ACCSCIENCE PUBLISHING

Journal of Chinese Architecture and Urbanism

ORIGINAL ARTICLE

Heritage-sensitive urban vitality assessment: A transferable framework for historic pedestrian districts

Jijiang Zhang, Faziawati Abdul Aziz*, and Mohd Fabian Hasna

Department of Landscape Architecture, Faculty of Design and Architecture, Universiti Putra Malaysia, Serdang, Selangor, Malaysia

Abstract

Historic districts present a "data–heritage paradox": their cultural and morphological complexity demands nuanced, context-aware modeling, yet the scarcity of labeled data hinders robust, cross-regional generalization of data-driven methods. To address this challenge, we propose a heritage-sensitive transfer learning (HSTL) framework aligned with the United Nations Educational, Scientific, and Cultural Organization's historic urban landscape approach. The framework integrates multi-source urban data—geographic information system/points of interest indicators, geotagged social media activity, and street view images—to construct a transparent neighborhood vitality index as the supervised learning target. Methodologically, HSTL enforces a structured separation between generic vitality drivers (X_g; e.g., density and accessibility) and heritage-sensitive attributes (X_h; e.g., façade chromatics and conservation zoning). A "heritage floor" constraint allocates a minimum aggregate weight of 30 percent to X_h to prevent high-variance generic features from overwhelming culturally salient signals. The feature blocks are standardized and processed separately prior to multimodal fusion, with knowledge transferred from a data-rich source domain (Beijing) to data-scarce target domains (Shanghai and Guangzhou) via fine-tuning. Using region-grouped cross-validation, HSTL achieves consistent target-domain gains, including approximately 25 percent reductions in mean squared error and ~0.15 increases in R^2 relative to target-only baselines. Ablation tests show that removing X_h decreases target-domain R^2 by 8–10 percent, confirming its functional contribution to vitality prediction. Overall, HSTL resolves the data–heritage paradox by providing an operationalizable, transferable approach for cross-regional vitality assessment that explicitly safeguards heritage-specific interpretability.

Keywords: Data–heritage paradox; Heritage-sensitive transfer learning; Historic district vitality; Domain adaptation; UNESCO historic urban landscape; Cross-regional generalization

***Corresponding author:**
Faziawati Abdul Aziz
(faziawati@upm.edu.my)

Citation: Zhang, J., Aziz, F. A. & Hasna, M. F. (2026). Heritage-sensitive urban vitality assessment: A transferable framework for historic pedestrian districts. *Journal of Chinese Architecture and Urbanism*. 8(4):025520101. https://doi.org/10.36922/JCAU025520101

Received: December 24, 2025

Revised: March 8, 2026

Accepted: April 7, 2026

Published online: May 19, 2026

Copyright: © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons AttributionNon-Commercial 4.0 International (CC BY-NC 4.0), which permits all non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Publisher's Note: AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

1.1. The vitality challenge in historic urban districts

Historic urban districts are critical repositories of cultural heritage, preserving

architectural legacies, collective memory, and place-based identities while continuing to function as living parts of contemporary cities (Jacobs, 1962; Montgomery, 1998). They also contribute to tourism economies, community cohesion, and place-based development strategies (Kalfas et al., 2024; Naheed & Shooshtarian, 2022). Yet accelerated urbanization and market-driven redevelopment have led to declining vitality in many historic districts, manifested in population loss, spatial degradation, and intrusive modernization that erodes both physical fabric and cultural meaning (Gu, 2024). Achieving long-term resilience, therefore, requires assessment approaches that reconcile vitality enhancement with heritage conservation objectives.

Vitality in historic districts is inherently multidimensional. Beyond pedestrian volume or economic intensity, it is shaped by tourism dynamics, cultural programming, adaptive reuse, and the experiential quality of heritage environments (Gehl, 2010; X. Li et al., 2025; Liu et al., 2024; Montgomery, 1998). Capturing these intertwined dimensions remains challenging for conventional evaluation methods. Expert-based scoring and field surveys, while valuable for qualitative insight, are often constrained by subjectivity, coarse spatial resolution, and weak cross-regional comparability (Hu et al., 2025; Zheng et al., 2025), and they struggle to reflect temporal fluctuations such as seasonal tourism cycles or localized functional transitions.

Recent urban analytics has increasingly leveraged multi-source digital traces—including mobility data, satellite imagery, street view images, and social media footprints—to quantify vitality at fine spatial and temporal scales (Calabrese et al., 2013; C. Fang et al., 2025; Niu & Silva, 2020). Computer vision has further enabled the extraction of perceptual streetscape attributes from street view imagery (X. Li et al., 2015; Sakamoto et al., 2023). However, when applied to historic districts, these approaches face a key limitation: heritage value is not fully reducible to generic physical or activity proxies, as it is intertwined with authenticity, cultural symbolism, and conservation status that are difficult to capture without domain-aware representations and specialized data (Q. Zhang et al., 2025).

A persistent bottleneck, therefore, remains: labeled vitality data and heritage-relevant annotations are unevenly distributed. While data-rich metropolitan cores can support data-intensive models, many historic districts—particularly outside major city centers—face systematic data scarcity, which constrains the transferability of deep learning-based approaches across regions.

Transfer learning provides a promising strategy to reduce this imbalance by adapting knowledge from data-rich source domains to data-scarce target contexts (Pan & Yang, 2009; Weiss et al., 2016). Yet most transfer–learning applications in urban studies are designed for generic urban environments and rarely reflect the cultural and regulatory constraints that define historic districts. In parallel, the United Nations Educational, Scientific, and Cultural Organization (UNESCO) recommendation on the historic urban landscape (HUL) emphasizes integrating heritage values with contemporary urban processes and managing heritage as part of dynamic urban change (Bandarin & Van Oers, 2012; UNESCO, 2011). Bridging data-driven vitality assessment with the HUL perspective motivates the present study.

1.2. Current status of research

Existing studies on historic-district vitality can be broadly categorized into two strands: international research on urban vitality and heritage conservation, and empirical studies emerging from China's urban regeneration context. The following subsections review these perspectives and identify the remaining research gaps addressed in this study.

1.2.1. International perspectives on urban vitality and heritage conservation

International research on urban vitality has evolved from foundational qualitative theories (e.g., Gehl, 2010; Jacobs, 1962; Lynch, 1964) toward data-driven and computational approaches. These foundational works established key relationships between vitality, land-use diversity, density, and street accessibility. Contemporary studies have operationalized vitality using large-scale digital traces, such as mobile-phone signaling (Calabrese et al., 2013), mobility, crowdsourced datasets (Niu & Silva, 2020), and social media activity (Iranmanesh & Alpar Atun, 2020), improving spatial granularity and temporal responsiveness.

However, most data-driven models of vitality have been developed for contemporary urban environments. Applying them to historic districts raises additional challenges because conservation constraints, symbolic meanings, and experiential qualities are often underrepresented in generic indicator systems. The UNESCO HUL framework reframes heritage conservation as a dynamic process embedded within broader urban systems, emphasizing the interaction between tangible and intangible heritage, morphology, and socio-economic vitality (Bandarin & Van Oers, 2012; UNESCO, 2011). Despite growing recognition of these principles, operationalizing HUL through quantitative, reproducible indicator systems remains limited and inconsistent, and many empirical vitality models still fail to adequately incorporate heritage-

sensitive attributes (Ginzarly, Houbart, et al., 2019; K. Fang & Wu, 2025; Veldpaus et al., 2013), a limitation that persists despite growing recognition of heritage's role in urban dynamics.

1.2.2. Research progress in China: From field surveys to multi-source data integration

In China, research on the vitality of historic districts has expanded alongside large-scale urban regeneration initiatives. Methodologically, studies have progressed from static field surveys toward multi-source data integration, incorporating points of interest (POI) datasets, heat maps, and mobility traces to capture activity patterns (Shi et al., 2020; J. Zhang & Zhao, 2024).

Empirical work has yielded insights into both the drivers and distinctive patterns of vitality in these contexts. Studies suggest that land-use mix, accessibility, and functional intensity vary spatiotemporally in Chinese historic districts (Li & Miao, 2025). More importantly, recent analyses reveal that the coupling between environmental quality and vitality exhibits unique spatial configurations in historic districts, diverging significantly from patterns observed in generic urban areas (Q. Zhang et al., 2025). Concurrently, value-oriented frameworks have been proposed to balance historical significance with contemporary use (Y. Zhang & Han, 2022).

Nevertheless, two persistent limitations constrain the field. First, many studies remain localized single-city case analyses, limiting cross-regional comparability and external validity (Kou et al., 2018). Second, advanced knowledge-transfer methodologies remain underutilized, particularly for historic districts outside major metropolitan cores, where data scarcity is pervasive (Gao & Xie, 2025). These constraints hinder the development of scalable tools for comparative heritage planning.

1.2.3. Research gaps and study objectives

Synthesizing international and domestic literature, three gaps motivate this study:

Gap 1: Insufficient heritage-sensitive representation. Existing vitality models frequently prioritize generic metrics (e.g., density, POI diversity, accessibility) while underrepresenting heritage-relevant attributes such as conservation context, architectural character, and symbolic meaning. This gap is consequential: recent empirical evidence suggests that vitality mechanisms in historic districts diverge fundamentally from those in generic urban environments (Wu et al., 2022; Yu et al., 2024; Q. Zhang et al., 2025), yet prevailing models—calibrated largely on generic urban data—remain ill-equipped to capture these contextual specificities.

Gap 2: Limited cross-regional transferability. Many frameworks are developed as single-city studies—often focusing on major metropolises—and exhibit limited generalizability across regions with differing cultural, morphological, and regulatory contexts. For instance, frameworks calibrated for well-documented districts in Beijing or Shanghai may not directly transfer to smaller historic towns in Western China, where data resources and conservation regimes differ substantially (Kou et al., 2018; Y. Zhang & Han, 2022).

Gap 3: Underutilization of transfer learning for heritage contexts. Although transfer learning is widely recognized as a solution to data scarcity (Pan & Yang, 2009; Weiss et al., 2016), its application to historic-district vitality assessment remains nascent. Existing heritage-related applications focus on building-scale recognition tasks (e.g., architectural classification, facade analysis) rather than district-scale vitality. Moreover, standard domain adaptation is often heritage-agnostic: it minimizes distributional discrepancy (e.g., feature alignment) without enforcing the explicit retention of heritage-relevant signals such as symbolism, zoning constraints, and architectural character (Kouw & Loog, 2018). Without heritage-sensitive adaptation mechanisms, transferred models may inadequately capture heritage-relevant drivers of vitality, potentially leading to diagnoses or intervention recommendations that conflict with conservation objectives.

1.3. Research objectives and operationalization

To address these interconnected gaps, this study proposes a heritage-sensitive transfer learning (HSTL) framework aligned with the UNESCO HUL perspective. The primary objective is to incorporate heritage dimensions into quantitative vitality assessment and enable robust knowledge transfer across data-scarce historic districts. To achieve this, heritage sensitivity is operationalized in two complementary ways:

(i) Feature level: The framework explicitly distinguishes between generic vitality indicators (X_g) and heritage-sensitive indicators (X_h). This structured separation enables the transparent representation of cultural and architectural attributes—such as façade chromatics, conservation zoning, and traditional land-use configurations—that are often homogenized in conventional models.

(ii) Modeling level: Transfer learning is employed to adapt knowledge from a data-rich source city (Beijing) to data-scarce target cities (Shanghai and Guangzhou). Critically, the model incorporates heritage-preserving mechanisms throughout the domain adaptation process to retain the distinctive signals embedded in

heritage-sensitive inputs. These mechanisms operate in two ways: feature grouping ensures that X_h and X_g are standardized separately before fusion, preventing high-variance generic features from statistically overwhelming heritage attributes; loss function design introduces constraints (e.g., minimum contribution thresholds or regularization penalties) that explicitly safeguard the predictive influence of X_h, ensuring that the model's optimization process does not erode culturally salient signals in pursuit of generic performance gains.

1.4. Significance and contributions

Building on this dual operationalization, the study makes three interconnected contributions to the field of urban heritage analytics:

Methodological contribution: structured heritage-sensitive feature design. The framework establishes a dual-block indicator protocol that explicitly distinguishes generic drivers (X_g) from heritage-sensitive attributes (X_h). This structured separation offers a reproducible basis for constructing heritage-aware vitality indices that remain comparable across diverse regional contexts.

Empirical contribution: validation of heritage attributes as functional vitality drivers. Through systematic transfer learning experiments across Beijing, Shanghai, and Guangzhou, the study demonstrates that incorporating X_h yields significant and consistent predictive gains in data-scarce target domains (~25 percent mean squared error [MSE] reduction, +0.15 R^2). Ablation tests further confirm that these gains stem from the functional contribution of heritage attributes, not merely from increased data dimensionality.

Theoretical contribution: operationalizing UNESCO HUL principles within a computational framework. By embedding heritage-preserving mechanisms into a transfer learning pipeline, the study provides a concrete, evidence-grounded instantiation of the HUL approach—demonstrating how heritage values can be quantitatively integrated into urban dynamics modeling while retaining cultural specificity.

2. Theoretical foundations and conceptual framework

Building on the gaps identified in Section 1—namely, (i) the insufficient incorporation of heritage-relevant signals in quantitative vitality assessment, (ii) limited cross-regional transferability, and (iii) the underutilization of transfer learning for heritage contexts—this section establishes the theoretical foundations of the proposed framework. We first conceptualize historic-district vitality as a dual-natured construct comprising both generic and heritage-sensitive dimensions (Section 2.1). We then motivate transfer learning as a theoretically grounded response to the data–heritage paradox that characterizes many historic districts (Section 2.2). Finally, we present a heritage-sensitive transfer mechanism (Section 2.3) that links these principles to the operational procedures and model design detailed in Section 3.

2.1. Conceptualizing historic district vitality: Generic and heritage-sensitive dimensions

Understanding vitality in historic districts requires reconciling two analytical logics: the generic drivers of urban activity and the culturally specific attributes that sustain heritage identity. In this study, historic-district vitality is understood as the composite intensity of urban activity across economic, social, cultural, and environmental dimensions, reflecting both immediate functional performance and longer-term resilience (Fu et al., 2025; Jacobs, 1962; Montgomery, 1998). While this multidimensional view is consistent with foundational urban theory, its application to heritage contexts requires further refinement because historic districts are shaped by conservation constraints, cultural symbolism, and place-specific identity mechanisms (Bandarin & Van Oers, 2012; UNESCO, 2011).

Crucially, vitality in historic districts exhibits a dual nature that distinguishes it from generic urban neighborhoods:

(i) Generic vitality indicators (X_g): Historic districts share universal drivers of urban activity with the wider city, such as pedestrian flow intensity, land-use/POI functional diversity, service accessibility, and digitally expressed engagement (e.g., social-media activity). These drivers—denoted as the generic indicator set X_g—are widely established in urban vitality research (Jacobs, 1962; Gehl, 2010).

(ii) Heritage-sensitive indicators (X_h): In parallel, vitality in historic districts is anchored in culturally and morphologically specific attributes that shape heritage experience and conservation logic, including conservation zoning, historic morphology, architectural character, and streetscape visual qualities (e.g., façade chromatics and texture). These factors—denoted as the heritage-sensitive indicator set X_h—align with the UNESCO HUL perspective, which emphasizes integrating tangible heritage fabric with intangible socio-cultural processes and contemporary urban dynamics (Bandarin & Van Oers, 2012; K. Fang & Wu, 2025; UNESCO, 2011).

Despite the conceptual appeal of the HUL framework, its translation into quantitative assessment remains underdeveloped. Ginzarly, Houbart, et al. (2019) conducted a systematic review revealing that while HUL-oriented studies emphasize value-based approaches conceptually, many revert to generic indicators when operationalized computationally, rather than developing context-specific heritage metrics. Veldpaus et al. (2013) similarly argued that urban heritage management requires indicators that integrate tangible fabric with intangible cultural processes—dimensions that are not reducible to conventional "generic neighborhood" metrics. Consistent with these critiques, conventional vitality assessments often conflate generic and heritage-sensitive dimensions or systematically underrepresent X_h, effectively treating historic districts as generic neighborhoods governed by uniform criteria (Kou et al., 2018). This "generic model" approach has been criticized for overlooking the distinctive spatial–functional logics and value systems that define historic urban areas (Pendlebury et al., 2009; Rodwell, 2008). Importantly, this theoretical critique is substantiated by empirical evidence showing that vitality drivers in heritage contexts differ fundamentally from those in contemporary urban centers. Wu et al. (2022) reported that heritage-related factors such as architectural authenticity, conservation regulations, and cultural symbolism can outweigh generic determinants (e.g., accessibility or functional density) in shaping vitality patterns. Yu et al. (2024) similarly found that architectural elements and façade-related characteristics exert a stronger influence on street vitality in historic districts than several commonly used generic metrics. Q. Zhang et al. (2025) further documented a heritage-specific spatial pattern—"high-quality concentration internally" alongside "high-vitality concentration externally"—which challenges the positive quality–vitality association often assumed in generic urban models. Together, these findings underscore the insufficiency of operationalizing HUL through generic indicators alone and motivate the explicit X_g – X_h distinction adopted in this study.

To operationalize the HUL principle within a computational setting, the present study therefore formalizes X_g and X_h as conceptually distinct indicator blocks at the theoretical level, providing a transparent basis for indicator selection (Section 3.2), composite index construction (Section 3.3), and heritage-preserving transfer learning (Section 3.4). This theoretical distinction—capturing the dual nature of historic-district vitality through separate X_g and X_h indicator blocks—is summarized in Table 1 and translated into an operational indicator taxonomy in Section 3.2 (Table 2).

Building on this theoretical distinction, generic indicators (X_g) capture universal drivers of urban activity that support cross-regional comparability, while heritage-sensitive indicators (X_h) capture culturally and morphologically specific attributes that define heritage identity and conservation context. This dichotomous classification—summarized in Table 1—provides the conceptual basis for the operational framework detailed in Section 3.

2.2. Transfer learning and the data–heritage paradox

Building on the conceptual distinction between X_g and X_h established in Section 2.1, a central methodological challenge arises: assessing both indicator blocks is particularly difficult in historic districts, where relevant data are often scarce, unevenly distributed, and heterogeneous across cities.

This study conceptualizes this challenge as the data–heritage paradox: historic districts require context-aware modeling to capture heritage-sensitive mechanisms of vitality, yet they often lack the labeled, harmonized data required by such models (Gao & Xie, 2025; Kou et al., 2018). Two inherent challenges in digital heritage research sustain this paradox. First, many heritage attributes (e.g., authenticity, conservation condition, cultural significance) require expert annotation, which makes large-scale labeled datasets costly and difficult to assemble (Fiorucci et al., 2020). Second, standardization is constrained by cross-regional variation in cultural interpretation and regulatory regimes, producing heterogeneous criteria that resist uniform computational modeling (Ginzarly, Pereira Roders, et al., 2019). As a result, models trained in data-rich settings often exhibit poor generalizability when deployed in new historical contexts, where differences in morphology, architectural typology, and socio-cultural practices induce domain shift.

Transfer learning offers a theoretically grounded response by enabling knowledge learned from a data-rich source domain to be adapted to data-scarce target domains (Pan & Yang, 2009; Weiss et al., 2016). Its relevance here rests on two complementary premises. First, relationships encoded in X_g—for example, connectivity–flow associations or functional mix–activity coupling—often exhibit partial structural regularities across cities and can be transferred to support cross-regional comparability. Second, indicators in X_h are inherently place-specific and must remain explicitly represented during adaptation to preserve cultural meaning and conservation interpretability. Accordingly, naïve cross-regional model deployment (e.g., applying a Beijing-trained model directly to Guangzhou) is expected to underperform under domain shift, whereas transfer

Table 1. Conceptual typology of generic and heritage-sensitive vitality indicators in historic districts

Dimension	Indicator category	Specific indicators (theoretical)	Key theoretical grounding
Generic vitality (X_g)	Activity intensity	Pedestrian flow volume; crowd density (diurnal/nocturnal)	Gehl (2010); Jacobs (1962)
	Functional mix	Land-use diversity; points of interest category richness; service facility density	Jacobs (1962); Montgomery (1998)
	Accessibility	Street-network connectivity; public transit proximity; walkability indices	Gehl (2010); Lynch (1964)
	Digital activity	Social-media check-in density; geotagged post frequency	Iranmanesh & Alpar Atun (2020); Niu & Silva (2020)
Heritage-sensitive vitality (X_h)	Conservation context	Heritage protection zoning; regulatory conservation status; United Nations Educational, Scientific and Cultural Organization/national heritage designation	Bandarin & Van Oers (2012); UNESCO (2011)
	Morphological character	Historic plot pattern; street layout integrity; urban grain continuity	Conzen (1960); Whitehand (2007)
	Architectural identity	Building typology; facade composition; roofscape features; materiality	Fang & Wu (2025); Yu et al. (2024)
	Visual and aesthetic quality	Streetscape chromatic richness; facade texture; enclosure ratio; sky view factor	Li et al. (2015); Sakamoto et al. (2023)
	Cultural anchoring	Proximity to heritage landmarks; density of cultural points of interest (museums, temples, historic sites); event/festival seasonality	Ginzarly, Pereira Roders, et al. (2019); Q. Zhang et al. (2025)

Notes: This table presents the theoretical classification of vitality indicators for historic urban districts, distinguishing between generic drivers of urban activity (X_g) and heritage-specific attributes (X_h) that shape cultural distinctiveness and conservation logic. The operationalization of these indicators—including specific data sources, measurement protocols, and empirical indicators used in this study—is detailed in Section 3.2 (see Table 2 for the full operational indicator system) and Section 3.3 (see Table 3 for the composite neighborhood vitality index construction)

learning can reduce mismatch through fine-tuning and adaptation strategies that update model parameters using limited target labels. This theoretical grounding motivates the proposed heritage-sensitive transfer mechanism, which retains X_h as an explicit feature block throughout the modeling pipeline (Section 3.4) rather than implicitly assuming that generic proxies suffice.

2.3. Conceptual framework: A heritage-sensitive transfer mechanism

Transfer learning has been widely adopted in urban analytics for tasks such as building energy prediction (G. Li et al., 2024), land-use classification (Zhou et al., 2017), and traffic forecasting. Recent studies have also begun exploring transfer learning for heritage-related tasks. Xu et al. (2025) achieved 93 percent accuracy in evaluating historical appearance integrity using pre-trained computer vision models for heritage-specific visual attributes. Llamas et al. (2017) applied transfer learning to classify architectural heritage buildings from street view images, while Grilli and Remondino (2019) used it for three-dimensional point cloud classification of historic structures.

However, most existing applications emphasize physical-feature recognition at the building scale rather than district-scale holistic vitality. Moreover, standard domain adaptation approaches are often heritage-agnostic: they reduce domain discrepancy through statistical alignment (e.g., feature distribution matching) without explicitly preserving heritage-specific signals such as cultural symbolism, conservation zoning, and architectural character (Kouw & Loog, 2018). Without such preservation, transferred models risk misidentifying the drivers of vitality or recommending interventions that conflict with conservation objectives.

Synthesizing the above principles, we propose an HSTL framework that connects theoretical requirements in heritage contexts to an implementable cross-regional learning pipeline. As illustrated in Figure 1, the framework departs from conventional "one-size-fits-all" feature pooling by adopting a feature-structured transfer mechanism:

generic and heritage-sensitive indicators are explicitly distinguished at the indicator-system level (X_g versus X_h) and handled differently during representation learning and domain adaptation.

Operationally, this conceptual mechanism maps directly to three interconnected components detailed in Section 3:

(a) Feature structuring (Section 3.2; Table 2). Multi-source indicators are organized into a generic block X_g and a heritage-sensitive block X_h, ensuring that heritage signals are not subsumed by high-variance generic metrics and remain explicitly interpretable.
(b) Supervised target construction (Section 3.3; Table 3). A neighborhood vitality index (NVI) is constructed as a transparent composite label from weighted indicators derived from X_g and X_h. This supervised target provides a consistent training signal across regions and is used to supervise prediction via $\mathcal{L}\left(NVI, \widehat{NVI}\right)$.
(c) Heritage-preserving transfer learning (Section 3.4; Figure 2). For each spatial unit j, indicator blocks are instantiated as standardized feature vectors $\mathrm{x}_{g,j}$ and $\mathrm{x}_{h,j}$, together with a streetscape embedding z_j extracted street view image (SVI). A multimodal predictor $f\left(\mathrm{x}_{g,j}, \mathrm{x}_{h,j}, \mathrm{z}_j\right)$ estimates $\widehat{NVI}_j$. During cross-regional adaptation, transferable representations associated with (X_g, z) are fine-tuned, whereas X_h is retained as an explicit input block to safeguard local conservation signals from being homogenized by domain generalization.

This conceptual framework thus establishes the theoretical blueprint for the methodological procedures detailed in Section 3. By addressing the data–heritage paradox through a feature-structured transfer mechanism, the framework ensures that cross-regional assessment remains technically robust while retaining heritage-relevant semantic signals essential for historic-district planning and conservation.

3. Methodology: A heritage-sensitive transfer learning framework for cross-regional vitality assessment

This section operationalizes the theoretical principles established in Section 2—namely, the dual nature of historic district vitality represented by generic vitality drivers (X_g) and heritage-sensitive attributes (X_h), and the role of transfer learning in addressing the data–heritage paradox—through a structured and reproducible methodological framework.

The proposed approach integrates three key components:

(i) A heritage-sensitive indicator system for capturing culturally specific spatial attributes (Section 3.2),
(ii) a transparent composite index construction for defining the supervised learning target (Section 3.3), and
(iii) a transfer-learning-based modeling strategy for cross-regional prediction (Sections 3.4–3.5).

Together, these components enable comparable vitality assessment across historic districts while preserving locally embedded cultural characteristics.

3.1. Problem definition and target variable construction

This study develops a transferable, heritage-sensitive analytical framework for assessing urban vitality across historic pedestrian districts characterized by heterogeneous spatial forms, cultural contexts, and varying levels of data availability. Formally, let $D = \{D_1, D_2, \ldots, D_K\}$ denote a set of historic districts located in different regions. The objective is to learn a predictive function that maps multi-source urban indicators to a comparable vitality outcome while preserving heritage-specific characteristics relevant to historic urban environments.

To this end, the NVI is defined as the supervised learning target (OECD, 2018; Y. Zhang & Han, 2022). The NVI represents the composite intensity of urban vitality at the neighborhood or street-segment level and integrates functional, social, and cultural dimensions of urban activity relevant to historic districts.

By providing a unified outcome variable, the NVI serves as the reference metric for model training and cross-regional validation. Explicitly defining the NVI prior to model construction ensures conceptual transparency and facilitates consistent comparison across geographically and culturally diverse urban contexts. The operational construction of the NVI, including indicator aggregation and weight determination, is detailed in Section 3.3.

3.2. Heritage-sensitive indicator system and feature design

To operationalize heritage sensitivity within the proposed framework, this study constructs a structured indicator system that integrates both generic urban vitality drivers and heritage-specific spatial attributes. The indicator design follows the principles of the UNESCO HUL framework while ensuring compatibility with multi-source urban data. The resulting indicator system provides the feature foundation for constructing the NVI and training the HSTL framework.

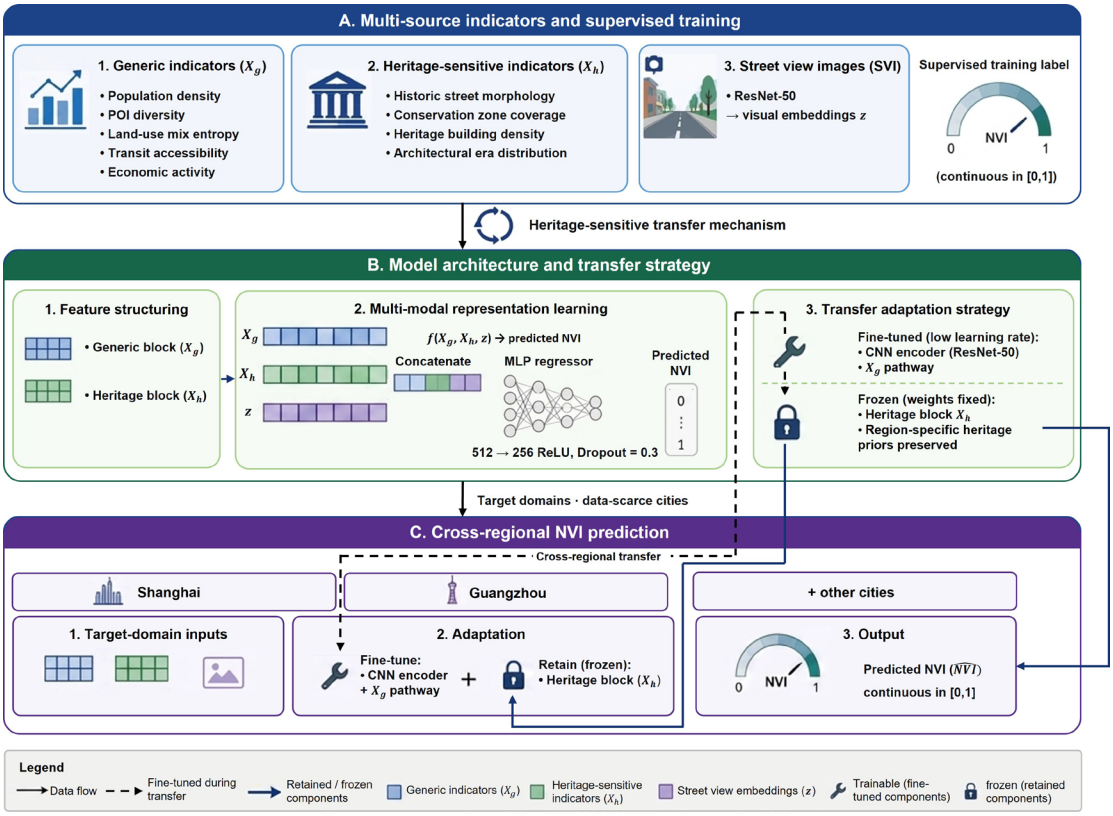


Figure 1. Conceptual diagram of the heritage-sensitive transfer learning framework for cross-regional vitality assessment. X_g = Generic vitality indicators; X_h = Heritage-sensitive indicators; z = SVI embeddings; NVI = Neighborhood vitality index (supervised target); $\widehat{NVI}$ = predicted NVI
Abbreviations: CNN: Convolutional neural network; POI: Points of interest; ReLU: Rectified linear unit; ResNet-50: Residual network with 50 layers; SVI: Street view images. Source: Diagram by the authors

3.2.1 Indicator categories

The indicator system is designed to balance universal drivers of urban vitality with heritage-sensitive contextual attributes, in accordance with the UNESCO HUL Recommendation (UNESCO, 2011). Four indicator categories are incorporated to capture complementary dimensions of vitality in historic urban environments:

- Pedestrian activity indicators, capturing observed human presence and movement intensity at the street-segment level;
- Building function and land-use indicators, derived from GIS-based land-use and POI datasets, to represent functional diversity and activity-supporting capacity;
- Social media activity indicators, utilizing georeferenced digital traces as proxies for place-based engagement and symbolic use of urban space (Noulas et al., 2012);
- Environmental and heritage context indicators, capturing spatial morphology and environmental characteristics relevant to historic districts, including street enclosure, greenery, conservation zoning, historic morphology, and façade-related attributes extracted from SVI (Gong et al., 2018) and remote-sensing proxies (e.g., green view index, normalized difference vegetation index).

To ensure comparability across heterogeneous indicators, all structured variables are standardized using z-score normalization. Group-wise standardization is applied separately to the generic indicator block (X_g) and the heritage-sensitive indicator block (X_h) to maintain unit-free comparability while preventing high-variance generic indicators from numerically dominating heritage-related signals.

3.2.2. Distinguishing generic and heritage-sensitive indicators

Heritage sensitivity is operationalized through explicit indicator structuring rather than implicit or post hoc weighting. Indicators are therefore classified into two theoretically grounded groups based on their relevance to

historic conservation contexts:

- Generic vitality indicators (X_g), representing functional and behavioral drivers commonly observed across urban environments (e.g., pedestrian activity, land-use diversity, and social media engagement);
- Heritage-sensitive indicators (X_h), capturing attributes specific to historic urban landscapes, including historical morphology, conservation regulations, architectural character, and culturally symbolic spatial elements.

This structured separation preserves heritage-related signals during cross-regional transfer learning, reducing the risk that culturally significant attributes are numerically overshadowed by high-variance generic indicators. The classification, data sources, operationalization strategies, and functional roles of all indicators are summarized in Table 2. Together, these indicators define the structured feature space used by the HSTL framework introduced in Sections 3.4–3.5.

3.3. Weight determination and structural allocation

To ensure methodological transparency and minimize subjective bias, indicator weights are determined through a structured three-stage protocol. This approach balances empirical variance structure with the conceptual requirement that heritage-sensitive attributes be substantively represented in vitality assessment, consistent with established composite index construction principles (OECD, 2018).

3.3.1. Three-stage weighting protocol

The three-stage weighting protocol partitions, weights, and normalizes indicators to balance generic and heritage-sensitive contributions.

(a) Stage 1: Group-level partitioning and heritage safeguarding

Indicators are first partitioned into two blocks: the generic vitality block (X_g) and the heritage-sensitive block (X_h), as defined in Table 2. To methodologically operationalize the UNESCO HUL principle and prevent heritage-related attributes from being numerically overshadowed by high-variance generic indicators (e.g., POI density), we introduce a heritage floor constraint. This constraint ensures that a minimum aggregate weight (e.g., 30 percent) is allocated to X_h, so that heritage attributes remain structurally represented in the composite index rather than being treated as marginal variables.

(b) Stage 2: Within-group weight estimation via principal component analysis

Within each indicator block, preliminary weights are estimated using principal component analysis (PCA). For X_g, weights are proportional to the loadings on principal components that explain at least 75 percent of cumulative variance. For X_h, squared loadings from a separate PCA are used to ensure balanced representation of distinct heritage dimensions, irrespective of raw variance magnitude (Jolliffe, 2002; Hastie et al., 2009).

(c) Stage 3: Normalization

All indicator weights are normalized to satisfy the standard probability constraint:

$$\sum_{i=1}^{n} w_i = 1, w_i \geq 0 \quad (1)$$

3.3.2. Summary of indicator weight distribution

Table 3 summarizes the empirically derived indicator weights estimated from the Beijing source domain, which serve as the baseline configuration for subsequent model training and cross-regional adaptation.

$$NVI = 0.65\, X_g + 0.35\, X_h \quad (2)$$

where

$$X_g = \sum_{i=1}^{8} w_i x_i, X_h = \sum_{j=1}^{8} w_j x_j \quad (3)$$

represent the aggregated scores of the generic vitality block and the heritage-sensitive vitality block, respectively, with weights normalized such that $\sum w_i = 1$ within each block. Indicator weights were derived through a three-stage protocol (Section 3.3.1):

(i) group-level partitioning with a heritage floor constraint ensuring a minimum 30 percent aggregate allocation to the heritage-sensitive block (X_h);
(ii) within-group weight estimation using principal component analysis; and
(iii) final normalization of indicator weights within each block.

The resulting configuration ($X_g = 0.65$, $X_h = 0.35$) serves as the baseline parameterization for model training in the Beijing source domain and subsequent cross-regional fine-tuning. The heritage floor operationalizes the UNESCO Historic Urban Landscape principle, ensuring that heritage-sensitive attributes remain substantively represented in the composite vitality index (OECD, 2018; UNESCO, 2011).

3.3.3. Sensitivity and robustness analysis

To assess whether the NVI is unduly sensitive to the weighting configuration, a perturbation-based sensitivity analysis was conducted. Each indicator weight was

independently perturbed within ± 20 percent of its baseline value (as reported in Table 3), followed by re-normalization to maintain $\sum_i w_i = 1$. This baseline configuration, derived from the Beijing source domain, provides a stable reference point for assessing the NVI's sensitivity to alternative weighting schemes.

For each perturbation scenario, NVI values were recalculated and evaluated in terms of (i) rank-order stability, measured by Spearman's ρ, and (ii) predictive stability when the perturbed NVI was used as the supervised target under the same cross-regional validation protocol described in Section 3.5.3.

Across perturbation scenarios, ranking stability remained high (ρ > 0.92), indicating that relative vitality patterns are largely preserved under moderate weight uncertainty. Model performance fluctuations were small, with typical changes of $|\Delta R^2| < 0.03$ and $|\Delta MSE| < 0.02$, suggesting that the learning pipeline is not overly sensitive to minor reallocation of indicator weights.

To further examine the functional role of heritage-sensitive indicators, the aggregate weight allocated to X_h was progressively reduced. When the heritage allocation dropped below approximately 20 percent, cross-regional predictive performance tended to deteriorate (e.g., $\Delta R^2 \approx -0.08$ in representative runs, with similar patterns observed across multiple random initialization seeds). This finding is consistent with the premise that heritage-sensitive attributes provide non-redundant information under domain shift—a result that directly supports the theoretical distinction between X_g and X_h established in Section 2.1 and the data–heritage paradox conceptualized in Section 2.2.

This analysis serves as a robustness diagnostic rather than a causal attribution and is complemented by the ablation comparisons reported in Section 4.3, which quantify the predictive contribution of heritage-sensitive attributes under cross-regional transfer.

3.4. Transfer learning architecture and adaptation strategy

Building on the heritage-sensitive indicator system established in Section 3.2 and the weighted NVI target defined in Section 3.3, this section details the transfer learning architecture for operationalizing cross-regional vitality prediction. The proposed framework integrates structured indicators (X_g, X_h) with visual features extracted from SVI within a multimodal fusion pipeline. Crucially, the architecture implements the heritage-preserving mechanism introduced conceptually in Section 2.3: parameters associated with domain-invariant components (X_g and SVI encoder) are fine-tuned during cross-regional transfer, while X_h is retained as an explicit input block using local target-domain data.

3.4.1. Model architecture

To enable cross-regional transferability while preserving local heritage specificity, a supervised transfer learning model is implemented that integrates streetscape image representations with structured multi-source indicators. SVI is processed using a convolutional neural network (CNN) backbone, while non-image indicators are treated as tabular inputs.

A 50-layer residual network (He et al., 2016), initialized with ImageNet pre-trained weights, was adopted as the CNN backbone because of its demonstrated effectiveness in urban scene understanding (Younesi et al., 2024; Zhou et al., 2017), favorable balance between representational capacity and computational efficiency, and widespread adoption in transfer learning applications (Goodfellow et al., 2016). The CNN produces a visual embedding vector $z_j \in \mathbb{R}^{2048}$ for each neighborhood or street segment j.

3.4.2. Multi-source feature integration and vector construction

For each spatial unit j, the corresponding feature vectors $\mathbf{x}_{g,j} \in \mathbb{R}^{d_g}$ and $\mathbf{x}_{h,j} \in \mathbb{R}^{d_h}$ are constructed by assembling the standardized indicator values from Table 2 within each block. The complete structured feature vector is then formed by concatenating the two blocks:

$$\mathbf{x}_{v,j} = \begin{bmatrix} \mathbf{x}_{g,j} \\ \mathbf{x}_{h,j} \end{bmatrix} \tag{4}$$

The structured vector $\mathbf{x}_{v,j}$ is concatenated with the visual embedding $\mathbf{z}_j$ extracted by the CNN backbone and passed to a lightweight multilayer perceptron regressor to predict $\widehat{NVI}_j$. This hybrid input structure preserves the analytical distinction between universal vitality drivers and heritage-specific attributes while enabling joint learning. Figure 2 illustrates the overall heritage-sensitive transfer learning pipeline.

3.5. Model training, optimization, and validation protocol

This section describes the model training and evaluation procedures adopted to ensure robust and generalizable cross-regional prediction of the NVI. It specifies the optimization objective and regularization strategy used during model training, followed by procedures for redundancy control among indicators and the cross-regional validation protocol designed to prevent spatial leakage and assess model transferability.

3.5.1. Training objective

For neighborhood j, model parameters θ are optimized by minimizing the regularized MSE (Hastie et al., 2009):

$$L(\theta)=\frac{1}{m}\sum_{j=1}^{m}(NVI_j-\widehat{NVI}_j)^2+\lambda\|\theta\|_2^2 \quad (5)$$

Optimization is performed using the adaptive moment estimation optimizer (Kingma & Ba, 2014) with early stopping to mitigate overfitting (Prechelt, 1998). Hyperparameters are tuned using validation folds.

3.5.2. Redundancy control

Pearson correlation filtering is applied using a threshold of $|r|>0.80$, following established guidelines for multicollinearity reduction in predictive modeling (Dormann et al., 2013). For highly correlated indicator pairs, the indicator with higher predictive relevance or lower missingness is retained, following principles of feature selection (Guyon & Elisseeff, 2003).

3.5.3. Cross-regional validation

Model performance is evaluated using a leave-one-city-out or region-grouped cross-validation strategy to prevent geographic leakage between training and testing sets (Brenning, 2012; Roberts et al., 2017). This approach is crucial for avoiding biased performance estimates in spatial prediction tasks, as geographic autocorrelation between training and testing sets can lead to substantial overestimation of model accuracy (Kaufman et al., 2012; Ploton et al., 2020). Mean performance metrics and standard deviations across folds are reported. Where data availability permits, an independent hold-out city is further used to assess cross-regional transferability.

3.6. Summary of the methodological contribution

By integrating a heritage-sensitive indicator system with a transfer learning–based modeling strategy, the proposed framework extends transfer learning from a purely computational technique to a context-aware analytical tool. The methodology enables urban vitality assessment that is both cross-regionally comparable and locally meaningful, providing a robust empirical foundation for evidence-based heritage conservation and urban regeneration decision-making.

4. Empirical validation: Cross-regional performance and heritage sensitivity

This section empirically evaluates the proposed HSTL framework using the multi-source dataset compiled from Beijing (source domain), Shanghai, and Guangzhou (target domains). The evaluation is designed to answer three interconnected questions: (i) Does HSTL improve predictive performance in data-scarce target domains relative to non-transfer baselines? (ii) Does the framework achieve cross-domain representation alignment while preserving heritage-specific signals? (iii) Are heritage-sensitive indicators functionally non-redundant, and does their structured preservation mechanism matter? To address these questions, we first describe data harmonization and preprocessing (**S**ection 4.1), then visualize representation alignment via *t*-distributed stochastic neighbor embedding (*t*-SNE) (Section 4.2), report predictive performance and ablation tests (Section 4.3), and finally synthesize the empirical insights in relation to the theoretical framework (Section 4.4).

4.1. Data collection, harmonization, and preprocessing

To empirically evaluate cross-regional vitality assessment in historic districts, we compiled a multi-source dataset from Beijing (source domain) and two target domains, Shanghai and Guangzhou. These cities exhibit distinct built-form typologies and conservation contexts—Beijing's orthogonal hutong fabric, Shanghai's Shikumen–Lilong pattern, and Guangzhou's arcade-street morphology—providing a rigorous setting to test domain shift in both generic urban dynamics and heritage-sensitive characteristics (Huang et al., 2022). The evaluation focuses on whether the proposed HSTL framework improves predictive performance in data-scarce target domains while retaining the functional contribution of heritage-sensitive inputs.

4.1.1. Data sources and feature harmonization

Data were assembled following the dual-structure indicator system theoretically defined in Section 2.1 and operationalized in Section 3.2, comprising generic vitality indicators (X_g) and heritage-sensitive indicators (X_h) (Chen et al., 2024). Structured indicators from heterogeneous sources were harmonized into a shared feature space prior to modeling. This harmonization involved spatial joining to consistent street-segment units, temporal aggregation to comparable time windows (e.g., weekday averages), and unit conversion to standardized metrics (e.g., density per 100 meters). To prevent numerical dominance of high-variance generic indicators, group-wise *z*-standardization was applied within X_g and X_h. Importantly, standardization parameters were estimated on the training split and subsequently applied to validation and test splits to avoid information leakage. SVI was sampled at 50-m intervals along street center lines in each district, yielding approximately 2,000 images per city. These were processed separately and used exclusively for

Table 2. Classification of generic (X_g) and heritage-sensitive (X_h) indicators for the neighborhood vitality index construction

Indicator category	Feature type	Representative variables	Primary data sources	Operationalization strategy	Functional contribution to the NVI	Theoretical rationale
Pedestrian activity	Generic (X_g)	Pedestrian density; movement frequency; temporal activity intensity	Vision-based pedestrian counts; observational datasets	Average daily pedestrian flow per street segment (normalized)	Primary vitality intensity component	Captures real-time human presence and spatial use patterns, widely recognized as a core proxy for urban vitality (Jacobs, 1962)
Building function/land use	Generic (X_g)	Functional mix; commercial frontage ratio; service density	Points of interest (POI) databases (e.g., Baidu Map, Amap); geographic information system (GIS) land-use datasets	Entropy-based functional diversity index; frontage proportion	Functional support component	Represents land-use diversity and service provision that sustain continuous urban activity (Jacobs, 1962)
Social media activity	Generic (X_g)	Geotagged post counts, engagement frequency, and check-in density	Geotagged social media platforms (e.g., Weibo [geotagged posts]; Dianping [location-based reviews])	Log-transformed check-in density per street segment	Interaction amplification component	Reflects public engagement and symbolic use of space, complementing physical activity indicators (Noulas et al., 2012)
Environmental context	Generic (X_g)	Street greenery (NDVI/GVI); enclosure ratio; sidewalk width	Remote sensing imagery; street view image (SVI); GIS datasets	Green pixel ratio; enclosure (D/H) ratio; street width metrics	Environmental modulation component	Environmental comfort and spatial enclosure influence pedestrian experience and indirectly affect vitality intensity (Ewing & Clemente, 2013)
Conservation zoning	Heritage-sensitive (X_h)	Heritage protection level, buffer presence, and regulatory designation	Urban planning registry; GIS heritage protection maps	Categorical protection-level scoring (e.g., grades 1–3)	Heritage constraint component	Encodes institutional heritage value and planning regulations shaping development intensity (UNESCO, 2011)
Historic morphology	Heritage-sensitive (X_h)	Plot pattern continuity; street-network persistence; building-age concentration	Historical GIS maps; cadastral records; SVI datasets	Morphological continuity index (e.g., persistence ratio)	Structural continuity component	Captures spatial structures inherited from historic development trajectories (UNESCO, 2011)
Architectural character	Heritage-sensitive (X_h)	Façade continuity; traditional material prevalence; chromatic consistency	SVI-based façade segmentation; material surveys	Proportion of consistent façade elements; façade color variance metrics	Symbolic identity component	Represents architectural coherence and cultural symbolism, defining historic district identity (Naheed & Shooshtarian, 2022)
Cultural and symbolic context	Heritage-sensitive (X_h)	Proximity to heritage landmarks, cultural nodes, and historically significant land uses	GIS heritage registries; cultural POI datasets	Inverse distance to heritage landmarks; node density metrics	Cultural salience component	Represents intangible cultural meanings and symbolic centrality that attract place-based activity (UNESCO, 2011)

Notes: This indicator framework operationalizes heritage sensitivity by explicitly distinguishing universal urban vitality drivers (X_g) from culturally significant attributes (X_h). The structured separation ensures that heritage-related signals remain explicitly represented during cross-regional transfer learning, preventing them from being numerically overshadowed by high-variance generic indicators. All indicators are standardized using group-wise z-score normalization prior to model training. The framework integrates insights from classic urban vitality theory (Jacobs, 1962) and the United Nations Educational, Scientific, and Cultural Organization historic urban landscape recommendation (UNESCO, 2011), with specific operational metrics adapted from Ewing et al. (2013) and Naheed and Shooshtarian (2022)

https://doi.org/10.36922/JCAU025520101

Table 3. Empirically derived indicator weights for the composite neighborhood vitality index (NVI)–Beijing source domain

Indicator block	Category	Specific indicator	Indicator weight	Category weight	Block weight
Generic vitality (X_g)	Pedestrian activity	Presence density	0.15	0.25	0.65
		Movement frequency	0.10		
	Building function	Functional mix	0.12	0.20	
		Service density	0.08		
	Social media activity	Check-in density	0.10	0.15	
		Post frequency	0.05		
	Environmental context	Street greenery (GVI)	0.03	0.05	
		Street enclosure (D/H ratio)	0.02		
Heritage-sensitive vitality (X_h)	Conservation zoning	Protection level (grade 1–3)	0.06	0.10	0.35
		Buffer zone presence	0.04		
	Historic morphology	Plot pattern continuity	0.05	0.10	
		Street-network persistence	0.05		
	Architectural character	Façade continuity	0.06	0.10	
		Chromatic consistency	0.04		
	Cultural/symbolic context	Landmark proximity	0.03	0.05	
		Cultural node density	0.02		

Notes: The composite NVI is defined as Equation (2)

visual feature extraction within the CNN backbone (Xu et al., 2024), a methodological approach that aligns with emerging workflows for constructing multimodal datasets in heritage studies (Bai et al., 2022).

The resulting dataset integrates structured indicators and visual imagery across all three cities, with consistent feature definitions and spatial units to enable cross-regional comparison.

Table 4 summarizes the data sources, representative features, spatial units, and preprocessing procedures used in this study.

4.1.2. Sample construction and domain split

The street segment was adopted as the primary spatial unit across all cities, a common and effective scale for analyzing urban form and pedestrian-oriented vitality (Harvey &

Aultman-Hall, 2016). Segments were delineated using a consistent road-network-based rule (e.g., segments defined as road links between intersections, with a minimum length of 50 meters) to ensure spatial comparability across domains, and all indicators were aggregated to this unit.

In total, approximately 5,000 spatial samples were constructed. As reported in Table 5, the dataset follows a 60/20/20 domain split: 60 percent from Beijing for representation learning in the source domain, and 20 percent each from Shanghai and Guangzhou to support target-domain fine-tuning and evaluation under data scarcity, a fundamental paradigm in transfer learning research (Pan & Yang, 2009). For fair comparison, identical target-domain splits were used for both the HSTL framework and the target-only baseline.

4.1.3. Preprocessing workflow and distributional diagnostics

The preprocessing workflow, summarized in Table 4, was conducted in three sequential stages:

(i) Data cleaning. Outliers in numeric indicators were handled using the interquartile range criterion, and missing values were imputed using the mean within the training split.
(ii) Standardization. Structured indicators were z-standardized group-wise within X_g and X_h using parameters learned from training data only.
(iii) Image augmentation. To enhance robustness and generalization of visual representations, SVI inputs were augmented using conservative transformations (random rotation within ± 15° and mild noise injection [Gaussian noise, σ = 0.01]), applied exclusively to image data during training, following common practice in deep learning for computer vision (Shorten & Khoshgoftaar, 2019).

Figure 3 provides a descriptive diagnostic of normalization for a representative generic indicator (pedestrian density).

4.2. Cross-regional representation alignment: visualization diagnostics

To qualitatively examine representation shifts across domains under the domain-adaptation framework, learned feature embeddings were visualized using *t*-SNE (Van der Maaten & Hinton, 2008). As shown in Figure 4, feature embeddings from different cities exhibit clear separation prior to adaptation (mean inter-domain centroid distance = 12.3), reflecting differences in urban morphology and heritage context.

After applying HSTL adaptation, embeddings display increased overlap and reduced inter-domain dispersion, with the mean centroid distance reduced from 12.3 to 4.7 in the *t*-SNE projected space, a 61.8 percent reduction—interpreted as a relative indicator rather than an absolute measure (Van der Maaten & Hinton, 2008), suggesting improved alignment of representations across domains.

4.3. Predictive performance and ablation tests

To evaluate the empirical effectiveness of the proposed HSTL framework, predictive performance and component-level contributions were systematically examined. Model performance was first compared against a target-only baseline across source and target domains using standard regression metrics. Subsequently, an ablation study was conducted to assess the functional contribution of heritage-sensitive indicators and the structural preservation mechanism introduced in Section 3.4.

4.3.1. Model comparison and evaluation metrics

Predictive performance was evaluated using MSE and adjusted R^2 as complementary regression metrics. Two configurations were compared:

- Baseline (target-only CNN): Trained solely on target-domain data without cross-regional transfer.
- HSTL model (ours): Trained in the source domain and adapted to target domains via fine-tuning with heritage-sensitive integration.

Uncertainty was quantified using bootstrap resampling (1,000 replicates) to compute 95 percent confidence intervals (Efron, 1992). Domain-wise results are reported in Table 6.

The relatively narrow confidence intervals (e.g., Shanghai HSTL: MSE [0.19, 0.23]) indicate stable estimation across bootstrap replicates, supporting the reliability of the observed performance gains.

Across both target domains, HSTL yields consistent improvements relative to the baseline. Gains are substantially larger in target domains (e.g., Shanghai $\Delta R^2 = +0.15$) than in the source domain, consistent with the intended role of transfer learning in addressing data scarcity.

4.3.2. Ablation study: Functional contribution of heritage features

To directly assess the functional role of heritage-sensitive inputs (X_h) and their preservation mechanism, an ablation study compared three configurations:

(i) Full HSTL framework: $X_g + X_h$ with heritage-preserving integration.
(ii) Generic-only model: X_g only, excluding X_h.

(iii) Homogenized model: $X_g + X_h$ pooled into a single feature block and jointly standardized prior to fusion, without explicit heritage preservation.

Results are summarized in Table 7.

These ablation results carry two important implications. First, the substantial performance drop in target domains when X_h is removed confirms that heritage attributes are not merely descriptive context but functionally non-redundant predictors of vitality—empirically substantiating the theoretical distinction established in Section 2.1. Second, the inferior performance of the homogenized model relative to full HSTL demonstrates that the manner of inclusion matters: heritage signals must be structurally preserved (via group-wise standardization and explicit retention) rather than simply pooled with generic indicators, validating the heritage-preserving mechanism designed in Section 3.4.

4.4. Interpretation of findings

The empirical results presented in Sections 4.1–4.3 yield three key insights that connect the proposed HSTL framework to the theoretical foundations outlined in Section 2 and the methodological design described in Section 3.

First, the data–heritage paradox is empirically resolvable through structured transfer learning. With approximately 1,000 samples per target city, HSTL achieves substantial MSE reductions (25–30 percent) and R^2 improvements (12–15 percent) relative to the target-only baseline (Table 6). These consistent performance gains across both target domains demonstrate that knowledge transferred from a data-rich source (Beijing) can effectively compensate for data scarcity in Shanghai and Guangzhou, provided that heritage-sensitive signals are explicitly preserved during adaptation. The relatively narrow confidence intervals (e.g., Shanghai HSTL: MSE [0.19, 0.23]) indicate stable estimation across bootstrap replicates, supporting the reliability of the observed gains.

Second, representation alignment is achieved without homogenizing heritage signals. The t-SNE visualizations (Figure 4) show that feature distributions across domains become more aligned after adaptation (inter-domain distance reduced by 61.8 percent)—yet the explicit retention of X_h as a separate input block ensures that local conservation signals are not erased in pursuit of cross-regional comparability. This balance directly operationalizes the UNESCO HUL principle of integrating heritage values with contemporary urban dynamics while preserving cultural specificity (UNESCO, 2011).

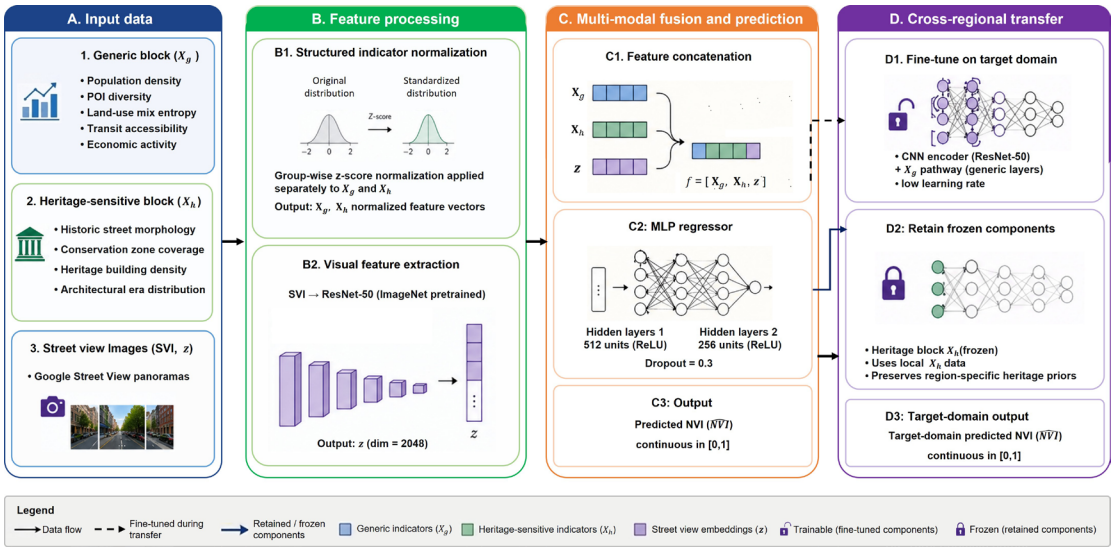


Figure 2. Architecture of the heritage-sensitive transfer learning pipeline for cross-regional prediction of the neighborhood vitality index (NVI). Note: Multi-source structured indicators are partitioned into generic (X_g) and heritage-sensitive (X_h) indicator blocks and standardized separately using group-wise z-score normalization. Street view images (SVI) are processed using a residual network with 50 layers (ResNet-50) backbone pretrained on ImageNet to extract visual embeddings $\boldsymbol{z}$. The three feature components are concatenated and passed to a multilayer perceptron regressor to predict NVI. During cross-regional transfer, parameters associated with domain-invariant components (the convolutional neural network [CNN] encoder and X_g-related layers) are fine-tuned on target-domain data, while X_h is retained as an explicit input block to preserve region-specific heritage signals
Abbreviations: POI: Points of interest; ReLU: Rectified linear unit. Source: Diagram by the authors

Third, heritage attributes are functionally non-redundant predictors of vitality, and their structured preservation is essential. The ablation results (Table 7) carry two important implications. Removing X_h (generic-only) reduces target-domain R^2 by 0.08 (10–11 percent relative drop), confirming that heritage indicators are predictive drivers rather than descriptive add-ons. Critically, the homogenized model—which includes X_h but without structured preservation—underperforms the full HSTL configuration in target domains (e.g., Shanghai: 0.72 versus 0.77). This demonstrates that the manner of inclusion matters: heritage signals must be structurally preserved (via group-wise standardization and explicit retention) rather than simply pooled with generic indicators, empirically validating the heritage-preserving mechanism designed in Section 3.4 and the dual-natured vitality concept established in Section 2.1.

Together, these findings position the HSTL framework as both a predictive tool and a diagnostic lens: it not only improves accuracy in data-scarce historic districts but also reveals which heritage attributes most strongly shape vitality patterns—information that can guide context-sensitive planning and conservation prioritization. By demonstrating that heritage sensitivity enhances rather than impedes cross-regional transferability, the study provides evidence-based support for integrating cultural heritage considerations into data-driven urban analytics.

5. Conclusion and future work

This study addressed the data–heritage paradox in historic district vitality assessment by developing modeling frameworks that are both cross-regionally transferable and culturally sensitive. In response, we proposed and validated an HSTL framework that integrates multi-source urban data with explicit preservation of heritage-specific attributes. This chapter synthesizes the key findings and contributions (Section 5.1), acknowledges limitations that contextualize the scope of inference (Section 5.2), and outlines productive directions for subsequent research (Section 5.3).

5.1. Key findings and contributions

This study addressed a critical methodological gap at the intersection of heritage conservation and urban analytics: how to assess vitality in historic districts across regions in a way that is both technically robust and culturally sensitive. In response to the data–heritage paradox articulated in Section 2.2, we developed and validated an HSTL framework that operationalizes the UNESCO HUL principle (UNESCO, 2011) by integrating heritage-related context with contemporary urban functions in a unified, reproducible modeling pipeline. The study makes three interconnected contributions:

(a) Conceptual contribution: formalizing a dual-structure taxonomy of historic-district vitality. We distinguished generic vitality drivers (X_g) from heritage-sensitive attributes (X_h)—a distinction that extends conventional "generic-only" indicator systems by explicitly incorporating conservation-relevant dimensions such as façade chromatics, historic morphology, and cultural symbolism (Table 2). This taxonomy provides a transparent template for heritage-aware urban analytics and is empirically validated by ablation results (Table 7): removing X_h reduces target-domain R^2 by 10–11 percent, confirming that heritage attributes are functionally non-redundant predictors rather than merely descriptive context.

(b) Methodological contribution: a heritage-preserving transfer architecture. We demonstrated that transfer learning can be adapted to heritage contexts through block-structured feature integration and group-wise standardization (Section 3.4). This design ensures that heritage signals are numerically preserved during domain adaptation—a critical advance over standard "heritage-agnostic" approaches that minimize distributional discrepancy without regard to cultural specificity. The architecture's effectiveness is evidenced by two findings: (i) substantial performance gains in data-scarce target domains (MSE reduction 25–30 percent, R^2 improvement 12–15 percent; Table 6), confirming that knowledge transfer can resolve the data–heritage paradox; and (ii) the inferior performance of the homogenized model in ablation tests (Table 7), which demonstrates that how heritage features are included matters—structural preservation, not mere inclusion, is essential for robust cross-regional generalization.

(c) Translational contribution: operationalizing UNESCO HUL within a computational framework. By embedding heritage indicators as explicit structural components of both the target index (NVI) and the modeling pipeline, the proposed approach operationalizes the HUL principle, bridging conservation theory and data-driven planning practice. The *t*-SNE visualizations (Figure 4) provide qualitative evidence that cross-regional comparability can be improved (inter-domain distance reduced by 61.8 percent) without erasing local distinctiveness—a balance that directly implements HUL's emphasis on integrating heritage values with contemporary urban dynamics while preserving cultural specificity.

Table 4. Summary of multi-source data inputs, spatial units, representative features, and preprocessing procedures used for cross-regional evaluation

Data type	Source	Representative features	Spatial unit	Preprocessing	Sample size	Domain
Street view images	Google Street View API; Baidu Maps API	Visual features (RGB); façade continuity; greenery (GVI)	50 m intervals along street centerlines	Resize to 224 × 224; pixel normalization; data augmentation (rotation, flip)	~2,000 per city	All domains
Pedestrian flow	Mobile operator data; vision-based counters	Pedestrian density; diurnal/nocturnal intensity	Street segment	Outlier removal; log transformation; *z*-score normalization	~1,000 (total)	All domains
POI/land use	Baidu Maps; Amap; GIS databases	Functional mix (entropy); service density; commercial frontage	Street segment (100 m buffer)	Geocoding; category mapping; density calculation; *z*-score normalization	~1,500 (total)	All domains
Social media	Weibo; Dianping (geotagged)	Check-in density; post frequency; engagement intensity	Street segment	Text filtering; deduplication; log transformation; *z*-score normalization	~1,200 (total)	All domains
Environmental context	Remote sensing (Landsat); GIS	NDVI; street enclosure (D/H ratio); sidewalk width	Street segment	Raster sampling; *z*-score normalization	~500 (total)	All domains
Heritage attributes	Urban planning registry; GIS heritage layers	Conservation zoning; protection level; proximity to landmarks	Street segment	Buffer analysis; categorical encoding; inverse distance calculation	N/A (vector data)	All domains

Notes: Group-wise *z*-score normalization was applied separately to generic (X_g) and heritage-sensitive (X_h) indicators, with normalization parameters estimated on the training split only to prevent data leakage. Sample sizes refer to the number of spatial units (street segments or grid cells) after preprocessing. Domain allocation (source versus target) is detailed in Table 5

Abbreviations: API: Application programming interface; GIS: Geographic information system; GVI: Green view index; NDVI: Normalized difference vegetation index; POI: Points of interest; RGB: Red, green, blue

Table 5. Domain-wise sample counts and allocation for source training and target-domain adaptation

Domain	City	Sample count	Percentage (%)
Source	Beijing	2,847	60
Target 1	Shanghai	1,023	20
Target 2	Guangzhou	1,034	20
Total		4,904	100

Notes: All samples are constructed at the street-segment level within historic districts across all three cities. The 60/20/20 split ensures sufficient source-domain representation learning while simulating realistic data scarcity in target domains. Identical target-domain splits are used across all model configurations, heritage-sensitive transfer learning, and baseline to ensure fair comparison

Together, these contributions position the HSTL framework as both a predictive tool and a diagnostic lens. For planners and heritage managers, the NVI enables differentiated diagnosis: distinguishing districts experiencing functional decline (low X_g) from those undergoing heritage character erosion (low X_h). This distinction supports more targeted interventions—e.g., differentiating revitalization of accessibility and services from conservation of architectural character and cultural salience—providing an evidence-based foundation for context-sensitive heritage management.

5.2. Limitations and boundary conditions

Several limitations should be acknowledged to contextualize the scope of inference and guide refinement:

(a) Geographic scope and institutional regimes. The framework is validated within Chinese historic-district contexts. Its generalizability to international settings with different regulatory systems and heritage governance models remains to be tested. Future work should apply the HSTL framework to diverse heritage regimes (e.g., European and Asian UNESCO sites) to assess robustness under varied cultural and institutional conditions.

(b) Index construction choices. The NVI is a constructed target. Although the weighting protocol is transparent and empirically grounded (Section 3.3), alternative weighting strategies—such as expert elicitation, participatory weighting, or multi-criteria decision approaches—may yield different vitality profiles. Examining the sensitivity of findings to these alternative constructions represents an important direction for future research.

(c) Platform and visibility biases. Reliance on third-party data sources (e.g., social media traces and street view imagery) may introduce representation bias related to digital visibility, platform usage, or demographic coverage (Boyd & Crawford, 2012; Tufekci, 2014). These biases could systematically underrepresent populations or activities with lower digital engagement, potentially affecting the generalizability of vitality assessments.

(d) Static assessment. The current implementation is primarily cross-sectional and does not explicitly model temporal dynamics such as seasonality, event-driven fluctuations, or long-term trends. This limits its suitability for real-time monitoring and forecasting applications. Extending the framework to incorporate longitudinal indicators would enable analysis of vitality trajectories and intervention effects over time.

These limitations, while contextualizing the scope of current inferences, also delineate productive directions for subsequent research, which are outlined in Section 5.3.

5.3. Future research directions

Building on the limitations identified above, future research can extend this work along four complementary trajectories:

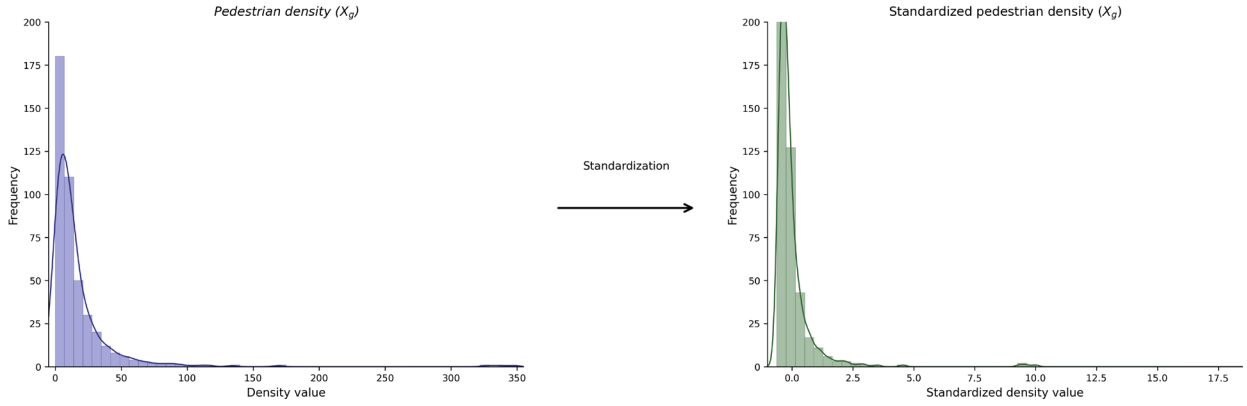


Figure 3. Distribution of a representative generic vitality indicator (pedestrian density, X_g) before and after preprocessing. Note: Histograms illustrate pedestrian density prior to preprocessing (left; skewness = 2.34, kurtosis = 5.67) and after preprocessing (right; skewness = 0.12, kurtosis = 2.89). The preprocessing pipeline substantially reduces distributional skewness and improves scale comparability for downstream modeling. This figure is presented as a descriptive diagnostic rather than inferential evidence. Pedestrian density serves as a primary intensity component within the generic vitality indicator block X_g (see Table 2). Source: Graphs by the authors

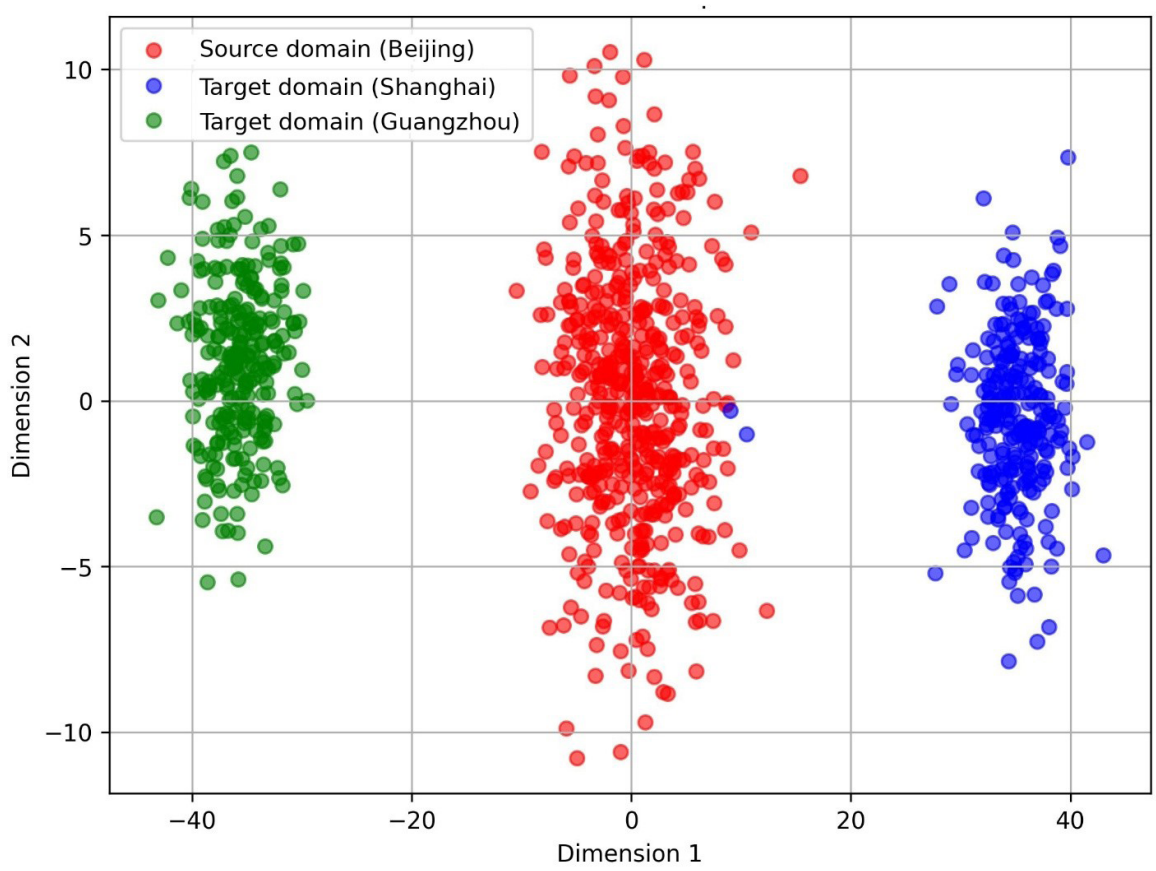


Figure 4. *t*-distributed stochastic neighbor embedding (*t*-SNE) visualization of learned feature representations across source and target domains under the heritage-sensitive transfer learning framework. Source: Plot by the authors

(a) Global validation across heritage regimes. The current framework is validated within Chinese historic-district contexts. To address the geographic scope limitation and test generalizability under differing conservation policies, cultural practices, and built-form typologies, future work should apply the HSTL framework to diverse international heritage settings, including European and Asian UNESCO sites. Such cross-cultural validation would assess the framework's robustness under varied regulatory regimes and heritage governance models.

(b) Enhanced domain adaptation mechanisms. While the current fine-tuning approach effectively mitigates domain shift, more advanced adaptation modules—such as adversarial domain adaptation (Ganin et al., 2017) or discrepancy-minimization objectives (e.g., maximum mean discrepancy (Gretton et al., 2012)—could further improve robustness under stronger source–target distribution shifts. These techniques may also partially address platform and visibility biases by learning more invariant representations across heterogeneous data sources.

(c) Spatiotemporal modeling of vitality dynamics. The current implementation is primarily cross-sectional, limiting its ability to capture temporal dynamics such as seasonality, event-driven fluctuations, or long-term trends. Extending the framework to incorporate longitudinal indicators and time-series modeling would enable analysis of vitality trajectories and intervention effects, supporting real-time monitoring and forecasting applications in heritage management.

(d) Participatory and value-sensitive artificial intelligence. To address both index construction choices and representation biases, future work should integrate principles of value-sensitive design (Friedman et al., 2002) by incorporating stakeholder and community inputs into the indicator definition and weighting. Participatory approaches such as expert elicitation, community workshops, or deliberative weighting could ensure that vitality assessments

Table 6. Predictive performance of the baseline and heritage-sensitive transfer learning framework across source and target domains

Model	Domain	MSE (95% CI)	Adjusted R^2 (95% CI)	Improvement
Baseline CNN	Beijing (source)	0.12 (0.11, 0.13)	0.85 (0.83, 0.87)	–
Transfer learning	Beijing (source)	0.10 (0.09, 0.11)	0.88 (0.86, 0.90)	ΔR^2 = + 0.03
Baseline CNN	Shanghai (target)	0.28 (0.25, 0.31)	0.62 (0.59, 0.65)	–
Transfer learning	Shanghai (target)	0.21 (0.19, 0.23)*	0.77 (0.74, 0.80)*	ΔR^2= + 0.15*
Baseline CNN	Guangzhou (target)	0.31 (0.28, 0.34)	0.58 (0.55, 0.61)	–
Transfer learning	Guangzhou (target)	0.23 (0.21, 0.25)*	0.73 (0.70, 0.76)*	ΔR^2 = + 0.15*

Notes: Performance is reported using MSE and adjusted R^2; values in brackets denote 95 percent bootstrap CIs (1,000 resamples). Improvements (ΔR^2) are reported relative to the baseline within each domain. Target-domain gains (Shanghai: +0.15; Guangzhou: +0.15) are substantially larger than source-domain gains (+0.03), indicating effective cross-regional transfer. An asterisk (*) indicates that the transfer-learning model shows a statistically supported improvement over the corresponding baseline model within the target domain, based on bootstrap estimation and the reported 95% CIs. Abbreviations: CI: Confidence interval; CNN: Convolutional neural network; MSE: Mean squared error

Table 7. Ablation results evaluating the functional contribution of heritage-sensitive indicators (X_h) and preservation mechanisms

Model configuration	Beijing (source) adjusted R^2	Shanghai (target) adjusted R^2	Guangzhou (target) adjusted R^2
Full model ($X_g + X_h$, heritage-preserving)	0.88	0.77	0.73
Generic-only (X_g only)	0.85	0.69	0.65
Homogenized ($X_g + X_h$, no preservation)	0.86	0.72	0.68
Performance drop (full versus generic-only)	−0.03	−0.08	−0.08
Relative drop (%)	3.4	10.4	11.0

Notes: Removing X_h (generic-only) reduces adjusted R^2 by 0.03 in the source domain and by 0.08 in both target domains—equivalent to relative drops of 10.4 percent (Shanghai) and 11.0 percent (Guangzhou). The homogenized model, which includes X_h but without structured preservation, underperforms the full heritage-sensitive transfer learning configuration in target domains (e.g., Shanghai: 0.72 versus 0.77), indicating that heritage-sensitive attributes must be structurally preserved—not merely included—to maximize cross-regional transferability. These findings empirically substantiate the theoretical distinction between X_g and X_h established in Section 2.1 and validate the heritage-preserving mechanism designed in Section 3.4

 https://doi.org/10.36922/JCAU025520101

reflect lived experience and place-based values rather than solely algorithmic proxies.

Ultimately, this study supports an emerging paradigm in which heritage conservation and urban vitality are treated as integrated dimensions rather than competing objectives. The findings suggest that computational methods, when designed with explicit heritage sensitivity, can enhance cross-regional comparability while retaining culturally meaningful signals—providing a reproducible basis for evidence-informed conservation and regeneration decisions.

Acknowledgments

The authors gratefully acknowledge the Department of Landscape Architecture, Faculty of Design and Architecture, Universiti Putra Malaysia, for their academic and infrastructural support to this research.

Funding

This research was funded by the Universiti Putra Malaysia Grant (Geran Putra IPS; grant number: GP-IPS-9772600).

Conflict of interest

The authors declare that they have no competing interests.

Author contributions

Conceptualization: Jijiang Zhang, Faziawati Abdul Aziz
Formal analysis: Jijiang Zhang
Funding acquisition: Faziawati Abdul Aziz
Investigation: Jijiang Zhang, Mohd Fabian Hasna
Methodology: All authors
Supervision: Faziawati Abdul Aziz
Visualization: Jijiang Zhang
Writing–original draft: Jijiang Zhang
Writing–review & editing: All authors

Ethics approval and consent to participate

Obtaining ethics approval is not applicable. This study did not involve any human participants, animals, or personally identifiable data.

Consent for publication

Not applicable. This study did not involve human subjects, identifiable personal data, or patient images.

Availability of data

The datasets used in this study include publicly available sources (e.g., Google Street View, OpenStreetMap, official statistical data) and project-generated composite indicators. While raw data are not publicly shared due to third-party license restrictions, processed data and methodology details can be made available upon reasonable request from the corresponding author.

References

Bai, N., Nourian, P., Luo, R., & Pereira Roders, A. (2022). Heri-graphs: A dataset creation framework for multimodal machine learning on graphs of heritage values and attributes with social media. *ISPRS International Journal of Geo-Information*, *11*(9), 469.

https://doi.org/10.3390/ijgi11090469

Bandarin, F., & Oers, R. van. (2012). *The historic urban landscape: Managing heritage in an urban century*. Hoboken, NJ: Wiley.

https://doi.org/10.1002/9781119968115

Boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society, 15*(5), 662–679.

https://doi.org/10.1080/1369118X.2012.678878

Brenning, A. (2012). Spatial cross-validation and bootstrap for the assessment of prediction rules in remote sensing: The R package sperrorest. In: *2012 IEEE International Geoscience and Remote Sensing Symposium* (pp. 5372–5375). IEEE.

https://doi.org/10.1109/IGARSS.2012.6352393

Calabrese, F., Diao, M., Di Lorenzo, G., Ferreira, J., & Ratti, C. (2013). Understanding individual mobility patterns from urban sensing data: A mobile phone trace example. *Transportation Research Part C: Emerging Technologies, 26*, 301–313.

https://doi.org/10.1016/j.trc.2012.09.009

Chen, J., Zhao, X., Wang, H., Yan, J., Yang, D., & Xie, K. (2024). Portraying heritage corridor dynamics and cultivating conservation strategies based on environment spatial model: An integration of multi-source data and image semantic segmentation. *Heritage Science, 12*(1).

https://doi.org/10.1186/s40494-024-01497-7

Conzen, M. R. G. (1960). *Alnwick, Northumberland: A study in town-plan analysis*. Institute of British Geographers.

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J. R. G., Gruber, B., Lafourcade, B., Leitão, P. J., Münkemüller, T., McClean, C., Osborne, P. E., Reineking, B., Schröder, B., Skidmore, A. K., Zurell, D., & Lautenbach, S. (2013). Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography, 36*(1), 27–46.

https://doi.org/10.1111/j.1600-0587.2012.07348.x

Efron, B. (1992). Bootstrap methods: Another look at the

jackknife. In: S. Kotz & N. L. Johnson (Eds.), *Breakthroughs in statistics* (pp. 569–593). New York, NY: Springer.

https://doi.org/10.1007/978-1-4612-4380-9_41

Ewing, R., Clemente, O., Neckerman, K. M., Purciel-Hill, M., Quinn, J. W., & Rundle, A. (2013). *Measuring urban design*. Washington, D.C.: Island Press/Center for Resource Economics.

https://doi.org/10.5822/978-1-61091-209-9

Fang, C., Zhou, L., Gu, X., Liu, X., & Werner, M. (2025). A data-driven approach to urban area delineation using multi-source geospatial data. *Scientific Reports, 15*(1), 8708.

https://doi.org/10.1038/s41598-025-93366-x

Fang, K., & Wu, Y. (2025). The importance of integrating historic context and activities into heritage district: Enhancing attraction and vitality. *npj Heritage Science, 13*(1), 434.

https://doi.org/10.1038/s40494-025-01862-0

Fiorucci, M., Khoroshiltseva, M., Pontil, M., Traviglia, A., Del Bue, A., & James, S. (2020). Machine learning for cultural heritage: A survey. *Pattern Recognition Letters, 133*, 102–108.

https://doi.org/10.1016/j.patrec.2020.02.017

Friedman, B., Kahn, P., & Borning, A. (2003). Value sensitive design: Theory and methods. *University of Washington Technical Report, 2*(8), 1–8.

Fu, J.-M., Tang, Y.-F., Zeng, Y.-K., Feng, L.-Y., & Wu, Z.-G. (2025). Sustainable historic districts: Vitality analysis and optimization based on space syntax. *Buildings, 15*(5), 657.

https://doi.org/10.3390/buildings15050657

Ganin, Y., Ustinova, E., Ajakan, H., Germain, P., Larochelle, H., Laviolette, F., Marchand, M., & Lempitsky, V. (2017). Domain-adversarial training of neural networks. In G. Csurka (Ed.), *Domain adaptation in computer vision applications* (pp. 189–209). New York, NY: Springer International Publishing.

https://doi.org/10.1007/978-3-319-58347-1_10

Gao, Z., & Xie, Y. (2025). A parameter-based multi-source transfer learning method for building load forecasting with sparse data scenarios. *Energy Reports, 13*, 4936–4947.

https://doi.org/10.1016/j.egyr.2025.04.050

Gehl, J. (2010). *Cities for people*. Island Press.

Ginzarly, M., Houbart, C., & Teller, J. (2019). The historic urban landscape approach to urban management: A systematic review. *International Journal of Heritage Studies, 25*(10), 999–1019.

https://doi.org/10.1080/13527258.2018.1552615

Ginzarly, M., Pereira Roders, A., & Teller, J. (2019). Mapping historic urban landscape values through social media. *Journal of Cultural Heritage, 36*, 1–11.

https://doi.org/10.1016/j.culher.2018.10.002

Gong, F.-Y., Zeng, Z.-C., Zhang, F., Li, X., Ng, E., & Norford, L. K. (2018). Mapping sky, tree, and building view factors of street canyons in a high-density urban environment. *Building and Environment, 134*, 155–167.

https://doi.org/10.1016/j.buildenv.2018.02.042

Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). *Deep learning*. Cambridge, MA: MIT Press.

Gretton, A., Borgwardt, K. M., Rasch, M., Scholkopf, B., & Smola, A. (2012). A kernel two-sample test. *Journal of Machine Learning Research, 13*, 723–773.

Grilli, E., & Remondino, F. (2019). Classification of 3D digital heritage. *Remote Sensing, 11*(7), 847.

https://doi.org/10.3390/rs11070847

Gu, C. (2024). The basics of urbanization. In C. Gu (Ed.), *China's urbanization* (pp. 1–38). New York, NY: Springer Nature.

https://doi.org/10.1007/978-981-97-3782-6_1

Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research, 3*, 1157–1182.

Harvey, C., & Aultman-Hall, L. (2016). Measuring urban streetscapes for livability: A review of approaches. *The Professional Geographer, 68*(1), 149–158.

https://doi.org/10.1080/00330124.2015.1065546

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning*. New York, NY: Springer.

https://doi.org/10.1007/978-0-387-84858-7

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 770–778). IEEE.

https://doi.org/10.1109/CVPR.2016.90

Hu, L., Liu, Y., & Yu, B. (2025). Evolution method of built environment spatial quality in historic districts based on spatiotemporal street view: A case study of Tianjin Wudadao. *Buildings, 15*(11), 1953.

https://doi.org/10.3390/buildings15111953

Huang, X., Gong, P., Wang, S., White, M., & Zhang, B. (2022). Machine learning modeling of vitality characteristics in historical preservation zones with multi-source data. *Buildings, 12*(11), 1978.

https://doi.org/10.3390/buildings12111978

Iranmanesh, A., & Alpar Atun, R. (2020). Reading the urban socio-spatial network through space syntax and geo-tagged twitter data. *Journal of Urban Design, 25*(6), 738–757.

https://doi.org/10.1080/13574809.2020.1814132

Jacobs, J. (1962). *The death and life of great American cities*. Ontario, Canada: Random House.

Jolliffe, I. T. (2002). *Principal component analysis* (2nd ed.). New York, NY: Springer.

Kalfas, D., Kalogiannidis, S., Ambas, V., & Chatzitheodoridis, F. (2024). Contribution of the cultural and creative industries to regional development and revitalization: A European perspective. *Urban Science, 8*(2), 39.

https://doi.org/10.3390/urbansci8020039

Kaufman, S., Rosset, S., Perlich, C., & Stitelman, O. (2012). Leakage in data mining: Formulation, detection, and avoidance. *ACM Transactions on Knowledge Discovery from Data, 6*(4), 1–21.

https://doi.org/10.1145/2382577.2382579

Kingma, D. P., & Ba, J. (2014). *Adam: A method for stochastic optimization*. *arXiv*. arXiv:1412.6980.

https://doi.org/10.48550/arXiv.1412.6980

Kou, H., Zhou, J., Chen, J., & Zhang, S. (2018). Conservation for sustainable development: The sustainability evaluation of the Xijie historic district, Dujiangyan City, China. *Sustainability, 10*(12), 4645.

https://doi.org/10.3390/su10124645

Kouw, W. M., & Loog, M. (2018). *An introduction to domain adaptation and transfer learning*. *arXiv*. arXiv:1812.11806.

https://doi.org/10.48550/arXiv.1812.11806

Li, G., Wu, Y., Yan, C., Fang, X., Li, T., Gao, J., Xu, C., & Wang, Z. (2024). An improved transfer learning strategy for short-term cross-building energy prediction using data incremental. *Building Simulation, 17*(1), 165–183.

https://doi.org/10.1007/s12273-023-1053-x

Li, H., & Miao, L. (2025). A study of the non-linear relationship between urban morphology and vitality in heritage areas based on multi-source data and machine learning: A case study of Dalian. *ISPRS International Journal of Geo-Information*, *14*(4), 177.

https://doi.org/10.3390/ijgi14040177

Li, X., Kozlowski, M., Salih, S. A., & Ismail, S. B. (2025). Evaluating the vitality of urban public spaces: Perspectives on crowd activity and built environment. *Archnet-IJAR: International Journal of Architectural Research*, *19*(3), 562–583.

https://doi.org/10.1108/ARCH-01-2024-0009

Li, X., Zhang, C., & Li, W. (2015). Does the visibility of greenery increase perceived safety in urban areas? Evidence from the Place Pulse 1.0 dataset. *ISPRS International Journal of Geo-Information*, *4*(3), 1166–1183.

https://doi.org/10.3390/ijgi4031166

Liu, J., Li, Y., Xu, Y., Zhuang, C. C., Hu, Y., & Yu, Y. (2024). Impacts of built environment on urban vitality in cultural districts: A case study of Haikou and Suzhou. *Land, 13*(6), 840.

https://doi.org/10.3390/land13060840

Llamas, J., Lerones, P. M., Medina, R., Zalama, E., & Gómez-García-Bermejo, J. (2017). Classification of architectural heritage images using deep learning techniques. *Applied Sciences, 7*(10), 992.

https://doi.org/10.3390/app7100992

Lynch, K. (1964). *The image of the city*. Cambridge, MA: MIT Press.

Montgomery, J. (1998). Making a city: Urbanity, vitality and urban design. *Journal of Urban Design, 3*(1), 93–116.

https://doi.org/10.1080/13574809808724418

Naheed, S., & Shooshtarian, S. (2022). The role of cultural heritage in promoting urban sustainability: A brief review. *Land, 11*(9), 1508.

https://doi.org/10.3390/land11091508

Niu, H., & Silva, E. A. (2020). Crowdsourced data mining for urban activity: Review of data sources, applications, and methods. *Journal of Urban Planning and Development, 146*(2), 04020007.

https://doi.org/10.1061/(ASCE)UP.1943-5444.0000566

Noulas, A., Scellato, S., Lambiotte, R., Pontil, M., & Mascolo, C. (2012). A tale of many cities: Universal patterns in human urban mobility. *PLoS ONE, 7*(5), e37027.

https://doi.org/10.1371/journal.pone.0037027

OECD. (2018). *Indicators for resilient cities* (OECD Regional Development Working Papers No. 2018/02). OECD Publishing.

https://doi.org/10.1787/6f1f6065-en

Pan, S. J., & Yang, Q. (2009). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering, 22*(10), 1345–1359.

https://doi.org/10.1109/TKDE.2009.191

Pendlebury, J., Short, M., & While, A. (2009). Urban World Heritage Sites and the problem of authenticity. *Cities, 26*(6), 349–358.

https://doi.org/10.1016/j.cities.2009.09.003

Ploton, P., Mortier, F., Réjou-Méchain, M., Barbier, N., Picard, N., Rossi, V., Dormann, C., Cornu, G., Viennois, G., & Bayol, N. (2020). Spatial validation reveals poor predictive performance of large-scale ecological mapping models. *Nature Communications, 11*(1), 4540.

https://doi.org/10.1038/s41467-020-18321-y

Prechelt, L. (1998). Early stopping—But when? In G. B. Orr & K.-R. Müller (Eds.), *Neural networks: Tricks of the trade* (Vol. 1524, pp. 55–69). Berlin Heidelberg: Springer .

https://doi.org/10.1007/3-540-49430-8_3

Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F., & Dormann, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic

structure. *Ecography, 40*(8), 913–929.

https://doi.org/10.1111/ecog.02881

Rodwell, D. (2008). *Conservation and sustainability in historic cities*. Hoboken, NJ: John Wiley & Sons.

Sakamoto, S., Kogure, M., Hanibuchi, T., Nakaya, N., Hozawa, A., & Nakaya, T. (2023). Effects of greenery at different heights in neighbourhood streetscapes on leisure walking: A cross-sectional study using machine learning of streetscape images in Sendai City, Japan. *International Journal of Health Geographics, 22*(1), 29.

https://doi.org/10.1186/s12942-023-00351-6

Shi, P., Xiao, Y., & Zhan, Q. (2020). A study on spatial and temporal aggregation patterns of urban population in Wuhan City based on Baidu Heat Map and POI data. *International Review for Spatial Planning and Sustainable Development, 8*(3), 101–121.

https://doi.org/10.14246/irspsda.8.3_101

Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data, 6*(1), 60.

https://doi.org/10.1186/s40537-019-0197-0

Tufekci, Z. (2014). Big questions for social media big data: Representativeness, validity and other methodological pitfalls. *Proceedings of the International AAAI Conference on Web and Social Media, 8*(1), 505–514. https://ojs.aaai.org/index.php/ICWSM/article/view/14517

UNESCO. (2011). *Recommendation on the historic urban landscape*. In *Records of the General Conference 36th Session*. UNESCO.

Van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research, 9*(11), 2579–2605.

Veldpaus, L., Pereira Roders, A. R., & Colenbrander, B. J. F. (2013). Urban heritage: Putting the past into the future. *The Historic Environment: Policy & Practice, 4*(1), 3–18.

https://doi.org/10.1179/1756750513Z.00000000022

Weiss, K., Khoshgoftaar, T. M., & Wang, D. (2016). A survey of transfer learning. *Journal of Big Data, 3*(1), 9.

https://doi.org/10.1186/s40537-016-0043-6

Whitehand, J. W. R. (2007). Urban morphology. *Urban Morphology, 11*(2), 103–107.

Wu, J., Lu, Y., Gao, H., & Wang, M. (2022). Cultivating historical heritage area vitality using urban morphology approach based on big data and machine learning. *Computers, Environment and Urban Systems, 91*, 101716.

https://doi.org/10.1016/j.compenvurbsys.2021.101716

Xu, J., Dai, Y., Cai, J., Qian, H., Peng, Z., & Zhong, T. (2025). Evaluation of urban street historical appearance integrity based on street view images and transfer learning. *ISPRS International Journal of Geo-Information*, *14*(7), 266.

https://doi.org/10.3390/ijgi14070266

Xu, J., Wang, J., Zuo, X., & Han, X. (2024). Spatial quality optimization analysis of streets in historical urban areas based on street view perception and multisource data. *Journal of Urban Planning and Development, 150*(4), 05024036.

https://doi.org/10.1061/JUPDDM.UPENG-4770

Younesi, A., Ansari, M., Fazli, M., Ejlali, A., Shafique, M., & Henkel, J. (2024). A comprehensive survey of convolutions in deep learning: Applications, challenges, and future trends. *IEEE Access, 12*, 41180–41218.

https://doi.org/10.1109/ACCESS.2024.3376441

Yu, B., Sun, J., Wang, Z., & Jin, S. (2024). Influencing factors of street vitality in historic districts based on multisource data: Evidence from China. *ISPRS International Journal of Geo-Information*, *13*(8), 277.

https://doi.org/10.3390/ijgi13080277

Zhang, J., & Zhao, X. (2024). Using POI and multisource satellite datasets for mainland China›s population spatialization and spatiotemporal changes based on regional heterogeneity. *Science of the Total Environment, 912*, 169499.

https://doi.org/10.1016/j.scitotenv.2023.169499

Zhang, Q., Cheng, T., Xu, P., & Jiang, X. (2025). Balancing heritage conservation and urban vitality through a multi-tiered governance strategy: A case study of Nanjing›s Yihe Road historic district, China. *Land, 14*(9), 1894.

https://doi.org/10.3390/land14091894

Zhang, Y., & Han, Y. (2022). Vitality evaluation of historical and cultural districts based on the values dimension: Districts in Beijing City, China. *Heritage Science, 10*(1), 137.

https://doi.org/10.1186/s40494-022-00776-5

Zheng, G., Ding, L., & Zheng, J. (2025). A multi-dimensional evaluation of street vitality in a historic neighborhood using multi-source geo-data: A case study of Shuitingmen, Quzhou. *ISPRS International Journal of Geo-Information*, *14*(7), 240.

https://doi.org/10.3390/ijgi14070240

Zhou, B., Lapedriza, A., Khosla, A., Oliva, A., & Torralba, A. (2017). Places: A 10 million image database for scene recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 40*(6), 1452–1464.

https://doi.org/10.1109/TPAMI.2017.2723009