

Article

A GloVe-Based POI Type Embedding Model for Extracting and Identifying Urban Functional Regions

Chengkun Zhang ¹, Liuchang Xu ², Zhen Yan ^{3,*} and Sensen Wu ¹

¹ School of Earth Sciences, Zhejiang University, 38 Zheda Road, Hangzhou 310027, China; zhangchengkun@zju.edu.cn (C.Z.); wusensengis@zju.edu.cn (S.W.)

² College of Information Engineering, Zhejiang A&F University, Hangzhou 311300, China; xuliuchang@zafu.edu.cn

³ Center for Agricultural and Rural Development (CARD), School of Public Affairs, Zhejiang University, 866 Yuhangtang Rd., Hangzhou 310058, China

* Correspondence: yanzhen@zju.edu.cn; Tel.: +86-571-5633-7015

Abstract: Points-of-interest (POIs) are an important carriers of location text information in smart cities and have been widely used to extract and identify urban functional regions. However, it is difficult to model the relationship between POIs and urban functional types using existing methods due to insufficient POIs information mining. In this study, we propose a Global Vectors (GloVe)-based, POI type embedding model (GPTEM) to extract and identify urban functional regions at the scale of traffic analysis zones (TAZs) by integrating the co-occurrence information and spatial context of POIs. This method has three main steps. First, we utilize buffer zones centered on each POI to construct the urban functional corpus. Second, we use the constructed corpus and GPTEM to train POI type vectors. Third, we cluster the TAZs and annotate the urban functional types in clustered regions by calculating enrichment factors. The results are evaluated by comparing them against manual annotations and food takeout delivery data, showing that the overall identification accuracy of the proposed method (78.44%) is significantly higher than that of a baseline method based on word2vec. Our work can assist urban planners to efficiently evaluate the development of and changes in the functions of various urban regions.

Keywords: urban functional regions; geo-text mining; points-of-interest; GloVe

Citation: Zhang, C.; Xu, L.; Yan, Z.; Wu, S. A GloVe-Based POI Type Embedding Model for Extracting and Identifying Urban Functional Regions. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 372. <https://doi.org/10.3390/ijgi10060372>

Academic Editor: Wolfgang Kainz

Received: 29 March 2021

Accepted: 28 May 2021

Published: 31 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

As cities develop, various types of functional regions that support residents in their daily life, and work, form gradually. The functional type of a region is determined by various factors, such as human economic activities, infrastructure, and human mobility trends [1]. Monitoring how functional regions and land use evolve across a city is of great practical significance to urban planners, and can be achieved by efficiently and accurately extracting and identifying the functional types of regions [2,3]. Remote sensing images with high spatial resolution have become a widely used data source for extracting urban functional regions because they accurately reflect the physical features of the land surface, such as its shape, texture, and spectral characteristics [4–6]. In addition to these natural attributes, however, the function of an urban region is strongly influenced by human social and economic activities, the characteristics of which cannot be adequately extracted from remote sensing images.

To overcome this limitation, urban computing studies have used various types of geospatial data with social and economic attributes, such as points of interest (POIs), social media data, smartphone signal data, and taxi trajectory data [3,7–11]. Of these geospatial data sources, the easy accessibility of POIs makes them advantageous for studying the structure and function of urban spaces [12–14]. However, although most existing

methods for extracting and identifying urban functional regions based on POIs utilize either POI statistics [12,14] or their relative spatial relationships [15], these play equally important roles in defining the functional type of a given local space [16]. For example, both the relationship between shopping malls and restaurants in local space and the co-occurrence information in global space are important when defining commercial regions.

In the process of rapid urban development, the human–land relationship (e.g., the spatial distribution of human activities, the intensity of land use) is constantly changing. In the era of big data, POIs have the beneficial characteristics of being lightweight and easy to access. Furthermore, human activities usually take place in POIs [12]. Therefore, POIs play an important role in the effective and accurate simulation and prediction of changing trends in the human–land relationship. With the objective of extracting and identifying urban functional regions, we propose a POI information mining method that takes into account both the global co-occurrence frequency of POIs and their geographical relevance in a specific geographic window, thus integrating POI statistics with spatial semantic features to construct a Global Vectors (GloVe)-based POI type embedding model (GPTEM), in which POI type vectors can more accurately express the role of POIs in urban spaces. The main contributions of this paper are as follows:

- Based on the advantages of existing topic model-based and neural network-based methods, the method outlined in this paper takes the global statistical information and the spatial context of POIs into account, and accurately captures the co-occurrence information and spatial semantic features of POI types using the POI type embedding model. The proposed method more accurately extracts large-scale and spatial feature information from POI datasets, thus improving on existing POI information mining methods.
- Based on the obtained POI type embedding, this method can capture sufficient human–land interaction information for each study unit in the extraction and identification of urban functional regions. Compared with the baseline method, it therefore enables more efficient and accurate modeling of the relationship among POIs, human activities, and urban functional types. Moreover, we qualitatively and quantitatively verify the results by comparing them to manual annotations and food takeout delivery data.

With the ever-changing trends in human–land relationships, the method proposed in our study can produce rapid and effective scientific simulations and predictions of urban spatial functions by collecting related data. This presents a great opportunity for city managers to promote future urban planning and smart governance in defining future cities.

2. Related Work

In the field of urban computing, topic model-based methods can be used to mine information hidden in data with socioeconomic attributes. Conventionally, physical characteristics of land surfaces extracted from remote sensing images have been widely used to classify land use type and extract urban functional regions [17,18]. However, the functional types of regions in a city are strongly influenced by the interaction between humans and the surface [19,20]. Hence, solutions that analogize data with socioeconomic attributes, such as POIs, social media data, and taxi trajectory data, to natural language are essential for mining the latent features of urban spaces.

Yuan et al. [14] applied the topic model and natural language processing (NLP) to analogize urban regions to documents, urban functions to topics, POI types to metadata, and human movement patterns to words; in doing so, they identified urban functional regions by combining POI and taxi trajectory data. In cities, however, highly precise dynamic mobility data are difficult to obtain. Gao et al. [12] extracted urban functional regions by analyzing POI co-occurrence patterns after combining social media check-in data with POIs and integrating the data into the latent Dirichlet allocation (LDA) model. LDA,

a topic model in the field of NLP, is typically used to distinguish words by mining the topic of the document and classifying the words into different topics [21], and to estimate the continuous representation of words in vector space. When applied to large geospatial datasets, LDA has the disadvantages of high computational cost and low efficiency [22]. Moreover, topic models such as LDA often ignore the linear relationship of words in vector space; that is, similar words should be closer to each other in vector space [23]. In addition, according to Tobler's first law of geography, closer geographical elements have stronger correlations. The contextual information of geospatial data is thus essential to a spatial semantic understanding. However, when using a topic model to extract urban functional regions, a large amount of spatial contextual information is lost [23].

With advances in machine learning, neural network-based methods are increasingly used to explore the relationships among POIs or other data with socioeconomic attributes in geographic space. Mikolov et al. [23,24] proposed the word2vec algorithm composed of two architectures, namely, CBOW and Skip-gram. This algorithm considers trained model parameters as vectorized representations of the input words and embeds each word in the document as a low-dimensional dense vector. Neural network-based approaches, such as word2vec, overcome the problem of a loss of spatial contextual information, which is encountered when using topic models.

Recently, urban computing researchers have used word2vec to extract the semantic features of data with socioeconomic attributes and to examine the urban spatial distribution pattern. For example, Yan et al. [25] proposed a method to capture the semantic information of place types by measuring similarities and relations between them. They achieved this by enhancing the spatial context to learn the word embedding of the place types. Using a word embedding model, Liu et al. [26] used POI data to examine the niche pattern of locations in a city, with the set of locations co-occurring with a central location defined as the niche of the central location, and quantitatively analyzed regional variability. Trajectory data [27], Twitter data [28] and user review data on travel websites [29] have also been combined with word embedding models to extract urban spatial features.

The aforementioned works have focused on examining the spatial features of cities at a micro level. However, as urban functional regions are the result of the interaction between humans and the land surface and the spatial accumulation of social and economic features in cities [3,30], the features of a single place type cannot completely describe the functional types of an urban region. Yao et al. [15] modeled the relationship between spatial distributions of POIs and land use types based on word2vec. Following the work of Yan et al. [25], Zhai et al. [31] identified functional types in various regions of a city by combining place2vec with POI data and clustering the POI feature vectors at the neighborhood area (NA) scale. Aside from neighborhood area, traffic analysis zones (TAZs) generated from OpenStreetMap road network data provide an alternative scale for applying POIs to the field of urban studies [32,33]. TAZs are also used to divide the urban space in this study. Different from the work of Wang et al. [33], which regard POIs as the basis for analyzing land use intensity, we extract and identify urban functional regions based on the information mined from POI. In addition, place2vec, proposed by Yan et al. [25], as an improvement on word2vec, performs well in word analogy tasks and better retains the contextual information of geospatial data. However, word2vec and its improved methods (e.g., place2vec) ignore the statistical information, such as global co-occurrence frequency, in geospatial data, which is not ideal for urban computing applications [16]. Co-occurrence patterns are crucial pieces of spatial feature information in urban regions. For example, if shopping malls and restaurants appear more frequently in a region, then it is likely to be annotated as a commercial region.

To solve this drawback of word2vec, Pennington et al. [16] proposed a word embedding model named Global Vectors (GloVe) that integrates the global statistical information on the basis of the local context window-based method, thus combining the advantages of both topic models and neural network-based models. Jeawak et al. [34] com-

bined Flickr (a photo sharing website) tag data and traditional structured data (e.g., temperature, land use) to embed geographical locations into vectors, verifying that GloVe is more effective than are traditional bag-of-words models in ecology-related tasks (e.g., predicting species distribution, soil types, and land cover). To the best of our knowledge, the feasibility of applying the GloVe model to urban computing remains to be proven. In this work, we use the architecture of the GloVe model to train POI type vectors.

3. Methodology

The method proposed in this study consists of the following sections: Construction of the urban functional corpus, Embedding the POI type in vector space, Identifying the urban functional type, and Evaluating the accuracy of the urban functional label.

3.1. Construction of the Urban Functional Corpus

In NLP, a corpus is a large number of processed text collections in a predetermined format [35], typically containing documents that can be used to simulate natural language on a large scale. Specifically, the contextual relationship between words in each document can be used to simulate the contextual relationship of words in real-world human languages. The distribution of POIs in cities is similar to the distribution of word frequencies in natural language corpora, both of which follow the power law distribution [25]. Therefore, the semantic relationship of POIs in urban space can be obtained using a word embedding model. Inspired by this concept, in this work, we analogize the study area to a corpus in natural language and call it the urban functional corpus. We then analogize the collection of POIs in the buffers centered on each POI to documents in natural language.

First, we allow the word embedding model to learn the contextual relationship between adjacent POI types within a certain range. In addition, we make the POI types that are closer to each other co-occur more frequently; this allows us to account for the proximity of geographic space in the model, enabling it to distinguish the relationships between POI types with different spatial distance. Each document contains an indefinite number of POIs to be analogized to words in natural language. We use the POI dataset on Gaode Map as the research object. To capture semantic information in richer detail when training POI type vectors, we include third-level POI types in the dataset in addition to both first- and second-level POI types. Table 1 shows examples of urban functional corpus. We can find that the POI types in the collection have similar effects on shaping the urban function types, and POIs of similar types (e.g., *government agencies* and *township-level governments and institutions*) co-occur frequently.

To reflect the distribution features of POIs in geographic space and preserve the positional relationship between adjacent POIs, we need to make the geographically adjacent POIs relatively close to each other when constructing the documents. However, POIs are distributed nonuniformly in urban space. In other words, the number of words in each document is uncertain, and the distribution of POIs is relatively dense in some regions. If documents are constructed based on the principle of global optimization, the time cost is extremely high, rendering the method unfeasible. Therefore, we refer to the shortest-path method proposed by Yao et al. [15] when constructing documents. We build buffer-based documents based on the greedy algorithm, which integrates a series of POIs in the buffer into documents that have actual geographic spatial meanings and can reflect the positional relationships between POIs to some extent. The algorithm flow is as follows:

- (1) We assume that the set of n POIs in a buffer is $A = \{P_1, P_2, \dots, P_n\}$. First, we calculate the distance between all POI pairs in the buffer. We then take the farthest POI pair as the end points of the path, namely P_s and P_e (e.g., *frontage courtyard* and *township-level governments and institutions* in the first item of Table 1). We record the path length as l (e.g., in the first item of Table 1, l is the spatial distance between *frontage courtyard* and *township-level governments and institutions* in the current step). The remaining

POIs in the buffer are included in the set of wait-to-insert. At this time, the set composed of the starting and ending points of the current path is $L = \{P_s, P_e\}$, and the wait-to-insert set is $W = \{P_w \mid P_w \in A - L\}$.

- (2) In this step, all POIs to be inserted are traversed. For each POI P_i in W , we calculate the distance l of the path after P_i is inserted into any segment in L (e.g., in the first item of Table 1, in the first traversal, there is only one segment between *frontage courtyard* and *township-level governments and institutions*; in the next traversal, when *government agencies* is inserted into L , there are two segments between *frontage courtyard* and *government agencies*, *government agencies* and *township-level governments and institutions*). This traversal is to ensure that POI P_i (where P_i can be any POI in wait-to-insert set W) is inserted into the proper position. Here, the “proper position” means that if P_i is inserted into the position between P_x and P_y in the path, the total length l (i.e., the sum of spatial distances between the POI pairs calculated in the order in L : in the first item of Table 1, in the first traversal, the total length l is the sum of spatial distances between *frontage courtyard* and P_i , and between P_i and *township-level governments and institutions*) of the set $L = \{P_s, \dots, P_x, P_i, P_y, \dots, P_e\}$ after insertion is the shortest among all possible options.
- (3) The previous step is repeated until the wait-to-insert set W is empty. The resulting shortest-path sequence is the buffer-based document.

Table 1. Examples of urban functional corpus.

Collections of POIs in the Buffers (Original Data)	The full Name of POI Corresponding to the Index
Jie Fu Xiang Fu Fu Fu Xiang Xiang Xiang Xiang Jian Zhi Jian Fu Xiang Ti Ti Gao Jian Zhi Si Er Wang Shou Wei Bin Chao Ju A Tian Yu Z Zhong Bin Ti Xi Te Bu Kuai Xi Chao Ju Chu Niao Hui Xi Fu Xue You	Jie: Frontage courtyard; Fu: Government agencies; Xiang: Township-level governments and institutions; Jian: Public prosecutors’ office; Zhi: Social security organization; Ti: Sporting Goods Store; Gao: Pastry shop; Si: Company; Er: Children’s Store; Wang: Internet technology; Shou: Mobile phone sales; Wei: Repair site; Bin: Hotel; Chao: Supermarket; Ju: Home building materials market; A: Furniture mall; Tian: Dessert shop; Yu: Entertainment venues; Z: KTV; Zhong: Chinese restaurant; Xi: Place for bathing and massage; Te: Specialty/Local restaurant; Bu: Fabric market; Kuai: Fast-food restaurant; Chu: Kitchen and bathroom market; Niao: Flower, bird, fish and insect market; Hui: Flower market; Xue: School; You: Kindergarten

Note: We replaced the Chinese characters in the original data with their corresponding Chinese phonetic alphabet.

3.2. Embedding the POI Type in Vector Space

In NLP, word embedding—first proposed by Bengio et al. [36]—is a language feature extraction technology that transforms words into numerical representations in a high-dimensional vector space [37]. The biggest advantage of word embedding models is that they do not require manually labelled data; instead, these models can quickly train unlabeled data using unsupervised methods at a low computing cost [23,24,38]. In contrast to word2vec and other word embedding models trained on local context windows, GloVe trains word embedding on the basis of the number of global word co-occurrences, making it possible to directly capture global statistical information in the corpus and to represent the semantics of words as vectors [16]. In addition, GloVe optimally uses the available

computing resources, making it easy to train large datasets quickly to efficiently obtain the semantic features of urban spaces. As described earlier, we analogize the urban space in our study area to a corpus in natural language, called an urban functional corpus, which contains a large number of POIs. The co-occurrence pattern of various types of POIs is of great significance to understanding the semantics in urban spaces. For example, Table 2 shows some statistical indexes of POI collections in the urban functional corpus. It is noted that the length of “sentences” (i.e., POI collections) in urban functional corpus differs greatly. For the long sentence (e.g., the POI collection with maximum length), GloVe can capture the global co-occurrence information for the whole sentence. Inspired by GloVe, this paper presents a GloVe-based POI type embedding model (GPTeM).

Table 2. Statistical indexes of POI collections in the urban functional corpus.

Maximum Length	Minimum Length	Average Length
782	1	65.73

Assuming that the co-occurrence matrix of POIs is X and that the element X_{ij} in X represents the number of occurrences of POI type j in the context of POI type i , $X_i = \sum_k X_{ik}$ represents the number of occurrences of any POI type in the context of POI type i , and $P_{ij} = P(j|i) = \frac{X_{ij}}{X_i}$ represents the probability that POI type j ap-

pears in the context of POI type i . When training POI type vectors based on urban functional corpora, the ratio of the co-occurrence probabilities of the POI types is more useful for identifying highly spatially related POI types than the co-occurrence probability itself. As examples, we take the following third-level POI types on Gaode Map: *College and University*, *Research Institution*, *Government and Institution below the Township Level*, *Other Agriculture*, *Forestry*, *Animal Husbandry and Fishing Base*, *Logistics Express*, and *Securities Company*. We assume that $i = \text{College and University}$, $j = \text{Government and Institution below the Township Level}$, and k is the other types of POI to be compared. In cities, *Research Institution* is usually spatially related to *College and University*, whereas its correlation with *Government and Institution below the Township Level* in urban space is relatively weak. In other words, when $k = \text{Research Institution}$, we expect P_{ik} to be much larger than P_{jk}

and therefore $\frac{P_{ik}}{P_{jk}}$ to be very large; conversely, when $k = \text{Other Agriculture, Forestry,}$

Animal Husbandry and Fishing Base, we expect this ratio is very small. In addition, for *Logistics Express* types that are spatially related or for the *Securities Company* unrelated to *College and University* and *Government and Institution below the Township Level*, for which the values of P_{ik} and P_{jk} are very close, the co-occurrence probability ratios are close to 1. Therefore, compared with the co-occurrence probability itself, the ratio of the co-occurrence probabilities is more helpful in distinguishing different types of POIs, and thus in more accurately obtaining the spatial semantic features of POIs in cities.

Moreover, the co-occurrence pattern of POIs in our study is symmetrical—that is, the roles of the POI type i and its contextual POI type j in the training process are interchangeable. In addition, considering that the vector space is linear and the final loss function should be as simple as possible, the generalized form of the model is as follows:

$$G((v_i - v_j)^T \widetilde{w}_k) = \frac{P_{ik}}{P_{jk}} = \frac{G(w_i^T \widetilde{w}_k)}{G(w_j^T \widetilde{w}_k)} \quad (1)$$

In Equation (1), v_i and v_j refer to the vectors of POI type i and j , respectively, \widetilde{w}_k represents the vector of POI type k in the context; and G represents a function of v_i , v_j , and \widetilde{w}_k . Setting $G = \exp$, to keep our model symmetrical, we introduce a deviation term b and obtain the following formula:

$$v_i^T \widetilde{w}_k + b = \log(X_{ik}) \quad (2)$$

Clearly, the co-occurrence matrix we obtain is sparse. To limit the value of our loss function when $X_{ik} = 0$ and to ensure that POI types overly small or large co-occurrence frequencies are not overrepresented [16], we define the loss function of the model as follows:

$$L = \sum_{i=1}^V \sum_{j=1}^V f(X_{ij}) (v_i^T \widetilde{w}_k + b - \log(X_{ij}))^2 \quad (3)$$

In Equation (3), V refers to the total number of POI types in the urban functional corpus, and the weight function $f(X_{ij})$ is defined as follows [16]:

$$f(x) = \begin{cases} \left(\frac{x}{100}\right)^{0.75} & \text{if } x < 100 \\ 1 & \text{otherwise} \end{cases} \quad (4)$$

The number of dimensions of the embedding vectors is an important parameter of the GPTM. Considering that the total number of POIs in the dataset is not as rich as natural language vocabulary, we set the dimensions of the POI type vectors to 70, which has been proven to be effective in related studies [25,31].

To visually analyze the spatial distribution features of the 70-dimensional POI type vectors, we apply the data dimensionality reduction technique for mining information from high-dimensional data [39,40]. The dimensionality reduction method used in this study is t-distributed stochastic neighbor embedding (t-SNE), which is a widely used unsupervised machine learning algorithm for dimensionality reduction [41]. Compared with other dimensionality reduction methods (e.g., principal component analysis), t-SNE better maintains the local structure, meaning that it can better map POI type vectors that are closer in high-dimensional vector space into the adjacent low-dimensional vector space [42].

3.3. Identifying the Urban Functional Type

We obtain all POI type vectors using the GloVe-based POI type embedding model. In related studies, the features of documents have usually been described by calculating the weighted average of words in the documents. In this study, the urban space study area is divided into a series of TAZs based on its administrative area boundaries and the first three levels of the road network. The functional features of a TAZ are then represented by the weighted average of all POI type vectors it contains, as follows:

$$TAZ_{vec_i} = \frac{\sum_{j=1}^N type(p_{i,j})}{N} \quad (5)$$

where TAZ_{vec_i} represents the feature vector of the i -th TAZ, $type(p_{i,j})$ refers to the vector of POI type j in the i -th TAZ, and N represents the total number of POIs in the i -th TAZ.

The TAZ feature vectors calculated using the above formula are very high-dimensional. K-means is an algorithm widely used in high-dimensional data clustering and has

the advantages of fast convergence and strong interpretability. However, in the original K-means algorithm, the initial clustering center needs to be determined manually, and different clustering centers may yield completely different clustering results. To solve this problem, this paper selects K-means++, an improved clustering algorithm, to extract the urban functional clusters based on TAZs. The basic principle of the K-means++ algorithm is to select the initial clustering center such that the distance between clusters is as far as possible and the final clustering result is as close as possible to the global optimal solution [43].

In the K-means++ algorithm, the choice of the number of clusters (K) strongly affects the quality of the clustering. Recent studies have used the silhouette coefficient to evaluate the clustering results [12,15]. This coefficient evaluates whether a data point is accurately classified by measuring its similarity to the category to which it belongs. To make the clustering results more objective and accurate than the methods of selecting the number of clusters based on experience or expertise, we evaluate the accuracy of the clustering by calculating the average silhouette coefficient for all data samples.

After obtaining the clustering results, actual spatial meanings must be assigned to each cluster to identify the urban functional regions; that is, the clusters must be annotated with urban function labels. Previous studies have typically annotated urban functions based on indicators related to the occurrence frequency (or density) of POI types (e.g., composition ratio, density) [15]. However, in urban spaces, some types of POIs may appear frequently, meaning that this frequency- or density-based approach may produce large errors. To avoid this problem, we evaluate the urban functions of the clustering regions by calculating the enrichment factors (EF) of the POI types in each clustering region, as follows:

$$EF_i^j = (N_i^j / N_i) / (N^j / N) \quad (6)$$

where EF_i^j refers to the EF of POI type j in the i -th cluster region, N_i^j represents the number of POI type j in the i -th cluster region, N_i represents the total number of POIs in the i -th cluster region, N^j represents the total number of POI type j in the city, and N represents the total number of POIs in the city.

3.4. Evaluating the Accuracy of the Urban Functional Label

To quantitatively verify the accuracy of our identification of urban functional regions, we need data with urban functional labels for each TAZ; that is, the accuracy of unsupervised clustering needs to be evaluated through a supervised task. Accordingly, to label the data as accurately as possible, with the urban land use plan issued by the government, 10 volunteers graduate students and experts with background knowledge of urban planning were invited to annotate the urban functional types in each TAZ according to the list of urban functional types identified by the proposed method.

Assigning some TAZs to a single urban functional type may be difficult, as different volunteers could judge the urban functional type of the same TAZ differently. Therefore, the volunteers were asked to annotate the urban functional types in each TAZ as two different urban functions, thus eliminating subjective interpretations as much as possible. Assuming that the first and second annotation by the i -th volunteer for the j -th TAZ are $F_i[TAZ_1^j]$ and $F_i[TAZ_2^j]$, respectively, we calculate the manual labeling output of the urban functional type in the j -th TAZ as follows:

$$Function_j = Mode\{w_1 F_i[TAZ_1^j], w_2 F_i[TAZ_2^j], \dots, w_1 F_{10}[TAZ_1^j], w_2 F_{10}[TAZ_2^j]\} \quad (7)$$

In Equation (7), $Function_j$ represents the final result of the manually annotated urban functional type, $Mode$ represents the weighted mode of the urban functional

type, and $w_1 + w_2 = 1$. Considering that Zhai et al. [31] demonstrated that the first manual annotation of an urban functional type is much more important than the second, we set $w_1 = 0.8$ and $w_2 = 0.2$.

To qualitatively evaluate our identification of urban functional regions, we use food takeout origin–destination (OD) data for the main urban area of our study area. Here, “takeout” refers to services that deliver food from restaurants to areas such as residential buildings, schools, and workplaces. The reason why we use food takeout OD data is that food takeout OD data can be a kind of supplementary data to verify the identification results from the side. Specifically, because takeout OD data show the start (origin) and end (destination) points of takeout deliveries ordered by residents, they can reflect the distribution of different urban functional types of regions in a city to some extent. These data can then be used to evaluate the accuracy of identification of residential regions, public service regions, daily life service regions, and other similar urban functional region types. In this study, the spatial distribution of the origins and destinations of takeout food are expressed through kernel density estimation, a nonparametric method used to estimate the probability density function. This is an effective means to estimate the density distribution of any discrete points and does not depend on the scale of the space division [44]. Thus, the kernel density estimation of takeout OD data can reveal the regions where takeout orders are concentrated to a certain extent, which is helpful when evaluating the results of identification of urban functional regions.

3.5. Overall Architecture

Figure 1 depicts the overall flowchart of the proposed method in the study, in which we perform the following four-stage extraction and identification method:

- (1) We construct an urban functional corpus based on the greedy algorithm using POI data and the buffer centered on each POI.
- (2) Based on the urban functional corpus, we obtain POI type vectors using the GPTEM and visualize the configuration pattern of the POI type vectors.
- (3) We cluster the TAZ feature vectors calculated using the POI type vectors and calculate the enrichment factor of each POI type in the clustered regions to extract the urban functional type of each region in the study area.
- (4) Finally, we quantitatively and qualitatively evaluate the identification results by comparing them to the urban functional labels of each TAZ annotated by volunteers and the takeout OD data.

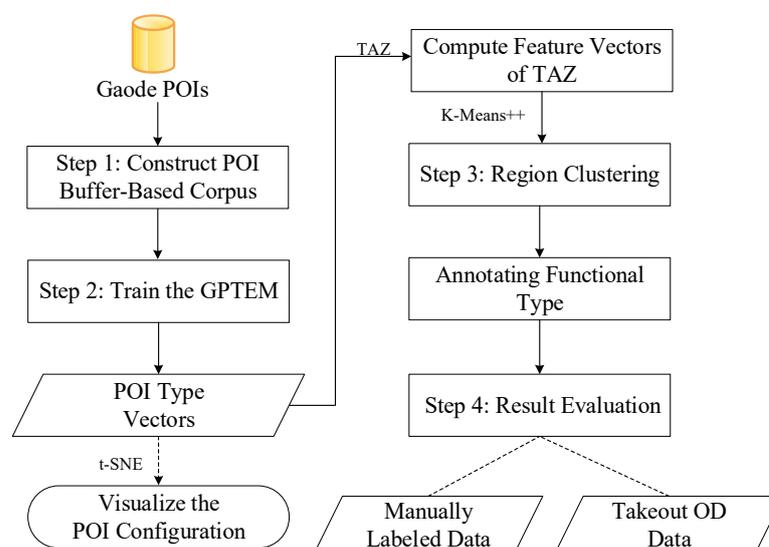


Figure 1. Flowchart of the proposed method in the study.

4. Study Area and Dataset

The study area is the main urban area of Hangzhou in Zhejiang province, southeastern China. Hangzhou is the political, economic, and cultural center of Zhejiang, and the central city in the southern part of the Yangtze River Delta. With the rapid development of the city over recent years, the urban functions of the main urban area have become highly heterogeneous and mixed. Considering the distribution of public activities, we take the eight major districts of Hangzhou, namely *West Lake*, *Gongshu*, *Jianggan*, *Xiacheng*, *Shangcheng*, *Binjiang*, *Xiaoshan*, and *Yuhang*, as our study area.

In this study, the POI dataset is obtained through the API (<https://restapi.amap.com/>) of the Gaode Map Service in 1 June 2020, one of the most widely used map service providers in China. An example request submitted to this service is <https://restapi.amap.com/v3/place/text?key=myKey&extensions=all&keywords=&types=010100&city=myCity&citylimit=true&offset=25&page=1&output=json> (where “myKey” and “myCity” are replaced according to actual needs. We accessed the service in 1 June 2020). This dataset contains more than 300,000 POI records with multiple category levels. Figure 2 shows a data sample of Gaode POIs. From the first-level POI types, we remove *Address Information* (which has negligible influence on urban function) and *Vehicle* (there are few POIs under *Vehicle*, and this does not affect evaluation of the accuracy of identification of urban functional regions). We include third-level POI types in the dataset because they reflect the nuances of urban functions more than the first and second-level POI types. For example, *Internet Technology* and *Metallurgy & Chemical Industry* are third-level POI types under the second-level POI type *Company*, which in turn is under the first-level POI type *Company & Enterprise*. The constructed POI dataset has 17 first-level POI types—*Catering Service*, *Road Ancillary Facilities*, *Famous Tourist Sites*, *Public Utilities*, *Company & Enterprise*, *Shopping Service*, *Transportation Facilities Service*, *Financial Insurance Service*, *Science Education and Cultural Services*, *Business Residence*, *Life Service*, *Event & Activity*, *Sports and Leisure Service*, *Access Facilities*, *Healthcare Service*, *Government Agency & Social Group*, and *Accommodation Service*—115 second-level POI types, and more than 400 third-level POI types.

```
{
  "longitude":120.1537,
  "latitude":30.170184,
  "name":"sticky rice",
  "adname":"Binjiang District",
  "smalleclass":"Chinese restaurant",
  "typecode":"050100",
  "midclass":"Chinese restaurant",
  "largeclass":"Catering Service"
}
```

Figure 2. A data sample of Gaode POIs.

Furthermore, by merging the administrative area boundary data with data on the first three levels of the road network, the study area is divided into 997 TAZs. To improve the validity and reliability of the results, small patches of area less than 500,000 square meters and TAZs without any POIs are merged into their neighboring TAZs.

5. Experiments and Results

5.1. Training and Visualizing POI Type Vectors

In our study area, 357,990 POIs are distributed across 997 TAZs. To avoid confusion between the POI types, we establish an index table giving each POI type a corresponding unique index value to replace the role of the POI type in the remainder of the analysis. Table 3 presents examples of this index table.

To construct the urban functional corpus, the buffer zone radius is set to 50 m. Before training the POI type vectors using the proposed model, we set the number of POI type

vector dimensions to 70, window size to 10, and epoch to 10. Thus, 476 third-level POI types are converted into POI type vectors that indicate the spatial context relationship. Table 4 presents examples of POI type vectors.

Table 3. Example entries in the index table.

Index Value	POI Type
Zhong	Chinese Restaurant
Gong	Factory
W	Residential Area

Note: We replaced the Chinese characters in the original data with their corresponding Chinese phonetic alphabet.

Through the dimensionality reduction method of t-SNE, all POI type vectors are mapped into three-dimensional semantic space, as shown in Figure 3. In the semantic space after dimensionality reduction, POI types with similar or related spatial semantics tend to be close to each other. For example, the subplot (a) of Figure 3 includes several types of financial POIs (e.g., HUAXIA Bank, ICBC, and Hang Seng Bank ATM), all of which play a similar role in urban space; hence, these POIs are relatively close to each other in the semantic space. Of note is that some of the closer POIs in the subplot (b) of Figure 3 belong to different first-level POI types. For example, Cinema belongs to Sports and Leisure Service, Infant Room belongs to Public Utilities, and Yungui Restaurant belongs to Catering Service. However, they usually appear together as they all meet people's daily leisure and entertainment needs. These accurate and nuanced results demonstrate the rationality of choosing third-level POI types in addition to first- and second-level types to construct the urban functional corpus.

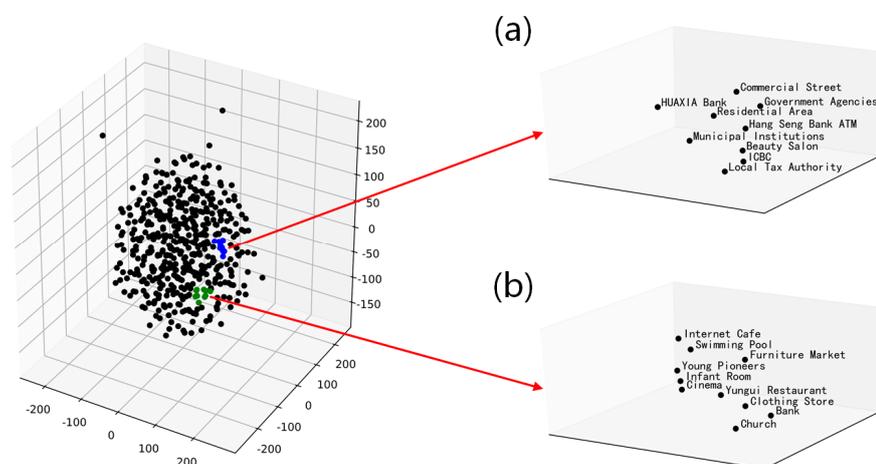


Figure 3. POI type vectors in three-dimensional semantic space, with example local types that are close to each other (a,b).

5.2. Clustering the Urban Regions

The feature vectors of TAZs are obtained by calculating the weighted average of the POI type vectors in any given TAZ. While clustering TAZ feature vectors using K-Means++, we experimentally determine the optimal K value (i.e., the number of clusters, which is the most influential parameter) by varying K from 4 to 15; the resulting TAZ feature vectors are clustered using K-Means++ and the corresponding silhouette coefficients calculated. Figure 4 shows that the silhouette score is highest when K = 7. Therefore, we set K = 7, with the resulting clustering shown in Figure 5.

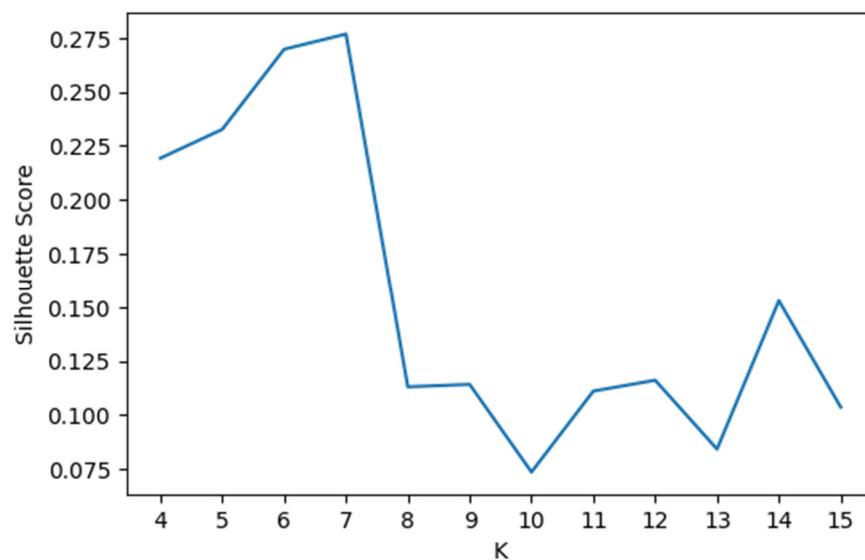


Figure 4. Silhouette score (y-axis) as a function of the number of clusters, K (x-axis).

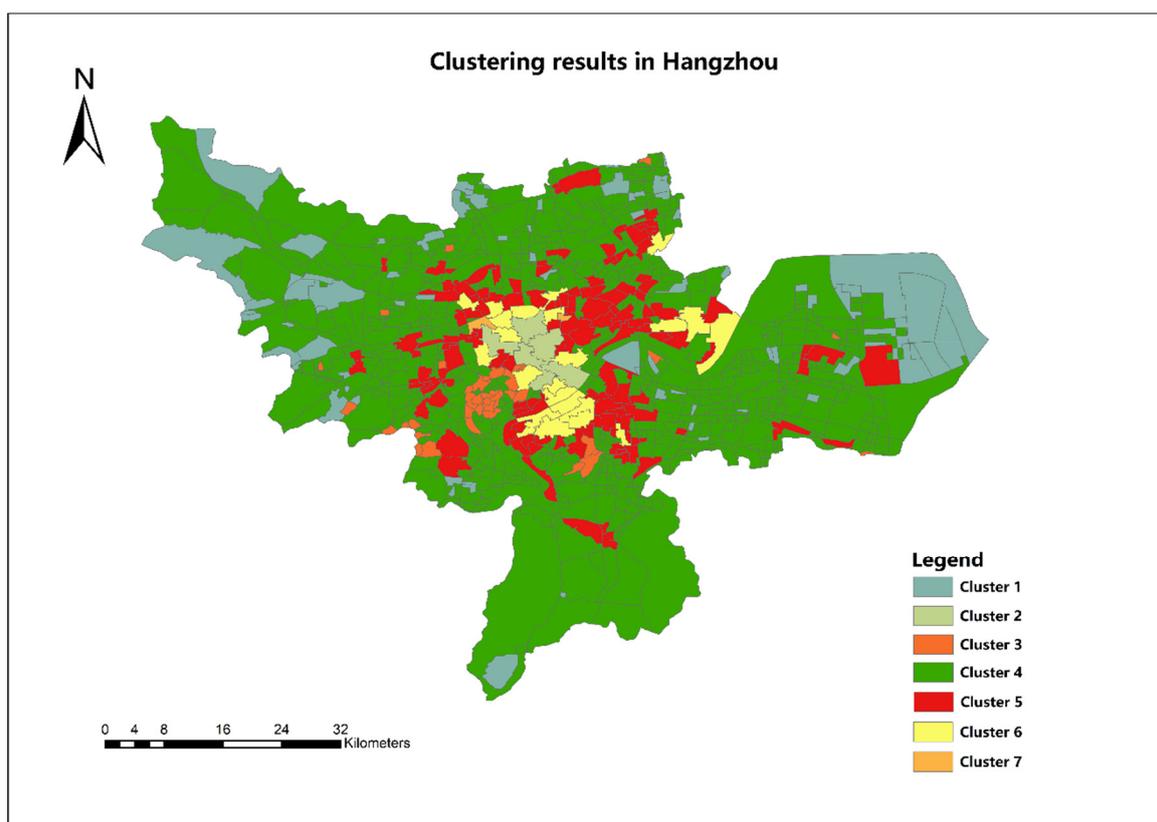


Figure 5. Clustering results for the study area (Hangzhou) obtained at $K = 7$.

5.3. Annotating the Urban Functional Type

To identify the urban functions of each cluster, from among the 476 three-level POI types, we list the top 20 POI types by EF in each cluster in Table 5:

Table 4. Examples of POI type vectors.

POI Type	POI Type Vector (Dimension = 70)
Middle School	(0.00137245, 0.00350524, ..., 0.00271256)
Bar	(0.00526796, 0.00108780, ..., 0.00437424)
International Organization Office	(0.00494723, 0.00629526, ..., -0.00390901)

Cluster 1: Agriculture, Forestry, Animal Husbandry, Fishing, and Industrial Regions. The EFs of the POI types of agriculture, forestry, animal husbandry, and fishery (e.g., Pasture = 188.27, Farm = 51.35, Fishing Park = 15.37) and industry (e.g., Factory = 14.01, Industrial Park = 10.62) are relatively high (see Table 5). In addition, the TAZs covered in this cluster area are mainly distributed away from downtown, indicating that the urban functions of this clustering area are mainly agriculture, forestry, animal husbandry, fishing, and industry.

Cluster 2: Commercial Regions. We define this cluster as commercial because it is concentrated in the most central and prosperous areas of the city. More importantly, POIs with high EFs in this clustering area are mainly commercial banks (e.g., Citibank = 4.18) and insurance firms (e.g., Xinhua Life Insurance Company = 4.18).

Cluster 3: Scenic, Daily Life Service, and Residential Regions. The TAZs of this cluster area are mainly distributed in the scenic areas of the West Lake (the most famous tourist attraction in the study area) and its surroundings. This cluster is annotated as a mixed urban functional area with scenic, daily life service, and residential features because the area shows a mixed POI distribution pattern, covering scenic spots (e.g., National Attraction) and POIs closely related to people's everyday lives (e.g., Tea House, or Service Center).

Cluster 4: Public Service, Agriculture, Forestry, Animal Husbandry, and Fishing Regions. Similar to Cluster 1, most of the TAZs in this cluster area are located on the outskirts of the city. Unlike Cluster 1, however, this cluster area covers one of the most important transportation hubs in the study area—the Hangzhou Xiaoshan International Airport—and related infrastructure providing public services. Some of the POI types with high EF rankings are similar to those in Cluster 1 (e.g., Fruit Base). However, several types of infrastructural POIs related to airports, transportation, and other public services (e.g., Waiting Room) are evident in the EF ranking table (see Table 5).

Cluster 5: Daily Life Service and Public Service Regions. The TAZs in this clustering area are mainly distributed in the areas surrounding the city center and contain colleges (e.g., Yuquan Campus, Zhejiang University, Zhejiang Provincial Institute of Socialism), railway stations (e.g., Hangzhou North Railway Station), and places of daily life and leisure (e.g., Qingzhiwu). In this clustering area, the EFs for POIs of daily life service (e.g., Walmart) and public service (e.g., Vehicle Pass Office = 3.27) are relatively high.

Cluster 6: Commercial and Daily Life Service Regions. Similar to Cluster 2, most of the TAZs in this cluster area are located in the city center. However, in addition to covering business and office premises (e.g., law firms, and banks), this cluster also contains facilities for dining, leisure, and other life services (e.g., Hangzhou World Trade Plaza), as reflected in Table 5. Specifically, the EFs of commercial categories (e.g., Audit Firm, and NCB) and POIs related to people's daily life (e.g., Carrefour, Fujian Restaurant, and Watsons) are relatively high.

Cluster 7: Residential and Daily Life Service Regions. This cluster area contains only two TAZs, mainly covering residential regions and daily life service facilities (e.g., Zijingang Community). Table 5 verifies the presence of such urban functional types in this area, as residential POIs and those related to people's daily lives, such as Furniture Mall, Flower Market, GOME, and Korean Restaurant, have relatively high EFs.

5.4. Evaluating Identification Accuracy

To verify the reliability of the results, we invited volunteers with background knowledge of urban planning to manually annotate each TAZ with its urban functional types on the basis of urban land use planning data (see Figure 6c). As described in Section 3.5, we take the mode of their labels as the final urban functional labels.

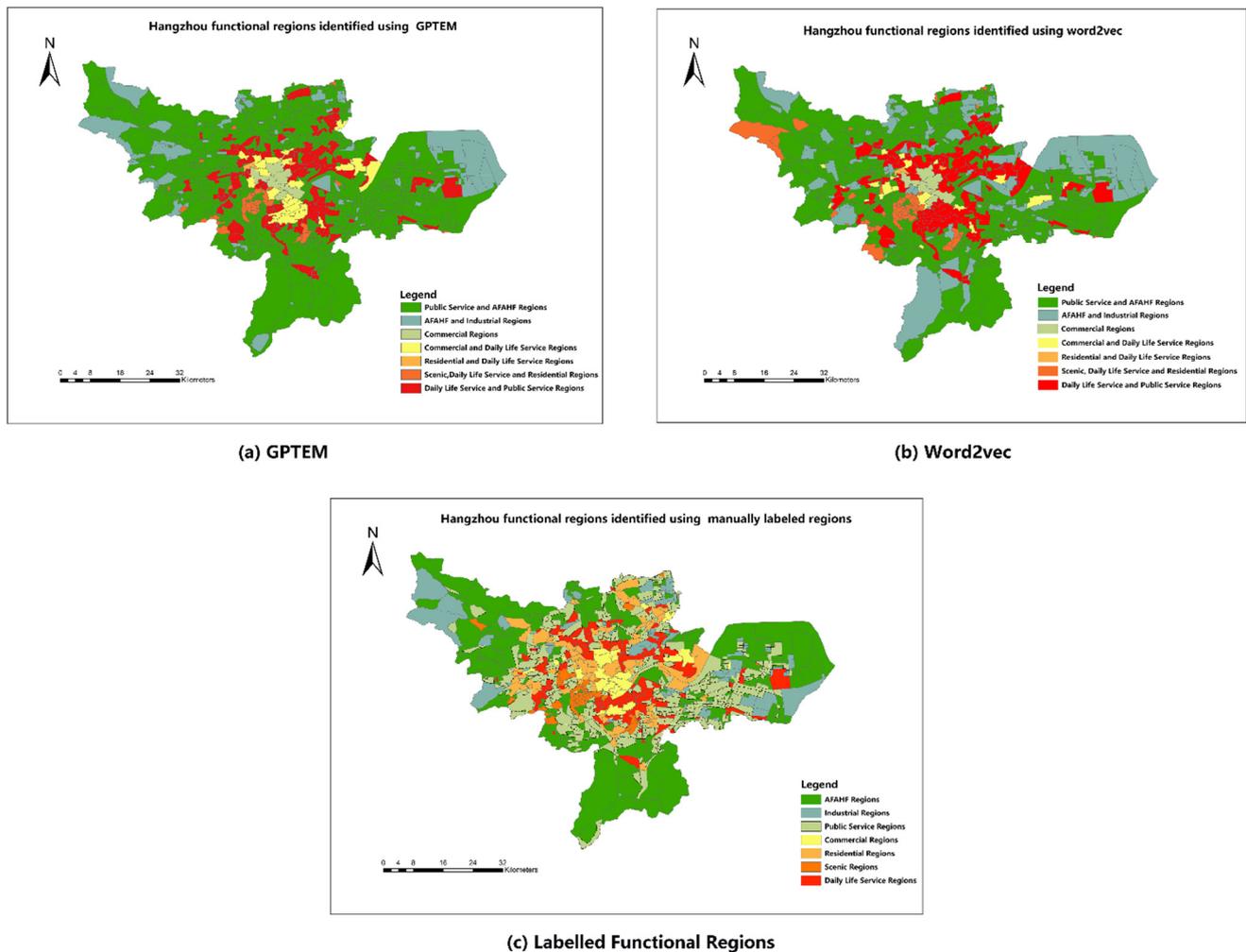


Figure 6. Functional regions identified using different GPTEM (a), word2vec (b), and (c) manually labelled regions. Note: the legend used is different in each panel.

For this evaluation, word2vec is used as the baseline method. To eliminate the interference of other factors (e.g., the effects of different POI type vector dimensions on training speed) in the result to the greatest extent possible, when training POI type vectors using word2vec, the hyperparameter values are kept consistent with those set in Section 5.1. Figure 6a, b present the extraction and identification results obtained using GPTEM and word2vec, respectively. According to the calculation results of the enrichment factor in Table 5, each cluster area may contain more than one urban functional type, consistent with a real-world city containing highly heterogeneous and mixed function types. However, as described in Section 3.4, the volunteers manually labelled each TAZ with a single urban functional type, leading to inconsistent legends in Figure 6a–c. For a cluster region that contains more than one urban functional type, we assume that the composition ratio of each urban functional type is the same when calculating the confusion matrix. Tables 6 and 7 illustrates the confusion matrixes of the identification results based on word2vec and GPTEM. To demonstrate the advantages of our GloVe-based method over word2vec, we compare the two models in terms of model training time and overall accuracy (Table

8). The training time with GPTEM is more than seven times shorter than with word2vec. In terms of overall accuracy, GloVe accounts for the relative co-occurrence patterns between POIs in the urban space, and its trained POI type vectors capture urban spatial semantic information more accurately, resulting in the higher overall accuracy of GPTEM. Clearly, the GloVe-based method outperforms the word2vec-based method.

Table 5. Top 20 POI types by EF in each cluster.

POI in Cluster 1	EF	POI in Cluster 2	EF	POI in Cluster 3	EF	POI in Cluster 4	EF	POI in Cluster 5	EF	POI in Cluster 6	EF	POI in Cluster 7	EF
Pasture	188.27	Citibank	4.18	International Organization Office	126.83	Fishing Ground	4.89	Hubei Restaurant	3.27	World Heritage	4.46	Foreign Institution	10.77
Freight Port	62.76	HSBC	4.18	Youth Hostel	37.3	Vegetable Base	4.89	Reebok Store	3.27	British Restaurant	4.46	Furniture Mall	10.34
Poultry Breeding Base	53.79	DBS	4.18	Tea House	34.1	Funeral Parlor	4.89	Shuttle Station	3.27	XiabuXiabuRestaurant	4.46	Shanghai Bank	9.42
Farm	51.35	Citibank ATM	4.18	Park Attraction Ticket Office	33.82	Brazilian Restaurant	4.89	Vehicle Pass Office	3.27	Facilities inside the Park	4.46	BMH Market	9.29
Other AFAHF Base	49.4	HSBC ATM	4.18	National Attraction	31.71	Racetrack	4.89	Border Crossing	3.27	Carrefour	4.46	HBM Market	8.85
Fruit Base	47.07	HK DBS Bank ATM	4.18	Famous Tourist Site	25.14	Waiting Room	4.89	Entrance	3.22	ENT Hospital	3.34	K&B Market	7.77
Famous Enterprise	34.86	Rugby Field	4.18	Upper First-Class Hospital	22.62	Airport Departure/Arrival	4.89	Exit	3	NCB ATM	2.97	Shanghai Bank ATM	6.85
Provincial Attraction	31.38	Brain Hospital	4.18	Viewpoint	21.14	Baggage Inquiry	4.89	Train Station	2.62	Other Asian Restaurant	2.67	Eye Hospital	5.26
Picking Garden	26.27	Tax Authority	4.18	Provincial Attraction	21.14	Campground	4.19	Walmart	2.46	Photographic Equipment Store	2.55	Flower Market	5.24
Highway Service Area	17.12	CDB	4.18	Botanical Garden	19.51	Fishing Park	3.99	Change Office	2.46	Fujian Restaurant	2.55	L&P Market	4.95
Cemetery	17.12	BEA	4.18	Foreign Institution	18.12	Cemetery	3.87	Platform	2.41	Wanning Supermarket	2.43	Video Shop	4.71
Fishing Park	15.37	Foreign Embassy	4.18	Chinese Vegetarian Restaurant	18.12	Resort	3.78	Zoo	2.18	Psychiatric Hospital	2.23	FBFI Market	4.7
Leisure Place	14.34	EIBC	4.18	Campground	18.12	Toll Station	3.7	Freight Port	2.18	Mediterranean Restaurant	2.23	Chaozhou Restaurant	4.43
Toll Station	14.24	HK HSB	4.18	French Cuisine Restaurant	15.85	Fruit Base	3.67	Refund	2.18	Audit Firm	2.23	Vehicle Management Agency	4.19
Factory	14.01	HK HSB ATM	4.18	Service Center	15.47	Airport	3.67	PBC	2.18	Mosque	2.23	Football Field	4.04
Outdoor Fitness Facility	11.17	BEA ATM	4.18	Port	15.14	Traffic Law Enforcement Station	3.6	Hualian Supermarket	2.1	NCB	2.23	Tennis Court	3.68
Industrial Park	10.62	Taikang Life Insurance Company	4.18	Ferry	14.63	Farm	3.56	Ticket Office	1.96	Watsons	1.99	GOME	3.43
Resort	9.91	Heliport	4.18	Memorial	14.59	Aquatic Center	3.56	Commodity Market	1.86	Puma Store	1.98	Korean Restaurant	3.08
Machinery & Electronics	9	Xinhua Life Insurance Company	4.18	Ticket Office	14.23	Picking Garden	3.53	Ticketing	1.82	Italian Restaurant	1.96	Restaurant	3.02
	8.19		3.34		13.88		3.49	PingAn	1.77	Shanghai	1.91	CEB ATM	2.9

Metallurgy & Chemical Industry	Chest Hospital	Temple& Tao- ist Temple	Poultry Breed- ing Base	Restaurant
--------------------------------------	-------------------	----------------------------	----------------------------	------------

Other AFAHF Base: Other Agriculture, forestry, animal husbandry, and fishing base; **HSBC:** The Hong Kong and Shanghai Banking Corporation Limited; **DBS:** Development Bank of Singapore; **CDB:** China Development Bank; **BEA:** The Bank of East Asia; **EIBC:** Export–Import Bank of China; **PBC:** People’s Bank of China; **PingAn:** Ping An Insurance (Group) Company of China; **NCB:** Nanyang Commercial Bank; **BMH Market:** Building material hardware market; **HBM Market:** Home building materials market; **K&B Market:** Kitchen and bathroom market; **L&P Market:** Lamps and porcelain market; **FBFI Market:** Flower, bird, fish, and insect market; **CEB:** China Everbright Bank.

Table 6. Confusion matrix of identification result based on word2vec.

	Public Service Regions	AFAHF Regions	Commercial Regions	Residential Regions	Industrial Regions	Scenic Regions	Daily Life Ser- vice Regions
Public Service Regions	0.153	0.416	0.000	0.000	0.416	0.007	0.008
AFAHF Regions	0.002	0.972	0.000	0.000	0.000	0.012	0.014
Commercial Regions	0.140	0.140	0.560	0.000	0.060	0.020	0.080
Residential Regions	0.274	0.226	0.021	0.021	0.211	0.005	0.242
Industrial Regions	0.256	0.226	0.000	0.000	0.476	0.006	0.036
Scenic Regions	0.010	0.115	0.000	0.000	0.115	0.750	0.010
Daily Life Service Regions	0.080	0.218	0.029	0.000	0.168	0.000	0.505

Table 7. Confusion matrix of identification result based on GPTEM.

	Public Service Regions	AFAHF Regions	Commercial Regions	Residential Regions	Industrial Regions	Scenic Regions	Daily Life Ser- vice Regions
Public Service Regions	0.992	0.003	0.000	0.000	0.003	0.001	0.001
AFAHF Regions	0.000	0.992	0.000	0.000	0.000	0.004	0.004
Commercial Regions	0.160	0.140	0.680	0.000	0.000	0.000	0.020
Residential Regions	0.452	0.263	0.021	0.032	0.000	0.011	0.221
Industrial Regions	0.315	0.298	0.000	0.000	0.369	0.000	0.018
Scenic Regions	0.146	0.125	0.000	0.000	0.000	0.708	0.021
Daily Life Service Regions	0.181	0.180	0.000	0.000	0.000	0.000	0.639

Table 8. Evaluation index of urban function identification.

Methods	Training Time (in Seconds)	Kappa	Overall Accuracy
word2vec	315.48	0.4059	0.4504
GloVe	38.99	0.7000	0.7844

Furthermore, using more than 3,400,000 takeout OD data points provided by Dianwoda Company and covering a one-month period (August 2017), we qualitatively evaluate the extraction and identification results. Figure 7 presents the results of kernel density estimation on the origin and destination of the takeout deliveries.

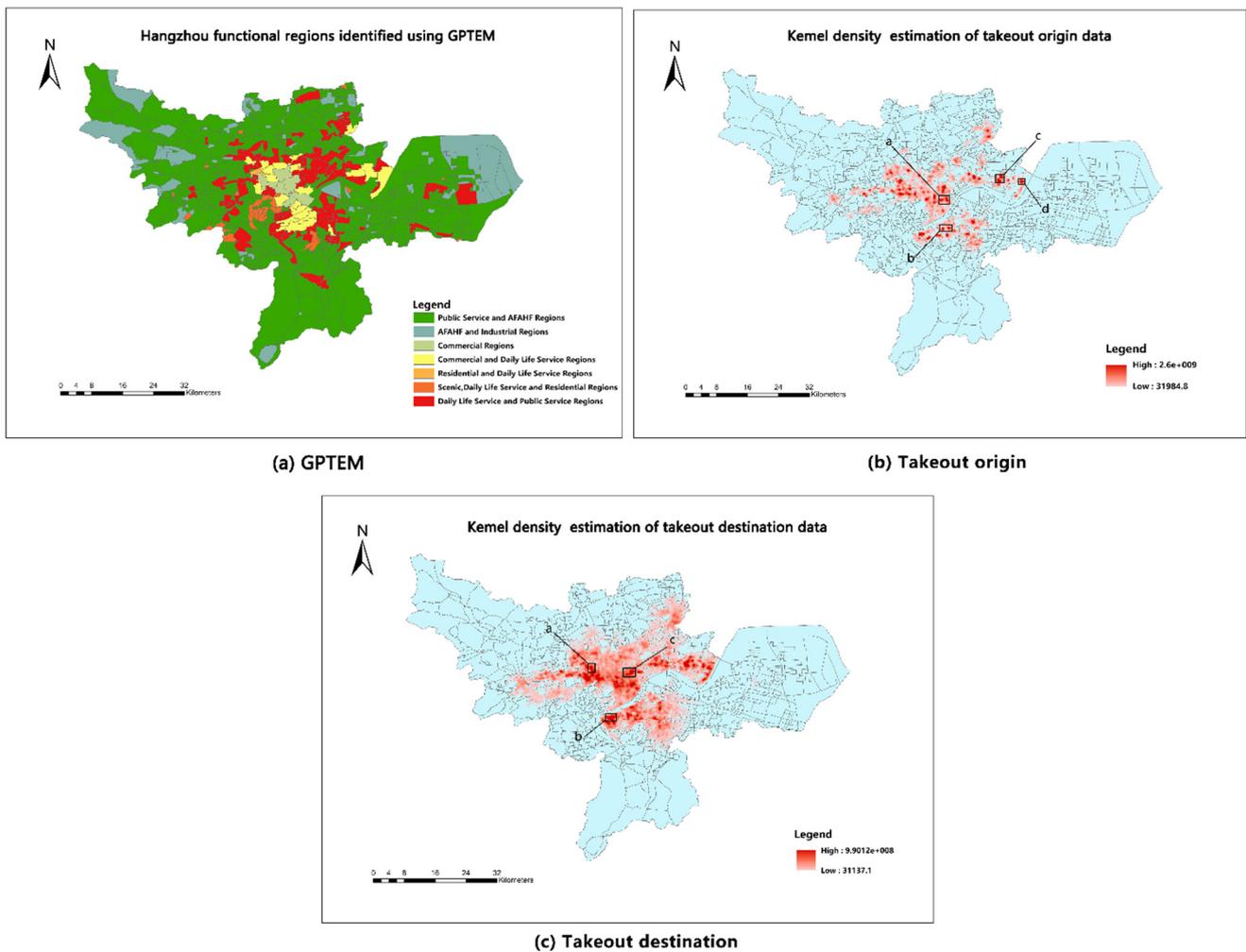


Figure 7. Functional regions identified using GPTEM (a) and kernel density estimation of takeout origin (b) and destination (c) data.

Comparing the results of the kernel density estimation in Figure 7b, c with the results of urban functional region identification in Figure 7a, the start points of takeouts (Figure 7b) are mainly concentrated in the commercial regions (e.g., area a. in Figure 7b) and daily life service regions (e.g., area b, c, and d in the subplot (b) of Figure 7). These two region types include most of the restaurants that provide takeout service. The end points of takeout deliveries are mainly concentrated in residential regions (e.g., area a. in the subplot (c) of Figure 7), commercial regions (e.g., area b in the subplot (c) of Figure 7), and public service regions (e.g., area c in the subplot (c) of Figure 7). Because the takeout data used in this study cover a long period of time (i.e., one month), places in commercial regions (e.g., companies, and banks) and public service regions (e.g., government organizations, hospitals, and schools) may be common destinations for takeout deliveries during working hours, whereas apartments in residential regions would be more common destinations at other times. Thus, the kernel density estimation in Figure 7 demonstrates the soundness of our identified urban functional regions.

5.5. Discussion

In this section, we first discuss the construction of the urban functional corpus. Unlike the construction of the corpus based on the TAZ in the research of Yao et al. [15] and the construction based on tuples consisting of the central POIs and its K-nearest POIs in the research of Zhai et al. [31], we aim to make the urban functional corpus more continuous and closer in structure to the corpus composed of natural language while taking into account geographical proximity. Therefore, we construct the urban functional corpus in units of buffers centered on each POI. As mentioned in Section 3.2, the buffer-based method enhances the spatial context relationship. Furthermore, as the Gaode POIs dataset contains relatively complete POI types, it has practical significance to the extraction and identification of urban functional regions.

In addition, according to the confusion matrix displayed in Table 7, we note that the proposed method is extremely accurate for the extraction of public service regions and agriculture, forestry, animal husbandry, and fishing (AFAHF) regions. The results based on the proposed method and manual annotations exhibit spatial homogeneity in these two regions, mainly because they are mostly located in the suburbs of the study area where the POIs are relatively homogeneous and sparse and the spatial features are relatively easy to capture. Meanwhile, there is obvious spatial heterogeneity in residential areas, mainly due to the fact that there are fewer POIs in residential categories. More importantly, in most cases, residential regions are mixed with other types of urban functional regions such as public service regions and daily life service regions, which makes it difficult to extract their spatial features and results in large classification errors.

6. Conclusions and Future Work

We proposed a GloVe-based POI type embedding model that accounts for the co-occurrence statistics of POIs and spatial context while training POI type vectors. Using the proposed model, we extract and identify urban functional regions using POIs at the scale of traffic analysis zones. Through comparison against a baseline method based on word2vec, we proved that our method outperforms the word2vec-based method in terms of training speed and overall accuracy.

To extract and identify urban functional regions with very high accuracy, conventional data (e.g., geodemographic data, questionnaire data) may be more reliable. The goal of our study is to help city planners monitor regional development after initial urban planning, and we have achieved this by providing a new method to mine information hidden in POIs. Given their easy accessibility, POIs are a low cost and effective data source for quickly evaluating the development and changes in the functions of various regions in a city.

This study has some limitations. First, because each POI can only represent one point in an urban space, we could not account for its influence on nearby areas. In the future, other types of data could be merged into the methodology presented in this study to make the urban functional corpus more practical by assigning weights to POIs. Second, this study introduces TAZ as the basic unit of analysis and thus is inevitably prone to the modifiable areal unit problem due to the zoning effect. In the future, the impact of different urban space division schemes on the results could be considered. Finally, our accuracy evaluation process is based on the judgment of volunteers and is hence subjective, the multiple annotations for each region notwithstanding. In the future, more labelled data can be introduced to make the accuracy evaluation more rigorous and reliable, and the proposed method can be refined accordingly to make it even more effective in meeting the needs of urban planners and managers.

Author Contributions: Conceptualization, Sensen Wu; Formal analysis, Liuchang Xu; Funding acquisition, Zhen Yan; Supervision, Zhen Yan; Writing – original draft, Chengkun Zhang. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Key Research and Development Project of China (Grant No. 2018YFB0505003), the National Natural Science Foundation of China (Grant No. 41922043, 41871287 and 42001323), the Research and Development Foundation of Zhejiang A&F University (Grant No. 2020FR064), the Open Research Fund of Zhejiang Provincial Key Laboratory of Resources and Environmental Information System, Zhejiang University (Grant No. 2020330101004109).

Data Availability Statement: The original POI data and food takeout delivery data provided by Gaode Map and Dianwoda Company can be accessed at <https://lbs.amap.com> and <https://www.dianwoda.com> respectively. Restrictions apply to the availability of these data, which were used under license for this study.

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

References

1. Karlsson, C. Clusters, functional regions and cluster policies. *JIBS CESIS Electron. Work. Pap. Ser.* **2007**, *84*, 3.
2. Regan, C.M.; Bryan, B.A.; Connor, J.D.; Meyer, W.S.; Ostendorf, B.; Zhu, Z.; Bao, C. Real options analysis for land use management: Methods, application, and implications for policy. *J. Environ. Manag.* **2015**, *161*, 144–152.
3. Zheng, Y.; Capra, L.; Wolfson, O.; Yang, H. Urban computing: Concepts, methodologies, and applications. *ACM Trans. Intell. Syst. Technol. (TIST)* **2014**, *5*, 1–55.
4. Blaschke, T.; Hay, G.J.; Kelly, M.; Lang, S.; Hofmann, P.; Addink, E.; Feitosa, R.Q.; van der Meer, F.; van der Werff, H.; van Coillie, F.; et al. Geographic object-based image analysis—Towards a new paradigm. *ISPRS J. Photogramm.* **2014**, *87*, 180–191.
5. Zhang, X.; Du, S.; Wang, Y. Semantic classification of heterogeneous urban scenes using intrascene feature similarity and inter-scene semantic dependency. *IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens.* **2015**, *8*, 2005–2014.
6. Zhong, Y.; Zhu, Q.; Zhang, L. Scene classification based on the multifeature fusion probabilistic topic model for high spatial resolution remote sensing imagery. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 6207–6222.
7. Hao, P.; Cheang, W.; Chiang, J. Real-time event embedding for POI recommendation. *Neurocomputing* **2019**, *349*, 1–11.
8. Liu, J.; Han, K.; Chen, X.M.; Ong, G.P. Spatial-temporal inference of urban traffic emissions based on taxi trajectories and multi-source urban data. *Transp. Res. Part C Emerg. Technol.* **2019**, *106*, 145–165.
9. Qiao, Y.; Luo, X.; Li, C.; Tian, H.; Ma, J. Heterogeneous graph-based joint representation learning for users and POIs in location-based social network. *Inform. Process. Manag.* **2020**, *57*, 102151.
10. Vich, G.; Marquet, O.; Miralles-Guasch, C. Green exposure of walking routes and residential areas using smartphone tracking data and GIS in a Mediterranean city. *Urban For. Urban Green.* **2019**, *40*, 275–285.
11. Zhang, L.; Sun, Z.; Zhang, J.; Kloeden, H.; Klanner, F. Modeling hierarchical category transition for next POI recommendation with uncertain check-ins. *Inform. Sci.* **2020**, *515*, 169–190.
12. 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.
13. Jiang, S.; Alves, A.; Rodrigues, F.; Ferreira, J., Jr.; 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.
14. Yuan, J.; Zheng, Y.; Xie, X. Discovering regions of different functions in a city using human mobility and POIs. In Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Beijing China, 12–16 August 2012.
15. 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.
16. Pennington, J.; Socher, R.; Manning, C.D. Glove: Global vectors for word representation. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar, 25–29 October 2014.
17. Banzhaf, E.; Netzband, M. Monitoring urban land use changes with remote sensing techniques. *Appl. Urban Ecol. Global Framew.* **2011**, *2011*, 18–32.
18. Hu, S.; Wang, L. Automated urban land-use classification with remote sensing. *Int. J. Remote Sens.* **2013**, *34*, 790–803.
19. Liu, Y.; Liu, X.; Gao, S.; Gong, L.; Kang, C.; Zhi, Y.; Chi, G.; Shi, L. Social sensing: A new approach to understanding our socio-economic environments. *Ann. Assoc. Am. Geogr.* **2015**, *105*, 512–530.
20. Song, J.; Tong, X.; Wang, L.; Zhao, C.; Prishchepov, A.V. Monitoring finer-scale population density in urban functional zones: A remote sensing data fusion approach. *Landsc. Urban Plan* **2019**, *190*, 103580.
21. Blei, D.M.; Ng, A.Y.; Jordan, M.I. Latent dirichlet allocation. *J. Mach. Learn. Res.* **2003**, *3*, 993–1022.
22. Chen, W.; Er, M.J.; Wu, S. PCA and LDA in DCT domain. *Pattern Recogn. Lett.* **2005**, *26*, 2474–2482.
23. Mikolov, T.; Chen, K.; Corrado, G.; Dean, J. Efficient estimation of word representations in vector space. *arXiv* **2013**, arXiv:1301.3781.

24. Mikolov, T.; Sutskever, I.; Chen, K.; Corrado, G.; Dean, J. Distributed representations of words and phrases and their compositionality. *arXiv* **2013**, arXiv:1310.4546.
25. Yan, B.; Janowicz, K.; Mai, G.; Gao, S. From itdl to place2vec: Reasoning about place type similarity and relatedness by learning embeddings from augmented spatial contexts. In Proceedings of the 25th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, Redondo Beach, CA, USA, 7–10 November 2017.
26. 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.
27. Zhang, J.; Li, X.; Yao, Y.; Hong, Y.; He, J.; Jiang, Z.; Sun, J. The Traj2Vec model to quantify residents' spatial trajectories and estimate the proportions of urban land-use types. *Int. J. Geogr. Inf. Sci.* **2021**, *35*, 193–211.
28. Shoji, Y.; Takahashi, K.; Dürst, M.J.; Yamamoto, Y.; Ohshima, H. Location2vec: Generating distributed representation of location by using geo-tagged microblog posts. In Proceedings of the International Conference on Social Informatics, Saint-Petersburg, Russia, 25–28 September 2018.
29. McKenzie, G.; Adams, B. A data-driven approach to exploring similarities of tourist attractions through online reviews. *J. Locat. Based Serv.* **2018**, *12*, 94–118.
30. Xu, L.; Du, Z.; Mao, R.; Zhang, F.; Liu, R. GSAM: A deep neural network model for extracting computational representations of Chinese addresses fused with geospatial feature. *Comput. Environ. Urban Syst.* **2020**, *81*, 101473.
31. Zhai, W.; Bai, X.; Shi, Y.; Han, Y.; Peng, Z.; 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.
32. 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.
33. Wang, S.; Xu, G.; Guo, Q. Street centralities and land use intensities based on points of interest (POI) in Shenzhen, China. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 425.
34. Jeawak, S.S.; Jones, C.B.; Schockaert, S. Embedding geographic locations for modelling the natural environment using flickr tags and structured data. In Proceedings of the European Conference on Information Retrieval, Cologne, Germany, 14–18 April 2019.
35. Ng, H.T.; Zelle, J. Corpus-based approaches to semantic interpretation in NLP. *AI Mag.* **1997**, *18*, 45.
36. Bengio, Y.; Ducharme, R.; Vincent, P.; Janvin, C. A neural probabilistic language model. *J. Mach. Learn. Res.* **2003**, *3*, 1137–1155.
37. Socher, R.; Perelygin, A.; Wu, J.; Chuang, J.; Manning, C.D.; Ng, A.Y.; Potts, C. Recursive deep models for semantic compositionality over a sentiment treebank. In Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing, Seattle, WA, USA, 18–21 October 2013.
38. Mikolov, T.; Yih, W.; Zweig, G. Linguistic regularities in continuous space word representations. In Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Atlanta, GA, USA, 9 June 2013.
39. Skupin, A.; Fabrikant, S.I. Spatialization methods: A cartographic research agenda for non-geographic information visualization. *Cartogr. Geogr. Inf. Sci.* **2003**, *30*, 99–119.
40. Sen, S.; Swoap, A.B.; Li, Q.; Boatman, B.; Dippenaar, I.; Gold, R.; Ngo, M.; Pujol, S.; Jackson, B.; Hecht, B. Cartograph: Unlocking thematic cartography through semantic enhancement. In Proceedings of the 22nd International Conference on Intelligent User Interfaces, IUI, Limassol, Cyprus, 13–16 March 2017.
41. Hinton, G. Visualizing Data using t-SNE. *J. Mach. Learn. Res.* **2008**, *9*, 2579–2605.
42. Liu, K.; Yin, L.; Lu, F.; Mou, N. Visualizing and exploring POI configurations of urban regions on POI-type semantic space. *Cities* **2020**, *99*, 102610.
43. Arthur, D.; Vassilvitskii, S. K-Means++: The Advantages of Careful Seeding. Eighteenth Acm-siam Symposium on Discrete Algorithms, New Orleans, LA, USA, 7–9 January 2007.
44. Deng, Y.; Liu, J.; Liu, Y.; Luo, A. Detecting Urban Polycentric Structure from POI Data. *Int. J. Geo-Inf.* **2019**, *8*, 283.