Frontiers of Architectural Research 15 (2026) 582–608

Higher Education Press

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.keaipublishing.com/foar

SOUTHEAST UNIVERSITY

RESEARCH ARTICLE

Machine learning-based study on factors influencing street vitality in urban fringe commercial districts: a case of Wuhan

Jingmei Shao [b], Yan Long [a,*], Xuejun Liu [c], Yuqiao Zheng [b], Yang Song [b], Jian Wang [g], Bo Liu [e], Jun Yang [f], Yujie Chen [h], Fan Zhang [d]

[a] *School of Art and Design, Wuhan University of Science and Technology, Wuhan 430065, China*
[b] *School of Urban Construction, Wuhan University of Science and Technology, Wuhan 430065, China*
[c] *School of Urban Design, Wuhan University, Wuhan 430072, China*
[d] *Wuhan Planning & Design Institute, Wuhan 430010, China*
[e] *College of Intelligent Construction, Wuchang University of Technology, Wuhan 430223, China*
[f] *School of Civil Engineering and Architecture, Wuhan Institute of Technology, Wuhan 430074, China*
[g] *School of Management, South-Central Minzu University, Wuhan 430074, China*
[h] *Guangzhou Urban Planning & Design Survey Research Institute Co., Ltd, Guangzhou 510060, China*

Received 20 February 2025; received in revised form 20 May 2025; accepted 11 June 2025

KEYWORDS
Street vitality;
Built environment;
Nonlinear relationships;
Machine learning;
Urban fringe area;
Commercial districts

Abstract Under the context of megacity decentralization, street vitality in urban fringe commercial districts becomes a key indicator for assessing population attraction and sustainability. However, most existing studies focus on central urban areas, with limited attention to the vitality characteristics of urban fringe streets and the nonlinear relationship between the built environment and street vitality. Methods for quantifying street vitality also remain underdeveloped. This study examines streets in commercial districts located in Wuhan's urban fringe areas. Street vitality is quantified using mobile signaling data. Based on multi-source data, a "5D" built environment indicator system is constructed. Machine learning combined with SHAP algorithms is employed to reveal the nonlinear effects and interaction mechanisms of built environment variables. Finally, hierarchical clustering is used to identify different types of street vitality. The results show that: (1) Location distance is the dominant factor influencing street vitality, with time-varying effects; (2) Built environment variables affect vitality in nonlinear ways; (3) There are interactive effects among indicators of different dimensions of built environment; (4) Streets with similar contribution patterns from built environment variables tend to have similar vitality levels. Vitality formation mechanisms vary by street type and show spatial clustering. Findings support refined urban fringe street design.

* Corresponding author.
E-mail address: longyan@wust.edu.cn (Y. Long).
Peer review under the responsibility of Southeast University.

https://doi.org/10.1016/j.foar.2025.06.010

1. Research background

Under the dual pressure of the functional saturation of the central city and the backward development of the urban fringe, the orderly disintegration of urban space to the fringe has become an important topic in the spatial reconstruction of megacities, and the urban fringe has gradually become the focus of urban planning and spatial reconstruction. As extensions of the city, urban fringe areas exhibit significant changes, strong development momentum, and characteristics of dynamism, cyclicity, and complexity (Liu, 2022). Studies have shown that enhancing the attractiveness of urban fringe areas to populations has become a key strategy for megacities and large cities to alleviate population pressure in central urban areas (Long et al., 2023). Street vitality has emerged as an essential indicator for evaluating population attractiveness and sustainable development potential (Li and Pan, 2023). As significant public spaces, streets in urban fringe areas play a crucial role in fostering social interactions, commercial activities and improving residents' quality of life. During the rapid urban expansion, urban fringe areas have gradually developed large-scale, functionally integrated commercial districts. The internal streets of these districts exhibit high functional mixing and strong potential for human activity, while also possessing clear boundaries and stable spatial structures—making them ideal settings for systematic and quantitative studies on the mechanisms of street vitality (Fei, 2019). Focusing on such spaces not only facilitates the exploration of vitality characteristics at the street scale but also highlights the unique contribution of commercial space to enhancing vitality in urban fringe areas, thereby providing valuable insights for the orderly renewal of fringe street spaces.

Streets serve as primary public spaces in cities, providing venues for walking, socializing, working, shopping and entertainment, thereby supporting urban diversity and vitality (Li et al., 2024a). Street vitality primarily pertains to the social vitality dimension of urban vitality, manifesting as various human activities in street spaces (Guo et al., 2021; Li and Lin, 2023). It can be measured by the number of people present on the street or the intensity of crowd activities (Gehl, 2001). The built environment of streets refers to the physical and spatial characteristics of the street and its surroundings, which provide spaces for and influence human activities. Existing research has confirmed that the formation and evolution of street vitality are closely linked to the built environment (Li and Pan, 2023), making the relationship between the built environment and street vitality a crucial topic in geography and urban planning (Liu et al., 2023). Numerous studies have explored different aspects of street vitality, including measurement methods, spatiotemporal variation characteristics, the construction of built environment indicator systems, and the mechanisms through which the built environment influences street vitality (Jiang et al., 2022; Wu et al., 2022a; Zou et al., 2023). However, most scholars have relied on traditional regression analyses to establish linear relationships (Gou and Wang, 2011; Li et al., 2021), thereby overlooking localized nonlinear relationships and interaction effects among built environment factors. Although some studies have begun to explore the nonlinear relationship between the built environment and street vitality (Guo et al., 2021; Wu et al., 2022b; Yang et al., 2023), this research is still in its early stages. Meanwhile, current measurements of street vitality lack high-precision methods based on high-frequency mobile data, such as mobile phone signaling, that can cover wider spatial areas. Moreover, street vitality research tends to focus on central urban areas or historical districts, with insufficient attention given to urban fringe areas.

Against this background, this study selects typical commercial districts in the urban fringe areas of Wuhan as the research area, using internal streets as the basic unit to examine how built environment factors influence street vitality within fringe commercial districts. First, mobile signaling data is used to measure street vitality, and a built environment indicator system is constructed based on the "5D" elements of the built environment. Then, XGBoost (eXtreme Gradient Boosting) and LightGBM (Light Gradient Boosting Machine) algorithms are employed to build models, while the SHAP (SHapley Additive exPlanations) interpretability algorithm is applied to explain the machine learning results. This enables an analysis of the nonlinear impacts of various built environment indicators on street vitality across different time periods, as well as the interaction effects among these indicators. Finally, hierarchical clustering is used to divide the types of street vitality, thus depicting the formation mechanism of street vitality in commercial districts in detail. On this basis, this study puts forward strategies to improve the vitality of the streets in the fringe areas, and provides reference for the refined design of the streets in the urban fringe areas of Wuhan.

2. Literature review

2.1. Urban fringe areas

In 1936, Louis (1936) formally introduced the concept of urban fringe areas from the perspective of urban ecology, building upon the theories of garden cities and concentric zone models. In 1980, Gu and Xiong (1989) introduced the concept of "urban fringe areas" to China. Since then, numerous scholars both domestically and internationally have conducted extensive research on related concepts, and it is generally believed that urban fringe areas are complex transitional zones that exhibit both urban and rural characteristics (Xie and Fu, 2023). Research on urban fringe areas primarily focuses on spatial structure and functional evolution (Sui and Lu, 2021), socio-economic characteristics (Zhao and Wan, 2021), ecological and environmental issues (Guan et al., 2022), land use and planning policies (Feng et al., 2022), and urban-rural integration and social governance (Luo et al., 2024), with the delineation of urban fringe areas being a key area of interest for scholars. Early definitions of urban fringe areas in China were primarily based on administrative boundaries (Cui and Wu, 1990). With the development of quantitative analysis methods, Li and Bai (2005) used a comprehensive analysis method to analyze the gradient changes of various indicators in urban space, thereby delineating the urban fringe area. Similarly, Ju et al. (2019) evaluated land use intensity using the impervious surface index and applied the maximum entropy threshold method to determine the inner and outer boundaries of the urban fringe area in Haikou. Zhou et al. (2017) adopted an indicator system comprising impervious surface cover, landscape fragmentation, and population density, using information entropy and sliding t-detection methods to identify the urban fringe area in Xi'an. The identification of urban fringe area boundaries provides a solid foundation for further research in this field.

A study (Xu et al., 2024) pointed out that vitality assessment provides a new perspective for addressing the issues in urban fringe areas. Based on the identification of the urban fringe area in Hangzhou, the study explored the vitality levels of Hangzhou's urban fringe from two dimensions: population vitality and spatial vitality. As a dynamic transitional area, the vitality assessment of the urban fringe is crucial for understanding the socio-economic characteristics, ecological conditions and residents' quality of life in this region. Therefore, strengthening the quantitative evaluation and research on urban fringe vitality is of significant theoretical and practical value for promoting sustainable urban development, optimizing urban spatial structure and advancing urban-rural integration. However, the research on vitality assessment in urban fringe areas is still noticeably underdeveloped.

2.2. Identification of commercial streets

Commercial districts in urban fringe areas play a key role in driving economic growth, alleviating pressure on core areas, and shaping the city's image. These districts not only attract investment and provide job opportunities, but also alleviate congestion in the core areas. Additionally, through planning and design, they enhance the appeal of fringe areas. Fully utilizing the potential of commercial districts is crucial for achieving sustainable development in urban fringe areas and enhancing the overall competitiveness of the city. Large-scale, integrated, and diverse commercial facilities in these districts attract significant foot traffic, increasing street vitality and becoming important support points for economic and social activities. Most studies use Points of Interest (POI) data to identify commercial districts, but various identification methods are employed. Wang et al. (2015) used basic unit division and K-means clustering for functional type classification. Cui et al. (2020) combined kernel density analysis with contour tree extraction to identify commercial centers. Yang et al. (2019) used intersection commercial kernel density and the three standard deviation method to extract commercial district boundaries. Numerous studies have shown that kernel density analysis is an effective and accurate method for identifying commercial districts. Therefore, this study employs kernel density analysis to identify the commercial district areas in urban fringe zones, using the streets within these areas as the research subjects.

2.3. Concept and measurement of street vitality

The concept of "vitality" was first introduced in geography for urban-scale studies. Jane Jacobs argued that urban vitality primarily stems from interpersonal interactions and the diversity generated by the interweaving of activity spaces within the city (Jacobs, 1961). Building on this, Maas (1984) emphasized the importance of social interaction and commercial-cultural resources for urban vitality. Streets, as fundamental spatial units connecting various urban spaces, play a crucial role in shaping urban form and public space. They serve multiple functions, including transportation, commercial activity, and social interaction, and are key spatial carriers of urban vitality (Gong et al., 2019; Huang et al., 2023). While traditional research in architecture and urban planning has mainly focused on visual aesthetics, spatial form, and traffic efficiency (Li et al., 2024a), the built environment itself does not directly generate vitality; rather, it provides the setting for human activities (Zhou et al., 2019). Therefore, understanding street vitality requires examining the interaction between physical space and human behavior (Yang et al., 2023). Although there is no unified definition of "street vitality," most studies regard it as the presence of dense populations and concentrated activities along the street (Niu et al., 2022; Wu et al., 2021), reflecting the attractiveness of the physical environment to people. In other words, street vitality can be seen as the outcome of interactions between the built environment and human behavior (Marcus, 2010; Lyu et al., 2025). It manifests externally in two dimensions: temporally, as the variation and continuity of crowd activities over time; and spatially, as the intensity and distribution of crowd activities in space (Wei and Wang, 2024). A highly vital street is thus characterized by spatiotemporal continuity in both human flow and activity.

The measure of street vitality should include the resident population density in a specific time period and a specific spatial range (Si et al., 2021). In the previous qualitative and descriptive research, field survey, questionnaire survey and cognitive map are usually used. With the widespread use of big data and the development of communication technologies, multi-source data, such as Baidu heat maps (Min and Ding, 2020), LBS data (Niu et al., 2019), mobile signaling data (Long and Zhou, 2016), and social media data (Hu et al., 2014), have gradually been applied to the measurement of street vitality. Compared to other data sources, mobile signaling data has the advantages of large-scale coverage, real-time availability, authenticity, accuracy, and relatively low cost, making it an ideal data source for street-level studies. Therefore, this study uses mobile phone signaling data to improve the accuracy of street vitality measurement, which makes up for the shortcomings of traditional methods in small sample size, small coverage and difficult to reflect the real dynamic activities of the crowd.

2.4. The influence of the built environment on street vitality

Research on how the built environment stimulates street vitality can be traced back to the 1960s. Jane Jacobs argued that density, diversity, small blocks, and old buildings collectively contribute to street vitality. Gehl (2001) found that factors such as street width and length, interface diversity, motor vehicle traffic volume, and indoor-outdoor elevation differences influence pedestrian activity. Katz (1994) examined the impact of compactness, walkability, functional mix, and building density on street vitality. The widely accepted notion that the built environment affects street vitality has guided the spatial design of streets. A substantial body of research has been dedicated to uncovering the relationship between the built environment and street vitality. In 2008, Cervero et al. developed the "3D" built environment indicator system, which includes density, diversity and design. Later, Ewing et al. expanded this framework by distance to transit and destination accessibility, forming the "5D" built environment indicator system, which has provided a strong foundation for built environment measurement in both domestic and international studies. For instance, Li and Pan (2023) constructed a "5D" built environment indicator system and found that while commercial and public service facilities are crucial for enhancing street vitality, a high degree of functional mixing does not significantly boost street vitality.

With advancements in computer vision technology, some researchers have incorporated three-dimensional built environment elements into indicator systems using street view data (Si et al., 2021). Visual perception plays a critical role in understanding space and significantly influences human activity within spatial environments. Existing studies have employed semantic segmentation models to process street view data from mapping platforms, extracting three-dimensional built environment features for assessing street environmental quality (Zhang et al., 2019), analyzing their correlation with vitality (Li and Lin, 2023; Li et al., 2022), and classifying street spaces (Gong et al., 2019). Therefore, this study utilizes street view data to quantify certain built environment indicators within the design dimension.

2.5. Machine learning methods for uncovering nonlinear relationships and synergistic effects

Advancements in machine learning and interpretability methods have made it possible to study the complex relationship between the built environment and human activities. Models such as XGBoost overcome the limitations of linear assumptions, providing diverse tools for exploring nonlinear relationships. The data-driven nature of machine learning models enables researchers to gain insights directly from empirical data without relying on prior assumptions. These models can handle high-dimensional data, making them crucial for uncovering new relationships between street vitality and the built environment and for revealing their intricate interactions. However, such models are often regarded as "black boxes", making it difficult to interpret internal variable contributions. To address this issue, Lundberg and Lee (2017) proposed the Shapley Additive Explanations (SHAP) framework, an interpretability tool for machine learning models. SHAP accurately explains the contribution of each feature variable to model predictions, offering both global model interpretations and local feature explanations. In the field of urban planning, numerous studies have applied machine learning methods. Zhang et al. (2024) revealed the complex nonlinear relationship between urban vitality and built environment indicators in both suburban new towns and central urban areas. Xiao et al. (2021) used multi-source data from 166 metro station areas in Shenzhen to construct a gradient boosting decision tree model, uncovering the nonlinear and synergistic effects of transit-oriented development (TOD) on urban vitality. Therefore, this study employs machine learning models to investigate the nonlinear relationship between street vitality and the built environment. By addressing the challenges of nonlinear effects, variable interactions, and multicollinearity—issues that traditional regression models struggle with—this approach enhances both predictive accuracy and interpretability.

3. Research scope and data

3.1. Research scope and study area

Wuhan (Fig. 1), the capital of Hubei Province, is composed of 13 administrative districts, 6 functional zones, and 153 streets, covering a total area of 8569.15 km^2. In 2021, the city's socio-economic development entered a phase of high-quality growth, with the goal of accelerating the construction of "five centers" to build a modernized Wuhan. At the same time, a new round of territorial spatial planning began, proposing the development of a spatial structure that emphasizes "optimizing the main city, strengthening the four secondary zones, integrating urban and rural areas, and promoting coordinated development," bringing new development opportunities to the urban fringe areas.

Based on this, the study uses changes in impervious surface area, landscape fragmentation, and population density to delineate the fringe area. The commercial districts in the urban fringe are identified based on POI kernel density analysis results. This area is characterized by the concentration of commercial facilities such as restaurants, shopping centers, and entertainment venues, with a relatively well-developed built environment that has the potential to attract people to stay. It is thus representative for studying the street vitality of Wuhan's urban fringe areas.

In this study, street space is regarded as the smallest analysis unit, and it is surrounded by buildings and plants on both sides of the street. Through topological treatment of road network, the road is segmented at intersections, and

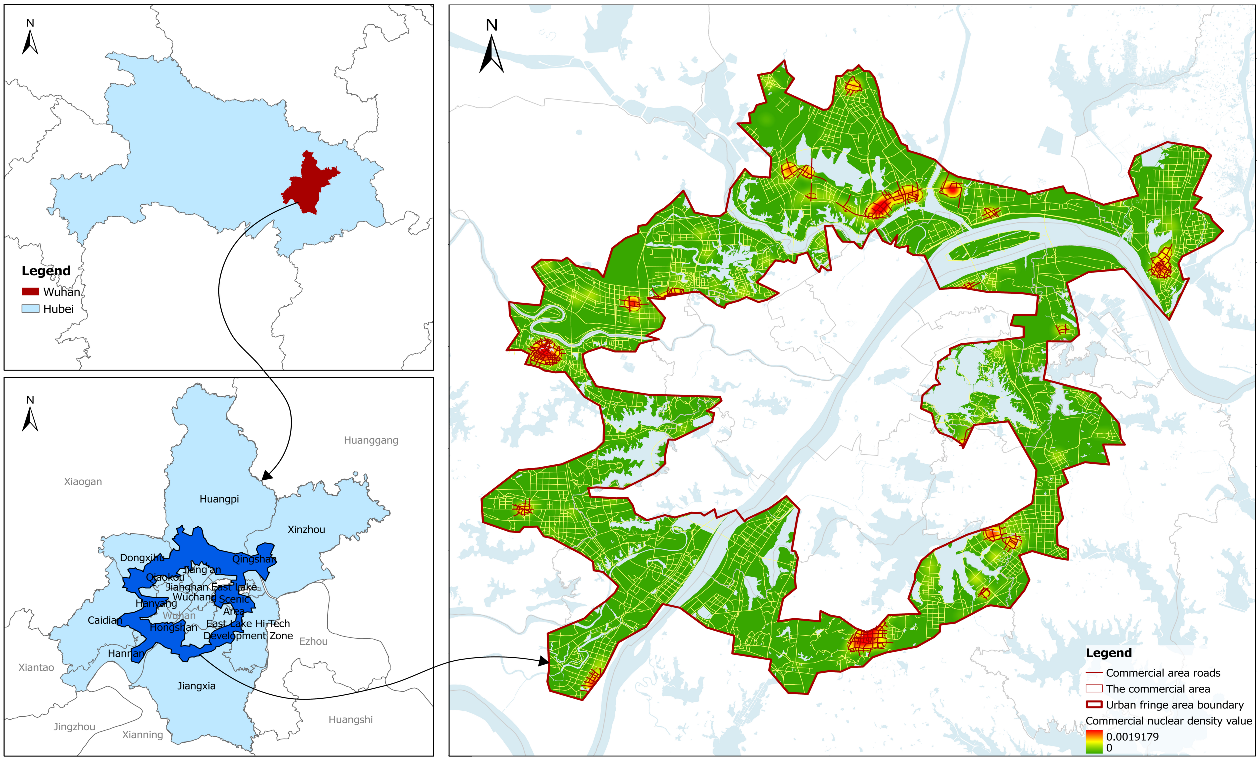

Fig. 1 Study area: urban fringe area of Wuhan.

the street segments between intersections are defined as street space units. Ultimately, a total of 791 street space units were built in the commercial district of the urban fringe.

3.2. Research data

The data used in this study mainly includes road network data, mobile signaling data, street view data, POI data, AOI data, building footprint data and transportation station data. The details, acquisition time, and data sources are provided in Table 1.

3.2.1. Road network data

This study obtained road network data for Wuhan, which includes road hierarchy attributes. The centerlines of primary, secondary, and tertiary roads within the study area were extracted and subjected to topological processing. Satellite imagery was referenced for verification and cross-checking, and road networks were interrupted at intersections. Previous studies have indicated that a 55 m buffer can encompass POI along streets (Long and Zhou, 2016) and is commonly used to calculate built environment indicators for streets. In this study, a 55 m buffer zone

Table 1 Data names and their sources.

Data name	Data content	Data source
Road network data	2023 Wuhan city road network data	OpenStreetMap
Mobile signaling data	September 2023 Wuhan city mobile signaling data with 130 m × 150 m resolution	China Unicom Operator
Street view data	May 2024 street view data for Wuhan's urban fringe areas	Baidu Street View Map Open Platform
POI data	2022 Wuhan city spatial data points for commercial, living, education, culture, medical, sports, and bus stations	Amap
AOI data	2024 Wuhan city land use attribute data for shopping, medical, living, and leisure areas	Baidu Map
Building footprint data	2024 building footprint lines and base area	Self-drawn images, Baidu Map
Transportation station data	Subway, bus, and BRT station data	Baidu Map

was created along both sides of the road centerline, thereby defining street space units.

3.2.2. Mobile signaling data

Mobile signaling data is generated through the exchange of information between mobile phone terminals and base stations, encompassing various types of information that can be used to infer user spatial distribution and trajectory patterns. After data processing, the average number of staying people in each time period for statistical grid cells is calculated. The number of people staying in each street unit and the staying population density are then computed based on the area ratio between street units and grid cells. This is used to represent the intensity of population activity, and appropriate measures have been taken to handle the privacy attributes of the data.

3.2.3. Street view data

Street view data is a type of geographic spatial data captured from a street-level perspective, encompassing images or videos of urban streets, buildings and other features. It provides an intuitive representation of the urban spatial layout and environment, collected by street view vehicles, drones and other equipment. This data is widely used in urban planning, transportation and other fields. Considering that Baidu Street View is frequently updated, 19,699 sampling points were selected along the road network of Wuhan's urban fringe area at 100 m intervals. Using Python, the URLs were constructed to call the Baidu Street View API to retrieve 19,699 street view images (512×512 pixels, JPG format). These images were then processed using the Deeplab v3+ semantic segmentation algorithm to calculate the area ratios of 19 factors.

DeepLabV3+ is an improved algorithm developed by Google in 2018 based on DeepLabV3, which integrates features using decoders and dilated convolutions to capture multi-scale information. This algorithm has been widely applied across various datasets. To train and optimize the DeepLabV3+ model for semantic segmentation in urban street scenes, especially in complex environments, this study used the Cityscapes dataset (including both fine and coarse annotations), enhancing the model's accuracy in recognizing and segmenting urban street scenes.

3.2.4. POI data and AOI data

The POI data consists of 111,991 open-source POI data points from Gaode, collected in 2022. In addition, the study also utilizes 24 types of AOI (Areas of Interest) data, which provide a more comprehensive reflection of the spatial characteristics and service capabilities of geographic entities. By combining these two datasets, a more holistic analysis of the spatial characteristics of the urban fringe area and their impact on street vitality can be conducted. Both POI and AOI data are converted into the WGS-84 coordinate system to ensure data consistency.

3.2.5. Building footprint data and transportation station data

The building footprint data represents the geographical shape and location of buildings in vector format, including boundary outlines and geographic coordinates. This data is sourced from Bing Maps and manually verified for accuracy. It is used to study the relationship between building-related indicators and street vitality. The transportation station data records the locations and information of public bus stops, subway stations and other transportation hubs. In this study, the distance between streets and transportation stations is used as a key indicator to evaluate street vitality.

4. Research methodology

4.1. Technical approach

This paper follows a technical approach of defining the research scope, data collection, indicator extraction, machine learning model prediction and explanation, and street vitality formation mechanism explanation, sequentially analyzing the complex nonlinear relationship between the built environment and street vitality (Fig. 2). The key steps are as follows: First, the study area is defined as the commercial districts in the urban fringe areas of Wuhan, based on which street units are delineated. Second, using multi-source big data, built environment features at the street level are extracted as independent variables across five dimensions—density, diversity, design, distance to transit, and destination accessibility. Meanwhile, mobile phone signaling data at the street scale are used to quantify street vitality as the dependent variable. Third, a non-linear prediction model is constructed using Python programming and machine learning techniques. Fourth, the SHAP method is employed to explain and visualize the model results. Fifth, targeted planning strategies are proposed based on the street vitality formation mechanism.

4.2. Method for defining the study area

4.2.1. Method for delimiting the edge zone

Drawing on previously published delimitation methods (Long et al., 2023), this study defines the urban fringe area using spatial overlay analysis, considering the inflection points of impervious surface ratio, landscape fragmentation and population density thresholds. First, the preliminary defined areas based on the inflection points of impervious surface ratio and landscape fragmentation are spatially overlaid. The overlapping areas are directly identified as the urban fringe zone, while the non-overlapping areas are further filtered and determined based on population density thresholds.

4.2.2. Method for identifying commercial districts

This study identifies commercial districts based on commercial POI data. The kernel density analysis tool in ArcGIS is used to calculate the commercial facility kernel density values within the urban fringe area, generating a kernel density surface. Then, the three standard deviation method is applied to extract the boundaries of commercial districts. Following the criteria in the "Wuhan Commercial Site Layout Plan (2016—2020)," which specifies that the area of a municipal commercial center should be greater than 25 ha, regions with an area smaller than 25 ha are excluded from the analysis.

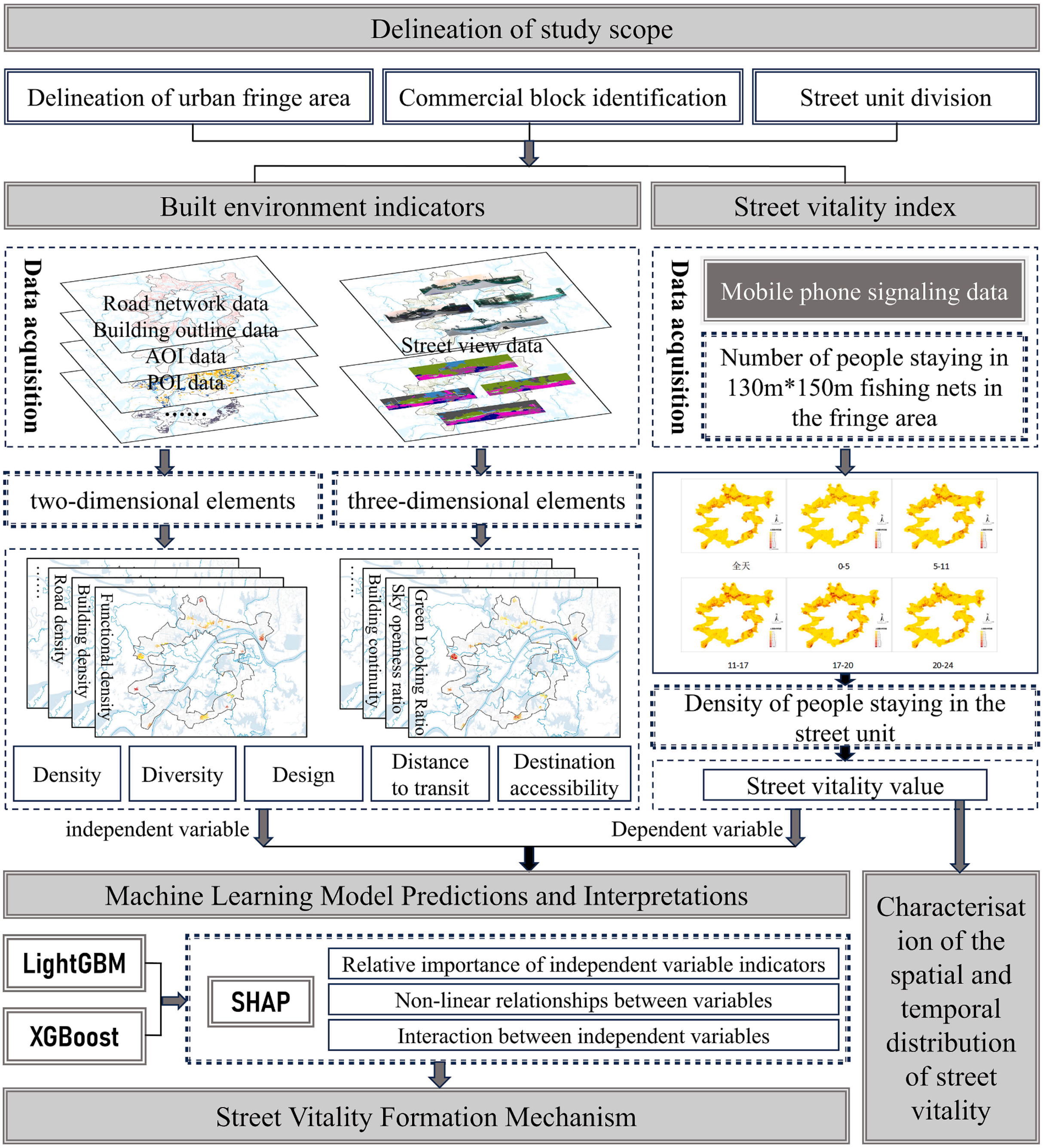

Fig. 2 Technical approach diagram.

4.3. Variable selection

On the level of street vitality measurement, this study takes mobile phone signaling data as the representation data of street vitality dependent variables. Although mobile phone signaling data has limited coverage of non-users and specific groups (e.g., children and the elderly), the Technical Guidelines for the Application of Mobile Signaling Data in the Field of Natural Resources recommend an expansion method based on operator market share and regional penetration rates to correct sample-population deviations and improve accuracy. With its high spatiotemporal resolution, dynamic tracking ability and broad coverage, mobile signaling data serves as a key and reliable source for street vitality research. The measurement of street vitality includes the density of staying populations within a specific time period and spatial range. Based on this, the study first extracts the number of people staying for more than half an hour in the 130 m $\times$ 150 m grids of the urban fringe areas of Wuhan for the whole day and each time segment using mobile signaling data. Then, the number of people staying in each street unit is calculated according to the proportion of spatial area, and the pedestrian flow density for each street unit is computed to represent the street vitality of each street unit. The larger the value, the more likely people are to engage in activities in that street, indicating stronger street vitality.

At the level of built environment measurement, this study develops a street-scale indicator system based on the "5D" dimensions of the built environment—Density, Diversity, Design, Distance to transit, and Destination accessibility. Specifically, the Density dimension captures the concentration of spatial resources such as functional facilities, buildings, and roads; Diversity reflects the mix and balance of functional types, indicating the complexity of land use; Design includes physical form, spatial interface, and traffic safety facilities, aiming to assess the overall pedestrian-friendly quality of the street environment; Distance to transit evaluates the spatial connectivity between street segments and transport facilities such as bus stops and subway stations; Destination accessibility measures the ease with which people can access public spaces and service facilities from the street, indirectly reflecting the built environment's support for daily activities. Based on prior research and data availability (Wu et al., 2022a, 2022b; Yang et al., 2023), a total of 21 built environment indicators were selected, encompassing both two-dimensional and three-dimensional spatial features of streets. Detailed calculation methods are provided in Table 2.

4.4. Model method

This study primarily employs two ensemble learning methods based on Gradient Boosting Decision Trees (GBDT)—LightGBM and XGBoost—along with the interpretable machine learning algorithm SHAP to explain the model outcomes. Compared to traditional linear regression models, GBDT-based algorithms offer significant advantages in modeling nonlinear relationships and feature interactions, enabling a more accurate depiction of the complex and heterogeneous associations between multidimensional built environment characteristics and street vitality (Li et al., 2024b). Moreover, in comparison to other ensemble methods such as Random Forests, LightGBM and XGBoost demonstrate superior performance in handling high-dimensional data, enhancing predictive accuracy, and improving computational efficiency (Chen and Guestrin, 2016; Ke et al., 2017), making them well-suited for this study's context involving multi-source spatial data and large-scale street-level samples. During model selection, we conducted performance comparisons among linear regression, Random Forest, and GBDT models. The results indicate that the GBDT models achieved higher predictive accuracy and were thus adopted as the primary modeling approach. The SHAP algorithm is used to interpret the predictions of the models by calculating each feature's contribution to the output. This enables quantification of variable importance, nonlinear relationships, and interaction effects, thereby providing more transparent and reliable analytical results.

4.4.1. Basic principles of LightGBM and XGBoost

In street vitality analysis, LightGBM and XGBoost, as gradient boosting tree algorithms, can effectively extract useful features from complex datasets and predict and evaluate street vitality levels. These algorithms improve the model's predictive ability by constructing multiple decision trees, making them outstanding in handling street data with complex nonlinear relationships (Li et al., 2024a; Liu et al., 2025).

LightGBM (Light Gradient Boosting Machine) uses a histogram algorithm to optimize feature processing. It discretizes continuous features into fixed bins, which reduces memory consumption and computational overhead. In street vitality analysis, this optimization enables rapid feature processing and model training on large-scale street datasets, such as high-frequency mobile signal data and POI data. LightGBM's leaf-wise growth strategy, which prioritizes splitting on the leaf nodes with the greatest gain, can better capture subtle differences in street vitality, particularly in data with complex spatial distribution characteristics. Furthermore, LightGBM's native support for categorical features helps to more precisely handle different types of street facilities or activity areas, further improving the accuracy of vitality analysis.

XGBoost (eXtreme Gradient Boosting) excels in street vitality analysis due to its efficiency and precision. XGBoost uses second-order Taylor expansion to optimize the loss function, providing higher prediction accuracy when handling street vitality data. By employing a greedy algorithm to select the best features for splitting, XGBoost can accurately capture the key factors affecting street vitality, such as traffic density, commercial activity, and green space distribution. XGBoost's weighted quantile method makes the feature partitioning process more efficient, especially when handling a large number of features, such as multi-dimensional street attribute data, significantly enhancing computational speed. In addition, XGBoost's ability to handle missing values allows it to automatically choose the optimal splitting method in the face of incomplete street data, ensuring the model's stability and reliability.

4.4.2. Explainable machine learning algorithm SHAP

The explainable machine learning algorithm SHAP (SHapley Additive exPlanations) is derived from the Shapley value in game theory, originally used to evaluate the contributions of each participant in a cooperative game. In machine learning, SHAP enhances model transparency by assigning a "contribution value" to each prediction made by the model, greatly improving the interpretability of the model (Muschalik et al., 2024).

The fundamental principle of SHAP is its ability to decompose the prediction results of a complex model into the contribution values of each feature. Specifically, SHAP takes into account the interaction effects between each feature and other features, calculating the independent contribution of each feature to the prediction outcome. This method not only provides a detailed explanation of the feature influence for each sample, but also allows researchers to gain deeper insight into how the model makes decisions in specific scenarios.

The formula for calculating SHAP values is as follows:

$$\varphi_i = \sum_{S \in N/\{i\}} \frac{|S|!(p-|S|-1)!}{M!} [f_x(S \cup \{i\}) - f_x(S)]. \tag{1}$$

Table 2 Built environment indicators and calculation methods.

Variable	Dimension	Indicator	Formula	Calculation method
Dependent variable		Street vitality	$V_i = \frac{P_i}{A_i}$	The street vitality value V_i is defined as the density of stayers within a street unit, calculated by dividing the number of stayers by the area of the street unit. Specifically, P_i represents the number of people staying in street unit i during the study period, extracted from mobile signaling data (with a dwell time $\geq$ 30 min), and A_i is the area of street unit i (m^2).
Independent variable	Density	Functional density	$FD_i = \frac{N_i}{A_i}$	Functional density FD_i represents the intensity of urban functions per unit area. N_i denotes the number of POIs within a 55 m buffer around street unit i, and A_i is the area of the 55 m buffer zone for street unit i (m^2).
		Building density	$BD_i = \frac{B_i}{A_i}$	Building density BD_i represents the proportion of land occupied by buildings within the street buffer zone, reflecting the intensity of spatial development around the street. B_i is the total building footprint area (m^2) within the 500 m buffer of street unit i, and A_i is the area of the 500 m buffer zone (m^2).
		Road density	$RD_i = \frac{L_i}{A_i}$	Road density RD_i reflects the extent of road coverage per unit area. A higher value indicates a denser road network, generally associated with better transportation accessibility and street connectivity. L_i is the total road length (in m) within the 500 m buffer of street unit i, and A_i is the area of the 500 m buffer zone (m^2).
	Diversity	Degree of business format mixing	$S = -\sum_{i=1}^{n} (p_{q,i} \ln p_{q,i})$	The functional diversity within the 55 m buffer of a street unit reflects the variety of business types and is measured using the Shannon Diversity Index. $p_{q,i}$ represents the proportion of POIs of type i within street unit q, and n denotes the number of distinct POI categories in the unit.
		Degree of place function mixing	$S = -\sum_{k=1}^{m} (p_{q,k} \ln p_{q,k})$	The functional mix within the 55 m buffer of a street unit is measured using the Shannon Diversity Index, reflecting land use diversity. $p_{q,k}$ denotes the proportion of AOI type k by area within unit q, and m represents the number of AOI categories in the unit.
	Design	Street length	$Length = L\_street$	L_street is the length of the center line of the street (m).
		Green looking ratio	$GLR_i = \frac{1}{n_i}\sum_{j=1}^{n_i}\frac{A^{(j)}_{\text{green}}}{A^{(j)}_{\text{total}}}$	The average proportion of vegetation across all sampled points within a street unit is calculated to obtain the overall green looking ratio (GLR_i), which reflects the street's greening level. n_i denotes the number of street view image sampling points in unit i; $A^{(j)}_{\text{green}}$ represents the pixel area of vegetation in the street view image at sampling point j; and $A^{(j)}_{\text{total}}$ is the total pixel area of the image at point j.
		Sky openness ratio	$SOR_i = \frac{1}{n_i}\sum_{j=1}^{n_i}\frac{A^{(j)}_{\text{sky}}}{A^{(j)}_{\text{total}}}$	The sky openness ratio (SOR_i) of a street unit is calculated by averaging the proportion of sky area across all sampled street view images within the unit, reflecting the degree of visual openness. Let n_i denote the number of sampling points in street unit i, $A^{(j)}_{\text{sky}}$ the pixel area of the sky in the street view image at sampling point j, and $A^{(j)}_{\text{total}}$ the total pixel area of that image.
		Relative pedestrian width	$RPW_i = \frac{1}{n_i}\sum_{j=1}^{n_i}\frac{A^{(j)}_{\text{walk}}}{A^{(j)}_{\text{road}}}$	The relative pedestrian width (RPW_i) of a street unit is calculated by averaging the ratio of the sidewalk area to the roadway area across all sampled street view images within the unit, reflecting the continuity of pedestrian space. Let n_i denote the number of sampling points in street unit i, $A^{(j)}_{\text{walk}}$ the pixel area of the sidewalk, and $A^{(j)}_{\text{road}}$ the pixel area of the roadway in the street view image at sampling point j.
		Pedestrian		The pedestrian continuity (PC_i) of a street unit is calculated by the standard deviation of

	continuity	$PC_i = StdDev\left(\left\{\frac{A_{\text{walk}}^{(j)}}{A_{\text{total}}^{(j)}}\right\}_{j=1}^{n_i}\right)$	the sidewalk area ratio across all sampled street view images within the unit, reflecting the continuity of pedestrian space. Let n_i denote the number of sampling points in street unit i, $A_{\text{walk}}^{(j)}$ the pixel area of the sidewalk at sampling point j, and $A_{\text{total}}^{(j)}$ the total pixel area of the street view image at sampling point j.
	Building continuity	$BC_i = StdDev\left(\left\{\frac{A_{\text{building}}^{(j)}}{A_{\text{total}}^{(j)}}\right\}_{j=1}^{n_i}\right)$	The building continuity (BC_i) of a street unit is calculated by the standard deviation of the building area ratio across all sampled street view images within the unit, reflecting the continuity of the street's building facades. Let n_i denote the number of sampling points in street unit i, $A_{\text{building}}^{(j)}$ the pixel area of the building at sampling point j, and $A_{\text{total}}^{(j)}$ the total pixel area of the street view image at sampling point j.
	Street aspect ratio	$SAR_i = \frac{1}{n_i}\sum_{j=1}^{n_i}\frac{A_{\text{building}}^{(j)}}{A_{\text{road}}^{(j)} + A_{\text{walk}}^{(j)}}$	The street aspect ratio (SAR_i) of a street unit is calculated by averaging the ratio of building area to the total road width (road + sidewalk) across all sampled points within the unit, reflecting the spatial compactness. Let $A_{\text{building}}^{(j)}$ denote the pixel area of the building at sampling point j, $A_{\text{road}}^{(j)}$ the pixel area of the road at sampling point j, and $A_{\text{walk}}^{(j)}$ the pixel area of the sidewalk at sampling point j.
	Traffic safety facilities proportion	$TSFP_i = \frac{1}{n_i}\sum_{j=1}^{n_i}\frac{A_{\text{fence}}^{(j)} + A_{\text{pole}}^{(j)}}{A_{\text{total}}^{(j)}}$	The traffic safety facilities proportion ($TSFP_i$) of a street unit is calculated by averaging the proportion of the area occupied by traffic safety facilities (railings + poles) across all sampled points within the unit, reflecting the level of traffic safety. Let $A_{\text{fence}}^{(j)}$ and $A_{\text{pole}}^{(j)}$ denote the pixel areas of the railing and pole regions, respectively, at sampling point j, and $A_{\text{total}}^{(j)}$ the total pixel area of the street view image at sampling point j.
	Vehicle interference index	$VII_i = \frac{1}{n_i}\sum_{j=1}^{n_i}\frac{A_{\text{vehicle}}^{(j)}}{A_{\text{total}}^{(j)}}$	The vehicle interference index (VII_i) of a street unit is calculated by averaging the proportion of the area occupied by motor vehicles across all sampled points within the unit, reflecting the extent to which motor vehicles interfere with pedestrian activities. Let n_i be the number of street view sampling points within street unit i, $A_{\text{vehicle}}^{(j)}$ the pixel area of motor vehicles in the street view image at sampling point j, and $A_{\text{total}}^{(j)}$ the total pixel area of that image.
Distance to transit	Bus stop density	$BSD_i = \frac{N_{\text{bus}}^{i}}{L_i}$	Bus stop density (BSD_i) represents the ratio of the number of bus stops within the 55 m buffer of a street unit to the length of the street, reflecting the spatial concentration of bus stops. N_{bus}^{i} denotes the number of bus stops within the 55 m buffer of street unit i, and L_i denotes the length of street unit i (in m).
	Distance to the nearest bus stop	$D_{\text{bus},i} = \min(d(c_i, s_j)),\ s_j \in S_{\text{bus}}$	Distance to the nearest bus stop ($D_{\text{bus},i}$) refers to the Euclidean distance from the centroid of a street segment's centerline to the nearest bus stop, reflecting the proximity of public transit access. c_i denotes the centroid coordinates of street unit i; s_j represents the coordinates of bus stop j, which belongs to the bus stop set S_{bus}; and $d(c_i, s_j)$ is the Euclidean distance (in m) between centroid c_i and stop s_j.
	Distance to the nearest subway station	$D_{\text{metro},i} = \min(d(c_i, s_k)),\ s_k \in S_{\text{metro}}$	Distance to the nearest subway station ($D_{\text{metro},i}$) refers to the Euclidean distance from the centroid of a street segment's centerline to the nearest metro or BRT station, indicating the proximity of public transit access. c_i denotes the centroid coordinates of street unit i; s_k represents the coordinates of metro or BRT station k, which belongs to the metro/BRT station set S_{metro}; and d is the Euclidean distance (in m) between centroid c_i and station s_k.
Destination	Intersection		Intersection density (ID_i) refers to the ratio of the number of intersections within the

(*continued on next page*)

Table 2 (*continued*)

Variable	Dimension	Indicator	Formula	Calculation method
	accessibility	density	$ID_i = \frac{N_{\text{int}}^i}{A_i}$	500 m buffer of a street segment to the area of that buffer, reflecting the connectivity of the surrounding street network. N_{int}^i is the number of intersections within the 500 m buffer of street unit i, and A_i is the area of the 500 m buffer (in km^2).
		Open space accessibility	$D_{\text{open},i} = \min(d(c_i, s_l)),\ s_l \in S_{\text{open}}$	Open space accessibility ($D_{\text{open},i}$) refers to the Euclidean distance from the centroid of a street segment's centerline to the nearest open space, such as green space or public squares, reflecting the ease with which residents can access open spaces. c_i denotes the centroid coordinates of street unit i; s_l represents the coordinates of open space l, which belongs to the open space set S_{open}; and d is the Euclidean distance (in m) between centroid c_i and open space s_l.
		Distance to the nearest entertainment facility	$D_{\text{ent},i} = \min(d(c_i, s_m)),\ s_m \in S_{\text{ent}}$	Distance to the nearest recreational facility ($D_{\text{ent},i}$) refers to the Euclidean distance from the centroid of a street segment's centerline to the nearest recreational facility, reflecting the convenience for residents to access leisure and entertainment public spaces. c_i denotes the centroid coordinates of street unit i; s_m represents the coordinates of recreational facility m, which belongs to the recreational facility set S_{ent}; and d is the Euclidean distance (in m) between centroid c_i and facility s_m.
		Location distance	$D_{\text{loc},i} = \min(d(c_i, s_n)),\ s_n \in S_{loc}$	Location distance ($D_{\text{loc},i}$) refers to the Euclidean distance from the centroid of a street segment's centerline to the nearest large shopping mall, reflecting the quality of the street unit's locational advantage. c_i denotes the centroid coordinates of street unit i; s_n represents the coordinates of large shopping mall n, which belongs to the set of large shopping malls S_{loc}; and d is the Euclidean distance (in m) between centroid c_i and mall s_n.

In the formula, N represents the set of all features in the training dataset, S is a subset of N, $f_x(S)$ is the model output for the subset S, $f_x(S \bigcup \{i\})$ is the model output when feature i is added to the subset S, p is the number of features, and x is the feature value.

The SHAP model prediction formula is as follows:

$$f(x) = g(z') = \varphi_0 + \sum_{j=1}^{M} \varphi_j z'_j \tag{2}$$

In the formula, $f(x)$ represents the predicted value of street vitality, $z'_j \in \{0,1\}^M$ indicates whether the corresponding feature is present in the decision path, M is the number of input features, φ_0 is the constant, and φ_j is the SHAP value of the feature j.

SHAP can accurately quantify the contribution of each built environment variable to the model's base street vitality value and output value, thus assessing its impact on vitality prediction. This study combines the relevant principles from published literature (Lundberg and Lee, 2017; Wu et al., 2022b; Yang et al., 2023) and uses SHAP to interpret the prediction results of the XGBoost and LightGBM models, quantifying the influence of various built environment variables on street vitality.

4.4.3. Model selection

Before model construction, this study first compared the performance of LightGBM, XGBoost, linear regression, and random forest. Taking the model selection based on the full-day sample data as an example, we used machine learning libraries such as Scikit-learn, LightGBM, and XGBoost in Python 3.7 to split the dataset into a training set (80%) and a testing set (20%), using Mean Squared Error (MSE) as the loss function for preliminary training. To improve model fitting performance and control overfitting risk, 5-fold cross-validation combined with grid search (GridSearchCV) was applied for hyperparameter optimization. After obtaining the optimal hyperparameters for each model, we trained the models separately and evaluated their performance using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and the coefficient of determination (R^2). Lower MAE and RMSE values and higher R^2 values indicate better model fit. The results (see Table 3) show that the XGBoost model performed best on the full-day dataset.

Subsequently, based on the optimal hyperparameters of the selected model (see Table 4) and in combination with the SHAP algorithm, a feature importance analysis was conducted to evaluate the nonlinear relationships and interaction effects of variables. To reduce the random error in the results, model training was repeated 10 times with different random seeds, and the average feature importance and its 95% confidence interval were calculated.

Table 3 Performance comparison.

Model name	R^2	MAE	RMSE
Linear Regression	0.3104	27655.16	40125.06
Random Forest	0.4979	16674.99	26135.61
XGBoost	0.5812	14639.71	23868.60
LightGBM	0.5584	15556.21	24510.85

Table 4 Optimal hyperparameters.

Parameter name	Value
n_estimators	225
max_depth	10
learning_rate	0.0323899
subsample	0.8
colsample_bytree	0.7
min_child_weight	3
gamma	0.4268848
lambda	0.0000711
alpha	0.0000027

In addition to the full-day sample, models for other time periods (00:00—05:00, 05:00—11:00, 11:00—17:00, 17:00—20:00, and 20:00—24:00) were also fitted using similar procedures. The results indicate that LightGBM performed best during the 00:00—05:00 and 05:00—11:00 time periods, while XGBoost showed superior performance in the remaining time periods.

5. Results analysis

5.1. Spatiotemporal variation characteristics of street vitality intensity

5.1.1. Variation characteristics of pedestrian density in different time periods

Based on the variation of the hourly staying population density in street units (Fig. 3), and taking into account the urban activity rhythm, residents' daily routines, traffic flow, and the temporal distribution characteristics of commercial activities, the vitality changes of the street's staying population in urban fringe areas are divided into five time periods (late night, morning, afternoon, evening, and night).

0:00—5:00 Late Night Period: During this period, the pedestrian density is low, primarily because most residents are sleeping, and commercial activities and traffic flow nearly stop. Nighttime economic activities are minimal, and except for some 24-h convenience stores or night-shift workers, the overall street vitality significantly decreases. Therefore, this period is defined as the valley period of street vitality.

5:00—11:00 Morning Period: This period corresponds to the morning commuting and peak work hours of residents. Between 5:00—7:00, some urban functions gradually resume, such as breakfast shops opening, and traffic flow begins to increase. From 7:00—9:00, the morning rush hour peaks. This period sees a gradual rise in street vitality driven by residents' travel activities. Before 11:00, with the start of morning work and commercial activities, pedestrian density continues to increase.

11:00—17:00 Afternoon Period: During this period, street vitality remains relatively stable, reflecting the daily routine of urban residents and commercial activities. Lunchtime sees a concentration of people in restaurants, leisure, and shopping activities. In the afternoon, the street enters a relatively calm period where pedestrian traffic and

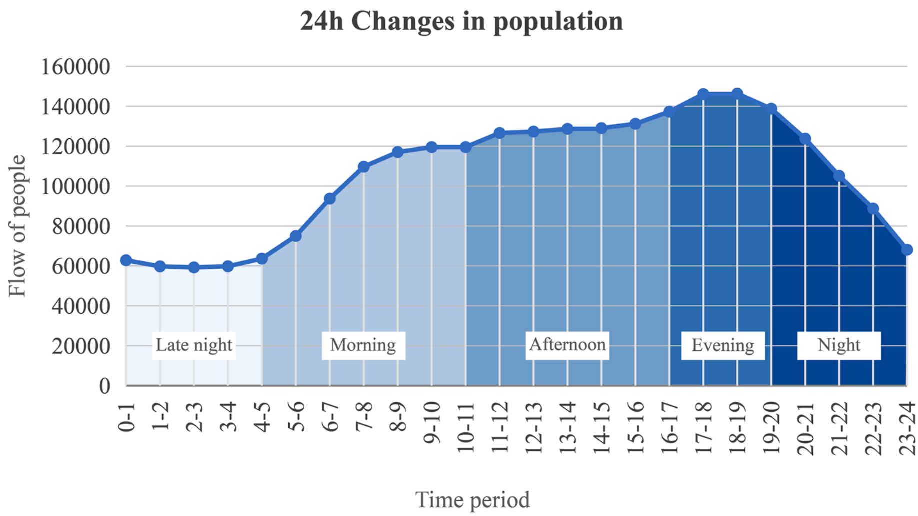

Fig. 3 24-Hour variation of pedestrian density.

commercial activities remain active, with daily life, work, and leisure activities dominating the street vitality.

17:00—20:00 Evening Period: This is the peak period for street vitality. It is closely related to residents finishing work, school, and the concentrated start of dining and entertainment activities. From 17:00—19:00, as vehicle and pedestrian flow increases, the street becomes busy and reaches its peak vitality for the day. From 19:00—20:00, pedestrian density gradually decreases.

20:00—24:00 Night Period: With the arrival of night, after 20:00, most residents reduce their activities, the demand for street functions decreases, pedestrian density drops, and street vitality rapidly declines.

This segmentation of time periods helps more precisely analyze the dynamic characteristics of street vitality over time. The late-night period represents the activity trough, with vitality rising quickly in the morning due to commuting and commercial activities. The afternoon is a relatively stable period in terms of life rhythm. The evening reaches its peak, while vitality decreases rapidly at night.

5.1.2. Spatial distribution characteristics of crowd activity intensity

Based on Wuhan's mobile signaling data, this paper visualizes and analyzes the average pedestrian density during different time periods, indirectly revealing the spatial distribution characteristics of street vitality (Fig. 4). After exploring the spatial representation of

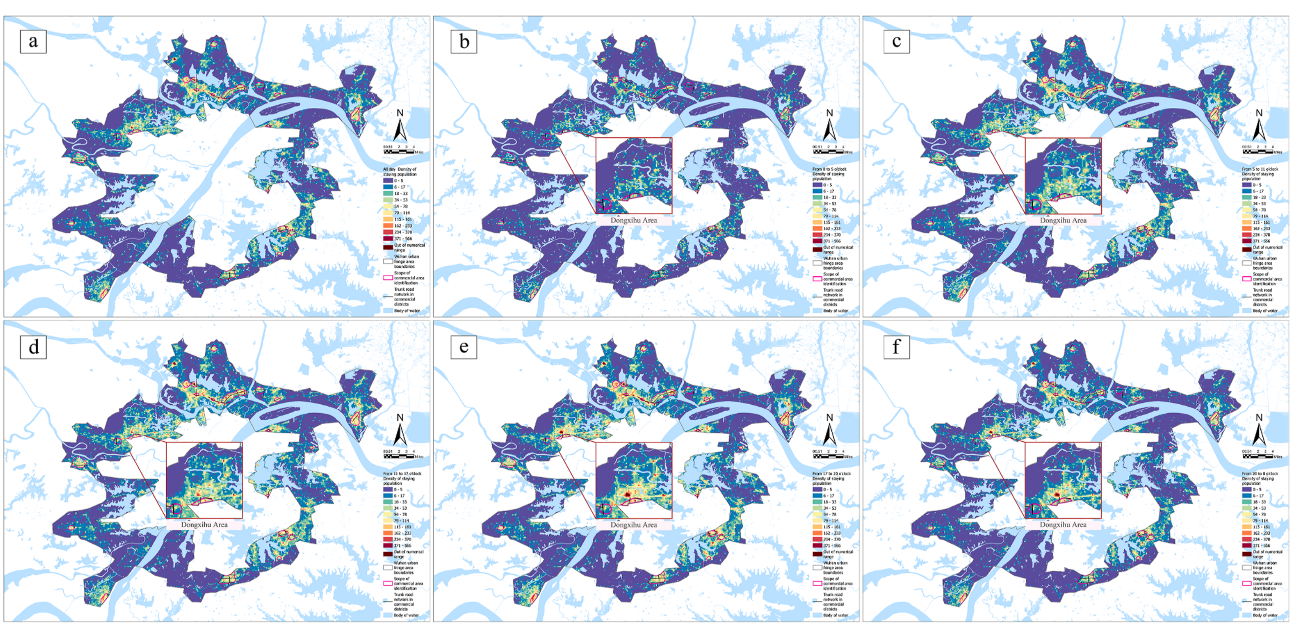

Fig. 4 Changes in staying population density at different time periods: (a) 24-h period, (b) 0:00—5:00, (c) 5:00—11:00, (d) 11:00—17:00, (e) 17:00—20:00, (f) 20:00—24:00.

street vitality for each time period, the following findings were made:

5.1.2.1. Static spatial distribution characteristics of staying population flow. The static spatial distribution of all-day staying population flow presents a "multi-area and multi-center" pattern, reflecting significant spatial heterogeneity and its close association with the functional structure of the built environment. High-value areas are mainly concentrated in the built-up areas, clustering along urban functional axes and key districts, locally forming several high-density patches and exhibiting a typical "point—axis—surface" agglomeration pattern. Among them, the "commercial—transportation" dual-core structure is particularly prominent, with multiple relatively high-value zones highly overlapping with commercial districts or large commercial facilities, indicating that commercial activities are a key driver of urban vitality. Meanwhile, large-scale public service facilities such as stadiums and exhibition centers also constitute important nodes for population aggregation. In addition, staying population flow shows a linear belt-like distribution along major roads, forming high-value patches at transportation hubs or road intersections. It is worth noting that some identified commercial districts have relatively low staying population flow, indicating that their actual vitality level may be low.

The spatial distribution of population flow in urban fringe areas also exhibits regional heterogeneity: in the north, Huangpi District is a high-value area, where high staying population zones are linearly distributed along Julong Avenue, showing concentrated range and clear boundaries; in the northwest, although Dongxihu District has a generally high population flow, its spatial distribution is more evenly spread in a planar pattern, forming only localized high-value points around nodes such as Wuhan Fifth Ring Sports Center; in the southwest, the staying population in Hannan District is centripetally clustered around Shamao New Town, reflecting a regionally centralized development pattern.

5.1.2.2. Dynamic spatial variation characteristics of staying population flow. The spatial distribution pattern of population flow in urban fringe areas remains relatively stable across different time periods, showing the characteristic of "generally consistent overall pattern, with significant local dynamic fluctuations," with population flow and activity range exhibiting rhythmic changes over time. Specifically, 0:00—5:00 is a period of low nighttime activity, with the lowest level of population movement; from 5:00—11:00, population flow increases rapidly and activity areas expand; between 11:00—17:00, population flow continues to increase, with activity areas further expanding and high-value aggregation zones forming in certain areas; from 17:00—20:00, population flow maintains a high level, but the activity range slightly contracts, showing stronger spatial agglomeration; from 20:00—24:00, population flow decreases rapidly, with activity areas significantly contracting. Overall, the dynamic variation process of staying population flow presents a typical rhythmic evolution pattern of "diffusion—agglomeration—peak—decline".

To sum up, the spatial agglomeration pattern of resident flow is highly related to the urban structure, and the rhythmic change in time is related to periodic daily activities. Static distribution reveals the structural characteristics of "multi-area and multi-center" and emphasizes the comprehensive function of commerce, transportation and public service. Dynamic evolution presents a typical change pattern of "diffusion-aggregation-peak-fading". The above analysis not only enriches the understanding of the spatial distribution of crowd activity intensity, but also provides basic support for the subsequent study of street vitality in commercial districts.

5.2. Relative importance of built environment indicators

The importance ranking of built environment indicators (Fig. 5) is based on the SHAP value calculation results. Specifically, the marginal contribution of each feature to the predicted value of street vitality across all samples is first computed by the model, yielding corresponding SHAP values. Then, the absolute value of each SHAP value is taken, and the mean is calculated for each feature. Sorting these mean absolute SHAP values in descending order produces the global importance ranking shown on the left side of Fig. 5. The right side of the figure presents the local effects of each indicator on street vitality, illustrating the actual influence of each feature across different street samples. The vertical axis represents the indicators, the horizontal axis represents the SHAP values, and the color indicates the magnitude of the feature value. Through the visualization results, the influence characteristics and time changes of built environment on street vitality are analyzed.

Looking at the overall picture for the whole day, as shown in Fig. 5(a), the top five indicators contributing the most to global importance are location distance, intersection density, open space accessibility, distance to the nearest subway station, and functional density. This reflects that the vitality of the fringe area depends on both the physical connection to the city center and the functional layout and public space configuration within the region. A well-developed road traffic system, high-density functional layout, and accessible open spaces can effectively offset the resource disadvantages of the fringe area compared to the core area, enhancing street vitality. In urban fringe areas, street vitality relies more on transportation connectivity and functional layout, while the influence of factors such as pedestrian continuity and building continuity is more limited. Fringe area residents tend to rely more on long-distance transportation, such as cars or subways. Among the five dimensions of "density," "diversity," "design," "distance to transit," and "destination accessibility," the indicators that have the most significant impact on street vitality are functional density, degree of business format mixing, green looking ratio, distance to the nearest subway station, and location distance. Further analysis reveals that location distance, intersection density, and open space accessibility generally show a

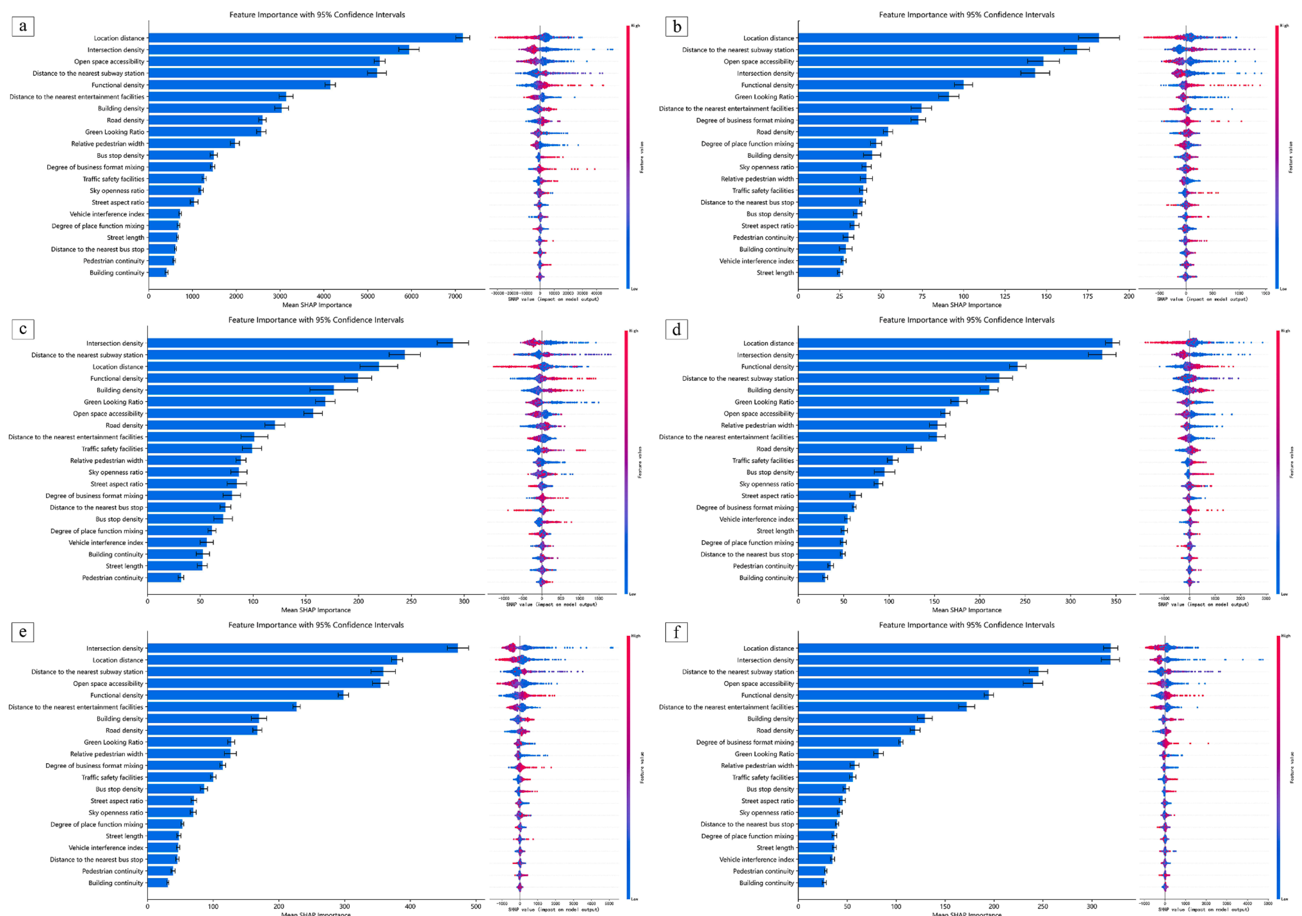

Fig. 5 Relative importance of built environment indicators: (a) 24-h period, (b) 0:00—5:00, (c) 5:00—11:00, (d) 11:00—17:00, (e) 17:00—20:00, (f) 20:00—24:00.

negative correlation with SHAP values, while functional density shows a positive correlation. The distance to the nearest subway station does not show a clear relationship with SHAP values at a macro level.

Looking at the analysis by time period, as shown in Fig. 5 (b)—(f), location distance and intersection density have a stable influence on street vitality across all time periods, while the contributions of distance to the nearest subway station, open space accessibility, and functional density fluctuate at different times, reflecting the varying spatial environment demands of people throughout the day. The impact of distance to the nearest subway station is more pronounced in the morning (5—11 a.m.), indicating that the subway plays a significant role during the commuting peak. Open space accessibility has a more prominent effect on street vitality during the evening hours (5—12 p.m.), reflecting the increased demand for leisure activities after work. The importance of functional density is more pronounced during the day (5 a.m.—5 p.m.), showing that during this time, people's demand for commercial, service, and other functional spaces increases, boosting street vitality. This may be related to the end of working hours, social activities, and consumption needs. Relative pedestrian width, street aspect ratio, and the proportion of traffic safety facilities are ranked relatively lower during each time period, but these factors still have some impact on the local environmental quality of streets and pedestrian experience. Although their global contribution is lower, they may have a local effect of promoting vitality in specific streets and specific time periods.

5.3. Non-linear relationship between built environment and street vitality

As shown in Fig. 6, each feature dependence plot corresponds to a built environment indicator, which meticulously depicts the nonlinear relationship between the built environment and street vitality. In each plot, each point represents a street unit sample, with the *x*-axis indicating the value of the variable, and the *y*-axis representing the SHAP value, i.e., the local impact of that variable on vitality. The GAM curve fitting the scatter plot is also included in the figure. The portion above zero indicates a positive effect on the dependent variable, while the portion below zero indicates a negative effect.

5.3.1. Density dimension

As shown in Fig. 6(a), the increase in functional density exhibits a clear positive correlation with the SHAP value of street vitality. At lower levels of functional density, the

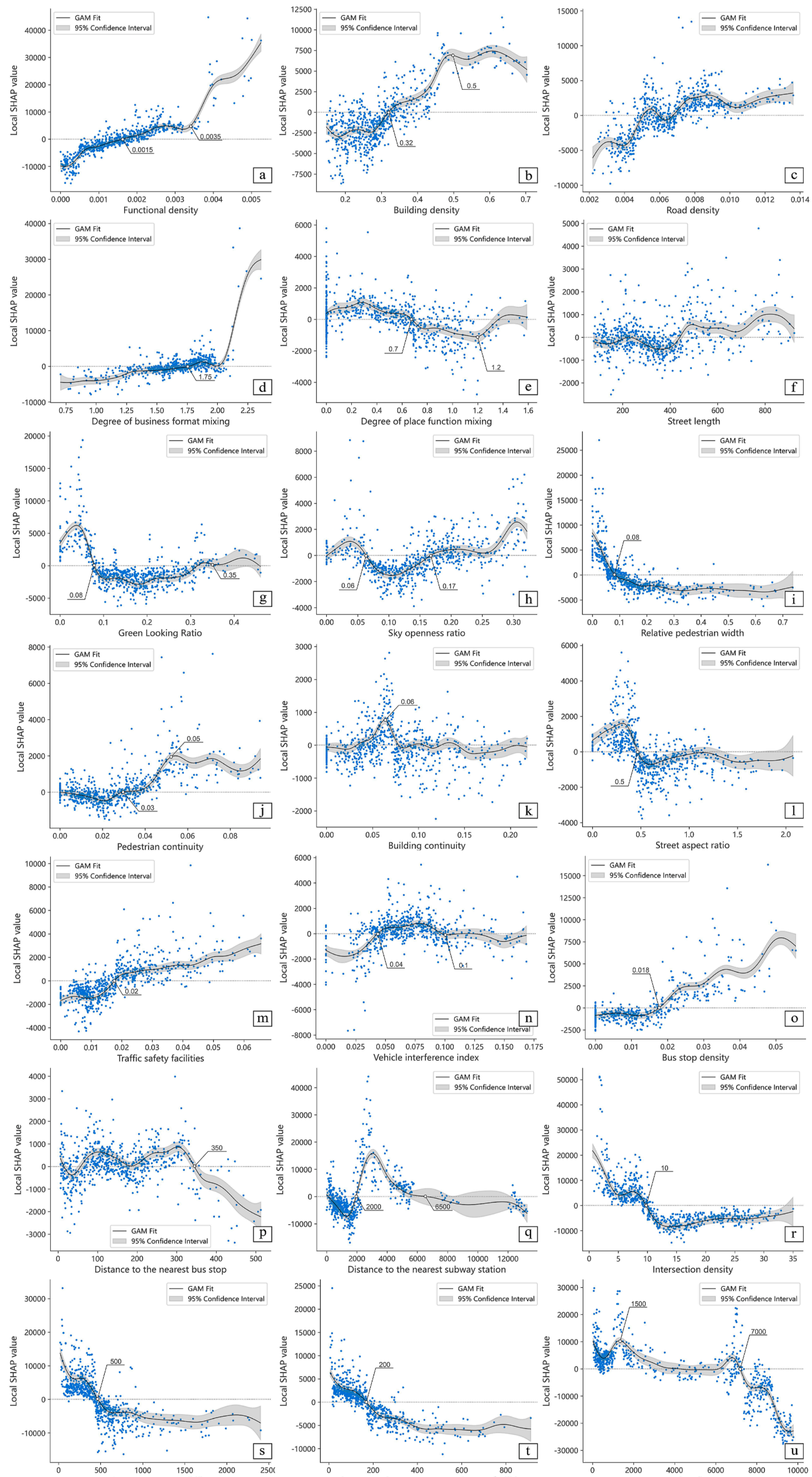

Fig. 6 Nonlinear effects of built environment indicators on street vitality.

SHAP value is generally negative, indicating that low functional density negatively impacts street vitality. When the functional density exceeds about 15 per hectare, the SHAP value becomes positive, beginning to exert a positive effect on street vitality. The promoting effect of functional density on street vitality gradually intensifies. Once functional density exceeds about 35 per hectare, the SHAP value increases rapidly, signifying a strong positive impact on street vitality. Additionally, the concentration of the scatter plot suggests that the functional density within commercial street areas in the fringe zone is generally at a low level, indicating that there is still room for improvement in this area's functional density.

As shown in Fig. 6(b), overall, building density exhibits a positive correlation with the SHAP value of street vitality. When building density is less than about 0.32, the SHAP value is negative, suppressing street vitality. When building density is between 0.32 and 0.5, as building density increases, the SHAP value rises, and its promoting effect on street vitality becomes more significant. However, when building density exceeds about 0.5, although the SHAP value remains positive, it begins to level off or even slightly decline. It indicates that at excessively high building densities, the positive effect on street vitality approaches saturation, and marginal diminishing returns may occur.

As shown in Fig. 6(c), the relationship between road density and the SHAP value exhibits a fluctuating upward trend, suggesting that the impact of road density on street vitality is somewhat complex. However, overall, as road density increases, street vitality is also enhanced, indicating that a moderate increase in road density can, to some extent, promote street vitality.

5.3.2. Diversity dimension

As shown in Fig. 6(d), the degree of business format mixing has a positive correlation with SHAP values. When the degree of business format mixing exceeds approximately 1.75, the SHAP value becomes positive, indicating a positive effect on street vitality. Especially when the degree of business format mixing exceeds about 2, this promoting effect significantly strengthens. Similar to functional density, the degree of business format mixing in the study area is generally low, suggesting that there is still considerable room for improvement in the business format mixing of streets in urban fringe areas.

As shown in Fig. 6(e), when the degree of place function mixing is low, the SHAP value is positive, contributing positively to street vitality. However, when the degree of place function mixing exceeds about 0.7, the SHAP value becomes negative, indicating that a higher degree of place function mixing suppresses street vitality. Nevertheless, as the mixing degree increases further, the curve slightly rises in the higher range (approaching 1.2), suggesting that under certain extremely high mixing degrees, place function mixing may have a positive impact on street vitality again. When the degree of place function mixing is low, the stronger functional specificity and concentration of people can promote the efficient operation of specific activities, thus positively influencing street vitality. However, when the functions are too mixed, people's interest in visiting may diminish, thereby suppressing street vitality. This result, however, may be influenced by limitations in the data and methodology. One limitation is the data itself. First, the AOI data coverage is incomplete, and certain functional categories in some areas are not recorded, leading to an underestimation of mixing degrees. AOI mainly reflects external outlines and cannot accurately represent land use types and internal functions; for example, a multifunctional area may be labeled as a single category. Second, AOI data does not include information on building height, meaning that buildings with small footprints but high floors (such as multi-story shopping centers) do not have their diverse functions reflected, thus underestimating their functional mixing. These factors result in deviations in the calculation of place function mixing and affect the prediction and actual relationship with street vitality.

5.3.3. Design dimension

As shown in Fig. 6(f), the trend line of street length fluctuates around 0 with SHAP values, indicating that the influence of street length on street vitality is somewhat unstable under different conditions. This fluctuation may be influenced by other factors, and the role of street length in street vitality is not significant.

As shown in Fig. 6(g), when the Green Looking Ratio (GLR) is less than 0.08, the SHAP value is greater than zero, indicating a positive effect on street vitality, particularly when the GLR is around 0.05, where this promoting effect is most significant. However, when the GLR exceeds a certain threshold, it suppresses street vitality. Research (Tong et al., 2020) has pointed out that high green visibility usually represents a high-quality street environment, but it does not necessarily guarantee street vitality. Excessively high GLR may lack necessary public service facilities and street functions, making it difficult to attract large crowds. However, when the GLR exceeds about 0.35, street vitality is again enhanced. This may be because streets near scenic areas or green spaces, where higher GLR brings comfort, attract a certain amount of foot traffic.

As shown in Fig. 6(h), when sky openness is low (approximately between 0 and 0.06), the SHAP value is greater than zero, indicating a promotion of street vitality. Lower sky openness is usually associated with dense building layouts or higher building heights, which help concentrate commercial and pedestrian activities, enhancing street vitality. When sky openness is in the range of 0.06—0.17, the SHAP value becomes negative, and the openness in this range suppresses street vitality. The reason could be that moderate openness leads to insufficient building density and functional support for the street. The space becomes more open but lacks vibrancy, failing to effectively gather foot traffic or commercial activities. When sky openness exceeds about 0.17, the SHAP value increases significantly, especially in the range of 0.25—0.30, where the promotion of street vitality is notably higher. This suggests that larger sky openness provides an open view and a comfortable public space environment, likely attracting more outdoor activities and social interactions, thereby enhancing street vitality. At this point, the street environment becomes more pleasant, offering greater potential for leisure, recreation, or commercial activities. Overall, the relationship between sky openness and street vitality follows a "U-shaped" curve. Both low and high sky openness contribute to enhancing street vitality, while moderate openness suppresses vitality.

As shown in Fig. 6(i), relative sidewalk width shows a negative correlation with SHAP values, indicating that a larger relative sidewalk width is not necessarily better. When the relative sidewalk width exceeds about 0.08, the SHAP value turns negative, exerting a suppressive effect on street vitality. If the sidewalk is excessively wide, it may alter its intended function. For example, overly wide sidewalks might be used for other purposes (such as for street vendors or bicycle lanes), reducing their contribution to street vitality. Additionally, overly wide sidewalks may make people feel less secure and disconnected from the street environment, thus reducing their willingness to linger. This decreased interaction between pedestrians and street buildings or shops results in a reduction of street vitality. In contrast, moderately wide sidewalks are better at promoting interpersonal interaction and social activity.

As shown in Fig. 6(j), an increase in pedestrian continuity has a positive effect on street vitality, meaning that higher pedestrian continuity generally strengthens street vitality. Continuous pedestrian zones facilitate pedestrian movement and increase interactions and activities, thus enhancing overall street vitality. When pedestrian continuity exceeds about 0.03, the SHAP value shifts from negative to positive and rises rapidly. However, this positive impact begins to stabilize when pedestrian continuity reaches about 0.05, and although pedestrian continuity still positively affects street vitality, its marginal effect weakens beyond this point.

As shown in Fig. 6(k), the scattered distribution of street unit sample points and the weak influence of building continuity on street vitality result in the fitting curve tending toward zero. The SHAP value reaches its peak around 0.06 of building continuity, showing a certain promotion of street vitality.

As shown in Fig. 6(l), when the street aspect ratio is less than approximately 0.5, the SHAP value is positive, first rising and then declining, overall having a positive effect on street vitality. In this case, the relatively small aspect ratio typically means that the street is wide and the buildings are relatively low. This environment may provide pedestrians with a good view and openness, improving the walking experience. A proper street aspect ratio creates a comfortable spatial scale that encourages more street interactions and commercial activities. However, when the aspect ratio exceeds about 0.5, the SHAP value turns negative, indicating that the street aspect ratio suppresses street vitality at this point. The buildings' height relative to the street becomes too high, narrowing the street space, which may cause larger shadow areas or obstruct views, creating a sense of spatial enclosure and discomfort for pedestrians. This environment reduces people's willingness to stay and engage in activities on the street, thereby suppressing street vitality.

As shown in Fig. 6(m), the proportion of traffic safety facilities has a clear positive correlation with the SHAP value. When the proportion of traffic safety facilities reaches about 0.02, the SHAP value shifts from negative to positive, indicating that the presence of traffic safety facilities enhances street vitality. As the proportion of safety facilities increases further, the SHAP value continues to rise, demonstrating that more and better-quality traffic safety facilities create a safer and more comfortable street environment, offering more space and opportunities for pedestrians and other non-motorized users, thus enhancing street vitality.

As shown in Fig. 6(n), when the vehicle interference index is at a low level (approximately between 0 and 0.04), the SHAP value is negative, indicating that vehicle interference suppresses street vitality at this point. Although vehicle interference is minimal, it may represent insufficient traffic flow, leading to fewer pedestrians and commercial activities, which implies lower street vitality. When the vehicle interference index is in the range of 0.04—0.1, the SHAP value turns positive, indicating a certain positive impact of vehicle interference on street vitality during this phase. This may be because a moderate level of vehicle activity means that the street has convenient traffic, attracting crowds to stay and more commercial activities, thus boosting vitality. When the vehicle interference index exceeds about 0.1, the SHAP value stabilizes and approaches zero, indicating that vehicle interference no longer has a significant effect on street vitality at higher levels.

5.3.4. Distance to transit dimension

As shown in Fig. 6(o), the density of bus stops exhibits a clear positive correlation with street vitality, especially when the density of bus stops exceeds a critical threshold, where the positive effect on street vitality becomes particularly significant. When the bus stop density exceeds about 18 stops per square kilometer, the SHAP value becomes positive, indicating a shift from an inhibitory effect to a promoting effect on street vitality. This suggests that once the bus stop density reaches a certain level, the coverage and convenience of public transportation begin to fully exert their influence. The density of bus stops is generally considered an important indicator of public transportation convenience. As the density of bus stops increases, a larger number of people can conveniently reach the street via the public transportation system. This convenience typically leads to higher foot traffic, thereby increasing the frequency of commercial and social activities.

As shown in Fig. 6(p), when the distance to the nearest bus stop is within about 350 m, the SHAP value is greater than 0, and street vitality typically improves. However, this improvement is relatively weak, possibly due to the lower population density, insufficient infrastructure, and commercial amenities in urban fringe areas, which limit the influence of bus stops. Bus stops are usually located in areas with concentrated foot traffic, and when the distance from a bus stop exceeds 350 m, the SHAP value turns negative, suppressing street vitality. The further the distance, the stronger the suppressing effect.

As shown in Fig. 6(q), when the distance to the nearest subway station is within about 2000 m, the SHAP value is less than 0, and there is an inhibitory effect on street vitality. Only a few streets very close to subway stations show an improvement in street vitality. However, when the distance to the subway station is between 2000 m and 6500 m, the SHAP value becomes greater than 0, showing a clear promoting effect on street vitality, especially around

3000 m, where this promoting effect peaks. Beyond about 6500 m, street vitality again experiences a negative impact. This phenomenon may be related to the subway's opening time and infrastructure conditions in the study area. Subway stations near the fringe area were relatively late in terms of construction and opening, and the surrounding vitality improvements may be constrained by delayed regional development and infrastructure, preventing a positive feedback loop between transportation convenience and street vitality in the short term. Moreover, many of the commercial districts in the study area are in older urban zones, where the level of surrounding infrastructure (such as commercial facilities and social venues) is relatively low. This incomplete support service cannot effectively attract large crowds, and the transportation convenience provided by the subway has not yet been directly converted into crowd activity and commercial prosperity. In contrast, some commercial districts, due to their inherent commercial appeal, can still promote street vitality even when the subway station is located at a greater distance.

5.3.5. Destination accessibility dimension

According to experience, high intersection density often means denser road network and stronger traffic connectivity, which should help to improve street vitality in theory. However, the research findings contradict this expectation. As shown in Fig. 6(r), intersection density in the urban fringe area negatively correlates with SHAP values. When intersection density is below about 10/km^2, SHAP values are positive, promoting vitality; above about 10/km^2, SHAP values become negative, suppressing vitality. Lower intersection density usually means fewer road divisions, allowing for a more continuous and smooth walking experience with fewer traffic lights, potentially increasing street usage and dwell time, thus enhancing vitality. In contrast, higher intersection density complicates traffic flow management and reduces pedestrian convenience, weakening the street's appeal. In Wuhan's urban fringe areas, older districts with high-density road networks and numerous intersections, originally designed for pedestrians and non-motorized traffic, have seen their positive impact on street vitality diminish due to aging and decentralized land use. In contrast, the road networks in newly developed areas, designed with modern planning principles, feature lower intersection density and structures more suited to car-dominated traffic. With concentrated land use, these areas attract people and activities, enhancing street vitality.

As shown in Fig. 6(s), open space accessibility is negatively correlated with SHAP values. The closer the distance to open spaces, the higher the accessibility. When the distance to open space is within about 500 m, the accessibility is higher, and the SHAP value is greater than 0, positively affecting street vitality. When the distance to open space exceeds 500 m, the SHAP value shifts from positive to negative, turning the influence on street vitality into a negative one. In urban design and planning, maintaining a reasonable distance between streets and open spaces can effectively enhance street vitality. Ensuring the distance between streets and parks and squares is within walking distance can not only promote daily activities, but also increase business and social activities by guiding people to flow.

As shown in Fig. 6(t), similar to open space accessibility, the distance to the nearest entertainment facilities has a critical threshold of 200 m when it comes to street vitality. When the distance to entertainment facilities is within about 200 m, the SHAP value is greater than 0, significantly promoting street activity. However, once the distance exceeds 200 m, the SHAP value becomes negative and gradually decreases, with the negative impact on street vitality becoming stronger. Entertainment and recreational facilities are key nodes for attracting foot traffic and can directly or indirectly influence pedestrian flow, commercial activities, and social interactions on the street. When these facilities are located closer to the street, more people are likely to stay or engage in activities on the street due to their convenience, thereby enhancing street vitality.

As shown in Fig. 6(u), location distance also exhibits a certain negative correlation with SHAP values. When the distance to a large shopping mall is within about 1500 m, SHAP values are at higher levels, and the positive impact on street vitality is more pronounced. A distance of 1500 m is considered a relatively acceptable walking or short-trip transportation range. When streets are within this range, it means that residents and visitors can easily access the shopping mall, and may also pass through and stop at the street along the way. This walking experience or short-distance travel increases pedestrian mobility and business opportunities on the street, thereby enhancing street vitality. When the distance to a large shopping mall exceeds about 7000 m, SHAP values turn negative, exerting a suppressive effect on street vitality. As a central node of the city, a shopping mall's influence is limited to a certain spatial range; beyond this range, the mall's impact on surrounding streets weakens, which in turn suppresses street vitality. Therefore, in urban planning, strategically locating large shopping malls is a key strategy for enhancing street vitality.

6. Key interaction effects between built environment variables

In SHAP, the interaction effect value explains the interaction between features by measuring the average contribution of each feature to the model prediction under different feature combinations. To further explore this, this study calculates the average absolute value of the SHAP interaction effects of two variables, identifying the variables with the strongest interaction effects. It analyzes how the influence of built environment variables on street vitality is affected by another variable. In the figure, the *x*-

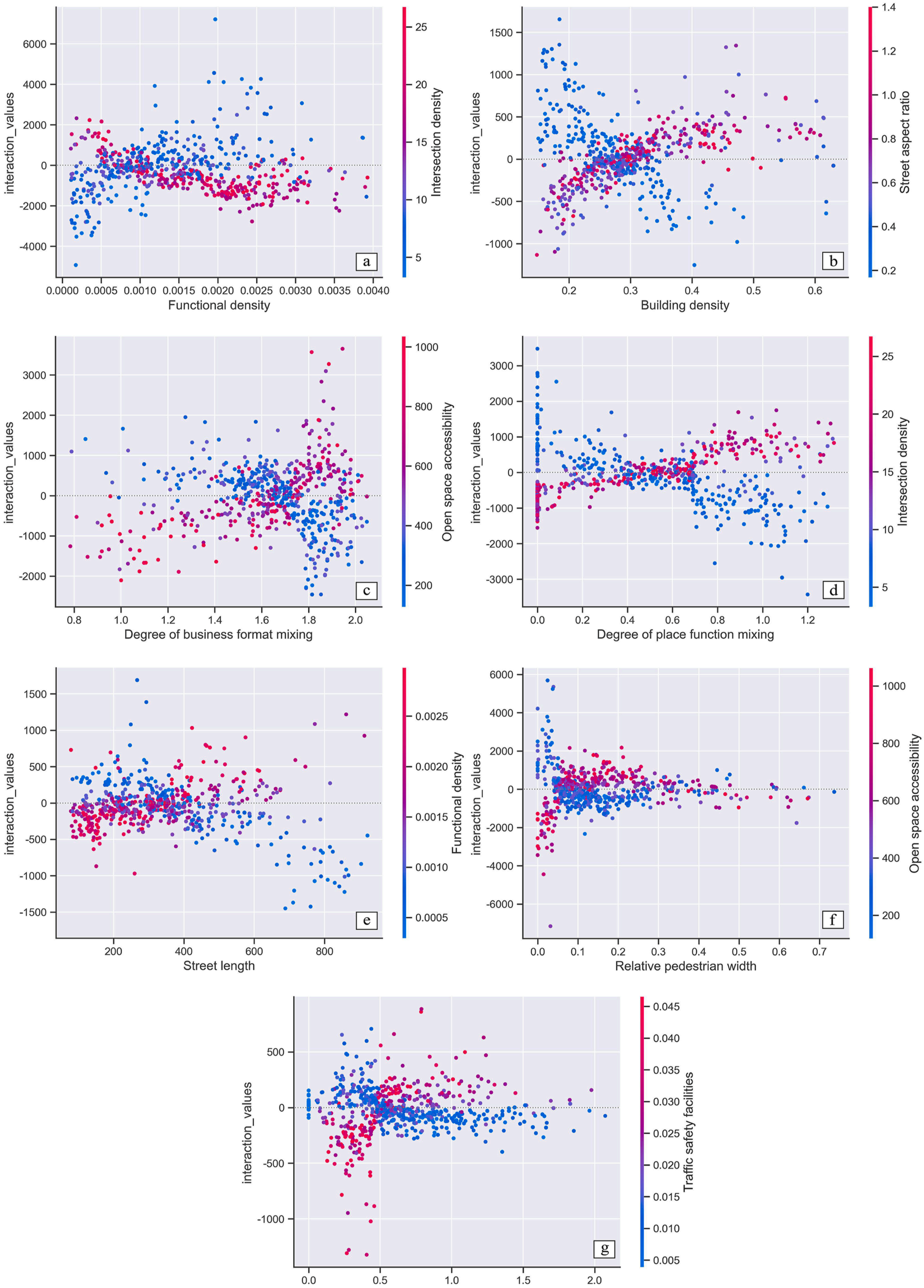

Fig. 7 Interaction effects of built environment indicators.

axis represents an independent variable, the color indicates the magnitude of the other independent variable that has a strong interaction with it, and the *y*-axis represents the SHAP interaction effect value between the two independent variables.

6.1. Function density and intersection density

As shown in Fig. 7(a), there is a reverse interaction effect between function density and intersection density, meaning that the impact of function density on street vitality varies under different intersection densities. At low intersection density (represented by the blue sample points), as function density increases, the SHAP interaction value also rises, gradually changing from negative to positive. This indicates that higher function density has a stronger positive effect on street vitality. For example, on Lumo Road in the East Lake High-tech Development Zone, the surrounding intersection density is low, which results in less traffic flow and fewer disturbances. Pedestrians can move more freely along the street, creating an environment that provides more space for various functions to thrive. In this setting, increasing function density can enhance the functional attraction of the street. However, there is diminishing marginal returns with increasing function density, meaning that higher density does not always lead to better outcomes.

In contrast, at high intersection density (represented by the red sample points), an increase in function density leads to a decrease in the SHAP interaction value, causing a negative impact on street vitality. For example, on Pingjiang East Road in Xinzhou District, the high number of intersections increases the possibility of traffic congestion and reduces pedestrian safety and comfort. Furthermore, the over-dispersed distribution of commercial and service functions makes it harder for them to concentrate, which weakens the flow of pedestrians and vehicles, thereby inhibiting functional clustering and the overall vitality of the street. In areas with high intersection density, excessively high function density increases the risk of overdevelopment and overuse. A moderate functional distribution, however, helps balance traffic and pedestrian flow, making the street network more suitable for the street's capacity and actual needs, thereby effectively enhancing street vitality.

6.2. Building density and street aspect ratio

As shown in Fig. 7(b), there is a reverse interaction effect between building density and street aspect ratio (height-to-width ratio of streets). When building density is low, a lower street aspect ratio results in a positive SHAP interaction value, promoting street vitality. In contrast, a higher street aspect ratio produces a negative SHAP interaction value, inhibiting street vitality. For example, on Wuzhong Street in Dongxihu District, the surrounding modern residential areas feature low building density and large spacing between buildings, creating an open environment. However, due to the taller buildings, the higher street aspect ratio creates a strong vertical spatial feeling, leading to a sense of oppression, which suppresses street vitality.

On the contrary, in the case of high building density, the high street aspect ratio SHAP interaction value is greater than 0, which promotes street vitality, otherwise it inhibits street vitality. For example, Sixian Road in Jiangxia District has a high density of surrounding buildings, and the building height is often low. At this time, the high aspect ratio of the street makes the street space more compact, but it does not produce a sense of oppression. Instead, it concentrates more functional facilities, creating an environment with dense crowds, compact space and rich functions, bringing a good sense of scale and comfortable walking experience to pedestrians, thus promoting street vitality. If the height and width of the street are relatively low, the space may appear too loose, which will weaken the vitality of the street.

6.3. Degree of business format mixing and open space accessibility

As shown in Fig. 7(c), the degree of business format mixing and open space accessibility exhibit an inverse interaction effect. When open space is within a 500 m range, an excessively high degree of business format mixing results in a negative SHAP interaction value, showing a suppressive effect on street vitality. Only when the degree of business format mixing is controlled below 1.7 can it effectively promote street vitality. For example, in Tenglong Avenue in Huangpi District, near Hohai, large commercial plazas along the street are primarily focused on dining and shopping formats. In this case, a lower degree of business format mixing can more effectively complement the open space, providing more concentrated commercial and service functions, thereby attracting foot traffic. Although a higher degree of business format mixing offers more options, it may lead to functional redundancy and dispersed foot traffic, weakening overall street vitality. Therefore, a lower degree of business format mixing is generally more advantageous in this situation. However, this does not imply complete uniformity—moderate diversity remains important.

When open space exceeds 500 m, residents can no longer easily walk to public spaces, and the street itself needs to take on more social interaction and activity functions. The interaction effect between the two increases with the degree of business format mixing. When the degree of business format mixing is high (greater than 1.7), the SHAP interaction value becomes positive, promoting street vitality. In this case, the street can offer more functions and options to compensate for the lack of open space, boosting vitality. The promoting effect reaches its peak when the degree of business format mixing is around 1.9. Beyond this level, diminishing marginal returns may occur.

6.4. Degree of place function mixing and intersection density

As shown in Fig. 7(d), the degree of place function mixing and intersection density exhibit a reverse interaction effect. When intersection density is below approximately 10 intersections/km^2, the street network is relatively sparse, and traffic mobility is low. In this case, the SHAP interaction value decreases as the degree of place function mixing increases. When the degree of place function mixing exceeds around 0.5, the SHAP interaction value shifts from positive to negative, suppressing street vitality. Streets with low intersection density, due to insufficient accessibility, may weaken street vitality. However, the concentration of single or dominant functions is more likely to form a "gathering effect," enabling people to quickly meet their needs in concentrated functional areas. This concentration of pedestrian flow enhances interaction and vitality on the street to some extent. This phenomenon is reflected in the streets of the Hanko North Commercial Trading Unit in the Houhu area. The Hanko North Commercial Trading Center, known as the "New Hanzheng Street," achieves a full range of commercial activities ("clothing, food, housing, travel, entertainment, shopping, and leisure") and differs significantly from typical commercial streets in scale, layout, functional positioning, and operation mode. The concentration of commercial functions in the Hanko North Trading Unit attracts a large crowd, compensating for the negative impact of low intersection density.

When intersection density is higher, i.e., above approximately 10 intersections/km^2, the street network becomes more complex, connectivity increases, and traffic mobility improves, making it easier for pedestrians and vehicles to pass through. In this case, the interaction value increases with the degree of place function mixing. When the degree of place function mixing exceeds around 0.5, the interaction value shifts from negative to positive, promoting street vitality. For example, Shufan Street in the old city area of Caidian District, with a high intersection density and a high degree of place function mixing, provides convenient services to local residents. In this high-density street network environment, the diverse building layouts align better with the dense transportation network. A high degree of place function mixing offers a variety of services and functions, attracting a diverse crowd, while the high intersection density facilitates the flow of people between different functional areas, thereby enhancing street vitality. This indicates that when intersection density is high, an increase in the degree of place function mixing can positively promote street vitality. When the degree of place function mixing is relatively high, intersection density should also be appropriately increased.

6.5. Street length and functional density

People perceive and react to space differently depending on the length of the street. As shown in Fig. 7(e), when street length is relatively short (less than about 400 m), a lower functional density corresponds to a positive SHAP interaction value, which promotes street vitality. However, higher functional density results in a negative SHAP interaction value, suppressing street vitality. On shorter streets, the sense of open space is more important, and lower functional density contributes to a better sense of openness, which can attract pedestrian flow. For example, the middle section of Maoyuan Road in Caidian District, which is short and separated by intersections, becomes overcrowded due to dense land use, and this results in a conflict with heavy traffic flow, leading to a reduction in street vitality.

On the other hand, when street length is longer (greater than about 400 m), the situation reverses: higher functional density corresponds to a positive SHAP interaction value, promoting street vitality. In longer streets, more functions are needed to attract and maintain pedestrian interest and flow. Increasing functional density effectively compensates for the monotony caused by the spatial distance. This phenomenon is quite common in urban street designs.

6.6. Relative pedestrian width and open space accessibility

As shown in Fig. 7(f), when the distance to open spaces is short (less than about 400 m), a lower relative pedestrian width corresponds to a positive SHAP interaction value, which promotes street vitality. As the relative pedestrian width increases, the interaction value decreases rapidly and approaches zero. This is because, when open spaces are within close proximity, a large number of pedestrians tend to go directly to these spaces instead of lingering on the surrounding streets for long periods. Therefore, in such cases, wider vehicular lanes are more beneficial for accessibility, and excessively wide pedestrian spaces are not necessary to accommodate more pedestrian activity. On the other hand, narrower pedestrian spaces create a "spatial compression" effect, encouraging more frequent interaction between pedestrians and the street, which increases pedestrian density and overall vitality.

When the distance to open spaces is larger (greater than about 400 m), the influence of relative pedestrian width on street vitality weakens. Lower relative pedestrian widths correspond to a negative SHAP interaction value, which suppresses street vitality. This could be because, when open spaces are farther away, people rely more on the street's inherent functions and characteristics. A lower relative pedestrian width may indicate a wider roadway and a narrower sidewalk, where the restricted walking space limits movement and activities, resulting in a poor walking experience. As the relative pedestrian width increases, the interaction value rises rapidly and eventually levels off. Therefore, the relative pedestrian width of streets should be set within an appropriate range to ensure balanced accessibility and pedestrian comfort.

6.7. Street aspect ratio and traffic safety facilities proportion

The street aspect ratio influences how pedestrians perceive and respond to space. As shown in Fig. 7(g), when the street aspect ratio is low (below about 0.5), a lower proportion of traffic safety facilities corresponds to a positive SHAP interaction value, which promotes street vitality. Conversely, an overabundance of traffic safety facilities tends to suppress vitality. This may be because a more open environment makes people feel comfortable and more willing to engage in spontaneous activities, whereas excessive safety features may disrupt this sense of openness.

When the street aspect ratio is high (above about 0.5), a higher proportion of traffic safety facilities corresponds to a positive SHAP interaction value, promoting street vitality. Conversely, a lack of such facilities may weaken vitality. This could be due to the increased sense of spatial tightness in areas with higher aspect ratios, which may reduce the feeling of safety. In such cases, traffic safety facilities help alleviate discomfort and make people more willing to engage in activities on the street.

7. Street vitality formation mechanism

Revealing the mechanisms of street vitality formation can provide valuable references for refined street design. This study uses hierarchical clustering to explore the impact mechanism of built environment variables on street vitality by analyzing the similar patterns of local effects. SHAP values of various indicators are used as clustering standards, and the similarity ranking of samples is based on the similarity in feature contribution.

As shown in Fig. 8 (top right), the street samples are ordered along the *x*-axis based on similarity, and the corresponding Shapley values are stacked along the *y*-axis. Positive Shapley values (red bars) represent positive contributions to vitality, while negative Shapley values (blue bars) suppress vitality. The final predicted value is the intersection of the red and blue bars. When the distance threshold is set to 40, three clusters are obtained, accounting for 8.47%, 46.9%, and 44.63%, respectively. Cluster 1 (high vitality—density-oriented) has the highest vitality prediction value, the fewest samples, and is mainly distributed in the Houhu area of Huangpi District. Cluster 2 (medium vitality—location-limited) has a moderate vitality prediction value and is mainly concentrated in the old town area of Caidian District and the Yangluo old town living unit in Xinzhou District. Cluster 3 (low vitality—transportation-limited) has the lowest vitality prediction value, the most samples, and is distributed across various fringe area commercial streets, with a more concentrated distribution in the Zifang old town area of Jiangxia District. A typical street is randomly selected from each cluster to further analyze the characteristics of each cluster (Fig. 8, bottom right).

Sample 1 is located in the middle section of Sheyang Street in Huangpi District, exhibiting a high vitality prediction value and belonging to the high vitality-density-oriented type. Its high functional density is the primary contributing factor to street vitality. The west side of the street is close to Shekou Market, with various types of businesses along the street. Although it is over 2000 m away from the subway station, it still has a positive influence, with the density of bus stations contributing secondarily. In the future, the advantage of high functional density should be leveraged by fine-tuning the zoning of street-side shops and services, attracting more crowds, increasing the density of nearby subway stations, and optimizing details like roadside parking and street furniture arrangements to enhance the street's quality.

Sample 2 is located in the northern section of Xinfulu in the old town of Caidian District, with a moderate vitality prediction value, belonging to the medium vitality-location-limited type. Functional density and accessibility to open spaces promote vitality, while distant location and high intersection density have a suppressive effect. For such streets, a community-level commercial center should

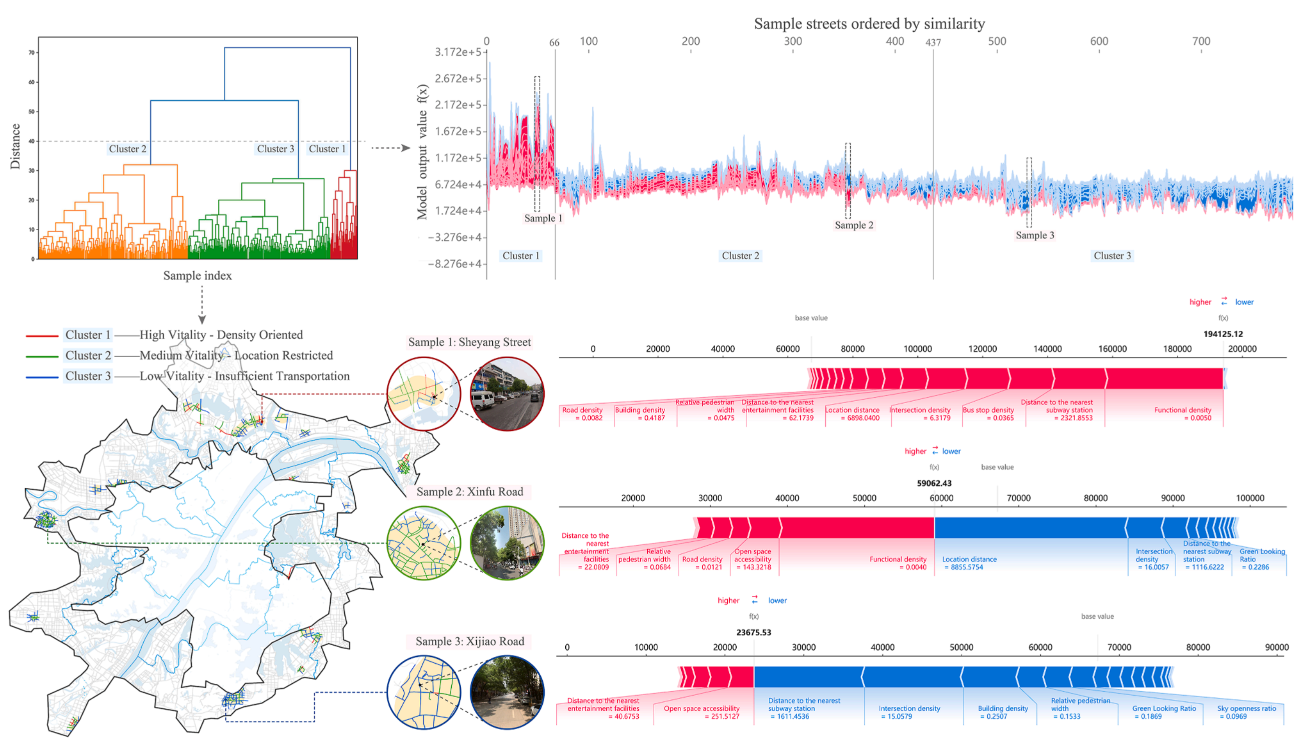

Fig. 8 Spatial distribution of street vitality types and typical streets.

be built within the area to compensate for location limitations, further increasing functional density and thereby boosting the overall vitality of the street.

Sample 3 is located in the northern section of Xijiao Road in the old town of Jiangxia District, with the lowest vitality prediction value, belonging to the low vitality-transportation-limited type. The distant subway station and dense intersection density are key factors suppressing street vitality. For streets in such areas, shuttle services to the subway stations can be added to reduce the adverse effects of the long distance. Additionally, optimizing intersection layouts by setting pedestrian safety islands, designing intersection channelization, and improving signage systems will improve pedestrian experience, enhance street vitality, and increase safety.

8. Conclusion

In this study, the streets in the commercial district of Wuhan fringe are taken as the research object, and the nonlinear relationship and synergistic effect between built environment and street vitality are discussed by using LightGBM, XGBoost and SHAP algorithms. First, kernel density analysis is used to define the boundaries of the commercial districts, and mobile signaling data is employed to quantify street vitality and analyze its spatiotemporal distribution characteristics. Next, a "5D" built environment indicator system is established, and modeling is performed for different time periods, with local interpretation provided by SHAP algorithms to reveal the relative importance of built environment indicators, their nonlinear relationships, and interaction effects on street vitality. In addition, hierarchical clustering is used to identify street vitality types, and optimization suggestions are proposed to enhance street vitality. The main conclusions of the study are as follows:

(1) In the urban fringe areas, street vitality is primarily influenced by factors such as location distance, intersection density, accessibility to open spaces, proximity to the nearest subway station, and functional density. These factors reveal that the vitality of streets in fringe areas relies not only on physical connections to the city center but also on the rational layout of functions and public spaces within the region. Specifically, functional density has a positive correlation with vitality, while location distance, intersection density, and open space accessibility show negative correlations. The impact of proximity to the subway station is more complex. During different time periods of the day, the influence of various variables on street vitality shows dynamic changes. In the morning during the peak commuting hours, the impact of subway station proximity is significant. In the evening, the importance of open spaces is highlighted, reflecting the increased demand for recreational spaces after work. During the day, the promoting effect of functional density on vitality is the greatest, indicating that people's demand for commercial and service functions increases during this period.

(2) The local dependence charts show that there is a complex nonlinear relationship between various built environment variables and street vitality. The intensity of each variable's impact on street vitality is not constant but changes dynamically, often exhibiting one or more turning points. Beyond these turning points, the local effect may shift from negative to positive or from positive to negative. For example, when functional density exceeds about 15 per hectare, the SHAP value becomes greater than 0, beginning to have a positive effect on street vitality, and the promoting effect of functional density on vitality gradually strengthens. Additionally, this influence may also have a marginal effect, meaning that after the variable reaches a certain threshold, further increases result in a diminishing effect on vitality. For instance, when building density exceeds about 0.5, the SHAP value remains positive but tends to stabilize, even showing signs of diminishing marginal returns.

(3) The local effect of one built environment variable can be amplified or diminished by changes in another variable, meaning there is an interaction effect between the two built environment indicators that jointly influence street vitality. For example, in areas with low intersection density, the greater the functional density, the stronger its positive impact on street vitality. However, in areas with high intersection density, an increase in functional density can negatively affect street vitality. Therefore, a moderate level of intersection density and functional distribution can allow the street network to better balance traffic and pedestrian flow, aligning more closely with the street's capacity and actual needs, thus effectively enhancing street vitality. When designing streets, it is important to consider the synergistic relationships between multiple dimensions or indicators, not just focusing on the impact of individual elements on street vitality, but also integrating the combined effects of various indicators.

(4) Samples with similar built environment variable contributions tend to show similar vitality levels and exhibit spatial clustering. Clustering results indicate that low-vitality streets dominate, with negative effects of the built environment generally outweighing the positive ones. Vitality formation mechanisms differ across street types, suggesting that targeted recommendations should be made based on the nonlinear relationships and interaction effects between built environment indicators and vitality. Streets with medium and low vitality should be prioritized. For example, the northern section of Xinfu Road can benefit from a community-level commercial center to address its distant location, while increasing functional density to enhance vitality. In contrast, the northern section of Xijiao Road needs a denser intersection layout and improved intersection design to boost vitality and pedestrian safety.

This study integrates multi-source data and machine learning algorithms to move beyond traditional linear frameworks, adhering to the objective patterns of spatial vitality formation. It systematically reveals the complex influence mechanisms of the built environment on street vitality in urban fringe areas, thereby enriching the theoretical depth of urban spatial vitality research. At the theoretical level, the study introduces high-resolution mobile phone signaling data at the street scale to enhance the spatial accuracy of vitality measurement, verifies the applicability of the "5D" built environment theory in fringe areas, and employs the SHAP algorithm to capture the nonlinear thresholds and interaction effects of variables—thereby expanding the explanatory framework of the built environment—street vitality relationship. At the practical level, the findings offer data support for urban renewal, traffic organization, and public space configuration, aiding the formulation of fine-grained, street vitality-oriented intervention strategies. Especially under resource constraints, identifying key variables and their effect boundaries can assist planning departments in implementing differentiated and precise spatial interventions for different street types—such as enhancing the functional density of low- and medium-vitality streets, optimizing intersection layouts, or improving public space design—thus achieving optimal resource allocation for vitality improvement. However, this study also has some limitations. First, the measurement method for street vitality still requires improvement, as allocating staying population based on spatial area ratios may introduce estimation bias. Second, the study is limited to commercial districts in fringe areas, without fully considering street typologies, boundary characteristics, and spatial diversity, so the generalizability of the results remains to be tested. Third, the study focuses on correlation analysis and has yet to fully uncover the causal mechanisms behind the formation of street vitality. Therefore, future research could expand in directions such as cross-city comparison, street classification, and causal modeling. Despite these limitations, this paper makes meaningful contributions in terms of vitality measurement methodology, theoretical development, and practical orientation.

Author contribution

Jingmei Shao, Yan Long, and Xuejun Liu contributed equally to this work. Jingmei Shao, Yan Long, and Xuejun Liu were responsible for the research design, model construction, and drafting of the initial manuscript. Yuqiao Zheng contributed to data processing and manuscript translation. Yang Song and Jian Wang participated in data collection and interpretation of the results. Bo Liu and Jun Yang provided technical guidance and critical revisions. Yujie Chen and Fan Zhang supervised the study and finalized the manuscript. All authors have read and approved the final version of the manuscript.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This research was funded by the Open Fund Project of Hubei Cultural Creativity Industrialization Design Research Center in 2023 (Grant No. HBCY2304), the 2023 Hubei Provincial International Science and Technology Cooperation Project (Grant No. 2023EHA032), the 2024 Hubei Construction Science and Technology Program (Grant No. JK2024047), and the 2024 Hubei Provincial Department of Education Humanities and Social Sciences General Project (Grant No. 24Y099). In addition, this research was also funded by the 2023 Hubei Provincial Construction Science and Technology Program (Grant No. 1656-095) and the School-level College Students' Innovation and Entrepreneurship Training Program of Wuhan University of Science and Technology (Grant No. 24116).

References

Chen, T., Guestrin, C., 2016. XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785—794.

Cui, G., Wu, J., 1990. The spatial structure and development of Chinese urban fringe. Acta Geogr. Sin. 45, 399—411.

Cui, T., Liu, J., Li, W., Xu, S., Luo, A., 2020. Research on extraction of multiple commerce center based on contour tree method. Sci. Surv. Mapp. 45, 150—155, 177.

Fei, B., 2019. Study on the Relationship Between Complex Architecture Interface and Surrounding Street Vitality in Commercial Centers: Chongqing Case Studies. Master's Thesis. Chongqing University.

Feng, C., Zhang, H., Xiao, L., Guo, Y., 2022. Land use change and its driving factors in the rural—urban fringe of beijing: a production—living—ecological perspective. Land 11, 314.

Gehl, J., 2001. Life Between Buildings: Using Public Space. The Danish Architectural Press, Cph.

Gong, Z., Ma, Q., Kan, C., Qi, Q., 2019. Classifying street spaces with street view images for a spatial indicator of urban functions. Sustainability 11, 6424.

Gou, A., Wang, J., 2011. SD method based street space vitality evaluation. Planners 27, 102—106.

Gu, C., Xiong, H., 1989. A Brief Discussion on the Study of Urban Fringe Areas, vol. 8, pp. 95—101.

Guan, Y., Li, X., Li, S., Sun, H., Liu, H., 2022. Effect of urban fringes green space fragmentation on ecosystem service value. PLoS One 17, e0263452.

Guo, X., Chen, H., Yang, X., 2021. An evaluation of street dynamic vitality and its influential factors based on multi-source big data. ISPRS Int. J. GeoInf. 10, 143.

Hu, Q., Wang, M., Li, Q.Q., 2014. Urban hotspot and commercial area exploration with check-in data. Acta Geod. Cartogr. Sinica 43, 314—324.

Huang, Z., Li, Z., Lang, W., 2023. Street space quality measurement based on multi-source big data and its impact on street vitality: a case study of Guangzhou's historic urban area. Shanghai urban plan. Rev 122—130.

Jacobs, J., 1961. The Death and Life of Great American Cities. Random House, New York.

Jiang, Y., Han, Y., Liu, M., Ye, Y., 2022. Street vitality and built environment features: a data-informed approach from fourteen Chinese cities. Sustain. Cities Soc. 79, 103724.

Ju, Q., Yin, H., Li, W., Zhou, P., 2019. Spatial scope of the urban-rural fringe in Haikou. Nat. Sci. J. Hainan Univ. 37, 180—185.

Katz, P., 1994. Toward an Architecture of Community. McGraw-Hill Education, New York.

Ke, G., Meng, Q., Finely, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T.-Y., 2017. LightGBM: a highly efficient gradient boosting decision tree. In: Presented at the Advances in Neural Information Processing Systems 30 (NIP 2017.).

Li, Jilong, Lin, S., Kong, N., Ke, Y., Zeng, J., Chen, J., 2024a. Nonlinear and synergistic effects of built environment indicators on street vitality: a case study of humid and hot urban cities. Sustainability 16, 1731.

Li, Jinze, Song, Z., Wen, J., Cai, C., Tang, P., 2024b. Exploring nonlinear relationship between built environment and street vitality using machine learning: a case Study of Ding Shu, China. In: Presented at the CAADRIA 2024: Accelerated Design, Singapore, pp. 375—384.

Li, K., Lin, Y., 2023. Exploring the correlation between streetscape and economic vitality using machine learning: a case study in the old urban district of Xuzhou, China. ISPRS Int. J. GeoInf. 12, 267.

Li, M., Pan, J., 2023. Assessment of influence mechanisms of built environment on street vitality using multisource spatial data: a case study in Qingdao, China. Sustainability 15, 1518.

Li, Q., Cui, C., Liu, F., Wu, Q., Run, Y., Han, Z., 2021. Multidimensional urban vitality on streets: spatial patterns and influence factor identification using multisource urban data. ISPRS Int. J. GeoInf. 11, 2.

Li, S., Bai, R., 2005. Determination of zone feature in fringe of a big city based on fuzzy overall evaluation. J. China Agric. Univ. 10, 99—104.

Li, Y., Yabuki, N., Fukuda, T., 2022. Exploring the association between street built environment and street vitality using deep learning methods. Sustain. Cities Soc. 79, 103656.

Liu, M., 2022. Research on the Relationship Between Job and Housing and its Influencing Factors of New City Space in the Fringe of Big City Based on multi-source Data -A Case Study of Huashan Ecological New City in Wuhan. Master's Thesis. Wuhan University of Technology.

Liu, S., Ge, J., Ye, X., Wu, C., Bai, M., 2023. Urban vitality assessment at the neighborhood scale with geo-data: a review toward implementation. J. Geogr. Sci. 33, 1482—1504.

Liu, W., Yang, Z., Gui, C., Li, G., Xu, H., 2025. Investigating the nonlinear relationship between the built environment and urban vitality based on multi-source data and interpretable machine learning. Buildings 15, 1414.

Long, Y., Lu, Z., Hu, S., Luo, S., Liu, Xi, Shao, J., Zheng, Y., Liu, Xuejun, 2023. Study on influencing factors and planning strategies of population spatial distribution in urban fringe areas from the perspective of built environment—the case of Wuhan, China. Land 12, 1739.

Long, Y., Zhou, Y., 2016. Quantitative evaluation on street vibrancy and its impact factors: a CaseStudy of Chengdu. New Archit 52—57.

Louis, H., 1936. Die geographische Gliederung vom Gross-Berlin. Bayerische Staatsbibliothek 146—171.

Lundberg, S.M., Lee, S.-I., 2017. A unified approach to interpreting model predictions. In: Proceedings of the 31st International Conference on Neural Information Processing Systems, NIPS'17. Curran Associates Inc., Red Hook, NY, USA, pp. 4768—4777.

Luo, X., Jin, X., Liu, X., Zhang, S., Ying, S., Zhou, Y., 2024. Mechanism and model of comprehensive land consolidation promoting urban-rural integration in peri-urban from the perspective of symbiosis theory. J. Nat. Resour. 39, 1053—1067.

Lyu, G., Angkawisittpan, N., Fu, X., Sonasang, S., 2025. Investigating the relationship between built environment and urban vitality using big data. Sci. Rep. 15, 579.

Maas, P.R., 1984. Towards a Theory of Urban Vitality. University of British Columbia.

Marcus, L., 2010. 'A Hymn to Movement': the 'City Symphony' of the 1920s and 1930s. Mod. Cult. 5, 30—46.

Min, Z., Ding, F., 2020. Analysis of temporal and spatial distribution characteristics of street vitality based on baidu thermal diagram:the case of the historical City of Nanchang City, Jiangxi Province. Urban Dev. Stud. 27, 31—36.

Muschalik, M., Baniecki, H., Fumagalli, F., Kolpaczki, P., Hammer, B., Hüllermeier, E., 2024. Shapiq: shapley interactions for machine learning.

Niu, T., Qing, L., Han, L., Long, Y., Hou, J., Li, L., Tang, W., Teng, Q., 2022. Small public space vitality analysis and evaluation based on human trajectory modeling using video data. Build. Environ. 225, 109563.

Niu, X., Wu, W., Li, M., 2019. Study on the impact of built environment on street vitality and its spatiotemporal characteristics based on LBS location data. Int. J. Urban Plan. 34, 28—37.

Si, R., Lin, Y., Xiao, Z., Ye, Y., 2021. Spatio-temporal analysis of built environment and street vitality relationship based on street-level imagery: a case study of Futian District, Shenzhen. Sci. Geogr. Sin. 41, 1536—1545.

Sui, C., Lu, W., 2021. Study on the urban fringe based on the expansion—shrinking dynamic pattern. Sustainability 13, 5718.

Tong, M., She, J., Tan, J., Li, M., Ge, R., Gao, Y., 2020. Evaluating street greenery by multiple indicators using street-level imagery and satellite images: a case study in Nanjing, China. Forests 11, 1347.

Wang, F., Gao, X., Xu, Z., 2015. Identification and classification of urban commercial districts at block scale. Geogr. Res. 34, 1125—1134.

Wei, H., Wang, G., 2024. Investigating the spatiotemporal pattern between street vitality in historic cities and built environments using multisource data in Chaozhou, China. J. Urban Plan. Dev. 150, 05024027.

Wu, W., Ma, Z., Guo, J., Niu, X., Zhao, K., 2022a. Evaluating the effects of built environment on street vitality at the city level: an empirical research based on spatial panel durbin model. Int. J. Environ. Res. Public. Health 19, 1664.

Wu, W., Ma, Z., Guo, J., Zhao, K., 2022b. Nonlinear effect of built environment on street vitality: a multi-source big data analysis based on XGBoost model. Chin. Landsc. Archit. 38, 82—87.

Wu, W., Niu, X., Li, M., 2021. Influence of built environment on street vitality: a case study of West Nanjing road in Shanghai based on Mobile location data. Sustainability 13, 1840.

Xiao, L., Lo, S., Liu, J., Zhou, J., Li, Q., 2021. Nonlinear and synergistic effects of TOD on urban vibrancy: applying local explanations for gradient boosting decision tree. Sustain. Cities Soc. 72, 103063.

Xie, B., Fu, B., 2023. The evolution of the concept of urban fringe areas and the reconstruction of classification systems. Arid Land Geogr. 46, 1903—1914.

Xu, Q., Zhou, X., Zhou, S., 2024. Range identification of urban fringe areas and evaluation of urban vitality based on multisource data: taking the main urban area of Hangzhou city as an example. Urban Dev. Stud. 31, 80—88.

Yang, D., Wang, X., Han, R., 2023. Nonlinear and synergistic effects of the built environment on street vitality: the case of Shenyang. Urban Plan. Forum 93—102.

Yang, J., Zhu, J., Sun, Y., Zhao, J., 2019. Delimitating urban commercial central districts by combining kernel density estimation and road intersections: a case Study in Nanjing City, China. ISPRS Int. J. GeoInf. 8, 93.

Zhang, L., Ye, Y., Zeng, W., Chiaradia, A., 2019. A systematic measurement of street quality through multi-sourced urban data: a human-oriented analysis. Int. J. Environ. Res. Public. Health 16, 1782.

Zhang, Z., Zhao, L., Zhang, M., 2024. Exploring non-linear urban vibrancy dynamics in emerging new towns: a case study of the Wuhan metropolitan area. Sustain. Cities Soc. 112, 105580.

Zhao, P., Wan, J., 2021. Land use and travel burden of residents in urban fringe and rural areas: an evaluation of urban-rural integration initiatives in Beijing. Land Use Policy 103, 105309.

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. Sustain. Cities Soc. 50, 101605.

Zhou, X., Liu, Y., Yang, H., 2017. Spatial recognition and boundary region division of urban fringe area in Xi'an City. J. Geo-Inf. Sci. 19, 1327—1335.

Zou, H., Liu, R., Cheng, W., Lei, J., Ge, J., 2023. The association between street built environment and street vitality based on quantitative analysis in historic areas: a case study of Wuhan, China. Sustainability 15, 1732.