

Identification and Classification of Urban PLES Spatial Functions Based on Multisource Data and Machine Learning

Jingying Fu ^{1,2}, Ziqiang Bu ^{1,2}, Dong Jiang ^{1,2,3} and Gang Lin ^{1,2,*}

¹ Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China

² College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100049, China

³ Key Laboratory of Carrying Capacity Assessment for Resource and Environment, Ministry of Natural Resources, Beijing 100101, China

* Correspondence: ling@reis.ac.cn; Tel.: +86-188-1175-9109

Abstract: Production space, living space, and ecological space (PLES) increasingly restrict and influence each other, and the urban PLES conflict significantly affects the sustainable development of a city. This study extracts multi-dimensional features from high-resolution remote sensing images, building vectors, points of interest (POI), and nighttime lighting data, and applies them to urban PLES feature recognition, dividing Ningbo into an agricultural production space, industrial and commercial production space, public living space, resident living space and ecological space. The specific research work was as follows: first, a convolutional neural network (CNN) was used to extract high-rise scene information from high-resolution remote sensing images; at the same time, through the geostatistical method, the building vector features, POI features, and night light features were extracted to express the economic and social characteristics of a city. Then, we used the nearest neighbor algorithm, decision-making tree algorithm, and random forest algorithm to train individual and combined features. Finally, random forest, which had the best training effect, was selected as the classifier in the fusion stage; as a result, the prediction accuracy rate reached 90.79%. The experimental results showed that the recognition model, based on multisource data and machine learning, had a good classification effect. Finally, we analyzed the current situation of the spatial distribution of PLES in Ningbo.

Keywords: multisource data; machine learning; PLES; random forest



Citation: Fu, J.; Bu, Z.; Jiang, D.; Lin, G. Identification and Classification of Urban PLES Spatial Functions Based on Multisource Data and Machine Learning. *Land* **2022**, *11*, 1824. <https://doi.org/10.3390/land11101824>

Academic Editor: Hossein Azadi

Received: 19 September 2022

Accepted: 14 October 2022

Published: 18 October 2022

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



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Since the beginning of the 21st century, the speed of urbanization in China has been accelerating. Due to rapid urbanization, human activities have interfered with land use, resulting in the depletion of resources, an increase in the environmental carrying capacity, and other problems [1,2]. A city is a complex regional system composed of different land uses such as vegetation, water bodies, and buildings. At present, as part of the urbanization of China's large and medium-sized cities, there have been widespread social problems such as population congestion, traffic congestion, shortages of water and power, severe domestic pollution, etc. The above problems will limit urbanization and reduce living standards [3,4]. The reason for these issues is the unreasonable layout of production and living space. In order to reduce costs and improve benefits, production departments form an agglomeration state. However, dense industrial areas such as blocks or strips will infringe on living and ecological space, reduce the quality of life for surrounding residents, and possibly cause pollution and other problems. At the same time, excessive real estate construction by developers will take up ecological space, so the conflict between PLES is increasingly serious. The disordered distribution of PLES in cities and the resulting deterioration of the environment pose serious challenges to regional development and urban sustainable development [5–7].

PLES refers to the functions or characteristics of different types of territorial space formed under the joint influence of its internal components and external factors on the production and living activities of residents. It is a complex community affected by economic, social, and geographical factors. Economic development has caused a drastic transformation of national spatial function, making it more diversified and complex [6,8,9]. Therefore, we need to explore the rules of PLES function distribution in cities, so as to improve people's living standards and provide a scientific basis for sustainable economic, social, and environmental development.

The study of land use/cover change in typical regions has been a hot topic in geographical science since the 1990s. Remote sensing is indispensable for the identification of urban land types. Scholars from all over the world have carried out studies on the acquisition of urban land use information based on remote sensing imagery [10–12]. For example, several national land cover mapping programs became operational in the 1990s and the early 2000s [13,14], Zeferino et al. used Landsat 8 satellite remote sensing images to study the impact of environmental data on urban land use [15]. In recent years, extraction methods based on high-resolution remote sensing images have been widely used in land use and land cover recognition [16–18]. The analysis of urban land use is usually carried out in three types of spatial units: pixel, object, and scene. Pixel-based and object-based units are commonly used to assess land cover and land use [19,20]. Scene based units are usually used to identify urban functional areas and urban land use patterns. Usually, road networks are used to segment land as basic analysis units. Many researches have applied scene-based classification methods, such as Latent Dirichlet Allocation (LDA) to extract physical features (such as features of spectrum, texture, and ground components) from high-resolution remote sensing images [21]. However, it explores the underlying semantic information of a block and ignores the spatial distribution relationship between objects. In high-density cities, many block may have the same physical attributes, but different functional attributes. Moreover, relying solely on remote sensing images ignores the socioeconomic characteristics of the blocks.

With the popularization of Internet technology and satellite positioning, a series of rich spatiotemporal data, such as point of interest data, social media data, and taxi trajectory data, have emerged. The emergence of these multisource data facilitates the identification of urban land structures. Many scholars have begun to apply emerging spatiotemporal data to classify and study urban space [22–24]. For example, Steiniger et al. tried to classify urban functional areas according to the morphological characteristics of buildings [25], while Tu et al. combined mobile phone data and social media data to infer urban functions and reveal hourly dynamics [26]. At the same time, there have been a number of studies that use emerging spatiotemporal data to classify urban spatial functions using a PLES classification system [27–30]. For example, Zhao et al., based on the POI data of Zhengzhou in 2007 and 2017, applied the quadrat proportion method and random forest to identify PLES within the city and achieved high classification accuracy [31]. Wang et al. applied POI data to construct a weight model of the “influence–spatial area,” established the corresponding relationship between POI data and PLES based on expert experience, quantitatively identified PLES in the Urumqi urban area, and analyzed its spatial distribution laws [32]. These studies used social and economic data such as POIs to identify and classify urban PLES, and achieved good classification results.

In summary, scholars from various countries have applied remote sensing data or emerging spatiotemporal data to classify urban functions, thereby gradually improving the accuracy and efficiency of urban functional area identification. However, the application of multisource data to existing research is limited to the use of POI data and remote sensing data, which does not allow for multimodal investigation of urban functional areas, or in-depth exploration of the scene characteristics, object characteristics, and socioeconomic characteristics of urban blocks. The combined three-dimensional (3D) building data are ignored in feature mining, but are important for accurately distinguishing between production space and living space. In addition, most of the current research on urban functional

areas divides cities into a series of functional areas, such as commercial areas, industrial areas, and residential areas. However, PLES, a classification system with planning significance, is rarely used to divide urban functional areas. Therefore, it is necessary to explore the use of multisource data and machine learning methods for urban PLES identification, which can help solve the problems of existing research being too subjective and the indistinguishability of production space and living space in urban areas, as well as provide a reference for other scholars studying urban PLES identification.

This study took Ningbo as the research area. By extracting scenes, objects, and economic characteristics contained in multisource data, the PLES function of the city can be identified by obtaining experience from the data itself. Then the ensemble learning model was applied to the multisource data fusion training to improve the prediction accuracy. Finally, the model with the highest training accuracy was selected to identify the PLES function of Ningbo's main urban area. We explored its spatial distribution law, and then put forward targeted optimization suggestions. The specific research contents are as follows: (1) feature extraction of city blocks based on multisource data: firstly, we used high resolution remote sensing image data, night light data, POI data, and building vector data to extract urban block features. Among them, high resolution remote sensing image data can reflect the texture, color, and other scene characteristics of blocks; night light data are an important reflection of urban night economy; POI data can intuitively map the details of social and economic activities in the area; building vector data can reveal the connections and differences between buildings, and can reflect the 3D characteristics of buildings. (2) PLES feature recognition model training and optimal model solving: by applying the nearest neighbor algorithm, decision tree algorithm, and random forest algorithm, different feature combinations extracted were trained, and then the optimal training model was screened out through iteration of each model parameter. (3) PLES identification and PLES distribution law analysis in Ningbo: The optimal model was applied to predict the urban spatial function, and the recognition results were merged. With the help of geographic information technology, the urban PLES distribution was visualized, the regional PLES distribution law was explored, and urban spatial optimization suggestions were put forward.

The mainstream methods for the identification of PLES can be divided into merging classification and quantitative measurement. The former belongs to qualitative research, and merges and classifies land use data to identify PLES mainly based on yearbook data, national land survey, and remote sensing image data. For the latter, more emphasis is placed on quantitative analysis, mainly using the spatial function value measurement function group to establish the land use function measurement system, and identifying the PLES through the quantitative measurement of the dominant functions of land use. The advantage of quantitative measurement method is that it can accurately identify the dominant function of PLES. The research in this paper belongs to the latter category. Therefore, according to the dominant function of land use, we established the following spatial classification system table of PLES (Table 1).

Table 1. PLES spatial classification system.

PLES		Land Use Type (Functional Description)
Primary Classification	Secondary Classification	
Production space	Agricultural production space	Paddy fields, dry land
	Industrial production space	Industry and mining, land use for transport construction
Living space	Urban living space	Urban land
	Rural living space	Rural residential land
Ecological space	Forest ecological space	Woodland, shrub land, open forest land, other woodland
	Grassland ecological space	Grassland with high, medium, and low coverage
	Aquatic ecological space	Canals, lakes, reservoir pit ponds, permanent glaciers and snowfields, tidal flats, bottomland
	Other ecological spaces	Sandy land, Gobi, saline alkali land, swamp land, bare land, bare rock land, and other unused land

According to the characteristics of urban PLES distribution, the production space of the PLES involved in the study was specifically divided into agricultural production space and industrial and commercial production space. Agricultural production space mainly refers to arable land, while industrial and commercial production space refers to industry, commerce, banking, and so on. Living space mainly includes residential living space and public living space. Residential living space mainly refers to villages and communities, while public living space refers to public facilities and areas used for scientific research and education, or other public services. Ecological space mainly includes green space, water, and unused land [8,33–35].

2. Materials and Methods

2.1. Study Area

The study area was Ningbo, which is an important port city on the southeastern coast of China. It is an economic center of gravity in the eastern part of Zhejiang Province and the southern flank of the Yangtze River Delta. To the north of Ningbo are Qiantang River and Hangzhou Bay, respectively; to the west and south are Shaoxing City and Taizhou City, respectively; and to the east is Zhoushan City across the sea (Figure 1). Ningbo has unique economic development advantages compared with inland cities, including unique transportation advantages and abundant marine resources. The port economy has shaped the urban spatial pattern of Sanjiang–Zhenhai–Beilun, which gathers the city's commercial and service activities and offers considerable residential functions. The rise of the port industry has brought about profound changes in the urban industrial structure and stimulated the construction of service industries and industrial and commercial supporting facilities in urban areas [36,37]. Therefore, it is necessary to explore the structure of Ningbo's territorial space utilization. At the same time, Ningbo's territorial spatial structure, with it being the southern economic center of China's Yangtze River Delta, has an important impact on the economic development of the entire urban agglomeration.

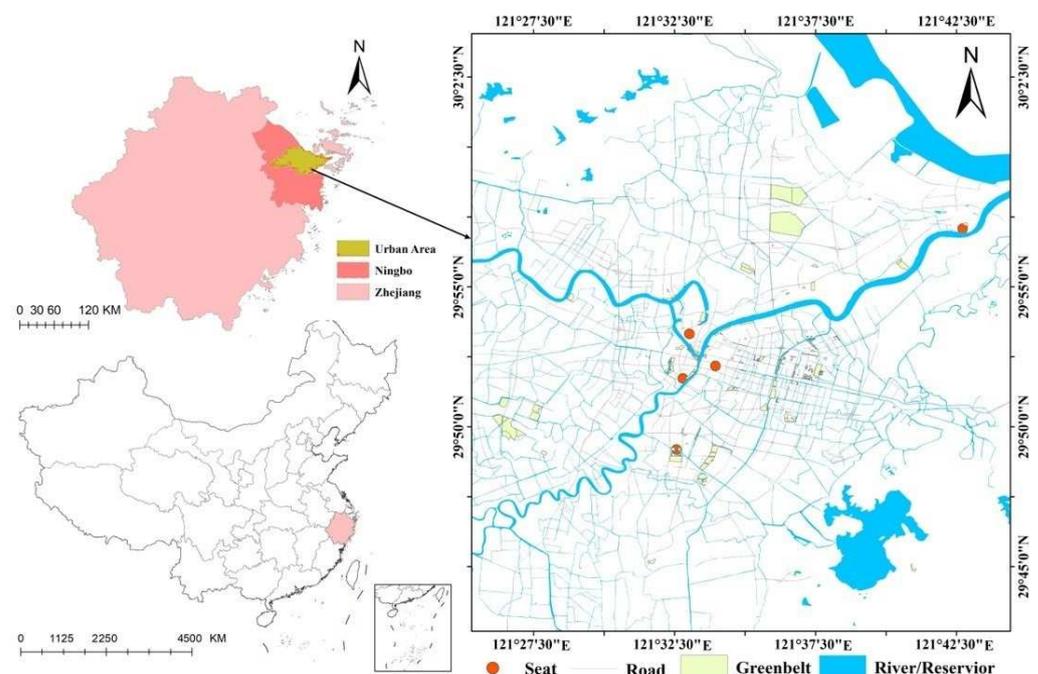


Figure 1. Location of study area.

Rapid population and economic growth have led to high demand for industrial and residential space, resulting in an unprecedented evolution of the PLES pattern, posing serious challenges to ecosystems. The environmental risks in Ningbo have increased significantly in the past 40 years, and in the next few decades, with ongoing economic

and social development, the demand for land and space resources will increase these risks. At the same time, resource bottlenecks and environmental pressures will become increasingly prominent. The contradiction is becoming more and more obvious [38–40]. For the sustainable development of Ningbo, rational planning of urban PLES layout should be considered.

2.2. Research Data

The data involved in this study mainly include road network data, high resolution remote sensing images, night light data, POI data, and building vector data: (1) Ningbo city road vector data can be downloaded from OpenStreetMap (<https://www.OpenStreetMap.org>, accessed on 12 February 2022). There are nine levels of roads in this road database; we used three main types of road data: primary, secondary, and tertiary, filtered and topologically edited to generate a road network capable of segmenting urban areas into many small cells. (2) The high-resolution remote sensing image data include the remote sensing image training data, namely the UC-Merced dataset, and the image data of Ningbo captured from Google Maps according to administrative divisions. We grabbed 18-level images, the spatial resolution of which was 1.07 m; the image format was tiff, and the image location information and translation and rotation parameters of the coordinate system were provided as .tfw files. (3) This paper captured the POI data of Ningbo in 2020 from the Gaode Map (<https://gaode.com/>, accessed on 12 February 2022) platform. After cleaning and classification, nearly 100,000 POI data records were obtained, including the location information of POIs and the type information of the point. The original POIs data was composed of 23 categories, including business housing, shopping, life services, catering services, science, education, and cultural services, etc. (4) Building vectors were from Baidu Map (<https://map.baidu.com/>, accessed on 12 February 2022); we obtained 66,072 in total. The data contain the area, length, number of floors, and other relevant characteristics for each building. (5) The night light data of Ningbo City were obtained from the official LuoJia-1 website. The LuoJia-1 satellite carries a high-sensitivity night light remote sensing camera and has night light imaging ability with a resolution of 130 m and a width of 260 km, which makes it helpful for analyzing the regional macroeconomic situation [41]. The experimental data samples in the study area are shown in Figure 2.

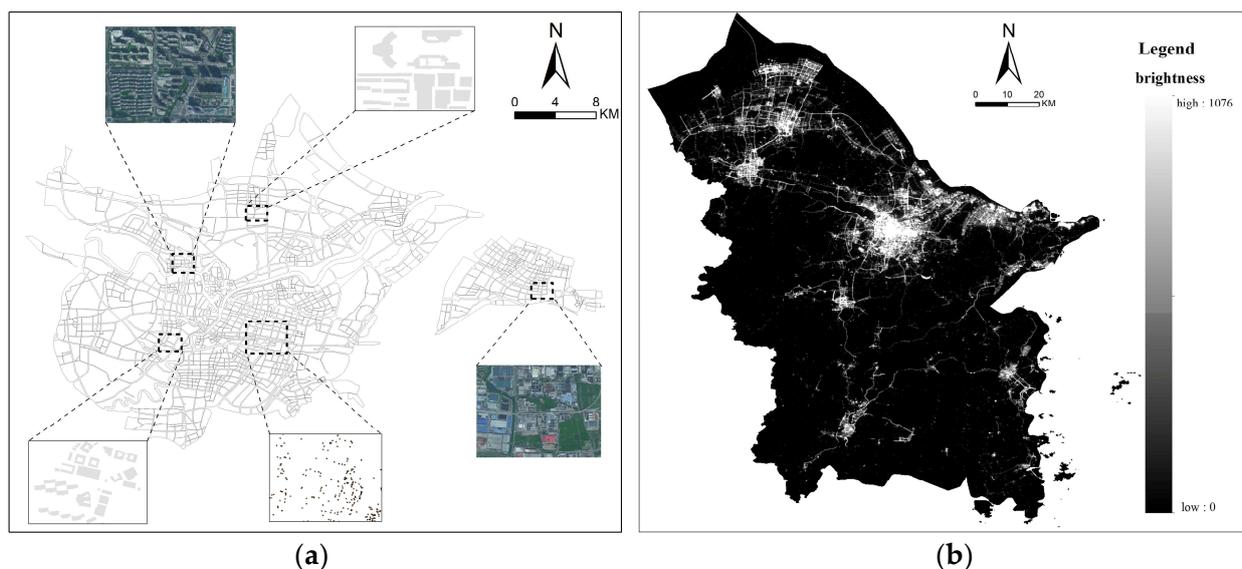


Figure 2. (a) The research units segmented by OpenStreetMap network, remote sensing images, POIs and building footprints in some sample units. (b) Night lights in Ningbo City.

2.3. Research Methods

In order to make full use of the rich features of multisource spatiotemporal data, this study integrated multisource data and machine learning models to build a comprehensive recognition model. The specific technical framework was as follows.

The first step was feature extraction of multisource data, including the following: (1) measuring physical properties of buildings based on planar vector data of buildings. We collected the area, perimeter, number of floors, and structure ratio of all buildings, and then counted the sum, average, and standard deviation of building features in each block (segmented by road network)—a total of 12 features. (2) We counted the number of various POIs in each block to characterize the characteristics of the social and economic activities of the block. (3) With the help of GIS grid division statistics method, we calculated 10 statistical values. such as the sum and standard deviation of nighttime light brightness, in each block to reflect the nighttime economic characteristics of the block. (4) Using CNN, we detected hidden spatial information from remote sensing images. The number of neurons in the output layer of the convolutional neural network determines the length of the extracted features. After extracting these features, we labeled the selected typical samples by visual interpretation. Then we used the nearest neighbor, decision tree, and random forest algorithms to train individual and combined features to compare the accuracy. Finally, the random forest algorithm, which had a better ensemble learning effect, was used to predict the functional area, and the weight of different features in the classification task was calculated to analyze the contribution of different features in the classification process. The method flowchart is shown in Figure 3.

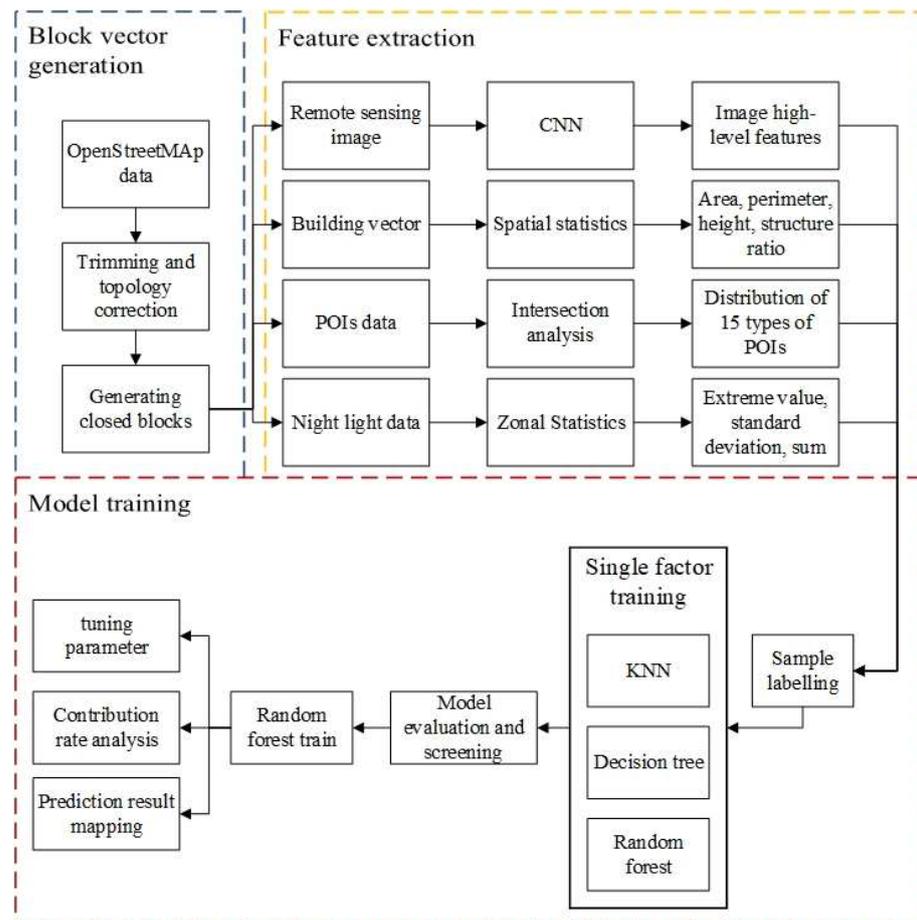


Figure 3. Technical framework.

The spatial statistics we conducted in the experiment are based on version 10.2 of ArcGIS software, in which the ArcMap platform can store, query, count spatial elements and calculate spatial data in the form of raster or vector. We used tools such as ‘ZonalStatisticsAsTable’ and ‘Spatial Join’ in ArcMap to assist spatial data processing.

2.3.1. Feature Extraction of Urban Blocks

According to the different dimensions of data features, we divided the multisource data features used in this study into scene features and socioeconomic features. Scene features refer to the neighborhood features reflected by high resolution images, and socioeconomic features were derived from the building data and POI data. Different extraction methods were adopted for different types of data.

(1) Scene feature extraction: we used CNN to extract the scene features of urban blocks. The common convolutional neural network model includes not only the input layer and the output layer, but also the convolutional layer, the pooling layer, and the fully connected layer. They are connected to each other by stacking and combining to form the whole model [42,43]. The convolutional layer uses the convolution kernel to use the output of the previous layer as the input to perform the convolution operation, and outputs the feature map obtained after the action of this layer, so as to use this mechanism to complete hierarchical feature extraction. There are multiple convolution kernels in the convolutional layer. The convolution kernel performs point multiplication with the corresponding part of the feature map through the sliding window method. After adding the bias, the feature extraction output of the convolution kernel and the local feature map value are obtained. When the entire region of the input feature map is passed, the convolution operation of this layer is completed (as in Equation (1)):

$$X_j^i = f(\sum_{j \in M_j} (X_j^{i-1} \times k_{j,i}^l) + b_j^l), \quad (1)$$

where X_j^i represents the j feature matrix output by the i layer network, M_j represents the output result set of the previous layer, $k_{j,i}^l$ represents the convolution kernel, b_j^l is the offset, and $f(x)$ represents the excitation function.

After the convolution operation is completed, a pooling operation—that is, downsampling, which is one of the commonly used methods in the field of image processing—is generally performed. The pooling layer also divides the input feature map into several subparts by sliding the window, and each subregion is nonlinearly pooled to obtain the output value. This method completes the downsampling of the input layer. Commonly used pooling methods include maximum pooling, which is the maximum value of the output subregion pixels; the average pooling principle calculates the average value of all pixel values in a region, with its value as the average output of the subregion. The pooling layer reduces the consumption of computing resources by the model through downsampling, and increases the robustness of the model. The features obtained after several convolution layers and pooling layers are used, and then the multidimensional matrix is expanded into a vector, which is classified by the fully connected layer. Each neuron in the fully connected layer will be calculated with all the neurons in the previous layer according to Equation (2) to obtain the output value of the neuron. Compared with the local connection of the convolutional layer, there are more parameters in the fully connected layer, which increases the difficulty of training.

$$y_i = f(W_i x_i + b_i), \quad (2)$$

where W_i is the weight of the neuron, x_i is the input of the neuron, and b_i is the bias.

In this study, the Inception ResNet V2 network was used for scene feature extraction of blocks—that is, the Inception structure is combined with the residual structure of ResNet [44]. The network structure is shown in Figure 4. As we can see, Inception ResNet V2 network uses several maximum down sampling layers for data dimensionality reduction before convolution, and then inputs data to different branches for convolution and

pooling operations. After that, features of different branches are channeled. The extracted different features are fused. The network continuously passes the features through Inception structure and ResNet structure, and then executes a mean downsampling operation to input the features into the Dropout layer for feature screening to prevent overfitting. Finally, we obtain the desired feature parameters through the SoftMax layer.

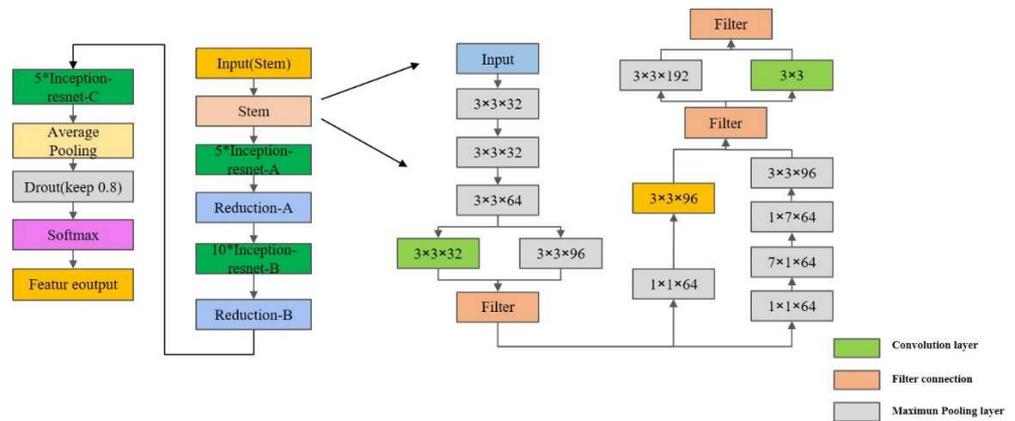


Figure 4. Inception ResNet V2.

In addition, we first used the UC-Merced dataset to train the Inception ResNet V2 model with pretrained weights. This dataset includes 21 scene types; each scene type has 100 images, the maximum size of each image is 256×256 images, and the maximum spatial resolution is 0.3 m. Some images in this dataset are shown in Figure 5.



Figure 5. Sample images of each scene in UC-Merced dataset ((a): cultivated land; (b): Road intersection; (c): residential land; (d): forest land).

(2) Extraction of social and economic characteristics: (a) there is a close relationship between the physical properties of a building and the function of a block. In the ArcGIS software, the area and perimeter of each building vector were calculated, and then the

structural ratio of the building vector was calculated through the ratio of the perimeter to the area. Finally, we used ArcGIS's spatial connection tool to obtain statistics such as the sum, mean, standard deviation, and other attributes of the building's area, perimeter, structure ratio, and number of floors corresponding to each block. (b) The functional types of urban PLES are not only related to geographical location, but also to a series of economic activities. This study used POIs to extract socioeconomic characteristics to enrich parcel information. First, the POIs were redivided into 15 categories: public facilities, catering services, education and cultural services, shopping services, companies and enterprises, medical services, accommodation services, commercial residential, living services, green space landscape, transportation facilities services, finance/insurance services, sports and leisure services, government agencies, and social organizations. Then, the number of different types of POIs in each block was counted through the spatial connection method in ArcGIS. Finally, the proportion of each type of POIs was calculated to characterize the socioeconomic characteristics of the neighborhood. (c) The distribution of POI data can represent the relationship between specific economic activity subjects and PLES types of urban blocks, but cannot express the overall economic characteristics of blocks and economic differences within blocks, while nighttime light data can reflect detailed nighttime economic characteristics. Using the city block vector in ArcGIS to perform regional statistics on the night light raster data, statistical information such as the mean, sum, and standard deviation of the night lights in each block can be obtained.

2.3.2. Construction of the Identification Model of Urban Spatial Functional Areas

The matching of multisource data is the basis for the construction of the urban spatial functional area identification model. The attempt and adjustment of different machine learning are the main means of model optimization in this paper.

(1) Training data preparation

The purpose of this research is to construct a scientific and effective ensemble learning method for urban PLES feature recognition, in which the most critical problem is the matching of multisource data. This paper extracts the scene, object, and economic features of blocks from remote sensing data and socioeconomic data. For the extraction of scene features, the problem of image size needs to be dealt with first. Classic convolutional neural networks have fixed inputs and outputs. Traditional AlexNet, GoogLeNet, and VGGNet all require fixed-size image input. Inception ResNet v2, used in this article, requires 229×229 image input; however, the size of the remote sensing images captured by the block vector is often much larger. How to cut the block image into a convolutional neural network model for feature extraction, and then match the extracted results to each block, becomes the core of the matching algorithm in this paper. We adopted the following solution: firstly, the block vector was rasterized according to the unique value (ID), and the image map with the same resolution, geographic coordinates, and projection information as the remote sensing image was obtained. Then, this image as a band was superimposed onto the high-resolution remote sensing image of Ningbo City to obtain a combined raster dataset. Then, a regular grid was used to trim this raster dataset to obtain 35,229 lattice images of size 229×229 . They can be extracted by a convolutional neural network, and the extracted features are recorded by the block ID corresponding to the mode of the newly added band. Finally, the image features corresponding to the same ID number were averaged to obtain the image features corresponding to each block.

The feature extraction of social perception data such as POI data and building vector directly extracted the corresponding features of each block through the spatial connection function of ArcGIS. For night light data, the zoning statistics function of ArcGIS was used to calculate the mean value, standard deviation, and mean value of night light brightness in each block. Finally, the ID number of the block was used as the unique block, and different features are connected to obtain the combined features.

(2) Machine learning approach

There are many feature dimensions extracted from multi-source data, while the number of samples is small, so we need to select a model that can tolerate outliers and is good at handling multi feature sample classification. According to the above requirements, we selected the nearest neighbor algorithm, decision tree algorithm, and random forest algorithm; they are easy to operate and understand for the classification of urban PLES spatial function types.

Nearest neighbor algorithm: first specify a training dataset, and then search for the nearest K samples from the training sample according to the latest input sample. All the samples were classified into this type [45]. The nearest neighbor algorithm is also a lazy learning algorithm—that is, it does not require the classifier to be trained with training samples before the prediction.

Decision tree: A tree structure (which can be binary or nonbinary) is used to make decisions. The basic process of decision making using a decision tree is to input all samples to be tested and their corresponding feature attributes to the root node of decision tree, then select output branches according to their values, and repeat this process until reaching the leaf node; the feature categories stored in the leaf node of the tree represent the analysis results [46]. The decision tree algorithm is widely used because of its low time complexity, its ability to process data with irrelevant features, and its ability to process small datasets.

Random forest algorithm: The component unit of random forest is decision tree. For classification problems, each decision tree is a classifier that learns and makes predictions independently. It synthesizes the results of all classifications and then starts voting to produce a more accurate classification. Each tree is trained with different training samples from the original data, accounting for about two-thirds of the original data; the rest is used to calculate the overall error and estimate the generalization performance of the tree [45]. In the process of data sampling, features are also randomly sampled to ensure the independence of the predictor, so the random forest system can process a large number of high-dimensional data and judge the interaction between various characteristics.

(3) Model training

First, the samples were preprocessed and labeled. Based on the remote sensing image base map of Ningbo, the block vector of Ningbo was labeled by a manual visual method. We obtained 328 samples, including 32 samples of agricultural production space, 99 samples of industrial and commercial production space, 157 samples of residential living space, 12 samples of public living space, and 28 samples of ecological space. Then the selected samples were randomly divided into a training set and a test set according to the ratio of 8:2. After comparing with different classifiers (nearest neighbor, decision tree), random forest was selected as the final classifier in the fusion stage. It has been proven that random forest has a better classification effect in ensemble learning.

3. Results

3.1. Classification Results of Multisource Features Based on Different Models

We analyzed the performance of each type of extracted features on the validation dataset when they participated in machine learning model training alone, and compared the accuracy of different machine learning models. Table 2 shows the validation accuracy of high-resolution remote sensing image data, points of interest, building vectors, and night lights on the nearest neighbor algorithm, decision tree algorithm, and random forest algorithm, respectively.

Table 2. PLES layout of Ningbo.

	Google Image	POIs	Buildings	Night Light
Nearest neighbor	68.10%	76.83%	56.93%	58.04%
Decision tree	72.88%	76.83%	60.36%	64.71%
Random forest	78.95%	71.91%	55.91%	69.44%

Looking longitudinally, among the different models, the accuracy of the nearest neighbor classifier is 68.10% when using image features alone, while the accuracy of decision tree and random forest are 72.88% and 78.95%, respectively. This shows that the nearest neighbor classification had lower accuracy, and the performance on the validation set was much lower than that of random forest, with a difference of 10.85% in accuracy. However, the nearest neighbor algorithm did not perform the worst on all data. On POI data, the nearest neighbor algorithm was more accurate than random forest. The classification accuracy of the nearest neighbor was relatively low overall, which may be because the nearest neighbor algorithm has a low prediction accuracy for rare categories when the samples are imbalanced. The accuracy of decision tree is between nearest neighbor and random forest, so it is more suitable for the classification of samples with multiple uncorrelated features than the nearest neighbor algorithm, and the training process and results of decision tree are interpretable.

From a horizontal perspective, among different data, the classification accuracy of building vector data alone is the lowest, and the lowest accuracy is only 55.91% among the three classifiers. According to the original data, there is no obvious distinction between the mean and standard deviation of the area, perimeter, structure ratio, and number of floors of blocks with different functional types. In other words, there is little difference between the mean values of the four features of building vector, which means it is not suitable for the classification of functional areas. This is because, in ecological space and agricultural production space, there were also scattered buildings such as houses and factories. When calculating their mean value, the denominator was the number of building vectors in the block. Therefore, in terms of the average, there is little difference in the architectural characteristics of different functional areas. Using only POIs can achieve 71.19% accuracy. It can be seen that the distribution of POI data had a strong correlation with the functional types of blocks, but its classification accuracy was not high in general. Shopping services include large supermarkets, shopping malls, special commercial blocks, etc. Such POI's dense blocks can be inferred as commercial districts, namely industrial and commercial production spaces. However, some convenience stores, personal product stores, and cosmetics stores were often close to residential areas, which led to a lower distinction between living space and production space. In addition, the machine learning model was prone to misclassification for rural residential areas and some blocks with missing POI data; because there were few POI data in these areas, it was easy to classify them as agricultural production space or ecological space.

3.2. Analysis of Comprehensive Prediction Results Based on Ensemble Learning

High-resolution remote sensing image data, POI data, night light data, and building vector data reflect the scene, object, and economic characteristics of each block, respectively. Through different combinations of multivariate features, we found that the prediction accuracy of building vector data was low, which reduced the overall accuracy of the ensemble learning model. Therefore, we used three different dimensions of high-resolution remote sensing image data, POI data, and night light data to apply the random forest algorithm with high training accuracy to comprehensively classify and predict urban PLES functions. In the final training, the training and validation datasets were split in a ratio of 8:2, and we used the grid search algorithm to tune the hyperparameters of the random forest. Among the many hyperparameters, the number of decision trees in the random forest had the greatest impact on the accuracy of the model. In the training stage of model, when the number of decision trees is less than 20, the model accuracy increases at a very fast rate. When the number of decision trees increases to 20, the accuracy of the model is improved to more than 82.5%, then the model fluctuates with the growth of the number of trees. The highest value of the entire curve appears on the horizontal axis when the number of decision trees is 72, and the model has the highest accuracy, reaching 90.79%.

The model constructed a total of 72 decision trees, with each decision tree predicted according to the characteristics of the sample. In this model, different features had different

effects on the results. As can be seen from Figure 6a, the contribution rate of image features was the highest, the features of POIs were second, and the contribution rate of building vectors was the lowest, which corresponds to the training accuracy of different data sources. Datasets that performed well in individual prediction also had a higher contribution rate in combined predictions. According to the confusion matrix of the prediction results of the ensemble learning model on the validation set (Figure 6b), the classification effect of the residential living space and production space, including agricultural production space and industrial and commercial production space, is good, and there is almost no wrong classification. However, the proportion of errors in the classification of public living space and ecological space is relatively high. Due to the small area and less distribution of these two types of space, which are basically surrounded by residential areas, a large proportion of samples are misclassified into residential areas or commercial areas, which also limits the validation accuracy of the whole model.

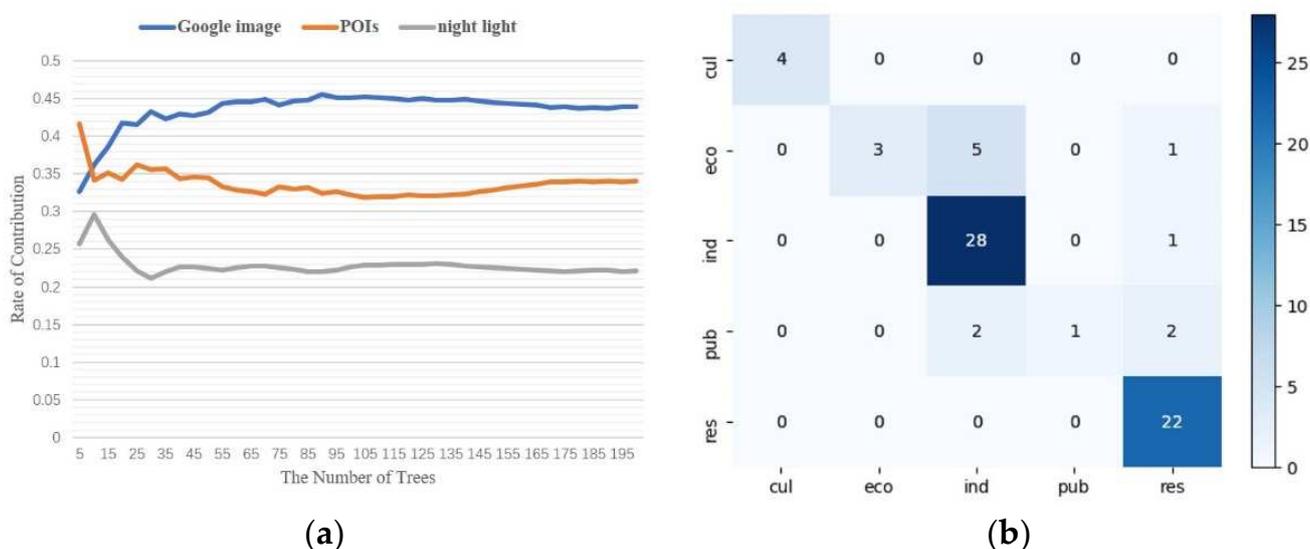


Figure 6. (a) Contribution rate of different data in the forecasting process. (b) Validation set confusion matrix.

Although high-resolution image data and POI data have a high contribution to the classification of urban functional areas, the features extracted from high-resolution remote sensing image data and POI data lose their interpretation significance for functional areas. Therefore, we counted the mean values of night light in different functional area types of blocks in the prediction results, and plotted box plots as shown in Figure 7 to test and explain the role of night light data in PLES partition. As can be seen from Figure 7, among the five types of functional areas, the mean value of nighttime light brightness in agricultural production space is the lowest, and the brightness value of industrial and commercial production space is also low. On the contrary, the living space including public living space and residents living space at night light brightness value is higher. This is consistent with the actual law: urban residents work in the living space during the day and return to the living space to rest at night. On the other hand, the luminance values of each block in agricultural living space, ecological space and residential living space are relatively high in outlier samples, which to some extent implies that the classification of urban functional areas cannot be based only on the luminance values of nighttime lights. At the same time, some outliers such as high-brightness blocks in agricultural production space and ecological space may be the basis for misclassification.

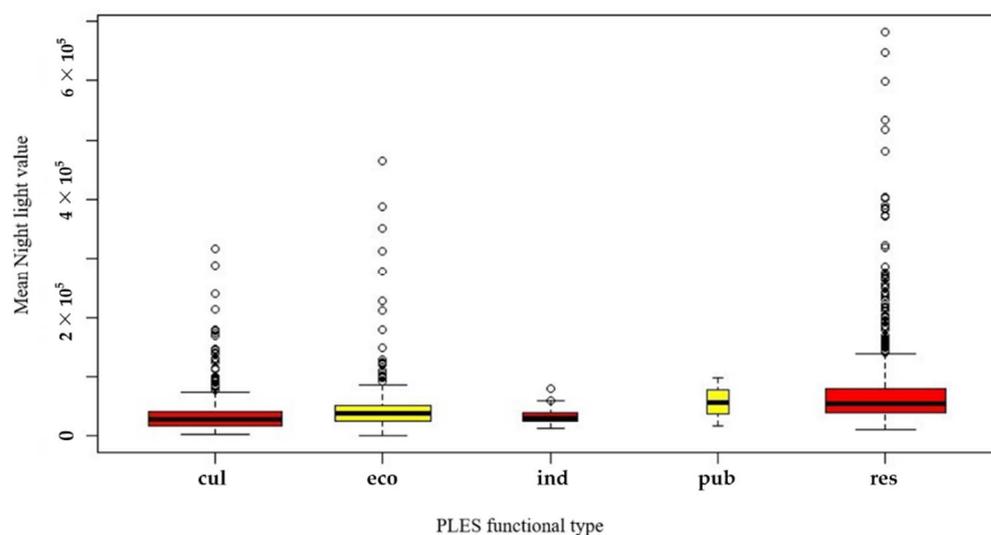


Figure 7. Mean night light value visualization of different functional areas. (Boxes of different colors represent different block types, box width represents the number of samples, scattered points on the same alignment with each box represent abnormal values, and two horizontal lines above and below the box represent quartile lines).

3.3. Analysis of the Spatial Pattern of PLES in Ningbo

The integrated learning model based on multisource spatiotemporal data identified the functional areas of Ningbo and then obtained the PLES layout map. As can be seen from Figure 8, the PLES function of the blocks in Ningbo was mainly residential space and industrial and commercial production space, while the public space and ecological space were scattered. The public space was mostly mixed with living space, in order to allow residents in the residential area to have rest and play opportunities nearby. The ecological space was mainly distributed in the suburbs or the urban–rural fringe, where there were fewer human activities and more green plants and water bodies, but the area was small and scattered. The residential space was mainly distributed in the ring road, in a cluster shape, surrounded by the production space. Industrial areas were mainly distributed on the edge of residential areas, which were generally regular plots, mainly to facilitate the communication between production departments and improve the efficiency of material transportation. In terms of its absolute location, the industrial zone was mainly distributed in the north of Ningbo, close to the port, which is convenient for foreign sales and raw material transportation. Beilun District, which is relatively independent of the main urban area, showed a similar pattern.

According to the proportion of each land type, 173 blocks were cultivated land, and the roads in the cultivated land area were sparse, which led to a large area of blocks with a functional area of cultivated land. The total area of cultivated land was 37,860 hectares, accounting for 47.9%. The ecological space covered 3075 hectares. The ecological space was mostly small blocks divided by roads and water networks, so the area was small, accounting for only 0.19%. The industrial and commercial production space covered an area of 1981 hectares, with the patches relatively large and concentrated; 764 blocks were residential space, with the distribution area dense with roads, resulting in relatively regular and finely divided blocks. However, the number of residential space blocks was the largest, exceeding the sum of several other functional area types, with an area of 8096 hectares. The public space block was small because the general public space was surrounded by residential or commercial areas. There were only 16 public spaces independently occupying the entire block, with an area of 175 hectares, and the area proportion was negligible.

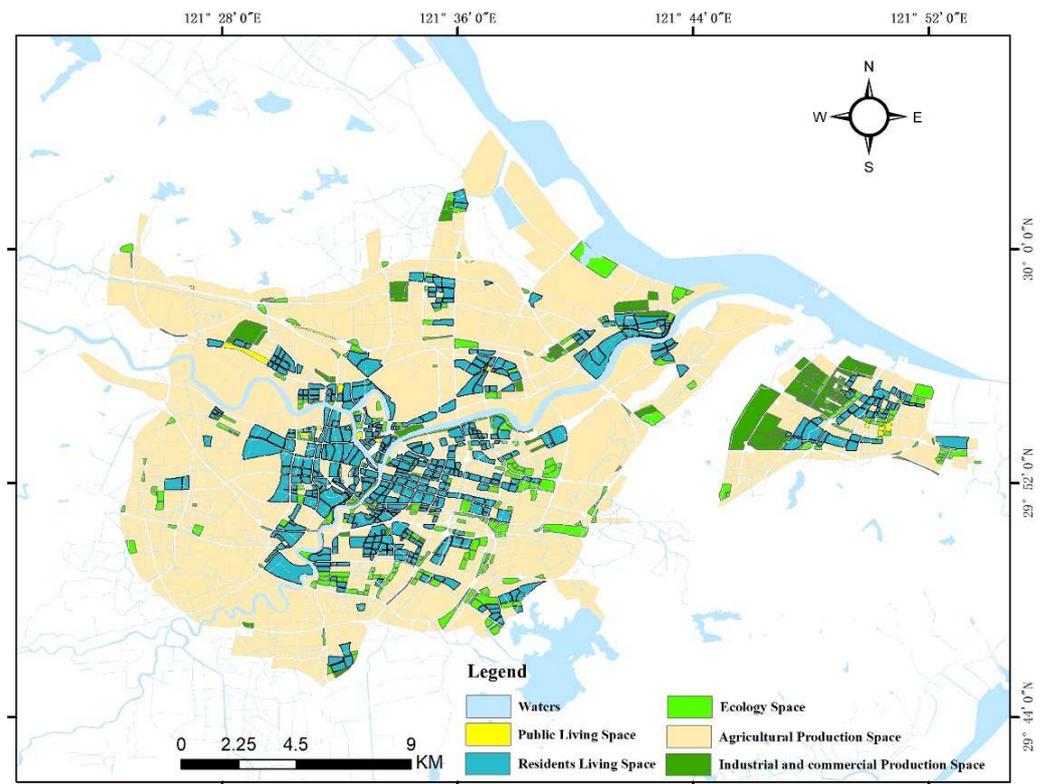


Figure 8. PLES layout of Ningbo City.

4. Discussions

The traditional land use classification ignores the functionality of land and the combination characteristics of different surface features, which is not conducive to the design of regional master planning [13,14]. In order to utilize emerging data and solve real-world problems in planning, this study proposed a functional classification method of urban PLES based on multi-source data and artificial intelligence. Taking Ningbo City as an example, the urban PLES classification was carried out, and a high classification accuracy was obtained, which verifies the accuracy of the method.

However, our article also has certain limitations. For example, in the scene feature extraction process of remote sensing images, only the common Inception-ResNet v2 model was used, instead of trying out different models and comparing the accuracy. Moreover, the UC-Merced dataset had few training samples, not enough to fully train such a deep neural network, which may lead to large errors in the scene feature extraction of blocks. In addition, in the phase of urban block functional area type recognition, due to the use of a coarser scale OSM road network, the number of blocks divided is limited, and the proportion of mixed functional areas is large, which affects the accuracy and credibility of the model. Therefore, in the future research, we will consider the identification of the mixing region and expand the samples of the mixing function for further experiments. Second, refine the classification system of function region, for example, the public living space can be subdivided into schools, hospitals, public facilities, public institutions, etc. At the same time, we will continue to collect refined land use data, complete the labeling of plots, and expand the sample size of each plot type. What is more, we will try to apply the method to more cities. Future research will try to improve the feature extraction process and try to use more effective machine learning algorithms to build a fusion learning model with higher identification accuracy.

5. Conclusions

The rapid development of remote sensing technology has led to scholarly interest in classifying urban PLES. At the same time, the emergence of social and economic big data such as POIs has greatly expanded the data basis for the identification of urban functional areas. Different from traditional functional area classification, this paper proposed a PLES classification method for urban functional areas, and added night light data and building data to make up for the lack of POI data for generalizing the overall economic characteristics of the block. We used a machine learning model integrating multisource data to identify functional areas, so that the comprehensive classification model could better reflect various characteristics of the block, and effectively improved the classification accuracy of the comprehensive prediction model. Our main findings were as follows:

- (1) We collected high-resolution remote sensing image data, POI data, building vector data, and night light data that reflected the characteristics of Ningbo's block scene and socioeconomic characteristics, and extracted multisource data features through convolutional neural network and GIS technology. The features extracted based on multisource data better reflected the characteristics of Ningbo city blocks from multiple dimensions, so were more effective for identifying urban PLES types.
- (2) We compared the accuracy of several commonly used machine learning models (nearest neighbor, decision tree, and random forest) for functional area classification. The random forest algorithm is a commonly used ensemble learning algorithm that can fuse features of different dimensions. Compared with the other two models, the prediction accuracy of random forest was greatly improved.
- (3) The accuracy of the respective training models of the multisource data and the contribution rate in the integrated prediction model were compared: because the actual building vectors in some areas were not marked, the training accuracy of the building vector data was low. In addition, the number of different training data features was not balanced. There were only 10 features of night lights, which was much lower than the number of features for high-resolution remote sensing image data and POIs, so the prediction accuracy of night light data was also relatively low, but slightly higher than that of building vector data. The best training results were from high-resolution image data, whose prediction accuracy was over 85%. This proves that high-resolution remote sensing image data reflect comprehensive features such as texture features and spectral features in urban areas, and the process of feature extraction using deep learning models was efficient. Furthermore, from the perspective of multisource feature fusion learning, the training accuracy of different data corresponded to the contribution rate in the final prediction model, and the data with higher training accuracy made a higher contribution in the fusion model. Conversely, data with lower feature training accuracy will perform worse in the fusion training process, such as building footprint data, and could even play a negative role in comprehensive prediction, reducing the degree of matching between feature vectors and labels. This shows that, in the random forests and ensemble learning models, it is not the case that the more features there are, the better the fit.
- (4) The random forest algorithm, which had the best training effect, was used to identify the PLES function in Ningbo. From the identification results, it can be seen that the PLES in Ningbo was mainly distributed in concentric circles from the inside to the outside. The center of the city was dominated by residential living space, while a large area of agricultural production space, and industrial and commercial production space, were distributed on the periphery of the city center. However, the public living space and ecological space were small in area, and their distribution was relatively scattered, without obvious rules.

Author Contributions: Z.B., D.J. and G.L. contributed to all aspects of this work; Z.B. conducted data analysis and J.F. wrote the main manuscript text. All authors reviewed the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by a grant Strategic Priority Research Program of the Chinese Academy of Sciences (Grant No. XDA19040305), National Natural Science Foundation of China (Grant No. 41971250), Youth Innovation Promotion Association (Grant No. 2018068), State Key Laboratory of Resources and Environmental Information System and Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences (Grant No. E0V00112YZ).

Data Availability Statement: The data presented in this study are available on request from the author.

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

References

1. Jiang, L.; Liu, Y.; Wu, S.; Yang, C. Analyzing ecological environment change and associated driving factors in China based on NDVI time series data. *Ecol. Indic.* **2021**, *129*, 107933. [[CrossRef](#)]
2. Boess, E.R.; Kørnø, L.; Lyhne, I.; Partidario, M.R. Integrating SDGs in environmental assessment: Unfolding SDG functions in emerging practices. *Environ. Impact Assess. Rev.* **2021**, *90*, 106632. [[CrossRef](#)]
3. Bennett, E.M.; Carpenter, S.R.; Caraco, N.F. Human impact on erodible phosphorus and eutrophication: A global perspective: Increasing accumulation of phosphorus in soil threatens rivers, lakes, and coastal oceans with eutrophication. *BioScience* **2001**, *51*, 227–234. [[CrossRef](#)]
4. Phalan, B.; Onial, M.; Balmford, A.; Green, R.E. Reconciling food production and biodiversity conservation: Land sharing and land sparing compared. *Science* **2011**, *333*, 1289–1291. [[CrossRef](#)] [[PubMed](#)]
5. Zou, L.; Liu, Y.; Wang, J.; Yang, Y.; Wang, Y. Land use conflict identification and sustainable development scenario simulation on China's southeast coast. *J. Clean. Prod.* **2019**, *238*, 117899. [[CrossRef](#)]
6. Zou, L.; Liu, Y.; Wang, J.; Yang, Y. An analysis of land use conflict potentials based on ecological-production-living function in the southeast coastal area of China. *Ecol. Indic.* **2021**, *122*, 107297. [[CrossRef](#)]
7. Yu, S.; Deng, W.; Xu, Y.; Zhang, X.; Xiang, H. Evaluation of the production-living-ecology space function suitability of Pingshan County in the Taihang mountainous area, China. *J. Mt. Sci.* **2020**, *17*, 2562–2576. [[CrossRef](#)]
8. Li, J.; Sun, W.; Li, M.; Meng, L. Coupling coordination degree of production 2020, living and ecological spaces and its influencing factors in the Yellow River Basin. *J. Clean. Prod.* **2021**, *298*, 126803. [[CrossRef](#)]
9. Wang, A.; Liao, X.; Tong, Z.; Du, W.; Zhang, J.; Liu, X.; Liu, M. Spatial-temporal dynamic evaluation of the ecosystem service value from the perspective of “production-living-ecological” spaces: A case study in Dongliao River Basin, China. *J. Clean. Prod.* **2022**, *333*, 130218. [[CrossRef](#)]
10. Forestier, G.; Puissant, A.; Wemmert, C.; Gançarski, P. Knowledge-based region labeling for remote sensing image interpretation. *Comput. 2012 Environ. Urban Syst.* **2012**, *36*, 470–480. [[CrossRef](#)]
11. Shahriari, M.; Bergevin, R. Land-use scene classification: A comparative study on bag of visual word framework. *Multimed. Tools Appl.* **2017**, *76*, 23059–23075. [[CrossRef](#)]
12. Blaschke, T.; Hay, G.J.; Kelly, M.; Lang, S.; Hofmann, P.; Addink, E.; Tiede, D. Geographic object-based image analysis—towards a new paradigm. *ISPRS J. Photogramm. Remote Sens.* **2014**, *87*, 180–191. [[CrossRef](#)] [[PubMed](#)]
13. Vogelmann, J.E.; Howard, S.M.; Yang, L.; Larson, C.R.; Wylie, B.K.; Van Driel, N. Completion of the 1990s National Land Cover Data Set for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources. *Photogramm. Eng. Remote Sens.* **2001**, *67*, 650–662.
14. Wulder, M.A.; White, J.C.; Cranny, M.; Hall, R.J.; Luther, J.E.; Beaudoin, A.; Dechka, J.A. Monitoring Canada's forests. Part 1: Completion of the EOSD land cover project. *Can. J. Remote Sens.* **2008**, *34*, 549–562. [[CrossRef](#)]
15. Zeferino, L.B.; de Souza, L.F.T.; do Amaral, C.H.; Fernandes Filho, E.I.; de Oliveira, T.S. Does environmental data increase the accuracy of land use and land cover classification? *Int. J. Appl. Earth Obs. Geoinf.* **2020**, *91*, 102128. [[CrossRef](#)]
16. Ienco, D.; Interdonato, R.; Gaetano, R.; Minh, D.H.T. Combining Sentinel-1 and Sentinel-2 Satellite Image Time Series for land cover mapping via a multi-source deep learning architecture. *ISPRS J. Photogramm. Remote Sens.* **2019**, *158*, 11–22. [[CrossRef](#)]
17. Chen, F.; Wang, K.; Van de Voorde, T.; Tang, T.F. Mapping urban land cover from high spatial resolution hyperspectral data: An approach based on simultaneously unmixing similar pixels with jointly sparse spectral mixture analysis. *Remote Sens. Environ.* **2017**, *196*, 324–342. [[CrossRef](#)]
18. Cheng, G.; Xie, X.; Han, J.; Guo, L.; Xia, G.S. Remote sensing image scene classification meets deep learning: Challenges, methods, benchmarks, and opportunities. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 3735–3756. [[CrossRef](#)]
19. Ma, L.; Li, M.; Ma, X.; Cheng, L.; Du, P.; Liu, Y. A review of supervised object-based land-cover image classification. *ISPRS J. Photogramm. Remote Sens.* **2017**, *130*, 277–293. [[CrossRef](#)]
20. Li, M.; Stein, A.; Bijker, W.; Zhan, Q. Urban land use extraction from Very High Resolution remote sensing imagery using a Bayesian network. *ISPRS J. Photogramm. Remote Sens.* **2016**, *122*, 192–205. [[CrossRef](#)]
21. 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](#)]
22. Liu, Y.; Liu, X.; Gao, S.; Gong, L.; Kang, C.; Zhi, Y.; Shi, L. Social sensing: A new approach to understanding our socioeconomic environments. *Ann. Assoc. Am. Geogr.* **2015**, *105*, 512–530. [[CrossRef](#)]

23. Chen, Y.; Liu, X.; Li, X.; Liu, X.; Yao, Y.; Hu, G.; Pei, F. Delineating urban functional areas with building-level social media data: A dynamic time warping (DTW) distance based k-medoids method. *Landsc. Urban Plan.* **2017**, *160*, 48–60. [[CrossRef](#)]
24. 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](#)]
25. Steiniger, S.; Lange, T.; Burghardt, D.; Weibel, R. An approach for the classification of urban building structures based on discriminant analysis techniques. *Trans. GIS* **2008**, *12*, 31–59. [[CrossRef](#)]
26. Tu, W.; Cao, J.; Yue, Y.; Shaw, S.L.; Zhou, M.; Wang, Z.; Li, Q. Coupling mobile phone and social media data: A new approach to understanding urban functions and diurnal patterns. *Int. J. Geogr. Inf. Sci.* **2017**, *31*, 2331–2358. [[CrossRef](#)]
27. Liu, C.; Wang, Y.; He, R.; Wang, C. An analysis framework for identifying and optimizing ecological-production-living space based on resident behavior. *J. Nat. Resour.* **2019**, *34*, 2113–2122. [[CrossRef](#)]
28. Wang, C.; Gao, F.; Zhou, J. The Spatio-temporal Differentiation Characteristics of “Production-Living-Ecological Space” in the Border Area of Guangzhou and Foshan with the Support of Multi-source Data: Taking the “Baiyun-Nanhai” Experimental Area as an Example. *Huazhong Archit.* **2022**, *40*, 119–125. [[CrossRef](#)]
29. Han, Z.; Shi, J.; Wu, J.; Wang, Z. Recognition Method of “The Production, Living and Ecological Space” based on POI Data and Quad-tree Idea. *J. Geo-Inf. Sci.* **2022**, *24*, 1107–1119.
30. Zhang, Y.; Luan, Q.; Xiong, C.; Liu, X. Spatial heterogeneity evaluation and zoning of production-living-ecological space based on multi-source spatial data. *Trans. Chin. Soc. Agric. Eng.* **2021**, *37*, 214–223. [[CrossRef](#)]
31. Zhao, H.; Wei, J.; Sun, D.; Liu, Y.; Wang, S.; Tan, J.; Miao, C. Recognition and spatio-temporal evolution analysis of production-living-ecological spaces based on the random forest model: A case study of Zhengzhou city, China. *J. Geomat.* **2018**, *40*, 945–957. [[CrossRef](#)]
32. Wang, Q.; Wang, C. Research on Spatial Identification and Distribution Characteristics of “Production—Living—Ecological” in Cities from the Perspective of Big Data—A Case Study of Urumqi City. *Resour. Dev. Mark.* **2022**, *38*, 142–147.
33. Ottitsch, A. Land use potential analysis Achenkirch. *Phyton* **1996**, *36*, 119–130.
34. Liu, D.; Ma, X.; Gong, J.; Li, H. Functional identification and spatiotemporal pattern analysis of production-living-ecological space in watershed scale: A case study of Bailongjiang Watershed in Gansu. *Chin. J. Ecol.* **2018**, *37*, 1490–1497.
35. Liao, L.; Dai, W.Y.; Chen, J.; Huang, W.; Jiang, F.; Hu, Q. Spatial conflict between ecological-production-living spaces on Pingtan Island during rapid urbanization. *Resour. Sci.* **2017**, *39*, 1823–1833.
36. Liang, H.; Guo, Z.; Wu, J.; Chen, Z. GDP spatialization in Ningbo City based on NPP/VIIRS night-time light and auxiliary data using random forest regression. *Adv. Space Res.* **2020**, *65*, 481–493. [[CrossRef](#)]
37. Ningbo Municipal Statistics Bureau. *Statistical Bulletin of National Economic and Social Development of Ningbo in 2020*; Ningbo Municipal Statistics Bureau: Ningbo, China, 2021.
38. Feng, Y.; Yang, Q.; Tong, X.; Wang, J.; Chen, S.; Lei, Z.; Gao, C. Long-term regional environmental risk assessment and future scenario projection at Ningbo, China coupling the impact of sea level rise. *Sustainability* **2019**, *11*, 1560. [[CrossRef](#)]
39. Feng, Y.; Yang, Q.; Tong, X.; Chen, L. Evaluating land ecological security and examining its relationships with driving factors using GIS and generalized additive model. *Sci. Total Environ.* **2018**, *633*, 1469–1479. [[CrossRef](#)]
40. Li, Y.; Wei, B.; Suo, A.; Zhang, Z.; Xu, Y.; Liang, Y. Spatial and Temporal Coupling Relationships of Coastline Exploitation and Environmental Carrying Safety in Ningbo, China. *J. Coast. Res.* **2020**, *36*, 1292–1301. [[CrossRef](#)]
41. Liu, Q.; Zhan, Q.; Li, J.; Yang, C.; Liu, W. Extracting Built-up Areas Using LuoJia-1A Nighttime Light Imagery in Wuhan, China. *Geomat. Inf. Sci. Wuhan Univ.* **2021**, *46*, 30–39. [[CrossRef](#)]
42. LeCun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **1998**, *86*, 2278–2324. [[CrossRef](#)]
43. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Commun. ACM* **2017**, *60*, 84–90. [[CrossRef](#)]
44. Nazir, U.; Khurshid, N.; Ahmed Bhimra, M.; Taj, M. Tiny-Inception-ResNet-v2: Using deep learning for eliminating bonded labors of brick kilns in South Asia. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Long Beach, CA, USA, 16–17 June 2019; pp. 39–43.
45. Hand, D.J.; Vinciotti, V. Choosing k for two-class nearest neighbour classifiers with unbalanced classes. *Pattern Recognit. Lett.* **2003**, *24*, 1555–1562. [[CrossRef](#)]
46. Moore, D.H. Classification and regression trees, by Leo Breiman, Jerome H. Friedman, Richard A. Olshen, and Charles J. Stone. brooks/cole publishing, monterey, 1984, 358 pages, \$27.95. *Cytometry* **1987**, *8*, 534–535. [[CrossRef](#)]