

Article

Block2vec: An Approach for Identifying Urban Functional Regions by Integrating Sentence Embedding Model and Points of Interest

Zhihao Sun ¹ , Hongzan Jiao ¹, Hao Wu ² , Zhenghong Peng ^{2,*} and Lingbo Liu ¹ 

¹ Department of Urban Planning, School of Urban Design, Wuhan University, Wuhan 430072, China; sun99xt@whu.edu.cn (Z.S.); jiaohongzan@whu.edu.cn (H.J.); Lingbo.liu@whu.edu.cn (L.L.)

² Department of Graphics and Digital Technology, School of Urban Design, Wuhan University, Wuhan 430072, China; wh79@whu.edu.cn

* Correspondence: pengzhenghong@whu.edu.cn; Tel.: +86-27-6877-3062

Abstract: Urban functional regions are essential information in parsing urban spatial structure. The rapid and accurate identification of urban functional regions is important for improving urban planning and management. Thanks to its low cost and fast data update characteristics, the Point of Interest (POI) is one of the most common types of open access data. It mainly identifies urban functional regions by analyzing the potential correlation between POI data and the regions. Even though this is an important manifestation of the functional region, the spatial correlation between regions is rarely considered in previous studies. In order to extract the spatial semantic information among regions, a new model, called the Block2vec, is proposed by using the idea of the Skip-gram framework. The Block2vec model maps the spatial correlation between the POIs, as well as the regions, to a high-dimensional vector, in which classification of urban functional regions can be better performed. The results from cluster analysis showed that the high-dimensional vector extracted can well distinguish the regions with different functions. The random forests classification result (Overall accuracy = 0.7186, Kappa = 0.6429) illustrated the effectiveness of the proposed method. This study also verified the potential of the sentence embedding model in the semantic information extraction of POIs.

Keywords: urban functional regions; point of interest; sentence embedding; spatial semantics; random forest



Citation: Sun, Z.; Jiao, H.; Wu, H.; Peng, Z.; Liu, L. Block2vec: An Approach for Identifying Urban Functional Regions by Integrating Sentence Embedding Model and Points of Interest. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 339. <https://doi.org/10.3390/ijgi10050339>

Academic Editor: Wolfgang Kainz

Received: 25 March 2021

Accepted: 14 May 2021

Published: 17 May 2021

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



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Cities are composed of various functions that describe human social activities and their employment of land [1,2], and can be divided into various functional regions, such as commercial, residential, industrial and open space. Urban functional regions are closely related to many urban structure studies, such as neighborhood vibrancy [3,4], travel distribution [5], urban mass transit [6] and urban energy consumption [7]. With the rapid urbanization in recent years, the urban function structure has become increasingly diverse and sophisticated. In addition, the evolution of the actual function of the region may be inconsistent with the planning intention of the land [8–10]. Thus, the fast and accurate identification of urban functional regions has become essential for improving urban planning and management [11–13].

Cadastral maps and censuses data are valuable sources of land use data as they explicitly reflect land use and contribute to land use management. However, there are extremely strict requirements for its update speed and update frequency, which is obviously not conducive to our real-time understanding of the urban land use structure. Remote sensing images [14–17] and radar/Lidar [18–20] have been used effectively for land use and land cover classification due to these can capture both spectral and textural properties of the land. However, it may be difficult for them to distinguish the categories closely

related to human social activities, because these data cannot capture functional interaction pattern, nor can they understand socioeconomic environments [21–25]. Therefore, the land cover categories for impervious surfaces usually include commercial, residential and industrial land.

To monitor and understand the potential information about human social activities, multi-source geographic data has been investigated to perceive human social activities and further infer the functions of the regions, such as mobile phone data [23,26–30], GPS trajectory data [31–33], smart card data [1,34], social media data [25,35,36] and Point of Interests (POI) data [2,9,10,33,34,36–38]. Wherein, POI data, as its inexpensive and fast to acquire from the internet, has significance advantage in representing reliable information about the location and type of urban activities (e.g., Shopping, Entertainment and Restaurant) [9,39]. Moreover, POI data can explicitly express semantic information about the urban built environment.

There is a growing body of literatures using POI data to identify urban functional regions. Considering strong relationships between functional regions with socioeconomic activities, Liu and Long [38] attempted to map functional patterns at the parcel level by generating various indicators based on frequencies of POI data. Yuan et al. [40] identified urban functional zones by using the POI data and taxi trajectory data in Beijing. However, due to the complexity of urban spatial structure, it is not enough to analyze functional patterns only using POI frequencies. Several neural language process (NLP) methods have been deployed to infer urban functional regions. In Gao et al. [2], the Latent Dirichlet Allocation (LDA) topic model was used to infer urban functional regions using POI and user check-in activities data. Chen et al. [41] compared the spatial organization of 25 cities based on the co-location patterns mining method. Word2vec algorithm was used to infer the spatial relationship of POIs in terms of urban functional region classification [10]. Place2vec algorithm, which considers spatial context based on the first law of geography [42], was used to identify the urban functional regions in Wuxi, China [8]. Conversely, spatial relationship exists not only between facility points (such as POI) but also between parcels. In other words, as the POI is related more to the other POIs that are geographically close to it according to the first law of geography [8,43], a parcel is also more related to the parcels that are geographically closer to it. However, the above methods are word embedding methods that only consider the spatial relationship between each POI, and few studies have explicitly addressed the spatially interacting relationships between parcels.

To address the gaps, a parcel-based approach, called Block2vec, was proposed to extract spatial information between parcels inspired by sentence embedding methods [44,45]. Based on the nearest neighbor method, the POI sequence and further sequence group were constructed for each parcel in Block2vec. The latent semantic feature extraction model was then built by using the skip-gram framework. Here, the Long Short Term Memory (LSTM) network [46] was deployed to build the Block2vec model, which was a one-to-many (central parcel to background parcels) and hierarchical model. Finally, the Block2vec model was tested and verified by a case study in Wuhan, China.

The remainder of the paper is structured as follows: Section 2 presents the study area, dataset. Then Section 3 introduces the method for the Block2vec model. Section 4 describes comparisons with experimental results. Section 5 discusses the advantages and limitations of the proposed method. Finally, Section 6 presents the conclusion and future work.

2. Study Area and Dataset

2.1. Study Area

The study area of this research is the main urban area of Wuhan, which is the capital of Hubei Province, China. The region consists of an area within the third ring road, Zhuankou, Wugang and Miaoshan, covering an area of 678 km². Divided by the Yangtze, the city is known as the ‘Three Towns of Wuhan’ with Hankou and Hanyang on the west bank, and Wuchang on the east. For this study, the region was divided into 2385 parcels according

to the road network data. Figure 1 shows the main urban area and POIs distribution in Wuhan.

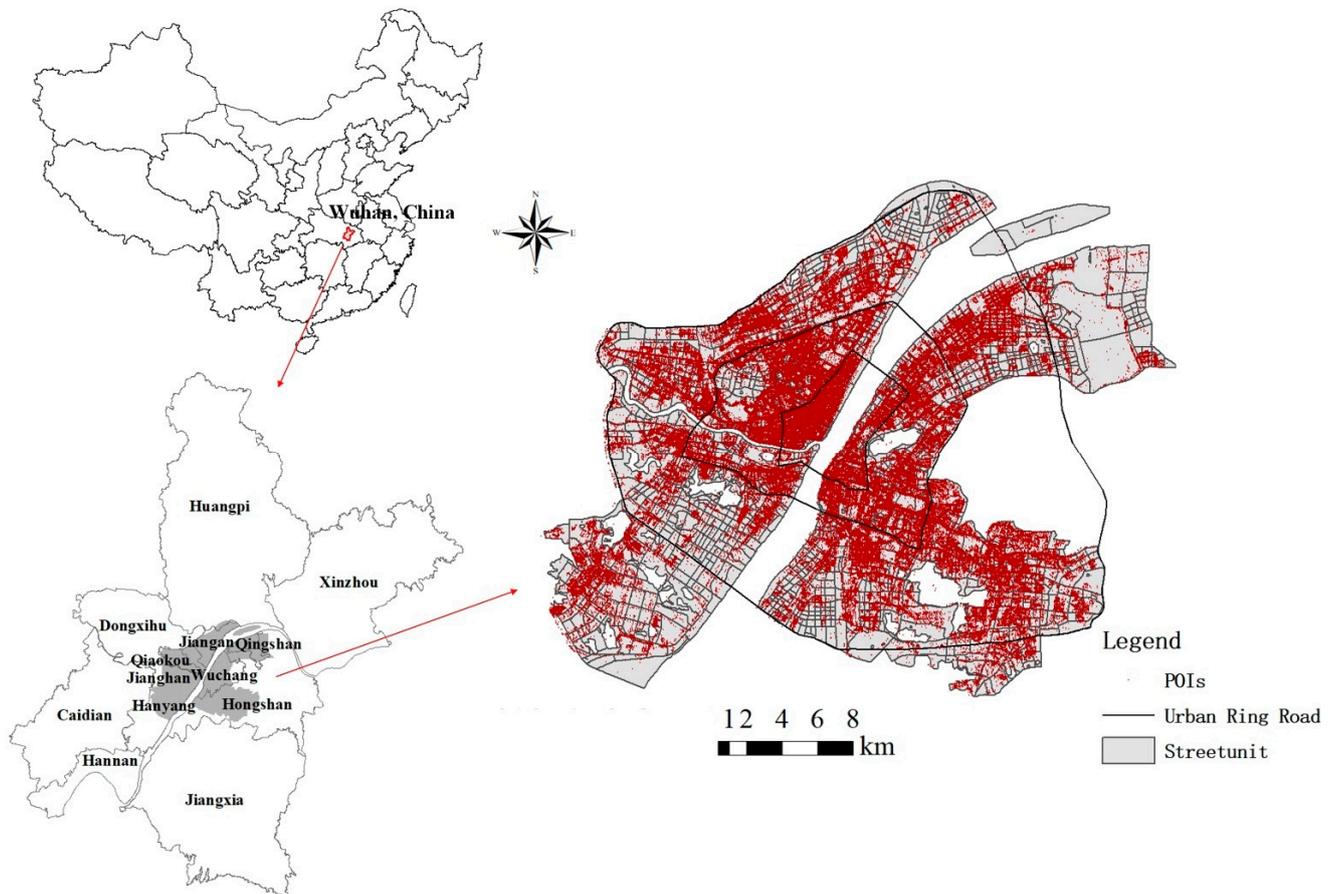


Figure 1. Location, administrative districts and POIs distribution of Wuhan. The main urban area was obtained from the Wuhan Natural Resources and Planning Bureau website (http://zrzyhgh.wuhan.gov.cn/zwgk_18/fdzdgk/ghjh/zzqgh/202001/t20200107_602858.shtml, accessed on 12 May 2017).

2.2. Dataset

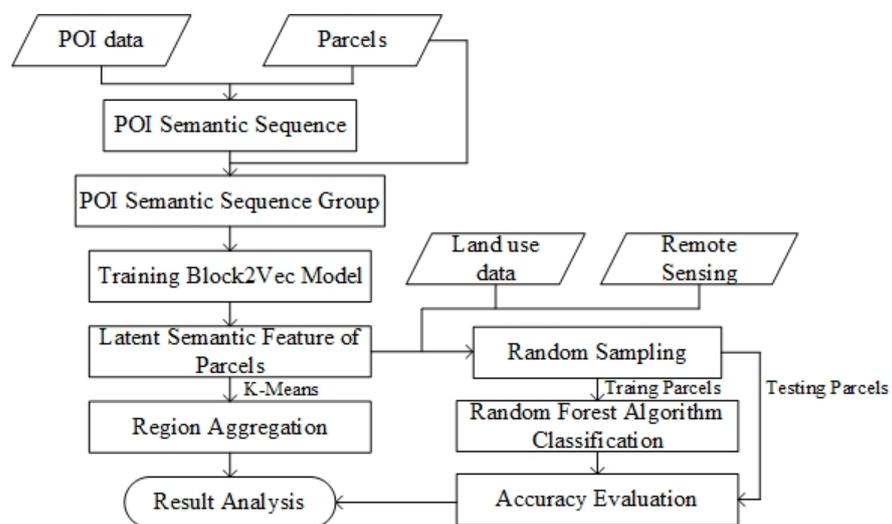
The POI data used in this study were obtained through the AutoNavi Development Platform (ADP) (<https://lbs.amap.com/api/webservice/guide/api/search>, accessed on 29 December 2016). For the study area, 537,375 POI records were collected in December 2016. Each POI record contains the geographic latitude and longitude of the POI and a classification category of multiple levels, of which there are 20 major categories of top-level and over 500 subcategories third-level. For example, as a primary school, its primary category is science and education cultural service, the secondary category is school and the tertiary category is a primary school. The detailed POI categories could be obtained through the website (<https://lbs.amap.com/api/webservice/download>, accessed on 14 February 2017). Among all the categories, the Address / Location was excluded in the later study because it could not explicitly express some human social activities. In Table 1, excluding address and location, the categories with the largest number of POIs are Shopping Mall, Catering Service and Living Service.

Table 1. The Proportions of primary categories in the study area.

Code	POI Category	Proportions	Code	POI Category	Proportions
1	Car service	1.18%	11	Tourism Attraction	0.37%
2	Car repair	0.24%	12	Residence	4.59%
3	Car sales	0.62%	13	Governmental and Public Organizations	1.77%
4	Motorcycle Service	0.04%	14	Science and Education	3.93%
5	Catering Service	16.60%	15	Transportation facilities	3.39%
6	Shopping Mall	23.79%	16	Bank/Financial	1.54%
7	Living Service	11.75%	17	Factory	6.50%
8	Sports and Recreation	2.37%	18	Road Facility	0.02%
9	Hospital	2.56%	19	Address and Location	15.78%
10	Accommodation Services	1.99%	20	Public Facility	0.96%

3. Methodology

The overall workflow of the proposed approach is shown in Figure 2. The main goal of the approach proposed is to extract the semantic information from POIs in a parcel, to better identify the function of the regions. Firstly, the POI data and Parcels were used to produce the POI semantic sequence for each parcel. Secondly, POI semantic sequence was grouped according to parcels by using the nearest neighbor method. Thirdly, the latent semantic feature extraction model was established using LSTM network. The model was trained using the POI semantic sequence groups and then mapped the semantic sequence into a high-dimensional latent semantics vector. Then, the K-Means algorithm was used to verify the discriminability and validity of the latent semantic features, and Random Forest Algorithm (RFA) were adopted to classify urban functional regions. Finally, the performance of the urban functional region classification was estimated based on its overall accuracy (OA) and Kappa score.

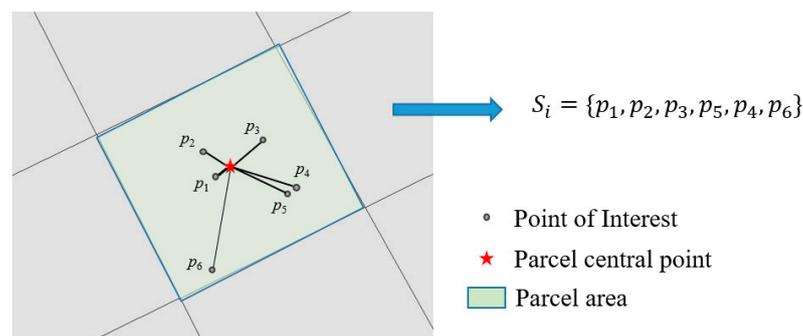
**Figure 2.** The workflow of the urban functional region classification.

3.1. Constructing Semantic Sequence for Each Parcel

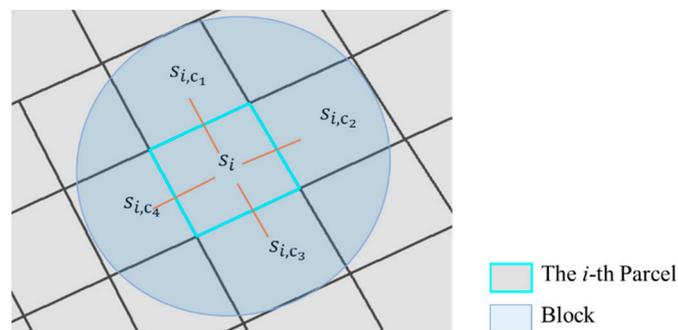
The function of a region is related to the integration of all types of activities there [2]. Generally speaking, there are multiple service facilities in one parcel, and different locations in the parcel have different spatial contact opportunities. According to the different locations, the POIs could be divided into two parts, including the part located closer to the road and the other part located in the parcel. The former serves the population in the adjacent parcels, while the latter will mainly service the population in this parcel.

In this study, the semantic sequence of POIs with specific order was constructed to express the different spatial contact opportunities in one parcel. Considering the spatial difference of the POIs above, POIs in a parcel could be sorted by order of the spatial

distance from each POI to the center of the parcel. For example, in Figure 3a, there are currently six POIs in the i -th parcel, where p_1 is the closest to the center point and p_6 is the closest to the road. Based on the distance to the center point, the semantic sequence S_i was constructed as $\{p_1, p_2, p_3, p_5, p_4, p_6\}$. In practice, parcels could have different numbers of POIs, which means that their POI sequences could have different lengths. In the next study, the LSTM layer requires a fixed number of input neurons. Therefore, the POI sequences with various lengths need to be proceeded to have a fix-length sequence. In this paper, the fixed length is set to the length that accumulates the percentage to 90%. Namely, if the POI length exceeds the fixed length in a parcel, the excess POIs would be removed. While if the length is less than the fixed length, the specific characters would be filled.



(a) Construction of POIs sequence for the i -th parcel.



(b) Typical spatial distribution and relationship between center parcel and context parcel.

Figure 3. Semantic sequence group for the center parcel S_i .

To fully mine spatial semantic relationships in POI data, it is necessary to not only consider the spatial relationship between POIs but also that between adjacent parcels. In natural language processing, a word or a sentence has two contextual relationships, forward and backward. However, in geospatial, there will be several different directional contexts. To simplify this problem, the contextual relationship of four adjacent parcels was considered as a block, which could be regarded as a contextual relationship. The typical spatial distribution of a block was shown in Figure 3b, where four nearest parcels (C_1, C_2, C_3, C_4) around the central parcel i were regarded as context parcels. Therefore, the Semantic sequence group for parcel i was defined as $[S_i, (S_{i,c_1}, S_{i,c_2}, S_{i,c_3}, S_{i,c_4})]$.

3.2. Latent Semantic Feature Extraction Model

Previous studies have shown that the *seq2seq* model can effectively extract the latent features of a sentence by using its context information [44,45,47,48]. Different from the word embedding method, the sentence embedding method represented by *seq2seq* models could perform the sentence embedding task better, because it can comprehensively capture the relevant characteristics of different words at the level of the sentence, rather than understand them at the level of words. Inspired by the above model, the POI sequence in

a parcel could be regarded as a sentence, with the k nearest parcels in its geospatial as its context parcels. As illustrated in Figure 3b, the most adjacent k value was set to 4.

In this study, the Skip-Gram model, which has been used in the skip thought vectors model [44], was applied to establish the latent semantic feature extraction model, which can be described by three parts: the encoder, decoder and objective function. As shown in Figure 4, an encoder was used to map the POI sequence of the central parcel to a latent semantic feature, and multiple decoders were used to generate POI sequences of context parcels.

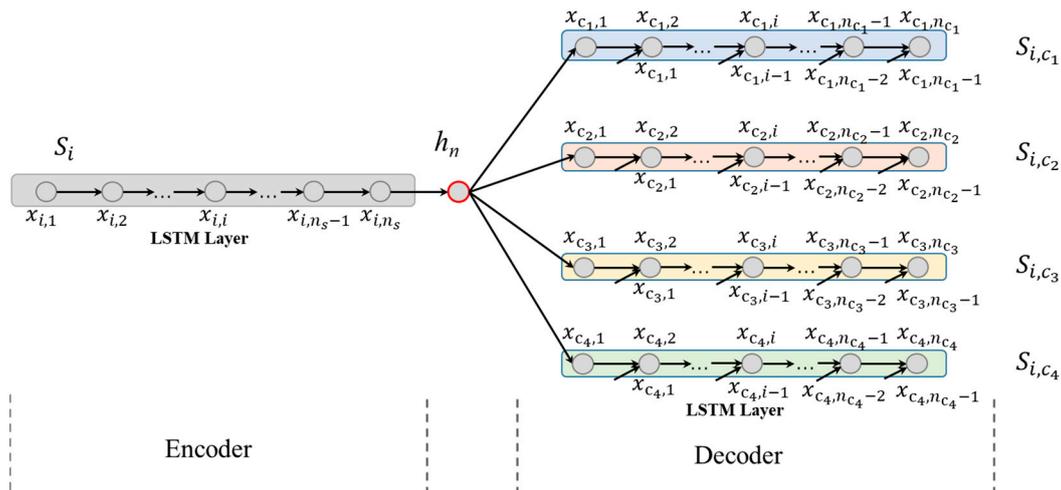


Figure 4. Latent semantic feature extraction model.

The LSTM layers were used to build the encoder of the model. For i -th parcel, let $\{x_1, x_2, \dots, x_n\}$ be the POI sequence in the S_i sequence, where n is the number of POIs in the sequence S_i . For each calculation step, the encoder calculates a hidden layer feature h_t , which can be regarded as a hidden expression for the sequence $\{x_1, x_2, \dots, x_t\}$. The hidden state h_n thus represents the entire sequence, namely, the latent semantic feature vector of S_i . To encode the sequence S_i , iterate over the following equations from the first POI in the POI sequence:

$$i_t = \sigma(W_{ii}x_t + b_{ii} + W_{hi}h_{(t-1)} + b_{hi}), \quad (1)$$

$$f_t = \sigma(W_{if}x_t + b_{if} + W_{hf}h_{(t-1)} + b_{hf}), \quad (2)$$

$$g_t = \tanh(W_{ig}x_t + b_{ig} + W_{hg}h_{(t-1)} + b_{hg}), \quad (3)$$

$$o_t = \sigma(W_{io}x_t + b_{io} + W_{ho}h_{(t-1)} + b_{ho}), \quad (4)$$

$$c_t = f_t * c_{(t-1)} + i_t * g_t, \quad (5)$$

$$h_t = o_t * \tanh(c_t), \quad (6)$$

where, i_t is the input gate, f_t is the forget gate, g_t is the update gate and o_t is the output gate, c_t is the cell state and h_t is the hidden state of the encoder at step t .

Four LSTM Layers were adopted to establish the decoder of the model, respectively. The network structure of each decoder is similar to that of the encoder. With the state h_n as a condition, four decoders then generate the POI sequences of context parcels.

Given a POI sequence group $[S_i, (S_{i,c_1}, S_{i,c_2}, S_{i,c_3}, S_{i,c_4})]$, the optimization objective function is the sum of log-probabilities for the context semantic sequences conditioned on the encoder representation:

$$O_S = \sum_{c=c_1,c_2,c_3,c_4} \sum_t \log P(x_c^t | x_c^{<t}, h_i), \quad (7)$$

where h_i denotes the hidden state of the sequence S_i , x_c^t denotes the predicted value for parcel c at step t .

In the model training process, the total objective was to minimize the sum of the above optimization objective functions over all sequence groups.

3.3. Identification of Urban Functional Regions Based on Latent Semantics

After training the above model, the encoder with the learned weights was used as a feature extractor to map the POI sequence semantics of each parcel to a latent semantic feature h_n . Theoretically, the more similar the POI semantic function between the parcels and their surrounding environment, the more they gather in the latent semantic space. Therefore, several classifiers could be trained to distinguish different regions' functions. As shown in Figure 5, K-Means and Random Forest Algorithm (RFA) were adopted to classify the parcels with different latent semantic features.

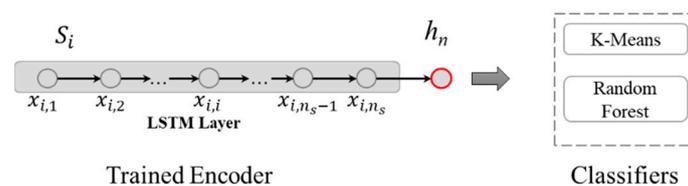


Figure 5. Classification diagram based on the trained encoder.

3.3.1. K-Means-Based Parcel Aggregation

To verify the discriminability and validity of the latent semantic features, the K-Means algorithm was used to aggregate the research parcels according to these features. The distance of similarity in vector space can be measured by various spatial distance calculation methods, such as Euclidean distance and Cosine distance. Since the feature dimension of the latent semantic space obtained in this paper is high, the cosine distance was adopted to measure the latent semantic feature vectors. Consequently, the cosine-distance-based K-Means clustering algorithm was applied to aggregate those parcels.

The silhouette score [49] was then used to evaluate how appropriate objects lie within their cluster. For the sample P_i , the average distance between P_i and other samples within the same cluster is defined as a and the average distance between P_i and samples within other clusters is defined as b , then the silhouette score is calculated as follows:

$$Score_{P_i} = \frac{b - a}{\max(a, b)} \quad (8)$$

It can be seen from the above formula that the value of the silhouette score ranges between $[-1, 1]$, and the closer to 1 the better the clustering performance is. Therefore, we calculate the average silhouette score of all samples as the evaluation of the K-Means clustering.

3.3.2. RFA-Based Parcel Classification

The unsupervised clustering analysis, however, only classifies the categories by the differences between the POI latent semantic features of different parcels. Due to the inexplicability of the extracted POI's latent semantics, it is difficult to assign and define the categories that are clustered by the cluster analysis. Therefore, the supervised classification method based on existing training samples is an essential part of our consideration.

Among them, the RFA is widely used in supervised classification because of its good adaptability to high-dimensional features and difficulty in over-fitting, and strong anti-noise ability [50,51]. Let the H_{ij} ($i \in [1, M], j \in [1, N]$) and Y_k ($k \in [1, K]$) be the latent features and land use types of parcel i , where M is the total number of parcels and N is the dimensions of the features and K is the total number of the types of regions' functions. Using the bagging method, samples with n ($n \leq N$) features were randomly selected from the N features, and then were used to build a decision tree. By the random combination

of k features, C decision trees were repeatedly built without pruning operations. Each decision tree predicted the result separately and then all the results were integrated. Even though a single decision tree may be over-fitting, this risk can be reduced by integrating the results of all decision trees. In this paper, RFA model implementation combines all results by averaging their probabilistic prediction, instead of letting each decision tree vote for a single class.

As mentioned above, the actual function of the region may not be consistent with the planners' intention. In this study, the samples were selected by using a prior information from multiple sources, including urban land use planning maps, remote sensing images and online maps. The urban land use planning maps can be obtained through the website (http://zrzyhgh.wuhan.gov.cn/zwgk_18/fdzdtk/ghjh/zzqgh/202001/t20200107_602858.shtml, accessed on 12 May 2017). The samples including the five types of functions: residential regions, commercial regions, business regions, open green spaces and industrial regions. The training samples were randomly divided into two equal-sized subsets, one used for the training and another one used for the testing. Then the model was trained using the training samples and the testing samples were used for the accuracy evaluation of the trained models. To ensure the robustness of the classification, the above-mentioned random forest classification was repeated 100 times, and then the average accuracy was used as the final evaluation result. Additionally, several state-of-the-art POIs semantic mining methods, such as term frequency-inverse document frequency (TF-IDF) [9], Latent Dirichlet Allocation (LDA) [52] and Word2vec [10] were used for comparison with our proposed method.

4. Results

In this study, 2315 research parcels contained a total of 537,375 POI data, while Tianxingzhou and a few parcels without POI were removed. Then, the three-level classification of POI types (496 types in total) was used to construct the POI sequences for parcels, which could have different lengths. Figure 6 shows the distribution of POI sequence length of parcel. It can be seen that the POI sequence length of most parcels is smaller. When the length is up to 500, the cumulative percentage reaches 91.69%. Therefore, this study sets the fixed length of the sequence to 500. Finally, sequence groups for each parcel were constructed as described in Section 3.1.

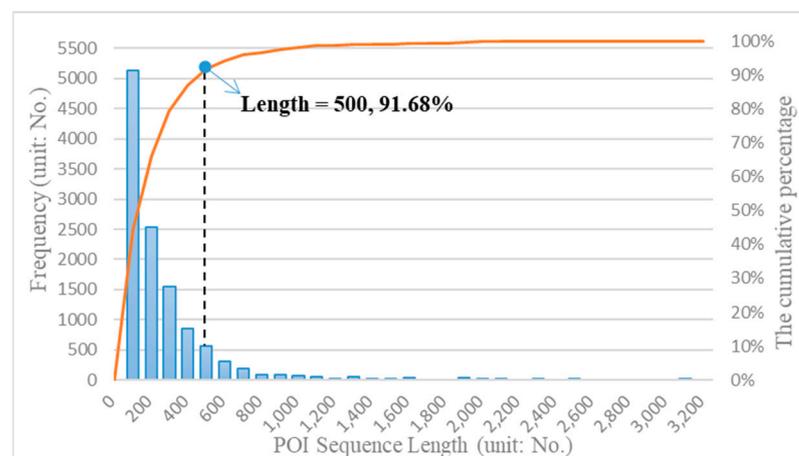


Figure 6. Distribution of POI sequence length of parcel. The POI sequence length is the number of POIs in the parcel.

Several modules, such as the scikit-learn module (an open-source machine learning tool, <https://scikit-learn.org/stable/>, accessed on 10 May 2019), the PyTorch module (an open-source machine learning and deep learning framework, <https://pytorch.org/>, accessed on 2 August 2018) and the Gensim module (an open-source topic modeling framework, <https://radimrehurek.com/gensim/>, accessed on 23 September 2019), were

adopted to construct and train the regional potential semantic extraction model described in the Section 3.2. The LSTM structure was adopted to the latent semantic feature extraction model. In this model, the number of layers of LSTMs was set to 1, the latent semantic feature dimension was set to 200, the mini-batch was set to 64 and the number of iterations was 100.

4.1. Identification of the Urban Functional Regions

4.1.1. Urban Functional Regions Aggregation by K-Means Algorithm

As illustrated above, owing to the similar latent semantics of their POI spatial sequences, parcels with the same functional semantics will be more closed in the latent semantic space than other functional parcels. The cosine-distance-based K-Means clustering algorithm was then performed to verify the discriminability and validity of the latent semantic features. As shown in Figure 7, when the clustering is two types, the silhouette score is the highest, then the silhouette score decreases gradually with the increase of the number of clusters. As a result, when $k = [2, 3, 4]$, the silhouette score could reach the top-3 values. Moreover, the local maximum is obtained when the number of clusters k is 6, 8 and 12.

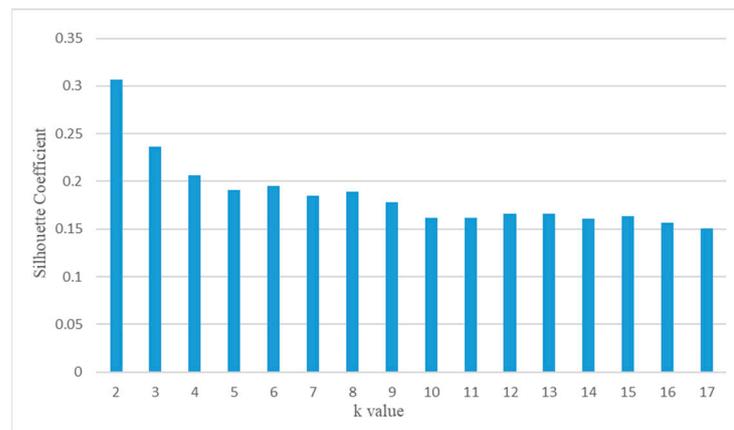


Figure 7. Silhouette score of K-Means clustering with different k values (k presents the number of clusters in cluster analysis).

Figure 8 maps the K-Means clustering results with different values of k :

When $k = 2$, by comparing with the remote sensing map and the land use map of comprehensive planning in Wuhan, we can find that the obvious circle structure can be observed in Figure 8a. Moreover, the clustering results divide the urban spatial function into a central area and an edge area, which may be related to the function and development level of the urban area in the center and suburbs.

When $k = 3$, the further division is performed compared to $k = 2$, and the circle structure still exists in Figure 8b; not only that, but the class 2 category at this time is more concentrated in the city center area, while class 3 is more concentrated in the city edge/peripheral area. Through the comparison of remote sensing maps and land use maps of comprehensive planning in Wuhan, the distribution of class 3 is consistent with the actual layout of various industrial areas in Wuhan.

When $k = 4$, the clustering map in Figure 8c is mainly to reclassify class 2 and class 3 when $k = 3$, which produces class 1, class 3 and class 4 at this time. Additionally, the class 2 in Figure 8c is basically consistent with class 1 in Figure 8b. Among them, class 1 is more concentrated, showing a partial patchy and point-like distribution. At the same time, through the comparison of remote sensing maps and urban land use maps of comprehensive planning, it is found that the distribution of class1 categories is consistent with the distribution of commercial areas in Wuhan.

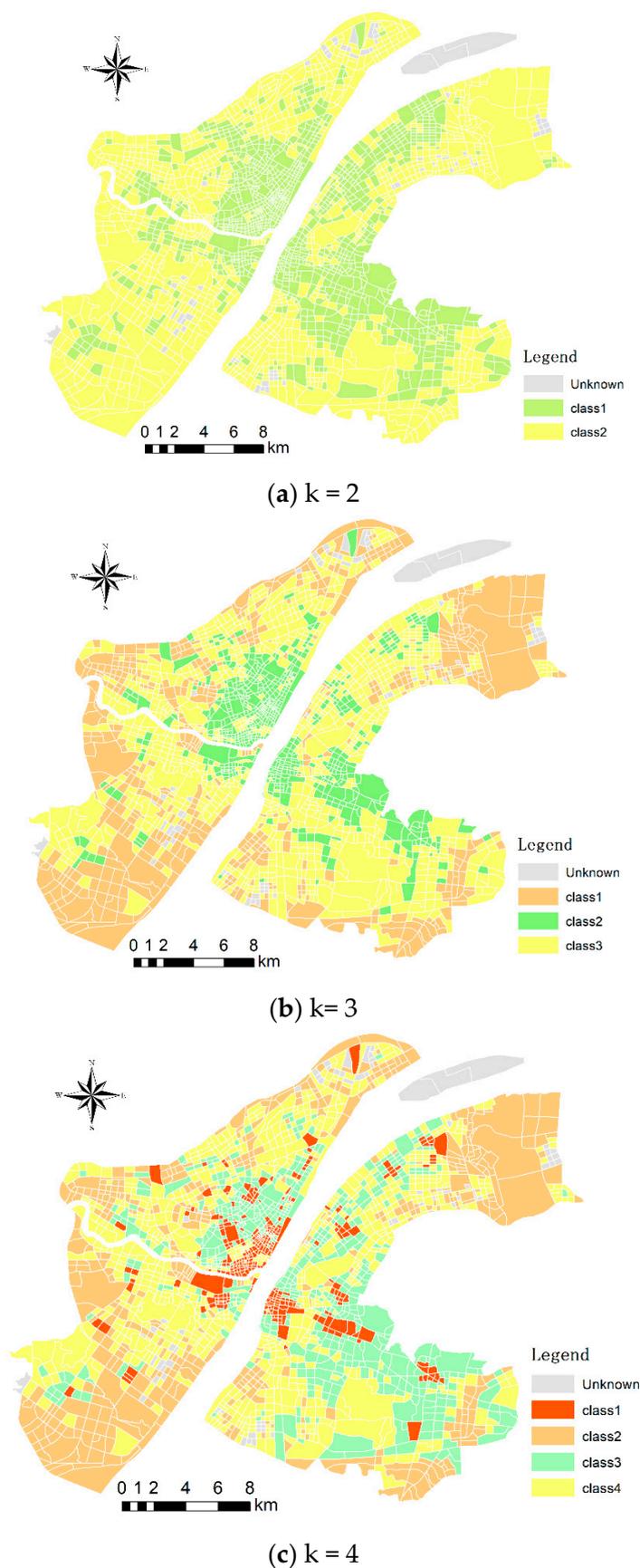


Figure 8. Results of K-Means clustering analysis of blocks using POI latent semantic features ($k = 2,3,4$).

4.1.2. Identification of Urban Functional Regions Based on Random Forest Algorithm

Using the unsupervised K-Means clustering method, it can be seen that the proposed method can effectively extract the latent semantic features of POI sequences. However, the unsupervised method cannot give an explicit definition of the classified categories, so the supervised classification method based on existing training samples is the necessary means that need to be adopted.

Based on the latent semantic feature vectors extracted from the above model, this paper uses the random forest algorithm to classify urban functional regions. At the same time, 96 samples have been randomly sampled. Additionally, some state-of-the-art methods, including TF-IDF [9], LDA [2] and Word2vec [43], are used to compare with our methods.

RFA model provided by the scikit-learn module library (<https://scikit-learn.org/stable/>, accessed on 10 May 2019) was adopted to classify urban functional regions, where the number C of decision trees is set to 200. The implementation of the Word2vec, LDA and TF-IDF models for comparison experiments was performed using the module provided by the gensim module library (<https://radimrehurek.com/gensim/index.html>, accessed on 23 September 2019), where the model parameter settings for each method are maintained consistent with previous literature.

To ensure the stability of results, each method was repeated 100 times. Table 2 provides an accurate assessment of urban functional region classification using different methods, and Figure 9 shows the results of urban functional region classification mapping using different methods.

Table 2. Accurate assessment of urban functional region classification via different methods.

Methods	Overall Accuracy	Kappa Score
Word2vec	0.6657 ± 0.0137	0.5769 ± 0.0173
TF-IDF	0.6486 ± 0.0254	0.5523 ± 0.0330
LDA	0.5972 ± 0.0196	0.5014 ± 0.0249
Block2vec	0.7186 ± 0.0186	0.6429 ± 0.0237

TF-IDF: term frequency-inverse document frequency. LDA: Latent Dirichlet Allocation.

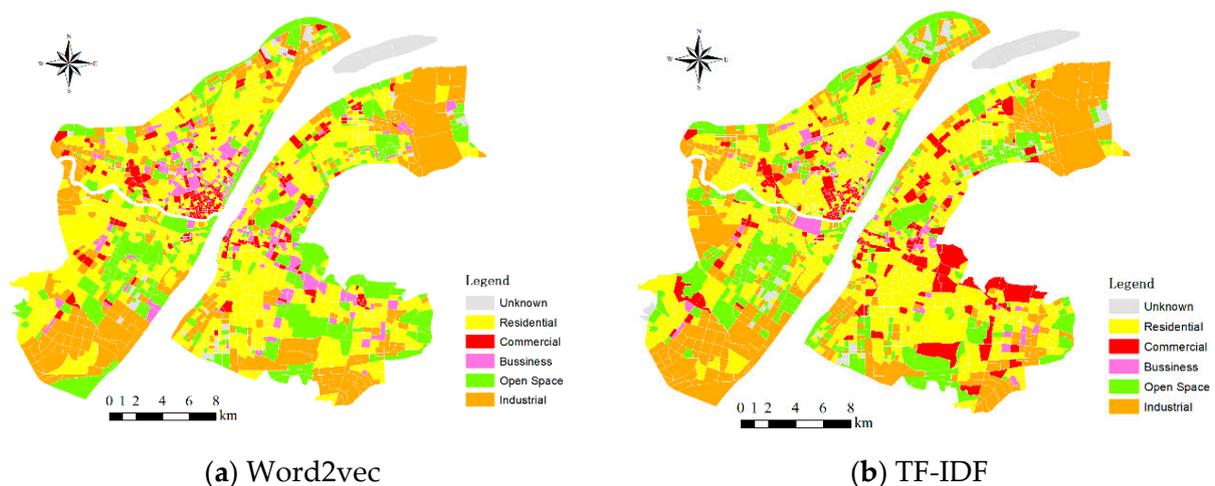


Figure 9. Cont.

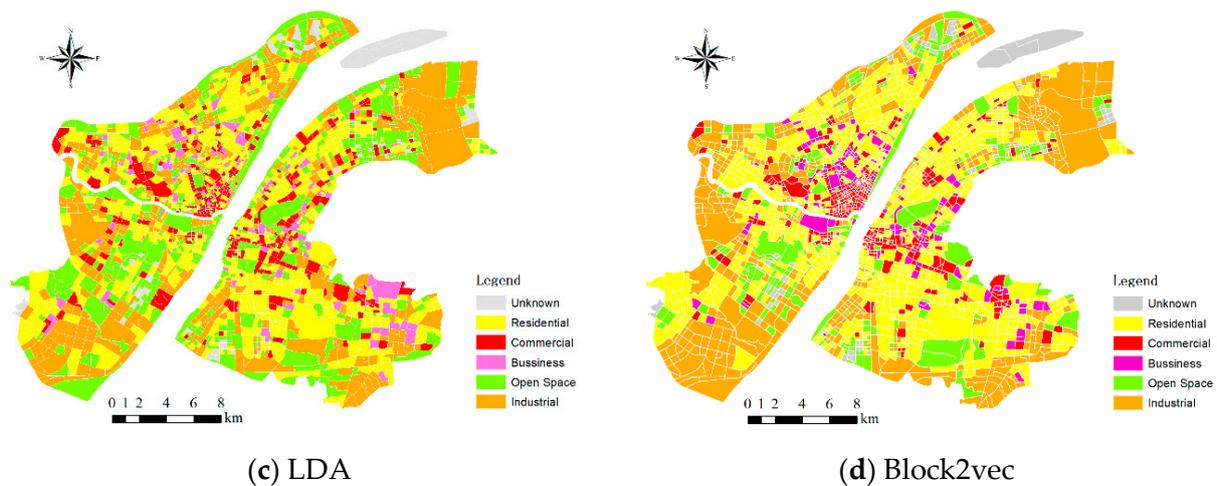


Figure 9. Urban functional region classification results via different methods.

Inconsistent with the previous studies, although the TF-IDF method only considers the quantitative features of POIs in the region, it still achieves relatively good classification accuracy compared to the LDA model. The Word2vec model, which considers both the quantitative features and spatial distribution features of POIs, has a higher classification accuracy than the TF-IDF and LDA because it considers both the frequency characteristics of the POI and the spatial relationship between the POIs. Compared with the above results, the proposed Block2vec achieved the highest classification accuracy and Kappa score.

Figure 10 shows the confusion matrixes of urban functional region classification via different methods. Compared with other methods, the proposed method (Figure 10d) has the highest accuracy in the classification of Residential, Commercial and Industrial regions, and the top-2 accuracy in the classification of Business. The Word2Vec method (Figure 10a) has the highest accuracy in the classification of Business and Open Space, while it is lower in the classification of Residential and Industrial. The results show that, considers the spatial relationship of the parcels, the feature extraction model that can effectively improve the classification accuracy of Residential, Commercial and Industrial, but cannot improve the classification accuracy of the Open Space.

To further verify the classification results based on the proposed model, three local regions were compared with Google map and land use map of comprehensive planning in Wuhan. Figure 11a is the central area of the city, and its actual function type is mainly based on the business and commercial regions. The results show that the distribution of the proposed model classification is consistent with the distribution of planning maps. Figure 11b is another central area of the city, and its commercial scale is smaller than that in Figure 11a. The business in this area's planning map is allocated from north to south, while the classification results of the proposed model are allocated from east to west. Figure 11c is a business and industrial area in the southeastern part of the city. It can be seen that the classification results of the proposed model are completely inconsistent with the planning map. Through the comparison with online maps, the distribution of the proposed model classification is more realistic. Even though most of the regions in Figure 11c is planned as industrial land, with the arrival of a large number of software technology companies, the functions of this regions have gradually transformed into business and commercial functions in this area. This confirms the previous research that the evolution of the actual function of the region may not be consistent with the planers' intention.

Word2vec	Residential	Commercial	Business	Open Space	Industrial
Residential	0.600	0.200	0.100	0.000	0.100
Commercial	0.100	0.800	0.000	0.000	0.100
Business	0.000	0.143	0.714	0.000	0.143
Open Space	0.000	0.000	0.125	0.750	0.125
Industrial	0.000	0.000	0.091	0.182	0.727

(a) Word2vec

TF-IDF	Residential	Commercial	Business	Open Space	Industrial
Residential	0.800	0.100	0.000	0.000	0.100
Commercial	0.100	0.700	0.100	0.100	0.000
Business	0.000	0.143	0.286	0.143	0.429
Open Space	0.125	0.000	0.000	0.625	0.250
Industrial	0.000	0.000	0.000	0.091	0.909

(b) TF-IDF

LDA	Residential	Commercial	Business	Open Space	Industrial
Residential	0.700	0.200	0.000	0.100	0.000
Commercial	0.200	0.500	0.100	0.100	0.100
Business	0.000	0.143	0.571	0.000	0.286
Open Space	0.250	0.000	0.000	0.625	0.125
Industrial	0.000	0.000	0.091	0.091	0.818

(c) LDA

Block2vec	Residential	Commercial	Business	Open Space	Industrial
Residential	0.900	0.000	0.000	0.100	0.000
Commercial	0.200	0.800	0.000	0.000	0.000
Business	0.000	0.143	0.571	0.000	0.286
Open Space	0.250	0.000	0.250	0.500	0.000
Industrial	0.000	0.000	0.000	0.000	1.000

(d) Proposed method

Figure 10. Confusion matrixes of classification results via (a) Word2Vec, (b) TF-IDF, (c) LDA and (d) proposed method.

4.2. The Influence of the Size of Latent Semantic Features

The latent semantic feature model used in this paper maps the POI sequence in a block to a latent high-dimensional semantic feature space, so the dimension of the latent semantic feature directly determines the semantic richness of the latent feature. If the dimension of the latent semantic feature is too low, it is difficult to obtain rich POI sequence semantics and lead to loss of information; however, too high a dimension may lead to information redundancy. Therefore, in this section, we try to analyze the ability of different size latent semantic features to identify and distinguish urban functional regions.

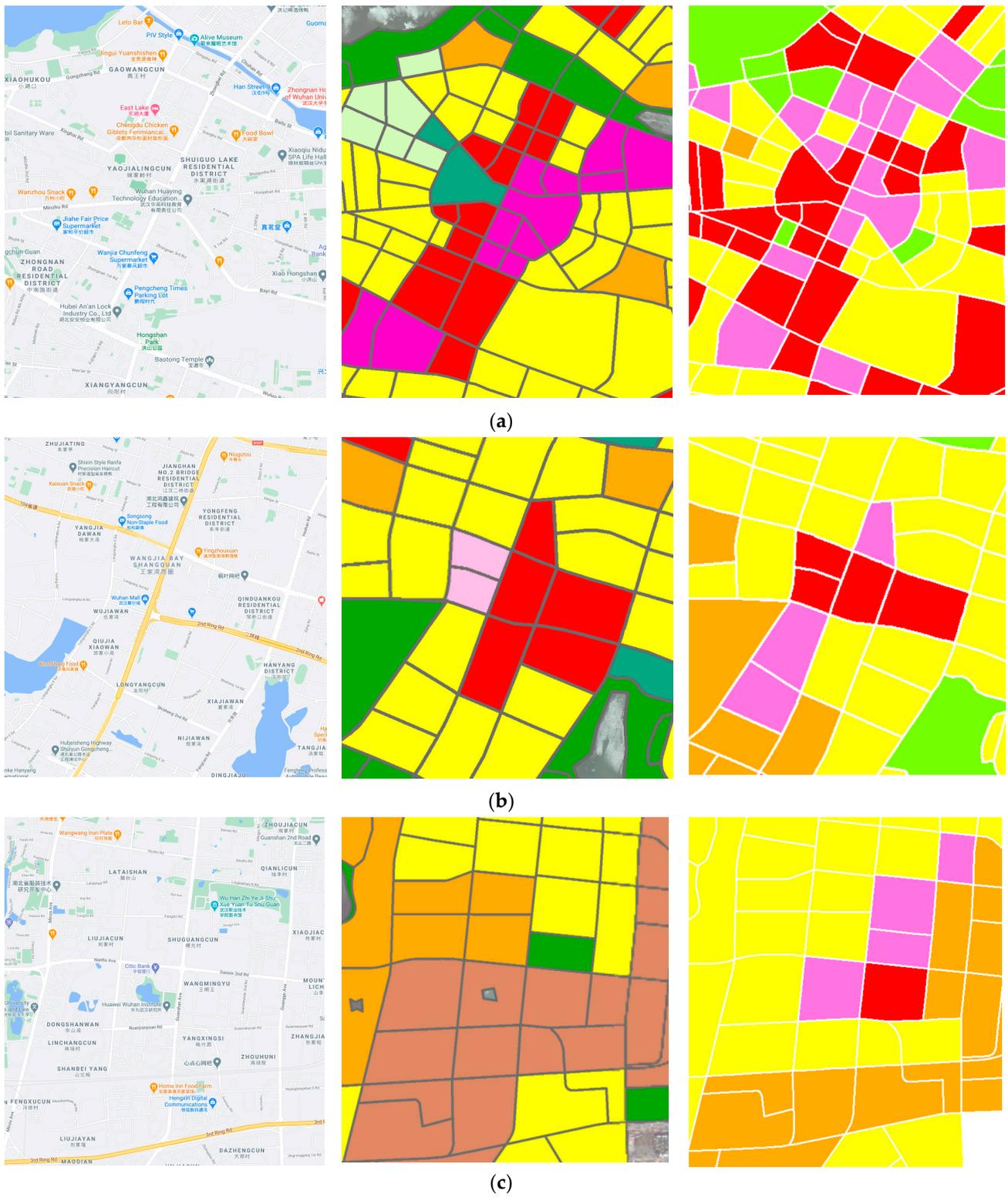


Figure 11. Comparison of local regions classification results. (a) The central area in Wuchang; (b) the central area in Hanyang; (c) the area southeast of Wuchang (The Google online maps are on the left; the planning maps are in the middle and our classification results are on the right).

Figure 12 shows the classification accuracy evaluation changes of urban functional regions when the model is set to different dimensions. It can be seen that when the hidden layer size is between 10–100, with the increase of hidden layer size, the latent semantics acquired are more and more abundant and the classification accuracy is higher. When the hidden layer size is between 100 and 250, the change of classification accuracy is not obvious as the size is improved. When the hidden layer dimension continues to increase to 300, the classification accuracy decreases due to the too high dimension of the latent features. Therefore, it is appropriate to set the latent semantic feature size of the hidden layer to 200 in this paper.

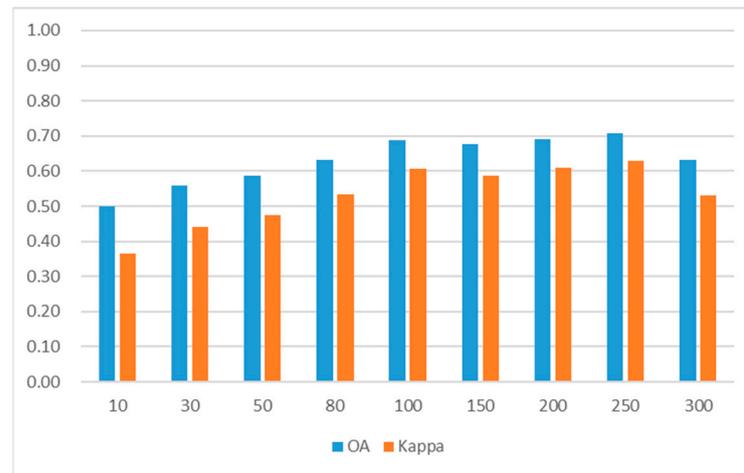


Figure 12. Changes in the accuracy of urban functional region classification under different latent semantic features.

5. Discussion

A timely and accurate urban functional regions map is conducive to urban management and urban planning. This study proposed an effective approach for the identification of urban functional regions by extracting latent semantic features of POIs in parcels. The proposed approach considers the following spatial relationships of POIs: (1) The spatial relationship of POIs in a parcel: There is an interdependent and competitive relationship between adjacent POIs in geographic space [53]. (2) The spatial accessibility varies in different areas of a parcel: Generally speaking, there are more public facilities in the areas near the streets due to the high accessibility for external contact, while the areas near the interior of the parcel have some unique facilities. (3) The relationship between parcels: Parcels with different functions are often close to or distant from each other due to their interdependent or competing relationships.

This study agrees with previous studies [2,9,10,42] that natural language processing (NLP) has a good advantage in extracting the semantic features of POIs. However, few studies have explicitly addressed the spatial correlations among parcels. In this study, considering relationships existing between POIs, the POI semantic sequence was built with specific order. Then, sequence group was constructed by considering the relationships existing between parcels (center parcel and context parcel). The LSTM network was used to extract the former, while the Encoder-Decoder structure was used to extract the latter. Consequently, the results achieved the highest accuracy (OA = 0.7186, Kappa = 0.6429), which indicates that our model can effectively extract the latent features for more accurate classification of the urban functional regions. Moreover, the result of the confusion matrix indicates that the proposed method could effectively improve the classification accuracy of the Residential, Commercial and Industrial regions. This reveals that those types of parcels have close spatial correlations, while they are less spatially connected to Open Space land. Furthermore, the comparison of local regions classification results verified

that the evolution of the actual function of the region may not be consistent with the planners' intention.

Classification accuracy for the four methods using POI data ranged from 0.5972 to 0.7186, which were close to the accuracy of relevant studies [8,30,35]. However, these were lower than that of remote sensing land use classification. The main reason is that the remote sensing images mainly classify the land based on the physical features of the land, and it has accumulated a large number of labeled datasets. Different from this, the classification based on the POIs was trained using the samples chosen by ourselves, with a small sample size and subject to subjective influence. At the same time, some regions' functions are affected by multiple human social activities.

It should be noted that this study has been examined only in the urban area, where there are abundant service facilities. Thus, it may be difficult to transfer this approach to areas with fewer POIs, such as suburbs and rural areas. Additionally, mixed-function type has not been mentioned as it is hard to artificially define this type [41,43]. Nevertheless, this paper innovatively proposed a parcel-based semantic extraction method, which outperformed other state-of-the-art methods reported in the paper in terms of its ability to extract POI semantics.

In addition, the research in this paper classified the parcels into five types of functional regions, which may conflict with standard urban land use classification. Some standard land use types are a mixture of various human social activities, which may not correspond exactly according to the land use types. There is also no unified definition standard in relevant studies [8,32,36]. However, this paper does not attempt to use the proposed method to define types to replace standard urban land use types. This research aims to provide a better data-driven method to quickly and accurately identify regional functions from POI data. Planners and government management thus can use this method to continuously and effectively observe and monitor changes in regional functions in the city.

Moreover, some functional regions can be subdivided. For instance, the residence could be divided into low-density and high-density residential land, which could be identified with remote sensing data [14]. However, it is difficult to distinguish them by the POIs alone. Incorporating other data, such as high-resolution remote sensing images and social media data, can effectively improve the ability to distinguish among more different types of urban functional regions.

6. Conclusions and Future Work

With rapid urbanization, the urban spatial structure of urban functional regions has become increasingly diverse and sophisticated. Therefore, it is necessary to produce a timely and accurate urban functional region map for urban management and urban planning. This study proposed an effective approach, called Block2vec, for the identification of urban functional regions by extracting latent semantic features of POI in parcels. First, a POI sequence and further sequence group were constructed for each parcel. Then, the POI sequence was mapped to a high-dimensional space by building a Block2vec model. Furthermore, the K-Means clustering and RFA classification were adopted to reveal the urban structures and to identify the functional types. Compared with other state-of-the-art methods (TF-IDF, LDA and Word2vec), the Block2vec method could obtain the highest accuracy (OA = 0.7186, Kappa = 0.6429). Furthermore, the proposed method has a significant improvement in the classification accuracy of residential, commercial and industrial land. The proposed method can help urban management and urban planners to understand the distribution of urban functional regions in a timely and accurate manner. At the same time, this study also verified the potential of the neural language process model in the semantic information extraction of POIs.

For future work, accumulating more study areas will help us to obtain more training samples of functional regions. Last but not least, incorporating other data, such as high-resolution remote sensing images and social media data, can effectively improve the ability to distinguish among more types of urban functional regions.

Author Contributions: Funding acquisition, Zhenghong Peng and Hao Wu; Methodology, Hongzan Jiao; Project administration, Zhenghong Peng; Validation, Hao Wu; Writing—original draft, Zhihao Sun; Writing—review & editing, Lingbo Liu. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, grant number 51978535 and 52078390.

Data Availability Statement: Not applicable.

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

References

- Zhong, C.; Huang, X.; Müller Arisona, S.; Schmitt, G.; Batty, M. Inferring building functions from a probabilistic model using public transportation data. *Comput. Environ. Urban Syst.* **2014**, *48*, 124–137. [CrossRef]
- Gao, S.; Janowicz, K.; Couclelis, H. Extracting urban functional regions from points of interest and human activities on location-based social networks. *Trans. GIS* **2017**, *21*, 446–467. [CrossRef]
- Jin, X.; Long, Y.; Sun, W.; Lu, Y.; Yang, X.; Tang, J. Evaluating cities' vitality and identifying ghost cities in China with emerging geographical data. *Cities* **2017**, *63*, 98–109. [CrossRef]
- Yue, Y.; Zhuang, Y.; Yeh, A.G.O.; Xie, J.Y.; Ma, C.L.; Li, Q.Q. Measurements of POI-based mixed use and their relationships with neighbourhood vibrancy. *Int. J. Geogr. Inf. Sci.* **2017**, *31*, 658–675. [CrossRef]
- Forghani, M.; Karimipour, F. Interplay between urban communities and human-crowd mobility: A study using contributed geospatial data sources. *Trans. GIS* **2018**, *22*, 1008–1028. [CrossRef]
- Yue, M.; Kang, C.; Andris, C.; Qin, K.; Liu, Y.; Meng, Q. Understanding the interplay between bus, metro, and cab ridership dynamics in Shenzhen, China. *Trans. GIS* **2018**, *22*, 855–871. [CrossRef]
- Zhang, M.; Zhao, P. The impact of land-use mix on residents' travel energy consumption: New evidence from Beijing. *Transp. Res. Part D Transp. Environ.* **2017**, *57*, 224–236. [CrossRef]
- Zhai, W.; Bai, X.; Shi, Y.; Han, Y.; Peng, Z.-R.; Gu, C. Beyond Word2vec: An approach for urban functional region extraction and identification by combining Place2vec and POIs. *Comput. Environ. Urban Syst.* **2019**, *74*, 1–12. [CrossRef]
- Yuan, J.; Zheng, Y.; Xie, X. Discovering regions of different functions in a city using human mobility and POIs. In Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining—KDD '12, Beijing, China, 12–16 August 2012.
- Yao, Y.; Li, X.; Liu, X.; Liu, P.; Liang, Z.; Zhang, J.; Mai, K. Sensing spatial distribution of urban land use by integrating points-of-interest and Google Word2Vec model. *Int. J. Geogr. Inf. Sci.* **2017**, *31*, 825–848. [CrossRef]
- Maat, K.; van Wee, B.; Stead, D. Land use and travel behaviour: Expected effects from the perspective of utility theory and activity-based theories. *Environ. Plan. B Plan. Des.* **2005**, *32*, 33–46. [CrossRef]
- Ellis, E.; Pontius, R. Land-Use and Land-Cover Change. The Encyclopedia of Earth. 2007. Available online: https://ecotope.org/people/ellis/papers/ellis_eoe_lulcc_2007.pdf (accessed on 30 September 2016).
- La Rosa, D.; Privitera, R. Characterization of non-urbanized areas for land-use planning of agricultural and green infrastructure in urban contexts. *Landsc. Urban Plan.* **2013**, *109*, 94–106. [CrossRef]
- Han, X.; Zhong, Y.; Zhao, B.; Zhang, L. Scene classification based on a hierarchical convolutional sparse auto-encoder for high spatial resolution imagery. *Int. J. Remote Sens.* **2017**, *38*, 514–536. [CrossRef]
- Zhong, Y.; Fei, F.; Liu, Y.; Zhao, B.; Jiao, H.; Zhang, L. SatCNN: Satellite image dataset classification using agile convolutional neural networks. *Remote Sens. Lett.* **2017**, *8*, 136–145. [CrossRef]
- Tao, C.; Pan, H.; Li, Y.; Zou, Z. Unsupervised spectral-spatial feature learning with stacked sparse autoencoder for hyperspectral imagery classification. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 2438–2442. [CrossRef]
- Li, J.; Huang, X.; Zhang, L. Semi-supervised sparse relearning representation classification for high-resolution remote sensing imagery. In Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS), Beijing, China, 10–15 July 2016.
- Mohammadimanesh, F.; Salehi, B.; Mahdianpari, M.; Gill, E.; Molinier, M. A new fully convolutional neural network for semantic segmentation of polarimetric SAR imagery in complex land cover ecosystem. *ISPRS J. Photogramm. Remote Sens.* **2019**, *151*, 223–236. [CrossRef]
- Li, Y.; Chen, Y.; Liu, G.; Jiao, L. A novel deep fully convolutional network for PolSAR image classification. *Remote Sens.* **2018**, *10*, 1984. [CrossRef]
- Tao, C.; Chen, S.; Li, Y.; Xiao, S. PolSAR land cover classification based on roll-invariant and selected hidden polarimetric features in the rotation domain. *Remote Sens.* **2017**, *9*, 660. [CrossRef]
- Tu, W.; Cao, J.; Yue, Y.; Shaw, S.L.; Zhou, M.; Wang, Z.; Chang, X.; Xu, Y.; 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**, *30*, 2331–2358. [CrossRef]
- Liu, Y.; Liu, X.; Gao, S.; Gong, L.; Kang, C.; Zhi, Y.; Chi, G.; Shi, L. Social sensing: A new approach to understanding our socioeconomic environments. *Ann. Assoc. Am. Geogr.* **2015**, *105*, 512–530. [CrossRef]

23. Pei, T.; Sobolevsky, S.; Ratti, C.; Shaw, S.L.; Li, T.; Zhou, C. A new insight into land use classification based on aggregated mobile phone data. *Int. J. Geogr. Inf. Sci.* **2014**, *28*, 1988–2007. [[CrossRef](#)]
24. Jia, Y.; Ge, Y.; Ling, F.; Guo, X.; Wang, J.; Wang, L.; Chen, Y.; Li, X. Urban land use mapping by combining remote sensing imagery and mobile phone positioning data. *Remote Sens.* **2018**, *10*, 446. [[CrossRef](#)]
25. 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](#)]
26. Toole, J.L.; Ulm, M.; González, M.C.; Bauer, D. Inferring land use from mobile phone activity. In Proceedings of the ACM SIGKDD International Workshop on Urban Computing, Beijing, China, 12 August 2012; pp. 1–8.
27. Ríos, S.A.; Muñoz, R. Land Use detection with cell phone data using topic models: Case Santiago, Chile. *Comput. Environ. Urban Syst.* **2017**, *61*, 39–48. [[CrossRef](#)]
28. Tu, W.; Hu, Z.; Li, L.; Cao, J.; Jiang, J.; Li, Q.; Li, Q. Portraying urban functional zones by coupling remote sensing imagery and human sensing data. *Remote Sens.* **2018**, *10*, 141. [[CrossRef](#)]
29. Mao, H.; Ahn, Y.Y.; Bhaduri, B.; Thakur, G. Improving land use inference by factorizing mobile phone call activity matrix. *J. Land Use Sci.* **2017**, *12*, 138–153. [[CrossRef](#)]
30. Caceres, N.; Benitez, F.G. Supervised land use inference from mobility patterns. *J. Adv. Transp.* **2018**, *2018*, 8710402. [[CrossRef](#)]
31. Pan, G.; Qi, G.; Wu, Z.; Zhang, D.; Li, S. Land-use classification using taxi GPS traces. *IEEE Trans. Intell. Transp. Syst.* **2013**, *14*, 113–123. [[CrossRef](#)]
32. Liu, X.; Kang, C.; Gong, L.; Liu, Y. Incorporating spatial interaction patterns in classifying and understanding urban land use. *Int. J. Geogr. Inf. Sci.* **2016**, *30*, 334–350. [[CrossRef](#)]
33. Wang, Y.; Gu, Y.; Dou, M.; Qiao, M. Using spatial semantics and interactions to identify urban functional regions. *ISPRS Int. J. Geo Inf.* **2018**, *7*, 130. [[CrossRef](#)]
34. Long, Y.; Shen, Z. Discovering functional zones using bus smart card data and points of interest in Beijing. In *Geospatial Analysis to Support Urban Planning in Beijing*; Springer: Cham, Switzerland, 2015; Volume 116, pp. 193–217. [[CrossRef](#)]
35. Frias-Martinez, V.; Frias-Martinez, E. Spectral clustering for sensing urban land use using Twitter activity. *Eng. Appl. Artif. Intell.* **2014**, *35*, 237–245. [[CrossRef](#)]
36. Zhang, Y.; Li, Q.; Tu, W.; Mai, K.; Yao, Y.; Chen, Y. Functional urban land use recognition integrating multi-source geospatial data and cross-correlations. *Comput. Environ. Urban Syst.* **2019**, *78*, 101374. [[CrossRef](#)]
37. Jiang, S.; Alves, A.; Rodrigues, F.; Ferreira, J.; Pereira, F.C. Mining point-of-interest data from social networks for urban land use classification and disaggregation. *Comput. Environ. Urban Syst.* **2015**, *53*, 36–46. [[CrossRef](#)]
38. Liu, X.; Long, Y. Automated identification and characterization of parcels with OpenStreetMap and points of interest. *Environ. Plan. B Plan. Des.* **2016**, *43*, 341–360. [[CrossRef](#)]
39. Yu, Z.; Capra, L.; Wolfson, O.; Yang, H. Urban computing: Concepts, methodologies, and applications. *ACM Trans. Intell. Syst. Technol.* **2014**, *5*, 1–55. [[CrossRef](#)]
40. Yuan, N.J.; Zheng, Y.; Xie, X.; Wang, Y.; Zheng, K.; Xiong, H. Discovering urban functional zones using latent activity trajectories. *IEEE Trans. Knowl. Data Eng.* **2015**, *27*, 712–725. [[CrossRef](#)]
41. Chen, Y.; Liu, X.; Li, X.; Liu, X.; Yao, Y.; Hu, G.; Xu, X.; 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](#)]
42. Yan, B.; Mai, G.; Janowicz, K.; Gao, S. From ITDL to Place2Vec—Reasoning about place type similarity and relatedness by learning embeddings from augmented spatial contexts. In Proceedings of the GIS ACM International Symposium on Advances in Geographic Information Systems, Redondo Beach, CA, USA, 7–10 November 2017.
43. Yao, Y.; Liang, H.; Li, X.; Zhang, J.; He, J. Sensing urban land-use patterns by integrating Google Tensorflow and scene-classification models. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **2017**, *XLII-2/W7*, 981–988. [[CrossRef](#)]
44. Kiros, R.; Zhu, Y.; Salakhutdinov, R.; Zemel, R.S.; Torralba, A.; Urtasun, R.; Fidler, S. Skip-thought vectors. In Proceedings of the NIPS’15: 28th International Conference on Neural Information Processing Systems, Montreal, QC, Canada, 8–13 December 2014; MIT Press: Cambridge, MA, USA, 2015.
45. Sutskever, I.; Vinyals, O.; Le, Q.V. Sequence to sequence learning with neural networks. In Proceedings of the NIPS’15: 28th International Conference on Neural Information Processing Systems, Montreal, QC, Canada, 8–13 December 2014; MIT Press: Cambridge, MA, USA, 2015.
46. Sundermeyer, M.; Schlüter, R.; Ney, H. LSTM neural networks for language modeling. In Proceedings of the 13th Annual Conference of the International Speech Communication Association, INTERSPEECH 2012, Portland, OR, USA, 9–13 September 2012.
47. Cho, K.; van Merriënboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; Bengio, Y. Learning phrase representations using RNN encoder-decoder for statistical machine translation. In Proceedings of the EMNLP 2014 Conference on Empirical Methods in Natural Language Processing, Doha, Qatar, 25–29 October 2014.
48. Wang, S.; Cao, J.; Chen, H.; Peng, H.; Huang, Z. SeqST-GAN: Seq2Seq generative adversarial nets for multi-step urban crowd flow prediction. *ACM Trans. Spat. Algorithms Syst.* **2020**, *6*, 22. [[CrossRef](#)]
49. Rousseeuw, P.J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **1987**, *20*, 53–65. [[CrossRef](#)]
50. Breiman, L. Random Forreest. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
51. Biau, G. Analysis of a random forests model. *J. Mach. Learn. Res.* **2012**, *13*, 1063–1095.

-
52. Liénou, M.; Maître, H.; Datcu, M. Semantic annotation of satellite images using latent dirichlet allocation. *IEEE Geosci. Remote Sens. Lett.* **2010**, *7*, 28–32. [[CrossRef](#)]
 53. Liu, X.; Andris, C.; Rahimi, S. Place niche and its regional variability: Measuring spatial context patterns for points of interest with representation learning. *Comput. Environ. Urban Syst.* **2019**, *75*, 146–160. [[CrossRef](#)]