Landscape and Urban Planning 251 (2024) 105171

Contents lists available at ScienceDirect

Landscape and Urban Planning

journal homepage: www.elsevier.com/locate/landurbplan

Research Paper

How does spatial structure affect psychological restoration? A method based on graph neural networks and street view imagery

Haoran Ma [a], Yan Zhang [a], Pengyuan Liu [b], Fan Zhang [c], Pengyu Zhu [d,*]

[a] *Institute of Space and Earth Information Science, The Chinese University of Hong Kong, Hong Kong, China*
[b] *Future Cities Lab Global, Singapore-ETH Centre, Singapore*
[c] *Institute of Remote Sensing and Geographical Information System, School of Earth and Space Sciences, Peking University, China*
[d] *Division of Public Policy, Hong Kong University of Science and Technology, Hong Kong, China*

HIGHLIGHTS

- The psychological restoration was evaluated through two different types of graphs at the street and city levels.
- Using sequential Street View Images to characterize urban spatial structure and explore its relationship to restoration.
- Spatial-dependent Graph Neural Network outperformed traditional models in model prediction.
- Spatial structure (street-level graphs) makes a significant contribution to the prediction of psychological restoration.

ARTICLE INFO

Keywords:
Restoration
Graph neural network
Street view image
Spatial structure
GeoAI
Attention restoration theory

ABSTRACT

The Attention Restoration Theory (ART) proposed four essential indicators (being away, extent, fascinating, and compatibility) for understanding urban and natural restoration quality. However, previous studies have overlooked the impact of spatial structure (the visual relationships between scene entities) and neighboring environments on restoration quality as they mostly relied on isolated questionnaires or images. This study introduces a spatial-dependent graph neural networks (GNNs) approach to address this gap and explore the relationship between spatial structure and restoration quality at a city scale. Two types of graphs were constructed: street-level graphs using sequential street view images (SVIs) to capture visual relationships between entities and represent spatial structure, and city-level graphs modeling the topological relationships of roads to capture the spatial features of neighboring entities, integrating perceptual, spatial, and socioeconomic features to measure restoration quality. The results demonstrated that spatial-dependent GNNs outperform traditional models, achieving an accuracy (Acc) of 0.742 and an F1 score of 0.740, indicating their exceptional ability to capture features of adjacent spaces. Ablation experiments further revealed the substantial positive impact of spatial structure features on the predictive performance for restoration quality. Moreover, the study highlighted the greater significance of naturally relevant entities (e.g., trees) compared to artificial entities (e.g., buildings) in relation to high restoration quality. This study clarifies the association between spatial structure and restoration quality, providing a new perspective to improve urban well-being in the future.

1. Introduction

The urban landscape, home to over half the world's population, is in a state of flux and is predicted to accommodate 75 % of the global populace by 2050 (Ritchie and Roser, 2018). This rapid urbanization necessitates a focus on mitigating the adverse impacts of the urban environment on human health, particularly in promoting physical health, managing stress, and preventing stress-related diseases (Akpınar, 2021; Cetin et al., 2021; Liu et al., 2020). Improving the urban built environment emerges as a broad-based solution, recognized for its potential to address these mental health issues (UNFPA, 2007; Keniger et al., 2013; Hough, 2014).

* Corresponding author.
E-mail addresses: 6210307146@stu.jiangnan.edu.cn (H. Ma), sggzhang@whu.edu.cn (Y. Zhang), pengyuan.liu@sec.ethz.ch (P. Liu), fanzhanggis@pku.edu.cn (F. Zhang), pengyuzhu@ust.hk (P. Zhu).

https://doi.org/10.1016/j.landurbplan.2024.105171
Received 29 November 2023; Received in revised form 15 July 2024; Accepted 16 July 2024
Available online 1 August 2024
0169-2046/© 2024 Elsevier B.V. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

Meanwhile, exposure to forests and green spaces has been confirmed to benefit human mental health, effectively reducing stress (Capaldi et al., 2014), improves mood (Berman et al., 2008), and restoring depleted cognitive resources (Akpınar, 2021; Liu et al., 2020), as supported by the Attention Restoration Theory (ART) (Kaplan, 1995; Hartig et al., 1991). However, previous empirical studies often relied on isolated questionnaire survey or single images as assessment methods and data resources (Burmil et al., 1999; Lindal and Hartig, 2013). Such approaches fail to consider the influence of spatial structure (i.e., connections within the scene) and geographic relevance (i.e., impacts outside the scene) on restorative quality. Consequently, the conclusions and empirical evidence from these studies may be limited and isolated. Thus, it is crucial to investigate these impacts to gain a comprehensive understanding of restorative environments.

Spatial structure, understood as the visual relationship between physical entities, has been linked to the layout of streets (Ashihara, 1986), landscape design (Cullen, 2012), and land use function (Zube, 1987), which in turn impact human perception (Lynch, 1984). Just as the first law of geography stated: "Everything is related to everything else, but near things are more related than distant things" (Tobler, 1970). Spatial structure therefore involves complicated interrelationships between these elements and their different forms could have entirely different outcomes or impacts. For example, Celikors and Wells (2022) showed that two images with similar visual elements could elicit different degrees of psychological restoration due to changes in their spatial structures. Although revealing the effects of spatial structures on psychological restoration can potentially enrich space interpretability, the rich and fine differences in spatial structures have been difficult to measure in earlier modeling work.

Street View images (SVIs) provide a valuable opportunity to capture spatial structure information from a human perspective. These images have high spatial resolution and provide sequential data (interval of 50 m) with rich urban information, which has been widely used in the study of urban form (Gong et al., 2019; Ito and Biljecki, 2021), visual perception (Biljecki and Ito, 2021), and health behavior (Fan et al., 2023; Rzotkiewicz et al., 2018). Meanwhile, using such data to assess urban restoration quality has also garnered increasing attention. Some studies have explored the potential of SVIs to predict urban restoration quality at the city level, demonstrating efficiency and accuracy (Han et al.,2023; Ma et al.,2023). Moreover, the sequential nature of these images provides opportunities to explore the intrinsic connections of spatial entities within street units (Zhang et al.,2023; Liang et al.,2023).

In recent years, methodological breakthroughs such as machine learning approaches have been broadly applied in urban studies. Among them, graph neural networks (GNNs) have shown significant advantages in capturing extrinsic relationships and predicting attributes in various fields, including traffic flow, urban population movement, and social perception (Liu and Biljecki, 2022; Zhang et al., 2023). Built upon these applications, our research aims to leverage on this approach and use graphs to represent urban spatial structures and assess their restoration quality. Specifically, we proposed a spatial-dependent GNNs approach to reveal the relation between spatial structure and restoration quality at a city scale. This approach involves two distinct types of graphs: the street-level graph and city-level graph. The street-level graphs capture visual relationships between entities, specifically the spatial structure, using sequential street view images within road units. In contrast, the city-level graph integrated urban variables, including socioeconomic, perceptual, and spatial features, to measure restoration quality, taking into account the surrounding environments. To the best of our knowledge, our proposed method of evaluating urban restoration quality and revealing the impact of spatial structure is groundbreaking.

The present study contributes significantly to the research landscape in three key ways:

- We proposed a spatial-dependent GNN method that effectively predicted the restoration quality on a city-level graph by capturing the neighboring space features.
- We verified the effect of spatial structure on the restoration quality of urban space and explored the internal spatial structure of different quality restoration spaces based on street-level graphs.
- We discovered that naturally related entities (e.g., trees) are more important than artificial entities (e.g., buildings) in the spatial structure of high-restoration quality space.

2. Related works

2.1. Restorative quality in urban environments

Attention Restoration Theory (ART) provided a framework for understanding the mental health benefits of environmental interactions. The psychological recovery and improvement of cognitive functioning after mental fatigue were referred to as *Restoration* in ART (Kaplan,1995). According to Kaplan and Kaplan (1989), to restore cognitive resources, the environment should possess the qualities of Being Away, Extent, Fascination, and Compatibility. Being away allowed us to distance ourselves from the routine fatigue of daily life. Extent refers to the expansiveness of the environment and the degree to which it invites exploration. Fascination is characterized by an environment's ability to attract our interest without consuming our attentional resources. Compatibility is the degree of alignment between an individual's needs or preferences and the environment's characteristics.

The advantages of interacting with natural environments were evident, such as reducing anxiety (Felsten, 2009), alleviating stress (Capaldi et al., 2014), improving mood (Berman et al., 2008), and performance in tasks requiring attention and working memory (Bratman et al.,2015; Stenfors et al.,2019). However, previous studies have focused primarily on the restorative benefits of natural environments and often compare them with urban environments (Hartig et al., 1991; Ulrich et al., 1991). Urban environments, on the other hand, were generally believed to deplete mental and attentional resources (Kaplan and Berman,2010; Schertz and Berman,2019). Such categorization can exacerbate the perceived contrast between "natural" and "urban" environments. Yet, urban elements such as green–blue spaces (Li et al., 2023) and walkable spaces (Han et al., 2023) have been associated with mental restoration. Some urban locations also possess restoration potential, such as art galleries (Clow and Fredhoi,2006), shopping centers, and cafes (Staats et al.,2016). However, these studies often treat urban environments as a homogeneous category, overlooking the variations in spatial structure and geographical relevance (Velarde et al.,2007).

Spatial elements can have different meanings in different contexts, and changes in spatial patterns result in variations in spatial structures. Historical studies and theories have suggested an association between human psychological perceptions and urban spatial structures (Lynch,1984; Ashihara,1986; Zube,1987). However, there is no documented evidence of a direct correlation between psychological restoration and spatial structures. Celikors and Wells (2022) proposed that similar visual properties could elicit different restoration judgments due to inherent spatial representation, highlighting the need for further investigation. Moreover, the lived experiences of city dwellers, which are shaped by the physical environment and are continuous, should be taken into account as the city is a spatial continuum and a system (Nordh et al.,2009; Lindal and Hartig,2013). Previous research often relied on non-sequential, non-geotagged data (Han et al.,2023; Ma et al., 2023), but similar and close things are more relevant in urban space, according to the first law of geography (Tobler, 1970). Thus, it's essential to consider both spatial structures and geographical relevance on an urban scale. However, traditional methods are often constrained by time and costs, focusing on specific locations and utilizing small sample sizes (Nordh et al.,2009; Lindal and Hartig,2013). This necessitates proposing a new research framework and data source.

2.2. Street view imagery in urban studies

In recent years, crowd-sourced data have been widely applied in urban studies, such as Street View Images (SVIs) (Biljecki and Ito, 2021). It contains rich urban information and is extensively used to extract environmental features (Tang and Long,2019; Wang et al., 2022; Zhou et al.,2019), build urban knowledge graphs (Zhang et al.,2023), analyze environmental health (Rzotkiewicz et al., 2018), and predict humans' perception (Zhang et al., 2018; Zhao et al., 2023). Zhang et al. (2018) pioneered the concept of establishing connections between SVIs and six human perceptions (namely, beautiful, boring, lively, depressing, wealthy, and safety) by calculating the visual perception of urban features through semantic segmentation. Subsequent research has confirmed the accuracy of this data in predicting perceptions of safety (Kang et al., 2023). Additionally, with the development of computer vision, such as image classification (Hu et al., 2020), semantic segmentation (Lauko et al.,2020), and object detection (Zhao et al.,2023), it has become efficient in analyzing large-scale SVIs to investigate urban issues. S1 Appendix, Table S1 shows the number of papers on different topics using street view images.

Moreover, SVIs have a sequential attribute and high spatial resolution, allowing researchers to study perceptual variations in cities on a large scale, in detail, and over time. For instance, Zhang et al. (2023) employed language information extracted from SVIs to capture the intrinsic and extrinsic relationship between scene entities to predict urban land functions. However, there have not been many cases in which data are used to evaluate the quality of urban restoration. In a recent study, Han et al. (2023) pioneered the exploration of a large-scale urban restoration quality assessment framework via 1,250 SVIs, achieving accurate results. Ma et al. (2023) also employed SVIs to explore the relationship between visual features and restoration quality on a campus scale. Table 1 summarizes 10 papers and methods that use street view to assess restoration quality. However, a significant limitation is that they only deal with non-sequential images (single images), which means that the relationship between scene entities on a road segment cannot be captured. Therefore, it is important to uncover the value of sequential attributes of SVIs, which entity relationships between streetscapes, represent spatial structures, and can further explore their relationship to restorative perception.

2.3. Spatially dependent graph neural networks

Graph Neural Networks (GNNs) have been shown to outperform traditional models in several areas, such as traffic flow prediction, urban population mobility, and social sensing (Liu and Biljecki, 2022). They are capable of handling structured non-Euclidean data, extracting spatial features from graphs for efficient learning (Zhang et al., 2021; Yao et al., 2021), and capturing the spatial dependency and heterogeneity of urban features (Liu et al., 2023; Liang et al., 2023). According to the first law of geography, city regions within a specific range may become increasingly similar due to the strong correlation between urban scenes and their neighboring areas (Tobler, 1970). Previous studies have found that the spatial relationship between neighbors, represented by graph theory, can identify high-level features. Zhang et al. (2023) constructed a city knowledge graph containing urban geographic information using SVIs and verified the feasibility and accuracy of GNNs in representing urban spatial structures.

GNNs can learn the deep representation of spatial relationships between adjacent scenes through aggregation algorithms, where each node can aggregate features from its neighbors (Defferrard et al., 2016). Based on the foundational principles of GNNs, several derived models have been developed, such as Graph Isomorphism Network (GIN) (Xu et al., 2019), Graph Convolutional Networks (GCN) (Kipf and Welling, 2016), Graph Attention Networks (GAT) (Veličković et al., 2017), and SAGE (Hamilton et al., 2017). Thanks to the powerful data organization capability and the ability to handle non-Euclidean data structures, GNNs can integrate various modalities of urban data into graph neural networks for downstream tasks. Examples of such data include SVIs (Liu et al., 2023), Points of Interest (POI) (Xu et al., 2022), land use (Liang et al., 2023), and social media data (Liu and De Sabbata, 2021), achieving state-of-the-art performance. For instance, Xu et al. (2022) combined visual features of cities with POI data for urban scene classification, improving the precision by 13 % compared to traditional methods. Although GNNs have the powerful ability to handle various urban tasks, there has been no research applying them to the study of the quality of urban restoration. Thus, our study is pioneering.

3. Methods and data

In this study, we proposed a graph-based framework to identify where had the highest or lowest restoration potential on city road units and explore what are the most important features that contributed to psychological restoration in the urban graph. The research framework consists of five parts (Fig. 1): 1) extracting features and embedding spatial structure (i.e., it constructed at the street level); 2) aggregating the potential urban features into a city-level graph, which was created based on OpenStreetMap (OSM); 3) labeling and enhancing dataset; 4) training and evaluating different types of classification models; and 5) conducting an overall analysis. This section will provide a detailed introduction to parts 1–4, and part 5 will be analyzed in the results section.

Our research area is within the third ring road of Wuhan (S1 Appendix, Fig. S1), a large city with a population of 11.21 million in Hubei province, which serves as a representative sample of the central region of China. For this study, we collected the OpenStreetMap (OSM) road network consisting of 5,075 road segments, and 64,750 panoramic SVIs sampling locations at 50-meter intervals were collected via Baidu Maps API from July 2022 to June 2023.

Table 1
Statistics of applied methods of street view image in restoration quality research.

Number	Methods						Research article
	Visual proportion	Object detection	Scene category	Visual feature	Spatial structure	other	
1	√						(Han et al., 2023)
2						√	(Guo et al., 2023)
3	√	√		√			(Ma et al., 2023)
4	√						(Wu et al., 2024)
5	√					√	(Chen et al., 2022)
6	√						(Zhao et al., 2020)
7		√	√				(Yin et al., 2022)
8	√						(Helbich et al., 2021)
9	√						(Meng et al., 2023)
10	√						(Hao et al., 2024)

Note: The symbol "×" indicates that not used, and "√" means the method is used.

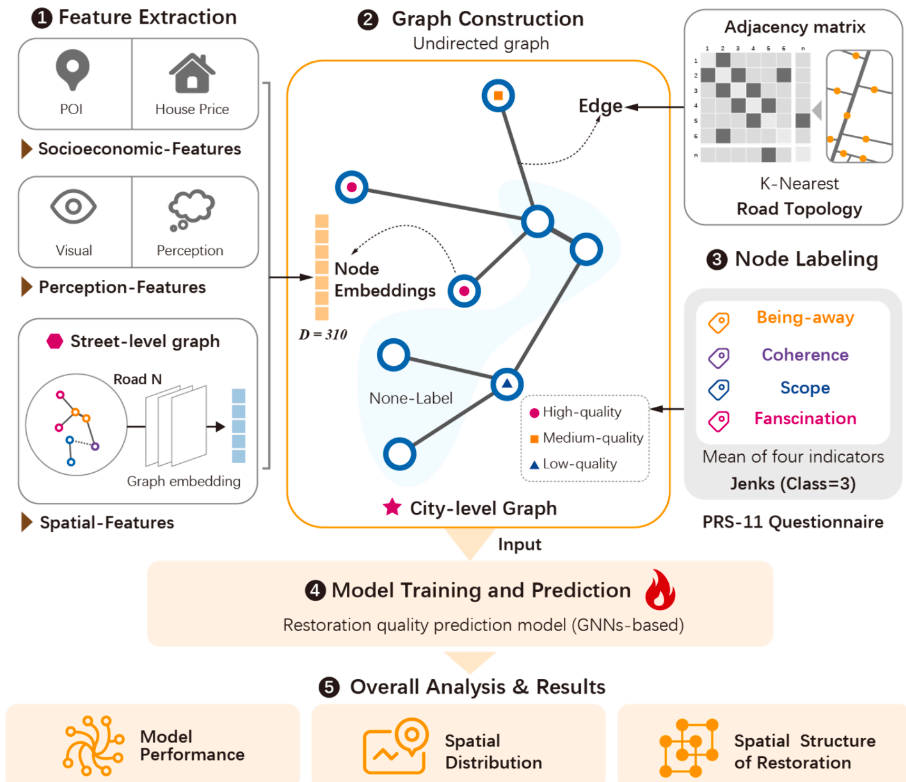


Fig. 1. Research framework. Our research consisted 5 parts: 1) extracting urban features related to psychology restoration, including socioeconomic, perceptional, and spatial features; 2) constructing a city-level graph by aggregating urban features into road units as nodes and road topology as edges; 3) labeling nodes by using the PRS-11questionnaire to score and label each road unit; 4) training the model and making prediction by treating the city-level graph as input to predict urban restoration quality through GNNs-based models; and 5) conducting an overall analysis, which including evaluating model performance, examining spatial distribution, and spatial structure of restoration quality.

3.1. Extracting features and embedding spatial structure

Our study categorized urban features into three types derived from various sources of urban data sources: 1) perceptual features (i.e., SVIs were used to infer human perception, which included object quantity, semantic proportion, and perceptual scores), 2) spatial features (i.e., SVIs were utilized to construct street-level graphs and obtain these embedded features), and 3) socioeconomic features (i.e., calculated from POI and housing prices). S1 Appendix, Table S2 shows the summary statistics for all variables.

Perceptual features. Urban space quality evaluations often extract physical components from SVIs, such as enclosure, greenery, openness, and safety (Tang and Long, 2019; Zhou et al., 2019). Previous studies have indicated that shallow visual features (i.e., pixel level) and deep visual features (i.e., semantic and object level) of images can affect perceptions of attention restoration (Ibarra et al., 2017; Celikors and Wells, 2022; Valtchanov and Ellard, 2015). We used OpenCV to calculate the pixel-level information of street view images (Zhao et al., 2023). For the extraction of deep visual features, including semantic segmentation at the semantic level to compute the proportion of physical elements and object detection at the object level to calculate the number of entities, we employed MaskFomer (Cheng et al., 2021) and DETR (Carion et al., 2020). MaskFomer is capable of segmenting 150 object categories, and DETR can detect 90 object categories, which have state-of-the-art performance. We used these models to extract visual features from SVIs.

Perceptions of cities also have an impact on the restoration of attention. More aesthetically appealing places tend to be more attractive, aligning with Kaplan's concept of a fascinating space that can restore attention resources (Berman et al., 2008). For the evaluation of perceptual scores, we used the Place Pulse 2.0 dataset, which includes 110,988 images from 56 cities in 28 countries, with six labels: depressing, boring, beautiful, safe, lively, and wealthy (Dubey et al., 2016). Based on examining its initial version, no significant cultural or individual preference biases were found, indicating its feasibility in global research (Salesses et al., 2013). We used a pre-trained model provided by Yao et al. (2019) to predict the perceptual scores of 64,750 SVIs in Wuhan City. Table 2 summarizes the models and algorithms used for feature extraction, and the dimension of all perception features is $D = 251$.

Spatial features (street-level graph). The spatial relationships between entities exhibit both intrinsic and extrinsic associations, resulting in a strong geographic relevance between similar elements in different urban spaces (Kang et al., 2018; Zhang et al., 2023). To effectively capture the structural relationships within urban scenes, we utilized sequential street views encompassed within each road unit (i.e., road segments). This approach enables us to capture and establish relationships between images, for example, identifying the same buildings in adjacent street views.

As shown in Fig. 2, we first created a spatial graph of urban scene entities. A panoptic segmentation model named Mask2Fomer (Cheng et al., 2022) was employed to capture the semantic and object urban

Table 2
Summary of feature extraction models and algorithms.

Features	Model	Dataset	Variables
Pixel-level	OpenCV	–	5 categories
Object-level	DETR (Carion et al., 2020)	COCO 2017	90 categories
Semantic-level	MaskFomer (Cheng et al., 2021)	ADE20K	150 categories
Perception	ResNet (Yao et al., 2019)	Place Pulse 2.0	6 categories

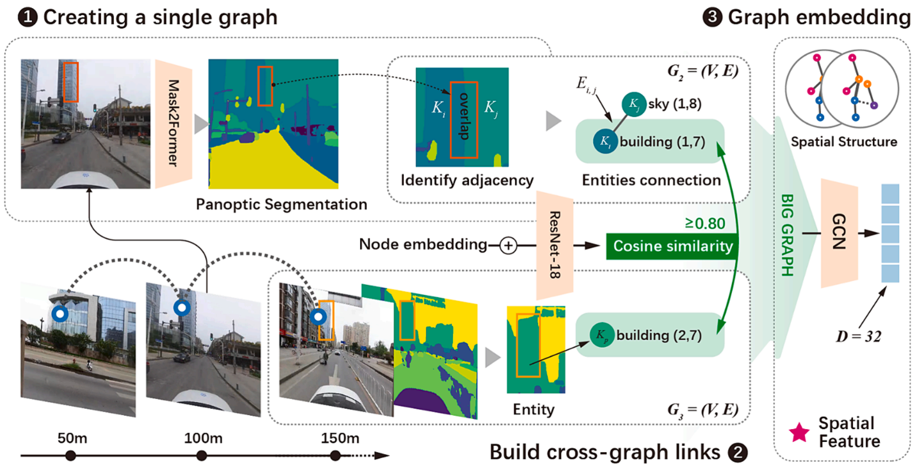


Fig. 2. Embedding the spatial structure. First, a spatial graph of entities was constructed by inferring the adjacency relationships between pixels of the entities. Second, cross-graph links were established by calculating the cosine similarity between corresponding entities (i.e., same category) in two graphs. ResNet-18 was employed for entity embedding. Finally, treating roads as the smallest units, GCN was utilized to transform the street structure into vectors.

features, which model trained on the Cityscape dataset including 19 urban-related categories. Compared to traditional segmentation models, Mask2Fomer had an ability to not only identify the instance features but label a different ID for each object. This means that we can easily calculate the relationships between entities by judging the adjacent pixels. Based on these results, we construct an undirected graph $G = (V, E)$ with K nodes, where V represents the set of nodes and E represents the set of edges. Specifically, for each node categories i, we identify the pixel edges of nodes with predicted categories as K_i. We then expand these pixel edges around two pixels and check if there is an overlap between the pixel edges of any two nodes. If there is an overlap or inclusion, we consider the nodes to be adjacent and add an edge (K_i, K_j) to the edge set $E_{i,j}$, indicating the adjacency between node K_i and node K_j. In the end, we acquired a spatial graph of urban entities for each SVI.

Next, sequential street view images from the same street were integrated to represent the spatial structure of the smallest units (road segments). Specifically, we used the ResNet-18 model (He et al., 2016) to calculate the node similarity in the same categories for each spatial graph in the same spatial unit. Each node will be embedded into a 1000-dimensional tensor, and then the cosine similarity between nodes of the same category will be calculated. If the similarity between two nodes in two graphs is higher than 0.80, these nodes will serve as bridges to connect the two graphs. Even though a building was detected in adjacent sequential street views (i.e., it had three or more segmentations), there was always a connection to the next one representing one building entity. Additionally, each integrated graph was embedded in a multi-dimensional vector ($D = 32$) using the GCN, representing the spatial structure of each road unit. The GCN utilized a two-layer convolution approach with hidden layer sizes of 128 computational units. The code is shared on GitHub.[1]

Socioeconomic features. In addition, we used POI data to examine service indicators and housing price data to examine the economic conditions of each road unit, these features have been proven to be related to the quality of environmental restoration (Subiza-Pérez et al., 2021; Samus et al., 2022; Luttik, 2000). Specifically, the POI data includes 23 types of functions, ranging from finance service to motorcycle service, which can be found in S1 Appendix, Table S1. We collected these data in June 2023 and used the Spatial Join tool to calculate the average POI density of each type and the average housing price (in yuan per square meter) in each road unit (i.e., the average value of a buffer zone based on the road centerline with a radius of 25 m). The dimension of these features is $D = 24$.

3.2. Features aggregation and graph construction (city-level graph)

The study area consists of 5,075 OSM road units, with the midpoint of each road serving as a node ($N = 5{,}075$) on the undirected graph, due to the road network shapes urban functions and traffic as the skeleton of the city (Hong and Yao, 2019). We aggregate the urban features mentioned in section 3.1 into the road units, and based on the road centerline, we created a buffer zone with a radius of 25 m, which can cover most of the urban road width. The Spatial Join tool was used to map the perceptual feature ($D = 251$), spatial features ($D = 32$), and socioeconomic features ($D = 24$) to the proper road units. The vector dimension of each road unit is $D = 310$.

Simultaneously, we employ the topology of the road network to create a spatial weight matrix, representing relationships between adjacent roads. Specifically, we use a matrix $n \times n$ (n is the number of all road segments), and A to express the adjacency relationships between the roads. If there is an adjacency relationship between streets i and j, $A_{i,j}$ is assigned a value of one; otherwise, $A_{i,j}$ is assigned a value of zero. We determine the road connection using the K-Nearest method (K=5), which signifies their adjacency relationships (i.e., 25,375 neighboring relationships). This method prevents dangling roads but may also categorize certain non-intersecting roads as nearby (Zhu et al., 2020).

3.3. Labeling and enhancing dataset

In this section, a graph-based dataset was created specifically for predicting the quality of urban restoration. The restoration quality prediction task was formalized as a three-class classification at a city-level graph, where each node represented a road. The restoration quality categories were labeled as high, medium, or low. In addition, to label the restoration quality categories for each road (i.e., node), we evaluated a substantial number of SVI samples using the Perceived Restorativeness Scale-11 (PRS-11) on a platform we developed. A total of 1,115 roads were labeled, with 80 % allocated for training and the remaining 20 % used as the test dataset (Table 3). Additional details on the sampling process can be found in S1 Appendix, Fig. S2.

To evaluate the restorative quality of the sampled SVIs, a survey was conducted using the Perceived Restorativeness Scale-11 (PRS-11). The original PRS-11 consists of 11 questions, with two questions related to scope, and three questions each for being away, coherence, and fascination, as defined by Pasini et al. (2014). Following the approach of Celikors and Wells (2022), we selected the four best question

[1] https://github.com/MMHHRR/Restoration_Topology.

Table 3
Statistical analysis of restoration quality data set.

Labels	Train data	Test data
Low-quality	347	87
Medium-quality	287	72
High-quality	258	64
Proportion	**0.80**	**0.20**
Total	**892**	**223**

descriptions for each indicator in our study. The specific illustrations and definitions of each question are provided in Table 4. Participants were instructed to select the image from our developed platform (Fig. 3) that best matched the description of the problem. Participants were given the option to click the Left or Right button, and the selected image would gain one point accordingly. If participants selected Neither or Both, neither image would gain any points, or both images would gain one point. Each image was evaluated for a duration of 30 s, and each image received a minimum of 20 evaluations (Celikors and Wells, 2022).

During a week-long online survey, we collected evaluation results from 120 participants, consisting of 70 women (average age = 24.658) and 50 men (average age = 26.059). The ethical aspects of the experiment were reviewed and approved by our university's institutional review board. To calculate the average score of the four indicators, representing the comprehensive restoration quality of the SVIs, we employed the Trueskill method (Herbrich et al., 2006). The scores were calculated on a scale ranging from 0 to 1. Subsequently, the final results were categorized into three classes as mentioned earlier, utilizing the Jenks Natural Breaks method. These class labels were assigned to the nodes of the city-level graph (Table 3).

3.4. Model training and evaluation

In this study, we used the Graph Isomorphism Network (GIN), which has demonstrated excellent performance on various benchmark datasets and graph tasks (Xu et al., 2019). A significant feature of GIN lies in its accurate bounding of the expressiveness of GNNs. The key equation for GIN updating node representations is as follows (Formula 1):

$$h_v^{(k)} = MLP^{(k)}\left((1+\varepsilon)^*h_v^{(k-1)} + \Sigma_{u\in\mathcal{N}(v)}h_u^{(k-1)}\right) \tag{1}$$

where $h_v^{(k)}$ represents the feature of node v at layer $k.\varepsilon$ is a learnable parameter, used to build self-loops for the central node. $\mathcal{N}(v)$ is the neighboring nodes for node v. $MLP^{(k)}$ denotes the Multi-Layer Perceptron at layer k, consisting of learnable parameters. GIN is one such example among many maximally powerful GNNs while being simple. For model structure, refer to S1 Appendix, Fig. S3.

In addition, we conducted several experiments to compare our approach with other graph models, including the Topology Adaptive Graph Convolutional Network (TAGCN) (Du et al., 2018), Attention-based Graph Neural Network (AGCN) (Thekumparampil et al., 2018), SAGE (Hamilton et al., 2017), Graph Attention Network (GAT) (Veličković et al., 2017), Simplifying Graph Convolutional Networks (SGC) (Wu et al., 2019), and Graph Convolutional Networks (GCN) (Kipf and Welling, 2016). In our study, the GCN model was configured with two different types of two-layer architectures. GCN 1 represented the hidden channels set to 32 and 16, while GCN 2 represented the hidden channels set to 64 and 32. Each of the graph-based models trained 500 epochs.

Furthermore, to verify the spatial-dependent ability possessed by GNN-based models, we also compared five machine learning models without any spatial weight, including Random Forest (RF) (Breiman, 2001), Decision Tree (DT) (Hastie et al., 2009), K-Nearest Neighbors (KNN) (Peterson, 2009), Radial Basis Function Support Vector Machines (RBF SVM) (Schwenker et al., 2001), and Gradient Boosting Decision Trees (GBDT) (Friedman, 2001). Additional parameter and setting details were provided in S1 Appendix, Table S3. The accuracy score and F1 score (Hossin and Sulaiman,2015) were used to assess the model performance, and GNNExplainer (Ying et al., 2019) was used for the interpretable analysis of the graph model, which made the model prediction process more understandable (for detailed methodology, please refer to the S1 Appendix, Fig. S4).

Table 4
Description of recovery quality evaluation questions (based on PRS-11 questionnaire).

Restorative quality	Definition	Function	Questions
Being-away	Absence of some aspect of life that is ordinarily present and presumably not always preferred.	Elimination of everyday distractors.	"To stop thinking about the things that I must get done I like to go to places like this."
Coherence	Interrelatedness of the immediately perceived elements.	Sufficient connectedness makes it possible to build a mental map and make sense of the environment.	"It is easy to see how things are organized in this place."
Scope	Constitution of a larger whole.	Sufficient scope makes building a mental map worthwhile by facilitating curiosity and a desire to be involved in the environment.	"This place is large enough to allow exploration in many directions."
Fascination	Mind-wandering via involuntary bottom-up attention.	Reducing mental fatigue by shifting from voluntary to involuntary attention.	"In this place, my attention is drawn to many interesting things."

Note: Definition and function of each restorative quality and measurement based on (Pasini et al. 2014, Celikors and Wells2022).

4. Results

4.1. Spatially dependent model performance and prediction results

According to the results presented in Fig. 4, the GIN demonstrated the highest performance in terms of classification (Acc = 0.749, F1 = 0.740) compared to other graph-based models. It was followed by GCN2 (Acc = 0.733, F1 = 0.711) and GCN1 (Acc = 0.715, F1 = 0.705) in terms of intra-group comparisons. On the other hand, among the traditional methods mentioned in Table 5, the RF achieved the best classification performance (Acc = 0.567, F1 = 0.440), followed by RGF SVM (Acc = 0.562, F1 = 0.407) and GBDT (Acc = 0.561, F1 = 0.405). These results indicate that graph-based models, which take into consideration spatial dependence, have higher classification accuracy compared to traditional methods. This result supports our hypothesis that incorporating spatial dependencies in graph networks can enhance the accuracy of the restoration measurement task (Zhang et al., 2023; Liu et al., 2023; Liu and Biljecki, 2022). Additional details were provided in S1 Appendix, Table S4, and Fig. S5.

The classification performance of the GIN model was notably satisfactory when compared to the other models, as evident from the analysis of the confusion matrix and T-SNE results (Fig. 5). The confusion matrix demonstrates the GIN model's capability to accurately identify all three categories of urban restoration quality in the test dataset. However, there were instances where certain roads within each category were misclassified as belonging to other categories. For instance, during the prediction of high-quality roads, 10 roads were incorrectly classified as low-quality.

We attribute this misclassification to the presence of data imbalance,

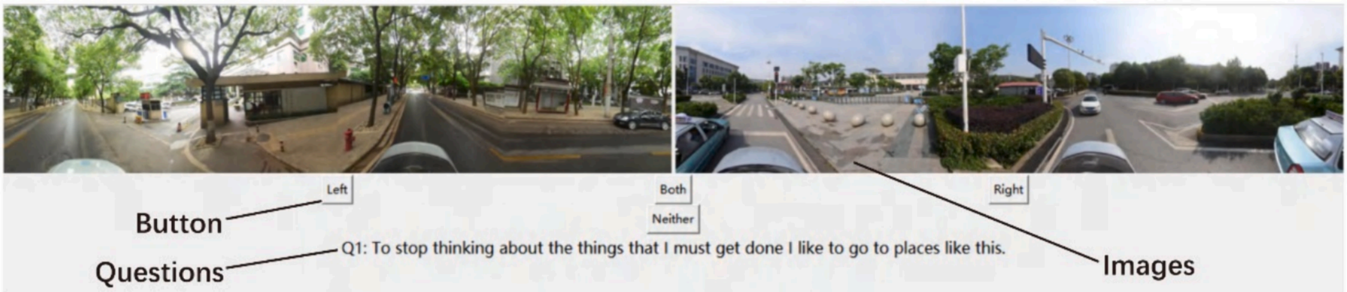


Fig. 3. Interface of urban restoration evaluation platform.

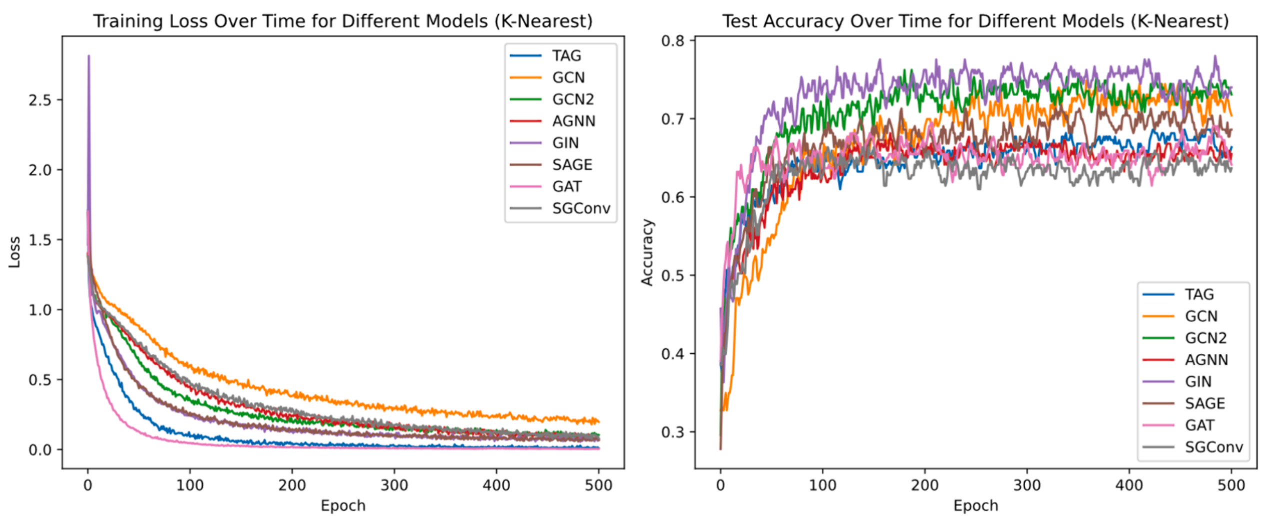


Fig. 4. Model performance. The training loss curves (left) and accuracy curves (right) of eight graph-based models were analyzed in this study. The spatial weight used for these models was K-Nearest, and the training process was conducted for 500 epochs.

Table 5
Model prediction performance.

Model	Spatial Weight	Accuracy (%)↑	F1 Score (%)↑	Time (s)↓
TAG (Du et al., 2018)	○	0.664	0.663	26.790
GCN 1 (Kipf and Welling, 2016)	○	0.715	0.705	8.570
GCN 2 (Kipf and Welling, 2016)	○	0.733	0.711	9.700
AGNN (Thekumparampil et al., 2018)	○	0.656	0.654	10.320
GIN (Xu et al., 2019)	○	**0.749**	**0.740**	11.260
SAGE (Hamilton et al., 2017)	○	0.687	0.686	7.620
GAT (Veličković et al.,2017)	○	0.657	0.656	42.720
SGC (Wu et al., 2019)	○	0.635	0.635	13.280
RF (Breiman, 2001)	×	0.567	0.440	20.790
DF (Hastie et al., 2009)	×	0.529	0.482	0.350
KNN (Peterson, 2009)	×	0.504	0.434	**0.160**
RGF SVM (Schwenker et al., 2001)	×	0.562	0.407	2.040
GBDT (Friedman, 2001)	×	0.561	0.405	80.370

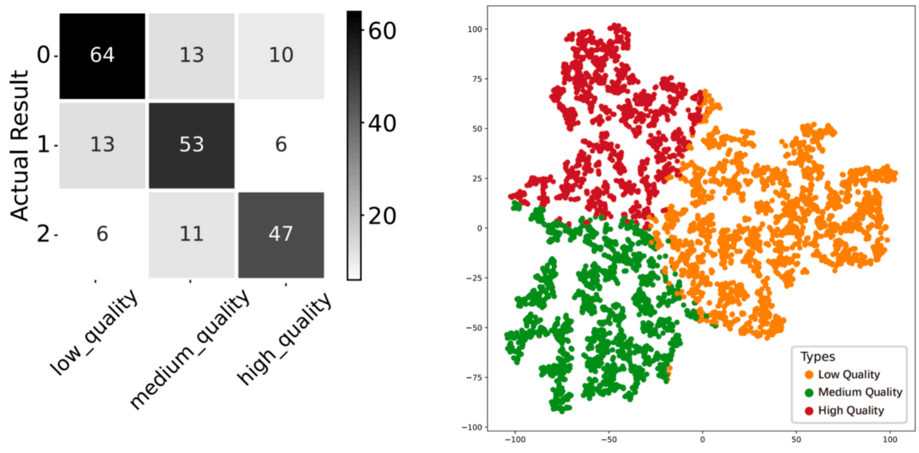
Note: GCN 1 represented the hidden channels set to 64 and 32, while GCN 2 represented the hidden channels set to 32 and 16. The symbol "○" indicates that spatial weight was considered, while "×" represents the absence of spatial weight.

as the medium and low-quality data instances are more prevalent. Moreover, we investigated to determine whether the model effectively learns visual classification features. By reducing the dimensionality of the GIN model's output layer, we discovered that the model possesses sufficient capability to classify restoration quality based on city features. Consequently, subsequent studies will be based on the GIN model.

4.2. Spatial distribution of restorative perception of urban streets

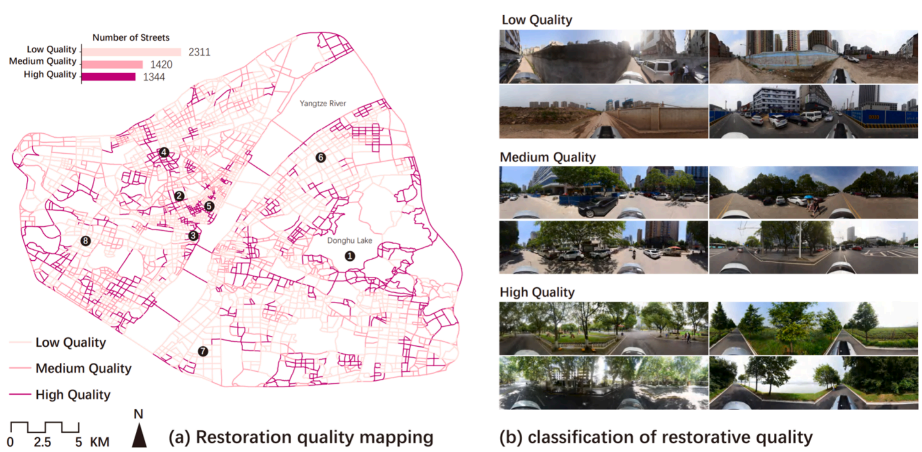
As shown in Fig. 6a, we mapped the predicted results using the GIN model (highly saturated lines indicate high restoration quality, otherwise the opposite). Within the third ring road of Wuhan, there were 1,344 high-quality restoration roads, 1,420 medium-quality restoration roads, and 2,311 low-quality restoration roads. High-quality restoration spaces have shown an aggregated pattern throughout the city, which may be related to the predictive model we used. GIN models consider not only their features but also the features of neighboring nodes.

High-quality restoration spaces were most prominent in 1, 2, 3, 4, and 5 areas (Fig. 6a). Among them, area 1 is the largest freshwater lake in Wuhan, namely Donghu Lake, with charming waterfront spaces. Waterfronts exhibited more significant restoration abilities, producing a cooling sensation in summer and providing a comfortable environment for sightseeing (Burmil et al., 1999). Interestingly, other high-quality restoration spaces were closely related to surrounding parks or green spaces, such as Zhongshan Park around 2, Hanyang Jiangtan Park around 3, and Houxianghe Park around 4. Urban parks or green areas provide abundant natural resources for attention restoration, and the benefits of interacting with nature are evident (Dadvand et al.,2015; Engemann et al., 2019). There were also means that urban greenery infrastructure had impacted the spatial structure of the neighboring streets. Surprisingly, although located in a high-density commercial area, 5 was still predicted to have high restoration quality. We believe



(a) Confusion matrix (b) Prediction results (T-SNE)

Fig. 5. Analysis of GIN model prediction results (spatial weight: K-Nearest, epoch = 500). (a) Confusion matrix of model prediction results. (b) T-SNE to reduce the dimension of the output layer of the GIN model.

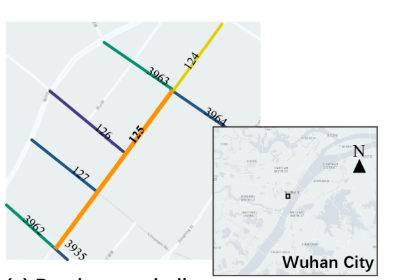


(a) Restoration quality mapping (b) classification of restorative quality

Fig. 6. Mapping distribution of Wuhan city spatial restoration quality. (a) Restoration quality mapping in Wuhan. (b) Classification of restoration quality, Source of the map: © OSM contributors.

this is related to the surrounding shopping centers and historic districts, which have been proven to promote attention restoration (Staats et al.,2 016; Fornara et al., 2009). Furthermore, we found that low-quality restoration spaces were mainly concentrated in residential areas, such as 6, 7, and 8. Residential areas occupy the largest proportion of land use in Wuhan, which is the most popular city in central China. However, monotonous residential spaces can easily feel boring, and tall buildings reduce the view distance, affecting the quality of restoration (Lindal and Hartig, 2013; Zhang et al., 2018).

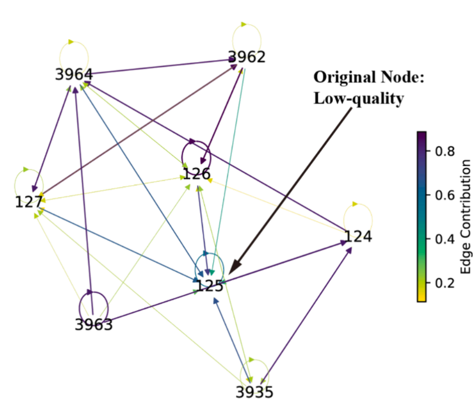
Based on the predicted results, we categorized the street scenes



(a) Road network diagram

124→125	126→125	127→125	3962→125
0.372	0.747	0.645	0.408
3963→125	3964→125	3935→125	
0.290	0.654	0.676	

(b) Nodes contribution



(c) Node relationship between its surrounding spaces

Fig. 7. Node relationship analysis based on spatially dependent GIN model. (a) Randomly selected road network diagram. (b) Nodes contribution for target node. (c) Node relationship between its surrounding entities.

according to different restoration qualities (Fig. 6b). We found that low-quality spaces lacked green vegetation and had high building densities. In medium-quality spaces, there was a higher presence of shrubs, which, to some extent, increased the restoration quality of the space. In high-quality spaces, the proportion of comfortable roads and the abundance of natural urban landscapes added to the attractiveness and charm of the space.

In addition, to capture the spatial dependencies in the GNN-based model, we employed GNNExplainer for interpretable analysis (mentioned in Section 3.4). A road unit was randomly chosen to elucidate the relationship between this particular space and its surrounding environment (Fig. 7). The analysis revealed that when street 125 was predicted as a low-quality restoration space, the adjacent spaces 126 (0.747) and 3935 (0.676) had the most significant impact on it, followed by space 3964 (0.654). In summary, the GNN-based model not only effectively considered the influence of neighboring spaces, but also provided a holistic perspective.

4.3. Relationship between spatial structure and restoration quality

As mentioned above, no studies have confirmed the role of spatial structure in restorative environments. We performed ablation experiments based on the GIN model, as shown in Table 6. It can be observed that in Experiment 1, considering all three classes of urban features together, the best classification performance was obtained (Acc = 0.749, F1 = 0.740). However, in Experiment 2, when we removed the spatial features only, the classification performance decreased significantly (Acc = 0.668, F1 = 0.667). To further confirm the impact of spatial features on the prediction results, experiments 3 and 4 were conducted to remove the perceptual and socioeconomic features, respectively. The final results confirmed that spatial features significantly affect the classification performance of the model (Experiment 3: Acc = 0.708, F1 = 0.705; Experiment 4: Acc = 0.722, F1 = 0.718). Experiment 5, which only considered spatial features, also exhibited good classification performance (Acc = 0.704, F1 = 0.703). In conclusion, the spatial features increased the prediction accuracy and significantly affected the model performance, suggesting that the spatial structure had a significant impact on the spatial restoration ability. Additional details can be found in the S1 Appendix and Table S5.

To further investigate the impact of spatial structure on the quality of urban restoration, we used GNNExplainer as a tool to open the "black box" of model prediction. Based on Experiment 5, we directly used a spatial graph (without embedding) as input to predict the restorative quality using GIN mode (i.e., graph classification). Thus, we were able to identify which entities significantly influence spatial restoration quality and determine their importance in the street-level graph. Fig. 8 presents the results obtained from two selected research areas in Wuhan city, representing spaces with low and high restorative quality, respectively. The results showed that artificial entities, such as sidewalks (0.543), fences (0.531), and buildings (0.526), had a higher contribution to the low-quality space. Conversely, entities such as the sky (0.503) and vegetation (0.491) demonstrated a greater effect on high-quality space, aligning with previous studies (Ma et al., 2023).

In addition, we utilized the betweenness centrality indicator to measure the importance of entities within each street-level graph. Nodes (or entities) with higher betweenness centrality values were regarded as having a greater influence or control over visual perception in the network (Brandes, 2001). Fig. 9 illustrates the calculation of the top 100 road units predicted as corresponding classes The results consistently revealed that spaces with high restorative quality typically exhibited higher betweenness centrality values for natural entities such as vegetation (0.144), terrain (0.129), and sky (0.121). These findings further support the notion that these natural elements play a significant role in enhancing the restorative quality of urban spaces.

5. Discussion

This study introduces a spatial-dependent GNN approach to predict urban restoration quality and reveal the relation between spatial structure and restoration space. By embedding spatial unit graphs through multiple sequential SVIs, the influence of spatial structure on restorative quality and the structural heterogeneity of restorative space are explored. Our findings suggest that the spatially dependent GNNs, through learning from the fusion of various features and geographic relationships, not only take into account the environment of specific locations but also provide a holistic perspective. This approach facilitates a comprehensive understanding of restoration characteristics at an urban scale, maximizing the consideration of interdependency between spaces. Therefore, our study fills a gap in understanding how spatial structure influences restoration quality.

Furthermore, we found that spatial structure strongly determines the restorative quality of urban environments. In spaces with low restoration quality, the importance of non-natural entities is significantly higher, while in spaces with high restoration quality, the opposite is true, with natural entities becoming more important. This is consistent with previous research, which shows a positive correlation between natural elements and restoration, together with various health benefits (Capaldi et al., 2014; Schertz and Berman, 2019).

In addition to the influence of spatial structure, our research has also revealed the distribution characteristics of high-quality restoration spaces in cities. Notably, urban waterfront spaces emerge conspicuously. Water views are considered positive restorative visual components, and people tend to walk or cycle in areas with abundant water views to reduce stress (Massoni et al., 2018). Furthermore, the characteristics of natural water can seamlessly integrate into the surrounding natural landscapes, enhancing aesthetic experiences and restoring attentional resources (Markevych et al., 2017; Roe et al., 2019). Therefore, waterfront landscape types should be given priority in urban design and redevelopment. Moreover, urban green spaces are closely associated with high restoration quality, and numerous studies have demonstrated the psychological benefits of urban green spaces, such as urban parks (Nordh et al., 2011, 2009). This research thus provides further evidence supporting the restorative potential of cities.

Additionally, while SVIs are emerging as an important data source in urban research (Tang and Long, 2019; Biljecki and Ito, 2021), there have been limited studies using them to investigate spatial restoration quality on an urban scale. SVIs demonstrate significant advantages in research. First, it has wide coverage, fast updates, and precise geographic coordinates. Second, it has a wealth of visual and spatial information, and when used to predict human perception, there is a minimal amount of

Table 6
Ablation experiment results based on GIN model.

Features	Experiment 1	Experiment 2	Experiment 3	Experiment 4	Experiment 5
Socioeconomic	○	○	○	×	×
Perceptual	○	○	×	○	×
Spatial	○	×	○	○	○
Accuracy (%)	0.749	0.668	0.708	0.722	0.704
F1 Score (%)	0.740	0.667	0.705	0.718	0.703

Note: The symbol "○" indicates that the features were retained, while "×" represents the removal of these features.

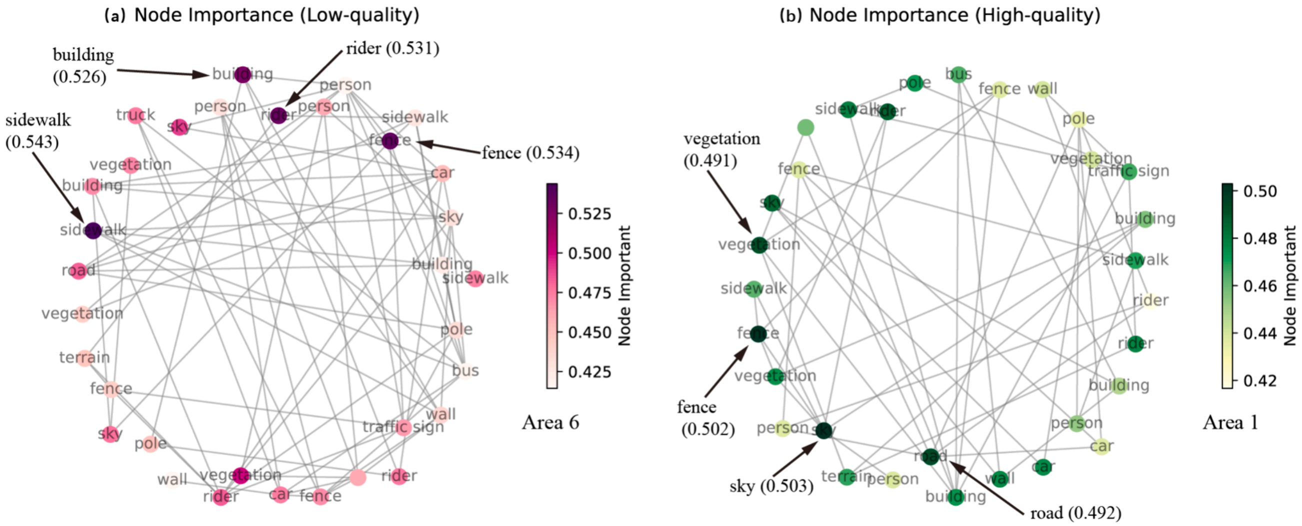


Fig. 8. Node contribution and spatial structure of different restorative spaces. (a) low quality restoration space and (b) high quality restoration space.

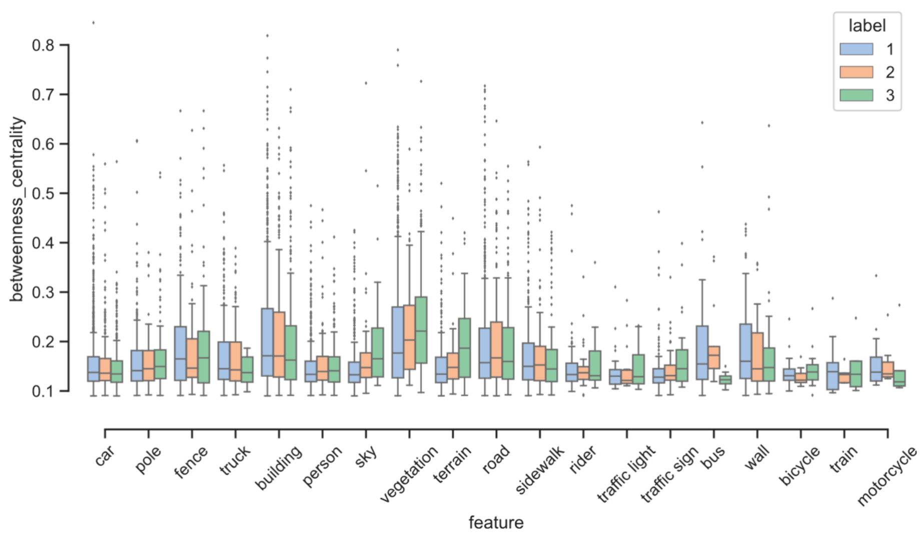


Fig. 9. The betweenness centrality of 19 types of spatial entities in three restorative spatial quality. "1″ represented low restoration quality, "2" represented medium restoration quality, and "3" represented high restoration quality.

data bias (Kang et al., 2023; Zhao et al., 2023; Ma et al., 2023). In this study, to predict urban restoration quality, a large number of SVIs were used to extract visual features and human perceptions, and embed the spatial structure in sequential scenes. The results proved that SVIs can accurately predict urban restoration quality and demonstrated the advantages of representing the spatial structure.

Simultaneously, this study has some limitations. First, urban environments are dynamic and influenced by factors such as human activities, urban functions, and traffic conditions (Kaplan and Herbert,1987; Quercia et al., 2014). Complex features, such as sound and temperature, can affect the restoration quality (Hartig et al., 2007; Qi et al., 2022; Ratcliffe, 2021). In the future, it is possible to expand to a broader range of dimensions by integrating multi-modal data or digital environments that incorporate these elements that potentially influence the quality of environmental restoration. Second, the city features captured by SVI data only reflect specific moments in urban scenes, thus exhibiting a temporal lag. Also, SVI is primarily captured from a driving perspective, lacking the perception of content from a human perspective (Biljecki and Ito, 2021). Finally, in a restoration environment, the structurally simple PRS-11 may weaken the accuracy of the assessment of environmental restoration quality. Future research could consider incorporating more diverse attention restoration questionnaires and questions to improve the precision of the results.

6. Conclusion

The long-standing discussion on the relationship between restorative quality and the physical environment lacks research on the impacts caused by diverse spatial structures and is scarce in efficient ways for measuring on an urban scale. Our study proposed a spatial-dependent GNN approach for solving these questions, which includes two types of graphs: street and city levels. This study made three contributions. First, we proposed a spatial-dependent prediction method for measuring urban restoration quality by capturing road topology relationships using graph neural networks and aggregating contextual features of cities as a city-level graph. Second, we used a novel graph approach to reveal spatial structure effects among different restoration qualities by capturing the intrinsic and extrinsic relationships between entities through sequential SVIs. Third, the study highlighted the greater significance of naturally relevant entities (e.g., trees) compared to artificial entities (e.g., buildings) in relation to high restoration quality, thereby enhancing the understandability of restorative spatial features. Overall, this study provides insight into healthy city construction, improves the interpretability of urban restoration spaces, and can be applied further to the design of healthy medium-scale spaces, such as communities or parks.

CRediT authorship contribution statement

Haoran Ma: Conceptualization, Methodology, Data analysis, Writing – Original draft preparation, Visualization, Investigation. **Yan Zhang:** Conceptualization, Methodology, Visualization, Investigation, Supervision, Writing – review & editing. **Pengyuan Liu:** Writing – review & editing. **Fan Zhang:** Visualization, Investigation, Writing – review & editing. **Pengyu Zhu:** Conceptualization, Methodology, Supervision, Writing – review & editing, Project manager, Direction guidance.

Data availability

The authors do not have permission to share data.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.landurbplan.2024.105171.

References

Akpınar, A. (2021). How perceived sensory dimensions of urban green spaces are associated with teenagers' perceived restoration, stress, and mental health? *Landscape and Urban Planning, 214*, Article 104185.

Ashihara, Y., 1986. The aesthetic townscape.

Berman, M. G., Jonides, J., & Kaplan, S. (2008). The cognitive benefits of interacting with nature. *Psychological science, 19*, 1207–1212.

Biljecki, F., & Ito, K. (2021). Street view imagery in urban analytics and gis: A review. *Landscape and Urban Planning, 215*, Article 104217.

Brandes, U. (2001). A faster algorithm for betweenness centrality. *Journal of mathematical sociology, 25*, 163–177.

Bratman, G. N., Daily, G. C., Levy, B. J., & Gross, J. J. (2015). The benefits of nature experience: Improved affect and cognition. *Landscape and Urban Planning, 138*, 41–50.

Breiman, L. (2001). *Random forests. Machine learning, 45*, 5–32.

Burmil, S., Daniel, T. C., & Hetherington, J. D. (1999). Human values and perceptions of water in arid landscapes. *Landscape and Urban Planning, 44*, 99–109.

C.A. Capaldi R.L. Dopko J.M. Zelenski The relationship between nature connectedness and happiness: A meta-analysis. Frontiers in psychology 2014 976.

Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., Zagoruyko, S., 2020. End-to-end object detection with transformers, in: Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part I 16, Springer. Pp. 213–229.

Celikors, E., & Wells, N. M. (2022). Are low-level visual features of scenes associated with perceived restorative qualities? *Journal of Environmental Psychology, 81*, Article 101800.

Cetin, M., Agacsapan, B., Cabuk, S. N., Senyel Kurkcuoglu, M. A., Isik Pekkan, O., Baran Argun, E., Dabanlı, A., Kucukpehlivan, T., Yilmazel, B., & Cabuk, A. (2021). Assessment of the ecological footprint of ondonir technical university–iki eylul campus. *Journal of the Indian Society of Remote Sensing, 49*, 2311–2327.

Cheng, B., Misra, I., Schwing, A. G., Kirillov, A., & Girdhar, R. (2022). Masked-attention mask transformer for universal image segmentation. In *in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 1290–1299).

Cheng, B., Schwing, A., & Kirillov, A. (2021). Per-pixel classification is not all you need for semantic segmentation. *Advances in neural information processing systems, 34*, 17864–17875.

Chen, C., Li, H., Luo, W., Xie, J., Yao, J., Wu, L., & Xia, Y. (2022). Predicting the effect of street environment on residents' mood states in large urban areas using machine learning and street view images. *Science of The Total Environment, 816*, Article 151605.

Clow, A., & Fredhoi, C. (2006). Normalisation of salivary cortisol levels and self-report stress by a brief lunchtime visit to an art gallery by ondon city workers. *Journal of Holistic Healthcare, 3*, 29–32.

Cullen, G. (2012). *Concise townscape*. Routledge.

Dadvand, P., Nieuwenhuijsen, M. J., Esnaola, M., Forns, J., Basagaña, X., Alvarez-Pedrerol, M., Rivas, I., López-Vicente, M., De Castro Pascual, M., Su, J., et al. (2015). Green spaces and cognitive development in primary schoolchildren. *Proceedings of the National Academy of Sciences, 112*, 7937–7942.

Defferrard, M., Bresson, X., Vandergheynst, P., 2016. Convolutional neural networks on graphs with fast localized spectral filtering. Advances in neural information processing systems 29.

Dubey, A., Naik, N., Parikh, D., Raskar, R., Hidalgo, C.A., 2016. Deep learning the city: Quantifying urban perception at a global scale, in: Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, PartI 14, Springer. Pp. 196–212.

Du, J., Zhang, S., Wu, G., Moura, J.M.F., Kar, S., 2018. Topology adaptive graph convolutional network.

Engemann, K., Pedersen, C. B., Arge, L., Tsirogiannis, C., Mortensen, P. B., & Svenning, J. C. (2019). Residential green space in childhood is associated with lower risk of psychiatric disorders from adolescence into adulthood. *Proceedings of the national academy of sciences, 116*, 5188–5193.

Fan, Z., Zhang, F., Loo, B. P. Y., & Ratti, C. (2023). Urban visual intelligence: Uncovering hidden city profiles with street view images. *Proceedings of the National Academy of Sciences, 120*, Article e2220417120. https://doi.org/10.1073/pnas.2220417120

Felsten, G. (2009). Where to take a study break on the college campus: An attention restoration theory perspective. *Journal of environmental psychology, 29*, 160–167.

Fornara, F., Troffa, R., et al. (2009). In *Restorative experiences and perceived affective qualities in different built and natural urban places. Revitalising Built Environments: Requalifying Old Places for New Uses* (pp. 1–10).

Gong, F. Y., Zeng, Z. C., Ng, E., & Norford, L. K. (2019). Spatiotemporal patterns of street-level solar radiation estimated using google street view in a high-density urban environment. *Building and Environment, 148*, 547–566.

Guo, W., Wen, H., & Liu, X. (2023). Research on the psychologically restorative effects of campus common spaces from the perspective of health. *Frontiers in Public Health, 11*.

Hamilton, W., Ying, Z., Leskovec, J., 2017. Inductive representation learning on large graphs. Advances in neural information processing systems 30.

Han, X., Wang, L., He, J., & Jung, T. (2023). Restorative perception of urban streets: Interpretation using deep learning and mgwr models. *Frontiers in Public Health, 11*, 1141630.

Hartig, T., Catalano, R., & Ong, M. (2007). Cold summer weather, constrained restoration, and the use of antidepressants in sweden. *Journal of Environmental Psychology, 27*, 107–116.

Hartig, T., Mang, M., & Evans, G. W. (1991). Restorative effects of natural environment experiences. *Environment and behavior, 23*, 3–26.

Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction* (volume 2). Springer.

Herbrich, R., Minka, T., Graepel, T., 2006. Trueskill™: a bayesian skill rating system. Advances in neural information processing systems 19.

Helbich, M., Poppe, R., Oberski, D., van Emmichoven, M. Z., & Schram, R. (2021). Can't see the wood for the trees? An assessment of street view-and satellite-derived greenness measures in relation to mental health. *Landscape and Urban Planning, 214*, Article 104181.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778.

Hong, Y., & Yao, Y. (2019). Hierarchical community detection and functional area identification with osm roads and complex graph theory. *International Journal of Geographical Information Science, 33*, 1569–1587.

Hossin, M., & Sulaiman, M. N. (2015). A review on evaluation metrics for data classification evaluations. *International journal of data mining & knowledge management process, 5*, 1.

Hough, R. L. (2014). Biodiversity and human health: Evidence for causality? *Biodiversity and conservation, 23*, 267–288.

Hu, C. B., Zhang, F., Gong, F. Y., Ratti, C., & Li, X. (2020). Classification and mapping of urban canyon geometry using google street view images and deep multitask learning. *Building and Environment, 167*, Article 106424.

Hao, N., Li, X., Han, D., & Nie, W. (2024). Quantifying the Impact of Street Greening during Full-Leaf Seasons on Emotional Perception: Guidelines for Resident Well-Being. *Forests, 15*, 119.

Ibarra, F. F., Kardan, O., Hunter, M. R., Kotabe, H. P., Meyer, F. A., & Berman, M. G. (2017). Image feature types and their predictions of aesthetic preference and naturalness. *Frontiers in Psychology, 8*, 632.

Ito, K., & Biljecki, F. (2021). Assessing bikeability with street view imagery and computer vision. *Transportation research part C: emerging technologies, 132*, Article 103371.

Kang, J., Körner, M., Wang, Y., Taubenböck, H., & Zhu, X. X. (2018). Building instance classification using street view images. *ISPRS journal of photogrammetry and remote sensing, 145*, 44–59.

Kang, Y., Abraham, J., Ceccato, V., Duarte, F., Gao, S., Ljungqvist, L., Zhang, F., Näsman, P., & Ratti, C. (2023). Assessing differences in safety perceptions using geoai and survey across neighbourhoods in stockholm, sweden. *Landscape and Urban Planning, 236*, Article 104768.

Kaplan, R., & Herbert, E. J. (1987). Cultural and sub-cultural comparisons in preferences for natural settings. *Landscape and urban planning, 14*, 281–293.

Kaplan, R., & Kaplan, S. (1989). *The experience of nature: A psychological perspective*. Cambridge University Press.

Kaplan, S. (1995). The restorative benefits of nature: Toward an integrative framework. *Journal of environmental psychology, 15*, 169–182.

Kaplan, S., & Berman, M. G. (2010). Directed attention as a common resource for executive functioning and self-regulation. *Perspectives on psychological science, 5*, 43–57.

Keniger, L. E., Gaston, K. J., Irvine, K. N., & Fuller, R. A. (2013). What are the benefits of interacting with nature? *International journal of environmental research and public health, 10*, 913–935.

Kipf, T. N., & Welling, M. (2016). Semi-supervised classification with graph convolutional networks. *International Conference on Learning Representations*.

Lauko, I. G., Honts, A., Beihoff, J., & Rupprecht, S. (2020). Local color and morphological image feature based vegetation identification and its application to human

environment street view vegetation mapping, or how green is our county? *Geo-spatial Information Science, 23*, 222–236.
Li, H., Browning, M. H., Rigolon, A., Larson, L. R., Taff, D., Labib, S., Benfield, J., Yuan, S., McAnirlin, O., Hatami, N., et al. (2023). Beyond "bluespace" and "greenspace": A narrative review of possible health benefits from exposure to other natural landscapes. *Science of The Total Environment, 856*, Article 159292.
Liang, X., Zhao, T., & Biljecki, F. (2023). Revealing spatio-temporal evolution of urban visual environments with street view imagery. *Landscape and Urban Planning, 237*, Article 104802.
Lindal, P. J., & Hartig, T. (2013). Architectural variation, building height, and the restorative quality of urban residential streetscapes. *Journal of environmental psychology, 33*, 26–36.
Liu, P., & Biljecki, F. (2022). A review of spatially-explicit geoai applications in urban geography. *International Journal of Applied Earth Observation and Geoinformation, 112*, Article 102936.
Liu, P., & De Sabbata, S. (2021). A graph-based semi-supervised approach to classification learning in digital geographies. *Computers, Environment and Urban Systems, 86*, Article 101583.
Liu, P., Zhao, T., Luo, J., Lei, B., Frei, M., Miller, C., & Biljecki, F. (2023). Towards human-centric digital twins: Leveraging computer vision and graph models to predict outdoor comfort. *Sustainable Cities and Society, 93*, Article 104480.
Liu, Q., Wu, Y., Xiao, Y., Fu, W., Zhuo, Z., van den Bosch, C. C. K., Huang, Q., & Lan, S. (2020). More meaningful, more restorative? linking local landscape characteristics and place attachment to restorative perceptions of urban park visitors. *Landscape and Urban Planning, 197*, Article 103763.
Luttik, J. (2000). The value of trees, water and open space as reflected by house prices in the netherlands. *Landscape and urban planning, 48*, 161–167.
Lynch, K. (1984). *Reconsidering the image of the city*. Springer.
Ma, H., Xu, Q., & Zhang, Y. (2023). High or low? exploring the restorative effects of visual levels on campus spaces using machine learning and street view imagery. *Urban Forestry & Urban Greening, 128087*.
Meng, Y., Li, Q., Ji, X., Yu, Y., Yue, D., Gan, M., Wang, S., Niu, J., & Fukuda, H. (2023). Research on campus space features and visual quality based on street view images: A case study on the chongshan campus of liaoning university. *Buildings, 13*, 1332.
Markevych, I., Schoierer, J., Hartig, T., Chudnovsky, A., Hystad, P., Dzhambov, A. M., De Vries, S., Triguero-Mas, M., Brauer, M., Nieuwenhuijsen, M. J., et al. (2017). Exploring pathways linking greenspace to health: Theoretical and methodological guidance. *Environmental research, 158*, 301–317.
Massoni, E. S., Barton, D. N., Rusch, G. M., & Gundersen, V. (2018). Bigger, more diverse and better? mapping structural diversity and its recreational value in urban green spaces. *Ecosystem Services, 31*, 502–516.
Nordh, H., Alalouch, C., & Hartig, T. (2011). Assessing restorative components of small urban parks using conjoint methodology. *Urban forestry & urban greening, 10*, 95–103.
Nordh, H., Hartig, T., Hagerhall, C., & Fry, G. (2009). Components of small urban parks that predict the possibility for restoration. *Urban forestry & urban greening, 8*, 225–235.
Pasini, M., Berto, R., Brondino, M., Hall, R., & Ortner, C. (2014). How to measure the restorative quality of environments: The prs-11. *Procedia-Social and behavioral sciences, 159*, 293–297.
Peterson, L. E. (2009). *K-nearest neighbor. Scholarpedia, 4*, 1883.
Qi, Y., Chen, Q., Lin, F., Liu, Q., Zhang, X., Guo, J., Qiu, L., & Gao, T. (2022). Comparative study on birdsong and its multi-sensory combinational effects on physio-psychological restoration. *Journal of Environmental Psychology, 83*, Article 101879.
Quercia, D., O'Hare, N. K., & Cramer, H. (2014). Aesthetic capital: What makes london look beautiful, quiet, and happy?, in. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing* (pp. 945–955).
Ratcliffe, E. (2021). Sound and soundscape in restorative natural environments: A narrative literature review. *Frontiers in psychology, 12*, Article 570563.
Ritchie, H., Roser, M., 2018. Urbanization. Our world in data.
Roe, J., Barnes, L., Napoli, N. J., & Thibodeaux, J. (2019). The restorative health benefits of a tactical urban intervention: An urban waterfront study. *Frontiers in built environment, 5*, 71.
Rzotkiewicz, A., Pearson, A. L., Dougherty, B. V., Shortridge, A., & Wilson, N. (2018). Systematic review of the use of google street view in health research: Major themes, strengths, weaknesses and possibilities for future research. *Health & place, 52*, 240–246.
Salesses, P., Schechtner, K., & Hidalgo, C. A. (2013). The collaborative image of the city: Mapping the inequality of urban perception. *PloS one, 8*, e68400.
Samus, A., Freeman, C., Van Heezik, Y., Krumme, K., & Dickinson, K. J. (2022). How do urban green spaces increase well-being? the role of perceived wildness and nature connectedness. *Journal of Environmental Psychology, 82*, Article 101850.
Schertz, K. E., & Berman, M. G. (2019). Understanding nature and its cognitive benefits. *Current Directions in Psychological Science, 28*, 496–502.
Schwenker, F., Kestler, H. A., & Palm, G. (2001). Three learning phases for radial-basis-function networks. *Neural networks, 14*, 439–458.
Staats, H., Jahncke, H., Herzog, T. R., & Hartig, T. (2016). Urban options for psychological restoration: Common strategies in everyday situations. *PLoS One, 11*, e0146213.
Stenfors, C. U., Van Hedger, S. C., Schertz, K. E., Meyer, F. A., Smith, K. E., Norman, G. J., Bourrier, S. C., Enns, J. T., Kardan, O., Jonides, J., et al. (2019). Positive effects of nature on cognitive performance across multiple experiments: Test order but not affect modulates the cognitive effects. *Frontiers in Psychology, 10*, 1413.
Subiza-Pérez, M., Pasanen, T., Ratcliffe, E., Lee, K., Bornioli, A., de Bloom, J., & Korpela, K. (2021). Exploring psychological restoration in favorite indoor and outdoor urban places using a top-down perspective. *Journal of Environmental Psychology, 78*, Article 101706.
Tang, J., & Long, Y. (2019). Measuring visual quality of street space and its temporal variation: Methodology and its application in the hutong area in beijing. *Landscape and Urban Planning, 191*, Article 103436.
Tobler, W. R. (1970). A computer movie simulating urban growth in the detroit region. *Economic geography, 46*, 234–240.
Thekumparampil, K.K., Oh, S., Wang, C., Li, L.J., 2018. Attention-based graph neural network for semi-supervised learning.
Ulrich, R. S., Simons, R. F., Losito, B. D., Fiorito, E., Miles, M. A., & Zelson, M. (1991). Stress recovery during exposure to natural and urban environments. *Journal of environmental psychology, 11*, 201–230.
UNFPA, U., 2007. State of world population 2007: Unleashing the potential of urban growth. United Nations Population Fund.
Valtchanov, D., & Ellard, C. G. (2015). Cognitive and affective responses to natural scenes: Effects of low-level visual properties on preference, cognitive load and eye-movements. *Journal of Environmental Psychology, 43*, 184–195.
Velarde, M. D., Fry, G., & Tveit, M. (2007). Health effects of viewing landscapes–landscape types in environmental psychology. *Urban forestry & urban greening, 6*, 199–212.
Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., Bengio, Y., 2017. Graph attention networks. arXiv preprint arXiv:1710.10903.
Wang, M., Chen, Z., Rong, H. H., Mu, L., Zhu, P., & Shi, Z. (2022). Ridesharing accessibility from the human eye: Spatial modeling of built environment with street-level images. *Computers, Environment and Urban Systems, 97*, Article 101858.
Wu, F., Souza, A., Zhang, T., Fifty, C., Yu, T., & Weinberger, K. (2019). Simplifying graph convolutional networks. *International conference on machine learning, PMLR.*, 6861–6871.
Wu, Y., Liu, Q., Hang, T., Yang, Y., Wang, Y., & Cao, L. (2024). Integrating restorative perception into urban street planning: A framework using street view images, deep learning, and space syntax. *Cities, 147*, Article 104791.
Xu, K., Hu, W., Leskovec, J., & Jegelka, S. (2019). How powerful are graph neural networks?. *International Conference on Learning Representations*.
Xu, Y., Jin, S., Chen, Z., Xie, X., Hu, S., & Xie, Z. (2022). Application of a graph convolutional network with visual and semantic features to classify urban scenes. *International Journal of Geographical Information Science, 36*, 2009–2034.
Yao, Y., Liang, Z., Yuan, Z., Liu, P., Bie, Y., Zhang, J., Wang, R., Wang, J., & Guan, Q. (2019). A human-machine adversarial scoring framework for urban perception assessment using street-view images. *International Journal of Geographical Information Science, 33*, 2363–2384.
Yao, Y., Zhang, J., Qian, C., Wang, Y., Ren, S., Yuan, Z., & Guan, Q. (2021). Delineating urban job-housing patterns at a parcel scale with street view imagery. *International Journal of Geographical Information Science, 35*, 1927–1950.
Yin, Y., Thwaites, K., & Shao, Y. (2022). Balancing street functionality and restorative benefit: Developing an expectation–current approach to street design. *Sustainability, 14*, 5736.
Ying, Z., Bourgeois, D., You, J., Zitnik, M., Leskovec, J., 2019. Gnnexplainer: Generating explanations for graph neural networks. Advances in neural information processing systems 32.
Zhang, F., Zhou, B., Liu, L., Liu, Y., Fung, H. H., Lin, H., & Ratti, C. (2018). Measuring human perceptions of a large-scale urban region using machine learning. *Landscape and Urban Planning, 180*, 148–160.
Zhang, Y., Chen, N., Du, W., Li, Y., & Zheng, X. (2021). Multi-source sensor based urban habitat and resident health sensing: A case study of wuhan, china. *Building and Environment, 198*, Article 107883.
Zhang, Y., Liu, P., & Biljecki, F. (2023). Knowledge and topology: A two layer spatially dependent graph neural networks to identify urban functions with time-series street view image. *ISPRS Journal of Photogrammetry and Remote Sensing, 198*, 153–168.
Zhao, T., Liang, X., Tu, W., Huang, Z., & Biljecki, F. (2023). Sensing urban soundscapes from street view imagery. *Computers, Environment and Urban Systems, 99*, Article 101915.
Zhao, J., Wu, J., & Wang, H. (2020). Characteristics of urban streets in relation to perceived restorativeness. *Journal of exposure science & environmental epidemiology, 30*, 309–319.
Zhou, H., He, S., Cai, Y., Wang, M., & Su, S. (2019). Social inequalities in neighborhood visual walkability: Using street view imagery and deep learning technologies to facilitate healthy city planning. *Sustainable cities and society, 50*, Article 101605.
Zhu, D., Zhang, F., Wang, S., Wang, Y., Cheng, X., Huang, Z., & Liu, Y. (2020). Understanding place characteristics in geographic contexts through graph convolutional neural networks. *Annals of the American Association of Geographers, 110*, 408–420.
Zube, E. H. (1987). Perceived land use patterns and landscape values. *Landscape ecology, 1*, 37–45.