

Article

Sampling Strategy for Detailed Urban Land Use Classification: A Systematic Analysis in Shenzhen

Mo Su ¹, Renzhong Guo ^{1,2,*}, Bin Chen ³, Wuyang Hong ^{1,4}, Jiaqi Wang ⁴, Yimei Feng ⁴ and Bing Xu ^{5,6,7}

¹ School of Resource and Environment Science, Wuhan University, Wuhan 430079, China; 2014102050019@whu.edu.cn (M.S.); 2015102050020@whu.edu.cn (W.H.)

² Research Institute for Smart Cities, Shenzhen University, Shenzhen 518060, China

³ Department of Land, Air and Water Resources, University of California, Davis, CA 95616-8627, USA; bch@ucdavis.edu

⁴ Shenzhen Urban Planning and Land Resource Research Center, Shenzhen 518034, China; 2014302590136@whu.edu.cn (J.W.); fengym@pnr.sz.gov.cn (Y.F.)

⁵ Ministry of Education Key Laboratory for Earth System Modeling, Department of Earth System Science, Tsinghua University, Beijing 100084, China; bingxu@tsinghua.edu.cn

⁶ Tsinghua Urban Institute, Tsinghua University, Beijing 100084, China

⁷ Center for Healthy Cities, Institute for China Sustainable Urbanization, Tsinghua University, Beijing 100084, China

* Correspondence: guorz@szu.edu.cn

Received: 3 April 2020; Accepted: 1 May 2020; Published: 8 May 2020



Abstract: A heavy workload is required for sample collection for urban land use classification, and researchers are in urgent need of sampling strategies as a guide to achieve more effective work. In this paper, we make use of an urban land use survey to obtain a complete sample set of a city, test the impact of different training and validation sample sizes on the accuracy, and summarize the sampling strategy. The following conclusions are drawn based on our systematic analysis in Shenzhen. (1) For the best classification accuracy, the number of training samples should be no less than 40% of the total number of parcels or no less than 5500 parcels. For the best labor cost performance, the number should be no less than 7% or no less than 900. (2) The accuracy evaluation is stable and reliable and requires validation sample numbers of no less than 10% of the total or no less than 1200. (3) Samples with a purity of 60–90% are preferred, and the classification effectiveness is better in samples with a purity greater than 90% under the same number. (4) If spatial equilibrium sampling cannot be carried out, sampling areas with complex land use patterns should be preferred.

Keywords: land use classification; field survey; samples; parcel segmentation; machine learning; land use mapping

1. Introduction

Urbanization has greatly changed our living environments, and more than half of the global population resides in urban areas [1]. China has undergone the fastest urbanization worldwide over the past three decades, and its artificial impervious area ranked first in 2015 [2]. For better urban planning, spatial governance, and sustainable development of urbanized areas in China, more up-to-date, detailed, and accurate land use classification is critically important.

Thus far, detailed urban land use classification in China has been performed only through field surveys [3,4]. Currently, only a few major cities, such as Shenzhen, Wuhan, and Chongqing, have detailed urban land use classifications at the entire city level [3,5–7]. This is an important task for the Third Terrestrial Survey of China [8].

Field surveys are time consuming and laborious, and researchers have long been committed to improving the efficiency of land use classification through remote sensing technology [9–20]. Gong and his colleagues were among the earliest researchers to use spatial-context information in addition to spectral data from satellite images to map urban land use categories, and their algorithms have been adopted in mapping global settlement areas [21]. However, because of the limitation of physical property measurements, the above-mentioned methods involving only spectral, texture, and structural features face challenges in effectively differentiating among residential, industrial, commercial, and service types of land uses.

In 2000, Zhang et al. proposed conducting urban land use classification by integrating GIS and remote sensing data [22]. In 2007, Goodchild noted that volunteered geographic information (VGI) can be used as a new data source for urban land use classification [23]. Information from OpenStreetMap (OSM), point of interest (POI), and social data, such as traffic trace data of individuals, taxis, and public transportation, can all be applicable to urban land use mapping [24–30]. VGI can be used as an important supplement to remotely sensed data in the detailed mapping of urban land use [31] and has since become a new focus area of research [32–40]. The most influential work was the mapping of essential urban land use categories (EULUC) in all cities in China by 70 researchers from more than 30 organizations [40].

Because it is impossible to determine the classification results simply through visual interpretation of images, the difficulty and workload of sample collection are increasing exponentially, representing a difficult challenge for most researchers. Researchers are in urgent need of sampling strategies as a guide to achieve more effective classification with relatively low labor costs. In the field of traditional land use/land cover, scholars have accumulated a large number of samples over a long time and quantitatively analyzed the impact of the sample number and other conditions on classification accuracy [41–45]. However, detailed urban land use classification is a new research focus; most studies use a limited number of sample units to test experimental classification methods, and no research results regarding the optimal sampling strategies have been reported [31,34,36,40,46–48].

In this study, we take advantage of the availability of an urban land use map of Shenzhen city that has been generated through a field survey of the entire city. By converting the map into a parcel-based land use map, we obtain a complete sample set for experiments with various sample sizes. Based on this map, we evaluated the impact of the sample size and land use mix of samples on the resulting classification accuracy.

2. Study Site and Method

2.1. Study Area and Data

Shenzhen is the most rapidly developing city in China. In 1979, Shenzhen was essentially a rural county bordering Hong Kong (Figure 1). By 2019, Shenzhen had more than 13 million permanent residents, and its per capita gross domestic product (GDP) ranked first in China [49–51]. Due to the high diversity and high precision of urban land use, complex land use types exist, such as villages surrounded by city blocks, golf courses, and large entertainment facilities. The high level of complexity and high land use intensity in Shenzhen provide a good opportunity for detailed urban land use classification experiments.



Figure 1. The study area.

2.2. Technical Process

Figure 2 shows a flowchart outlining the methodology used in this study, including the following four major procedures: first, parcel segmentation with road networks, water, and impervious layers; second, collection of training and validation samples; third, multisource feature extraction; and fourth, classification and mapping. All datasets used in this study are summarized in Table 1.

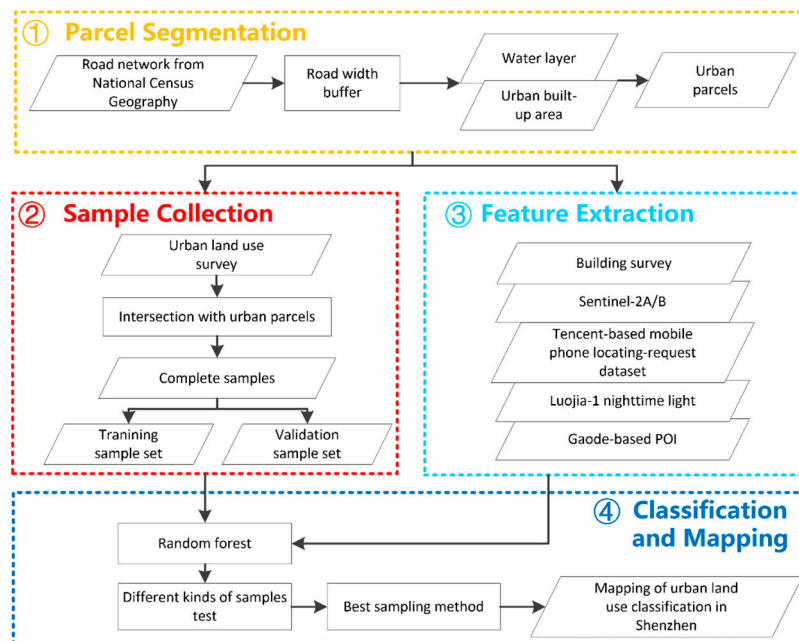


Figure 2. Flowchart of the research methods.

Table 1. Research data.

Data Type	Name	Year	Content	Resolution	Usage
Remote Sensing	Sentinel 2A/B	2018	Spectral data	10 m	Feature extraction
	Luojia 1	2018	Night light	130 m	Feature extraction
VGI	OSM	2018	Centerline	Multiple-scale	Parcel segmentation
	Gaode Map	2018	POI	Multiple-scale	Feature extraction
	Tencent social big data	2018	Tencent mobile-phone locating-request	1000 m	Feature extraction
Field Survey	Urban land use survey	2018	Urban land use parcel	Submeter	Sample collection
	National geographical condition survey	2015	Centerline, road level, width, water area	Submeter	Parcel segmentation
	Road survey	2018	Centerline, width	Submeter	Parcel segmentation
	Building survey	2018	Building base map and height	Submeter	Feature extraction
	Extent of built-up area	2016	Built-up area	Submeter	Parcel segmentation

2.3. Detailed Urban Land Use Classification System

In 2007, China issued the first formal land use classification standard, which was revised in 2017. This standard includes residential land, commercial and service land, industrial and mining storage land, public administration and public service land, and transportation land [52]. The city of Shenzhen developed a local classification system to supplement the national system [53]. In this study, based on the national and Shenzhen classification schemes, we develop the Shenzhen Urban Land Use Classification system (SULUC), which includes 5 Level I classes and 18 Level II classes (Table 2). The SULUC is basically consistent with the standard used in the Third Terrestrial Survey of China, and some Level II classes are even more detailed.

Table 2. Descriptions of the Shenzhen Urban Land Use Classification system (SULUC).

Level I	Level II	Descriptions
01 Residential	0101 Urban residential (UR)	Land used for residential housing and related facilities
	0102 Urban village (UV)	Original rural resident housing currently mostly surrounded by city blocks
02 Commercial	0201 Business and finance (BF)	Commercial and service land used for business operations
	0202 Recreational (Rec)	Cinema, recreational land, tourism land with less than 65% coverage of green space
	0203 Golf course (GC)	Golf course and service housing and facilities
	0204 Storage (Sto)	Land used for stockpiling and temporary storage for distribution
	0205 Other commercial (OC)	Retail, wholesale, production and sales, services, and entertainment land
03 Industrial	0301 Industrial (Ind)	Land used for production, product processing, manufacturing, machine repair, and other related facilities
04 Transportation	0401 Road (Roa)	Transportation land
	0402 Stations (Sta)	Land used for service facilities, such as stations, transfer stations, parking facilities
	0403 Airports (Air)	Civil or military airports
	0404 Harbor (Har)	Land used for harbors or related facilities
05 Public service	0501 Governmental office, media, and press (GO)	Land used for governmental offices, social organizations, broadcasting, film making, and publishing agencies
	0502 Instructional and research (IR)	Land used for instruction, research, design, surveying, testing, environmental assessment, extension, etc.
	0503 Medical, health, and social welfare (MH)	Land used for medical, healthcare, disease control, recovering, emergent saving facilities, philanthropic institution, etc.
	0504 Sports and cultural facilities (SC)	Public stadiums and training facilities
	0505 Parks and green space (PG)	Parks, zoos, gardens, squares, and other green space for recreation
	0506 Public infrastructure (PI)	Land used for public infrastructure

2.4. Parcel Segmentation

We used the road network from the 2018 special road survey to divide the Shenzhen area into land parcels using the following major procedures: first, a road buffer was generated using the road centerline and width; second, the road buffer zone was used to divide Shenzhen area into land parcels; third, water surface data from the National Geographical Condition Survey were used to exclude parcels of water; fourth, parcels within the built-up area were extracted, and the purpose of this step was to exclude farmland, forestland, bare land, and other categories that do not belong to SULUC. In fact, there were approximately 200 parcels that did not belong to SULUC in the built-up area, accounting for approximately 2% of the total number of parcels, which had little impact on the overall accuracy of classification. The land parcels were divided into 12,965 land parcels (Figure 3).

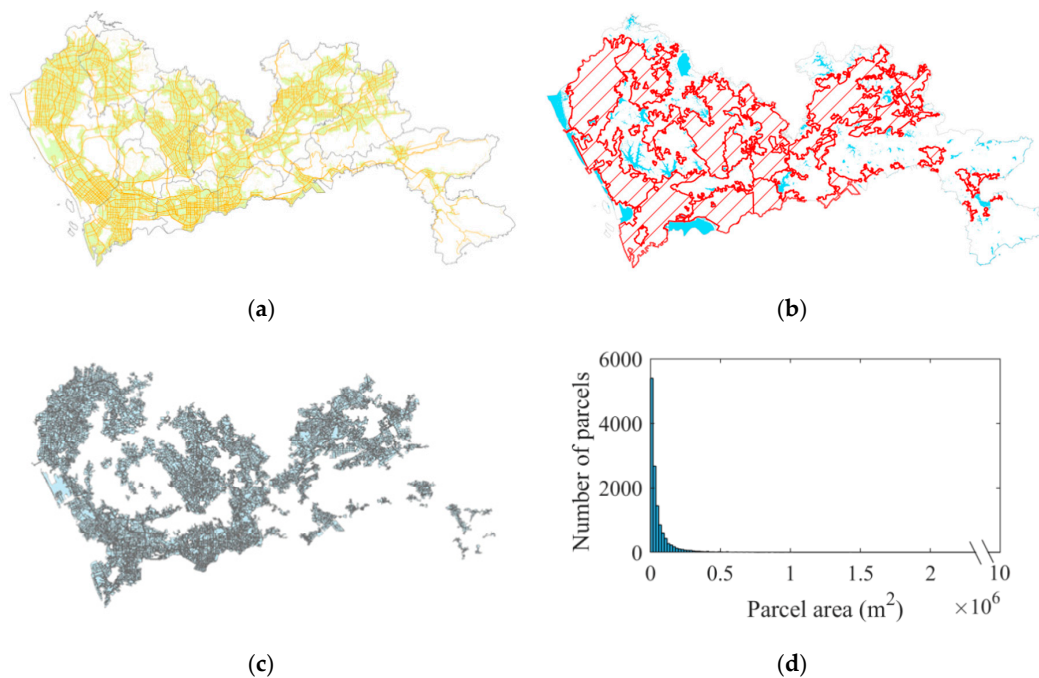


Figure 3. Maps of parcel segmentation and a histogram of the parcel area: (a) road network; (b) built-up area and water; (c) map of parcels; (d) histogram.

The average size of a parcel was approximately 6 ha, which was approximately four times greater than the land parcel size in the field survey. More than 100 parcels were superlarge land parcels exceeding 50 ha. These superlarge parcels included villages in cities and large tracts of factories with no obvious roads (Figure 4). These areas were located in the less developed part of the city, and using the road network-based land partition method as a quick land partition strategy should be improved in the future.

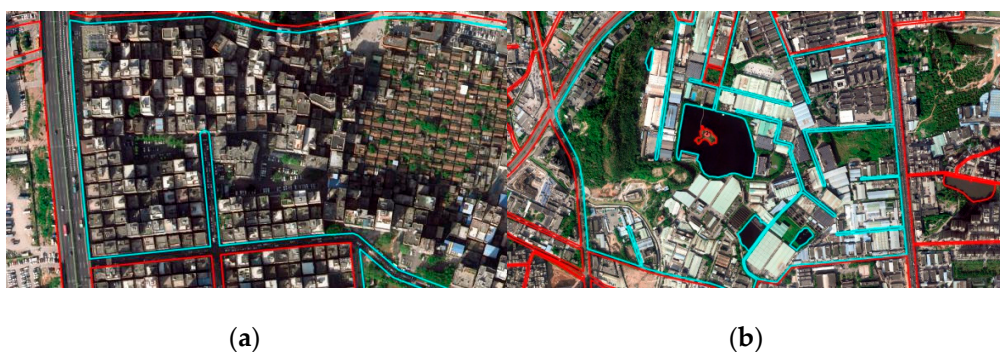


Figure 4. Examples of superlarge parcels: (a) village surrounded by the city; (b) large tracts of factories.

2.5. Feature Extraction

We used the following five types of features in the parcel-level land use classification based on Sentinel-2A/B images, Tencent mobile-phone locating-request (MPL) data, LuoJia-1 nighttime light images, Gaode POI data [40,54], and building surveys:

2.5.1. Multispectral Features from Sentinel-2A/B Imagery

We used the constellation Sentinel-2A/B images from January 1 to December 31, 2018, from the Copernicus Open Access Hub to extract the multispectral features. We first calculated the normalized difference vegetation index (NDVI) of each pixel. We further used the pixel-based maximum NDVI values as a quality index to merge the whole-year images. Then, we calculated the mean and standard deviations of the blue, green, red, and near-infrared bands, NDVI, and normalized difference water index (NDWI) in each urban parcel.

2.5.2. Human Activity Features from Tencent MPL Data

We used the MPL dataset from November 1 to November 30, 2018, from Tencent, Inc. to track the dynamics of the population distribution. MPL records are produced by retrieving the real-time locations of active mobile-phone users as they use Tencent's location-based services (LBS). We aggregated the 5 min MPL records per 8 h on weekdays and weekends, which represented the geographic pattern of the human distribution during three temporal periods (12 a.m.–8 a.m., 8 a.m.–4 p.m., and 4 p.m.–12 a.m.).

2.5.3. Nighttime Light Features from LuoJia-1 Nighttime Light Imagery

We used LuoJia-1 nighttime light images acquired from June to December 2018, and the spatial resolution of these images was 130 m. For each urban parcel, we calculated the mean value of the digital number.

2.5.4. POI Features from Gaode

We used POI data from Gaode, Inc. in 2018. Each POI record consists of the name, location coordinates, and POI type, such as catering, retailing, automobile, accommodation, recreation, public facility, transportation, culture and media, and so forth. For each urban parcel, we calculated the total number of all POI and the total number and proportion of each type of POI within that parcel.

2.5.5. Building Features from Survey Data

We used building survey data consisting of the base area, stories, and average story height of each building in Shenzhen. We further aggregated these data into parcel levels to calculate the number of stories, the sum of the building height, and the average building height.

The specific features are summarized in Table 3.

Table 3. Summary of the features used in the parcel-level mapping of SULUC.

Data Source	Features
Sentinel-2A/B	Mean and standard deviation of blue, green, red, near-infrared bands, NDVI, and NDWI
Tencent-based MPL	Mean of 8 h active population during weekdays and weekends
LuoJia-1 nighttime light	Mean of digital number values
Gaode-based POI	Total number of all POI and total number and proportion of each type of POI within each parcel
Building survey	Number of stories, sum of building height, and averaged building height

2.6. Training and Validation Samples

Since the land survey data covering the entire city of Shenzhen are accessible, we possessed an accurate reference dataset for training and validating the sample collection. Quality assurance of the

field survey data was determined following a procedure of in situ photographing and by interviewing the land managers to record the condition of the land use operation. The data were sample-verified and quality-checked by a series of indoor processes to ensure that the results were consistent with field survey standards. Therefore, the field surveyed land use served as a reliable source of reference in this study.

Because parcels resulting from field surveys differ from parcels resulting from segmentation, within each land parcel, we obtained the statistics of the areal proportion of different land use types through a spatial intersect operation with the GIS software system. The land use category with the largest proportion was assigned to the land parcel (Table 4). A sliver polygon removal operation was applied to polygons less than 1000 m² in area.

Table 4. Land use category extraction from the field survey.

		UR	BF	Sto	OC	Ind	Roa	Sta	IR	PG	PI	Primary Land Use
Parcel Code	F00003	0%	0%	0%	0%	7%	17%	0%	0%	76%	0%	PG
	F00025	14%	3%	3%	3%	63%	5%	4%	5%	0%	0%	Ind
	F00387	55%	0%	0%	0%	7%	17%	0%	0%	0%	21%	UR

Through the above operation, we obtained a complete coverage reference land use dataset with proportional records of different land use types. An advantage of this dataset is that all land parcels can be used for training or validation. Therefore, we refer to this reference sample set as complete samples, and the number of parcels in each category is shown in Figure 5. Under the complete samples, the accuracy of the sample is equivalent to that of the field survey.

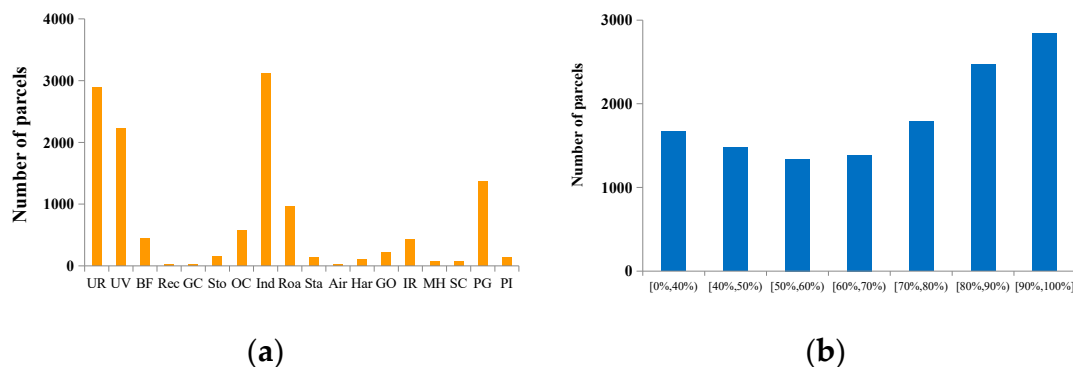


Figure 5. Number distribution of parcels: (a) histogram of parcels with different categories; (b) histogram of parcels with different purities.

The complete samples can reflect the land use mixing status. We used purity to quantify the land mixed-use level of the parcel. The higher the purity of the parcel, the lower the mixing level of land use. In the complete sample, we started with 100% purity and divided it into 10 groups according to each 10% decrease and combined 0–40% into one group. The number of each group is shown in Figure 5.

2.7. Classifier

Since 1996, machine learning has been widely used in the field of remote sensing classification. Many scholars have found that machine learning can obtain results with a higher precision than traditional parameter classifiers in processing complex data with a high-dimensional feature space [47,55–58]. In particular, random forest (RF) is widely used by scholars. RF is a machine-learning algorithm consisting of a large ensemble of regression trees that has shown great efficiency and robustness in both computational cost and model performance [46–48]. We applied the training parcels with the extracted features to produce a parcel-level mapping of urban land use classification in Shenzhen with RF.

3. Experimental Tests and Results

3.1. The Impact of the Sample Size

We set up two experiments. The first experiment tested the influence of different training sample sizes on accuracy. From the complete sample, 30% of the stratified random sampling was used as validation samples, and the remaining samples were used as training samples. The number of training samples decreased by 1% each time, and each decrease repeated randomly sampled k times. The second experiment tested the influence of the different validation sample sizes on the accuracy evaluation. From the complete sample, 35% of the stratified random samples were used as training samples, and the remaining samples were used as validation samples. The number of validation samples decreased by 1% each time, and each decrease repeated randomly sampled k times. For $k = 5$, the accuracy of each classification and the average accuracy are shown in Figure 6.

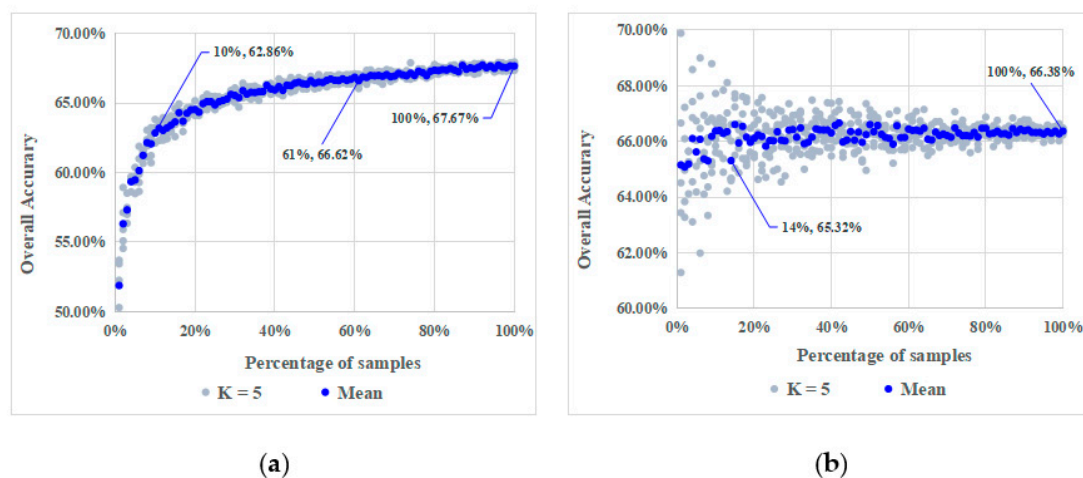


Figure 6. Impact of the sample size on accuracy: (a) impact of different training sample sizes on the classification accuracy; (b) impact of different validation sample sizes on the classification accuracy.

We define stable accuracy as a classification accuracy of the reduced samples no greater than 1% compared with that of all samples. Experiment One shows that the relationship between the number of samples and accuracy follows the rule of stable classification with limited samples (Gong, Liu, et al., 2019). The classification accuracy kept stable until the number of training samples was reduced to 61% of all training samples (5540, accounting for 40% of all urban parcels). When the number was reduced to 10% (908, approximately 7% of all urban parcels), the classification accuracy began to significantly decline.

Experiment Two shows that as the number of validation samples decreases, the range of the accuracy evaluation results increases. Considering the average accuracy as the measurement, when the number of validation samples was reduced to 14% of all validation samples (1178, approximately 9% of all urban parcels), the accuracy evaluation results were no longer stable.

In summary, to obtain stable and reliable classification results, the training samples need at least 40% of the total number of parcels or no less than 5500. At least 10% of the total number of parcels is required for the validation samples or no less than 1200. If the labor force is insufficient, the high-cost performance scheme requires the training samples to be at least 7% of all parcels or no less than 900. In this situation, the maximum accuracy loss was not greater than 7%.

3.2. Impact of the Sample Purity

In this experiment, the influence of the sample purity on the classification accuracy was tested. Currently, in most research concerning urban land use classification, the level of mixed land use is not high, and the training samples always have high purity [31,39,40]. The mixed-use level of land in

Shenzhen is high, and there are many low-purity parcels. Therefore, it is necessary to study whether it is reasonable to select high-purity samples as training samples (Figure 7).

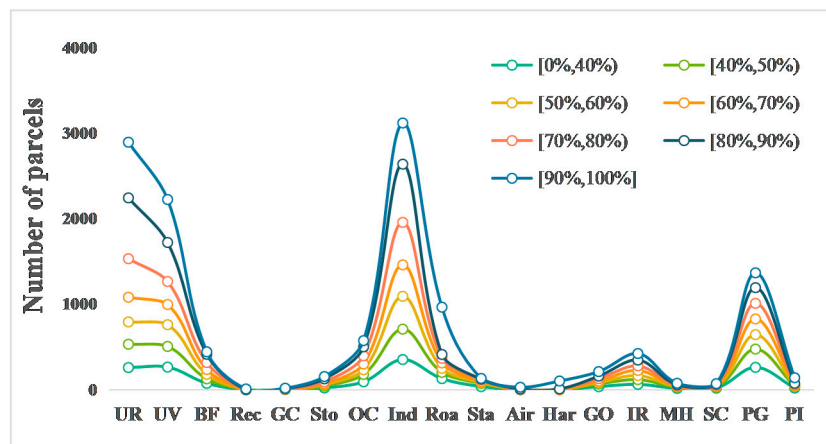


Figure 7. Sample numbers of different urban land use classifications with different purities.

We selected seven categories of 11,034 parcels for the test. The specific categories included urban residential, urban village, business and finance, storage, other commercial, industrial, instructional and research, parks and green space.

Among them, 30% of the stratified random sampling was used as validation samples, and the remaining samples were used as the mixed-purity [0,100%] sample set. Then, we divided the mixed-purity set into high purity ($\geq 90\%$), medium purity (60–90%), and low purity ($\leq 60\%$). Finally, we randomly selected the same number of training samples from the above four sets, and the results are shown in Figure 8.

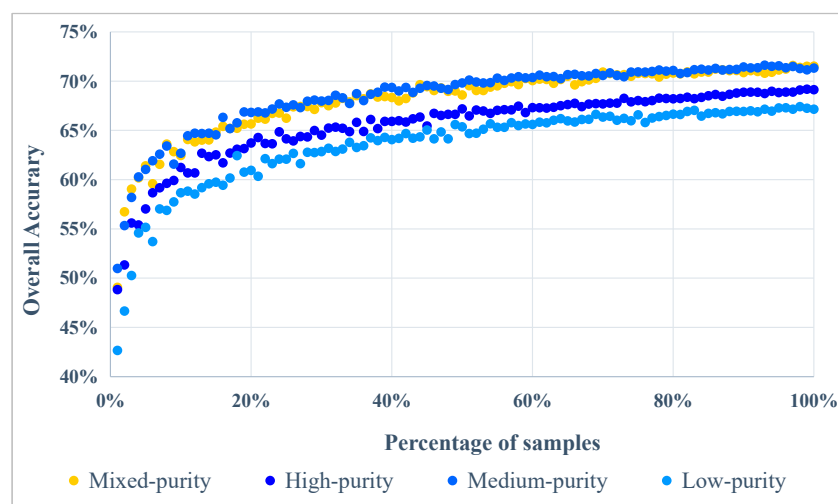


Figure 8. Overall accuracy of samples with different purities.

The experimental results show that under the same number of conditions, the classification accuracy of the mixed-purity samples was equal to that of the medium-purity samples and higher than that of the high-purity samples. The classification accuracy of the low-purity samples was the lowest. These results show that for a study area with a high land use mixing level, the representativeness of high-purity samples is not enough, which could lead to accuracy loss. The classification features of the low-purity samples are all mixed; thus, it is difficult for the classifier to learn effectively. The classification effect of the medium-purity samples is representative and can be used as the principle of sample collection.

3.3. Impact of the Sample Spatial Distribution

In this experiment, the influence of the sample space distribution on accuracy was tested. We divided Shenzhen into three zones: the original special zone, former Bao'an, and former Longgang. The original special zone included Luohu District, Futian District, Nanshan District, and Yantian District. Former Bao'an included current Bao'an District, Longhua District, and Guangming District. Former Longgang included the current Longgang District, Pingshan District, and Dapeng District. The same numbers of training and validation samples were randomly selected from the three regions for the cross experiment, and the accuracy was calculated with the training samples from the original special zone, former Bao'an, former Longgang, and the validation samples from the three regions (Figure 9).

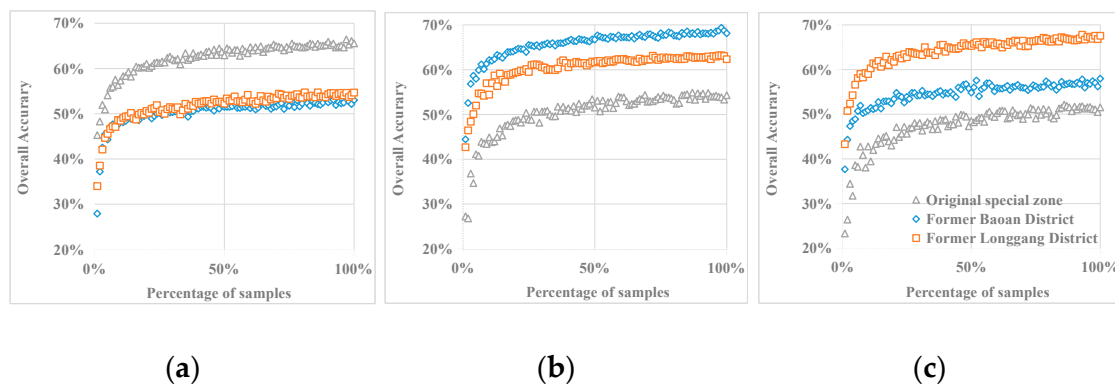


Figure 9. Classification accuracy of the sample migration experiments: (a) original special zone; (b) former Bao'an District; (c) former Longgang District.

The experimental results show that land use in different areas in a single city also has heterogeneity and that an uneven spatial distribution of samples could cause accuracy loss. In this experiment, the original special zone was the old special economic zone, which has good planning control and orderly land development. Former Bao'an is a labor-intensive industrial agglomeration area with inefficient and extensive land use. Former Longgang is restricted by ecological protection due to location factors, and its density is relatively low. There are differences in the representativeness of the three samples, and the classification accuracy of other areas is significantly reduced.

From the perspective of sample migration capacity, the more diverse the regional urban land use model, the stronger the migration capacity. In former Bao'an, Guangming is a relatively less developed area of Shenzhen, and Bao'an Qianhai center is the most important economic center. Therefore, multiple internal development stages coexist in former Bao'an, land use is extremely complex, and the migration capacity is strong. Due to the high level of overall urban development, the original special zone has low representativeness and a weak migration capacity.

3.4. Mapping of SULUC in Shenzhen

At the beginning, local professional urban land use surveyors were invited to choose training samples from the complete sample set according to their knowledge and experience. They generated 1163 high-purity samples. Four-fold cross-validation was adopted to optimize the land use classifier and the classifier was applied to the complete sample set for accuracy assessment. The overall accuracy for the Level I categories was 62%, and 55% for Level II categories. Then, we took the best sampling strategy in terms of the above-mentioned experiments and selected 5028 samples of medium purity as the training samples. Its frequency distribution was similar to that of the complete sample set (Figure 10). Using the same parcels, features, and classifier, the overall accuracy for Level I categories reached 76%, and that for Level II categories reached 71% (Tables 5 and 6). The accuracy was improved by approximately 15% under the optimal sampling strategy, shown in Figure 11.

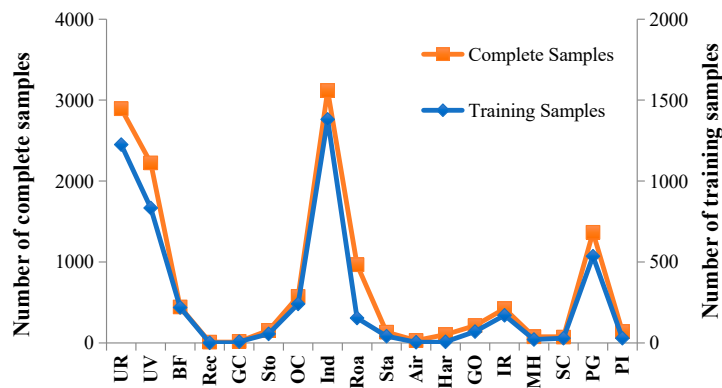


Figure 10. Frequency distribution of training samples and the complete sample set.

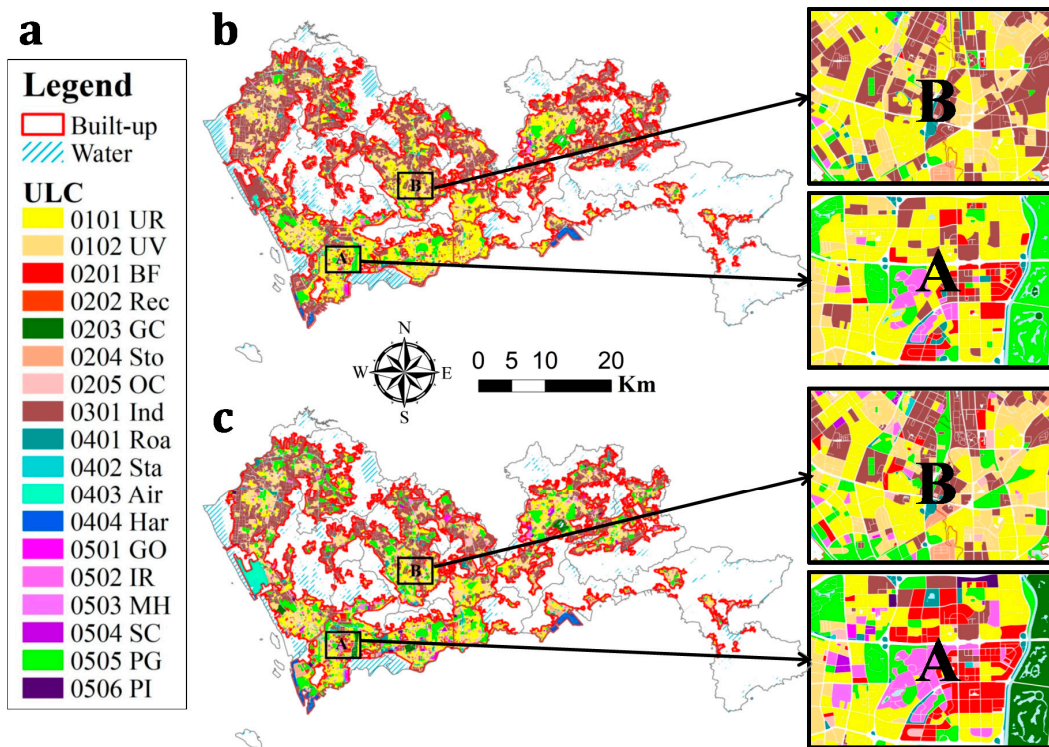


Figure 11. Mapping of SULUC: (a) legend; (b) mapping of SULUC; (c) complete samples.

Regarding Level I categories, major discrepancies were clustered in residential and industrial land, and the misclassification of other land use types to residential and industrial land accounted for over 50% of each of the misclassified categories. Regarding Level II categories, major discrepancies were clustered in the urban residential, industrial, and parks and green space land. For example, urban residential land was primarily misclassified as industrial land, industrial land was primarily misclassified as urban villages, and parks and green space land was primarily misclassified as urban residential, industrial, and road areas.

Table 5. Confusion matrix of SULUC Level I.

Level I	Overall Accuracy = 75.94%						Kappa = 66.24%		
	Residential	Commercial	Industrial	Transportation	Public Service	Total	PA (%)	UA (%)	
Residential	4748	23	259	40	48	5118	92.77%	76.59%	
Commercial	376	415	326	28	58	1203	34.50%	88.30%	
Industrial	459	10	2557	23	69	3118	82.01%	72.48%	
Transportation	82	7	97	930	116	1232	75.49%	72.54%	
Public service	534	15	289	261	1195	2294	52.09%	80.42%	
Total	6199	470	3528	1282	1486	12965			

Table 6. Confusion matrix of SULUC Level II.

Level II	Overall Accuracy = 70.91%															Kappa = 64.33%						
	UR	UV	BF	Rec	GC	Sto	OC	Ind	Roa	Sta	Air	Har	GO	IR	MH	SC	PG	PI	Total	PA	UA	
Residential	UR	2469	155	7	0	0	0	4	204	33	0	0	1	0	4	0	0	17	0	2894	85.31%	68.07%
	UV	173	1854	0	0	0	0	1	186	8	0	0	0	1	0	0	1	0	2224	83.36%	77.31%	
	BF	142	9	153	0	0	0	2	118	15	0	0	0	0	1	0	0	4	0	444	34.46%	86.44%
Commercial	Rec	1	0	0	7	0	0	0	0	1	0	0	0	0	0	0	1	0	10	70.00%	100.00%	
	GC	0	0	0	0	7	0	0	0	0	0	0	0	0	0	0	12	0	19	36.84%	100.00%	
	Sto	4	2	0	0	0	43	0	101	4	0	0	0	0	0	0	1	0	155	27.74%	97.73%	
Industrial	OC	120	58	6	0	0	1	160	210	14	0	0	0	0	1	0	0	5	0	575	27.83%	90.40%
	Ind	149	212	4	0	0	0	3	2706	28	0	0	0	0	7	0	0	9	0	3118	86.79%	65.82%
	Roa	53	5	1	0	0	0	0	74	801	0	0	2	0	1	0	0	28	1	966	82.92%	63.02%
Transportation	Sta	21	6	1	0	0	0	2	53	23	11	0	1	0	1	0	0	14	0	133	8.27%	100.00%
	Air	0	0	0	0	0	0	0	5	0	0	24	0	0	0	0	0	1	0	30	80.00%	100.00%
	Har	0	0	0	0	0	0	0	6	2	0	0	95	0	0	0	0	0	103	92.23%	95.96%	
Public service	GO	77	13	3	0	0	0	0	46	4	0	0	0	66	2	0	0	2	0	213	30.99%	100.00%
	IR	117	32	1	0	0	0	1	70	11	0	0	0	0	0	0	4	0	424	44.34%	81.74%	
	MH	26	6	1	0	0	0	0	14	4	0	0	0	0	2	24	0	0	77	31.17%	100.00%	
	SC	17	3	0	0	0	0	1	24	6	0	0	0	0	6	0	12	3	0	72	16.67%	100.00%
	PG	237	42	0	0	0	0	3	259	309	0	0	0	0	15	0	0	502	0	1367	36.72%	82.70%
Total	3627	2398	177	7	7	44	177	4111	1271	11	24	99	66	230	24	12	607	73	12965			

We compared the difference between the mapping of SULUC and land surveys in terms of the urban land use structure (Figure 12). Most commercial and public services lands are not correctly classified and are basically misclassified as residential and industrial, which is critical for improving accuracy in the future.

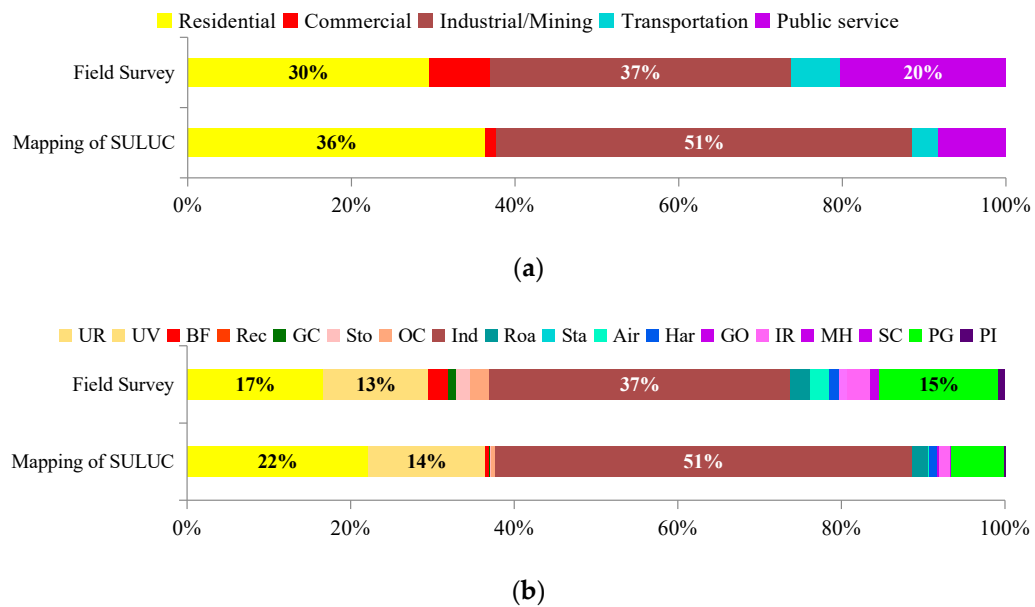


Figure 12. Comparison of the urban land use structure between the mapping of SULUC (upper) and complete samples (bottom).

From the perspective of the feature contribution rate, the most important feature is building height information, followed by POI and Sentinel 2A/B multispectral information (Figure 13). In the MPL data, the Luojia-1 nighttime light feature contribution rate is very low, mainly because the original spatial resolution of these data is low, which is not suitable for high-resolution urban land use classification tasks.

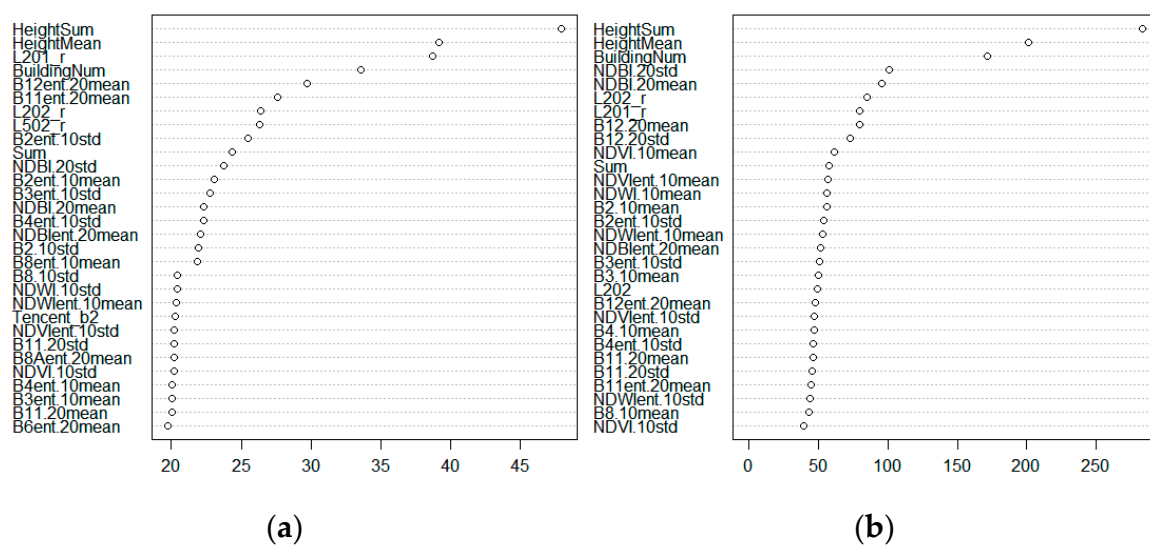


Figure 13. Feature contribution rate: (a) mean decrease in accuracy; (b) mean decrease in Gini.

4. Discussion

Mixed land use is a big obstacle to improving classification accuracy. Current results show that misclassifications of low-purity parcels were much more than those of high-purity parcels. The lower the purity of the parcel, the worse the classification accuracy (Figure 14). The reasons are as follows:

1. Due to the high scarcity of land, commercial, transportation, and public facilities in high-density cities such as Shenzhen often exist in the form of nonindependent land occupation. In this case, the features mentioned above may not be sufficiently significant compared with those in other cities.
2. There is more and more three-dimensional utilization of land use. For example, a business center generated by urban renewal could have a commercial center on its low floors and high-quality housing on the top floors; thus, this center is both commercial and urban residential. Additionally, government agencies could rent some commercial buildings for office space, and in this situation, the building is both for commercial use and public service use. In the above cases, it is unreasonable to assign only one category to a parcel. A possible solution is to assign multiple categories to a parcel through a probability method.

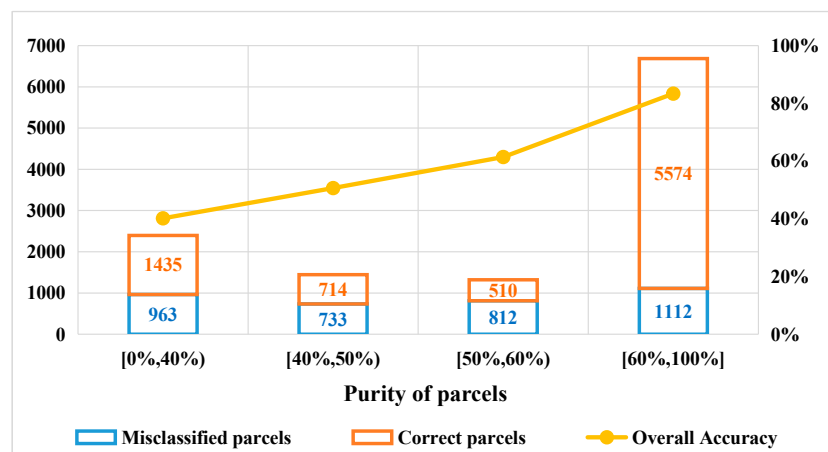


Figure 14. Number of misclassifications and classification accuracy of parcels with different purities.

The methodology of the parcel segmentation and feature extraction can be improved:

1. The segmentation of parcels is not detailed enough. Because road segmentation technology is not suitable for the underdeveloped areas of the road network in the city, this results in superlarge parcels which contain multiple land use categories. In the future, image segmentation can be introduced to segment the superlarge parcels generated by road segmentation.
2. The POI information collection from commercial companies is biased, resulting in unsatisfactory classification results. In the future, POI information from official electronic maps can be combined with POI information from commercial institutions to enhance the classification accuracy.

Given the opportunity that Shenzhen has a complete set of ground truth of land use samples, it makes it possible to design a series of experimental tests to investigate the impact of sample quantity and quality on detailed land use classification performance. We have further checked the availability of data in different cities around the world. The multispectral and nighttime light remote sensing data used in this paper can be obtained globally. Global road network data can also be accessible through OpenStreetMap. However, the major challenge of this study was to collect sufficient land use samples. Fortunately, Shenzhen has just conducted an urban land use survey, and we could obtain its complete sample set from the survey results. Similar research can be conducted in other cities in China after the completion of the Third Nationwide Land Survey of China. In other areas, the cadastral data could be

considered as a source of samples in similar experiments to demonstrate whether the conclusions are representative throughout the world.

5. Conclusions

In the process of detailed urban land use classification based on multisource remote sensing, VGI, and machine learning, we studied how to improve the classification accuracy by optimizing the number and purity of the samples and summarized the optimal sampling strategy. The main conclusions are as follows:

1. **Quantity strategy.** To acquire the best classification accuracy in a single city, it is necessary to collect training samples of no less than 40% of the total number of urban parcels or no less than 5500 in number. If limited labor is available for sample selection, it is recommended to collect no less than 7% of the total parcels of training samples or no less than 900 samples. Further reduction in the number could cause a significant loss of accuracy. To ensure the stability and reliability of the accuracy evaluation results, it is necessary to collect no less than 10% of the total parcels of validation samples or no less than 1200. Notably, if the principle of stratified random sampling is followed, the impact of the number of validation samples on the accuracy evaluation is limited. Even if the number of validation samples is reduced to 1% of the total, the maximum accuracy evaluation loss is not greater than 8%.
2. **Purity strategy.** Using only high-purity samples could cause a certain loss of accuracy. It means that there is no need to collect only high-purity parcels as training samples. The better strategy is to prioritize using samples with a purity between 60% and 90%. It is worth noting that random sampling without considering purity can also obtain ideal accuracy results, but there are great difficulties in identifying low-purity mixed land, which could require more work.
3. **Spatial distribution strategy.** The spatial distribution of the samples should be as balanced as possible, as unbalanced sampling will cause a significant accuracy loss even if in a single city. The samples have the ability to migrate. When spatial equilibrium sampling is not allowed, priority should be given to areas with complex land use patterns, which can provide better classification results.

Author Contributions: Conceptualization, R.G.; methodology, B.C. and M.S.; formal analysis, M.S.; visualization, M.S.; investigation, J.W. and Y.F.; data curation, M.S. and B.C.; writing—original draft preparation, M.S. and B.C.; supervision, R.G.; resources, B.X.; Writing—Review & edition, W.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: This research was supported by the Shenzhen Urban Planning and Land Resource Research Center. We express our appreciation of our colleagues in the laboratory for their constructive suggestions and comments.

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

References

1. United Nations, Department of Economic and Social Affairs PD. *World Urbanization Prospects: The 2018 Revision (ST/ESA/SER.A/420)*; United Nations: New York, NY, USA, 2018.
2. Gong, P.; Li, X.; Zhang, W. 40-Year (1978–2017) human settlement changes in China reflected by impervious surfaces from satellite remote sensing. *Sci. Bull.* **2019**, *64*, 756–763. [[CrossRef](#)]
3. Cheng, J.; Turkstra, J.; Peng, M.; Du, N.; Ho, P. Urban land administration and planning in China: Opportunities and constraints of spatial data models. *Land Use Policy.* **2006**, *23*, 604–616. [[CrossRef](#)]
4. Su, M.; Zhou, B.; Liao, Q.; Hu, W.; Hong, W. Thinking on the construction of the Whole Domain Digital Current Data. *Urban Plan.* **2018**, *42*, 97–104.
5. Xiao, J. Remote sensing survey of urban land use in Chongqing. *Remote Sens. Land Resour.* **1995**, *64*, 7–14.

6. Huang, A. Large scale urban land use survey based on Remote Sensing Information. *Subtrop. Soil Water Conserv.* **2009**, *21*, 59–63.
7. Su, M.; Liao, Q.; Luo, G.; Wei, X. Design and implementation of land survey technology for planning and land data fusion. *Sci. Technol. Manag. Land Resour.* **2013**, *30*, 93–99.
8. *Technical Regulations for the Third Nationwide Land Survey*; Ministry of Natural Resources of the People's Republic of China (Ed.) Geological Press: Beijing, China, 2019; Volume TD/T 1055-2019.
9. Anderson, J.R.; Hardy, E.E.; Roach, J.T.; Witmer, R.E. *A Land Use and Land Cover Classification System for Use with Remote Sensing Data*; United States Government Office: Washington, DC, USA, 1976; p. 964.
10. Gong, P.; Howarth, P.J. Performance analyses of probabilistic relaxation methods for land-cover classification. *Remote Sens. Environ.* **1989**, *30*, 33–42. [[CrossRef](#)]
11. Gong, P.; Howarth, P.J. Graphical approach for the evaluation of land-cover classification procedures. *Int. J. Remote Sens.* **1990**, *11*, 899–905. [[CrossRef](#)]
12. Gong, P.; Howarth, P.J. Land-use classification of SPOT HRV data using a cover-frequency method. *Int. J. Remote Sens.* **1992**, *13*, 1459–1471. [[CrossRef](#)]
13. Gong, P.; Marceau, D.J.; Howarth, P.J. A Comparison of spatial feature extraction algorithms for land-use classification with SPOT HRV data. *Remote Sens. Environ.* **1992**, *40*, 137–151. [[CrossRef](#)]
14. Treitz, P.M.; Howarth, P.J.; Gong, P. Application of satellite and GIS technologies for land-cover and land-use mapping at the rural-urban fringe: A case study. *Photogramm. Eng. Remote Sens.* **1992**, *58*, 439–448.
15. Gong, P. Reducing boundary effects in a kernel-based classifier. *Int. J. Remote Sens.* **1994**, *15*, 1131–1139. [[CrossRef](#)]
16. Adams, J.B.; Sabol, D.E.; Kapos, V.; Filho, R.A.; Roberts, D.A.; Smith, M.O.; Gillespie, A.R. Classification of multispectral images based on fractions of endmembers: Application to land-cover change in the Brazilian Amazon. *Remote Sens. Environ.* **1995**, *52*, 137–154. [[CrossRef](#)]
17. Ben-Dor, E.; Levin, N.; Saaroni, H. A Spectral Based Recognition of the Urban Environment Using the Visible And Near-Infrared Spectral Region (0.4-1.1 μM). A Case Study Over Tel-Aviv, Israel. *Int. J. Remote Sens.* **2001**, *22*, 2193–2218. [[CrossRef](#)]
18. Gao, F.; de Colstoun, E.B.; Ma, R.; Weng, Q.; Masek, J.G.; Chen, J.; Pan, Y.; Song, C. Mapping impervious surface expansion using medium-resolution satellite image time series: A case study in the Yangtze River Delta, China. *Int. J. Remote Sens.* **2012**, *33*, 7609–7628. [[CrossRef](#)]
19. Johnson, B.A. High-resolution urban land-cover classification using a competitive multi-scale object-based approach. *Remote Sens. Lett.* **2013**, *4*, 131–140. [[CrossRef](#)]
20. Yu, L.; Liang, L.; Wang, J.; Zhao, Y.; Cheng, Q.; Hu, L.; Liu, S.; Yu, L.; Wang, X.; Zhu, P.; et al. Meta-discoveries from a synthesis of satellite-based land-cover mapping research. *Int. J. Remote Sens.* **2014**, *35*, 4573–4588. [[CrossRef](#)]
21. Pesaresi, M.; Ehrlich, D.; Caravaggi, I.; Kauffmann, M.; Louvrier, C. Toward Global Automatic Built-Up Area Recognition Using Optical VHR Imagery. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2011**, *4*, 923–934. [[CrossRef](#)]
22. Zhan, Q.; Molenaar, M.; Gorte, B. Urban land use classes with fuzzy membership and classification based on integration of remote sensing and GIS. *Int. Arch. Photogramm. Remote Sens.* **2000**, *33*, 1751–1760.
23. Goodchild, M.F. Citizens as sensors: the world of volunteered geography. *GeoJournal* **2007**, *69*, 211–221. [[CrossRef](#)]
24. Soto, V.; Frías-Martínez, E. Automated land use identification using cell-phone records. In Proceedings of the 3rd ACM International Workshop on MobiArch, HotPlanet 11, Bethesda, MD, USA, 28 June 2011; pp. 17–22.
25. Cranshaw, J.; Schwartz, R.; Hong, J.I.; Sadeh, N. The Livelihoods Project: Utilizing Social Media to Understand the Dynamics of a City. In Proceedings of the 6th International AAAI Conference on Weblogs and Social Media, Dublin, Ireland, 4–7 June 2012.
26. Yuan, J.; Zheng, Y.; Xie, X. Discovering regions of different functions in a city using human mobility and POI. In Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD 2012, Beijing, China, 12–16 August 2012; pp. 186–194.
27. Long, Y.; Shen, Z. Discovering Functional Zones Using Bus Smart Card Data and Points of Interest in Beijing. In *Geospatial Analysis to Support Urban Planning in Beijing*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 193–217.

28. Song, X.; Pu, Y.; Liu, D.; Feng, Y. Mining the functional attributes of urban area by using pedestrian trajectory. *J. Surv. Mapp.* **2015**, *44*, 82–88.
29. Chen, S.; Tao, H.; Li, X.; Zhuo, L. Identification of urban functional areas based on latent semantic information — GPS spatio-temporal data mining of Guangzhou floating car. *J. Geogr.* **2016**, *71*, 471–483.
30. Liu, X.; Long, Y. Automated identification and characterization of parcels with OpenStreetMap and points of interest. *Environ. Plan. B Urban Anal. City Sci.* **2016**, *43*, 341–360.
31. 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](#)]
32. Schultz, M.; Voss, J.; Auer, M.; Carter, S.; Zipf, A. Open land cover from OpenStreetMap and remote sensing. *Int. J. Appl. Earth Obs. Geoinf.* **2017**, *63*, 206–213. [[CrossRef](#)]
33. Zhang, W.; Li, W.; Zhang, C.; Hanink, D.M.; Li, X.; Wang, W. Parcel-based urban land use classification in megacity using airborne LiDAR, high resolution orthoimagery, and Google Street View. *Comput. Environ. Urban Syst.* **2017**, *64*, 215–228. [[CrossRef](#)]
34. Zhang, Y.; Li, Q.; Huang, H.; Wu, W.; Du, X.; Wang, H. The Combined Use of Remote Sensing and Social Sensing Data in Fine-Grained Urban Land Use Mapping: A Case Study in Beijing, China. *Remote Sens.* **2017**, *9*, 865. [[CrossRef](#)]
35. Cao, R.; Zhu, J.; Tu, W.; Li, Q.; Cao, J.; Liu, B.; Zhang, Q.; Qiu, G. Integrating aerial and street view images for urban land use classification. *Remote Sens.* **2018**, *10*, 1553. [[CrossRef](#)]
36. Chen, W.; Huang, H.; Dong, J.; Zhang, Y.; Tian, Y.; Yang, Z. Social Functional Mapping Of Urban Green Space Using Remote Sensing And Social Sensing Data. *ISPRS J. Photogramm. Remote Sens.* **2018**, *146*, 436–452. [[CrossRef](#)]
37. Grippa, T.; Georganos, S.; Zarougui, S.; Bognounou, P.; Diboulo, E.; Forget, Y.; Lennert, M.; Vanhuyse, S.; Mboga, N.; Wolff, E. Mapping Urban Land Use at Street Block Level Using OpenStreetMap, Remote Sensing Data, and Spatial Metrics. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 246. [[CrossRef](#)]
38. Luo, N.X.; Wan, T.L.; Hao, H.X.; Lu, Q.K. Fusing High-Spatial-Resolution Remotely Sensed Imagery and OpenStreetMap Data for Land Cover Classification Over Urban Areas. *Remote Sens.* **2019**, *11*, 88. [[CrossRef](#)]
39. Zhang, Y.; Li, Q.; Tu, W.; Mai, K.; Yao, Y.; Chen, Y. Functional urban land use recognition integrating multi-source geospatial data and cross-correlations. *Comput. Environ. Urban Syst.* **2019**, *78*, 1–11. [[CrossRef](#)]
40. Gong, P.; Chen, B.; Li, X.; Liu, H.; Wang, J.; Bai, Y.; Chen, J.; Chen, X.; Fang, L.; Feng, S.; et al. Mapping essential urban land use categories in China (EULUC-China): preliminary results for 2018. *Sci. Bull.* **2020**, *65*, 182–187. [[CrossRef](#)]
41. Van Niel, T.G.; McVicar, T.R.; Datt, B. On the relationship between training sample size and data dimensionality: Monte Carlo analysis of broadband multi-temporal classification. *Remote Sens. Environ.* **2005**, *98*, 468–480. [[CrossRef](#)]
42. Demir, B.; Erturk, S. Increasing hyperspectral image classification accuracy for data sets with limited training samples by sample interpolation. In Proceedings of the 4th International Conference on Recent Advances in Space Technologies 2009, Istanbul, Turkey, 11–13 June 2009; pp. 367–369.
43. Goncalves, L.M.S.; Fonte, C.C.; Carrao, H.; Caetano, M. Improving image classification accuracy: A method to incorporate uncertainty in the selection of training sample set. In Proceedings of the 9th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, Leicester, UK, 20–23 July 2010; pp. 261–264.
44. Li, C.; Jie, W.; Lei, W.; Hu, L.; Peng, G. Comparison of Classification Algorithms and Training Sample Sizes in Urban Land Classification with Landsat Thematic Mapper Imagery. *Remote Sens.* **2014**, *6*, 964–983. [[CrossRef](#)]
45. Gong, P.; Liu, H.; Zhang, M.N.; Li, C.C.; Wang, J.; Huang, H.B.; Clinton, N.; Ji, L.Y.; Li, W.Y.; Bai, Y.Q.; et al. Stable classification with limited sample: transferring a 30-m resolution sample set collected in 2015 to mapping 10-m resolution global land cover in 2017. *Sci. Bull.* **2019**, *64*, 370–373. [[CrossRef](#)]
46. Georganos, S.; Grippa, T.; Vanhuyse, S.; Lennert, M.; Shimoni, M.; Wolff, E. Very High Resolution Object-Based Land Use-Land Cover Urban Classification Using Extreme Gradient Boosting. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 607–611. [[CrossRef](#)]
47. Maxwell, A.E.; Warner, T.A.; Fang, F. Implementation of machine-learning classification in remote sensing: An applied review. *Int. J. Remote Sens.* **2018**, *39*, 2784–2817. [[CrossRef](#)]
48. Ruiz Hernandez, I.E.; Shi, W. A Random Forests classification method for urban land-use mapping integrating spatial metrics and texture analysis. *Int. J. Remote Sens.* **2018**, *39*, 1175–1198. [[CrossRef](#)]

49. Statistics Bureau of Shenzhen Municipality; Survey Office of the National Bureau of Statistics in Shenzhen. *Shenzhen Statistical Yearbook 2019*; China Statistics Press: Beijing, China, 2019.
50. Statistics Bureau of Beijing Municipality; Survey Office of the National Bureau of Statistics in Beijing. *Beijing Statistical Yearbook 2019*; China Statistics Press: Beijing, China, 2019.
51. Statistics Bureau of Shanghai Municipality. The GDP of Shanghai in 2019. Available online: <http://tjj.sh.gov.cn/html/sjfb/202001/1004392.html> (accessed on 9 February 2020).
52. *Current Land Use Classification*; General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, China National Standardization Management Committee (Ed.) Standards Press of China: Beijing, China, 2017; Volume GB/T 21010-2017.
53. Shenzhen Municipal Commission of Planning and Land Resources. *Notice on Printing and Distributing the Code for Investigation of Land Change in Shenzhen (for Trial Implementation)*; Shenzhen Municipal Commission of Planning and Land Resources: Shenzhen, China, 2015.
54. Chen, B.; Song, Y.M.; Jiang, T.T.; Chen, Z.Y.; Huang, B.; Xu, B. Real-Time Estimation of Population Exposure to PM2.5 Using Mobile- and Station-Based Big Data. *Int. J. Environ. Res. Public Health*. **2018**, *15*, 573. [[CrossRef](#)]
55. Hansen, M.; Dubayah, R.; Defries, R. Classification trees: An alternative to traditional land cover classifiers. *Int. J. Remote Sens.* **1996**, *17*, 1075–1081. [[CrossRef](#)]
56. Atkinson, P.M.; Tatnall, A.R.L. Introduction Neural networks in remote sensing. *Int. J. Remote Sens.* **1997**, *18*, 699–709. [[CrossRef](#)]
57. Friedl, M.A.; Brodley, C.E. Decision tree classification of land cover from remotely sensed data. *Remote Sens. Environ.* **1997**, *61*, 399–409. [[CrossRef](#)]
58. Giacinto, G.; Roli, F.; Bruzzone, L. Combination of neural and statistical algorithms for supervised classification of remote-sensing images. *Pattern Recognit. Lett.* **2000**, *21*, 385–397. [[CrossRef](#)]



© 2020 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 (<http://creativecommons.org/licenses/by/4.0/>).