Article

Comparison of Machine-Learning Methods for Urban Land-Use Mapping in Hangzhou City, China

Wanliu Mao ^{1,2}, Debin Lu ¹, Li Hou ¹, Xue Liu ¹ and Wenzhe Yue ^{1,*} ¹ Department of Land Management, Zhejiang University, Hangzhou 310058, China;

11922077@zju.edu.cn (W.M.); ludebin@zju.edu.cn (D.L.); houli@cau.edu.cn (L.H.); xueliu@zju.edu.cn (X.L.)

² Zhejiang Academy of Surveying and Mapping, Hangzhou 311100, China

* Correspondence: wzyue@zju.edu.cn

Received: 30 July 2020; Accepted: 29 August 2020; Published: 31 August 2020

Abstract: Urban land-use information is important for urban land-resource planning and management. However, current methods using traditional surveys cannot meet the demand for the rapid development of urban land management. There is an urgent need to develop new methods to overcome the shortcomings of conventional methods. To address the issue, this study used the random forest (RF), support vector machine (SVM), and artificial neural network (ANN) models to build machine-learning methods for urban land-use classification. Taking Hangzhou as an example, these machine-learning methods could all successfully classify the essential urban land use into 6 Level I classes and 13 Level II classes based on the semantic features extracted from Sentinel-2A images, multi-source features of types of points of interest (POIs), land surface temperature, night lights, and building height. The validation accuracy of the RF model for the Level I and Level II land use was 79.88% and 71.89%, respectively, performing better compared to SVM (78.40% and 68.64%) and ANN models (71.30% and 63.02%). However, the variations of the user accuracy among the methods depended on the urban land-use level. For the Level I land-use classification, the user accuracy was high, except for the transportation land by all methods. In general, the RF and SVM models performed better than the ANN model. For the Level II land-use classification, the user accuracy of different models was quite distinct. With the RF model, the user accuracy of educational and medical land was above 80%. Moreover, with the SVM model, the user accuracy of the business office and educational land classification was above 75%. However, the user accuracy of the ANN model on the Level II land-use classification was poor. Our results showed that the RF model performs best, followed by SVM model, and ANN model was relatively poor in the essential urban land-use classification. The results proved that the use of machine-learning methods can quickly extract land-use types with high accuracy, and provided a better method choice for urban land-use information acquisition.

Keywords: urban land use; machine-learning method; multi-source data; Hangzhou

1. Introduction

Land-use/land-cover classification is the basis for Land-Use and Land-Cover Change research. However, it can no longer meet the requirements for efficiency and accuracy of land-use classification and complex information acquisition through traditional visual interpretation and mathematical statistics in rapidly urbanizing areas [1]. With the rapid development of China urbanization, how to quickly obtain information regarding urban land use is considered a hot topic. With powerful adaptive and self-learning parallel information-processing capabilities, machine-learning methods have been successfully applied and developed in many fields. For example, Paoletti et al. [2] developed a new, highly efficient fulfillment of support vector machine (SVM) using the high computational power of graphics processing units to lessen the execution time of the storage and processing of hyperspectral

images. Lee et al. [3] established multiple methods of simple/multiple linear regression, random forests (RF), and support vector regression to estimate canopy nitrogen weight of corn fields in south-west Ontario, Canada, showing that both machine-learning models showed a much higher accuracy than linear regression. Meanwhile, machine-learning methods such as artificial neural networks [4–6], support vector machines [7–9], and random forests [10–12] are also widely applied in land-cover classification. Here are some further examples. Zhao et al. [13] compared the classification effects of machine-learning methods on land cover in typical mountainous areas, and found that decision tree had the highest classification accuracy but poor stability, random forest had good stability and fast training speed, and support vector machines classified fast but required detailed feature parameters. When quantifying land-cover changes in a Mediterranean environment, Fragou et al. [14] exploited the support vector machines classifier to classify the natural landscape of Landsat Thematic Mapper images in different years, and the overall accuracy was all around 90%. Chakhar et al. [15] found that support vector machines and nearest neighbor methods had the best balance between robustness and efficiency among 22 non-parametric classification algorithms for classifying irrigated crops in a semiarid region. LaRocque et al. [16] combined multi-source remote-sensing images, and used random forest to successfully obtain 11 types of wetland mapping in Southern New Brunswick, Canada. Morell-Monzó et al. [17] applied the random forests algorithm to effectively identify and quantify abandoned agricultural plots using Sentinel-2 and airborne images, which provided an implementation method for mapping citrus and other crops in highly fragmented areas. It can be seen that the application of machine-learning methods in land-cover classification is relatively mature.

However, land-use classification is different from land-cover classification, which is reflected by the fact that land cover focuses on natural attributes, while land use focuses more on social attributes. Although these two attributes have some similarities, in urban areas, more emphasis is placed on land-use patterns and conditions. Remote-sensing data can fully display the natural attributes of ground components, but they cannot show the socio-economic attributes caused by human activities, due to the high degree of similarity in spectrum and texture among different types of land use in cities. The complexity of the socio-economic attributes makes urban land-use classification more challenging.

With the advent of the big-data era, mobile phone record data, floating car data, social media data, and other data with temporal and spatial characteristics is constantly being produced, which can effectively exhibit features about human activities and provide new ideas for urban land-use classification. There has been research [18] that has used similarity measures and threshold methods to classify urban land use in Beijing based on remote-sensing images and points of interest (POIs) data, showing that the overall accuracy of the Level I and Level II land-use classification was 81.04% and 69.89%, respectively. Obviously, adding socio-economic data can improve the land-use classification accuracy. However, it is relatively unusual for research to classify urban land use combining the image data and Internet open data, and little research has been conducted in large cities. For example, Liu et al. [19] integrated high-speed rail images and multi-source social media data, and used probabilistic topic models and SVM to effectively obtain urban land-use classification information, with an overall accuracy of 86.5% and a Kappa coefficient of 0.828. Gong et al. [20] released the first mapping results of essential urban land-use categories (EULUC) in China for 2018. Based on the basic pattern of classification generated mainly by road network from the OpenStreetMap (OSM), this research combined the data from Sentinel-2 image, POIs, and other multi-source data, and the random forest model was applied to classify the land use in cities of China. Afterwards, some researchers made further in-depth studies on parcel segmentation [21], sample selection [22], and feature selection [23,24]. However, the applicable conditions of different models are distinct because of the dissimilar urban development degree and complexity in different regional cities. Therefore, the application of multiple classification models for comparative research can improve the credibility of the results and provide a reference to subsequent research.

Impressed by the study of EULUC [20], this research took Hangzhou as the study area and extracted classification features from remote-sensing data, social media data, and other Internet open data,

then used three typical machine-learning methods: RF, SVM, and artificial neural network (ANN) to classify urban land use in Hangzhou (2018). It aimed to compare the classification accuracy of different machine-learning methods in urban land use. In addition, the information regarding classification results can also provide decision-making reference for urban land-use planning and management.

2. Study Area

Hangzhou is the provincial capital of Zhejiang province, and regarded as the economic, cultural, science, and educational center in Hangzhou metropolitan area. Hangzhou is also one of the central cities in the region of Yangtze River Delta located between $29^{\circ}11' - 30^{\circ}34'N$ and $118^{\circ}20' - 120^{\circ}37'E$. Despite the population growth slowdown in some cities in China, in recent years, the migrant population in Hangzhou still maintains a fast-growing trend. The city's gross domestic product (GDP) was RMB 1537.3 billion, and the permanent resident population was 10.36 million, with an urbanization rate of 78.5% in 2019. Rapid urbanization has brought significant changes to the land-use pattern of Hangzhou.

3. Data Sources and Methods

Figure 1 presents a flowchart outlining the methodology used in this study, which consists of three major parts. First, the study used the road network and water body to generate parcels, and based on the classification feature data sets extracted from the multi-source data, combined with field sampling photos and Google Street Views to form various land-use classification data sets. Then, a random method was used to divide the training samples and the validation samples. Next, RF, SVM, and ANN were chosen to classify urban land use, and the classification results and accuracy were further compared and analyzed.

Figure 1. Technique flow chart of land-use classification.

3.1. Parcel Generation and Sample Selection

3.1.1. The Data of Impervious Surface, Road, Water Body, and Parcel Generation

A hierarchical approach was developed to delineate urban boundaries based on 30-m resolution impervious surface data [25]. For the research of the EULUC [20], a basic network of parcel segmentation for the impervious surface were generated by the major roads and minor roads from OpenStreetMap (<https://www.openstreetmap.org/#map=4/36.96/104.17>) and the water layer from the 10-m resolution global land-cover map based on Sentinel-2 data [26]. To further refine urban parcels, a more accurate data of the road network and water body extracted from the monitoring data of Zhejiang province's geographic conditions was integrated to generate the final urban parcels. A total of 11,212 parcels were obtained in Hangzhou, with an average parcel area of 21.01 hectares (Figure 2(b₁,b₂)). Compared with EULUC (Figure 2(a₁,a₂)), the parcel segmentation was more precise.

Figure 2. (a1) The example of parcel segmentation for EULUC; (a2) The example of parcel image for EULUC; (b1) The example of parcel segmentation for this study; (b2) The example of parcel image for this study.

3.1.2. Parcel Classification and Sample Selection

Based on the land-use classification system of EULUC and the actual characteristics of regional mapping in Hangzhou, 6 Level I land-use classification (Residential, Commercial, Industrial, Transportation, Public management and service, Non-construction) and 12 Level II land-use classification (Residential, Business office, Commercial service, Industrial, Road, Transportation station, Airport, Administrative, Educational, Medical, Sport and cultural, Park and green space, Non-construction) were formed, as shown in Table 1. In particular, the category of non-construction

land was newly added to the classification, which mainly includes the land approved but not built in the city and the cultivated land in suburban areas.

Table 1. Types of land use and sample size.

Level I	Level II	Number of Samples
01 Residential	0101 Residential	362
02 Commercial		197
	0201 Business office	99
	0202 Commercial service	98
03 Industrial	0301 Industrial	217
04 Transportation		26
	0401 Road	-
	0402 Transportation station	23
	0403 Airport	3
05 Public management and service		255
	0501 Administrative	46
	0502 Educational	115
	0503 Medical	25
	0504 Sport and cultural	22
	0505 Park and green space	47
06 Non-construction	0601 Non-construction	70

In the selection of samples location and purity, to enhance the stability of the samples as much as possible, only parcels with a balanced spatial distribution and a purity of 80% or more can be selected as samples. Meanwhile, in order to avoid excessive differences in the number of samples of different land-use types, a total record of 1127 samples were finally obtained through field surveys and Google Street Views (photos shown in Figure 3). The number of samples for each land-use types is shown in Table 1. In addition, these samples were randomly divided into training data sets and testing data sets at a ratio of 7:3. According to the randomly selected training samples and testing samples, the confusion matrix was used to test the accuracy of the Level I and II land-use classification results obtained by different methods.

Figure 3. Photos of Level II land-use types collected by field surveys and Google Street Views.

3.2. Feature Extraction

3.2.1. Image Features

The Sentinel-2 satellite image has been used in many applications owing to the fine resolution both in time and space [15,27]. The 2018 Sentinel-2A satellite images were selected to extract the multispectral features, which were downloaded from the website of <https://scihub.copernicus.eu/>. After the atmospheric correction, the four bands of blue, green, red and near-infrared bands, and normalized vegetation index ($NDVI = (NIR - Red) / (NIR + Red)$), normalized water index ($NDWI = (Green - NIR) / (Green + NIR)$) all with 10 meters spatial resolution were calculated in each parcel to further obtain the corresponding index mean, standard deviation, and information entropy. Among them, the information entropy characterizes the image texture features, which can measure the randomness of the image information and represent the complexity of the image.

3.2.2. Land Surface Temperature

The land surface temperature (LST) reflects the social and economic activities of human beings to a certain extent. Previous studies have shown that there are significant disparities in the LST of different urban land-use types [28]. The GEE platform was used to extract the Landsat 8 image of Hangzhou city [29,30]. Due to poor images quality caused by cloud pollution in 2018, the Landsat 8 images of May 2017, as close as possible to 2018, finally were selected instead for the LST retrieval. The radiometric correction equation was used to calculate the pixel value of the thermal band of Landsat 8 image as the radiant temperature, which was corrected by the specific emissivity next [31,32]. In the end, the LST of Hangzhou was obtained and further calculated the mean and standard deviation of the LST in each parcel.

3.2.3. POIs

The POIs data in 2018 were acquired from Gaode Maps (<https://lbs.amap.com/api/webservice/guide/api/Search>). Each POIs data contains a series of information such as name, location coordinates, and city function categories. According to the Level II land-use classification, all POIs were reclassified into 12 corresponding types. In addition, the number, proportion, and total of each POIs type were calculated in each parcel, respectively.

3.2.4. Building Height

Buildings height data was consisted of the outline and height of each building in Hangzhou, which was acquired via the Baidu Map API (<http://api.map.baidu.com/staticimage/v2>). the average height of the buildings in each parcel was calculated further to aggregate these data into parcel levels.

3.2.5. Night Lights

Night lights are closely related to human economic activities, and the application of research in urban development and population has become more and more extensive [33–35]. The LuoJia-1 nighttime lights (NTLs) with a resolution of 130 meters in Hangzhou, 2018 were downloaded from the website of <http://59.175.109.173:8888/app/login.html>. Based on the DN value of the NTLs pixel, the mean and standard deviation of the DN value in each parcel were calculated. The features used in this study were summarized in Table 2.

Table 2. Summary table of features.

Data Source	Year	Features	Variables
Sentinel-2A	2018	Mean of blue, green, red, near-infrared bands, NDVI, NDWI	b2mean, b3mean, b4mean, b8mean, NDVImean, NDWmean
		Standard deviation of blue, green, red, near-infrared bands, NDVI, and NDWI	b2sd, b3sd, b4bsd, b8sd, NDVIsd, NDWIsd
		Mean of entropy of blue, green, red, near-infrared bands, NDVI, NDWI	enb2mean, enb3mean, enb4mean, enb8mean, enNDVImean, enNDWImean
		Standard deviation of entropy of blue, green, red, near-infrared bands, NDVI, NDWI	enb2sd, enb3sd, enb4bsd, enb8sd, enNDVIsd, enNDWIsd
Landsat 8	2017	Mean of Temperature	TEMmean
		Standard deviation of Temperature	TEMsd
Gaode-based POIs	2018	Total number of all POIs	Sum
		Total number of each type of POIs	101, 201, 202, 301, 402, 403, 501, 502, 503, 504, 505
		Proportion of each type of POIs	P101, P201, P202, P301, P402, P403, P501, P502, P503, P504, P505
Height from Baidu	2018	Mean of Height	Hmean
Luoja-1 nighttime lights	2018	Mean of digital number values	LJmean
		Standard deviation of digital number values	LJsd

3.3. Methods

3.3.1. Random Forest

The RF model is an integrated learning algorithm proposed by Breiman in 2001 [36–38]. It can increase the diversity of classification trees and enhance the performance of a single classification tree or regression tree by putting back sampling and randomly changing the combination of predictive variables in the evolution of different trees. The modeling steps are as follows: first, using bootstrap sampling technology to extract X_i training sets from the original data set, the size of each training set is about 2/3 of the original data set, and the remaining ($X-X_i$) samples form the out-of-bag data (out-of-bag, OOB). Second, the regression tree of each X_i training set was not pruned and allowed to grow freely. Randomly select m predictor variables at each node, and among these random variables, the optimal feature was select for node segmentation according to the principle of minimum Gini coefficient. Third, predict new data through the feedback information regarding X_i regression trees, and the classification result is determined by voting on the output results of each classification decision tree. In the process of random forest classification, three custom parameters need to be defined to optimize the model: the number of spanning trees ($n_estimators$), the number of predictors used to split the node at each node ($max_features$), and the minimum number of leaves ($min_samples_leaf$). The three parameters can be determined by the error rate of the data outside the bag.

3.3.2. Support Vector Machine

SVM is a new machine-learning method developed based on statistical learning theory and the principle of structural risk minimization [39]. Compared with the traditional learning methods, it has the characteristics of high accuracy, fast calculation speed, and strong generalization ability, which is widely used in image and land classification mapping [19,40]. The basic idea of SVM classification is to transform the input space into a high-dimensional feature space through nonlinear transformation, and then further find the optimal hyperplane (OHP) in this new high-dimensional feature space. The optimal hyperplane can not only correctly classify all the training samples, but also maximize the distance between the points closest to the classification plane namely it can maximize the classification interval to segregate the different classes. Simultaneously, the most crucial thing for classification using

SVM is the choice of kernel function and the solution of kernel parameters [39]. In this study, the radial basis function was used as the kernel function, and the grid search method was used to determine the optimal penalty coefficient C and the classification interval. Finally, set penalty coefficient C to 5, gamma to 0.01, and the classification rule adopted the “one-versus-rest” classifier.

3.3.3. Artificial Neural Network

The study used the Back-Propagation Network, which is the most widely used ANN model and was proposed by Rumelhart et al. in 1985 [41]. It is a multi-layer feed-forward neural network trained in accordance with the error back-propagation algorithm. A typical BP neural network structure includes an input layer, a hidden layer (consisting of one or more layers) and an output layer. The adjacent layers are connected by weights, and the network learning process of information consists of two processes of forward and backward propagation. The input information is forwarded through the activation function [42,43], and then the corresponding output is acquired. The output result needs to be compared with the target output. On condition that the error exceeds the predetermined value, it will be transferred to back-propagation. Meanwhile, the error signal will be fed back from the output layer to the hidden layer and the input layer, and the connection weights between the nodes (neurons) of each layer will be adjusted according to the error [44]. Repeat this process until the signal error reaches the allowable error range, to achieve the purpose of classification. The input layer of ANN is the number of neurons, i.e., the number of input variables. It depends on the complexity of the hidden layer problem. The number of neurons in the output layer is the number of output variables. The neural network activation functions include identity, logistic, tanh, and relu functions, and the solver for weight optimization include lbfgs, SGD, and adam functions (https://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPClassifier.html). In the study, the logistic function was used as the activation function, and the lbfgs function was used to optimize the weights.

The above classification calculations were all implemented using the Scikit-Learn package in Python 3.7.

4. Results

4.1. Classification Results

The results of land-use classification using RF, SVM and ANN models were various in the Level I land use (Figure 4). The areas of the three land types of residential, industrial, and public management and service land accounted for 18.67%, 35.66%, and 33.85% with RF, 26.61%, 41.80%, 25.62% with SVM, and 32.52%, 38.52%, 23.64% with ANN. On the contrary, both the commercial and transportation land accounted for less than 4%. Although the area percentages were different, residential, industrial, and public management and service land parcels accounted for a large proportion in the Level I land use, no matter what methods were used.

Figure 4. The area percentage of Level I land-use types in the mapping results.

The results of Level II land-use classification also showed significant differences (Figure 5). For the business office and commercial service land, the proportions were 1.06% and 2.66% with RF, 0.57% and 4.17% with SVM, and 0.58% and 4.40% with ANN. For the results of public management and service land, the areas of educational, parks and green space, and administrative land occupied a large proportion, which can be classified relatively well. The areas of the three types accounted for 6.65%, 10.32%, and 4.98% with RF, 11.68%, 5.45%, and 0.82% with SVM, and 12.46%, 3.68%, and 0.62% with ANN.

Figure 5. The area percentage of Level II land-use types in the mapping results.

Overall, the results of Level II land-use classification using three methods showed high consistency in space that the commercial service and business office land were mainly distributed in the central urban areas, and that industrial and park and green space land mostly located in the suburbs. However, in the specific areas, the classification results were quite different. Selecting the downtown areas of Hangzhou as the comparison area, it showed that the classification results of RF and SVM were relatively close, while the ANN results were quite different (Figure 6).

Figure 6. The spatial distributions of the Level II land use in Hangzhou, 2018. (a) Impervious surface in urban boundaries; (b) Detailed zoomed-in view map of the downtown area.

4.2. Accuracy Assessment

On the whole, the validation accuracy, training accuracy, and Kappa coefficient of the Level I land-use classification were 79.88%, 92.88%, 0.738 with RF, 78.40%, 82.34%, 0.720 with SVM, and 71.30%, 99.8%, 0.624 with ANN (Figure 7; Table 3). Obviously, the classification results with RF were the best, with a high degree of consistency and high reliability. The second was SVM, whose classification accuracy was close to RF. The ANN had a poor classification effect compared with the formers. From the results of the Level I land-use classification, first, residential and industrial land can be identified well with the three methods. The user accuracy and product accuracy of the two types land classification

were more than 80% with RF and SVM, and more than 75% with ANN. Secondly, the user accuracy and product accuracy of public management and service land classification were more than 70% with all the methods. The classification errors were mainly classified as residential land, and the commission error was above 10%. Thirdly, commercial land had a better classification effect with RF, with user accuracy reaching 86.27%, and 78.43% with SVM, 60.78% with ANN. Finally, the classification effect of non-construction and transportation land were poor. Combining the high-resolution remote-sensing images, we find that the transportation land such as stations has a poor regularity. Meanwhile, it is mostly mixed with other land types, resulting in low sample purity and small sample size, which may be one of the main reasons for low classification accuracy. The non-construction land is mainly the land approved within the city but not built, and farmland in the suburbs. These parcels are lack of POIs data, and the spectral features and texture features of the images are less prominent, resulting in low classification accuracy.

Figure 7. Confusion matrix and Receiver Operating Characteristic (ROC) curve for the Level I land-use classification.

Table 3. Accuracy comparison of the Level I land-use classification.

Level I	Random Forest		Support Vector Machine		Artificial Neural Network	
	Product Accuracy	User Accuracy	Product Accuracy	User Accuracy	Product Accuracy	User Accuracy
Residential	84.96%	83.48%	85.96%	85.22%	77.68%	75.65%
Commercial	70.97%	86.27%	68.97%	78.43%	58.49%	60.78%
Industrial	81.25%	85.25%	85.00%	83.61%	76.47%	85.25%
Transportation	20.00%	14.29%	18.18%	28.57%	16.67%	14.29%
Public management and service	87.14%	73.49%	83.10%	71.08%	73.49%	73.49%
Non-construction	66.67%	76.19%	62.50%	71.43%	56.25%	42.86%
Validation accuracy	79.88%		78.40%		71.30%	
Training accuracy	92.88%		82.34%		99.80%	
Kappa coefficient	0.738		0.720		0.624	

In summary, RF and SVM had better classification effects on residential, commercial, industrial, public management and service, and non-construction land. Moreover, ANN had better classification effects on residential, industrial, and public management and service land.

It can be seen from Figure 8 and Table 4 that the results of the Level II land-use classification were similar to those of the Level I. The overall classification effect of the RF model was still the best, whose validation accuracy, training accuracy, and Kappa coefficient were 71.89%, 91.74%, and 0.664, respectively. The second was SVM. Its validation accuracy, training accuracy, and Kappa coefficient were 68.64%, 81.83% and 0.630, respectively. Although the training accuracy of ANN reached 99.8%, its validation accuracy and Kappa coefficient were both low (63.02% and 0.559), showing obvious over fitting phenomenon. It was mainly due to the increased complexity of the Level II land-use types, and the overall classification accuracy varied greatly.

Figure 8. Confusion matrix and ROC curve of the Level II land-use classification.

Table 4. Accuracy comparison of the Level II land-use classification.

Level II	Random Forest		Support Vector Machine		Artificial Neural Network	
	Product Accuracy	User Accuracy	Product Accuracy	User Accuracy	Product Accuracy	User Accuracy
Residential	79.80%	79.80%	80.46%	70.71%	77.55%	76.77%
Business office	61.11%	68.75%	60.00%	75.00%	51.16%	68.75%
Commercial service	64.52%	62.50%	50.00%	56.25%	53.85%	43.75%
Industrial	82.76%	73.85%	85.96%	75.38%	75.00%	78.46%
Transportation station	40.00%	33.33%	33.33%	33.33%	40.00%	33.33%
Airport	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Administrative	50.00%	37.50%	42.86%	37.50%	26.32%	31.25%
Educational	77.27%	82.93%	77.50%	75.61%	73.33%	53.66%
Medical	88.89%	88.89%	46.15%	66.67%	66.67%	66.67%
Sport and cultural	33.33%	40.00%	33.33%	60.00%	12.50%	20.00%
Park and green space	66.67%	50.00%	66.67%	50.00%	43.75%	58.33%
Non-construction	55.17%	80.00%	62.96%	85.00%	43.75%	35.00%
Validation accuracy	71.89%		68.64%		63.02%	
Training accuracy	91.74%		81.83%		99.60%	
Kappa coefficient	0.664		0.630		0.559	

From the perspective of the specific types of the Level II land-use classification, the educational and medical in the public management and service land had a higher classification accuracy. Among the methods, RF was the best of the two land-use type classifications, with user accuracy reaching 88.89% and 82.93%, and product accuracy reaching 88.89% and 77.27% respectively. In contrast, the SVM and ANN had lower classification accuracy. Secondly, for park and green space, the user accuracy of the three methods was about 50%. The classification errors of park and green space were mainly divided into non-construction land. This is because the image characteristics of cultivated land in non-construction land is relatively similar to park and green space, which interferes with the classification of park and green space, resulting in low classification accuracy. The classification accuracy of sport and cultural and administrative land in the three methods were low. Except for the user accuracy of sport and cultural land classification with SVM of 60%, the other accuracies were below 50%. Combining with the high-resolution remote-sensing images, these two types of land use are mostly mixed with other land-use types, especially the administrative land, in addition its area ratio is small. In the Level II classification of commercial service land, the user accuracy of business office land was higher than that of commercial service land, and the user accuracy was all above 68% with the three methods, while that of commercial service land only reached above 60% with RF. The three methods all misclassified commercial service land into residential land or business office land, caused by the mixed use of commercial service, residential, and business office land mostly.

In summary, RF had a better classification effect on educational and medical land. The SVM had a better classification effect on business office and educational land. In addition, ANN had a poor overall effect on the Level II land-use classification.

5. Discussion

The more accurate road network and water body data from government departments was used to further generate a basic of parcel segmentation for urban land-use mapping, and the number of testing samples was greatly increased in this paper. Compared with the study of EULUC in China [20], the research system had an improvement, as well as urban land-use mapping results more detailed. Moreover, the accuracy testing shows the reliability of the data in the evaluation of the effect of urban land-use classification. This paper, same as other studies [20–24], mainly adopted the testing method of cross validation, using validation accuracy, Kappa coefficients, user accuracy, and producer accuracy, which are widely used in land-use/land-cover mapping to evaluate the classification accuracy. In addition, we also showed the training accuracy (Tables 3 and 4), ROC curve and Area Under

the Curve (AUC) value (Figures 7 and 8) of different model classifications, supporting the results of accuracy evaluation of the classification. Overall, the designed accuracy evaluation system can explain the accuracy of land-use classification results. With the random forest model, the validation accuracy of the Level I and Level II of land-use classification was 79.88% and 71.89% respectively in this paper. Compared with published studies using random forests model for urban land-use classification, it was close to the overall accuracy of the Level I and Level II classification of Shenzhen (75.94%, 71%) [22], and lower than that of Hangzhou (82%, 78%) [20], Ningbo (87.58%, 73.53%) [21], Lanzhou (83.75%, 76.25%) [23] and Nanjing (86.1%, 80%) [24]. Meanwhile, the classification accuracy of the random forest model was generally better than that of the SVM and the artificial neural networks models in this article, which was consistent with the results of comparison of machine-learning methods for land-cover classification in the complicated terrain regions proposed by Gu et al. [1]. Synthesizing the research results of Ningbo, Shenzhen, Lanzhou, Nanjing, and other regions to have obtained nice classification results, it showed that the random forest model had good robustness and applicability in land-use mapping of different cities. Finally, there were also cases where the effect of land-use classification in individual cities or regions was poor [20], because the selection of features [21,23,24] was an essential factor that affected the accuracy of land-use classification.

In addition, this study also found that the main factor affecting the accuracy of land-use classification was the high degree of urban land-use mixing, which will cause the purity of the land-use classification parcels to decrease [22], and making it difficult to determine the correct category. Regarding that how to improve the classification accuracy of mixed types, we suggest: First, increase the types of mixed land use, such as mixed commercial and residential land. In recent years, it has gradually become a common phenomenon to improve land-use efficiency through land-use mixing in some first-tier cities in China [45]. The second is to further generate the parcels. Apparently, the road network and water body data used in this study was not enough, so we need to combine other methods or data. Tu et al. [21] used an object-based segmentation approach to generate basic urban land-use classification parcels, but it needs to pay attention to the scale of the basic parcels. If the parcel area is too small, the map spots may be incomplete and lose the attribute feature of the land-use type itself, resulting in the texture features of some land-use types becoming inconspicuous, which will affect the classification accuracy. In addition to further mining the basic parcel segmentation, feature selection, and other factors, we can also consider the advantages of multiple classification methods, combining multiple methods for urban land-use mapping and information extraction in future research.

6. Conclusions

On a more precise basis of parcel segmentation of urban areas, this paper compared and analyzed the accuracy of the RF, SVM, and ANN models in urban land-use classification in Hangzhou, providing practical methods for urban land-use classification, as well as better method selection. The main conclusions are as follows:

(1) In general, RF had the best effect on urban land-use classification, followed by SVM, and ANN was comparatively poor. In the Level I land-use classification, the training accuracy, validation accuracy, and Kappa coefficient with RF were 92.88%, 79.88%, and 0.738, respectively; 82.34%, 78.40%, and 0.720 with SVM, and 99.80%, 71.30%, 0.624 with ANN. In the Level II land-use classification, the training accuracy, validation accuracy, and Kappa coefficient with random forest were 91.74%, 71.89%, and 0.664, respectively, 81.83%, 68.64%, and 0.630 with SVM, and 99.60%, 63.02%, and 0.559 with ANN.

(2) For the Level I land use, the accuracy of the land-use classification was high except for transportation, with the user accuracy below 30% by all the methods. Among them, the user accuracies of residential and industrial land classification were basically above 80%, and the user accuracy of commercial service and public management service land classification were basically above 70%.

(3) For the Level II land use, the classification accuracies of different models for dissimilar land-use types were quite distinct. In general, the Level II of the public management and service land had a better classification effect with RF, which showed that the user accuracy of educational and medical

land was above 80%. Moreover, the Level II of the commercial service land classification had a better effect with SVM, reflected in the user accuracy of business office land classification of 75%. Meanwhile, the classification effect of SVM in the educational land was also fine, with user accuracy of 76%. In addition, the Level II classification effect of the ANN was poor.

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

Funding: This work has been supported by the National Natural Science Foundation of China (No. 41671533 and 41871169), the Fundamental Research Funds for the Central Universities (No. 2017XZA216).

Acknowledgments: The authors would like to thank Peng Gong and Bin Chen for their constructive comments, and providing the mapping results of EULUC in Hangzhou. In addition, we also express our appreciation to Haoxuan Xia, Jinhui Xiong, and other team members for providing field verification data.

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

References

- Gu, X.; Gao, X.; Ma, H.; Shi, F.; Liu, X.; Cao, X. Comparison of Machine Learning Methods for Land Use/Land Cover Classification in the Complicated Terrain Regions. *Remote Sens. Technol. Appl.* **2019**, *34*, 59–69.
- Paoletti, M.E.; Haut, J.M.; Tao, X.; Miguel, J.P.; Plaza, A. A New GPU Implementation of Support Vector Machines for Fast Hyperspectral Image Classification. *Remote Sens.* **2020**, *12*, 1257. [[CrossRef](#)]
- Lee, H.; Wang, J.; Leblon, B. Using Linear Regression, Random Forests, and Support Vector Machine with Unmanned Aerial Vehicle Multispectral Images to Predict Canopy Nitrogen Weight in Corn. *Remote Sens.* **2020**, *12*, 2071. [[CrossRef](#)]
- Xiu, L.; Liu, X. Current Status and Future Direction of the Study on Artificial Neural Network Classification Processing in Remote Sensing. *Remote Sens. Technol. Appl.* **2003**, *18*, 339–345.
- Jia, Y. Application of Artificial Neural Network to Classification of Multi-source Remote Sensing Imagery. *Bull. Surv. Mapp.* **2000**, *7*, 7–8.
- Hou, H.; Hou, J.; Huang, C.; Wang, Y. Retrieve Snow Depth of North of Xinjiang Region from ARMS 2 Data based on Artificial Neural Network Technology. *Remote Sens. Technol. Appl.* **2018**, *33*, 241–251.
- Liu, Y.; Wang, L.; Zhang, B.; Men, J. Scene-level land use classification based on multi-features soft-probability cascading. *Trans. Chin. Soc. Agric. Eng.* **2016**, *32*, 266–272.
- Zhang, F.; Li, W.; Lu, L.; Zhang, Q.; Kang, L. Technologies of extracting land utilization information based on SVM method with multi-window texture. *J. Remote Sens.* **2012**, *16*, 67–78.
- Chen, Y.; Zhang, T.; Dou, P.; Dong, L.; Chen, H. Error Sources and Post Processing Method for Land Use/cover Change Estimation of Dongguan City based on Landsat Remote Sensing Imagery with SVM. *Remote Sens. Technol. Appl.* **2017**, *32*, 893–903.
- Liu, Y.; Du, P.; Zheng, H.; Xia, J.; Liu, S. Classification of China Small Satellite Remote Sensing Image based on Random Forests. *Sci. Surv. Mapp.* **2012**, *37*, 194–196.
- Yao, Y.; Liang, H.; Li, X.; Zhang, J.; He, J. Sensing Urban Land-Use Patterns by Integrating Google Tensorflow and Scene-Classification Models. *arXiv* **2017**, arXiv:1708.01580. [[CrossRef](#)]
- Wang, M.; Zhang, X.; Wang, J.; Sun, Y.; Jian, G.; Pan, C. Forest Resource Classification based on Random Forest and Object Oriented Method. *Acta Geod. Cartogr. Sin.* **2020**, *49*, 235–244.
- Zhao, D.; Gu, H.; Jia, Y. Comparison of Machine Learning Method in Object-based Image Classification. *Sci. Surv. Mapp.* **2016**, *41*, 181–186.
- Fragou, S.; Kalogeropoulos, K.; Stathopoulos, N.; Louka, P.; Srivastava, P.K.; Karpouzas, S.P.; Kalivas, D.P.; Petropoulos, G. Quantifying Land Cover Changes in a Mediterranean Environment Using Landsat TM and Support Vector Machines. *Forests* **2020**, *11*, 750. [[CrossRef](#)]
- Chakhar, A.; Ortega-Terol, D.; Hernández-López, D.; Ballesteros, R.; Ortega, J.F.; Moreno, M.A. Assessing the Accuracy of Multiple Classification Algorithms for Crop Classification Using Landsat-8 and Sentinel-2 Data. *Remote Sens.* **2020**, *12*, 1735. [[CrossRef](#)]

16. LaRocque, A.; Phiri, C.; Leblon, B.; Pirotti, F.; Connor, K.; Hanson, A. Wetland Mapping with Landsat 8 OLI, Sentinel-1, ALOS-1 PALSAR, and LiDAR Data in Southern New Brunswick, Canada. *Remote Sens.* **2020**, *12*, 2095. [\[CrossRef\]](#)
17. Morell-Monzó, S.; Estornell, J.; Sebastiá-Frasquet, M.-T. Comparison of Sentinel-2 and High-Resolution Imagery for Mapping Land Abandonment in Fragmented Areas. *Remote Sens.* **2020**, *12*, 2062. [\[CrossRef\]](#)
18. 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. [\[CrossRef\]](#)
19. Liu, X.; He, J.; Yao, Y.; Zhang, J.; Liang, H.; Wang, H.; Hong, Y. Classifying urban land use by integrating remote sensing and social media data. *Int. J. Geogr. Inf. Sci.* **2017**, *31*, 1675–1696. [\[CrossRef\]](#)
20. 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. [\[CrossRef\]](#)
21. Tu, Y.; Chen, B.; Zhang, T.; Xu, B. Regional Mapping of Essential Urban Land Use Categories in China: A Segmentation-Based Approach. *Remote Sens.* **2020**, *12*, 1058. [\[CrossRef\]](#)
22. Su, M.; Guo, R.; Chen, B.; Hong, W.; Wang, J.; Feng, Y.; Xu, B. Sampling Strategy for Detailed Urban Land Use Classification: A Systematic Analysis in Shenzhen. *Remote Sens.* **2020**, *12*, 1497. [\[CrossRef\]](#)
23. Zong, L.; He, S.; Lian, J.; Bie, Q.; Wang, X.; Dong, J.; Xie, Y. Detailed Mapping of Urban Land Use Based on Multi-Source Data: A Case Study of Lanzhou. *Remote Sens.* **2020**, *12*, 1987. [\[CrossRef\]](#)
24. 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. [\[CrossRef\]](#)
25. Gong, P.; Li, X.; Wang, J.; Bai, Y.; Chen, B.; Hu, T.; Liu, X.; Xu, B.; Yang, J.; Zhang, W.; et al. Annual maps of global artificial impervious area (GAIA) between 1985 and 2018. *Remote Sens. Environ.* **2020**, *236*, 111510. [\[CrossRef\]](#)
26. Gong, P.; Liu, H.; Zhang, M.; Li, C.; Wang, J.; Huang, H.; Clinton, N.; Ji, L.; Li, W.; Bai, Y.; 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. [\[CrossRef\]](#)
27. Pałasz, K.W.; Zawadzki, J. Sentinel-2 Imagery Processing for Tree Logging Observations on the Białowieża Forest World Heritage Site. *Forests* **2020**, *11*, 857. [\[CrossRef\]](#)
28. Yue, W.; Xu, J. Impact of Human Activities on Urban Thermal Environment in Shanghai. *Acta Geogr. Sin.* **2008**, *63*, 247–256.
29. Liu, X.; Zhou, Y.; Yue, W.; Li, X.; Liu, Y.; Lu, D. Spatiotemporal patterns of summer urban heat island in Beijing, China using an improved land surface temperature. *J. Clean. Prod.* **2020**, *257*, 120529. [\[CrossRef\]](#)
30. Liu, X.; Yue, W.; Yang, X.; Hu, K.; Zhang, W.; Huang, M. Mapping Urban Heat Vulnerability of Extreme Heat in Hangzhou via Comparing Two Approaches. *Complexity* **2020**, *2020*, 1–16. [\[CrossRef\]](#)
31. Lin, P.; Li, X.; Yang, X.; Xiao, L. Accuracy Analysis on the Urban Surface Temperature Evaluation by Use of Landsat 8 Data. *J. Fujian Norm. Univ.* **2018**, *34*, 16–24.
32. Sobrino, J.A.; Jiménez-Muñoz, J.C.; Paolini, L. Land surface temperature retrieval from LANDSAT TM 5. *Remote Sens. Environ.* **2004**, *90*, 434–440. [\[CrossRef\]](#)
33. He, C.; Shi, P.; Li, J.; Chen, J.; Pan, Y.; Li, J.; Zhuo, L.; Ichinose, T. Restoring urbanization process in China in the 1990s by using non-radiance-calibrated DMSP/OLS nighttime light imagery and statistical data. *Chin. Sci. Bull.* **2006**, *51*, 1614–1620. [\[CrossRef\]](#)
34. Wang, L.; Fan, H.; Wang, Y. Improving population mapping using LuoJia 1-01 nighttime light image and location-based social media data. *Sci. Total Environ.* **2020**, *730*, 139148. [\[CrossRef\]](#)
35. Ren, Z.; Liu, Y.; Chen, B.; Xu, B. Where Does Nighttime Light Come From? Insights from Source Detection and Error Attribution. *Remote Sens.* **2020**, *12*, 1922. [\[CrossRef\]](#)
36. Breiman, L.; Friedman, J.; Stone, C.J.; Olshen, R.A. *Classification and Regression Trees*; Chapman & Hall (Wadsworth, Inc.): New York, NY, USA, 1984.
37. Liu, S.; Zhu, H. Object-oriented Land Use Classification based on Ultra-high Resolution Images Taken by Unmanned Aerial Vehicle. *Trans. Chin. Soc. Agric. Eng.* **2020**, *36*, 87–94.
38. Ma, H.; Gao, X.; Gu, X. Random Forest Classification of Landsat 8 Imagery for the Complex Terrain Area based on the Combination of Spectral, Topographic and Texture Information. *J. Geo-Inf. Sci.* **2019**, *21*, 59–71.
39. Cortes, C.; Vapnik, V. Support Vector Network. *Mach. Learn.* **1995**, *20*, 273–297. [\[CrossRef\]](#)

40. Zhang, Y.; Liu, J.; Wan, L.; Qi, S. Land Cover/Use Classification Based on Feature Selection. *J. Coast. Res.* **2015**, *73*, 380–385. [[CrossRef](#)]
41. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning Representations by Back-propagating Errors. *Nature* **1986**, *323*, 533–536. [[CrossRef](#)]
42. Aitkenhead, M.J.; Aalders, I.H. Classification of Landsat Thematic Mapper Imagery for Land Cover Using Neural Networks. *Int. J. Remote Sens.* **2008**, *29*, 2075–2084. [[CrossRef](#)]
43. Lin, F.R.; Shaw, M.J. Active Training of Backpropagation Neural Networks Using the Learning by Experimentation Methodology. *Ann. Oper. Res.* **1997**, *75*, 105–122. [[CrossRef](#)]
44. Buscema, M. Back propagation neural networks. *Subst. Use Misuse* **1998**, *33*, 233–270. [[CrossRef](#)] [[PubMed](#)]
45. Zheng, H.; Wu, C.; Zheng, S.; Zhuo, Y.; Zhang, Q. The Spatial Consistency between Compact City and Mixed Land Use Development: A Case Study of Shanghai. *China Land Sci.* **2016**, *30*, 35–42.

© 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/>).