Article

An Approach of Identifying and Extracting Urban Commercial Areas Using the Nighttime Lights Satellite Imagery

Xuzhe Duan, **Qingwu Hu** *, **Pengcheng Zhao**, **Shaohua Wang and Mingyao Ai**

School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430070, China; duanxz@whu.edu.cn (X.D.); pengcheng.zhao@whu.edu.cn (P.Z.); shwang@whu.edu.cn (S.W.); aimingyao@whu.edu.cn (M.A.)
* Correspondence: huqw@whu.edu.cn

check for updates

Abstract: Urban commercial areas can reflect the spatial distribution of business activities. However, the scope of urban commercial areas cannot be easily detected by traditional methods because of difficulties in data collection. Considering the positive correlation between business scale and nighttime lighting, this paper proposes a method of urban commercial areas detection based on nighttime lights satellite imagery. First, an imagery preprocess model is proposed to correct imageries and improve efficiency of cluster analysis. Then, an exploratory spatial data analysis and hotspots clustering method is employed to detect commercial areas by geographic distribution metric with urban commercial hotspots. Furthermore, four imageries of Wuhan City and Shenyang City are selected as an example for urban commercial areas detection experiments. Finally, a comparison is made to find out the time and space factors that affect the detection results of the commercial areas. By comparing the results with the existing map data, we are convinced that the nighttime lights satellite imagery can effectively detect the urban commercial areas. The time of image acquisition and the vegetation coverage in the area are two important factors affecting the detection effect. Harsh weather conditions and high vegetation coverage are conducive to the effective implementation of this method. This approach can be integrated with traditional methods to form a fast commercial areas detection model, which can then play a role in large-scale socio-economic surveys and dynamic detection of commercial areas evolution. Hence, a conclusion can be reached that this study provides a new method for the perception of urban socio-economic activities.

Keywords: urban commercial area; nighttime lights satellite imagery; exploratory spatial data analysis; urban hot spots detection

1. Introduction

Commercial areas in a city are areas, districts, or neighborhoods primarily composed of commercial buildings, such as a downtown, central business district, financial district, "Main Street", commercial strip, or shopping center. Fifth Avenue in Manhattan, Champs Elysees in Paris, and Causeway Bay in Hong Kong can all be considered as typical commercial areas. The scale of urban commercial areas is an objective reflection of a city's economic development and an important part of its comprehensive competitiveness. Rapid and accurate extraction of urban commercial areas can provide an important basis for guiding urban economic layout, give full play to the social benefits and overall functions of commercial areas, accelerate the urbanization process, and maintain the sustainable development of the national economy [1]. To be specific, detecting the commercial areas accurately can help the government to plan the distribution of city functions and effectively promote the diversified development of the

city. It can help individual investors select the best position for business operation. It can also provide good research samples for sociological research departments. However, the existing commercial area extraction approach has many defects and shortcomings, which make it difficult to bring a commercial area's superiority into full play. This paper aims to explore an effective approach of identifying and extracting urban commercial areas by using the nighttime lights satellite imagery.

The current research on commercial areas detection mainly includes three methods: the survey method, the reference method, and the volunteered geographic information (VGI) method [2]. The survey method is the most direct way to determine the scope of the commercial area. This method determines the commercial area by making sample consumers fill out questionnaires and obtaining spatial connections between their homes and the business centers that they regularly shop at [3]. As the amount of sample data are large enough, the credibility of the results is relatively high. However, due to the large amount of manpower, material resources, and time cost, the feasibility of this method is low. The reference method determines the commercial area by referring to the existing commercial area of a certain kind of market or region [4]. In order to get as close as possible to the actual situation of the area where the unknown commercial area is located, the model can be modified according to the differences in the operation scale of the market, as well as the differences in resident population distribution, urban construction, traffic facilities, and other aspects. For instance, Reilly (1931) proposed the law of retail gravitation, which stated that the number of customers that a city could attract from the surrounding cities was directly proportional to the city's population and was inversely proportional to the square of the distance between the two cities [5]. Westbrook and Black (1985) proposed a multifactor model for the determination of commercial areas based on the factors that are difficult to quantify, such as goodwill and opportunity costs [6]. Ingene and Ghosh (1990) and McLafferty and Ghosh (2010) proposed three methods to estimate commercial areas, including the spatial proportion approach, the central region law, and the space sales law [7,8]. The reference method has high data availability and does not need a lot of labor costs, so it is a feasible method for the commercial area detection. However, the selection of the method variables and the construction of the model must fully refer to the reality, thus the test of goodness for fit is needed to ensure the credibility of the model and its results [2]. The VGI method uses the geographic information that is provided voluntarily by individuals to detect the commercial area [9,10]. The main sources of VGI data include car GPS tracks, private mobile phone signaling and social network check-in, and semantic information. For example, based on the "check-in" information with the user's location attributes, researchers performed exploratory spatial data analysis and spatial clustering analysis on VGI, and then obtained the location and scope of the commercial area [10–13]. However, the data owners often do not provide users' location information in public since the data involves personal privacy. Therefore, this method is limited by the acquisition of data, and its research is still scarce [2].

In view of the defects and deficiencies of the existing detection methods for commercial areas, this paper attempts to propose a more efficient and accurate approach to identify and extract commercial areas by using the nighttime lights satellite data. There were two kinds of nighttime lights satellite data sources that used to be applied in scientific research: Defense Meteorological Satellite Program/the Operational Linescan System (DMSP/OLS) and National Polar-orbiting Partnership/Visible Infrared Imaging Radiometer Suite (NPP/VIIRS). The spatial resolution of the former is usually about 2700 m whereas the latter is higher and can reach 740 m [14]. Even so, such resolution is not sufficient to provide adequate support for the detection and extraction of urban commercial areas. With the launch of Luojia No.1 01 nighttime lights satellite in June 2018, huge breakthroughs have been made in the spatial resolution—using complementary metal-oxide-semiconductor (CMOS) sensor, its spatial resolution can reach 130 m. With a width of 250 km × 250 km, it can map a complete global luminous image within 15 days under ideal conditions [15]. These characteristics not only meet the accuracy requirements of urban commercial areas research but also provide strong support for the dynamic detection and extraction of urban commercial areas updates.

Different from daytime remote sensing, nighttime light remote sensing provides a unique perspective on human social activities, thus it has been widely used for spatial data mining of socioeconomic domains [14]. The nighttime lights satellite can observe the night lights from the perspective of the whole city, and the brightness of the imagery reflects the illumination degree of the corresponding area at night. The area with strong night lighting is usually accompanied by large population distribution, and it is more likely to be the region with concentrated commercial development, which is what we call the "urban commercial area". With that guideline, from Luojia No.1 nighttime lights satellite imageries, we select four good imageries with obvious light features, few noise points, and high image quality as our research data source. Through data preprocessing, exploratory spatial data analysis (ESDA), and spatial clustering analysis, we can obtain high-value clustering hotspots of imagery brightness values, and by building standard ellipses on hotspots of clusters, we can obtain the general spatial distribution of urban commercial areas. Furthermore, we take the result of ESDA as a basis for evaluation, comparing the results of commercial areas detection "at different periods" and "in different regions" to analyze the effectiveness of using nighttime lights satellite imagery for commercial areas detection.

2. Materials and Methods

Firstly, for the imageries chosen to be the dataset, linear transformation of brightness values is required to make each pixel's brightness value within the range of 0–255 for the convenience of subsequent processing. Secondly, exploratory spatial data analysis is used to measure the spatial correlation, the spatial structure, and the global distribution pattern of each imagery's brightness values and to determine the optimal pattern of hotspot detection and commercial area clustering. Thirdly, cluster analysis is applied to identify hot spots, cold spots, and spatial outliers with statistical significance in each imagery. Finally, by measuring the geospatial distribution of clustering analysis results, the spatial characteristics of statistically significant clustering geographic elements (commercial areas' range, central change trend, and direction development trend) are obtained. The specific algorithm flowchart is shown in Figure 1.
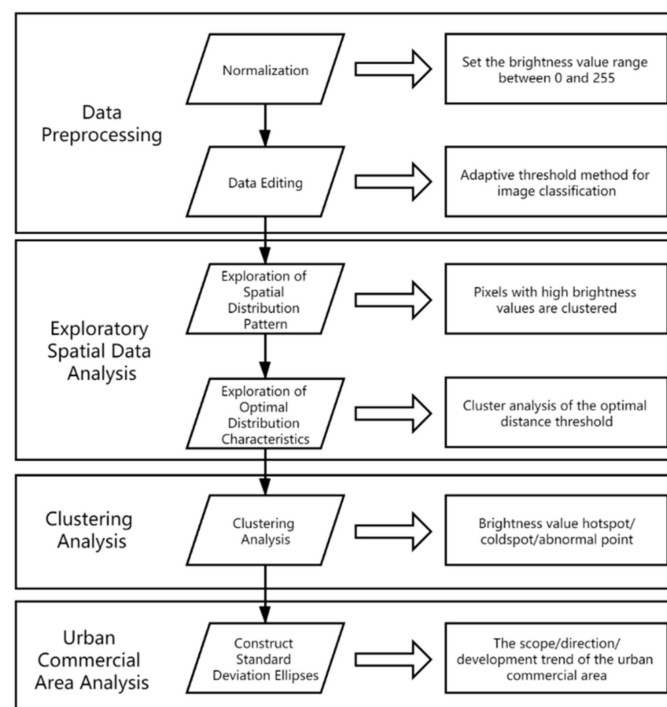
Figure 1. The algorithm flowchart for the urban commercial area identification and extraction based on the nighttime lights satellite imagery.

2.1. Data Preprocessing

The data preprocessing include the cropping of imageries, the normalization of brightness values, and the editing of imagery data.

Firstly, this paper uses the vector layer of the study area to crop the imageries, so as to remove the water areas in the city and the parts that are outside the study area.

Next, in order to unify the brightness values of the three imageries into an identical range, this paper uses the method of normalizing each imagery and stretching the brightness values linearly to the range of 0–255 for the next step of imagery processing, as shown in Equation (1)

$$G_t(i, j) = \left\lfloor 255 \times \frac{G(i, j) - G_{min}}{G_{max} - G_{min}} \right\rfloor,$$ (1)

where $G(i, j)$ represents the brightness value of location (i, j) of the original imagery; G_{min} refers to the minimum value of brightness in the original imagery; G_{max} represents the maximum; $G_t(i, j)$ represents the brightness value of location (i, j) after linear transformation; and the $\lfloor\ \rfloor$ symbol refers to the rounding operation of the internal value to ensure that the brightness value obtained is an integer.

Last, we shave the "exception pixels" of the imagery. In nighttime light imagery, there are many raster pixels with low brightness values or even close to 0. These raster pixels correspond to areas without light in the imagery and should not participate in ESDA and cluster analysis. However, these pixels will inevitably be calculated in the following processing. Therefore, we call these pixels "exception pixels". This paper uses the OTSU binarization method proposed by the Japanese scholar OTSU to divide the image into two parts, with "exception pixels" marked as 0 and other pixels marked as 1. The areas with the segmentation result of 1 are used as the study regions, whereas the ones of 0 as the nonstudy regions.

The OTSU algorithm is a nonparametric and unsupervised method of automatic threshold selection for picture segmentation that assumes that image pixels can be divided into the background part and the foreground part according to the threshold [16]. The algorithm exhaustively searches for the threshold that minimizes the intraclass variance, defined as a weighted sum of variances of the two classes, as shown in Equation (2)

$$\sigma_\omega^2(t) = w_0(t)\sigma_0^2(t) + w_1(t)\sigma_1^2(t),$$ (2)

where weights w_0 and w_1 are the probabilities of the two classes separated by a threshold t, and σ_0^2 and σ_1^2 are variances of these two classes, respectively. The class probability $w_{0,1}(t)$ is computed from the $L (L = 256)$ bins of the histogram: $w_0(t) = \sum_{i=0}^{t-1} p(i)$ and $w_1(t) = \sum_{i=t}^{L-1} p(i)$.

For two classes, minimizing the intraclass variance is equivalent to maximizing interclass variance, as shown in Equation (3)

$$\sigma_b^2(t) = w_0(t)w_1(t)[\mu_0(t) - \mu_1(t)]^2,$$ (3)

which is expressed in terms of class probabilities ω and class means μ, where the class means $\mu_0(t)$ and $\mu_1(t)$ are: $\mu_0(t) = \sum_{i=0}^{t-1} ip(i)/w_0(t)$ and $\mu_1(t) = \sum_{i=t}^{L-1} ip(i)/w_1(t)$.

The larger the threshold value, the larger the proportion of the background in the segmentation result, which will cause the brightness value of the image in some commercial areas to be assigned a value of 0. On the contrary, the smaller the threshold, the larger the proportion of foreground in the segmentation result, which will lead to poor clustering effect of image brightness value and make it difficult to detect the center of commercial areas. Taking the imagery of the area inside the third ring road of Wuhan on 14 June 2018 as an example, the threshold obtained by the OTSU algorithm is 16, and the binarization result is shown in Figure 2.
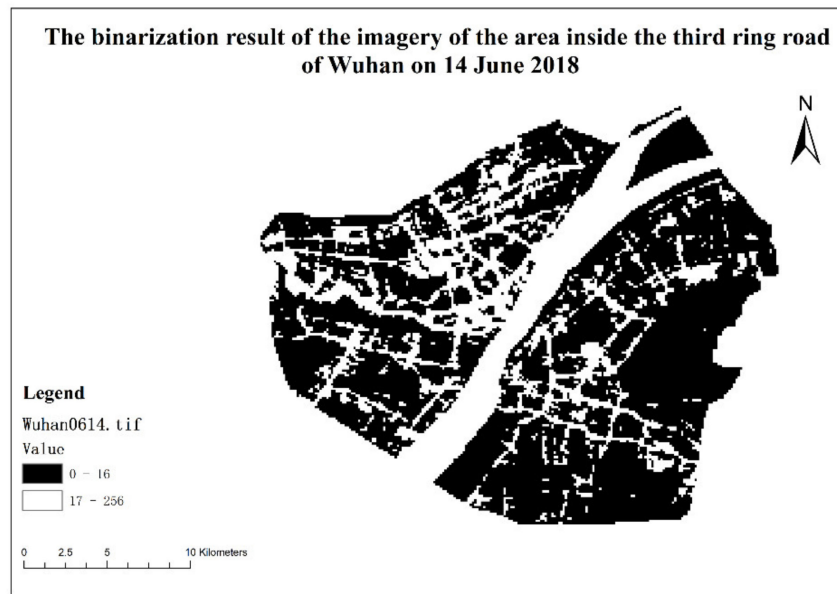
Figure 2. The binarization result of the imagery of the area inside the third ring road of Wuhan on 14 June 2018.

2.2. Exploratory Spatial Data Analysis

The spatial autocorrelation refers to the potential interdependence between observations of some variables in the same distribution area. It is an indicator used to measure the degree of interdependence between the data at a certain location and the data at other locations [17–20]. In this paper, the Global Moran's I statistical method is adopted to measure the global spatial autocorrelation of brightness values in the study area. The Ripley's K statistical method is used to explore the most characteristic spatial distribution pattern of brightness values.

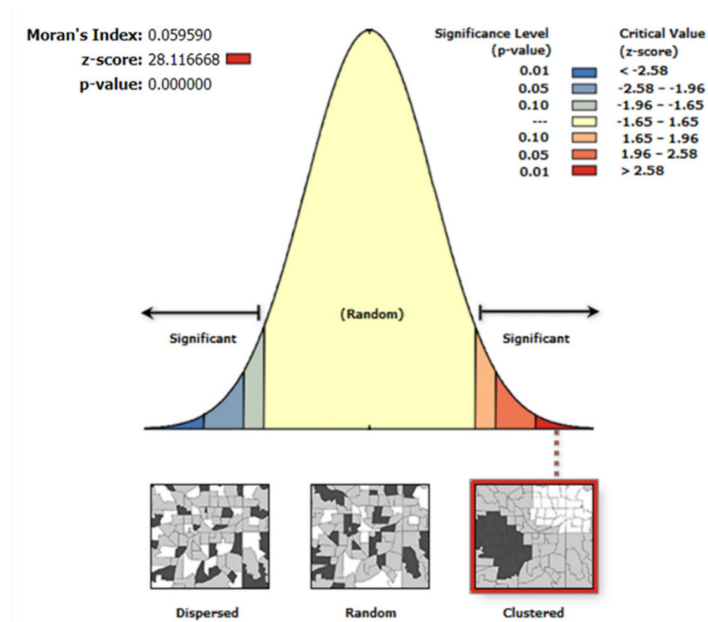### 2.2.1. Exploration of Spatial Distribution Pattern

The Moran's I statistical method is an algorithm that measures spatial autocorrelations based on both the locations and values of elements. A Moran's I that is higher than 0 reflects positive spatial correlations. With the index growth, the spatial correlations will also become more evident. Instead, a Moran's I below 0 reveals negative spatial correlations. The lower the index, the greater the spatial differences are. Positive spatial correlations mean that the data are spatially clustered, whereas negative spatial correlations indicate that the data are spatially discrete. In this paper, the pixels with an OTSU result of 1 in study imagery are used as the datasets to analyze the spatial distribution patterns. Global Moran's I statistics are evaluated according to Equation (4)

$$\mathrm{I} = \left[\mathrm{n} \sum\nolimits_{i=1}^{n} \sum\nolimits_{j=1}^{n} \omega_{i,j} z_i z_j \right] / \left[S_0 \sum\nolimits_{i=1}^{n} z_i^2 \right], \tag{4}$$

where n is total number of pixels of the image after removing the 0 values of the OTSU result, z_i is the deviation between the brightness value of the pixel at location i and its mean, $\omega_{i,j}$ is the spatial weight between the pixels i and j, and S_0 is the sum of all spatial weights where the expression is $S_0 = \sum_{i=1}^{n} \sum_{j=1}^{n} \omega_{i,j}$. The scores (z_1) are determined by $z_1 = (I - EI) / \sqrt{VI}$, where $EI = -1/(n-1)$, $V[I] = E[I^2] - EI^2$.

We use the spatial autocorrelation tool in ArcGIS 10.1 software (ArcGIS 10.1, 2012) to process imagery brightness values. For conceptualization of spatial relationships, we use INVERSE_DISTANCE, for the distance method, we use EUCLIDEAN_DISTANCE, and for standardization, we select NONE. Figure 3 is the spatial distribution patterns of brightness values calculated based on Global Moran's I statistics on the Wuhan imagery of 14 June 2018. It may be observed that the z-score value used

for the test is 28.116668 times the standard deviation and all significantly above 2.58, which means that the *p* value of its null hypothesis (the likelihood value) is below 0.01, and the requirement for *p* value is met at the confidence level of 99%. This indicates that the global spatial autocorrelation of brightness values distribution pattern conforms to the statistical characteristics of typical clustering models. Therefore, cluster analysis may be performed on urban commercial areas.

Figure 3. Results of global spatial autocorrelation analysis of the June imagery of Wuhan (generated by ArcMap 10.1).

2.2.2. Exploration of Optimal Distribution Characteristics

The spatial distance is the most important concept for spatial analysis, and varying results can be obtained when various analyses are performed on different spatial distances [21]. In this paper, Ripley's K function is used to explore the most statistically significant spatial autocorrelation of brightness values, as shown in Equation (5)

$$\mathrm{L}(\mathrm{d}) = \sqrt{\left[A \sum_{i=1}^{n} \sum_{j=1, j\neq i}^{n} k_{i,j} \right] / \left[\pi n(n-1) \right]}, \tag{5}$$

where *d* is the distance that determines the value within a certain distance, *n* is the total number of the pixels whose OTSU result is 1, *A* is the pixel location or frequency, and the weight $k_{i,j}$ is the influence of the elements within a certain distance *d*, when the factor *i* and factor *j* are in the distance *d*, the weight $k_{i,j}$ is 1, or the weight $k_{i,j}$ is 0.

For example, take the nighttime lights satellite imagery obtained on 14 June 2018 within the third ring road of Wuhan city and assess the optimal structure pattern of brightness values. We use the multidistance spatial cluster analysis tool in ArcGIS 10.1 software (ArcGIS 10.1, 2012) to process the imagery brightness values. We use the default 10 times as the number of distance changes, select 99_PERMUTATIONS as the compute confidence envelope, and use MINIMUM_ENCLOSING_RECTANGLE as the study area method. The results are shown in Table 1 and Figure 4.

Table 1. Results of Ripley's K analysis of the imagery obtained on 14 June 2018 of Wuhan city.

OBJECT ID	Expected K	Observed K	Diff K
1	0.006979	0.008704	0.001728
2	0.013953	0.016967	0.003014
3	0.020929	0.024651	0.003722
4	0.027905	0.032208	0.004303
5	0.034881	0.039444	0.004563
6	0.041858	0.046294	0.004436
7	0.048834	0.052962	0.004128
8	0.055810	0.059453	0.003643
9	0.062786	0.065545	0.002759
10	0.069763	0.071415	0.001653
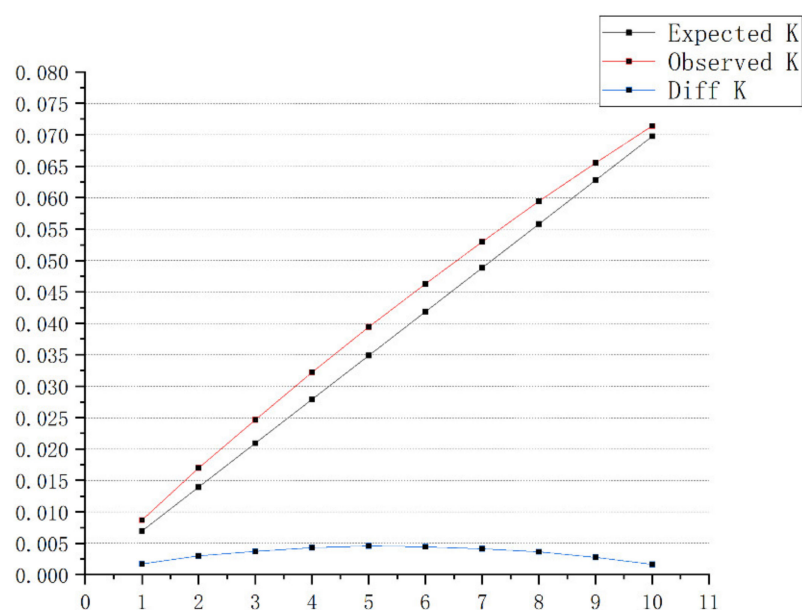
Figure 4. Ripley's K multidistance mode analysis result with nighttime lights satellite imagery.

In Table 1, *ExpectedK* is the expected K value, which refers to the expected value that is calculated through the random hypothesis. *ObservedK* represents the observed K value, which is determined by the actual data. *DiffK* reveals the difference between the observed K value and expected K value. It can be concluded that the data is clustered if *DiffK* is positive; whereasthe data is discrete. In Figure 4, the blue line is the expected value of K value, whereas the red line is the observed value of K value. As can be seen, the blue line is always above the red line, indicating that under the spatial distribution mode (distance mode) participating in the analysis, elements in the study area are always distributed in the clustering mode, which verifies the analysis results based on the Global Moran's I statistical method of global autocorrelation of brightness values.

It can be found from Figure 4 that the difference between the observed value and the expected value reaches the maximum in the distance modes numbered 5 and 6, and the distance characteristic value of the most obvious spatial clustering mode is determined to be 0.004563. Thus, the exploratory spatial cluster analysis can be performed according to the value of the optimal spatial distribution characteristic.

2.2.3. Local Spatial Autocorrelation Clustering and Hot Spot Detection

As the global spatial autocorrelation only uses a single value to reflect global autocorrelation, it is not easy to find spatial correlation patterns in different regions. The local spatial autocorrelation can analyze the distribution characteristics of the local spatial system and better explore the local characteristics of spatial data [22]. The local spatial autocorrelation is often expressed in the form of

Moran Scatter Plot, Local Indicators of Spatial Association (LISA), and Anselin Local Moran's I [17,23]. In this paper, the Anselin Local Moran's I method is adopted to explore the local spatial autocorrelation characteristics of brightness values, as shown in Equation (6)

$$I_i = \frac{x_i - \overline{X}}{S_i^2} \sum\nolimits_{j=1, j \neq i}^{n} \omega_{i,j}\left(x_i - \overline{X}\right), \tag{6}$$

where n is the total number of the pixels whose OTSU result is 1, $(x_i - \overline{X})$ is the difference between the brightness value of the pixel and its mean, $\omega_{i,j}$ is the spatial weight between location i and j, and $S_i^2 = [\sum_{j=1, j \neq i}^{n} \omega_{i,j}]/[n-1] - \overline{X}^2$. The scores (z) are determined by Equation (7) as follows:

$$z_{I_i} = [I_i - E[I_i]]/[\sqrt{V[I_i]}], \tag{7}$$

where $E[I_i] = -[\sum_{j=1, j \neq i}^{n} \omega_{i,j}]/[n-1]$ and $V[I_i] = E[I_i^2] - E[I_i]^2$. If the z of an element is a relatively high positive value, it means that surrounding elements assume similar values that are high or low. In the calculations, statistically significant (0.05 indicates significance level) high-value clustering is represented by *HH*, whereas statistically significant low-value clustering is denoted by *LL*.

We use the cluster and outlier analysis tool in ArcGIS 10.1 software (ArcGIS 10.1, 2012) to identify statistically significant hotspots, cold spots, and spatial outliers. The local autocorrelation clustering is carried out on the data in the area within Wuhan's third ring road according to the method in this paper, and 321 pixels with high value clustering characteristics are obtained. Figure 5 shows the high-value clustering feature points of the June imagery of Wuhan.
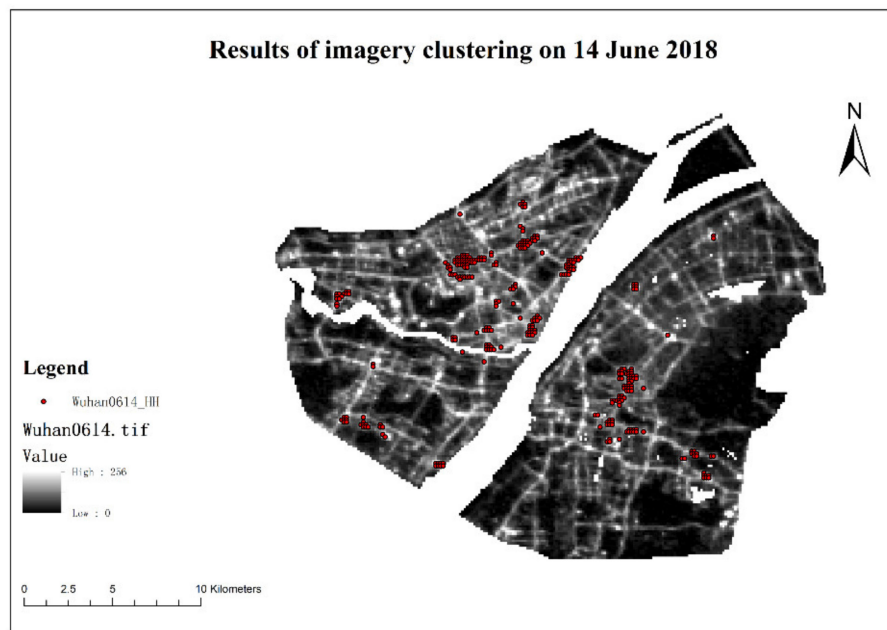
Figure 5. Results of clustering of the June imagery of Wuhan.

2.2.4. Geographical Distribution Measure of Commercial Areas

The high-value hotspots obtained by local autocorrelation cluster detection can be regarded as the center of urban commercial areas. This paper uses the method of constructing standard deviation ellipse to further determine the scope, center change, and development trend of the urban commercial area.

The standard deviation ellipse can summarize the spatial characteristics of geographic features, which include the central tendency, the dispersion, and the directional trends. The generated result is

an ellipse with three key parameters, namely, the center of the circle, the axial length of XY, and the angle of rotation. The axial length of the standard deviation ellipse can be calculated by Equation (8):

$$SDE_x = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \overline{X})^2}{n}}, \; SDE_y = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \overline{Y})^2}{n}}, \tag{8}$$

where x_i, y_i are the coordinates of the pixel location, $(\overline{X}, \overline{Y})$ is the mean center of points, n is the total number of the pixels whose OTSU result is 1, SDE_x is the semimajor axis of the standard deviation ellipse, and SDE_y is the semiminor axis of the standard deviation ellipse.

The angle of rotation is calculated in Equation (9) as follows:

$$\tan\theta = (A + B)/C, \tag{9}$$

where $A = \left(\sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} y_i^2\right)$, $B = \sqrt{\left(\sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} y_i^2\right)^2 + 4\left(\sum_{i=1}^{n} \overline{x}_i\overline{y}_i\right)^2}$, $C = 2\sum_{i=1}^{n} \overline{x}_i\overline{y}_i$, and $\overline{x}_i\overline{y}_i$ is the standard deviation between the cluster elements and the center of the ellipse [24–26].

We use the directional distribution tool in ArcGIS 10.1 software (ArcGIS 10.1, 2012) to process high-value cluster points. 2_STANDARD_DEVIATION is selected as the standard deviation level. According to the above equation, we obtain the distribution scope and direction of the urban commercial areas as shown in Figure 5. The construction of a standard deviation ellipse should contain at least three feature points. However, the number of feature points in some areas that are *HH* in all three imageries is less than three. For this case, this paper uses a fixed radius of 0.005 decimal degrees to generate circular buffer for these feature points, and takes the buffer as the scope of the commercial area, as shown in Figure 6.
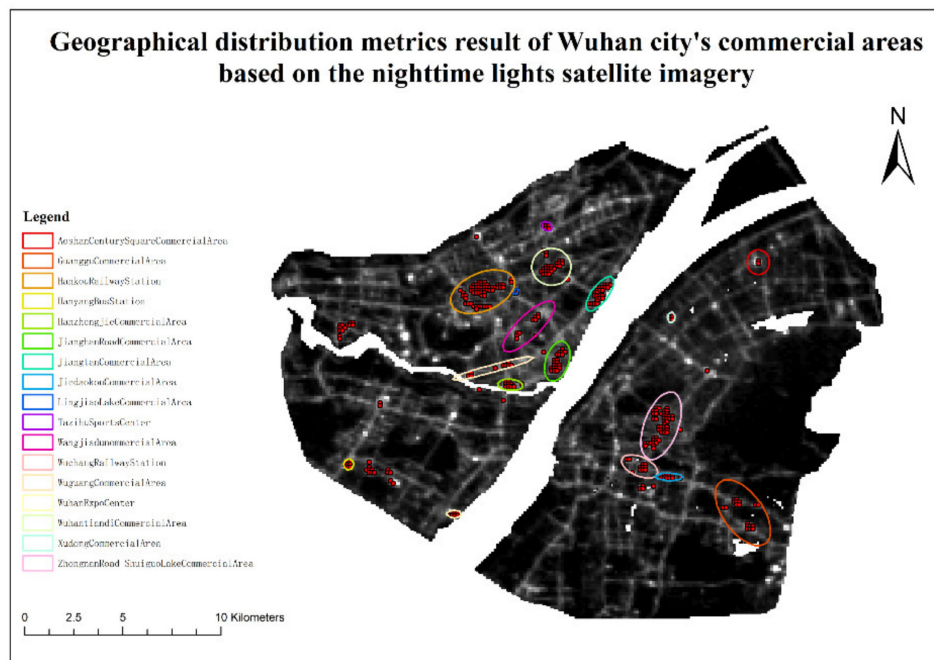
Figure 6. Geographic distribution metrics result of Wuhan city's commercial area based on the nighttime lights satellite imagery.

3. Results

Shenyang and Wuhan are two typical cities in north and south China. They have similar development level in urban scale and economic strength. This paper chooses the nighttime lights satellite imageries of these two cities for experiments. We list the green coverage rates in built-up

areas of the two cities [27,28], 14 large commercial areas in the region within the third ring road of Wuhan [29–31] and 11 in the corresponding region in Shenyang [32,33], in Table 2, which shows all the data and information used in the experiment.

Table 2. Commercial areas and green coverage rate of the two cities.

| City	Imagery Period	Commercial Area	Green Coverage Rate
Wuhan	14 June 2018 and 15 September 2018	Zhongnan Road commercial area, Wuguang commercial area, Zhongjiacun commercial area, Guanggu commercial area Jiedaokou commercial area, Hanzhengjie commercial area, Wangjiadun commercial area, Simenkou commercial area, Xudong commercial area, Aoshan Century Square commercial area, Lingjiao Lake commercial area, Jiangtan commercial area, Wuhantiandi commercial area, JianghanRoad commercial area	39.47%
Shen-yang	15 September 2018 and 17 March 2019	Jinlang commercial area, Xita commercial area, Beihang commercial area, Taiyuanjie commercial area, Wuai commercial area, Nanta commercial area, Zhongjie commercial area, Tiexi commercial area, Beizhan commercial area, Aoti commercial area, Changjiangjie commercial area	38.88%

3.1. The Commercial Areas Detection

By adopting the method proposed in this paper to detect and extract commercial areas, we can obtain 12 commercial areas located within the third ring road of Wuhan City (we use only one ellipse to represent the Zhongnan Road commercial area and the Shuiguo Lake commercial area because they are too close to each other). These commercial areas not only include the old ones in the traditional sense, such as the Jianghan Road commercial area, the Wuguang commercial area, and the Jiangtan commercial area, but also include the emerging ones, such as the Wuhantiandi commercial area, the Lingjiao Lake commercial area, and the Wangjiadun commercial area [29–31].

However, there are some incorrect extraction results in high value clustering hotspots. This is mainly because in addition to commercial areas, there are also some areas in the city with abundant night lights, such as Hankou railway station and Hanyang bus station, which are important transportation hubs in Wuhan. They are important areas for population gathering even at night and need a lot of lights to ensure the normal operation of railway and highway traffic. Another example is the Wuhan Expo Center in Hanyang district. As the largest exhibition venue in central China, it often needs to undertake various international and domestic exhibitions, which also produces a lot of lights at night. These factors lead to the inability to distinguish the commercial areas from the high-light areas of the noncommercial areas based only on the brightness values of the imageries.

From the perspective of the spatial distribution characteristics of each commercial area, its central trend, distribution range, and direction trend basically accord with the relevant research results of commercial areas in Wuhan. Taking the Zhongnan Road-Shuiguo Lake commercial area as an example, the commercial area extracted by using the method proposed in this paper is an ellipse with a semimajor axis in the direction of Zhongnan Road and Zhongbei Road, reaching Qingyuzui subway station in the north and Zhongnan No.2 Road in the south. In addition, as an important part of the Shuiguo Lake commercial area, Han Street is a pedestrian street whose extension direction is perpendicular to that of Zhongbei Road, so the length of the semiminor axis of the extracted commercial area mainly depends on the length of the Han Street. The map of this area is shown in Figure 7. The final detection

result is an ellipse that is approximately round, which conforms to the basic situation of the Zhongnan Road-Shuiguo Lake commercial area [34,35]. Overall, the analysis results of Wuhan commercial areas based on the brightness value of nighttime lights satellite imagery are in line with the objective facts and reflect the spatial pattern and distribution characteristics of commercial and economic distribution centers in Wuhan.
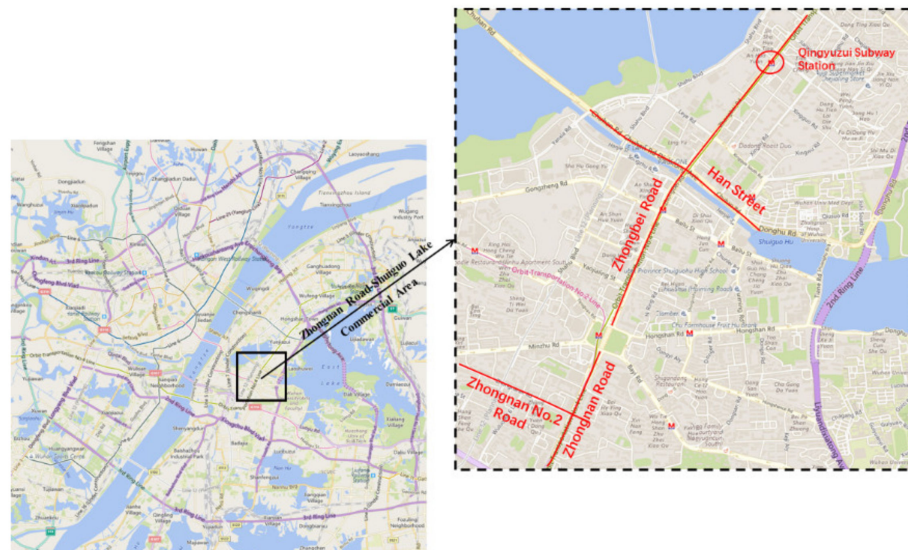

Figure 7. The map of the Zhongnan Road–Shuiguo Lake commercial area (base map: Bing Map).

In order to further verify the effectiveness of the proposed method, we conduct the commercial areas detection on the three imageries within the third ring road of Wuhan on 15 September 2018, and the corresponding area of Shenyang on 10 September 2018 and 17 March 2019, respectively. All four Global Moran's I statistics are shown in Table 3. These results indicate that the global spatial autocorrelation of brightness values distribution pattern conforms to the statistical characteristics of typical clustering models. This means that these four imageries can be used to detect commercial areas.

Table 3. The Global Moran's I statistics.

Imagery	Moran's Index	z-Score	p-Value	Is Clustered
Wuhan, 14 June 2018	0.059590	28.116668	0.000000	Yes
Wuhan, 15 September 2018	0.031851	4.293452	0.000018	Yes
Shenyang, 10 September 2018	0.034785	14.722704	0.000000	Yes
Shenyang, 17 March 2019	0.300103	39.895744	0.000000	Yes

We combine the imagery of Shenyang on 17 March 2019 and the existing commercial areas information of Shenyang to construct standard deviation ellipses for the high-brightness clustering points and obtain the Shenyang's commercial areas detection results. By analyzing the results, we find that large-scaled commercial areas in Shenyang City like the Jinlang commercial area, the Zhongjie commercial area, and the Xita commercial area are detected. In addition, the direction of the standard deviation ellipses and the actual extension of the commercial areas are also generally consistent, which further validates the effectiveness of the proposed method [32,33]. The detection results are shown in Figure 8.
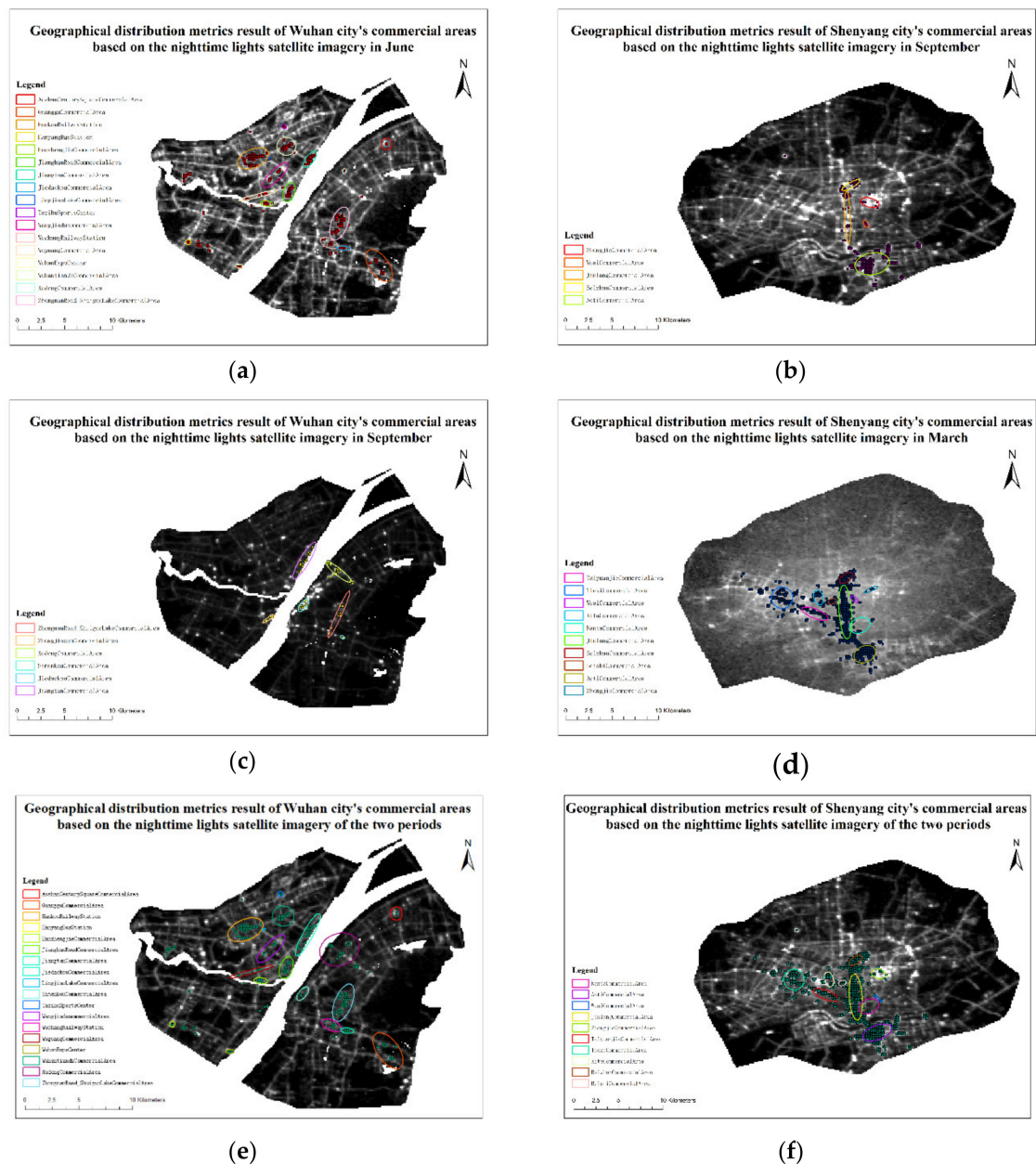
Figure 8. Geographical distribution metrics results: (**a**) results of detection in Wuhan on 14 June 2018; (**b**) results of detection in Shenyang on 10 September 2018; (**c**) results of detection in Wuhan on 15 September 2018; (**d**) results of detection in Wuhan on 17 March 2019; (**e**) results of the combination of (**a**) and (**c**); and (**f**) results of the combination of (**b**) and (**d**).

Figure 8a,c shows the commercial areas detected in Wuhan on 14 June 2018 and 15 September 2018. Figure 8b,d shows the commercial areas detected in Shenyang on 10 September 2018 and 17 March 2019. We also combine the detection results from two periods in the same region to ensure the comprehensiveness of the results, as shown in Figure 8e,f. Combining the detected commercial areas with the existing ones, we can create two accuracy tables, as shown in Tables 4 and 5.

Table 4. Detection accuracy table of Wuhan's imagery.

Commercial Area	Imagery on 14 June 2018	Imagery on 10 September 2018	Combination of the Two Imageries
Zhongnan Road commercial area	√	√	√
Wuguang commercial area	√		√
Zhongjiacun commercial area		√	√
Guanggu commercial area	√		√
Jiedaokou commercial area	√	√	√
Hanzhengjie commercial area	√		√
Wangjiadun commercial area	√		√
Simenkou commercial area		√	√
Xudong commercial area	√	√	√
Aoshan Century Square commercial area	√		√
Lingjiao Lake commercial area	√		√
Jiangtan commercial area	√	√	√
Wuhantiandi commercial area	√		√
JianghanRoad commercial area	√		√
Detection Accuracy	85.7%	42.9%	100.0%

Table 5. Detection accuracy table of Shenyang's imagery.

Commercial Area	Imagery on 10 September 2018	Imagery on 17 March 2019	Combination of the Two Imageries
Jinlang commercial area	√	√	√
Xita commercial area		√	√
Beihang commercial area			
Taiyuanjie commercial area		√	√
Wuai commercial area	√	√	√
Nanta commercial area		√	√
Zhongjie commercial area	√	√	√
Tiexi commercial area		√	√
Beizhan commercial area	√	√	√
Aoti commercial area	√	√	√
Changjiangjie commercial area		√	√
Detection Accuracy	45.5%	90.9%	90.9%

3.2. Sensitivity Analysis

This paper uses the OSTU algorithm to determine the imagery binarization threshold. However, different binarization algorithms may result in different imagery segmentation thresholds. In order to analyze the effect of this commercial area detection method on different segmentation thresholds, a sensitivity analysis needs to be performed.

As different imagery binarization algorithms are executed, it can be found that the absolute value of the difference between the threshold obtained by the OSTU algorithm and the one by other algorithms is within 2. Therefore, for the imagery segmentation threshold calculated by the OSTU algorithm, we use the values obtained by adding 2 and subtracting 2 as the new thresholds to binarize the imagery. Then, the respective Global Moran's Is are calculated, as shown in the following table.

Based on the data in Table 6, we can draw a sensitivity analysis line chart about the binarization threshold, as shown in Figure 9.

It can be found from the figure that as the threshold value fluctuates, the Global Moran's I will also change, but the overall change is small. This effectively shows that the method proposed in this paper has low threshold sensitivity for imagery segmentation and many different imagery binarization algorithms can be used to pre-process the nighttime lights satellite imagery.

Table 6. The Global Moran's I with different thresholds.

Imagery	Threshold	Moran's Index
	13	0.024876
Wuhan, 14 June 2018	16	0.059590
	19	0.051476
	44	0.050771
Wuhan, 15 September 2018	47	0.031851
	50	0.028270
	26	0.041921
Shenyang, 10 September 2018	29	0.034785
	32	0.026875
	11	0.342360
Shenyang, 17 March 2019	14	0.300103
	17	0.296880
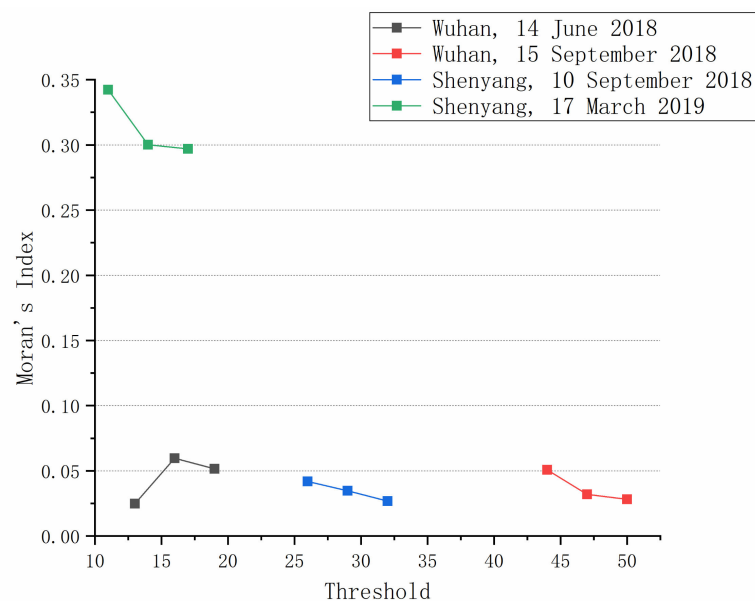
Figure 9. Sensitivity analysis line chart about the binarization threshold

4. Discussion

4.1. Detection Result in Different Periods

Comparing the results of ESDA of Wuhan imageries on 14 June 2018 and 15 September 2018, we find that the global Moran's I and the commercial area detection accuracy in June (0.059590 and 85.7%) are higher than that in September (0.031851 and 42.9%), and compared with the ESDA results on 10 September 2018 and 17 March 2019 in Shenyang, the global Moran's I and the commercial area detection accuracy in March (0.300103 and 90.9%) are higher than that in September (0.034785 and 45.5%) too. For the same region, the value of the global Moran's I can reflect the clustering degree of high-value points, and thus reflect the quality of the commercial areas detection results, that is, the detection accuracy. We believe that this is closely related to the local weather characteristics—whether Wuhan in June or Shenyang in March, the weather conditions in both places are so harsh that many outdoor recreational activities cannot be carried out, leading more residents to choose shopping malls for leisure activities [36]. Therefore, compared with September when the weather conditions are more suitable for outdoor activities, hot June in Wuhan and cold March in Shenyang tend to boost the business for local commercial areas. Thus, the global Moran's I and the detection accuracy calculated during these periods are also relatively high.

From Table 5, we find that although the number of commercial areas detected from the imagery of Wuhan on 15 September 2018 is small, the Simenkou commercial area and the Zhongjiacun commercial area missed by the imagery of Wuhan on 14 June 2018 are successfully detected. If the detection results of the two periods are combined, all the large commercial areas in the region of the Wuhan third ring road can be detected. Thus, it is worthwhile to use imageries from multiple time periods in the process of conducting commercial areas detection.

Based on the above comparison, we believe that in the process of using nighttime lights satellite imagery for commercial areas detection, we ought to analyze the weather conditions of the study area primarily, and try to select imageries of the period with poor weather conditions, such as winter and summer, for experiments. The detection results of these periods are often more consistent with the actual situation and have higher accuracy. At the same time, we may also select as many imageries as possible from the same area for research, and integrate the detected commercial areas to obtain more comprehensive results.

4.2. Detection Result in Different Regions

Combining the results of ESDA of Wuhan imageries with that in Shenyang imageries (as shown in Table 7), we can use the accuracy as the objective evaluation basis for the quality of the commercial areas detection, to explore the law of spatial variation of clustering degrees in the detection.

Table 7. Detection accuracy table of Shenyang's imagery.

Region	Imagery	Detection Accuracy	Green Coverage Rate
Wuhan City	14 June 2018	85.7%	39.47%
	15 September 2018	42.9%	
	Combination	100.0%	
Shenyang City	10 September 2018	45.5%	38.88%
	17 March 2019	90.9%	
	Combination	90.9%	

We use the combination of detection results in two periods in Wuhan and Shenyang (Figure 8e,f) to represent the final commercial area detection results in the corresponding regions. By comparing these two results, we can discuss the influence of geospatial factors on the commercial area detection.

Although there are many variables caused by different research areas, few of them can really affect the distribution of brightness value of imageries. We believe that the influence of green coverage in different areas on the light brightness value in the imagery is relatively significant. The green coverage rate in built-up areas can well reflect the exuberance of trees (especially street trees) in a city. According to the statistical yearbook (Wuhan Municipal Bureau of Statistics 2018, Shenyang Municipal Bureau of Statistics 2018), green coverage rate in built-up areas within the third ring road of Wuhan is slightly higher than that of Shenyang. For the two results, the detection accuracy shares the same rule—the accuracy in the area within the third ring road of Wuhan is marginally higher than that of Shenyang. According to analysis, we believe that areas with a higher green coverage rate have more street trees, which may reduce the interference of urban lighting equipment (especially street lamps) to the commercial area detection to a certain extent [37]. Therefore, the accuracy obtained by detection for the region within the third ring road of Wuhan is higher.

Therefore, we can conclude that the method of using the nighttime lights satellite imagery for commercial areas detection proposed in this paper is more suitable for the regions with a relatively high vegetation coverage rate in urban built-up areas. Thus, it is necessary to analyze the feasibility and accuracy of the detection based on the local green coverage rate when we use this method to detect commercial areas.

4.3. Results Analysis

Based on the above discussion, we can gain some enlightenment when selecting experimental data. There are two factors that need be taken into consideration. Firstly, when it comes to selecting of imagery, weather conditions should be noted. Furthermore, we should do it while it is poor, because the harsh weather can ensure that in the area of interest, the indoor commercial areas will be more frequently visited. Consequently, it is easier to detect and extract hot spots. Secondly, in terms of the selection of research areas, we ought to select areas with a relatively high green coverage rate in urban built-up areas to reduce the interference of street lamp on the detection of commercial areas.

By the time of writing, Luojia No.1 01 has transmitted limited imageries of Wuhan City and Shenyang City that could be used for research. In the imageries that we process, errors in pixel brightness values and inconsistent registration of different imageries exist inevitably, which may lead to certain deviations in the actual number and range of commercial areas extracted [38]. On one hand, we still need to screen out the better quality and larger time span imageries from the imageries transmitted from Luojia No.1 01 in the future for testing and analysis. On the other hand, although the brightness value of the imagery has been normalized during the preprocessing, this normalization is only linear. We can perform radiation calibration in preprocessing to convert the imageries' brightness value to absolute radiation brightness, so as to adapt to the need for nonlinear normalization.

5. Conclusions

As the latest generation of nighttime lights satellite, Luojia No.1 01 well supplements the spatial resolution deficiency of the two kinds of nighttime lights satellites (DMSP/OLS and NPP/VIIRS) that have been widely used before and provides a new research idea for the detection and extraction of urban commercial areas. This paper proposes a method that takes the nighttime lights satellite imageries transmitted by Luojia No.1 01 as the data source, makes use of the characteristic of the city commercial areas being more prominent in night lighting, carries out cluster analysis on the brightness value of the imageries, constructs the standard deviation ellipses, and then extracts the distributions and scopes of the urban commercial areas. Compared with the traditional measurement and analysis method of urban commercial areas, the method in this paper uses nighttime lights satellite imagery for operation, which has the characteristics of objective data, good real-time performance and high accuracy, and provides an intuitive and reliable reference for relevant departments to make urban planning and administrative decisions. At the same time, the dynamic changes of the commercial area can be obtained by using the combination of imageries from different periods to provide a basis for further research on the development and changes of the city.

References

1. Marshall, A. *Principles of Economics*; Macmillan Publishers: London, UK, 1890.
2. Jin, M. *A Study on Market Zone Partition by Using Hierarchical Weighted Voronoi Diagram Model: A Case of Hangzhou*; Zhejiang University Press Co., Ltd.: Hangzhou, China, 2016.
3. Lu, C.; Wang, H.; Liu, L. Retail district range compare model construction and Shanghai districts demonstration. *Econ. Geogr.* **2015**, *35*, 133–137.
4. Yang, L.; Zhu, H.; Wu, J. Market area analysis in retailing based on GIS. *Remote Sens. Technol. Appl.* **2003**, *18*, 144–148.

5.　Reilly, W.J. *The Law of Retail Gravitation*; Knickerbocker Press: New York, NY, USA, 1931.

6.　Westbrook, R.A.; Black, W.C. A motivation-based shopper typology. *J. Retail.* **1985**, *61*, 78–103.

7.　Igene, C.A.; Ghosh, A. Consumer and producer behavior in a multipurpose shopping environment. *Geogr. Anal.* **1990**, *22*, 70–93. [CrossRef]

8.　McLafferty, S.L.; Ghosh, A. Multipurpose shopping and the location of retail firms. *Geogr. Anal.* **2010**, *18*, 215–226. [CrossRef]

9.　GoodChild, M.F. Citizens as sensors: The world of volunteered geography. *GeoJournal* **2007**, *6*, 211–221. [CrossRef]

10.　Hu, Q.; Wang, M.; Li, Q. Urban hotspot and commercial area exploration with check-in data. *AGCS* **2014**, *43*, 314–321.

11.　Zhao, W.; Li, Q.; Li, B. Extracting hierarchical landmarks from urban POI data. *J. Remote Sens.* **2011**, *15*, 973–988.

12.　Wang, M.; Hu, Q.-W.; Li, Q.-Q.; Qin, L.-K. Extracting hierarchical landmark from check-in data. *Chin. J. Comput.* **2016**, *39*, 405–413.

13.　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]

14.　Li, D.; Li, X. An overview on data mining of nighttime light remote sensing. *AGCS* **2015**, *44*, 591–601.

15.　Luojia No.1 01 Satellite Data and Application Service. Available online: http://59.175.109.173:8888/index.html (accessed on 1 September 2019).

16.　Otsu, N. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern. Syst.* **1979**, *9*, 62–66. [CrossRef]

17.　Aneslin, L. Interactive techniques & exploratory spatial data analysis. In *Geographical Information Systems: Principles, Techniques, Management and Applications*, 2nd ed.; Paul, A.L., Michael, F.G., Eds.; Wiley Publisher: Hoboken, NJ, USA, 2005.

18.　Cliff, A.; Ord, J. *Spatial Processes: Models and Applications*; Pion Ltd.: London, UK, 1981.

19.　Geng, X.; Du, X.; Hu, P. Spatial clustering method based on raster distance transform for extended objects. *AGCS* **2009**, *38*, 162–167.

20.　Getis, A.; Ord, J. The Analysis of spatial association by use of distance statics. *Geogr. Anal.* **1992**, *24*, 189–206. [CrossRef]

21.　Guo, A.; Li, Y.; Yang, J. Research on spatial and temporal differences of urban architectural landscape based on multi-distance spatial cluster. *Bull. Surv. Mapp.* **2017**, *2*, 102–105.

22.　Tiefelsdorf, M.; Boots, B. A Note on the extremities of Local Moran's Iis and their impact on Global Moran's I. *Geogr. Anal.* **1997**, *29*, 248–257. [CrossRef]

23.　Aneslin, L. Local Indicators of Spatial Association—LISA. *Geogr. Anal.* **1995**, *27*, 93–115. [CrossRef]

24.　Robert, S.Y. The standard deviational ellipse: An updated tool for spatial description. *Geogr. Ann. Ser. B* **1971**, *53*, 28–39.

25.　Wang, B. Theories and methods for soil grain orientation distribution on SEM by standard deviational ellipse. *Chin. J. Geotechn. Eng.* **2009**, *31*, 1082–1087.

26.　Wang, B.; Shi, W.; Miao, Z. Confidence analysis of standard deviational ellipse and its extension into higher dimensional euclidean space. *PLoS ONE* **2015**, *10*, e0118537.

27.　Pan, J. Wuhan Municipal Bureau of Statistics. In *Wuhan Statistical Yearbook 2018*; China Statistics Press: Beijing, China, 2018.

28.　Li, C.; Tian, X. Shenyang Municipal Bureau of Statistics. In *Shenyang Statistical Yearbook 2018*; China Statistics Press: Beijing, China, 2018.

29.　Zhao, W. Functions of a central city& development of Wuhan urban agglomerations. *Wuhan Univ. J. Philos. Soc. Sci.* **2005**, *58*, 300–305.

30.　Zheng, X. *Spatial Location Analysis of Large Retail Stores in Wuhan*; Huazhong University of Science & Technology Press Co., Ltd.: Wuhan, China, 2005.

31.　Li, J. *A Research to the Market Analysis and Development Trend of Wuhan's Commercial Real Estate*; Central China Normal University Press: Wuhan, China, 2014.

32.　Shen, H. *Study on the Reasonable Scale of Commercial Real Estate in Shenyang*; Shenyang Jianzhu University: Shenyang, China, 2015.

33. Wu, F. *A Research on the Problems and Strategies of Retail Enterprises in Trading Area of Shenyang Golden Corridor*; Shenyang University: Shenyang, China, 2018.

34. Jiang, H. *Study on the Temporal and Spatial Characteristics of Urban Residents' Travel Based on Taxi Data*; Wuhan University Press: Wuhan, China, 2019.

35. Li, Z.; Sun, Y.; Lin, S.; Guo, Y. Spatial reproduction and place promotion in central China: A case study of Chu Milky Street of Wuhan. *Urban Rural Plan.* **2019**, *2*, 60–69.

36. Dashdorj, Z.; Sobolevsky, S. Impact of the spatial context on human communication activity. In Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers, Osaka, Japan, 9–11 September 2015; pp. 1615–1622.

37. Lu, X. Study about effects of street trees on LED road lighting. *China Light Lighting* **2016**, *2*, 4–8.

38. Zhong, L.; Liu, X. Application potential analysis of LJ1-01 new nighttime light data. *Bull. Surv. Mapp.* **2019**, *7*, 132–137.