distribution and characteristics of these spaces.

We utilize the NTL data in conjunction with a U-net neural network to delineate the spatial range of the PRD urban agglomeration. The process begins with the establishment of training samples, where the sample labels are derived from the officially announced spatial boundaries of the urban agglomeration in previous years, ensuring a certain level of accuracy in the sample labels. Following this, the training samples are annotated and divided into test, training, and validation sets, culminating in the identification of the urban

spatial range of the PRD for the year 2022, as shown in Figure 4. By calculating the average NTL brightness values in different regions, it is found that the core area's average brightness is 75.4, while the peripheral area's is 42.8, with a brightness gradient decreasing by 2.3 units per kilometer, indicating a significant brightness attenuation. Moran's I index is 0.72, indicating significant spatial autocorrelation in the NTL data, and the Getis–Ord Gi* statistic identifies Guangzhou and Shenzhen as high-brightness hotspot areas. The identified spatial extent of the urban agglomeration covers an area of 8491.26 square kilometers. The average connectivity index between cities within the urban agglomeration is 0.78, and the average shortest path length is 45.6 km, demonstrating an efficient transportation network. The boundary clarity index is 0.45, and the diffusion coefficient in the central region of Huizhou is 1.2, indicating a clear outward diffusion trend in that area.

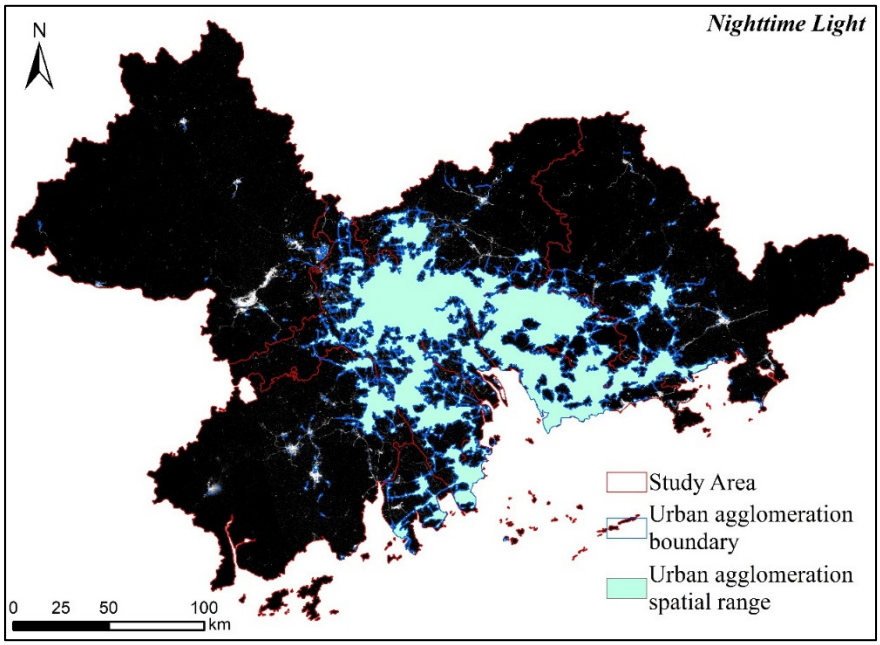

Figure 4. The urban agglomeration range of the PRD identified using NTL data.

3.2. Urban Agglomeration Spatial Identification Based on Weibo Sign-In Data

While NTL data reflect the disparities in economic development levels within urban agglomerations, focusing on macro-level spatial identification, Weibo sign-in data offer insights into urban agglomeration spaces from the perspectives of individual behavior and social interaction. Weibo sign-in data can reveal the patterns of population movement within urban agglomerations, aiding in the identification of highly active commercial districts, tourist attractions, transportation hubs, and their distribution within the urban agglomeration. Moreover, Weibo sign-in data can also uncover the interaction relationships between cities within the urban agglomeration. Drawing upon existing research on the fusion of NTL data with big data on urban agglomerations, we employ wavelet transform to fuse NTL data with social media data. The fused data exhibit a smaller spatial coverage compared to NTL data alone and display significant variations across different regions.

In the study of identifying the spatial extent of the PRD urban agglomeration using U-net neural networks combined with NTL data and Weibo data, training samples are labeled and divided into test, training, and validation sets. The final identification results for the urban agglomeration in 2022 are shown in Figure 5. The spatial extent identified through data fusion covers an area of 7993.08 square kilometers, 498.18 square kilometers less than the 8491.26 square kilometers identified using only NTL data, a reduction of 5.9%. The average brightness of the core area is 78.6, while that of the peripheral area is 40.3, with a brightness gradient decreasing by 2.6 units per kilometer. The Moran's I index after data fusion is 0.75, higher than the 0.72 obtained with only NTL data, indicating stronger

spatial autocorrelation. The Getis–Ord Gi* statistics show that high-brightness hotspots identified through data fusion are concentrated in Guangzhou, Shenzhen, and Dongguan, while low-brightness cold spots are located in Zhaoqing and Jiangmen. The spatial extent accuracy identified through data fusion is 99.1%, slightly lower than the 99.9% obtained using only NTL data, but the Kappa coefficient is 0.87, slightly higher than the 0.85 of the NTL data alone, indicating a higher consistency. The average connectivity index between cities within the urban agglomeration identified through data fusion is 0.82, higher than the 0.78 from NTL data alone, and the average shortest path length is 43.2 km, shorter than the 45.6 km identified using only NTL data, reflecting a higher connectivity efficiency. The boundary clarity index is 0.42, lower than the 0.45 from NTL data alone, indicating clearer boundaries. The diffusion coefficient for Huizhou and other areas is 1.3, higher than the 1.2 from NTL data alone, indicating more fine-grained brightness points and human activity features. Overall, the spatial extent identified through data fusion is more concentrated in the core area, reflecting a higher population density and more advanced urban infrastructure and services, making the characteristics of city group synergy and integration within the urban agglomeration more evident. These quantitative and spatial analyses demonstrate the advantages of the data fusion identification method in urban agglomeration spatial-extent identification, making the research results more persuasive and scientifically robust.

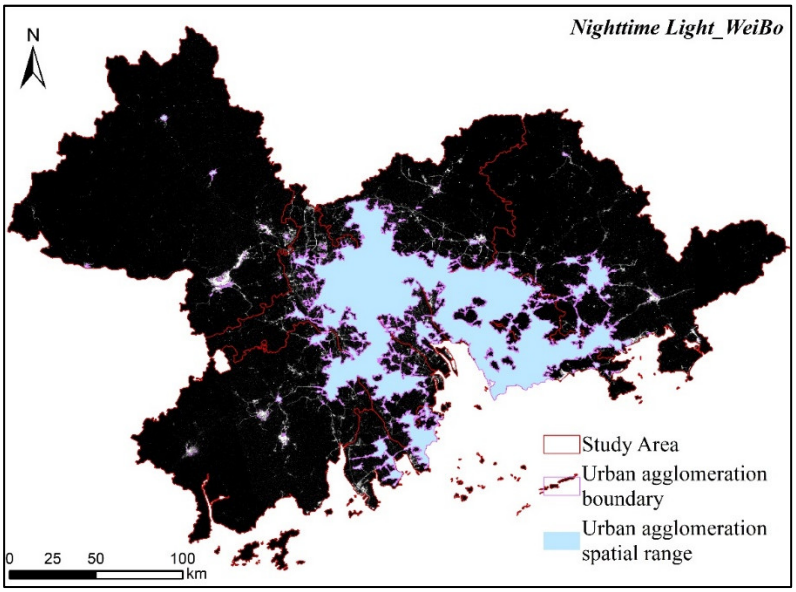

Figure 5. Fusion results of NTL data and social media data.

From the perspective of the spatial extent identified using the two data sources, there are significant differences between the urban agglomeration spaces identified through data fusion and those identified using only NTL data. Firstly, NTL data, due to their unique attribute of nighttime light brightness, determine the influence range of regions solely based on brightness values. In highly developed regions with closely spaced cities like the PRD, this method results in NTL data exhibiting a relatively concentrated spatial pattern, overlooking some finer internal differences within the urban agglomeration. Secondly, Weibo sign-in data reflect population data in different areas within the urban agglomeration but lack deeper socio-economic information, such as the purpose of user activities, satisfaction, or interactions with others, limiting a comprehensive understanding of the socio-economic dynamics of the urban agglomeration. In the spatial extent of the urban agglomeration identified through the fusion of Weibo sign-in data, Weibo data highlight the peripheral areas of cities. In contrast, the NTL–Weibo data fusion retains the characteristics of Weibo data while incorporating NTL data features, resulting in a diminished urban spatial extent near the main built-up areas of major cities within the urban agglomeration

and a strengthened urban development cluster between cities. This leads to a more fragmented urban agglomeration space identified through data fusion, reflecting the actual spatial situation of the urban agglomeration.

These differences indicate that the urban agglomeration space identified through data fusion is more detailed and fragmented, reflecting spatial heterogeneity. The data fusion method provides a richer and more detailed spatial extent of urban agglomerations, aiding researchers and decision makers in better understanding the distribution of human activities and the micro-patterns of urbanization within these areas, thereby effectively improving the accuracy of spatial identification. These differences demonstrate the heterogeneity in development levels among different cities within an urban agglomeration, reflecting the complex socio-economic dynamics and patterns of human activities. This enables the data fusion method to more comprehensively depict the spatial structure of urban agglomerations.

3.3. Accuracy Verification and Comparative Analysis

To verify and conduct a comparative analysis of the urban agglomeration spatial results identified before and after data fusion and to examine the differences between the results of this study and those of previous research, this study selects 5000 random pixel verification points. Through the aid of high-resolution imagery data from Google Earth, it is confirmed on-site that 1551 of these random verification points are located within the urban agglomeration spatial range, while 3449 are outside. The confusion matrix determined using the random pixel verification points is shown in Table 1. The accuracy is the percentage of all random pixel verification points that are successfully verified, while the Kappa coefficient is a measure of the consistency of the verification results. A Kappa coefficient closer to 1 indicates better verification results.

Table 1. Confusion matrix accuracy verification.

Data		Urban	Rural	Accuracy	Kappa
NTL	Urban	1153	398	85.38%	0.6468
	Rural	333	3116		
NTL_WB	Urban	1369	182	92.38%	0.8234
	Rural	199	3250		

As indicated in Table 1, the accuracy of urban agglomeration space identification using NTL data is 85.38%, with a Kappa coefficient of 0.6468. After data fusion, the accuracy of urban agglomeration space identification improves to 92.38%, with a Kappa coefficient of 0.8234. The validation results demonstrate that the accuracy of identifying the spatial range of urban agglomerations using a fusion of NTL data and social media data is enhanced by 7% compared to using NTL data alone. Furthermore, the Kappa coefficient increased by 0.1766, indicating that data fusion yields more accurate results in identifying the spatial range of urban agglomerations.

Comparing the spatial results of urban agglomerations identified using different datasets (as shown in Figure 6), there is a noticeable difference between the spatial ranges identified using NTL data and NTL_WB data. Consequently, we select four points with significant differences for the highlighted comparison and analysis. A detailed comparison of the identification results reveals that the urban agglomeration space identified using NTL data is larger due to its spill-over effect, and areas such as transportation routes, airports, and ports are included within the urban agglomeration space due to higher light-intensity values. However, after fusing Weibo sign-in data, the identified spatial range is smaller in these areas due to less population interaction, and the use of Weibo sign-in data is less frequent in suburban and urban edge areas, leading to a more fragmented identification of urban agglomeration spaces. Overall, the comparison of results identified using NTL data and NTL_WB data shows that data fusion identifies a smaller spatial range of urban agglomerations, with this trend being more evident in areas with fewer urban

clusters. This indicates that in the case of Weibo data, they tend to provide the identification of urban agglomeration spaces from the perspectives of individual behavior and social interaction rather than having clear boundaries like economic and land data. Thus, the results identified through data fusion are more capable of distinguishing the spatial ranges of urban agglomerations.

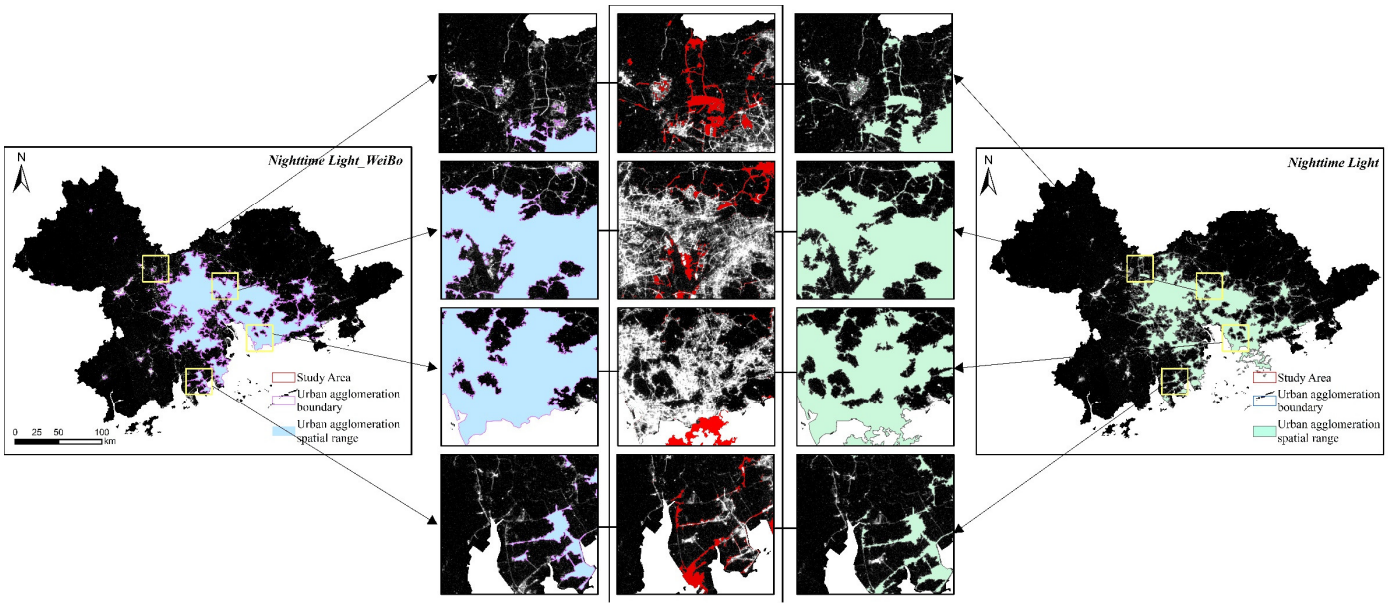

Figure 6. Comparative analysis of urban agglomeration space identification. (The middle column shows the difference in urban spaces identified using NTL_Weibo and NTL, represented by the result of NTL_Weibo minus NTL).

4. Discussion

This study, premised on the spatial differences within urban agglomerations, employs a fusion of Weibo data and NTL data, utilizing the U-net neural network to identify the spatial range and distribution characteristics of the PRD urban agglomeration. Furthermore, a comparative analysis is conducted between the urban agglomeration ranges identified after data fusion and those identified using solely NTL data.

The current identification of urban agglomeration spaces primarily utilizes data on land, population, and economy [46,47]. Such data, in earlier research, indeed facilitated the delineation of urban agglomeration boundaries and their specific spatial ranges, aiding in understanding the scale, population density, and level of economic activities of different urban agglomerations [48]. Analyzing the results of the urban agglomeration spaces identified in these studies, it is an undeniable fact that the spatial ranges of urban agglomerations are expanding. However, these datasets are typically collected by governmental agencies at fixed time points, implying that the data are not in real time. Given that urban agglomerations are dynamically changing, traditional data collection methods struggle to monitor these changes in real time. This increasingly complicates the task of clearly defining the spatial ranges of urban agglomerations with traditional research data and methods [49]. Thus, identifying urban agglomeration spaces through appropriate methods and approaches is evidently crucial for coordinated development within urban agglomerations and for achieving their sustainable development [50]. Current research on urban spatial analysis using NTL data is becoming increasingly abundant, including NPP, DMSP, and Luojia-01. DMSP data, with their longer time series, are suitable for historical studies and trend analysis, although they have lower spatial resolution. In contrast, Luojia-01 data offer a higher spatial resolution, providing more detailed urban brightness information. This study, employing a fusion of social media data and NTL data for identifying urban agglomeration spaces, demonstrates that the accuracy of urban agglomeration space

identification reached 92.38%, an improvement of 7% over the accuracy achieved with NTL data alone. The Kappa coefficient also increased by 0.1766, significantly enhancing the precision in identifying the spatial ranges of urban agglomerations. By fusing other data sources, the U-net neural network model can be trained separately using DMSP and Luojia-01 data to analyze and validate the correctness of this study's results. A comprehensive comparison of the identification results from different data sources shows that the characteristics identified using DMSP and Luojia-01 data, such as urban agglomeration spatial area, spatial range distribution, brightness distribution, and spatial autocorrelation, reveal differences in the detail capture and spatial heterogeneity. Compared with current research results, the data fusion method can more comprehensively depict the spatial structures of urban agglomerations, providing richer and more detailed spatial extents. This helps researchers and decision makers better understand the distribution of human activities and the micro-patterns of urbanization within urban agglomerations, thereby effectively improving the accuracy of spatial identification.

From our analysis of features following the fusion of NTL and WB data, several key observations emerge. Firstly, social media data provide a more comprehensive array of information, including population distribution, activity hotspots, and community interactions, which can enrich the understanding of the spatial differences and characteristics of urban agglomerations [51]. Secondly, while NTL data reflect urban construction and economic activity levels, social media data unveil details about population distribution and activities. The fusion of these datasets allows for the integration of social media and NTL data, thereby yielding more accurate and comprehensive results in the identification of urban agglomeration spaces [52]. By leveraging these diverse types of data, we can more effectively identify the spatial range of urban agglomerations, circumventing some of the difficulties and errors inherent in traditional methods [53]. In summary, the method of fusing social media data enables the acquisition of more comprehensive and precise results in identifying urban agglomeration spaces. By combining the unique features and strengths of different datasets, we can better capture the dynamics and detailed spatial information on urban agglomerations. This approach yields the accurate identification of urban agglomeration spaces, providing important references and guidance for urban planning, decision making, and development.

Although the study of identifying the spatial range of urban agglomerations is not a brand new topic, many studies have analyzed the identification and delineation of the spatial range of urban agglomerations in different urban agglomerations in China [54]. However, this study proposes a new way of fusing NTL data with social media data to accurately identify the spatial range of urban agglomerations based on the work of predecessors by conducting a detailed analysis of the PRD urban agglomeration. This introduces a fresh perspective and solution to the study of urban agglomeration spaces, offering a straightforward and widely applicable method that holds significant practical value and prospects for application.

5. Conclusions

The identification of urban agglomeration spatial ranges has always been foundational to urban agglomeration studies. A clearer delineation of urban agglomeration spaces can facilitate the optimization of resource allocation, thereby promoting regional balanced development. This study, based on the group differences within urban agglomeration spaces as reflected by NTL data and Weibo sign-in data, adopts a data fusion approach to integrate NTL and Weibo data. It employs a U-net neural network for the accurate identification of urban agglomeration spaces, with validation through a random pixel confusion matrix. This study develops a method for identifying urban agglomeration spaces by fusing NTL data with Weibo sign-in data, improving the accuracy of identified urban agglomeration spaces by 7% and increasing the Kappa coefficient by 0.1766 compared to using NTL data alone. Furthermore, a comparison of identification characteristics before and after data fusion reveals a significant enhancement in the results of urban

agglomeration space identification through data fusion compared to those identified using NTL data only. The accurate delineation of the PRD urban agglomeration's spatial range through data fusion has significant practical implications for coordinated development within the urban agglomeration. Moreover, it provides a scalable case study for other urban agglomeration research, offering feasible decision-making support for urban agglomeration planning and development.

This study employs data fusion techniques to identify and analyze the spatial range of the PRD urban agglomeration, yet there remain some limitations. Firstly, in terms of the study data, Weibo sign-in data exhibit an imbalance in user activity, with younger individuals more inclined to use Weibo sign-ins compared to older adults and children, leading to data bias issues. Additionally, Weibo data involve concerns regarding user privacy and data security, necessitating anonymization and protection measures [55]. Secondly, the data fusion process requires the consideration of consistency and reliability between different data sources to ensure the accuracy and credibility of the fusion results. The third limitation is the analysis's reliance on a single-time snapshot, which could introduce temporal anomalies. Therefore, future work will include a comparative analysis over a longer time series to further delineate the advantages and characteristics of combining NTL data with Weibo data.

Author Contributions: Conceptualization, Yongwang Cao, Song Liu, and Zaigao Yang; methodology, Yongwang Cao; software, Yongwang Cao and Song Liu; validation, Yongwang Cao; formal analysis, Yongwang Cao and Zaigao Yang; resources, Yongwang Cao and Song Liu; data curation, Yongwang Cao; writing—original draft preparation, Yongwang Cao and Zaigao Yang; visualization, Yongwang Cao; funding acquisition, Song Liu and Zaigao Yang. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Municipal Social Science Foundation of Guangzhou, grant numbers 2023GZGJ27 and 24QN004.

Data Availability Statement: Data are contained within the article.

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

References

1. Fang, C.; Yu, D. Urban agglomeration: An evolving concept of an emerging phenomenon. *Landsc. Urban Plan.* **2017**, *162*, 126–136. [CrossRef]
2. Tan, X.; Huang, B. Identifying Urban Agglomerations in China Based on Density-Density Correlation Functions. *Ann. Am. Assoc. Geogr.* **2022**, *112*, 1666–1684. [CrossRef]
3. Huang, Y.; Hong, T.; Ma, T. Urban network externalities, agglomeration economies and urban economic growth. *Cities* **2020**, *107*, 102882. [CrossRef]
4. Ruan, L.; He, T.; Xiao, W.; Chen, W.; Lu, D.; Liu, S. Measuring the coupling of built-up land intensity and use efficiency: An example of the Yangtze River Delta urban agglomeration. *Sustain. Cities Soc.* **2022**, *87*, 104224. [CrossRef]
5. Wang, W.; Li, B.; Su, F.; Jiang, Z.; Chen, S. Identifying Ecological Security Patterns Meeting Future Urban Expansion in Changsha-Zhuzhou-Xiangtan Urban Agglomeration, China. *Remote Sens.* **2023**, *15*, 3141. [CrossRef]
6. Li, S.; Guo, H.; Sun, Z.; Liu, Z.; Jiang, H.; Zhang, H. Patterns of Typical Chinese Urban Agglomerations Based on Complex Spatial Network Analysis. *Remote Sens.* **2023**, *15*, 920. [CrossRef]
7. Liu, Y.; Zhang, X.; Pan, X.; Ma, X.; Tang, M. The spatial integration and coordinated industrial development of urban agglomerations in the Yangtze River Economic Belt, China. *Cities* **2020**, *104*, 102801. [CrossRef]
8. Li, L.; Ma, S.; Zheng, Y.; Xiao, X. Integrated regional development: Comparison of urban agglomeration policies in China. *Land Use Policy* **2022**, *114*, 105939. [CrossRef]
9. Cottineau, C.; Finance, O.; Hatna, E.; Arcaute, E.; Batty, M. Defining urban clusters to detect agglomeration economies. *Environ. Plan. B Urban Anal. City Sci.* **2019**, *46*, 1611–1626. [CrossRef]
10. Wu, B.; Yang, C.; Wu, Q.; Wang, C.; Wu, J.; Yu, B. A building volume adjusted nighttime light index for characterizing the relationship between urban population and nighttime light intensity. *Comput. Environ. Urban Syst.* **2023**, *99*, 101911. [CrossRef]
11. Sharma, N.; Shastri, S.; Shastri, S. Does urbanization level and types of urban settlements matter for child stunting prevalence in India? Empirical evidence based on nighttime lights data. *Cities* **2023**, *140*, 104388. [CrossRef]
12. Peng, J.; Lin, H.; Chen, Y.; Blaschke, T.; Luo, L.; Xu, Z.; Hu, Y.n.; Zhao, M.; Wu, J. Spatiotemporal evolution of urban agglomerations in China during 2000–2012: A nighttime light approach. *Landsc. Ecol.* **2020**, *35*, 421–434. [CrossRef]

13. 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]
14. 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]
15. 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]
16. Priyashani, N.; Kankanamge, N.; Yigitcanlar, T. Multisource Open Geospatial Big Data Fusion: Application of the Method to Demarcate Urban Agglomeration Footprints. *Land* **2023**, *12*, 407. [CrossRef]
17. Anejionu, O.C.D.; Thakuriah, P.; McHugh, A.; Sun, Y.; McArthur, D.; Mason, P.; Walpole, R. Spatial urban data system: A cloud-enabled big data infrastructure for social and economic urban analytics. *Future Gener. Comput. Syst.* **2019**, *98*, 456–473. [CrossRef]
18. Hong, A.; Baker, L.; Prieto, C.R.; Duminy, J.; Buswala, B.; Guan, C.; Ravindranath, D. Reconciling big data and thick data to advance the new urban science and smart city governance. *J. Urban Aff.* **2023**, *45*, 1737–1761. [CrossRef]
19. Zhen, F.; Cao, Y.; Qin, X.; Wang, B. Delineation of an urban agglomeration boundary based on Sina Weibo microblog 'check-in' data: A case study of the Yangtze River Delta. *Cities* **2017**, *60*, 180–191. [CrossRef]
20. 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]
21. Zhang, W.; Derudder, B.; Wang, J.; Shen, W.; Witlox, F. Using Location-Based Social Media to Chart the Patterns of People Moving between Cities: The Case of Weibo-Users in the Yangtze River Delta. *J. Urban Technol.* **2016**, *23*, 91–111. [CrossRef]
22. Wang, Y.; Fu, X.; Jiang, W.; Wang, T.; Tsou, M.; Ye, X. Inferring urban air quality based on social media. *Comput. Environ. Urban Syst.* **2016**, *66*, 110–116. [CrossRef]
23. Ye, Z.; Newing, A.; Clarke, G. Understanding Chinese tourist mobility and consumption-related behaviours in London using Sina Weibo check-ins. *Environ. Plan. B Urban Anal. City Sci.* **2021**, *48*, 2436–2452. [CrossRef]
24. Bagan, H.; Borjigin, H.; Yamagata, Y. Assessing nighttime lights for mapping the urban areas of 50 cities across the globe. *Environ. Plan. B Urban Anal. City Sci.* **2019**, *46*, 1097–1114. [CrossRef]
25. Fragkias, M.; Lobo, J.; Seto, K.C. A comparison of nighttime lights data for urban energy research: Insights from scaling analysis in the US system of cities. *Environ. Plan. B Urban Anal. City Sci.* **2017**, *44*, 1077–1096. [CrossRef]
26. Huang, Y.; Li, J.; Wu, G.; Fei, T. Quantifying the bias in place emotion extracted from photos on social networking sites: A case study on a university campus. *Cities* **2020**, *102*, 102719. [CrossRef]
27. Peng, Z.; Wang, R.; Liu, L.; Wu, H. Exploring Urban Spatial Features of COVID-19 Transmission in Wuhan Based on Social Media Data. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 402. [CrossRef]
28. Zhang, X.; Sun, Y.; Zheng, A.; Wang, Y. A New Approach to Refining Land Use Types: Predicting Point-of-Interest Categories Using Weibo Check-in Data. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 124. [CrossRef]
29. Yang, X.; Gao, J.; Zheng, X.; Fang, M.; Tang, L.; Zhang, X. Point-of-interest detection from Weibo data for map updating. *Trans. GIS* **2022**, *26*, 2716–2738. [CrossRef]
30. Liu, X.; Song, Y.; Wu, K.; Wang, J.; Li, D.; Long, Y. Understanding urban China with open data. *Cities* **2015**, *47*, 53–61. [CrossRef]
31. Myint, S.W.; Zhu, T.; Zheng, B. A Novel Image Classification Algorithm Using Overcomplete Wavelet Transforms. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 1232–1236. [CrossRef]
32. Peng, J.; Hu, Y.n.; Liu, Y.; Ma, J.; Zhao, S. A new approach for urban-rural fringe identification: Integrating impervious surface area and spatial continuous wavelet transform. *Landsc. Urban Plan.* **2018**, *175*, 72–79. [CrossRef]
33. 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]
34. Zhou, Y.; He, X.; Zhu, Y. Identification and evaluation of the polycentric urban structure: An empirical analysis based on multi-source big data fusion. *Remote Sens.* **2022**, *14*, 2705. [CrossRef]
35. Drăguţ, L.; Csillik, O.; Eisank, C.; Tiede, D. Automated parameterisation for multi-scale image segmentation on multiple layers. *ISPRS J. Photogramm. Remote Sens.* **2014**, *88*, 119–127. [CrossRef]
36. 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]
37. Deng, H.; Li, H. Identification of Urban Spatial Structure of Pearl River Delta Urban Agglomeration Based on Multisource Spatial Data. *J. Urban Plan. Dev.* **2023**, *149*, 05023010. [CrossRef]
38. Miller, S.; Straka, W.; Mills, S.; Elvidge, C.; Lee, T.; Solbrig, J.; Walther, A.; Heidinger, A.; Weiss, S. Illuminating the Capabilities of the Suomi National Polar-Orbiting Partnership (NPP) Visible Infrared Imaging Radiometer Suite (VIIRS) Day/Night Band. *Remote Sens.* **2013**, *5*, 6717–6766. [CrossRef]
39. Chen, B.; Zhu, Y.; He, X.; Zhou, C. Analysis of High-Quality Tourism Destinations Based on Spatiotemporal Big Data—A Case Study of Urumqi. *Land* **2023**, *12*, 1425. [CrossRef]
40. Delibasoglu, I.; Cetin, M. Improved U-Nets with inception blocks for building detection. *J. Appl. Remote Sens.* **2020**, *14*, 044512. [CrossRef]
41. Sultonov, F.; Park, J.; Yun, S.; Lim, D.; Kang, J. Mixer U-Net: An Improved Automatic Road Extraction from UAV Imagery. *Appl. Sci.* **2022**, *12*, 1953. [CrossRef]

42. 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]
43. Cai, L.; Wang, H.; Sha, C.; Jiang, F.; Zhou, W. The Mining of Urban Hotspots Based on Multi-source Location Data Fusion. *IEEE Trans. Knowl. Data Eng.* **2021**, *35*, 2061–2077. [CrossRef]
44. Chen, Y.; Deng, A. Using POI Data and Baidu Migration Big Data to Modify Nighttime Light Data to Identify Urban and Rural Area. *IEEE Access* **2022**, *10*, 93513–93524. [CrossRef]
45. Silván-Cárdenas, J.L.; Wang, L. Sub-pixel confusion–uncertainty matrix for assessing soft classifications. *Remote Sens. Environ.* **2007**, *112*, 1081–1095. [CrossRef]
46. Han, H.; Li, H. Coupling Coordination Evaluation between Population and Land Urbanization in Ha-Chang Urban Agglomeration. *Sustainability* **2020**, *12*, 357. [CrossRef]
47. Melo, P.C.; Graham, D.J.; Noland, R.B. A meta-analysis of estimates of urban agglomeration economies. *Reg. Sci. Urban Econ.* **2008**, *39*, 332–342. [CrossRef]
48. Huang, Y.; Liao, R. Polycentric or monocentric, which kind of spatial structure is better for promoting the green economy? Evidence from Chinese urban agglomerations. *Environ. Sci. Pollut. Res. Int.* **2021**, *28*, 57706–57722. [CrossRef]
49. Pinch, S.; Sunley, P. Do Urban Social Enterprises Benefit from Agglomeration? Evidence from Four UK Cities. *Reg. Stud.* **2015**, *50*, 1290–1301. [CrossRef]
50. Kong, L.; Wu, Q.; Deng, J.; Bai, L.; Chen, Z.; Du, Z.; Luo, M. Assessing Regional Development Balance Based on Zipf's Law: The Case of Chinese Urban Agglomerations. *ISPRS Int. J. Geo-Inf.* **2023**, *12*, 472. [CrossRef]
51. Tian, X.; Batterham, P.; Song, S.; Yao, X.; Yu, G. Characterizing Depression Issues on Sina Weibo. *Int. J. Environ. Res. Public Health* **2018**, *15*, 764. [CrossRef]
52. Cai, J.; Huang, B.; Song, Y. Using multi-source geospatial big data to identify the structure of polycentric cities. *Remote Sens. Environ.* **2017**, *202*, 210–221. [CrossRef]
53. Zhan, Y.; Sui, L.; Wang, M.; Huang, J.; Zhu, J.; Chen, D.; Fan, J. Multiperspective Evaluation Model for the Spatial Distribution of Public Service Facilities Based on Service Capability and Subjective Preferences. *J. Urban Plan. Dev.* **2022**, *148*, 04022015. [CrossRef]
54. Huang, L.; Wang, J.; Cheng, H. Spatiotemporal changes in ecological network resilience in the Shandong Peninsula urban agglomeration. *J. Clean. Prod.* **2022**, *339*, 130681. [CrossRef]
55. Kong, L.; Liu, Z.; Pan, X.; Wang, Y.; Guo, X.; Wu, J. How do different types and landscape attributes of urban parks affect visitors' positive emotions? *Landsc. Urban Plan.* **2022**, *226*, 104482. [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.