# Understanding Place Characteristics in Geographic Contexts through Graph Convolutional Neural Networks

**Article** *in* Annals of the American Association of Geographers · January 2020

DOI: 10.1080/24694452.2019.1694403

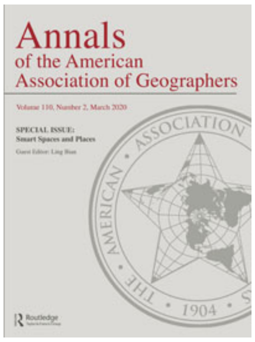


Di Zhu
University of Minnesota Twin Cities


Fan Zhang
Massachusetts Institute of Technology


Yaoli Wang
Peking University


Ximeng Cheng
Peking University




Taylor & Francis
Taylor & Francis Group

Annals of the American Association of Geographers

ISSN: 2469-4452 (Print) 2469-4460 (Online) Journal homepage: https://www.tandfonline.com/loi/raag21

# Understanding Place Characteristics in Geographic Contexts through Graph Convolutional Neural Networks

**Di Zhu, Fan Zhang, Shengyin Wang, Yaoli Wang, Ximeng Cheng, Zhou Huang & Yu Liu**

# Understanding Place Characteristics in Geographic Contexts through Graph Convolutional Neural Networks

Di Zhu,[*,†] Fan Zhang,[*,‡] Shengyin Wang,[*] Yaoli Wang,[*] Ximeng Cheng,[*] Zhou Huang,[*] and Yu Liu[*,†]

[*]Institute of Remote Sensing and Geographic Information Systems, Peking University
[†]Beijing Key Lab of Spatial Information Integration and Its Applications, Peking University
[‡]Senseable City Laboratory, Massachusetts Institute of Technology

Inferring the unknown properties of a place relies on both its observed attributes and the characteristics of the places to which it is connected. Because place characteristics are unstructured and the metrics for place connections can be diverse, it is challenging to incorporate them in a spatial prediction task where the results could be affected by how the neighborhoods are delineated and where the true relevance among places is hard to identify. To bridge the gap, we introduce graph convolutional neural networks (GCNNs) to model places as a graph, where each place is formalized as a node, place characteristics are encoded as node features, and place connections are represented as the edges. GCNNs capture the knowledge of the relevant geographic context by optimizing the weights among graph neural network layers. A case study was designed in the Beijing metropolitan area to predict the unobserved place characteristics based on the observed properties and specific place connections. A series of comparative experiments was conducted to highlight the influence of different place connection measures on the prediction accuracy and to evaluate the predictability across different characteristic dimensions. This research enlightens the promising future of GCNNs in formalizing places for geographic knowledge representation and reasoning. *Key Words: big geodata, graph convolutional neural networks, place characteristic, place connection, spatial prediction.*

推断一个地理位置的未知属性既取决于观测的属性，也取决于其连接对象位置的特征。由于位置特征无具体结构且位置连接的指标是多元的，如果空间预测任务的结果可能受相邻区域绘图方式的影响，很难识别不同位置真正的相关性，也就很难将这些位置特征纳入其中。为了弥补这一缺陷，我们引入图卷积神经网络(GCNN）将位置建模为图谱，将每个位置确认为一个节点，将位置特征编码为节点特征，并将位置连接显示为连接节点的边。GCNN 通过优化图神经网络层之间的权重，来获取相关地理环境的知识。我们在北京城市区域设计了一项案例研究，根据观测到的属性和特定地理位置的连接，预测未观测到的位置特征。我们通过一系列比较实验，揭示了不同位置连接测量对预测准确性的影响，评估不同特征维度的可预测性。本研究表明，GCNN 在地理知识表述与推理的位置确认方面具有广阔的应用前景。 *关键词: 大地理数据，图卷积神经网络，位置特征，位置连接，空间预测。*

Inferir las propiedades desconocidas de un lugar se fundamenta tanto en los atributos observables como en las características de los lugares con los cuales aquel está conectado. Debido a que las características del lugar no están estructuradas y las métricas de las conexiones del lugar pueden ser diversas, es todo un reto incorporarlas en una tarea de predicción espacial, donde los resultados podrían afectarse por el modo como están delineados los vecindarios y donde la verdadera relevancia que hay entre los lugares es difícil de establecer. Para zanjar esta dificultad, introducimos el gráfico de las redes neurales convolucionales (GCNNs) para modelar los lugares como un gráfico, donde cada lugar es formalizado como un nódulo, las características del lugar son codificadas como rasgos nodales y las conexiones del lugar se representan como los bordes. Los GCNNs capturan el conocimiento del contexto geográfico relevante optimizando los pesos entre las capas del gráfico de redes neurales. Se diseñó un estudio de caso en el área metropolitana de Beijing para predecir las características no observadas del lugar con base en las propiedades observadas y las conexiones específicas del lugar. Se condujo una serie de experimentos comparativos para destacar la influencia de las medidas de diferentes conexiones del lugar sobre la exactitud de la predicción, y para evaluar la predictibilidad a través de diferentes dimensiones de las características. Esta investigación ilumina el prometedor futuro de los GCNNs para formalizar lugares para la representación y razonamiento del conocimiento geográfico. *Palabras clave: big geodata, características del lugar, conexión del lugar, gráfico de las redes neurales convolucionales, predicción espacial.*

This article has been republished with minor changes. These changes do not impact the academic content of the article.
*Annals of the American Association of Geographers*, 110(2) 2020, pp. 408–420 © 2020 by American Association of Geographers
Initial submission, March 2019; revised submissions, May, July, September, and November 2019; final acceptance, November 2019
Color versions of one or more figures in the article can be found online at www.tandfonline.com/raag
Published by Taylor & Francis, LLC.

The geographical concept of *place* is often used as "a portion of space" (Agnew and Duncan 2014) within which people carry out habitual aspects of their lives, such as recreation, work, and sleep (Goodchild 2011). The perception of a place is a comprehensive integration of location names, emotional feelings, and other properties (Shamai 1991; Adams and McKenzie 2013). The term *place characteristics* encompasses a broad range of properties used to depict a place that are important in describing the uniqueness of a specific environment. By place characteristics, we are referring not only to the features such as place names and types but also to the socioeconomic properties, human activities, and perceptions that can be measured, such as liveliness, greenness, or transport convenience (Rattenbury and Naaman 2009; Zhang et al. 2018).

Because places are not isolated but are connected to each other in many ways (Nystuen and Dacey 1961; Gould 1991; Noronha and Goodchild 1992), the contextual information for a place (i.e., its connection to other places) is crucial to understand its characteristics. These connections link a set of places to a network that indicates the predefined geographic contexts for the places (Kwan 2007). We use *place connections* to represent the measures between places, which could be both physical and social, such as distance, adjacency, and spatial interactions. Intuitively, the predictability of a place characteristic should be higher when choosing more appropriate connection measures. For example, the connection between places via taxi origin–destination flows can help identify the land-use characteristics of the places (X. Liu et al. 2016).

The prediction of a place's unknown characteristic relies on both the place's observed characteristics and the characteristics of the places to which it is connected. Despite various ways of measuring the connections, the "true causally relevant" contexts can be very complex (Golledge 2002; Kwan 2012). The prediction could be affected by how the geographic context is defined. A.-X. Zhu et al. (2018) suggested measuring the similarity of the locations' geographic configurations in the covariate space as the connections used for spatial prediction. In most cases, however, the researcher cannot be certain that the connections used in the study are appropriate. Further, as various place characteristics can now be sensed from multisource user-generated big geodata (Y. Liu et al. 2015; Jenkins et al. 2016; MacEachren 2017), it is even more difficult to incorporate both high-dimensional and unstructured places' characteristics and diverse place connections in a spatial prediction model. Attempts to simplify the knowledge of relevant contexts as predefined mathematical functions cannot adequately model its complex nature. Methods such as geographically weighted regression use kernel functions to consider connections but only focus on nearby observations in space (Fotheringham, Brunsdon, and Charlton 2003), which could be arbitrary when long-range relevance is nonnegligible and would require much effort to be applied to a non-Euclidean situation (Lu et al. 2014).

Recently, there has been a surge of interest in graph convolutional neural networks (GCNNs) for learning graph-structured data where the range of connection varies (Bruna et al. 2014). To effectively process the connection information, GCNNs generally follow an aggregation scheme where each node aggregates characteristics of its neighbors to learn a deep representation of the contextual information (Defferrard, Bresson, and Vandergheynst 2016). This powerful technique is able to capture both the long-range and short-range relationships through its neural network weights. Clearly, it is suitable for modeling a graph of connected places.

This research introduces the use of GCNNs to model connected places where each place is represented as a node, place characteristics are the node features to be computed, and place connections are represented as the graph edges. The graph convolution can effectively learn from the graph structures and node features to understand the place characteristics in a geographic context. The objective of this study was to investigate the feasibility of incorporating place connections to predict place characteristics. In a case study of the Beijing metropolitan area, we took advantage of GCNNs in formalizing, reasoning, and understanding places. Three scenarios were designed to consider different connection types. A series of comparative experiments revealed the influence of place connections on predicting place characteristics.

## Methodology

### Building the Place-Based Graph

Assuming a set of places $P$ where each element $p_i$ means a place, a place characteristic $X$ can be represented as a feature vector $[x_1, x_2, \ldots]$ observed on $P$, where $x_i$ denotes the values for the $i$th dimension of $X$. A set $C$ includes various connection types between places, such as distance, topological adjacency, and spatial interactions. $c \in C$ refers to a certain connection metric. Then, a place-based graph $G = (V, E^{(c)})$ is constructed to connect places as a graph. Each place $p_i$ is formalized as a node $v_i \in V$ in G, and the place characteristic $X$ is encoded as the node feature $X_k \in X$ on every $v_k \in V$. The place connections in type $c$ are represented as a set of edges $E^{(c)}$, where $e_{ij} = (v_i, v_j, a_{ij}) \in E^{(c)}$ is the edge between $p_i$ and $p_j$ and $a_{ij}$ is the weight of the edge. Given $c$, $E^{(c)}$ represents the connection information in the place-based graph G.

The problem we addressed is illustrated in Figure 1. There are two kinds of place characteristics: $X$, which is easier to obtain (e.g., visual characteristic), and $Y$ (e.g., functional characteristic), which is more difficult to obtain but is vital to residents' daily lives, activities, and perceptions. Often, we had the information of both visual and functional characteristics for certain sampled places but only the visual characteristics for unsampled places. Using the information collected, we were able to predict the functional characteristics for unsampled places based on their observed visual characteristics and the characteristics of their connected places given $E^{(c)}$. As illustrated in Figure 1, $Y_1$, $Y_4$, and $Y_5$ were the place characteristics to be predicted for the unsampled places $p_1$, $p_4$, and $p_5$, respectively, but we only had the ground truth $Y_2$ and $Y_3$ for the sampled places $p_2$ and $p_3$, respectively.

## Predicting Place Characteristics Using GCNNs

The GCNN model was designed as a multilayer neural network structure to learn the node features and connection information of an input graph. By computing the differences between the model outputs and the expected output, a GCNN model iteratively updates its layer-wise neural network weights through a large amount of training and eventually approximates the predictability of place characteristics given certain place connections. The details of how to train a GCNN model can be found in the Appendix.

When using an $m$-layer GCNN model to predict the place characteristic $Y$ of unsampled places (i.e., $Y_1$, $Y_4$, and $Y_5$ in Figure 1), the place characteristic $X$ of both the sampled and unsampled places was used as the input node features, and then we selected a metric of place connections $E$ to build the place-based graph. For each training iteration, the weights in the GCNN were optimized using a back-propagation method by computing the prediction loss between the outputs on $p_2$, $p_3$ and the observed $Y_2$, $Y_3$. A fully trained GCNN model would generate a final output of node features that are the most probable values for $Y_1$, $Y_4$, and $Y_5$ given the place connections.

Because the results could be affected by how the geographic contexts are defined, GCNNs can learn from both graph structures and node features to capture the "relevant context" underlying $E^{(c)}$ and to facilitate the prediction of unobserved $Y$. GCNNs fit a neural network model that considers place characteristics and connections among all of the places to make the most credible prediction. The advantage of using GCNNs is threefold. First, all of the places are connected as a graph, so the connections among unsampled places can be explicitly modeled. Second, GCNNs do not require a predefined kernel function (Fotheringham, Yang, and Kang 2017), so they could

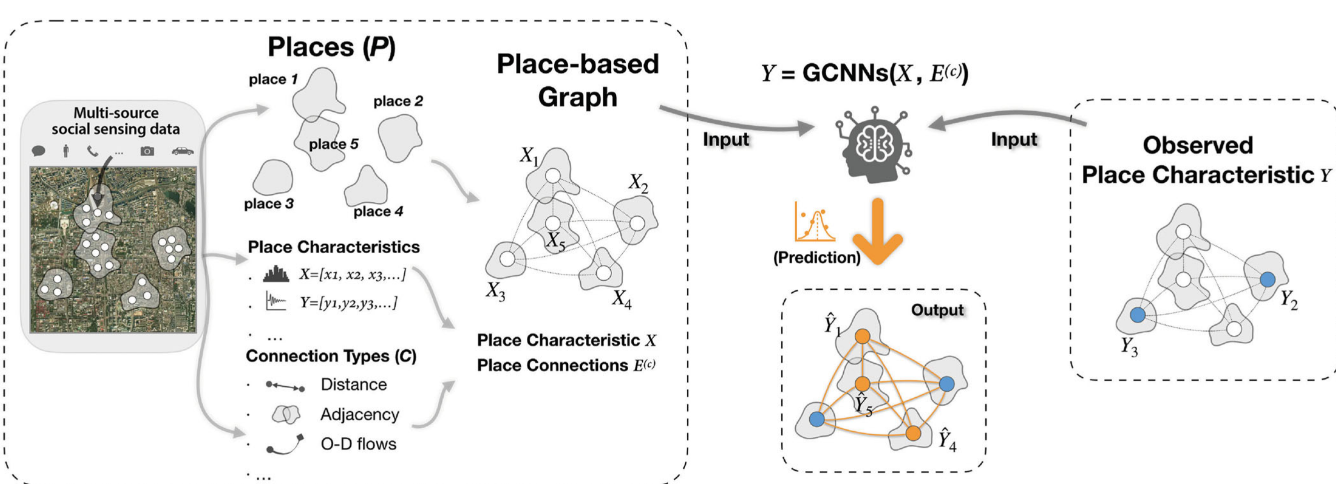



**Figure 1.** Building a place-based graph and using GCNNs to predict place characteristics. *Note:* O–D = origin–destination; GCNN = graph convolutional neural network..

learn both the long-range and short-range relationships in the place-based graph. Third, because a GCNN is designed as a nonparametric neural network, it does not assume any form of the spatial function to be fitted; for example, linear equations in most spatial regression models (Anselin 2010).

## Case Study

The study area for this research was Beijing, the capital of China. The fast-developing economy has made this city a giant metropolitan area where the flexible behaviors of its residents have led to diverse functional characteristics common in urban places regarding dining, residences, transportation, business, and so on. Understanding the heterogeneous pattern of place characteristics and how it is formed in the complex geographic context is key to enhancing the quality of life for residents and building a smart city.

A case study was designed to leverage GCNNs in predicting the unobserved characteristics for places based on their observed characteristics and connection to other places. The boundaries of 203 places within the study area were identified as the study units. Two types of characteristics were collected for the places and represented as feature vectors. One was the visual features as the input characteristic $X$; the other was the functional features as the place characteristic $Y$ to be predicted. Three scenarios of incorporating place connections—*self-only*, *adjacency*, and *spatial interaction*—were examined to build different place-based graphs. By randomly selecting some places as the sampled places and others as the unsampled places, experiments were conducted to evaluate (1) the influence of different place connection scenarios on the prediction accuracy and (2) how the predictability of $Y$ varies across different dimensions.

### Data Preparation

***Delineating Place Boundaries.*** To predict the characteristics of these places, a data set with 243,065 points of interest (POIs) was collected in 2016 from Baidu, a major source of location data. Each POI was labeled with a place name that reflected residents' common perceptions of the location. To identify the boundaries of places, we adopted a kernel density estimation method to compute the POI densities for each place name in space (Wang, Liu, and Chen 2018) and to depict the intuitive boundary of the place name. For each place name, the corresponding POI kernel densities were normalized into [0, 1] as the membership functions indicating to what extent an area belongs to this place (Gao et al. 2017; Wu et al. 2019). Then, a place's boundary was delineated based on a membership threshold.

We adopted a threshold of 0.5 to delineate both the core and peripheral area of the place names (Wang, Liu, and Chen 2018). Figure 2 shows the 203 places extracted within the study area. These delineated polygons cover the urban areas where names are broadly known and used in the locals' daily lives. We used these places as the study units in the following experiments.

***Quantifying Place Characteristics.*** Two kinds of place characteristics were collected. One is human perception of place locales (MacEachren 2017); that is, the visual characteristics. The other one is the land-use attributes of a place in terms of human activities (Lansley and Longley 2016); that is, the functional characteristics. These characteristics were treated as visual features and functional features in the GCNN model. Visual features and functional features characterized places from physical and social viewpoints, respectively. In this work, we used the visual features as the input to predict the functional features of places.

With regard to visual features, a total of 987,635 street view images taken in 2016 were collected from Tencent (Zhang et al. 2018). The images are photos describing the visual environment of the urban streets. We used a feature extractor (He et al. 2016) to derive place-based visual features from the images. The feature extractor has been proven efficient in scene perception modeling (Zhou et al. 2018). Each street view image was transformed into a 512-dimensional feature vector to represent human perceptions of a place. Visual features of a place were further obtained by taking the average of the vectors on all images within a place. Figure 3A shows that the visual features varied greatly for two places in the study area.

For functional features, we used a check-in data set collected from a social media platform named Sina Weibo. The data contained the annual number of check-ins for each POI. Within the study area, we obtained 4,099,016 check-in records for the 203


**Figure 2.** The 203 places in the Beijing metropolitan area. Some representative places are highlighted, including historic sites (A) the Summer Palace and (B) Tian'anmen; residential areas (C) Hepingli and (D) Shilipu; and commercial districts (E) Guomao and (F) Zhongguancun. For other maps in this article, we omit the north arrows and map scales for simplicity.

places. Check-in records with different activity labels were aggregated into seven functional characteristic types to represent seven different functional feature dimensions. Because the check-in intensities are heavy-tail distributed, we computed the logarithmic numbers of the check-in activities to the base 10 as the feature values to reduce the skewness of data. The seven functional feature dimensions included (along with their value range in parentheses) dining (0, 4.80], residence (0, 4.42], transport (0, 4.67], business (0, 4.72], recreation (0, 4.78], medical (0, 4.11], and outdoors (0, 4.24]. The spatial distribution of the functional features is plotted in Figure 3B.

***Measuring Place Connections.*** To test the effect of place connections in predicting place characteristics, three scenarios of incorporating place

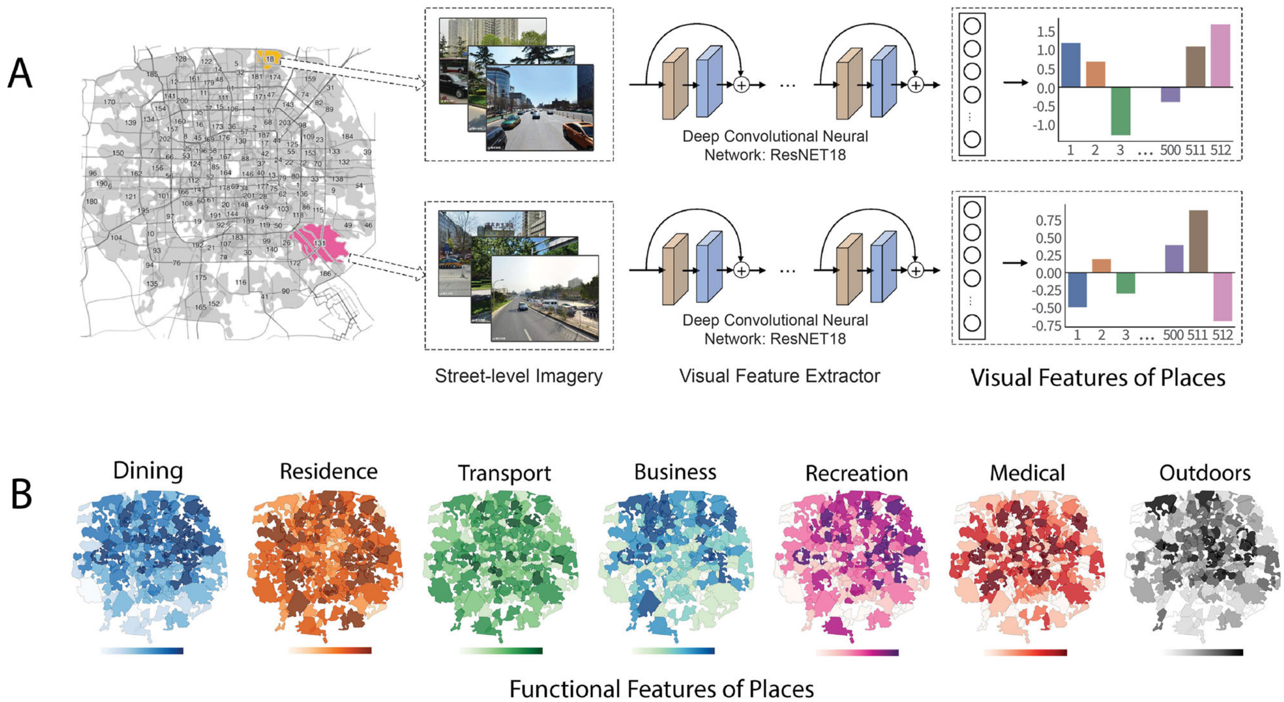



**Figure 3.** Place characteristics represented as feature values. (A) Visual features derived from street view images. (B) Seven functional feature dimensions sensed from social media check-in data; a darker hue denotes a higher value.

connections were examined to build different place-based graphs. The first scenario, self-only, considered no contextual information of a place, which means that we predicted the characteristics of the unsampled place without considering place connections. Self-only connection indicates that a place is connected only to itself. The second scenario, adjacency, incorporated connections in the prediction but was limited to only those connections between spatially adjacent places. The third scenario, spatial interaction, quantified both the short-range and long-range interaction volumes among places where the data were derived from a taxi origin–destination (O–D) data set (D. Zhu et al. 2017). Details about how these place connections were represented in matrix form can be found in the Appendix.

### A GCNN Model to Predict Places' Functional Features

The GCNN model required some places to be the training places, and others were the test places. The training places were those sampled with both visual features $X$ and functional features $Y$. The test places were places that have observed visual features $X$ but do not have observations for functional features $Y$. The task of predicting $Y$ for test places can be formulated as a semisupervised learning problem in the graph where we use the $X$ of both training and test places but the $Y$ of only training places to learn the place-based graph, as mentioned in the Methodology section.

Of the 203 places, we used 40 for training and 163 for testing. Figure 4A shows this initialization of the training and test places. A GCNN model was designed (Figure 4B) to predict the functional features for the test places. Seven submodels were developed, one for each of the seven functional feature dimensions. Each submodel was a two-layer GCNN that outputs the values for one functional dimension.

To report the prediction accuracy of a fully trained GCNN model, we used the lowest mean absolute percentage error (MAPE) on the test places. Let $N(T)$ be the total number of places in the test set $T$; the MAPE is calculated as $\frac{100\%}{N(T)}\sum_{i\in T}\frac{|\hat{Y}_i - Y_i|}{Y_i}$, where $\hat{Y}_i$ and $Y_i$ are the predicted and ground truth feature value of place $i$, respectively.

## Results and Discussion

### The Influence of Place Connection Types

Three scenarios of place connections—self-only, adjacency, and spatial interaction—were used in our

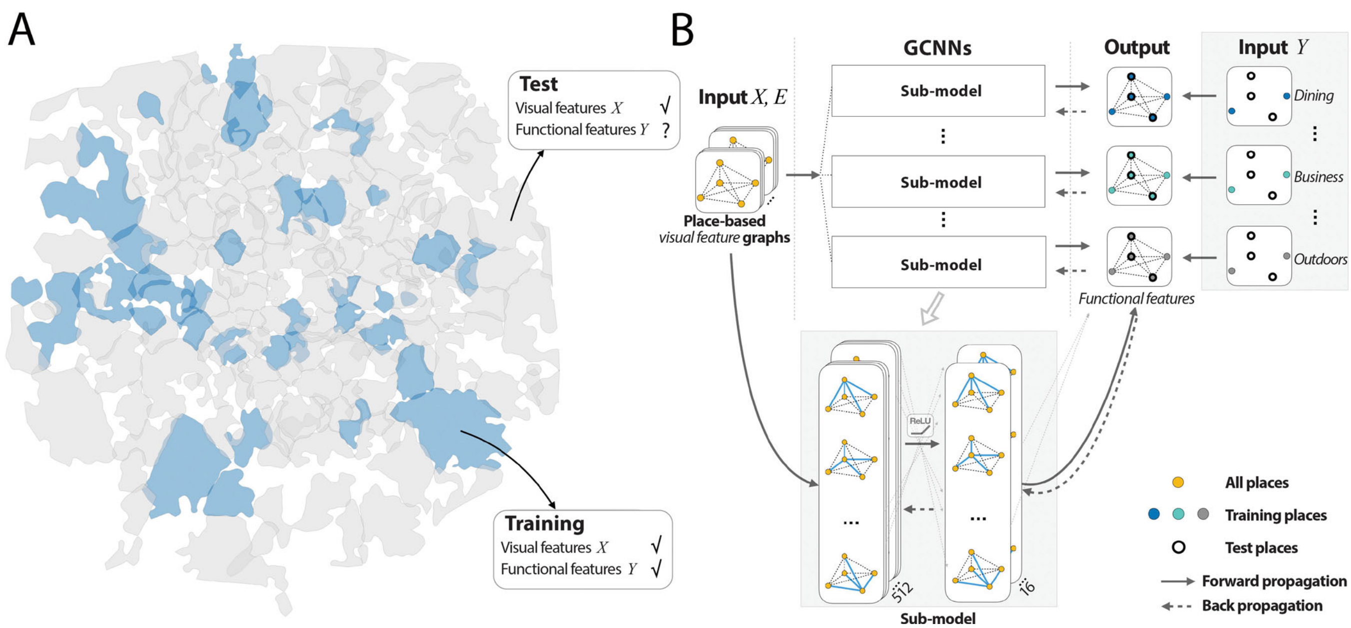



**Figure 4.** (A) An example of the training and test places initialization. (B) The GCNN model used to predict the seven functional feature dimensions for test places in a graph. *Note:* GCNN = graph convolutional neural network.

proposed GCNN model to predict the functional features of places. For each of the seven feature dimensions, place connections had a significant influence on the prediction accuracy. The GCNN model always achieved the best performance given the spatial interaction scenario. To highlight the influence of place connection metrics, we only report and discuss the results for the dining dimension in this section. Similar findings can be applied to the other six functional dimensions.

We simulated fifty different initializations of training and test places. The results of the fifty experiments on dining are plotted in Figure 5A. Self-only place connections show a median MAPE of 51.03 percent and a standard deviation of 0.0451; adjacency obtains the poorest median MAPE of 80.11 percent and a standard deviation of 0.1761; and the spatial interaction scenario displays the best overall accuracy with a median MAPE of only 16.24 percent and a standard deviation of 0.0148. Figure 5B shows the results for one of the fifty experiments. We found that neither self-only nor adjacency could achieve a predicted pattern similar to the real one, whereas for spatial interaction the GCNN achieved a predicted pattern strongly correlated with the real pattern.

The results of self-only indicated that using features alone cannot predict the functional features very well without considering the place connections. Because we did not consider any contextual information in the GCNN model in the self-only scenario, the prediction worked very similarly to a multivariate nonspatial regression. The fact that adjacency had the worst result indicates that it is unwise to use Euclidean metrics such as adjacency to predict the functional features, because adjacency cannot accurately describe the meaningful connections between places. Moreover, as spatial interaction can better reflect human activities and the actual connections among places, it outperformed the other two scenarios in predicting places' functional features.

### Variation across Functional Feature Dimensions

Because the performance of spatial interaction was better than the other two scenarios, the subsequent analysis focused on the spatial interaction scenario. When evaluating the prediction results of all seven functional feature dimensions, we conducted experiments based on multiple training ratios (from about 10 percent to 90 percent) of all of the places. Table 1 presents the results of the fifty parallel experiments. For each of the training ratios and feature dimensions, an average value of MAPE and a standard deviation of the fifty experiments were calculated for the test places to report the prediction accuracy and stability of the GCNN model, respectively.

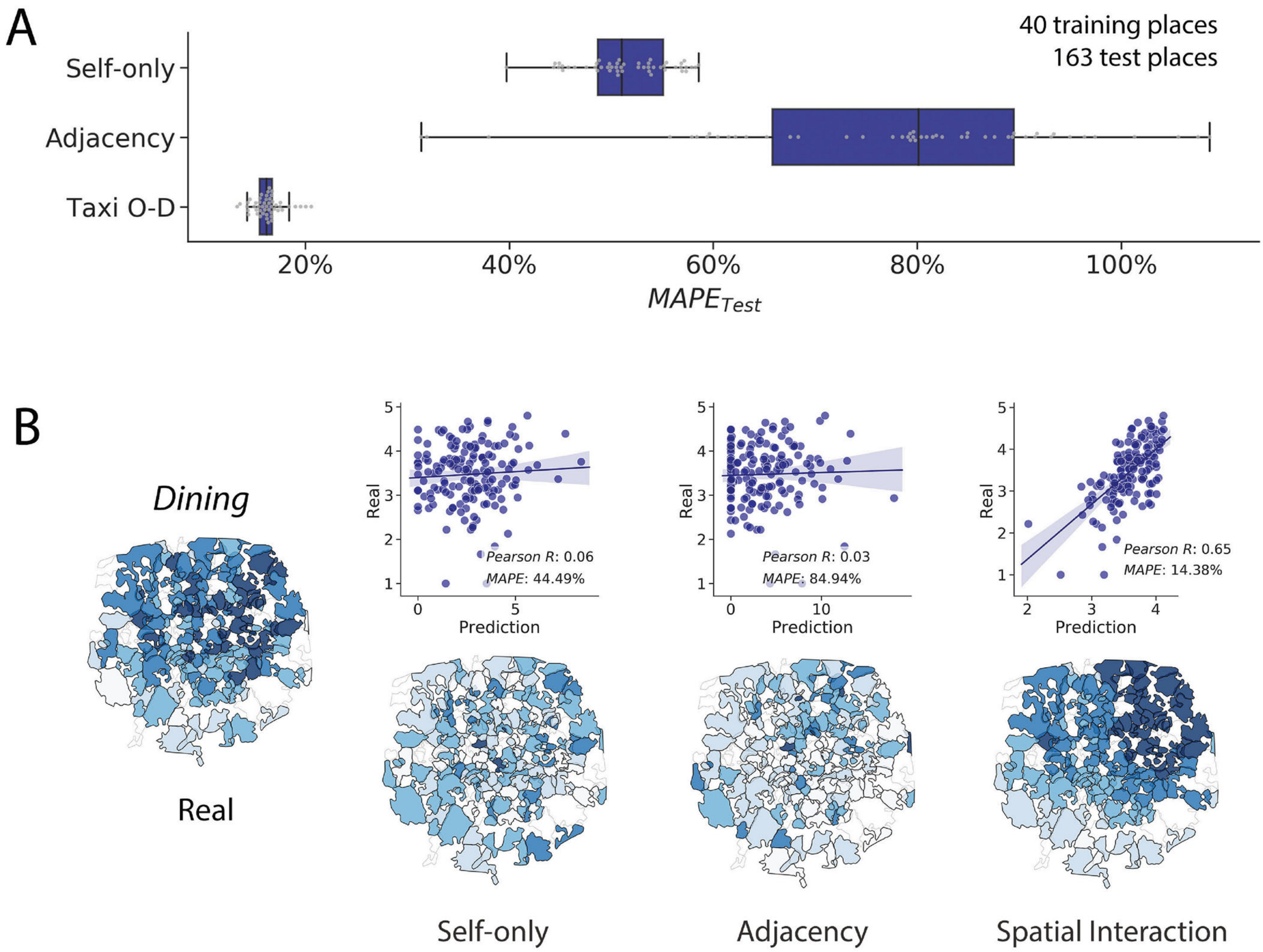



**Figure 5.** Different place connection scenarios resulted in different prediction accuracies. (A) The MAPE box plots of the fifty experiments on the dining dimension. (B) The predicted patterns on test places in one of the fifty experiments. Only the test places are colored to help interpret the model's performance. *Note:* O–D = origin–destination; MAPE = mean absolute percentage error.

According to the MAPEs in Table 1, dining and residence were better predicted than the other five dimensions, with average MAPEs lower than 20 percent. Medical was the most difficult dimension to predict and did not reach an average MAPE lower than 50 percent. We found that for a given training ratio, the accuracy remained stable, which means that the GCNN model can learn the knowledge for prediction even with only 10 percent observations. The standard deviation slightly rose when the training ratio was increased. This was because the test places of the fifty experiments were randomly selected but kept the same training ratio and the lesser test places led to a higher instability of the prediction. We checked the results of the best-fitted GCNN model under the 60 percent (120/203) training ratio for each functional feature dimension and visualized their corresponding patterns on test places in Figure 6. First, dining and residence achieved quite good predicted patterns compared to the real patterns, indicating that taxi O–D flows can help in understanding these two functional characteristics. Second, business, recreation, and transport showed medium-level predictabilities, which implies that using taxi O–D flows as the place connections is still not satisfactory when trying to predict these three functional characteristics. Third, outdoors and medical had inferior predicted patterns, meaning that the predictability of these two characteristics is very low even when taxi O–D flows are used. Because the functional characteristics of places are closely related to residents' activity patterns, some activities, such as hanging around in a park or going to a local hospital, might have very little relationship with taxis (Gong et al. 2016).

**Table 1.** The predictions of different functional feature dimensions on the test places

| # Train | # Test | Training ratio (%) | Average prediction accuracies of fifty parallel experiments in the spatial interaction scenario | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Dining | | Residence | | Transport | | Business | | Recreation | | Medical | | Outdoors | |
| | | | MAPE (%) | SD | MAPE (%) | SD | MAPE (%) | SD | MAPE (%) | SD | MAPE (%) | SD | MAPE (%) | SD | MAPE (%) | SD |
| 20 | 183 | 10 | 16.26 | 0.0157 | 15.94 | 0.0113 | 34.10 | 0.0214 | 33.32 | 0.0331 | 35.81 | 0.0270 | 51.24 | 0.0695 | 41.16 | 0.0430 |
| 40 | 163 | 20 | 16.31 | 0.0148 | 15.41 | 0.0146 | 34.72 | 0.0233 | 34.74 | 0.0417 | 36.56 | 0.0372 | 53.57 | 0.0636 | 42.18 | 0.0473 |
| 60 | 143 | 30 | 15.46 | 0.0171 | 15.42 | 0.0155 | 33.98 | 0.0252 | 34.42 | 0.0385 | 36.44 | 0.0320 | 52.32 | 0.0700 | 40.58 | 0.0305 |
| 80 | 123 | 40 | 15.32 | 0.0178 | 15.45 | 0.0209 | 33.72 | 0.0271 | 33.33 | 0.0288 | 35.72 | 0.0349 | 51.94 | 0.0664 | 41.56 | 0.0442 |
| 100 | 103 | 50 | 15.29 | 0.0225 | 14.98 | 0.0241 | 34.46 | 0.0312 | 33.97 | 0.0404 | 35.48 | 0.0351 | 51.88 | 0.0643 | 40.68 | 0.0406 |
| 120 | 83 | 60 | 15.08 | 0.0216 | 15.17 | 0.0311 | 34.76 | 0.0389 | 34.35 | 0.0395 | 35.13 | 0.0376 | 50.50 | 0.0677 | 40.73 | 0.0484 |
| 140 | 63 | 70 | 15.19 | 0.0273 | 14.99 | 0.0295 | 34.74 | 0.0474 | 33.62 | 0.0407 | 35.43 | 0.0511 | 53.58 | 0.0757 | 41.16 | 0.0586 |
| 160 | 43 | 80 | 15.57 | 0.0391 | 14.75 | 0.0446 | 33.76 | 0.0586 | 34.51 | 0.0587 | 36.79 | 0.0699 | 51.46 | 0.0850 | 41.41 | 0.0658 |
| 180 | 23 | 90 | 16.09 | 0.0590 | 14.62 | 0.0566 | 31.81 | 0.0858 | 33.61 | 0.0879 | 35.17 | 0.0962 | 50.84 | 0.1121 | 42.26 | 0.0942 |

*Note*: MAPE = mean absolute percentage error.

Additional data are needed to further investigate what kinds of spatial interactions are more appropriate in predicting these place characteristics.

As was implied in A.-X. Zhu et al. (2018) and Kwan (2012), the definition of explanatory covariates and geographic contexts could affect the findings on places. The predictability of a place's unknown characteristic could be affected by both the place's observed characteristics and the definition of its contextual places. Because our GCNN model was designed to capture the "relevant context" in its trainable neural network weights, the predictability of a task actually reflects to what extent a place is related to its connected places given a connection metric. The predictability could be higher when using suitable place connections and more informative explanatory characteristics, because the predictability is governed by the underlying relevance. For example, the results in Figure 5 reveal that the spatial interaction context is much more helpful in uncovering the relevance of places in terms of dining features than the adjacency context.

## Conclusions

Predicting a place's characteristics is challenging, as the place is related to the characteristics of other places, given a measure of place connections. Researchers cannot be certain whether the connections adopted are appropriate to reflect the "relevant" geographic context of places. In addition, as the place characteristics sensed from big geodata could be unstructured and the metric of place connections could be diverse, it is difficult to incorporate both of them in a spatial model. This research introduces the idea of using GCNNs to bridge the gap by effectively learning from the place characteristics and place connections through its trainable neural network weights and helping explore the predictability of a place characteristic in a geographic context.

The place characteristics and place connections are integrated into a place-based graph. Then, a GCNN model is used to make a credible prediction for unobserved place characteristics in a graph of connected places. The advantage of GCNNs is threefold. First, all places are connected as a graph, so that the connections among unsampled places can be explicitly modeled. Second, GCNNs can learn both the short-range and long-range relationships in the place-based graph to capture the complex relevance in a

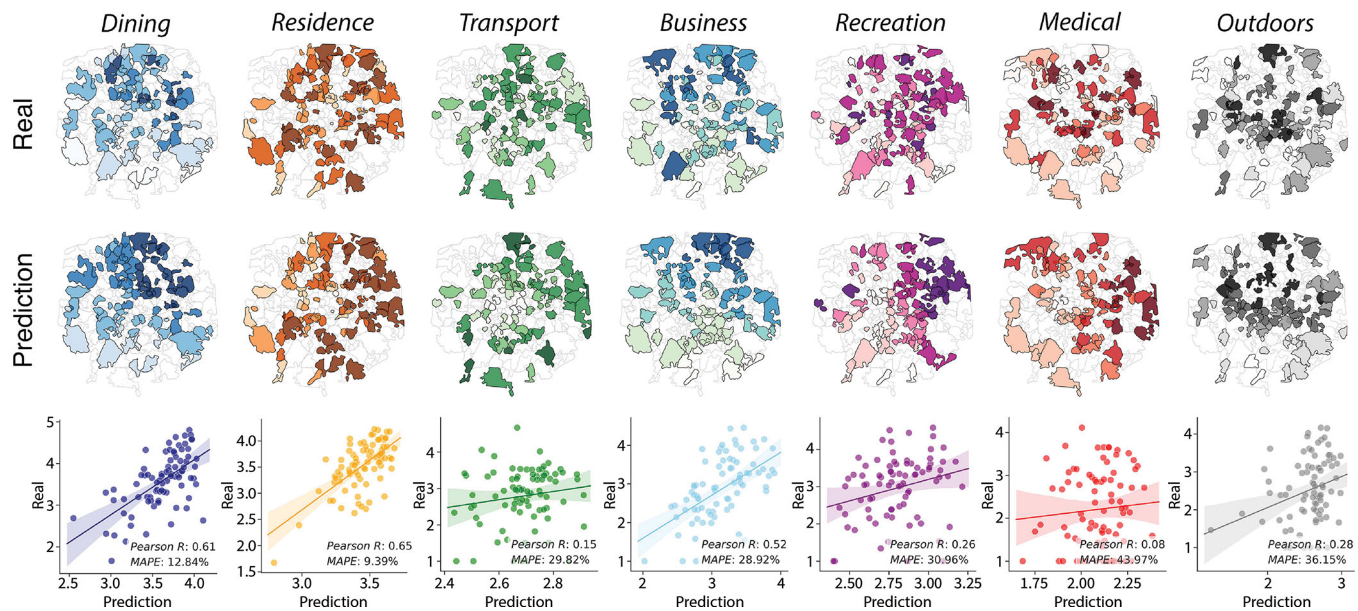



**Figure 6.** The best predicted pattern for each functional feature dimension using spatial interactions and a 60 percent training ratio. *Note:* MAPE = mean absolute percentage error.

geographic context. Third, the GCNN model is constructed as a nonparametric neural network and so does not assume any form of the spatial function to be fitted. An empirical study conducted in the Beijing metropolitan area showed that GCNNs are able to estimate the predictability of place characteristics in different geographic contexts.

Our contributions are as follows. First, we formalized places as a place-based graph to consider both place characteristics and place connections. Second, a semi-supervised GCNN model was proposed to predict a place's unobserved properties based on its observed properties and its contextual places. Third, the architecture of a GCNN model is flexible and powerful enough to encode both graph structure and node features, an ability that could be further exploited in future work. Moreover, with the recent advances in place-based knowledge graphs (Chen et al. 2018), introducing the knowledge of geographic contexts captured by black-box GCNN methods to such white-box methods could help uncover more geographic knowledge.

## Acknowledgments

Our deep and sincere thanks go to Dr. Ling Bian, Dr. A-Xing Zhu, and the anonymous reviewers for their constructive comments, which greatly improved the content and clarity of this article. Author Yu Liu served as corresponding author for this article.

## Funding

This research was supported by the National Key Research and Development Program of China (Grant 2017YFB0503602) and the National Natural Science Foundation of China (Grants 41625003, 41830645, 41771425, 41971331, and 41901321).

## ORCID

Di Zhu http://orcid.org/0000-0002-3237-6032
Fan Zhang http://orcid.org/0000-0002-3643-018X
Ximeng Cheng http://orcid.org/0000-0001-9923-7240
Zhou Huang http://orcid.org/0000-0002-1255-1913
Yu Liu http://orcid.org/0000-0002-0016-2902

## References

Adams, B., and G. McKenzie. 2013. Inferring thematic places from spatially referenced natural language descriptions. In *Crowdsourcing geographic knowledge*, ed. D. Sui, S. Elwood, and M. Goodchild, 201–21. Springer.

Agnew, J. A., and J. S. Duncan. 2014. *The power of place: Bringing together geographical and sociological imaginations*. London and New York: Routledge.

Anselin, L. 2010. Thirty years of spatial econometrics. *Papers in Regional Science* 89 (1):3–25. doi: 10.1111/j.1435-5957.2010.00279.x.

Bruna, J., Z. Wojciech, S. Arthur, and L. Yann. 2014. Spectral networks and locally connected networks on graphs. In *International Conference on Learning Representations (ICLR2014)*, ed. Y. Bengio and Y. Lecun. Banff, AB, Canada: Conference Track Proceedings. Accessed December 10, 2019. http://arxiv.org/abs/1312.6203.

Chen, H., M. Vasardani, S. Winter, and M. Tomko. 2018. A graph database model for knowledge extracted from place descriptions. *ISPRS International Journal of Geo-Information* 7 (6):221. doi: 10.3390/ijgi7060221.

Defferrard, M., X. Bresson, and P. Vandergheynst. 2016. Convolutional neural networks on graphs with fast localized spectral filtering. In *Advances in neural information processing systems*, ed. D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett, 3844–52. Barcelona: Annual Conference on Neural Information Processing Systems. Accessed December 10, 2019. http://papers.nips.cc/paper/6081-convolutional-neural-networks-on-graphs-with-fast-localized-spectral-filtering

Fan, R. K. C. 1997. *Spectral graph theory*. Providence, RI: American Mathematical Society.

Fotheringham, A. S., C. Brunsdon, and M. Charlton. 2003. *Geographically weighted regression: The analysis of spatially varying relationships*. Hoboken, NJ: Wiley.

Fotheringham, A. S., W. Yang, and W. Kang. 2017. Multiscale geographically weighted regression (MGWR). *Annals of the American Association of Geographers* 107 (6):1247–65. doi: 10.1080/24694452.2017.1352480.

Gao, S., K. Janowicz, D. R. Montello, Y. Hu, J. Yang, G. McKenzie, Y. Ju, L. Gong, B. Adams, and B. Yan. 2017. A data-synthesis-driven method for detecting and extracting vague cognitive regions. *International Journal of Geographical Information Science* 31 (6):1245–71. doi: 10.1080/13658816.2016.1273357.

Golledge, R. G. 2002. The nature of geographic knowledge. *Annals of the Association of American Geographers* 92 (1):1–14. doi: 10.1111/1467-8306.00276.

Gong, L., X. Liu, L. Wu, and Y. Liu. 2016. Inferring trip purposes and uncovering travel patterns from taxi trajectory data. *Cartography and Geographic Information Science* 43 (2):103–14. doi: 10.1080/15230406.2015.1014424.

Goodchild, M. F. 2011. Formalizing place in geographic information systems. In *Communities, neighborhoods, and health*, 21–33. Berlin: Springer.

Gould, P. 1991. Dynamic structures of geographic space. In *Collapsing space and time: Geographic aspects of communication and information*, ed. S. D. Brunn and T. R. Leinbach, 3–30. London: HarperCollins.

He, K., X. Zhang, S. Ren, and J. Sun. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, ed. Computer Vision Foundation, 770–78. Las Vegas: IEEE. Accessed December 10, 2019. https://www.cv-foundation.org/openaccess/content_cvpr_2016/papers/He_Deep_Residual_Learning_CVPR_2016_paper.pdf

Jenkins, A., A. Croitoru, A. T. Crooks, and A. Stefanidis. 2016. Crowdsourcing a collective sense of place. *PLoS One* 11 (4):e0152932. doi: 10.1371/journal.pone.0152932.

Kipf, T. N., and M. Welling. 2017. Semi-supervised classification with graph convolutional networks. *arXiv Preprint*. arXiv:1609.02907.

Kwan, M. P. 2007. Mobile communications, social networks, and urban travel: Hypertext as a new metaphor for conceptualizing spatial interaction. *The Professional Geographer* 59 (4):434–46. doi: 10.1111/j.1467-9272.2007.00633.x.

Kwan, M. P. 2012. The uncertain geographic context problem. *Annals of the Association of American Geographers* 102 (5):958–68. doi: 10.1080/00045608.2012.687349.

Lansley, G., and P. A. Longley. 2016. The geography of Twitter topics in London. *Computers, Environment and Urban Systems* 58:85–96. doi: 10.1016/j.compenvurbsys.2016.04.002.

LeCun, Y., Y. Bengio, and G. Hinton. 2015. Deep learning. *Nature* 521 (7553):436–44. doi: 10.1038/nature14539.

Liu, X., C. Kang, L. Gong, and Y. Liu. 2016. Incorporating spatial interaction patterns in classifying and understanding urban land use. *International Journal of Geographical Information Science* 30 (2):334–50. doi: 10.1080/13658816.2015.1086923.

Liu, Y., X. Liu, S. Gao, L. Gong, C. Kang, Y. Zhi, G. Chi, and L. Shi. 2015. Social sensing: A new approach to understanding our socioeconomic environments. *Annals of the Association of American Geographers* 105 (3):512–30. doi: 10.1080/00045608.2015.1018773.

Lu, B., M. Charlton, P. Harris, and A. S. Fotheringham. 2014. Geographically weighted regression with a non-Euclidean distance metric: A case study using hedonic house price data. *International Journal of Geographical Information Science* 28 (4):660–81. doi: 10.1080/13658816.2013.865739.

MacEachren, A. M. 2017. Leveraging big (geo) data with (geo) visual analytics: Place as the next frontier. In *Spatial data handling in big data era*, ed. C. Zhou, F. Su, F. Harvey and J. Xu, 139–55. Berlin: Springer.

Noronha, V. T., and M. F. Goodchild. 1992. Modeling interregional interaction: Implications for defining functional regions. *Annals of the Association of American Geographers* 82 (1):86–102. doi: 10.1111/j.1467-8306.1992.tb01899.x.

Nystuen, J. D., and M. F. Dacey. 1961. A graph theory interpretation of nodal regions. *Papers of the Regional Science Association* 7 (1):29–42. doi: 10.1007/BF01969070.

Rattenbury, T., and M. Naaman. 2009. Methods for extracting place semantics from Flickr tags. *ACM Transactions on the Web* 3 (1):1. doi: 10.1145/1462148.1462149.

Shamai, S. 1991. Sense of place: An empirical measurement. *Geoforum* 22 (3):347–58. doi: 10.1016/0016-7185(91)90017-K.

Wang, S., Y. Liu, and Z. Chen. 2018. Representing multiple urban places' footprints from dianping.com data. *Acta Geodaetica et Cartographica Sinica* 47 (8):1105–13.

Wu, X., J. Wang, L. Shi, Y. Gao, and Y. Liu. 2019. A fuzzy formal concept analysis-based approach to uncovering spatial hierarchies among vague places extracted from user-generated data. *International Journal of Geographical Information Science* 33 (5):991–1016. doi: 10.1080/13658816.2019.1566550.

Zhang, F., B. Zhou, L. Liu, Y. Liu, H. H. Fung, H. Lin, and C. Ratti. 2018. Measuring human perceptions of a large-scale urban region using machine learning. *Landscape and Urban Planning* 180:148–60. doi: 10.1016/j.landurbplan.2018.08.020.

Zhou, B., A. Lapedriza, A. Khosla, A. Oliva, and A. Torralba. 2018. Places: A 10 million image database for scene recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 40 (6):1452–64. doi: 10.1109/TPAMI.2017.2723009.

Zhu, A.-X., G. Lu, J. Liu, C.-Z. Qin, and C. Zhou. 2018. Spatial prediction based on Third Law of Geography. *Annals of GIS* 24 (4):225–40. doi: 10.1080/19475683.2018.1534890.

Zhu, D., and Y. Liu. 2018a. Modelling irregular spatial patterns using graph convolutional neural networks. *arXiv Preprint*. arXiv:1808.09802.

Zhu, D., and Y. Liu. 2018b. Modelling spatial patterns using graph convolutional networks. In *10th International Conference on Geographic Information Science (GIScience2018)*. Germany: Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik. doi: 10.4230/LIPIcs.GISCIENCE.2018.73

Zhu, D., N. Wang, L. Wu, and Y. Liu. 2017. Street as a big geo-data assembly and analysis unit in urban studies: A case study using Beijing taxi data. *Applied Geography* 86:152–64. doi: 10.1016/j.apgeog.2017.07.001.

DI ZHU is a PhD Candidate in GIScience at the Institute of Remote Sensing and Geographic Information System, School of Earth and Space Sciences, Peking University, Beijing 100871, China. E-mail: patrick.zhu@pku.edu.cn. His research interests include geospatial modeling, applied artificial intelligence, and urban social sensing.

FAN ZHANG received his PhD from The Chinese University of Hong Kong in 2017. E-mail: zhangfan@link.cuhk.edu.hk. His research interests are in GIScience, spatial data mining, and urban studies.

SHENGYIN WANG is a PhD student at Peking University, Beijing 100871, China. E-mail: sywang092@pku.edu.cn. Her research interests concentrate in GIScience and place formalization.

YAOLI WANG is a Postdoctoral Researcher at Peking University, Beijing 100871, China. E-mail: wangyaoli@pku.edu.cn. Her expertise is in transportation network analysis and optimization.

XIMENG CHENG is a PhD Candidate in GIS at the Institute of Remote Sensing and GIS, Peking University, Beijing 100871, China. E-mail: chengximeng@pku.edu.cn. His research interests are in spatiotemporal data mining, social sensing, machine learning, and deep learning.

ZHOU HUANG is an Associate Professor in the Institute of Remote Sensing and GIS, School of Earth and Space Sciences, Peking University, Beijing 100871, China. E-mail: huangzhou@pku.edu.cn. His research interests are geospatial data mining and high-performance geocomputing.

YU LIU is Professor of GIScience in the Institute of Remote Sensing and Geographical Information Systems, School of Earth and Space Sciences, Peking University, Beijing 100871, China. E-mail: liuyu@urban.pku.edu.cn. His research concentration is in GIScience and big geodata.

## Appendix

The use of traditional convolutional neural networks (CNNs; LeCun, Bengio, and Hinton 2015) could be problematic when the data are not structured in the regular spatial domain (Defferrard, Bresson, and Vandergheynst 2016). GCNNs are facilitated by the demand of generalizing well-established CNNs to the irregular spatial domain. Graph Fourier transform is needed to transform a place-based graph into the spectral domain (Fan 1997), such that a graph convolutional filter can be defined to learn the connection information among places and support the prediction of place characteristics.

Given a place-based graph $G = (V, E)$ with $n$ places, let $\tilde{X}$ be the node feature matrix of place characteristic $X$ and $\tilde{E} \in \mathbb{R}^{n \times n}$ be the matrix form of place connections' intensities, a general form of forward propagation $h(\cdot)$ between the $l$th and $(l+1)$ th hidden layer in a GCNN is

$$\tilde{X}^{l+1} = h(\tilde{X}^l, \tilde{E}) = \sigma\left(\tilde{D}^{-\frac{1}{2}}\tilde{E}\tilde{D}^{-\frac{1}{2}}\tilde{X}^l W^l\right), \tag{1}$$

where $\tilde{D}$ is the diagonal degree matrix with $\tilde{D}_{ii} = \sum_j \tilde{E}_{ij}$, $W^l$ is the layer-wise weights in neural networks, and $\sigma(\cdot)$ is a certain kind of activation function, for which we choose ReLU activation in this work to facilitate the nonlinear function approximation.

The normalized Laplacian matrix $\tilde{D}^{-\frac{1}{2}}\tilde{E}\tilde{D}^{-\frac{1}{2}}$ contains the preset connection information among places, and the trainable weights $W^l$ enable GCNNs to approximate the predictability of characteristics in the geographic context defined by $\tilde{E}$. Further explanations of Equation 1 can be found in Defferrard, Bresson, and Vandergheynst (2016), Kipf and Welling (2017), and D. Zhu and Liu (2018), where a Chebyshev polynomial was suggested to simplify and compute the weights in GCNNs. We do not include it here, as it is outside the scope of this article, but we do apply similar methods and design a trainable GCNN model in the study.

In the case study, three scenarios of place connections, self-only, adjacency, and spatial interactions, were examined as the $\tilde{E}$. In the self-only scenario, connections were represented as a diagonal matrix, with all diagonal elements equal to one. In the adjacency scenario, connections were represented as a binary matrix that contained the basic adjacency information of places; that is, adjacent places were connected and the others were not. In the spatial interactions scenario, the connections were represented as a dense symmetric matrix with each element indicating the number of taxi O–D flows for a pair of places.