

## Graph convolutional networks for street network analysis with a case study of urban polycentricity in Chinese cities

Ding Ma, Fangning He, Yang Yue, Renzhong Guo, Tianhong Zhao & Mingshu Wang

**To cite this article:** Ding Ma, Fangning He, Yang Yue, Renzhong Guo, Tianhong Zhao & Mingshu Wang (29 Feb 2024): Graph convolutional networks for street network analysis with a case study of urban polycentricity in Chinese cities, International Journal of Geographical Information Science, DOI: [10.1080/13658816.2024.2321229](https://doi.org/10.1080/13658816.2024.2321229)

**To link to this article:** <https://doi.org/10.1080/13658816.2024.2321229>



© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 29 Feb 2024.



Submit your article to this journal [↗](#)



Article views: 381



View related articles [↗](#)



View Crossmark data [↗](#)



RESEARCH ARTICLE



OPEN ACCESS



# Graph convolutional networks for street network analysis with a case study of urban polycentricity in Chinese cities

Ding Ma<sup>a</sup> , Fangning He<sup>b</sup>, Yang Yue<sup>a</sup>, Renzhong Guo<sup>a</sup>, Tianhong Zhao<sup>c</sup> and Mingshu Wang<sup>d</sup>

<sup>a</sup>Research Institute for Smart Cities, School of Architecture and Urban Planning, Shenzhen University and Guangdong-Hong Kong-Macau Joint Laboratory for Smart Cities, Key Laboratory of Urban Land Resources Monitoring and Simulation, Ministry of Natural Resources, Shenzhen, PR China; <sup>b</sup>Lyles School of Civil Engineering, Purdue University, West Lafayette, IN, USA; <sup>c</sup>College of Big Data and Internet, Shenzhen Technology University, Shenzhen, PR China; <sup>d</sup>School of Geographical and Earth Sciences, University of Glasgow, Glasgow, UK

## ABSTRACT

Graph theory effectively explains urban structures *via* street–street connectivity. However, systematic comparisons of street structures across cities remain challenging. This study employs graph convolutional networks (GCNs) to analyze street network structures. A two-branch GCN was used as the backbone to extract comparable features among street networks. The proposed approach was used to examine the structures of different urban road networks in a case study of polycentricity prediction across 298 Chinese cities. The model transformed approximately 4.5-million street segments into natural streets to create urban street graphs, which were subsequently analyzed to extract local and global embeddings. The extracted embeddings – with a portion labeled with a known urban polycentricity score – were used to predict the score for each city through a single-layer perceptron (SLP) model. Our results show consistency between the predicted polycentricity scores based on the derived street embeddings and those based on the population. Thus, the proposed GCN-based method can effectively predict the complexity and interconnection of street networks in different cities. This innovative integration of GCNs into urban studies demonstrates that deep learning techniques can analyze and comprehend the intricate patterns of street networks on a large scale.

## ARTICLE HISTORY

Received 31 August 2022

Accepted 16 February 2024

## KEYWORDS

Graph convolutional network; deep learning; street network; urban structure

## 1. Introduction

Streets enable the movement of people, goods and information, as well as shape the infrastructure layout and land use patterns, thereby realizing various urban functions. The intersections and connections of streets form a ‘network.’ A well-developed street

**CONTACT** Fangning He [fangninghe@alumni.purdue.edu](mailto:fangninghe@alumni.purdue.edu); Mingshu Wang [Mingshu.Wang@glasgow.ac.uk](mailto:Mingshu.Wang@glasgow.ac.uk)

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

network supports economic growth, fosters social cohesion and facilitates emergency responses during crises (Sharifi 2019; Wang et al. 2019). Studying street networks is imperative for creating sustainable and livable spaces that cater to the needs of individuals and communities. Accordingly, various street network representations have been devised to scrutinize spatial configurations (Penn et al. 1998; Marshall et al. 2018), which establish a direct link between physical space and people.

With the increasing popularity of graph theory and complex networks, the network structure of urban streets from a topological perspective (street–street connectivity) has been of primary interest to scholars from various disciplines through network centrality-based measures such as betweenness (Crucitti et al. 2006; Kirkley et al. 2018). However, comparing these network centrality parameters for individual streets across cities is difficult because no geometrically or statistically uniform layout for each urban street network exists (Xue et al. 2022). Deep learning techniques have recently exhibited their ability to bridge this gap and have attracted significant attention from researchers in both, the computer science and geospatial communities (Zhang et al. 2019). Graph convolutional networks (GCNs) and graph autoencoders (GAEs) can learn deep representations of street–street connections, employ graph convolutional layers to embed the graph into a latent vector space, and map the derived embeddings to a reconstruction of the input street network. The learned embedding from street networks can be viewed as a specific type of representation learning (Bengio et al. 2013) in urban studies that focuses on mapping urban configurations to continuous vector representations that can be nonlinearly fitted with related urban metrics. However, deriving embeddings from large-scale street networks that can balance high dimensionality, structural integrity and computational efficiency is challenging (Wang et al. 2016).

In this study, we propose a novel semi-supervised GCN framework that is designed to characterize urban structures across 298 prefectural-level Chinese cities for a comprehensive comparative analysis of large-scale urban street networks. Our framework leverages a two-branch graph neural network (GNN) structure. The first branch is based on an unsupervised autoencoder model that is used to learn the intrinsic patterns and embeddings from the input graph data. This autoencoder model enables us to explore the underlying structures of urban streets without requiring extensively labeled data. The second branch of the model incorporates embeddings from the first branch and is trained using a subset of labeled data to predict the morphological indices that fit the requirements of different urban analysis purposes. In addition, to ensure satisfactory performance and scalability, we rely on urban street graphs constructed from approximately 4.5 million street segments and employ a strategic subsampling approach to preserve the global graph structure and significantly reduce the computational load created by large street graphs. To showcase the feasibility of our framework, we conducted a case study predicting the index of urban polycentricity, i.e. a nuanced structural property describing the presence of multiple centers within a city (Meijers 2008), using embeddings extracted from street networks.

This study contributes to the literature in two key areas. Methodologically, the developed GCN-based framework highlights the potential of GNNs for street network analysis. It represents a significant advancement over traditional approaches by offering the unique capability to adaptively learn graph embeddings suitable for diverse tasks, all

without altering the fundamental architecture of the network. Furthermore, this novel framework design addresses the challenge of preserving the structure and handling nonlinearity in conventional street network analysis approaches, e.g. centrality-based analysis. Empirically, our case study of polycentricity prediction not only effectively demonstrated the efficacy of the framework but also illuminated the intricate relationship between the structure of street networks and urban polycentricity as informed by population. The predicted polycentric structures of Chinese cities revealed by our approach are potentially useful for policymakers because they enable a detailed comparison of the current urban landscape, facilitate the simulation of future developments and guide decisions based on existing or planned street configurations.

The remainder of this article is organized as follows: [Section 2](#) reviews the related literature. The proposed methodological framework is introduced in [Section 3](#). [Section 4](#) presents the results, with a focus on the subsampled street graphs for 298 cities, the GCN model performance and the correlation between the measured and learned urban polycentricity scores. The proposed methods, derived results, and implications for urban studies are discussed in [Section 5](#). [Section 6](#) provides the conclusions and outlines future research directions.

## 2. Related work

### 2.1. Complex structure of urban street networks

A city is not a tree but a complex network that embodies many interconnected urban elements (Alexander 1965). A typical example is a network of urban streets that can be naturally modeled as graph representations. Dual graphs (with the nodes as streets and links as intersections; Porta et al. 2006) and their related centrality measures are the most common methods for street structural configurational analysis (Crucitti et al. 2006; Kirkley et al. 2018; Wang and Debbage 2021; Wu et al. 2024). Empirical studies have demonstrated that urban street networks tend to possess a similar structure to that of many other real-world complex networks (e.g. scale-free and/or small-world properties) in terms of the statistical distribution (e.g. long-tailed or power-law) of graph measures (Batty 2008; Boeing 2017). In addition to the statistics for a single quantity, previous studies have identified both nonlinear and linear relationships among the street structural complexity, urban socioeconomic and environmental factors (such as GDP, population and CO<sub>2</sub> emissions) (Lu et al. 2016; Lan et al. 2019; Zhang et al. 2022) and urban life (Huang et al. 2022). Critics in the space syntax literature (e.g. Ma et al. 2019) have noted that the complex structure of street networks is hardly viewed in terms of geometric primitives, such as individual street segments or junctions that are conventionally adopted in geographical information system (GIS) models because they have a similar number of neighboring nodes or links. Instead, the network is formed by natural streets, each of which denotes a group of adjacent segments with strong continuity (see details in [Section 3.2](#)). This enables the complex network structure to be grasped, which further aids in deriving street hierarchies that are useful for many applications, such as traffic monitoring (Liu et al. 2019) and map generalization (Yu et al. 2020).

## 2.2. GCNs and their applications in urban studies

Among the different deep learning architectures, convolutional neural networks (CNNs) are the most commonly used for analyzing regularly structured data such as images (Defferrard et al. 2016). However, the use of these networks in the analysis of irregularly structured data, such as graphs remains challenging (Bronstein et al. 2017). Motivated by CNNs, two potential solutions, namely, spectral and spatial approaches have been investigated. The conceptual basis of the spectral-based approach is the graph Fourier transform, which resembles the Fourier transform of a 1D signal (Zhu and Rabbat 2012). In contrast, the spatial approach performs convolution operations directly on the graph vertices by aggregating the information from neighboring nodes in a convolutional manner. The spectral and spatial GCN approaches have both been adopted to address various types of graph-structured spatial vector data owing to their impressive ability to learn discriminative features from input graph-structured data. The potential of GCNs for characterizing urban spatial patterns based on regular and irregular geometric configurations has been demonstrated.

In urban studies, spatial vector graphs can be constructed from connections that link different types of spatial units, such as equally partitioned grids, self-defined boundaries based on data clusters, and individual spatial objects. For example, Yao et al. (2021) predicted the spatial flow distribution among  $1 \text{ km} \times 1 \text{ km}$  grids, Zhu et al. (2020) connected a set of point-of-interest (POI) clusters and applied a GCN to infer place characteristics, and Yan et al. (2019, 2021) proposed a GCN approach for the pattern cognition and classification of a group of individual buildings. In addition to networks of point or polygonal units, networks of linear units (i.e. road networks) have been used extensively for urban applications such as traffic forecasting (Yu et al. 2018). However, adapting emerging complex network structures for a comprehensive understanding of the spatial configurations among various cities remains a challenge. Xue et al. (2022) predicted the spatial homogeneity using segment-based centrality measures from a grid-level road network across 30 cities and used this GCN-based measure to determine urban structure and socio-economic performance. Despite this remarkable effort to link street structures with urban performance, it is worth discussing whether parts of streets in grids are sufficiently representative because no visual or statistical direct self-similarity proof between grid-level street segments and entire streets is available in the literature.

Another approach is to effectively sample the street network for each city to boost computational efficiency. Pooling or downsampling is commonly applied to grid-like image data in CNNs. Atwood and Towsley (2016) pioneered pooling with diffusion-based convolutions in GCNs, demonstrating the significance of pooling *via* graph coarsening for merging similar nodes. Subsequently, Ying et al. (2018) introduced a differentiable pooling operation for GCNs, which enables a hierarchical reduction in the graph size while preserving crucial data structures. Building on this concept, Lee et al. (2019) proposed a self-attention mechanism for graph pooling to identify and preserve the most salient nodes within a graph. However, whether the pooling strategy of existing methods can be directly applied to street data remains uncertain owing to the uniquely complex geometric and topological characteristics of such data.

Existing approaches usually require annotated datasets to train models. However, collecting dense semantic labels is usually very labor-intensive and time-consuming,

and the quality and adequacy of the labeled dataset significantly affect the performance of the learned model. Several semi-supervised and unsupervised approaches have been proposed to address the demand for strong supervision. GAEs have been used extensively to learn the efficient coding of unlabeled graph data. Kipf and Welling (2016a, 2016b) introduced a GAE with a GCN encoder and a simple inner product decoder, which exhibited competitive link prediction performance in several citation network datasets compared to other nondeep learning models. Inspired by the above research efforts, we propose a novel semi-supervised architecture to measure urban polycentricity from road networks using the GAE and single-layer perceptron (SLP) models.

### 2.3. Urban polycentricity

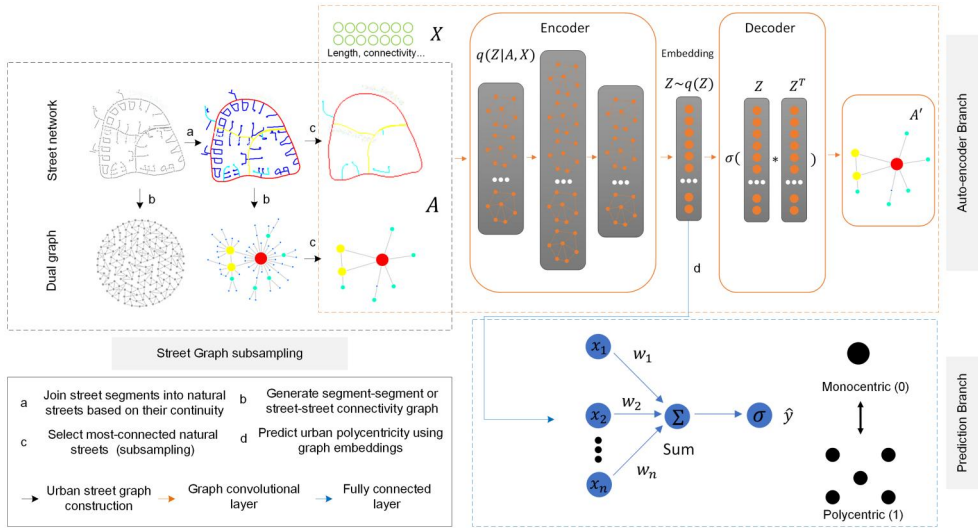
Over the past several decades, rapid urbanization has fostered the development of polycentric urban regions (PURs), which have been increasingly recognized in recent studies (Derudder et al. 2021; Harrison et al. 2022; Thomas et al. 2022; Derudder et al. 2022). The concept of urban polycentricity is often employed to assess the extent of such development quantitatively. It is a multifaceted concept that encompasses at least two dimensions: geometrical aspects (i.e. the morphological form) and topological relationships (i.e. functional links), as described by Burger and Meijers (2012). However, most of the existing related research has focused mainly on the morphological dimension of urban polycentricity. This approach typically involves examining the distribution and intensity of urban elements and activities, as highlighted in a recent study by Thomas et al. (2022).

Gridded data, particularly population grid data and nighttime imagery, have historically been the primary tools for assessing PURs. For example, Liu and Wang (2016) and Liu et al. (2018) used LandScan population grids to capture a holistic view of polycentric development in Chinese cities and their surrounding regions. In addition, location-based social media data have emerged as a novel source for analyzing urban polycentricity. Lv et al. (2021) demonstrated this application by performing a multi-scale PUR analysis in Chinese cities, incorporating data on POIs at the city level and check-in densities of street blocks within city centers.

Although numerous studies have leveraged various geospatial data to deduce aspects of urban polycentricity (e.g. Liu and Wang 2016; Wang 2021; Taubenböck et al. 2017; Liu et al. 2018; Volkmann and Münter 2022), a notable gap remains in the literature regarding the use of large-scale street network data and their inherent network topologies. Therefore, this study aims to illustrate the potential of the proposed GCN-based methodology for analyzing and interpreting street networks. Specifically, we present a case study that employs this innovative approach to predict urban polycentricity using street network data.

## 3. Methodology

The proposed framework includes three main steps: subsampling urban street networks, encoding and decoding subsampled urban street features, and predicting urban polycentricity (Figure 1). The full urban street network across almost 300 cities,



**Figure 1.** (Color online) Overall framework of GCN-based analysis of urban street structure includes three modules: the gray dashed box indicates the creation of subsampled street graphs based on the street-street topology, the orange dashed box shows the extraction of graph embeddings through the GAE model and the blue dashed box represents the prediction of urban polycentricity via the SLP model.

which is the input for the GCN, was too large to compute. Therefore, each urban street network needed to be generalized, and a fixed-size street network sample had to be used (Section 3.1). Thereafter, we implemented a GCN as a feature extractor to encode each subsampled street graph with a comprehensive embedding layer, resulting in a set of multidimensional vectors (embeddings) that could be restored by the decoder (Section 3.2). The extracted embeddings for each city further helped us to answer the question ‘How polycentric are these cities?’ The embeddings were divided into training and test sets. Based on the training set labeled with empirically measured urban polycentricity values, we applied the SLP model to predict the urban polycentricity for the test set (Section 3.3). Finally, the loss function was optimized for the proposed models (Section 3.4).

### 3.1. Subsampling city-scale street graph

Urban roads are naturally modeled as networks or graph representations based on their spatial relationships (e.g. intersections). Roads are conventionally stored as segments in GIS systems, in which they are separated at each street junction where at least three polylines intersect. Subsequently, a dual graph can be generated based on the segment–segment topology, in which the nodes are segments and the links intersect. This type of topology has been widely adopted for many urban applications, such as navigation, and has been effectively integrated into mainstream GIS software for spatial analysis (e.g. the Network Analyst toolbox in ArcGIS). However, uncovering the underlying street structure as a segment, which can be regarded as a geometric primitive with little meaningful information in reality, may be challenging (Ma et al.



2019). As shown in Figure 1, the dual graph of street segments that is visualized in a circular layout is mechanistic or less vivid because the connectivity for each segment is similar. The segment–segment topology can be transformed into a street–street topology for an organic structure. We group and join a set of neighboring segments with two conditions, namely, the same name and good continuity (small deflection angles, e.g. less than  $45^\circ$ ), into a ‘stroke’ (Thomson and Richardson 1999) or ‘natural street’ (Ma et al. 2019) to avoid the generation of meaningless streets. Specifically, we first merge neighboring segments with the same names. Because not all segments have name attributes in reality (database incompleteness), we further consider good continuity for the processed and unprocessed segments together for the final natural street generation. According to Gestalt psychology, a natural street (street) matches human visual cognition and is more meaningful than a segment alone.

Apart from human perception, street–street topology enables a complex network perspective to be adopted to explore the network of streets. Some street properties (e.g. street lengths and connectivities) appear to have a hierarchical structure. As illustrated in Figure 2, approximately 80% of the streets are less connected, whereas the other 20% are well connected, among which approximately 1% are very well connected. Thus, we can observe that the transformation from segments into streets also leads to a profound statistical change in the geometric and topological properties of the street, that is from a normal distribution to a heavy-tailed distribution. Heavy-tailed distribution statistics imply an imbalanced, hierarchical structure that enables us to conduct street network subsampling effectively. Subsequently, well-connected streets can be selected to represent the essence of the entire network to a significant extent. In addition, drawing on the idea of pyramid representation in image processing (Adelson et al., 1984), we adopt the power of two as the scale factor to select streets with the highest connectivity values as the input to the GCN for training and testing, as denoted by Equation (1).

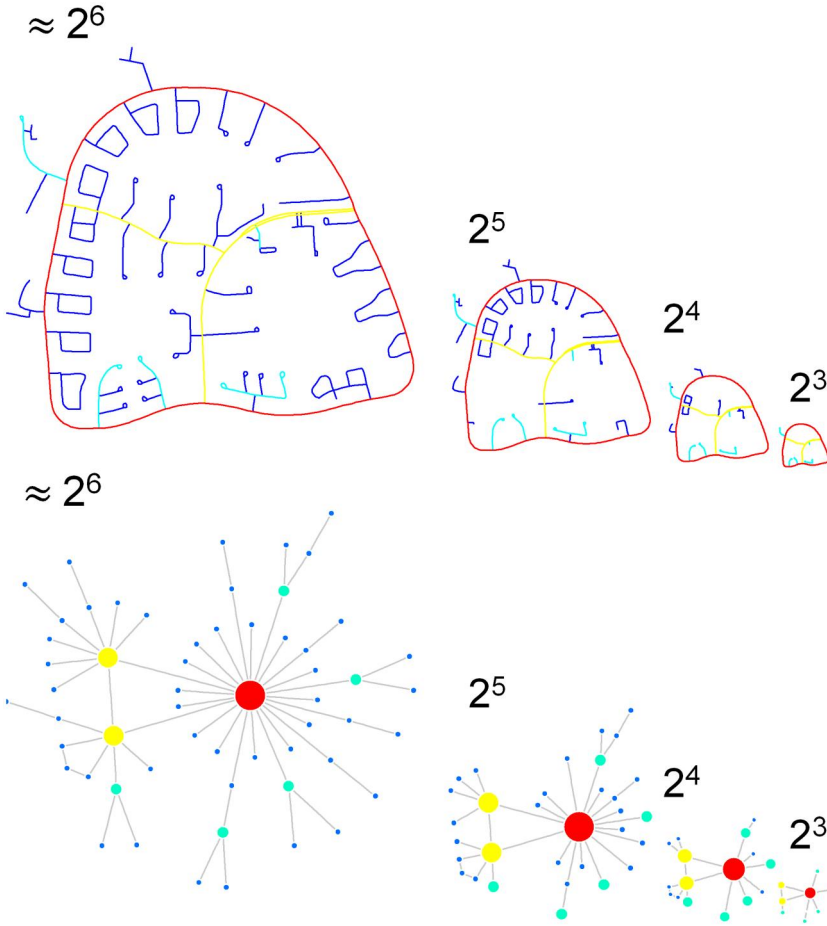
$$\text{of subsampled streets} = 2^{\text{level}} \quad (1)$$

To perform street subsampling, we first rank all streets in decreasing order according to connectivity and select the top  $2^{\text{level}}$  streets as the subsample. An example is depicted in Figure 2. A street network with reduced resolution can be obtained through simple ranking and selection. The resulting ‘network pyramid’ can be adopted as a data structure that efficiently supports the following convolutional operations through representation at multiple reduced scales. As an urban-scale street network usually consists of thousands of streets and the number of streets differs among cities, the level at which the street network should be subsampled must first be decided. This level applies to all cities because an equal-sized subsample for each city is required as the input to the GCN in the subsequent steps for computation and comparability.

### 3.2. GAE model

Assume a graph  $G = (V, A, X)$ , where  $V$  is the set of streets,  $A \in R^{n \times n}$  is the adjacency matrix, and  $X$  is a feature matrix that consists of two features comprising the length





**Figure 2.** Subsampling process of illustrative natural street representation and its dual graph layout. Note: The street network consists of 57 natural streets ( $2^{5.87} \approx 2^6$ ), the spatial intersection relationship of which helps to derive the dual graph. We visualize the importance of the street using the connectivity measure with a color spectrum, where red, yellow, cyan and blue represent the streets most connected to the poorest connected streets. Thereafter, subsampling of the network can be performed at a reduction ratio of 1/2 by selecting the top 32 ( $2^5$ ), 16 ( $2^4$ ) and 8 ( $2^3$ ) most connected streets. Although the number of streets is continuously halved, the main street structure is maintained throughout the series. Source: Adapted from Figure 1 in Ma et al. (2019). Copyright © 2019 by Pion. Reprinted with permission from SAGE Publications, Ltd.

and number of connected streets for each street (both values are normalized). The Laplacian matrix of the specified graph-structured data can be represented as  $\mathcal{L} = D - A$ . In this mathematical representation,  $\mathcal{L}, D$  is the diagonal node degree matrix, and the diagonal element  $D_{ii} = \sum_j A_{ij}$  is the sum of the elements of row  $i$  in the adjacency matrix  $A$ . The graph Laplacian matrix  $\mathcal{L}$  can be further normalized as  $\mathcal{L} = I_N - D^{-\frac{1}{2}} A D^{-\frac{1}{2}} = U \Lambda U^T$ , where  $U$  is the matrix of eigenvectors of the normalized graph Laplacian and  $\Lambda$  is the diagonal matrix of eigenvalues. Assuming that  $x$  is the signal vector that is defined on the nodes of graph  $G$ , the graph convolution can be defined as the multiplication of  $x$  with a filter  $g_\theta$  according to the convolution

theorem in the frequency domain, as outlined in Equation (2):

$$y = g_{\theta} * x = g_{\theta}(\mathcal{L})x = g_{\theta}(U\Lambda U^T)x = Ug_{\theta}(\Lambda)U^T x, \quad (2)$$

where  $g_{\theta}(\Lambda)$  is a function of the eigenvalues of  $\mathcal{L}$ . Therefore, this convolutional structure is impractical for large-scale graphs because the eigendecomposition of the normalized Laplacian matrix  $\mathcal{L}$  may be very computationally expensive. To address this problem, Hammond et al. (2011) introduced a fast approximation of  $g_{\theta}(\Lambda)$ , which is a truncated expansion in terms of the Chebyshev polynomials  $T_k(x)$  up to the  $K$ th order:

$$g_{\theta}(\Lambda) = \sum_{k=0}^K \theta_k T_k(\tilde{\Lambda}). \quad (3)$$

In the approximation,  $\tilde{\Lambda} = \frac{2\Lambda}{\lambda_{\max}} - I_N$  is the rescaled matrix of  $\Lambda$ , where  $\lambda_{\max}$  denotes the maximum eigenvalue of  $L$  and  $I_N$  is an identity matrix of size  $N$ ;  $\theta_k \in R^k$  is a vector of polynomial coefficients, which is to be learned in the training process; and  $T_k(x)$  is recursively defined as  $T_k(x) = 2xT_{k-1}(x) - T_{k-2}(x)$ , with  $T_0(x) = 1$  and  $T_1(x) = x$ . Kipf and Welling (2016a, 2016b) truncated the Chebyshev polynomial to the first order (i.e.  $K=2$  in Equation (3)) as a special variant, leading to the following simplified convolution:

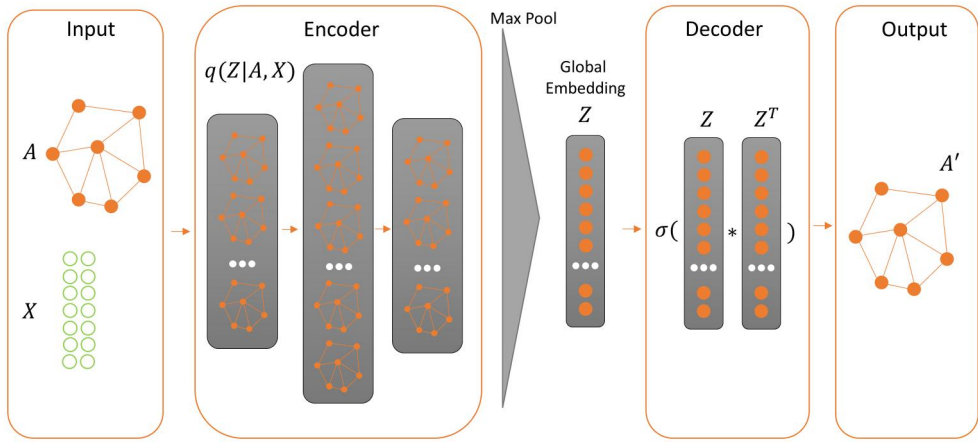
$$g_{\theta} * x \approx \theta(I_N + D^{-\frac{1}{2}}AD^{-\frac{1}{2}})x \approx \theta\left(D^{-\frac{1}{2}}\tilde{A}D^{-\frac{1}{2}} - \frac{1}{2}\right)x, \quad (4)$$

where  $\tilde{A} = A + I_N$  is equivalent to adding self-loops to the original adjacency matrix of the graph and  $\tilde{D}$  is the diagonal degree matrix of  $\tilde{A}$ . Consequently, we can generalize the simplified graph convolution to an input layer of feature  $X^{(l)} \in R^{N \times C}$  (i.e. a  $C$ -dimensional feature vector for every node) and  $F$  filters as follows:

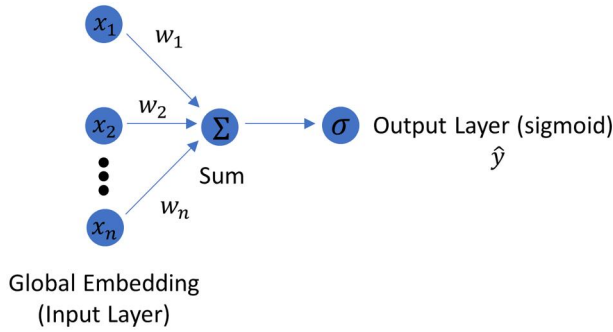
$$X^{(l+1)} = \sigma\left(\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}} - \frac{1}{2}X^{(l)}\Theta^{(l)}\right), \quad (5)$$

where  $\Theta^{(l)}$  is a matrix of  $C \times F$  parameters (i.e. a layer-specific trainable weight matrix),  $X^{(l+1)} \in R^{N \times F}$  is the convolved matrix, which is also the input feature for the next graph convolutional layer  $l+1$ , and  $\sigma$  denotes the activation function, such as  $\text{ReLU}(\cdot) = \max(0, \cdot)$ .

Given the graph convolution operations that are defined in the frequency domain, we can establish an autoencoder comprising five parts: the input, encoder, embedding, decoder and output, as illustrated in Figure 3. In particular, the encoder receives the graph features and adjacency matrices and processes them through multiple stacked graph convolutional layers, as shown in Equation (5). This procedure facilitates information aggregation from interconnected nodes across the graph to generate features for every individual node, that is local embeddings. Following the encoding process, a max pooling layer is employed to extract the overall features of a particular graph from the embedded node-level features. This approach is inspired by the objective of model permutation invariance; that is, different orders of nodes in the graph should produce identical embeddings. The max pooling layer, which is a straightforward symmetric function, accepts  $n$  vectors as the input and delivers an output vector that is unaffected by the input order to ensure order invariance. This concept of deriving global features has been used extensively in deep neural networks,



**Figure 3.** Architecture of the proposed GAE model.



**Figure 4.** (Color online) Architecture of the SLP-based model for predicting urban polycentricity.

such as PointNet (Qi et al. 2017) for point-cloud processing, given that point clouds typically consist of points in varying orders. After deriving the global features, the graph-level embeddings (i.e. global embeddings) are fed into the decoder to realize the reconstruction, which is expected to recover the graph structure in the original data. In this study, inspired by the work of Kipf and Welling (2016a, 2016b), the decoder recovers adjacency matrix  $A'$  as a product of embedding  $Z$  and its transpose. Owing to this architecture, the model can be optimized by minimizing the mean squared error (MSE) between adjacency matrix  $A$  and reconstructed adjacency matrix  $A'$ . The parameters of the convolution kernels and the biases of the activation functions are updated through backpropagation.

### 3.3. SLP-based prediction

We adopt an SLP model to predict urban polycentricities with graph-level embedding, which can be learned from the proposed autoencoder. As illustrated in Figure 4, the SLP model uses the embedding vector from the autoencoder as its input. Subsequently, the aggregate of the weighted embedding values is passed through a sigmoid function to map the prediction value to the range of 0 to 1. Notably, in this

study, the SLP model is trained in a supervised manner, with the city polycentricity values calculated by Liu and Wang (2016) used as a reference for our training data. During the training process, the reference values of a subset of cities are used to refine the predictive model, while the autoencoder network is concurrently trained on all the graph data in a purely unsupervised manner. In this regard, the proposed framework aligns with the definition of a semi-supervised framework, given that all available graphs are used for the unsupervised training of the autoencoder model, whereas only a portion of the reference values is used to train the SLP prediction model. The concurrent optimization of both branches emphasizes the semi-supervised nature of the model. The SLP model parameters are fine-tuned by minimizing the MSE between the predicted and reference values.

### 3.4. Loss function

We address graph-based urban polycentricity prediction using a semi-supervised learning approach owing to the limited number of reference labels. In this study, we express the final loss for model training as the sum of two losses that are derived separately from the two proposed branches:

$$\mathcal{L} = \mathcal{L}_{GACE} + \lambda \mathcal{L}_P, \quad (6)$$

where the unsupervised loss  $\mathcal{L}_{GACE}$  is the MSE between adjacency matrix  $A$  and reconstructed adjacency matrix  $A'$  from the GAE model. The supervised loss  $\mathcal{L}_P$  denotes the MSE between the predicted and available reference values. In addition, the trade-off parameter  $\lambda$  is used to balance the supervised and unsupervised losses during training. In this study, the value of  $\lambda$  is set to 0.0001 to prevent overfitting of the prediction branch.

## 4. Case study and experimental results

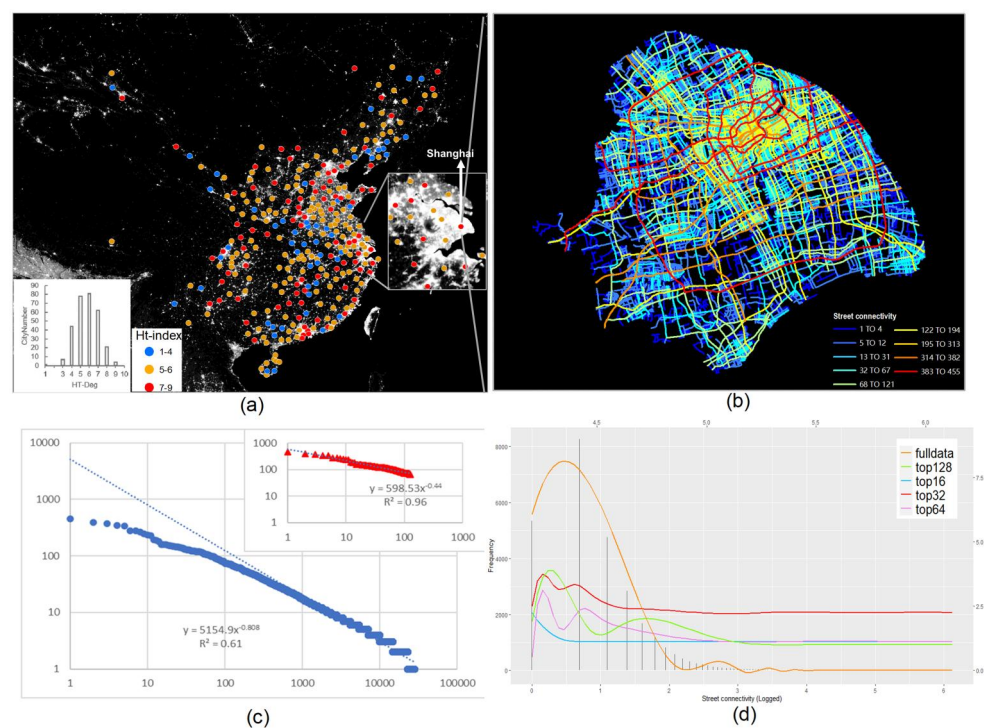
### 4.1. Data sources and processing

Four datasets were used in this study: (1) the national street network, (2) 298 city administrative borders, (3) 2019 national GDP statistics and (4) population-based urban polycentricity scores. The national street dataset was sourced from OpenStreetMap (OSM) and originally included 4,419,603 segments that were strictly separated at the vertices where at least three segments intersected. Each city boundary was further adopted as the unit for data processing. We extracted city-level street segments and transformed them into natural streets. Note that, on average, more than 70% of the name information was absent in the raw segment data, and the situation improved to 50% or less for the top cities. The national GDP data for 2019 were sourced from the published editions of the China Statistical Yearbook 2019. The population-based polycentricity scores of all Chinese cities were adopted from Liu and Wang (2016) and calculated based on the formula in Green (2007), resulting in a polycentricity score ranging from 0 to 1 (a higher value indicates that the city is more polycentric, 0 denotes a total lack of urban polycentricity where there is only one urban center in a

city, and 1 denotes complete urban polycentricity, with several urban centers of the same size in the city).

4.2. Subsampled street networks

Before discussing the experimental results, we briefly introduce the ht-index (Jiang and Yin 2014) to characterize the heavy-tailed distributions of the street measures. Data with a heavy-tailed distribution often exhibit an unbalanced pattern between large and small ratios, that is, very few large values *versus* many small values. Interestingly, this imbalance may recur within the data. The number of recurrences has been termed the ‘ht-index’ and further developed as an effective means for quantifying the extent of imbalance or complexity of data. A higher ht-index indicates a more complex street structure because more street hierarchies can be reflected. We computed the ht-index for each of the 298 cities in terms of the street connectivity using the generated natural streets. The ht-index values ranged from 1 to 9, and almost all the cities had an ht-index  $\geq 3$ . As shown in Figure 5(a), urban streets with relatively high complexity (red dots) are distributed evenly across the western, central and eastern regions. In addition, a general pattern is observed in which important cities such as provincial capital cities tend to have higher ht indices (above 5). Among these cities, Shanghai has the most complex street structure, with an ht-index of



**Figure 5.** (a) Ht-index of street connectivities across 298 Chinese cities, (b) complex structure of the Shanghai street network (ht-index = 9), (c) statistical similarity between the top 128 street connectivities and all the data examined by the power law and (d) histogram. Note: The nighttime image from 2020 of 298 cities in Panel (a) demonstrates the transition of moderate-to-intensive economic development from the western to the eastern regions in China to a certain extent.

nine. Given this high complexity, a clear visual binary division of street connectivities could be observed (Figure 5(b)), that is the scarcity of well-connected streets (in red and yellow) *versus* the number of poorly connected streets (in blue). This partitioning enabled us to identify the street hierarchies effectively, leading to successful subsampling of large-scale street networks.

We selected the 128 most highly connected streets as a subsample of each original urban street network. As in the case of streets in Shanghai, we first conducted simple power-law fitting for the connectivity values of the entire street and the selected subset. By taking the double logarithm of the x- and y-axes, we observed that the distributions of the two datasets were comparable (nearly straight lines; Figure 5(c)). In addition, we counted the frequencies of each logged value and plotted them in a histogram to analyze the statistical patterns from one subset to another. As a further step, a trendline (using curve fitting) was created for this bar, where the entire pattern appeared very right-skewed. Notably, the trendlines for the top 16, 32 and 64 values were distant from the pattern, similar to the whole, until the peak of the top 128 values occurred (green line in Figure 5(d)). This statistical pattern was also confirmed in other cities. Table 1 and Figure 6 present the representative top 128 streets in 20 selected Chinese cities from statistical and geographical perspectives, respectively. Unlike a subsample of streets within equally partitioned grids, the subsample in this study (the most connected and longest roads) covered the entire urban area. Interestingly, of the 20 cities with different socioeconomic statuses, each subsample (only 1% of streets) had an ht-index above or equal to 3, accounting for no less than 50% of the structural complexity of all streets (in most cases). In addition, we observed that the power-law distributions of the subsampled and full street networks were similar across cities, highlighting the ability of the ht-index to delineate subtle differences in street hierarchy.

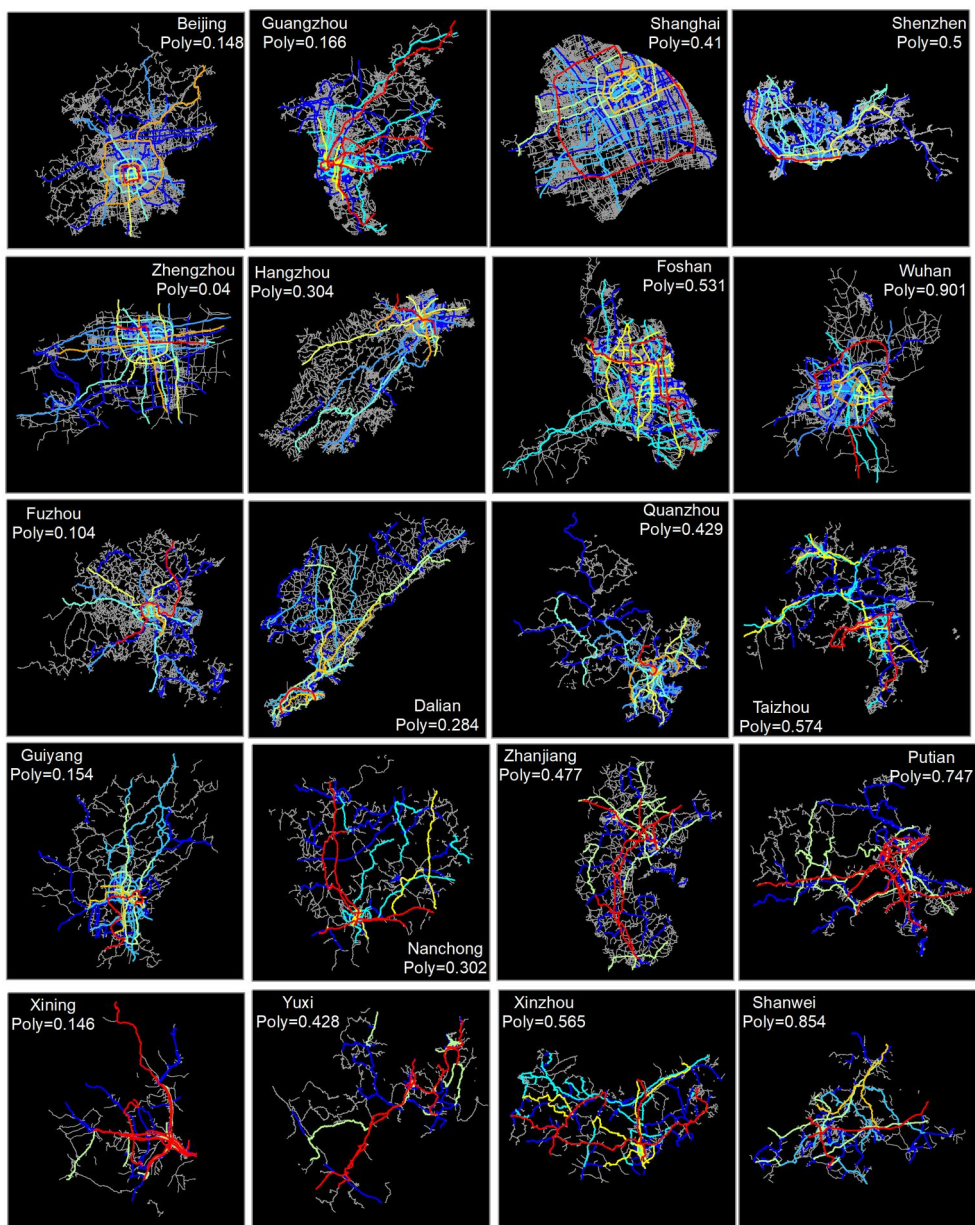
**Table 1.** Ht-index for subsample and all streets for 20 Chinese cities.

City	# Streets	Ht (all)	Ht (top)
Beijing	41,829	7	3
Guangzhou	28,215	8	4
Shanghai	34,924	9	5
Shenzhen	18,938	8	4
Zhengzhou	6014	7	4
Hangzhou	24,948	9	3
Foshan	16,053	8	4
Wuhan	12,615	6	3
Fuzhou	8744	8	4
Dalian	7512	8	5
Quanzhou	4414	7	6
Taizhou	6517	7	4
Guiyang	3147	6	5
Nanchong	2769	7	4
Zhanjiang	4061	8	5
Putian	1787	6	4
Xining	1811	5	3
Yuxi	1643	7	4
Xinzhou	3856	8	4
Shanwei	1758	6	5

Note: #Number.

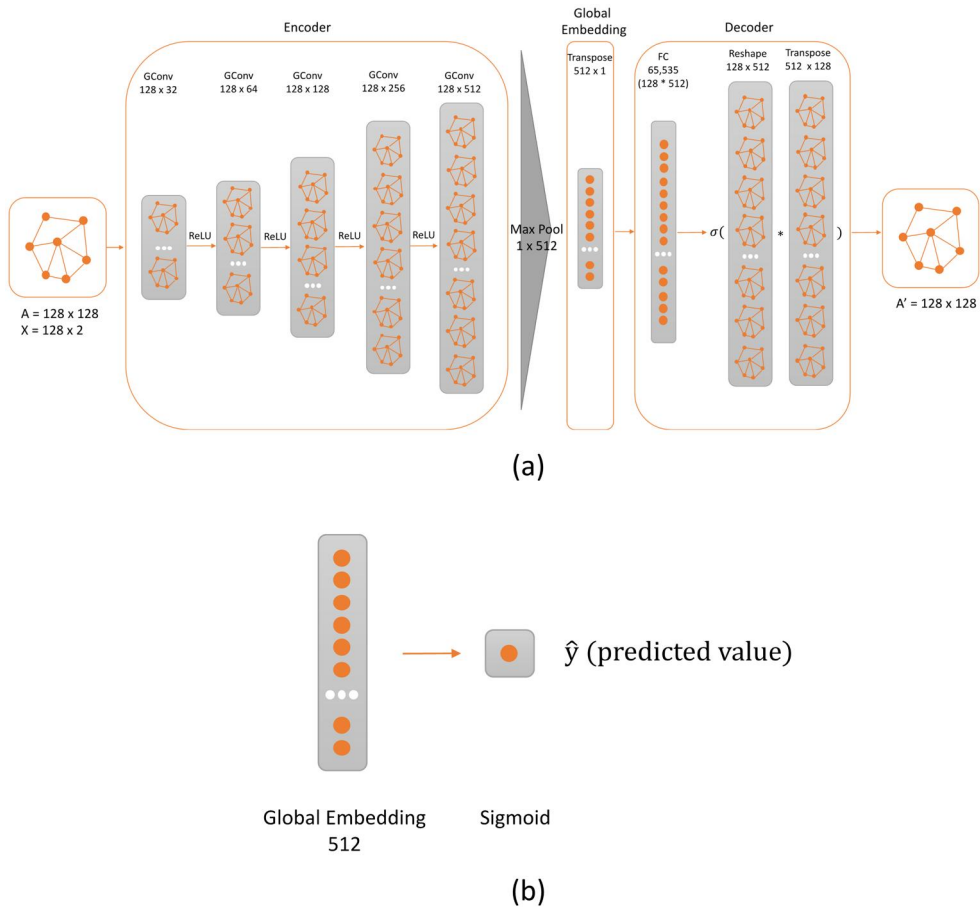
Ht: ht-index; all: all streets; top: top 128 streets





**Figure 6.** Layout of the top 128 streets for 20 Chinese cities at different development levels. Note: Poly = population-based polycentricity scores previously measured by Liu and Wang (2016). The hierarchy of the cities from top to bottom is aligned with the row sequence, e.g. the first row indicates the cities at the top hierarchy determined by urban GDP, and the fifth row is the fifth hierarchy. For each row (cities in the same hierarchy), we ordered the cities in ascending order from monocentric (left) to polycentric (right).





**Figure 7.** Architecture of (a) GAE and (b) SLP models for urban polycentricity prediction.

#### 4.3. Model training and parameter settings

The architecture and parameters of the autoencoder model are depicted in Figure 7(a). The encoder of the proposed framework contains five convolutional layers, with dimensions of 32, 64, 128, 256 and 512. After the fifth graph convolutional layer, the node-level graph features with dimensions of  $128 \times 512$  are fed into the max pooling layer to derive a  $512 \times 1$  vector, which is the global-level feature for the involved graph. For the decoder, the  $512 \times 1$  graph coding is first scaled to a  $65,536 \times 1$  vector and then reshaped to a  $128 \times 512$  matrix  $Z$ , with dimensions identical to those of the output of the final graph convolutional layer in the encoder. Finally, adjacency matrix  $A'$  is reconstructed as the product of matrix  $Z$  and its transpose. Simultaneously, the 512-dimensional global embeddings are used as the inputs for predicting the urban polycentricity. As shown in Figure 7(b), a sigmoid function is used to obtain a predicted value from 0 to 1. This value represents the predicted urban polycentricity of the input street-street connectivity graph.

In the experiment, the autoencoder and SLP models were simultaneously optimized using the Adam optimizer with a learning rate of 0.0001. All 298 city graphs were

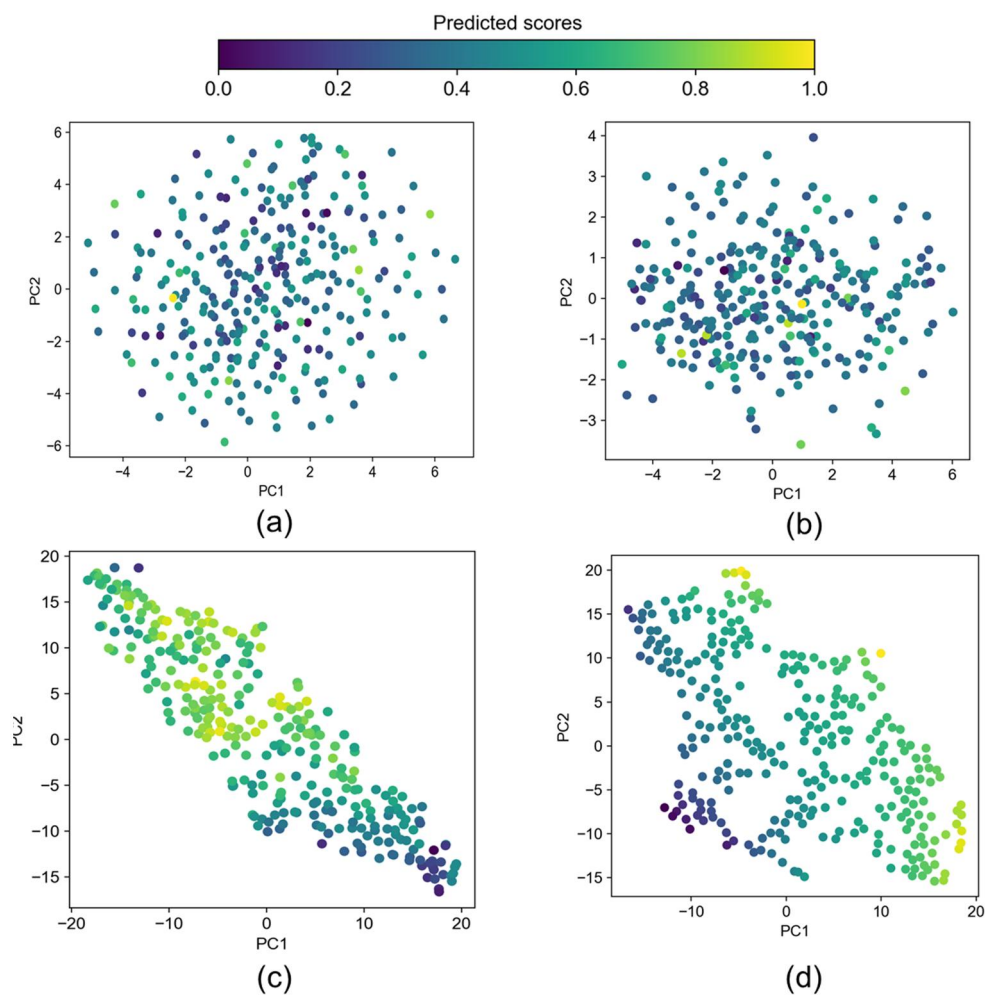
used for unsupervised training of the GAE model. A total of 159 cities, each of which had a referenced polycentricity score greater than 0, were used for training and testing (i.e. 104 (approximately 65%) for training and 55 (approximately 35%) for testing). Note that a total of 139 cities with referenced polycentricity scores of 0 were excluded from the training process. The minibatch size was set to 32, and the models were trained for 1000 epochs. The area under the receiver operating characteristic curve (AUC) was 0.934, and the average precision of the trained GAE model was 0.952.

#### 4.4. Comparison with baseline models

We selected the following baseline models, which all produced 128-dimensional representations across cities, to verify the importance of the subsampled streets in relation to urban polycentricity and to compare the performance of our proposed GAE model.

- GraphSAGE, which was proposed by Hamilton et al. (2017), is a powerful method for learning low-dimensional vector representations of nodes in graphs. This is achieved by sampling neighboring nodes and aggregating their features to obtain node representations. In this study, we constructed an end-to-end model as a baseline with a similar structure to that of the proposed GAE. The mean aggregation technique was employed for information aggregation.
- Node2Vec is a powerful unsupervised graph representation technique that transforms nodes in graphs into low-dimensional vectors (Grover and Leskovec 2016). We employed Node2Vec to embed each street into an eight-dimensional feature and combined it with a two-dimensional (2D) street attribute feature, which yielded a  $128 \times 10$  feature matrix for each city. Thereafter, principal component analysis (PCA) was used for dimensionality reduction to obtain 128-dimensional features for each city. The walk length and number of walks sampled for each node were set to 10, and the actual context size considered for the positive samples was five.
- The vanilla GCN was inspired by the foundational GCN approach introduced by Kipf et al. (2016a, 2016b). We implemented five GCN layers, with dimensions that were sequentially configured as 32, 64, 128, 256 and 1, for the encoder of the GAE. Notably, the final GCN layer in the encoder produced a  $128 \times 1$  output dimension, which was an aggregated feature that was achieved by retaining only one feature per node in the graph. These aggregated features acted as global embeddings and were employed to rebuild the adjacency matrix during the decoder stage.

Street graphs should ideally exhibit a distribution pattern in which graphs of the same type are adjacent, whereas graphs of different types are separated from the coding space. We mapped the 128-dimensional coding representation of each city to the two principal axes that were derived *via* PCA. Figure 8 depicts the correlation between the distribution of cities in the encoding space and their predicted polycentricity scores for each trained network. The results of our proposed method show that cities with lower predicted polycentricity scores are positioned in the lower region of the plot, whereas those with higher scores are grouped in the upper part of the graph. This pattern signifies the effective distribution of cities according to their polycentricity.



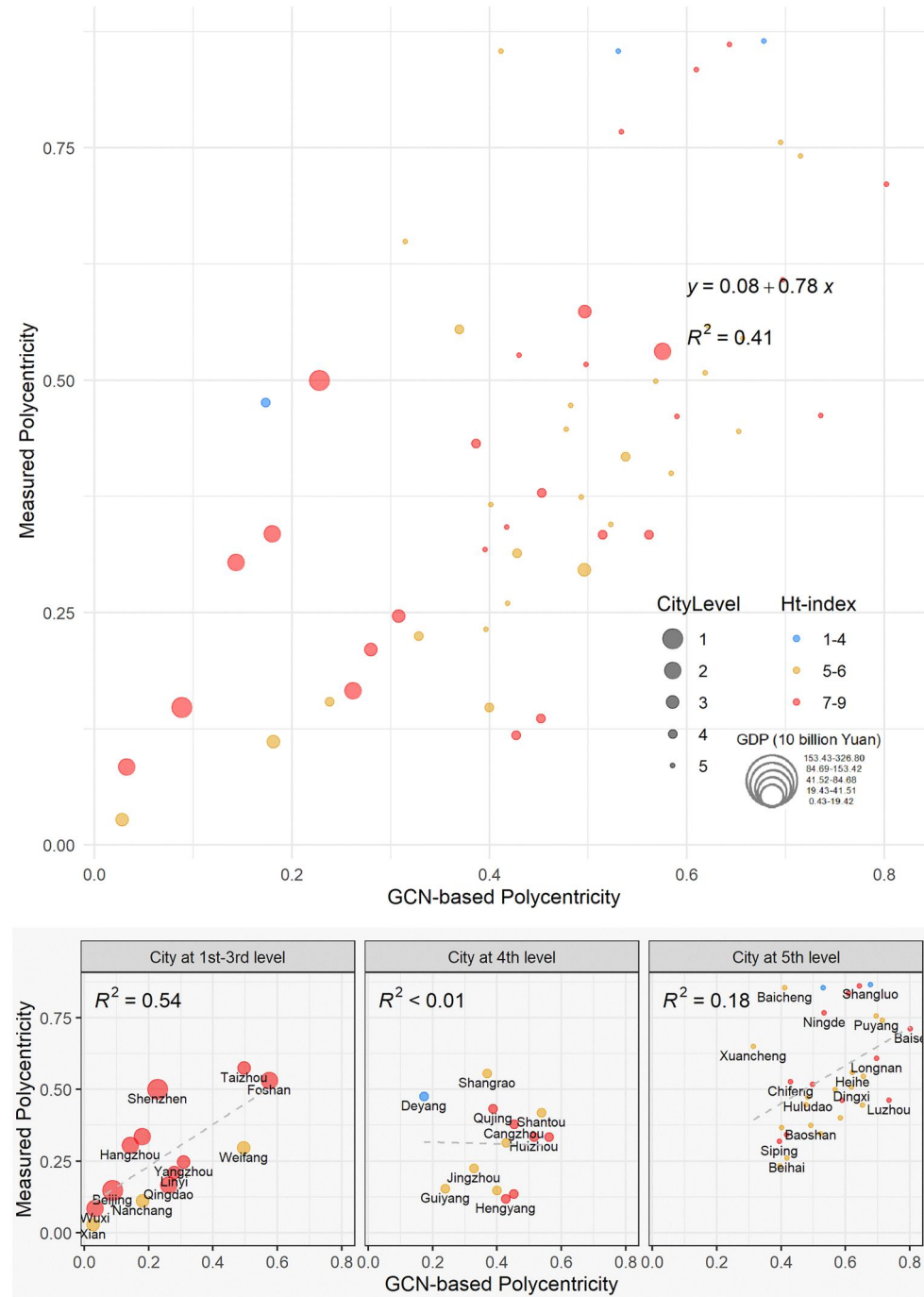
**Figure 8.** Visualization of derived embeddings from (a) vanilla GCN, (b) Node2Vec, (c) GraphSAGE, and (d) proposed GAE model in a 2D space. Note: The dot color was rendered using the predicted polycentricity scores for 298 cities.

**Table 2.** Performance of our approach and baseline models.

	$R^2$	MAE	RMSE
Node2Vec	0.12	0.18	0.24
GraphSAGE	0.23	0.21	0.27
Proposed GAE	0.41	0.14	0.17
Vanilla GCN	0.017	0.23	0.29

Three regression-based metrics, namely, the  $R^2$ , mean absolute error (MAE) and root MSE (RMSE), were adopted to determine the correlation between the predicted and observed polycentricity scores. We used data from only nonmonocentric cities (those with at least one subcenter) to perform the correlation test. As a result, 55 polycentric cities with diverse economic statuses and street complexities were selected. [Table 2](#)

presents the results for the metrics. Similar to the comparison of the embedding plots, it can be observed that our proposed approach outperforms the baseline models in the three measures. The performance was assessed by GraphSAGE and Node2Vec (as



**Figure 9.** Scatterplot of the predicted polycentric values and previously measured values for 55 Chinese cities at the different levels defined by the urban GDP.

indicated by the  $R^2$  values in Table 2). Noticeable distinctions in performance among the different models emerged in terms of the MAE. The proposed GAE exhibited remarkable performance, with an MAE of 0.14. In contrast, GraphSAGE and Node2Vec yielded higher MAE values of 0.21 and 0.18, respectively. The RMSE metric showed a comparable trend. The GAE achieved an RMSE of 0.17, whereas GraphSAGE and Node2Vec attained values of 0.27 and 0.24, respectively.

#### **4.5. Correlations between measured polycentricity scores and predicted scores**

Figure 9 shows the correlation between the polycentricity predicted using the proposed GAE model and the previously measured scores. The correlation coefficient was 0.41 (Table 2), which was significant at the 0.01 level (two-tailed). We divided the 55 cities into three groups according to their urban GDP to conduct an in-depth investigation. Specifically, the cities in the top hierarchies had relatively low polycentricity scores owing to the existence of leading dominant centers. In contrast, the polycentricity profiles for the cities in the bottom class differed because the scores were distributed evenly along the value range. The correlation results varied across the different groups. High predictability was observed within cities in the top three hierarchies (approximately 0.5) for urban GDP and street complexity. In contrast, cities at lower levels were found to have very low predictability (below 0.2). We also visualized the ht-index values for each city. Note that the groups based on urban GDP and the ht-index were not the same but overlapped to a certain degree. A general trend between the two groups was that cities with a better economic status tended to have a more complex street structure (but not always).

### **5. Discussion**

This study confirms the effectiveness of a deep learning framework for large-scale comparative urban analysis through network representation learning and reveals the underlying connection between street network structures and population-based polycentric urban structures. Like in other graph-structured data, the difficulty of learning street network representations lies in structure preservation and high nonlinearity. To address this issue, the proposed architecture first preserves the inherently complex street structure by iteratively joining adjacent, homonymic and continuous segments into streets with a deep hierarchy; that is, numerous less-connected streets, few well-connected streets and some in between. In this way, the street network exhibits a form of self-similarity, in which parts reflect the characteristics of the whole network (Zhang et al. 2022). Namely, well-connected streets are representative of the entire network structure. This recognition of self-similarity is instrumental in the subsampling process and offers fresh perspectives on the concept of scale in geographic representation learning, particularly within spatial networks (Yuan and McKee 2022). Each city possesses geometrical, topological and societal properties that differ among cities. The graph embeddings of well-connected streets that are extracted through multiple convolutional layers can map these attributes in a highly nonlinear latent space (128-dimensional) and help characterize urban similarities across different street layouts.

The superior performance exhibited by the GCN-based approach compared with that of other popular GNN models, such as Node2Vec and GraphSAGE, can be attributed more to the specific downstream task than to the model itself. The desired embeddings are expected to accurately represent the structural complexity of the subsampled street graph and its related street features for predicting urban polycentricity. The underperformance of Node2Vec is due to its lack of consideration of node features. Like Node2Vec, which yields inferior results, the vanilla GCN is inherently designed for node-level (rather than graph-level) feature extraction. The selected  $128 \times 1$  global embeddings that are derived by retaining only one feature per node from the final GCN layer in the encoder do not optimally represent the graphs examined in our study. These results also indicate the superior efficacy of the proposed max pooling layer in aggregating the global representation of the graph from embedded node-level attributes. However, GraphSAGE primarily emphasizes local structures by acquiring an aggregator that aggregates features from neighboring nodes (first or second order). Nevertheless, when the subsampled graph structure is known and representative, the GAE enables a more global perspective (higher-level order) by leveraging the relational information among the important nodes and their features in the graph. The characteristics of GCNs are more in line with the requirements of the task at hand and thus ultimately contribute to the superior performance of the proposed architecture.

Our analysis excludes cities that are characterized by a single population center during the prediction phase. This exclusion stems from the fact that monocentric cities inherently have a polycentricity score of 0 (Green 2007), which renders them unsuitable for correlation analysis. Although Liu and Wang (2016) assigned a polycentricity score of 0 to these cities based on population data, our findings indicate the potential for slight polycentricity (i.e. marginally above 0) in these areas. This discrepancy could be attributed to the limited resolution of the LandScan gridded population data (1 km  $\times$  1 km), which may not capture subtle urban centers as effectively as street network data. Street network data not only complement but also enhance our understanding of urban polycentricity by revealing more nuanced urban structures. Consequently, the integration of street- and population-based findings in this study offers a more detailed and comprehensive view of urban structural dynamics.

Cities with a higher GDP tend to have more complex street structures, and their predicted polycentricity scores are more consistent with the empirically measured population-informed urban polycentricity than those of other cities. Cities with better economies, such as those along the eastern Chinese seaboard (e.g. the Pearl River and Yangtze River Deltas; Figure 5), have experienced intensive decentralization and marketization processes (Liu et al. 2018); thus, they naturally have at least one populous location where people gather for various activities. As cities evolve, these concentrations of people and activities lead to more roads locally and various densities in the road network globally, resulting in the emergence of street complexity. The street complexity measured using the ht-index in this study reflects the variation in street-street connections, which is widely considered to be the driver of network community structure (Fortunato and Newman 2022). In other words, a larger ht-index indicates greater variation in connectivity values across streets and a greater chance of the

street network having communities in which streets are strongly connected. Our results show that more developed cities tend to have more complex street structures (larger ht-index values). A higher correlation between the predicted and measured polycentricity values among these cities also strengthens the potential relevance of complex street configurations to the polycentric urban structure.

## 6. Conclusions

This study successfully demonstrated the efficacy of a GCN-based approach for analyzing cross-city street network structures using a two-branch, semi-supervised architecture. The methodological innovation of using natural streets underscores the evolving discourse in street network analyses (Marshall et al. 2018). The study demonstrated that the representation learning of the subsampled streets across cities, constituting approximately 1% of the total streets, can capture the global structural characteristics of the entire street network. The case study on predicting urban polycentricity informed by population data in 298 Chinese cities based on the proposed GCN-based approach relies on the unique integration of deep learning with urban street network analysis. A key observation was that more economically developed cities exhibit more intricate street network structures, reflecting their dynamic urban evolution and the resulting diverse densities in street networks. This complexity, as measured using the ht-index, correlates with variations in street–street connections and the likelihood of polycentric urban forms being developed. This study also highlights that street network data can complement and enhance the nuanced understanding of urban structures more than traditional data sources, such as the gridded population data of LandScan.

This study lays the foundations for future research. The observed alignment between the predicted polycentricity scores from street networks and population-based data opens new avenues for research in urban studies. This synergy invites further investigation into the interconnected dynamics of street layouts and population distributions. However, as Derudder (2021) noted, network analyses do not always provide clear-cut interpretations, potentially limiting their practical applicability in urban planning and policy formulation. To bridge this gap, future research could incorporate explainable AI techniques into the proposed methodology to enhance the transparency and comprehensibility of urban polycentricity predictions from street networks for urban policymakers, planners and designers. Future research could also expand this methodological framework by applying other urban metrics and investigating their potential in different geographical contexts, thereby further enriching the understanding of urban structure and dynamics and facilitating the development of more livable and sustainable urban environments.

## Acknowledgments

The authors would like to thank the editors Prof. May Yuan and Prof. Stefano De Sabbata and the anonymous reviewers for their constructive comments to greatly improve the quality of the article.



## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Funding

This study is supported and funded by the National Natural Science Foundation of China (No. 42001180 and No. 42071357). We would like also to acknowledge the financial support from the Guangdong Science and Technology Strategic Innovation Fund (the Guangdong-Hong Kong-Macau Joint Laboratory Program, No. 2020B1212030009) and Shenzhen Key Laboratory of Digital Twin Technologies for Cities (No. ZDSYS20210623101800001).

## Notes on contributors

**Dr. Ding Ma** is an assistant professor of urban informatics at the Department of Urban Informatics, Shenzhen University. His research interests concentrate on complex network analysis and the use of GeoAI in modeling urban structure and dynamics. He contributed to the conceptualization, methodology, formal analysis, writing – original draft, writing – review and editing.

**Dr. Fangning He** holds a PhD in Geomatics at Purdue University and currently serves as a computer vision scientist at NVIDIA. His research interests include 3D reconstruction from imagery and LiDAR, as well as photogrammetry, high-definition map creation, and a range of other 2D and 3D mapping techniques. He contributed to the conceptualization, methodology, writing – original draft, writing – review and editing.

**Dr. Yang Yue** is a professor at the Department of Urban Informatics, Shenzhen University. Her research interests include urban spatial modeling, social computing and sustainable development evaluation. She contributed to the methodology, supervision and writing – review and editing.

**Dr. Renzhong Guo** is a member of the Chinese Academy of Engineering and a professor at the Department of Urban Informatics, Shenzhen University. His research includes smart cities, geographical information engineering, 3D city modeling, and spatial planning of national territories. He contributed to the supervision and project administration.

**Dr. Tianhong Zhao** is a lecturer at the College of Big Data and Internet, Shenzhen Technology University. His research interests include GNNs and spatiotemporal big data analysis. He contributed to visualization, formal analysis and writing – review and editing.

**Dr. Mingshu Wang** is a senior lecturer (tenured associate professor) at the University of Glasgow. His research combines geospatial data and computational methods to understand urban systems and development. He contributed to the conceptualization, methodology, investigation, writing – original draft, writing – review and editing.

## ORCID

Ding Ma  <http://orcid.org/0000-0001-9328-9584>

Mingshu Wang  <http://orcid.org/0000-0001-5260-3143>

## Data and codes availability statement

The data and codes that support the findings of this study are openly available in Figshare at <https://doi.org/10.6084/m9.figshare.20761462.v1>.

## References

- Alexander, C., 1965. A city is not a tree. *Architectural Forum*, 122 (1), 58–62.
- Adelson, E.H., et al. 1984. Pyramid methods in image processing. *RCA Engineer*, 29 (6), 33–41.
- Atwood, J. and Towsley, D., 2016. Diffusion-convolutional neural networks. In: *Advances in neural information processing systems*, 2001–2009.
- Batty, M., 2008. The size, scale, and shape of cities. *Science*, 319 (5864), 769–771.
- Bengio, Y., Courville, A., and Vincent, P., 2013. Representation learning: a review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35 (8), 1798–1828.
- Boeing, G., 2017. OSMnx: new methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Computers, Environment and Urban Systems*, 65, 126–139.
- Bronstein, M.M., et al., 2017. Geometric deep learning: going beyond Euclidean data. *IEEE Signal Processing Magazine*, 34 (4), 18–42.
- Burger, M. and Meijers, E., 2012. Form follows function? Linking morphological and functional polycentricity. *Urban Studies*, 49 (5), 1127–1149.
- Crucitti, P., Latora, V., and Porta, S., 2006. Centrality measures in spatial networks of urban streets. *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics*, 73 (3 Pt 2), 036125.
- Defferrard, M., Bresson, X., and Vandergheynst, P., 2016. Convolutional neural networks on graphs with fast localized spectral filtering. In: *Advances in neural information processing systems*, 3844–3852.
- Derudder, B., 2021. Network analysis of ‘urban systems’: potential, challenges, and pitfalls. *Tijdschrift Voor Economische en Sociale Geografie*, 112 (4), 404–420.
- Derudder, B., et al., 2021. Measuring polycentric urban development: the importance of accurately determining the ‘balance’ between ‘centers. *Cities*, 111, 103009.
- Derudder, B., et al., 2022. Polycentric urban regions: conceptualization, identification and implications. *Regional Studies*, 56 (1), 1–6.
- Fortunato, S. and Newman, M.E., 2022. 20 years of network community detection. *Nature Physics*, 18 (8), 848–850.
- Green, N., 2007. Functional polycentricity: a formal definition in terms of social network analysis. *Urban Studies*, 44 (11), 2077–2103.
- Grover, A. and Leskovec, J., 2016. node2vec: Scalable feature learning for networks. In: *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*. San Francisco, CA: Association for Computing Machinery, 855–864.
- Hamilton, W., Ying, Z., and Leskovec, J., 2017. Inductive representation learning on large graphs. In: *Advances in neural information processing systems*, 1025–1035.
- Hammond, D.K., Vandergheynst, P., and Gribonval, R., 2011. Wavelets on graphs via spectral graph theory. *Applied and Computational Harmonic Analysis*, 30 (2), 129–150.
- Harrison, J., et al., 2022. Governing polycentric urban regions. *Territory, Politics, Governance*, 11 (2), 213–221.
- Huang, J., et al. 2022. A city is not a tree: a multi-city study on street network and urban life. *Landscape and Urban Planning*, 226, 104469.
- Jiang, B. and Yin, J., 2014. Ht-index for quantifying the fractal or scaling structure of geographic features. *Annals of the Association of American Geographers*, 104 (3), 530–540.
- Kipf, T.N. and Welling, M., 2016a. Variational graph auto-encoders. *arXiv:1611.07308*.
- Kipf, T.N. and Welling, M., 2016b. Semi-supervised classification with graph convolutional networks. *arXiv:1609.02907*.
- Kirkley, A., et al., 2018. From the betweenness centrality in street networks to structural invariants in random planar graphs. *Nature Communications*, 9 (1), 2501.
- Lan, T., Li, Z., and Zhang, H., 2019. Urban allometric scaling beneath structural fractality of road networks. *Annals of the American Association of Geographers*, 109 (3), 943–957.
- Lee, J., Lee, I., and Kang, J., 2019. Self-attention graph pooling. In: *Proceedings of the 36th international conference on machine learning*. Long Beach, CA: PMLR, 3734–3743.
- Liu, K., Gao, S., and Lu, F., 2019. Identifying spatial interaction patterns of vehicle movements on urban road networks by topic modelling. *Computers, Environment and Urban Systems*, 74, 50–61.

- Liu, X., Derudder, B., and Wang, M., 2018. Polycentric urban development in China: A multiscale analysis. *Environment and Planning B: Urban Analytics and City Science*, 45 (5), 953–972.
- Liu, X. and Wang, M., 2016. How polycentric is urban China and why? A case study of 318 cities. *Landscape and Urban Planning*, 151, 10–20.
- Lu, Z., et al., 2016. Fractal dimensions of metropolitan area road networks and the impacts on the urban built environment. *Ecological Indicators*, 70, 285–296.
- Lv, Y., et al., 2021. Detecting the true urban polycentric pattern of Chinese cities in morphological dimensions: A multiscale analysis based on geospatial big data. *Cities*, 116, 103298.
- Ma, D., et al., 2019. Why topology matters in predicting human activities? *Environment and Planning B: Urban Analytics and City Science*, 46 (7), 1297–1313.
- Marshall, S., et al., 2018. Street network studies: from networks to models and their representations. *Networks and Spatial Economics*, 18 (3), 735–749.
- Meijers, E., 2008. Measuring polycentricity and its promises. *European Planning Studies*, 16 (9), 1313–1323.
- Penn, A., et al., 1998. Configurational modelling of urban movement networks. *Environment and Planning B: Planning and Design*, 25 (1), 59–84.
- Porta, S., Crucitti, P., and Latora, V., 2006. The network analysis of urban streets: a dual approach. *Physica A: Statistical Mechanics and Its Applications*, 369 (2), 853–866.
- Qi, C.R., et al., 2017. Pointnet: deep learning on point sets for 3d classification and segmentation. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. Honolulu, HI: IEEE, 652–660.
- Sharifi, A., 2019. Resilient urban forms: a review of literature on streets and street networks. *Building and Environment*, 147, 171–187.
- Taubenböck, H., et al., 2017. Measuring morphological polycentricity-A comparative analysis of urban mass concentrations using remote sensing data. *Computers, Environment and Urban Systems*, 64, 42–56.
- Thomas, R., Schmidt, S., and Siedentop, S., 2022. Toward comparative polycentricity scores: assessing variations in regional delineation and subcenter identification. *Environment and Planning B: Urban Analytics and City Science*, 49 (6), 1597–1611.
- Thomson, R.C. and Richardson, D.E., 1999. The ‘good continuation’ principle of perceptual organization applied to the generalization of road networks. In: *Proceedings of the ICA 19th international cartographic conference*. Ottawa, Canada: ICA, 1215–1223.
- Volgmann, K. and Münter, A., 2022. Understanding metropolitan growth in German polycentric urban regions. *Regional Studies*, 56 (1), 99–112.
- Wang, W., et al., 2019. Local floods induce large-scale abrupt failures of road networks. *Nature Communications*, 10 (1), 2114.
- Wang, M., 2021. Polycentric urban development and urban amenities: evidence from Chinese cities. *Environment and Planning B: Urban Analytics and City Science*, 48 (3), 400–416.
- Wang, M. and Debbage, N., 2021. Urban morphology and traffic congestion: longitudinal evidence from US cities. *Computers, Environment and Urban Systems*, 89, 101676.
- Wang, D., Cui, P., and Zhu, W., 2016. Structural deep network embedding. In: *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*. San Francisco, CA: Association for Computing Machinery, 1225–1234.
- Wu, C., et al., 2024. Machine learning-based characterisation of urban morphology with the street pattern. *Computers, Environment and Urban Systems*, 102078.
- Xue, J., et al., 2022. Quantifying the spatial homogeneity of urban road networks via graph neural networks. *Nature Machine Intelligence*, 4 (3), 246–257.
- Yan, X., et al., 2019. A graph convolutional neural network for classification of building patterns using spatial vector data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 150, 259–273.
- Yan, X., et al., 2021. Graph convolutional autoencoder model for the shape coding and cognition of buildings in maps. *International Journal of Geographical Information Science*, 35 (3), 490–512.
- Yao, X., et al., 2021. Spatial origin-destination flow imputation using graph convolutional networks. *IEEE Transactions on Intelligent Transportation Systems*, 22 (12), 7474–7484.

- Ying, Z., et al., 2018. Hierarchical graph representation learning with differentiable pooling. In: *Advances in neural information processing systems*, 4805–4815.
- Yu, B., Yin, H., and Zhu, Z., 2018. Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting. In: *Proceedings of the twenty-seventh international joint conference on artificial intelligence*. Stockholm, Sweden: IJCAI, 3634–3640.
- Yu, W., et al., 2020. Road network generalization considering traffic flow patterns. *International Journal of Geographical Information Science*, 34 (1), 119–149.
- Yuan, M. and McKee, A., 2022. Embedding scale: new thinking of scale in machine learning and geographic representation. *Journal of Geographical Systems*, 24 (3), 501–524.
- Zhang, H., Lan, T., and Li, Z., 2022. Fractal evolution of urban street networks in form and structure: a case study of Hong Kong. *International Journal of Geographical Information Science*, 36 (6), 1100–1118.
- Zhang, S., et al., 2019. Graph convolutional networks: a comprehensive review. *Computational Social Networks*, 6 (1), 11.
- Zhu, D., et al., 2020. Understanding place characteristics in geographic contexts through graph convolutional neural networks. *Annals of the American Association of Geographers*, 110 (2), 408–420.
- Zhu, X., and Rabbat, M., 2012. Approximating signals supported on graphs. In: *Presented at the 2012 IEEE international conference on acoustics, speech and signal processing (ICASSP)*. Kyoto, Japan: IEEE, 3921–3924.