



Article Study on the Spatial Classification of Construction Land Types in Chinese Cities: A Case Study in Zhejiang Province

Lin Dong¹, Jiazi Li¹, Yingjun Xu^{2,3,4}, Youtian Yang¹, Xuemin Li¹ and Hua Zhang^{2,3,4,*}

- ¹ Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China; 201921051152@mail.bnu.edu.cn (L.D.); 202021051160@mail.bnu.edu.cn (J.L.); 201921051167@mail.bnu.edu.cn (Y.Y.); lixuemin@mail.bnu.edu.cn (X.L.)
- ² Academy of Disaster Reduction and Emergency Management, Beijing Normal University, Beijing 100875, China; xyj@bnu.edu.cn
- ³ Key Laboratory of Environmental Change and Natural Disaster, Ministry of Education, Beijing Normal University, Beijing 100875, China
- ⁴ State Key Laboratory of Surface Processes and Resource Ecology, Beijing Normal University, Beijing 100875, China
- * Correspondence: zhanghua2011@bnu.edu.cn; Tel.: +86-15120098201

Abstract: Identifying the land-use type and spatial distribution of urban construction land is the basis of studying the degree of exposure and the economic value of disaster-affected bodies, which are of great significance for disaster risk predictions, emergency disaster reductions, and asset allocations. Based on point of interest (POI) data, this study adopts POI spatialization and the density-based spatial clustering of applications with noise (DBSCAN) algorithm to accomplish the spatial classification of construction land. Zhejiang province is selected as a study area, and its construction land is divided into 11 land types using an accurate spatial classification method based on measuring the area of ground items. In the research, the POI dataset, which includes information, such as spatial locations and usage types, was constructed by big data cleaning and visual interpretation and approximately 620,000 pieces in total. The overall accuracy of the confusion matrix is 76.86%, which is greatly improved compared with that constructed with EULUC data (61.2%). In addition, compared with the official statistical data of 11 cities in Zhejiang Province, the differences between the calculated spatial proportions and statistics are not substantial. Meanwhile, the spatial characteristics of the studied land-use types are consistent with the urban planning data but with higher accuracy. The research shows that the construction land in Zhejiang Province has a high degree of land intensity, concentrated assets, and high economic exposure. The approach proposed in this study can provide a reference for city management including urbanization process, risk assessment, emergency management and asset allocation.

Keywords: land-use type; POI; DBSCAN; socioeconomic risk

1. Introduction

Urban buildings are the most important places in which human social and economic activities occur, and exposure analyses of land-use types and economic values are of great significance for disaster risk assessments and emergency mitigation strategies [1]. With the rapid development of urbanization, urban land use, and spatial structures are undergoing significant changes, and the highly segmented division of labor in social industries imposes more refined and efficient requirements for urban spatial planning [2–4]. However, the traditional identification method of land-use types based on remote sensing and actual observation data requires considerable time and economic costs, and the data quality and update speed are also subject to the interference of weather and human-related factors [5]. In recent years, the extensive use of big data, such as Global System for Mobile (GSM), Global Position System (GPS), Point of Interest (POI), Smart Card Data (SCD), taxi trajectory



Citation: Dong, L.; Li, J.; Xu, Y.; Yang, Y.; Li, X.; Zhang, H. Study on the Spatial Classification of Construction Land Types in Chinese Cities: A Case Study in Zhejiang Province. *Land* **2021**, *10*, 523. https://doi.org/ 10.3390/land10050523

Academic Editors: Robert Gilmore Pontius, Jr. and Jesper M. Paasch

Received: 24 February 2021 Accepted: 11 May 2021 Published: 13 May 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 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/). and night light data has made it possible to acquire high-quality and large-scale spatial data quickly [6], which promotes the study of existing urban structure and functional areas.

Extracting land cover by remote sensing has been extensively used [7]. Land cover maps represent spatial information on different types of physical coverage of the Earth's surface, e.g., forests, grasslands, croplands, lakes, wetlands [8]. Different from land cover, land use is widely understood as a man-made adjustment of the land surface and the types usually include farmlands, parks, residential areas, etc. [9]. At present, scholars have conducted in-depth studies on the identification of urban land-use types and their interaction rules by using multisource data [10,11]. Among these studies, POI data are proven to be very practical in urban zoning studies due to their advantages of rich semantics, short acquisition cycles, low costs, fast updating speeds, objectivity and strong current situations [12,13]. For example, POI density characteristics are used for similarity measurements [14], and when combined with remote sensing data, a wide range of land distribution patterns can be identified [15]. Among them, Essential Urban Land Use Categories in China (EULUC-China) data obtained by using the random forest law have been widely applied [14]. Integrating POI with Word2Vec [16] and Place2Vec [17], the semantic model can be used to deduce the training method to detect urban land use distributions. Based on Latent Dirichlet Allocation (LDA) topic modeling technology, a statistical framework has been developed to identify regional functions and explore potential POI topics [18]. Using POI as metadata and combining human mobility information to develop the framework of discovers regions of different functions (DROF), the relationship between regional land-use types and POI can be expressed in the form of a function [19]. POI and mobility data are utilized to identify urban functional areas by head/tail breaks and kernel density clustering [20], among other methods. Studies have attempted to research the relationship between POI information and urban land-use types from different perspectives. These studies include not only mining the POI attribute information and the distribution model of density frequency [21] but also mining the complementary relationships between POI data and other data types which include remote sensing, social media, road network and building footprint [22,23].

Clustering algorithms have been widely applied in the field of urban land-use type recognition [16,17,24,25]. There are two main methods depending on whether the attribute characteristics of geographical entities have obvious spatial correlation or not. In the case of spatial correlation, clustering is carried out based on the spatial autocorrelation, suitable for researches that location information is dominant. In another case, clustering focuses on individual differences and correlation analyses of the element's own attributes. The Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm is a representative density-based clustering algorithm [26]. A cluster in DBSCAN is defined as the largest set of points connected by density, and its shape can be arbitrary. Due to the density in the noise area is lower than that in any cluster [27,28], the DBSCAN algorithm can effectively identify noise points in the research data. To identify regional industrial distributions and spatial structures under the spatial proximity effect, DBCSAN can be used to mine the spatial characteristics of ground-based objects in aggregation distributions.

To summarize, existing studies on the spatial classification of construction land types are exploring different approaches, but are still in infancy considering the availability of the classification results. There are few categories, and the spatial classification standards for specific use types are not unified. Based on POI data resources and measured data of the spatial occupation area of various land-use types, combined with inherent social and economic characteristics, such as residential complexes and commercial complexes, this study carries out spatial research on land-use types of construction land to provide support for accurate social and economic risk assessments and other work. This paper is structured as follows: Section 2 explains the sources of main data and classification methods; Section 3 presents the classification results and analyzes the accuracy in different ways; Section 4 discusses objectives and limits of the innovative methods; Section 5 concludes the significances and applications of this research.

2. Materials and Methods

2.1. Data Source

2.1.1. POI and Classification System

The POI data used in this study are derived from the Baidu Map Services (http: //lbsyun.baidu.com/ (accessed on 9 May 2021)), which is one of the most popular and largest web map service providers in China. The original point data total, approximately, 700,000 pieces of original data, including 41 first-level (Level 1) categories, such as shopping areas, medical services and government agencies and several second-level (Level 2) POI categories, which are organized by industry affiliation (http://lbsyun.baidu.com/ index.php?title=lbscloud/poitags (accessed on 9 May 2021)). According to the Code for classification of urban and rural land use and planning standards of development land (GB50137-2011) [29], the POI data are reclassified into 9 Level 1 categories and 8 Level 2 categories. The POI data are classified as Level 2 categories when they cannot be accurately described using Level 1 categories. The attributes of POI data include longitude, latitude, name, address, type, administrative region, and other indicators.

2.1.2. Nighttime Light Image

The nighttime light data represent visible light (e.g., city lights, fishing-fleet lights, fires) captured by remote sensing satellites in cloudless conditions at night. The nighttime light data are obtained from the NPP/VIIRS (Suomi National Polar-Orbiting Partnership's Visible Infrared Imaging Radiometer Suite) data provided by the National Geophysical Data Center (NGDC) of the United States. The 2018 data are used in this study with a spatial resolution of 500 m.

2.1.3. Impervious Land Data

Artificial impervious areas are predominant indicators of human settlements, which are mainly man-made structures that are composed of any material that impedes or prevents natural infiltration of water into the soil [30]. The urban impervious land data are selected from GAIA data [31], a yearly dynamic raster product with a spatial resolution of 30 m. The data are clipped and merged to obtain spatial range data of the impervious land in Zhejiang Province in 2018.

2.1.4. Road Network

The Zhejiang road network data are obtained from the Baidu Electronic Map and include national roads, provincial roads, county roads, township roads, and village roads.

2.1.5. Land-Use Data

The land-use data of non-construction land in Zhejiang Province are selected from the global land cover data of FROM-GLC [32] with a spatial resolution of 30 m. The raster data are clipped to obtain the land cover of non-construction land in Zhejiang Province.

2.2. Classification Method

The overall methodological procedures of the research are shown in Figure 1. First, the original POI system is reclassified according to the Current land use classification (GB/T 21010-2017) [33], and remote sensing images are used to sample tests and estimate the occupation area of each type of ground objects. Second, the impervious land data are optimized by nighttime remotely sensed data to obtain the construction land scope of the study area and the ground areas are served as the basis to get POI grids of the construction land. Then, DBSCAN spatial clustering is carried out for the POI grids, and the parcels composed of closed road networks are taken as a unit [34] to calculate the distribution and proportion of all types of grids in the block. Finally, the land use type of per parcel is judged by statistical results and classification rules, and a spatial distribution map of urban construction land use types, including single types and urban complexes, is obtained.



Figure 1. Methodological procedures of the research.

2.2.1. Measurement of Floor Area by Visual Interpretation

According to the Code for classification of urban and rural land use and planning standards of development land and in accordance with the principle of universality and with the consistency of the POI classification [35,36], the original POI is reclassified, and a new classification system of urban construction land use types based on the reclassified POI is established. A sampling survey is used to extract approximately 100 samples from each category of target features under the Level 2 system, and remote sensing images are used to estimate the ground area. The overall trend of the original data of the sample area presented a positively skewed distribution, and the logarithmic trend showed a normal distribution (Figure A1).

The expectation of a normal distribution is taken as the estimated land area of the corresponding category of ground objects and as the basis and foundation of the POI spatialization to match the POI location information and land area scale. The corresponding relationship between the POI reclassification system and the area estimation results is shown in Table 1.

Land-Use Type		Code	РОІ Туре	Ground Area	95% Confidence Interval
Nonconstruction		01	Agricultural, forestry, animal husbandry and fishery bases	-	-
Residential	02		Living quarters	-	-
Industrial		03	Companies, factories, industrial parks, well-known enterprises	5921	4320-8118
Administrative and public services	04	41	Government agencies, social institutions, service-oriented venues	14,737	9665–22,470
		42	General hospitals, emergency centers	2078	806–5361
		43	Schools and research institutes	4112	2485-6802
Commercial and business facilities	05	51	Convention and exhibition centers, commercial venues	1860	1388–3494
		52	Comprehensive markets, shopping malls and large commercial places	816	621–1073
		53	Commercial service institutions such as banks and gas stations	543	423–698

Table 1. POI classification schemes and estimated ground areas.

5 of 14

Land-Use Type	Code POI T		РОІ Туре	Ground Area	95% Confidence Interval
Logistics and warehouse		06	Logistics storage	98	74–130
Street and	07	71	Wharves, airports, railway stations and transit ports	3336	2177–5232
transportation		72	Parking lots, long-distance bus stations, subways, etc.	463	366–587
Municipal utilities		08	Emergency shelters	2710	1989-3695
Green space	09		Scenic spots, parks and squares	3183	2013-5034

Table 1. Cont.

2.2.2. Assigning the Spatial Attributes of Urban Construction Land

As for any point *i*, the nuclear density of the POI, which is P_i , is calculated as follows:

$$P_i = \frac{1}{n\pi R^2} \times \sum_{j=1}^n K_j \left(1 - \frac{D_{ij}^2}{R^2}\right)^2 \tag{1}$$

where K_j is the weight of point *j*, D_{ij} is the Euclidean distance between spatial points *i* and *j*, *R* is the bandwidth of the calculation rule area ($D_{ij} < R$), and *n* is the number of points *j* in the calculation area. A 500-m bandwidth is selected according to the existing research results [37] to obtain the grid result of the kernel density calculation. Due to the agglomeration and scale effects of cities and towns, the POI density in urban built-up areas is higher than that in suburban and rural nonconstruction land areas.

The NTL and POI comprehensive index method [38,39] can be used to optimize the method of directly using nighttime lights and solve the problems of light overflow and low resolution. Based on the premise that the nighttime light data and POI data have the same weight, the average value method is selected to integrate the positive correlation between the two datasets and urban construction land, eliminate the transient noise of the nighttime light data, and correct the abnormal light value points. The calculation formula is as follows:

$$POINTL_i = \sqrt{P_i \times NTL_i} \tag{2}$$

where $POINTL_i$ is the NTL and POI composite index, P_i is the kernel density value of point *i*, and NTL_i is the brightness value of the nighttime light at point *i*. Finally, by referring to the characteristics of the study area, the threshold value $POINTL_i$ was changed to 13 on the basis of the NTL and POI composite index method [31,32,34] to extract the initial urban construction land scope, and the intersecting operation was carried out with the data of the impervious land area. Then, the final extraction results of construction land are obtained after verification and modification of the original results.

2.2.3. Density-Based Spatial Clustering of Applications with Noise

Because some categories are not suitable for clustering, like street and transportation, a cluster feasibility determination is performed for all land-use types except residential land to identify clustering targets. All land-use types with aggregation and distribution characteristics are selected and shown in Table 2, including the following 7 types: type 03 industrial areas, type 41 government agencies, social institutions, service-oriented venues, type 43 schools and scientific research institutions, type 51 convention and exhibition centers, commercial venues, type 52 comprehensive markets, shopping malls, and other large entertainment and leisure places, type 53 commercial service institutions such as banks and gas stations, type 06 logistics storage facilities and warehouses. The points of these categories are abundant in quantity and relatively dense in distribution, showing the characteristics of aggregation distribution.

	Code	РОІ Туре	Cluster	Eps	MinPts
03		Companies, factories, industrial parks, well-known enterprises	\checkmark	3	5
04	41	Government agencies, social institutions, service-oriented venues	\checkmark	4	6
04	42	General hospital, emergency center		-	-
	43	Schools and research institutes	\checkmark	4	3
05	51	Convention and exhibition centers, commercial venues	\checkmark	3	5
05	52	Comprehensive market, shopping mall and other large commercial places	\checkmark	4	10
	53	Commercial service institutions such as Banks and gas stations	\checkmark	4	4
	06	logistics storage	\checkmark	20	3
07	71	71 Wharves, airports, railway stations and transit ports		-	-
	72	Parking lots, long-distance bus stations, subways, etc.	×	-	-
	08	Emergency shelter	×	-	-
	09	Scenic spots, parks and squares	×	-	-

Table 2. Clustering parameters.

Density-based spatial clustering of applications with noise (DBSCAN) is based on the concepts of reachable density and connected density. The core idea of DBSCAN is that the number of adjacent grids in a given radius neighborhood (Eps) is greater than a given threshold (MinPts). The neighborhood shape is determined by the raster distance, dist(m,n), between two rasters, m and n. Eps and MinPts are regarded as key parameters in the clustering method, and these parameter values directly affect the quality of the clustering results. Referring to the improved, adaptive, fast AF-DBSCAN clustering algorithm put forward by Zhou Zhiping [40], 50*50 quadrangle areas are selected to fit the K-nearest neighbor (KNN) distance distribution curve by polynomial fitting, and the curve inflection point is solved to obtain the optimal value of Eps. According to the statistical distribution characteristics of the data, the mathematical expectation of the number of Eps objects in each grid is calculated to obtain MinPts as follows:

$$MinPts = \frac{1}{n} \sum_{i=1}^{n} P_i$$
(3)

where *n* is the number of grids in the region and P_i is the number of points in point *i*'s Eps neighborhood. The land-use type, clustering basis and parameter setting used in this calculation are shown in Table 2.

As the regions composed of clusters are not closed, many small patches are generated in the same region by clustering, on the basis of identifying the boundary patches and internal patches in the study region, a buffer algorithm is used to establish a buffer zone for patches inside the same region, and the patches are resampled to a uniform 30-m grid to obtain a complete clustering region efficiently and quickly. The buffer distance is shown in Equation (4).

$$L_i = D_i \times \frac{Eps_i}{60} \tag{4}$$

In the formula, L_i is the buffer distance in the raster data of the area of land-use type *i* and D_i and Eps_i are the actual raster resolution and Eps parameter values of type *i*, respectively.

2.2.4. Classification Method Based on Parcel Unit

Considering a 30-m grid is inappreciable for the whole settlements, parcel is chosen as the elementary unit to the classification of urban construction land. After extending disconnected roads using tools provided by ArcGIS which is a map processing software, the vector surface of the road network is obtained by setting buffer zones for different levels of roads by referring to the Highway Route Design Specification (JTG D20-2017) [41]. Then, parcels within the range are identified by the road network closing rules.

Due to the low accuracy of the direct spatialization of type 02 residential land's POI, the exclusive method is adopted to divide the no-data land (without grids of type 03–09) into type 02 residential land grids and type 09 green space grids. The target parcel of the no-data land can be divided into the 02 residential land type if there are POI residential land data in it; otherwise, the parcel can be divided into the 09 green space type. The POI proportion parameter CR is constructed to represent the proportion of the number of grids of each type in the target parcel to the total number of grids. The detailed classification rules are shown in Figure 2, wherein CR1, CR2, and CR3 represent the categories corresponding to the first, second and third CR ranking from large to small, respectively, and FC is the final category of the block determination.



Figure 2. The classification process of the construction land use type based on road network blocks.

The concept of urban complexes is introduced to describe the mixed land use type. Urban complexes are an organic combination of a variety of urban functions based on building groups, including functional aggregation, land use intensification and urban economic aggregation [42]. Urban complexes are introduced into land-use classification systems to more accurately describe the land-use type characteristics of mixed land-use areas [43]. A residential complex refers to a single building or building group that is dominated by residential functions and is composed of spaces with different functions, including several complex functions such as commercial, public services, and transportation functions. In this classification system, residential complexes represent a mixed type, composed of residential buildings, which are the main land type, commercial, and business facilities, administrative public service facilities, and roads, streets, and transportation facilities.

Commercial complexes are multiuse modern urban commercial facilities dominated by commercial functions that are gradually moving towards a comprehensive form integrated with various public functions, such as transportation, culture, community service, and business offices. Commercial complexes are characterized by complex functions, intensive scales and integrated operations [44] and are the product of the mixing of functions and spaces. Commercial complexes correspond to more mature mixed land-use types that are dominated by commercial and business facilities supplemented by administrative and public service facilities, residential buildings, logistics storage facilities and warehouses and replenished by industrial facilities.

3. Classification Accuracy and Results

3.1. Verification of Classification Accuracy

To verify the accuracy of the spatial classification results of urban construction landuse types, a general verification based on the obtained confusion matrix and a spatial distribution verification comparing with remote sensing images are carried out. Taking 30-m grids of each land-use type as the basic units, a confusion matrix is constructed by counting the total number of matching grids and nonmatching grids of each land-use type. Since there is no one-to-one correspondence between residential or green space areas and their POIs, the matrix calculation was not involved. The analysis results show that the overall accuracy of the classification is 76.86%, and the user's accuracy values of the various categories are shown in Table 3. In the confusion matrix, the rows are the map categories and the columns are the reference categories.

Code	03	04	05	06	07	08	Total	UA (%)
03	330237	13088	23856	5790	290	643	373904	88.32
04	16006	148311	51560	3599	118	351	219945	67.43
05	20152	31226	141245	2692	253	90	195658	72.19
06	13689	5902	14507	57379	57	35	91569	62.66
07	7	21	23	7	282	3	343	82.22
08	180	82	92	1	76	1356	1787	75.88
ACC ¹ (%)								

Table 3. Confusion matrix for Level I classification.

¹ ACC represents the overall accuracy and UA represents user's accuracy.

Comparing against the China Statistical Yearbook of 2018, released by China's National Bureau of Statistics and including statistical data about various land-use types, the differences between the obtained percentages of land-use types and the percentages of yearbook-based land-use types are calculated as obtained percentage minus yearbook percentage for each land-use type for each region. Figure A2 is a boxplot of the differences. For the percentage differences, the mean absolute deviation is 4.13 percentage points. Two typical regions of Jianggan District in Hangzhou are selected to show detailed classification results as examples. The classification result map is compared with remote sensing images with road networks (supported by the Baidu Electronic Map), and the analysis shows that the classification result is basically accurate and matches the actual ground-based object distributions (Figure 3). Jianggan District, Hangzhou has good urban function agglomeration; industrial land is distributed along the river, including the Panasonic Hangzhou Industrial Park, Foxconn Qiantang Science and Technology Industrial Park, Beifang Industrial Park, etc. In the northern higher-education zone, there are many universities and research institutes which present spatial agglomeration characteristics; residential buildings and residential complexes are distributed continuously, including the Longhu-Yanlanshan Community, Xinyuan-Jinshacheng Community, and Xiasha Mingdu Community, and are interspersed with public service land and commercial service land to ensure the perfection of the residential functions. Figure 3 shows the comparison of the classification results.







(a)

(b) Legend

Green space

Residential complex



Administration and public services: A1-Hangzhou Dianzi University(Xiasha campus); A2-Zhejiang Sci-Tech University(Xiasha campus); A3-Zhejiang Police Vocational Academy(Xiasha campus). Industrial: B1- Master Kong Beverage Co. Ltd, Hangzhou Dingzheng Packing Material Co., Ltd. Green space: C1-Hangzhou Fire Theme Park. Commericial complex Commericial complex:D1-Yisha Pioneer Park,Advanced Medical Optics(Hangzhou) Co.,Ltd,etc;D2-Haida Office Building,Guomai Nonconstruction land Science Park, Weimian Medical Technology Co., Std, Pod inns, etc.

Figure 3. Comparison of classification results and remote sensing images of typical regions. (a) Comparison map of sample 1 in Jianggan District, Hangzhou City; (b) Comparison map of sample 2 in Jianggan District, Hangzhou.

3.2. Results of the Construction Land Classification

Using the classification model, the urban construction land in Zhejiang Province is extracted and divided into 10 types: residential, industrial, administrative and public service, commercial and business facilities, logistics and warehouse, streets and transportation facilities, municipal utilities, green spaces, residential complexes, and commercial complexes. Classification maps of the central regions of 11 cities in Zhejiang Province are shown in Figure 4.

By the end of 2017, the Department of Natural Resources of Zhejiang Province announced that the construction land area was 13,116.67 square kilometers. Within our study, we calculated that the construction land area of the whole province is 12,714.19 square kilometers, which fits within the official data; the road area is 1339.93 square kilometers and is divided into 101,400 blocks. Based on the spatial classification results obtained in this paper, data regarding the land area proportion of each land-use type in 11 cities in Zhejiang Province are statistically obtained (Table 4).



Figure 4. Classification results of 11 cities in Zhejiang Province (covering nonconstruction land).

02	03	04	05	06	07	08	09
Res	Ind	PA	CF	LW	ST	MF	GS
27.05	31.98	5.29	7.14	3.15	10.09	0.47	14.80
26.83	31.39	6.49	7.65	0.27	8.21	0.74	18.09
28.33	35.01	6.29	6.23	0.60	8.25	0.63	14.67
25.57	27.69	8.89	11.69	2.08	12.48	0.26	11.34
20.69	32.36	12.82	9.91	0.17	10.64	1.18	12.23
22.94	36.21	7.10	6.30	1.71	10.76	0.21	14.76
25.30	33.51	11.21	7.62	1.30	0.45	0.66	19.97
23.40	36.51	6.81	7.16	0.49	10.42	0.56	14.65
27.81	29.90	13.43	9.38	0.89	11.08	0.23	7.27
22.01	27.52	15.11	8.57	0.16	10.90	0.43	15.29
37.97	24.40	7.93	4.56	0.92	11.01	0.10	13.10
	02 Res 27.05 26.83 28.33 25.57 20.69 22.94 25.30 23.40 27.81 22.01 37.97	02 03 Res Ind 27.05 31.98 26.83 31.39 28.33 35.01 25.57 27.69 20.69 32.36 22.94 36.21 25.30 33.51 23.40 36.51 27.81 29.90 22.01 27.52 37.97 24.40	02 Res 03 Ind 04 PA 27.05 31.98 5.29 26.83 31.39 6.49 28.33 35.01 6.29 25.57 27.69 8.89 20.69 32.36 12.82 22.94 36.21 7.10 25.30 33.51 11.21 23.40 36.51 6.81 27.81 29.90 13.43 22.01 27.52 15.11 37.97 24.40 7.93	02 Res 03 Ind 04 PA 05 CF 27.05 31.98 5.29 7.14 26.83 31.39 6.49 7.65 28.33 35.01 6.29 6.23 25.57 27.69 8.89 11.69 20.69 32.36 12.82 9.91 22.94 36.21 7.10 6.30 25.30 33.51 11.21 7.62 23.40 36.51 6.81 7.16 27.81 29.90 13.43 9.38 22.01 27.52 15.11 8.57 37.97 24.40 7.93 4.56	02 Res03 Ind04 PA05 CF06 LW27.0531.985.297.143.1526.8331.396.497.650.2728.3335.016.296.230.6025.5727.698.8911.692.0820.6932.3612.829.910.1722.9436.217.106.301.7125.3033.5111.217.621.3023.4036.516.817.160.4927.8129.9013.439.380.8922.0127.5215.118.570.1637.9724.407.934.560.92	02 Res 03 Ind 04 PA 05 CF 06 LW 07 ST 27.05 31.98 5.29 7.14 3.15 10.09 26.83 31.39 6.49 7.65 0.27 8.21 28.33 35.01 6.29 6.23 0.60 8.25 25.57 27.69 8.89 11.69 2.08 12.48 20.69 32.36 12.82 9.91 0.17 10.64 22.94 36.21 7.10 6.30 1.71 10.76 25.30 33.51 11.21 7.62 1.30 0.45 23.40 36.51 6.81 7.16 0.49 10.42 27.81 29.90 13.43 9.38 0.89 11.08 22.01 27.52 15.11 8.57 0.16 10.90 37.97 24.40 7.93 4.56 0.92 11.01	02 Res 03 Ind 04 PA 05 CF 06 LW 07 ST 08 MF 27.05 31.98 5.29 7.14 3.15 10.09 0.47 26.83 31.39 6.49 7.65 0.27 8.21 0.74 28.33 35.01 6.29 6.23 0.60 8.25 0.63 25.57 27.69 8.89 11.69 2.08 12.48 0.26 20.69 32.36 12.82 9.91 0.17 10.64 1.18 22.94 36.21 7.10 6.30 1.71 10.76 0.21 25.30 33.51 11.21 7.62 1.30 0.45 0.66 23.40 36.51 6.81 7.16 0.49 10.42 0.56 27.81 29.90 13.43 9.38 0.89 11.08 0.23 22.01 27.52 15.11 8.57 0.16 10.90 0.43 37.97 24.40 7.93 4.56

Table 4. Statistics on the proportion of classified area in the city (%).

4. Discussion

The POI data, a popular kind of social media data, is convenient both in generation and application. One of the main reasons why the applications of POI are limited is that there are no attributes about spatial scale. To solve the key problem, this research puts forward the concept of POI spatialization. The POI samples are obtained randomly by categories, a relatively accurate ground area is determined by the distribution of area values. Then, the POI is converted to an object with spatial scale in the form of raster. Due to various categories and individual differences, 100 samples from each category cannot fully reflect the overall trend, though can basically ensure the needs of the research. There is no doubt that the spatialization method is cost-effective, especially when compared with acquiring building footprint data. Additionally, the research can be improved in the directions of adding sample categories, increasing the amount of samples and improving probability distribution.

Urban complexes refer to the urban buildings with convergent and intensive functions. More and more complexes derived from the urban complexes appear in the modern economic system. In the process of determining the categories of construction land regions, there is a widely occurring question that too many kinds of buildings are clustered and multiple functions are interrelated in a region. With the traditional land use regulations, such regions can hardly be grouped into a single category. So we introduce the types of commercial complexes and residential complexes, which helps many parcels have a more suitable mixed-use category. The design and evaluation of the urban complexes should be included in future urban plannings and classification studies.

5. Conclusions

Based on POI spatial location and occupation area data, this research adopts raster spatial clustering, block constraints, and other methods to identify and divide urban construction land use types. On the basis of retaining the diversity and richness of parcel information to the greatest extent, this research introduces inherent socioeconomic spatial characteristics of residential complexes and commercial complexes to optimize the mixed use land classification system of urban construction land. The results of research show the following:

The POI dataset have accurate location information and rich usage type information. After cleaning, its data quality, data size, and excellent coupling ability with other data ensure its extremely high availability. Spatialization based on the statistics of the ground area helps the POI data better reflect the industrial distribution and urban layout.

The spatial classification method of construction land using spatial clustering and block constraints can ensure the natural constraints of the classification unit and fully exploit the spatial characteristics of the data. It is an effective method with strong applicability, accuracy and efficiency, and portability. The accuracy of the overall obfuscation matrix is 76.86%, which is greatly improved compared with that of the EULUC, and the calculated proportions do not deviate substantially from official statistics of all land-use types.

The spatial distribution of construction land-use types complements and improves overall urban planning and social and economic statistics, and the high-precision classification results can be used to better analyze the land-use structure. Residential complexes and commercial complexes are densely distributed in the central city regions, while industrial land is mostly distributed in the periphery of the cities, showing obvious marginal clustering, which can promote further research. Furthermore, rationality of functional zones should be paid more attention in the urban management because inappropriate plans can lead to heavy traffic, poor environmental conditions, and lack of emergency response.

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

Funding: This research was funded by the National Key Research and Development Program of China, grant number 2017YFA0604903 and 2017YFA0604904, which are funded by the Ministry of Science and Technology of the People's Republic of China.

Data Availability Statement: The collection and preprocessing of data are in Section 2.1. The data used are primarily reflected in the article. Other relevant data are available from the authors upon request.

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



Appendix A

Figure A1. Sampling area distribution of industrial land after logarithm.



Figure A2. Box plots.

References

- 1. Greenberg, M.; Goldstein, B.D.; Anderson, E.; Dourson, M.; Landis, W.; North, D.W. Whither Risk Assessment: New Challenges and Opportunities a Third of a Century After the Red Book. *Risk Anal.* **2015**, *35*, 1959–1968. [CrossRef] [PubMed]
- 2. The World Bank. Available online: https://documents.worldbank.org/en/publication/documents-reports/documentdetail/22 9791468159607825/economics-of-coastal-zone-adaptation-to-climate-change (accessed on 15 January 2021).
- Wu, J.; Han, G.; Zhou, H.; Li, N. Economic development and declining vulnerability to climate-related disasters in China. *Environ. Res. Lett.* 2018, 13, 034013. [CrossRef]
- 4. Adger, W.N. Vulnerability. Glob. Environ. Chang. 2006, 16, 268-281. [CrossRef]
- 5. Li, G.; Sun, S.; Fang, C. The varying driving forces of urban expansion in China: Insights from a spatial-temporal analysis. *Landsc. Urban Plan.* **2018**, *174*, 63–77. [CrossRef]
- 6. Neuhaus, F. Emergent Spatio Temporal Dimensions of the City; Springer: Cham, Switzerland, 2015; pp. 13–36. [CrossRef]
- 7. Foody, G.M. Status of land cover classification accuracy assessment. *Remote Sens. Environ.* 2002, 80, 185–201. [CrossRef]
- 8. Copernicus Global Land Services. Available online: https://land.copernicus.eu/global/products/lc (accessed on 20 March 2021).
- 9. Xie, H. Towards Sustainable Land Use in China: A Collection of Empirical Studies. *Sustainability* **2017**, *9*, 2129. [CrossRef]
- 10. Zhong, Y.; Su, Y.; Wu, S. Open-source data-driven urban land-use mapping integrating point-linepolygon semantic objects: A case study of Chinese cities. *Remote Sens. Environ.* **2020**, *247*, 111838. [CrossRef]
- 11. Soliman, A.; Soltani, K.; Yin, J. Social sensing of urban land use based on analysis of twitter users' mobility patterns. *PLoS ONE* **2017**, *12*, e0181657. [CrossRef]
- 12. McKenzie, G.; Janowicz, K.; Gao, S. Poi pulse: A multi-granular, semantic signature-based information observatory for the interactive visualization of big geosocial data. *Cartographica* 2015, *50*, 71–85. [CrossRef]
- 13. Gong, P.; Chen, B.; Li, X. Mapping essential urban land use categories in china (EULUC-China): Preliminary results for 2018. *Sci. Bull.* **2020**, *65*, 182–187. [CrossRef]
- 14. Hu, T.; Yang, J.; Li, X.; Gong, P. Mapping Urban Land Use by Using Landsat Images and Open Social Data. *Remote Sens.* **2016**, *8*, 151. [CrossRef]
- 15. Li, X.; Hu, T.; Gong, P.; Du, S.; Chen, B.; Li, X.; Dai, Q. Mapping Essential Urban Land Use Categories in Beijing with a Fast Area of Interest (AOI)-Based Method. *Remote Sens.* **2021**, *13*, 477. [CrossRef]
- 16. Yao, Y.; Li, X.; Liu, X.; Liu, P.; Liang, Z.; Zhang, J.; Mai, K. Sensing spatial distribution of urban land use by integrating points-of-interest and google word2vec model. *Int. J. Geogr. Inf. Sci.* **2017**, *31*, 825–848. [CrossRef]
- 17. Zhai, W.; Bai, X.; Shi, Y. Beyond word2vec: An approach for urban functional region extraction and identification by combining place2vec and pois. *Comput. Environ. Urban Syst.* **2019**, *74*, 1–12. [CrossRef]
- Gao, S.; Janowicz, K.; Couclelis, H. Extracting urban functional regions from points of interest and human activities on locationbased social networks. *Trans. GIS* 2017, 21, 446–467. [CrossRef]
- Yuan, J.; Zheng, Y.; Xie, X. Discovering Regions of Different Functions in a City Using Human Mobility and POIs. In Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Beijing, China, 12–16 August 2012; Association for Computing Machinery: New York, NY, USA, 2012. [CrossRef]
- 20. Gu, Y.; Jiao, L.; Dong, T. Spatial distribution and interaction analysis of urban functional areas based on multi-source data. *Geomat. Inf. Sci. Wuhan Univ.* **2018**, 43, 1113–1121. (In Chinese) [CrossRef]
- 21. Xu, S.; Qing, L.; Han, L.; Liu, M.; Peng, Y.; Shen, L. A New Remote Sensing Images and Point-of-Interest Fused (RPF) Model for Sensing Urban Functional Regions. *Remote Sens.* 2020, *12*, 1032. [CrossRef]
- 22. Song, J.; Lin, T.; Li, X.; Prishchepov, A.V. Mapping Urban Functional Zones by Integrating Very High Spatial Resolution Remote Sensing Imagery and Points of Interest: A Case Study of Xiamen, China. *Remote Sens.* **2018**, *10*, 1737. [CrossRef]
- 23. Sun, J.; Wang, H.; Song, Z.; Lu, J.; Meng, P.; Qin, S. Mapping Essential Urban Land Use Categories in Nanjing by Integrating Multi-Source Big Data. *Remote Sens.* 2020, *12*, 2386. [CrossRef]
- 24. Zhou, Y.; Smith, S.J.; Elvidge, C.D.; Zhao, K.; Thomson, A.; Imhoff, M. A cluster-based method to map urban area from DMSP/OLS nightlights. *Remote Sens. Environ.* **2014**, 147, 173–185. [CrossRef]
- 25. Schwarz, N. Urban form revisited—Selecting indicators for characterising European cities. *Landsc. Urban Plan.* **2010**, *96*, 29–47. [CrossRef]
- Ester, M.; Kriegel, H.; Sander, J. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96), Portland, OR, USA, 2–4 August 1996; AAAI Press: Portland, OR, USA, 1996.
- 27. Schubert, E.; Sander, J.; Ester, M. Dbscan revisited, revisited: Why and how you should (still) use dbscan. *ACM Trans. Database Syst.* **2017**, *42*, 1–21. [CrossRef]
- 28. Shen, J.; Hao, X.; Liang, Z. Real-time superpixel segmentation by dbscan clustering algorithm. *IEEE Trans. Image Process.* **2016**, *25*, 5933–5942. [CrossRef] [PubMed]
- 29. Ministry of Housing and Urban-Rural Development of People's Republic China. *Code for Classification of Urban and Rural Land Use and Planning Standards of Development Land (GB50137-2011);* China Architecture & Building Press: Beijing, China, 2011; (In Chinese).
- 30. Elvidge, C.D.; Baugh, K.E.; Zhizhin, M.; Hsu, F.-C. Why VIIRS data are superior to DMSP for mapping nighttime lights. In Proceedings of the Asia-Pacific Advanced Network, Honolulu, HI, USA, 13–16 January 2013.

- Gong, P.; Li, X.; Wang, J. Annual maps of global artificial impervious area (gaia) between 1985 and 2018. *Remote Sens. Environ.* 2020, 236, 111510. [CrossRef]
- 32. Gong, P.; Wang, J.; Yu, L. Finer resolution observation and monitoring of global land cover: First mapping results with Landsat TM and ETM+ data. *Int. J. Remote Sens.* 2013, 34, 2607–2654. [CrossRef]
- 33. Ministry of Natural Resources of the People's Republic of China. *Current Land Use Classification (GB/T 21010-2017)*; Standards Press of China: Beijing, China, 2017; (In Chinese).
- 34. Liu, X.; Long, Y. Automated identification and characterization of parcels with openstreetmap and points of interest. *Environ. Plan. B Plan. Des.* **2016**, *43*, 341–360. [CrossRef]
- 35. Zhang, L. Research on poi classification standard. Bull. Surv. Mapp. 2012, 10, 82–84. (In Chinese)
- 36. Xu, H.; Zhu, J.; Wang, Z. Exploring the spatial pattern of urban block development based on poi analysis: A case study in Wuhan, China. *Sustainability* **2019**, *11*, 6961. [CrossRef]
- 37. Li, X.; Xue, X. Estimation method of nighttime light images'electric power consumption based on the boston matrix. *Geomat. Inf. Sci. Wuhan Univ.* **2018**, *43*, 1994–2002. (In Chinese)
- Li, F.; Yan, Q.; Zou, Y.; Liu, B. Improving the Accuracy of Built-up Area Extraction Based on Nighttime Light Data and POI: A Case Study of NPP/VIIRS and Luojia 1-01 Nighttime Light Images. Wuhan Univ. Nat. Sci. 2019, 1–14. (In Chinese) [CrossRef]
- 39. Silverman, B.W. Density Estimation for Statistics and Data Analysis. Technometrics 2012, 29, 495. [CrossRef]
- 40. Zhou, Z.; Wang, J.; Zhu, S. An improved adaptive and fast af-dbscan clustering algorithm. *CAAI Trans. Intell. Syst.* **2016**, *11*, 93–98. (In Chinese)
- 41. CCCC FIRST HIGHWAY CONSULTANTS. *The Highway Route Design Specification (JTG D20-2017);* China Communication Press: Beijing, China, 2018. (In Chinese)
- 42. Hu, M.Z.; Fan, Y.H.; Zhao, Q.Q. A Study of HOPSCA Development under the Intensive Land Use. *Appl. Mech. Mater.* 2011, 71, 589–593. [CrossRef]
- 43. Qin, Z.; Yu, Y.; Liu, D. The effect of hopsca on residential property values: Exploratory findings from Wuhan, China. *Sustainability* **2019**, *11*, 471. [CrossRef]
- 44. Jiang, H. Commercial Complex and Urban Public Space. Mod. Urban Res. 2009, 24, 48–52. (In Chinese)