

Article

Not peer-reviewed version

Measuring Perceptions of Walkable Streetscapes in Cultural Heritage Contexts

[Hessameddin Maniei](#)^{*}, [Elham Mehrinejad Khotbehsara](#), [Dietwald Gruehn](#)

Posted Date: 5 May 2026

doi: 10.20944/preprints202605.0233.v1

Keywords: streetscape perception; pedestrian streetscapes; green infrastructure; semantic segmentation; historic urban renewal strategies; diverse urban context; street width



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC, OpenAlex.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Measuring Perceptions of Walkable Streetscapes in Cultural Heritage Contexts

Hessameddin Maniei ^{1,*}, Elham Mehrinejad Khotbehsara ² and Dietwald Gruehn ¹

¹ Research Group Landscape Ecology and Landscape Planning (LLP), Department of Spatial Planning, TU Dortmund University, 44227 Dortmund, Germany

² School of Engineering, University of Southern Queensland, Toowoomba, Australia

* Correspondence: author: hessameddin.maniei@tu-dortmund.de

Abstract

This study examines pedestrian perceptions of streetscapes in Isfahan's cultural heritage site by integrating deep learning-based image segmentation with urban morphological analysis. Using a U-Net model applied to First-Person Pedestrian View (FPPV) images, five perceptual indices (imageability, enclosure, human scale, greenness, and walking index) were quantified to assess their influence on pedestrian experience. Street width was explicitly incorporated as a morphological variable to examine its relationship with perceptual qualities using spearman correlation analysis and visual trend analysis using Pearson correlation. The results reveal consistent relationships between visual composition and perceptual outcomes, particularly strong associations between imageability, enclosure, and vegetation structure, as well as trade-offs between enclosure and sky visibility. In contrast, variables such as human scale and walking index show weak or negligible associations with street width, suggesting that pedestrian presence and activity patterns in heritage contexts are more strongly influenced by landscape elements, water features, and spatial continuity than by dimensional factors alone. Findings highlight how urban renewal strategies, such as streetscape enhancement and cultural preservation, shape pedestrian movement and spatial perception. Segmentation-based analysis achieved an accuracy of 83% in classifying dominant streetscape elements, offering a robust alternative to traditional survey-based methods. This study contributes a data-driven framework for assessing pedestrian streetscapes, emphasizing morphological continuity, human-scale design, and green infrastructure as critical determinants of walkability. It also identifies key challenges, including fragmented spatial morphology and inconsistent urban furniture placement, which affect pedestrian comfort and use of space. These findings support evidence-based policy and design strategies for optimizing historic urban streetscapes, with implications for balancing heritage conservation and modern pedestrian needs. Future research may refine perceptual metrics and extend the approach across diverse urban contexts.

Keywords: streetscape perception; pedestrian streetscapes; green infrastructure; semantic segmentation; historic urban renewal strategies; diverse urban context; street width

1. Introduction

The built environment attribute of walkability determines how pedestrianized streets develop into active and inclusive cities that support social interaction. The sustainable method of public space development through walkability has gained increasing acceptance because it creates spaces that enable both transportation and community engagement [1,2]. The design of streetscapes functions as a fundamental element of urban planning because it determines how public areas appear and function as spaces for movement and social interaction [3,4]. The design of pedestrianized streets directly impacts how people experience urban life and their overall quality of life [5–7]. The physical characteristics and sensory aspects of these spaces determine how people walk and interact with

others while making the areas appealing for walking and other activities [8–10] emphasizes the need for streets designed with people in mind, because such human-oriented planning methods create better living conditions and improved built environments. Research has proven that urban design elements directly impact the walkability performance of street environments [8,11,12].

In recent years, the relationship between walkability and urban design has received renewed attention, especially with the integration of artificial intelligence (AI) tools in planning practice and the increasing focus on pedestrian-friendly design [13,14]. Studies have demonstrated that deep learning techniques can effectively predict urban aesthetics [15] and evaluate how greenery influences user satisfaction in historic environments, in contrast to more modern urban contexts [14]. Among these methods, semantic segmentation has emerged as a powerful tool for quantifying physical components of streetscapes, many of which are closely linked to walkability and, by extension, public health outcomes [16].

Urbanisation is accelerating at a pace that places significant strain on both the environmental integrity and visual harmony of our streets. In many cases, traditional urban planning approaches fail to respond adequately to the needs of sustainability and resident satisfaction. Although technologies such as artificial intelligence (AI), deep learning, and the Internet of Things (IoT) have introduced new ways to improve urban management, their application in heritage-rich or culturally sensitive locations has been slow to progress [14]. Recent innovations in computer vision and machine learning, however, show promise in offering scalable and unbiased tools for analysing urban form. These technologies enable a broader and more consistent evaluation of how spatial elements influence everyday experiences in cities. This is something that conventional surveys and manual audits alone struggle to capture.

Deep learning has become an important tool in streetscape analysis, often making use of Google Street View (GSV) images to capture features linked to urban quality and walkability. Semantic segmentation models such as U-Net and DeepLab V3+, commonly trained on datasets like Cityscapes, allow pixel-level classification of street elements and support large-scale analysis of urban form [13]. While these approaches work well in general or vehicle-focused settings, they are still rarely applied to pedestrianised heritage streets. In addition, the combination of segmentation outputs with morphological analysis to assess perceptual factors such as imageability, enclosure, human scale, greenness, and pedestrian density remains underdeveloped. A recent review on walkable environments through the lens of Space Syntax identified a lack of research combining spatial analysis and participatory approaches with technology-based methods, including computer vision and machine learning, to achieve a more comprehensive understanding of pedestrian perception and behaviour, particularly regarding how visual structure and spatial configuration shape walkability within urban environments [17].

Then, this study extends the findings of a previous investigation [3], which examined spatial configuration and walkability in Isfahan's heritage streets using syntactic and perceptual analyses. While the earlier work focused on network accessibility and spatial structure, the present research advances this framework by integrating semantic segmentation to quantify the perceptual and visual dimensions of pedestrian experience. More specifically, the study seeks to answer the following research questions:

1. In what ways can semantic image segmentation be applied to capture and quantify perceptual qualities of pedestrianized streets?
2. How do patterns of urban morphology correspond with pedestrian movement and spatial behavior in heritage areas?
3. Which design strategies can strengthen walkability and cultural identity while improving the overall quality of historic streetscapes?

The outcomes aim to advance the discourse on pedestrian-oriented urban planning and propose a transferable framework for assessing streetscapes in comparable heritage contexts.

2. Research Background and Theoretical Framework

2.1. Streetscape in Historic Space

Urban streets within heritage sites are multifunctional systems that combine historical identity with contemporary urban functions. They are critical components of the built environment [18], facilitating mobility [19], economic exchange [20], and social interaction [21], while also preserving architectural and cultural heritage [22]. The design of these streetscapes directly affects measurable parameters such as pedestrian flow, enclosure ratio, and green infrastructure coverage, which in turn influence walkability [23].

Rapid urbanization introduces structural challenges such as fragmented morphology, inconsistent street furniture placement, and reduced human scale, all of which require data-driven design evaluation. In heritage contexts, it is essential to balance conservation goals with performance-driven design approaches that support accessibility, functional integration, and long-term resilience.

In the historic centre of Isfahan, Iran, streets such as Chahar Bagh and Sepah serve as notable examples of pedestrian-only corridors. Their pedestrianisation has reshaped spatial enclosure and circulation patterns, offering a valuable context for analysing both perceptual and morphological changes. Despite these interventions, the broader accessibility of the area continues to rely heavily on motorised transport, with travel modes comprising approximately 40 percent car use, 25 percent metro, 20 percent bus, and only 15 percent walking [3]. This modal split suggests existing gaps in infrastructure that support active transportation. Previous studies have demonstrated that streets with coherent enclosure, integrated greenery, and thoughtfully placed seating not only enhance user comfort but also improve observed pedestrian density and the overall efficiency of movement.

2.2. Perceptual Qualities and Assessment of Urban Streets

Researchers identified four paradigms for studying perceptions such as aesthetics: the expert paradigm, the psychophysical paradigm, the cognitive paradigm, and the experiential paradigm [24,25], incorporating both qualitative and quantitative evaluation methods. While these methods provide useful insights, expert judgments can sometimes be subjective and may not fully represent the preferences of the wider public [15]. In this regard, perceptual assessment focuses on how people experience streetscapes through measurable qualities such as enclosure, imageability, complexity, transparency, and human scale [26]. These dimensions relate to neighborhood perception via street connectivity, density, and accessibility, shaping perceived safety and social cohesion [8]. Empirical evidence further suggests that functional land-use exposure influences emotional place meanings: the results show that exposure to everyday urban facilities such as grocery shops is negatively correlated with place identity [3], while locations associated with leisure activities are negatively correlated with place attachment [27]. Specific features, especially greenery, contribute to urban livability and comfort, consistent with previous urban design studies [28], while extensive work has linked walkability with public health outcomes [29], reinforcing the value of human-centered street design [30].

Design elements strongly influence human experience. Enclosure is associated with pedestrian comfort and a sense of security [5]; environmental modifiers such as shading and shelter reinforce both perceived enclosure and safety. Narrower cross-sections or strong vertical edges (building façades, trees) can elevate perceived spatial quality and legibility [31], aligning with human scale, which supports orientation and psychological comfort. Collectively, enclosure influences perceptions of intimacy, confinement, and overall livability [32].

However, perceptual qualities have often been applied in assessment projects without significant adaptation and change, whether through objective or subjective approaches that feature six pioneers on urban design focused on walkability and pedestrian issues [8]. Methods such as public surveys, while more inclusive, can face practical challenges in large-scale data collection, whereas expert-based evaluations may be subject to inconsistent judgments.

2.3. History of Research by Semantic Segmentation on Urban Streetscapes

With advances in computational methods, quantitative techniques have emerged as scalable alternatives to traditional assessments. Approaches such as visual preference surveys, GIS-based modelling, and, more recently, deep learning applied to street-level imagery allow researchers in recent years to systematically evaluate streetscapes at large scales [33].

Beyond algorithmic development, studies have integrated subjective evaluations (e.g., pedestrian surveys and expert questionnaires) with objective computational methods such as syntactic mapping [34], wayfinding analysis [34], and image segmentation [13,33]. These combined approaches have provided valuable insights into the relationship between urban morphology, perceptual qualities, and pedestrian experience [15], although their application to heritage contexts remains limited [3].

Street view imagery offers a uniquely grounded perspective that reflects how individuals perceive and interact with the built environment. It captures critical design features such as enclosure, greenness, signage, and infrastructure that shape people's experience of walkability, comfort, and safety. Increasingly, machine learning techniques are being used in this field, offering alternatives to traditional observational methods that are often time-consuming and subjective. For example, Fang et al. [35] showed that machine learning can enhance the methodological robustness and practical application of urban analysis frameworks. Building on these approaches, Ma et al. [15] introduced an end-to-end model to evaluate urban aesthetics, revealing positive associations between perceived greenness, spatial openness, walkability, and overall visual quality. Other researchers, such as Nagata et al. [16], have found strong relationships between segmented visual components and walkability performance, reinforcing the relevance of deep learning in fine-grained analysis of the pedestrian environment.

Recent work continues to highlight the value of computer vision and machine learning in walkability assessment. For instance, Jeon and Woo [33] employed panoramic imagery to evaluate walkability in subsidised housing neighbourhoods across Seoul. Their results underscored the effectiveness of semantic segmentation in detecting spatial disparities in streetscape quality, revealing how such techniques can uncover urban design inequities that might otherwise go unnoticed.

Position of this Study

Building on the existing body of research, this study introduces semantic segmentation within the context of cultural heritage streetscapes. Unlike earlier studies that relied on Google Street View imagery, the present research employs First-Person Pedestrian View (FPPV) images to address the absence of street-level data in Iran. A U-Net-based segmentation framework is implemented to classify key streetscape elements such as buildings, trees, sky, and street furniture at the pixel level. Based on these outputs, perceptual indices including imageability, enclosure, human scale, and pedestrian density are derived to quantify how individuals perceive and experience heritage environments. To clarify the position of the current study and the methodological steps adopted, Figure 1 illustrates the research workflow, linking theoretical foundations from previous studies with data collection, deep learning-based analysis, and spatial-statistical evaluation.

This study advances the application of segmentation models by shifting the analytical focus from general urban environments to heritage preservation and pedestrian perception. It offers a novel methodological contribution to the sustainable transformation of historic streetscapes. While DeepLabv3+ is typically optimised for large-scale datasets and high-performance computing contexts, U-Net is more effective in scenarios where annotated data are limited. This makes it especially suitable for analysing pedestrian-oriented urban environments, providing accurate segmentation results even with small training datasets.

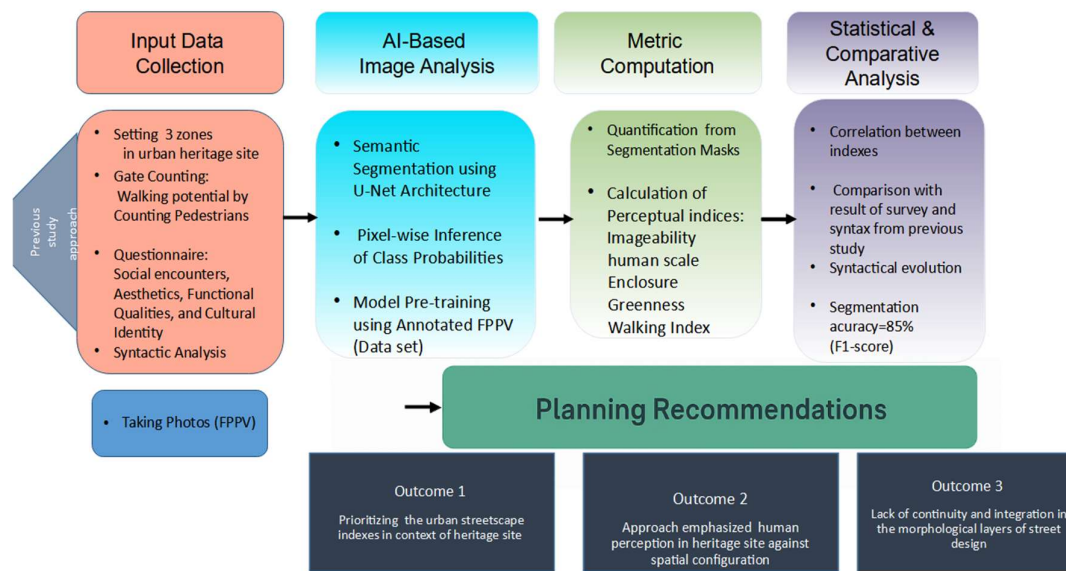


Figure 1. Workflow of the research framework, illustrating data collection (building on approaches used in previous studies in first step), analytical procedures, and applied spatial techniques (deep learning and statistical analysis).

3. Materials and Methods

This study builds upon a prior investigation of the same case study [3], which employed a combination of subjective and objective methods to evaluate street environments. Subjective assessments were based on participant surveys and expert evaluations of key aspects such as comfort, safety, accessibility, and visual quality. Concurrently, objective analyses utilised syntactic indicators derived from Geographic Information Systems (GIS) to measure spatial integration and connectivity. Collectively, these methods offered a comprehensive perspective on how spatial design influences walkability and pedestrian experience.

The earlier research underscored the importance of spatial configuration and network-level accessibility in shaping movement behaviour within heritage settings. It revealed that streets with higher syntactic integration tend to encourage more pedestrian activity, particularly when supported by features such as legible enclosures, tree cover, and seating. However, the emphasis remained largely on structural attributes, with limited capacity to systematically assess visual or perceptual qualities. This limited the scalability and objectivity of the findings.

To address these limitations, the present study incorporates semantic segmentation using a convolutional neural network known as U-Net to analyse First Person Pedestrian View (FPPV) images collected from pedestrianised streets in Isfahan. In contrast to previous approaches that relied on Google Street View, this method accommodates regions where such data is unavailable. The U-Net model enables pixel-level classification of key streetscape elements such as buildings, trees, sky, and street furniture. From these outputs, perceptual indices including imageability, enclosure, greenery, and human scale are derived to better understand how pedestrians experience heritage environments.

This approach advances streetscape analysis by offering a replicable and data-driven framework that reduces dependence on subjective evaluations. It allows for the systematic quantification of visual variables in urban design, particularly in culturally sensitive contexts. While DeepLabv3+ may be more appropriate for large-scale datasets, U-Net provides a more effective solution where annotated data is limited. By shifting the focus from general urban streets to culturally significant

heritage sites, this study introduces a novel methodological lens for supporting sustainable and pedestrian-oriented transformation of historic urban spaces.

3.1. Study Area

The study area is Isfahan, Iran, with emphasis on its UNESCO-listed Naqsh-e Jahan Square and the pedestrianized streets of Chahar Bagh and Sepah (Figure 2). These streets represent the historic urban core where cultural preservation and modern renewal intersect.

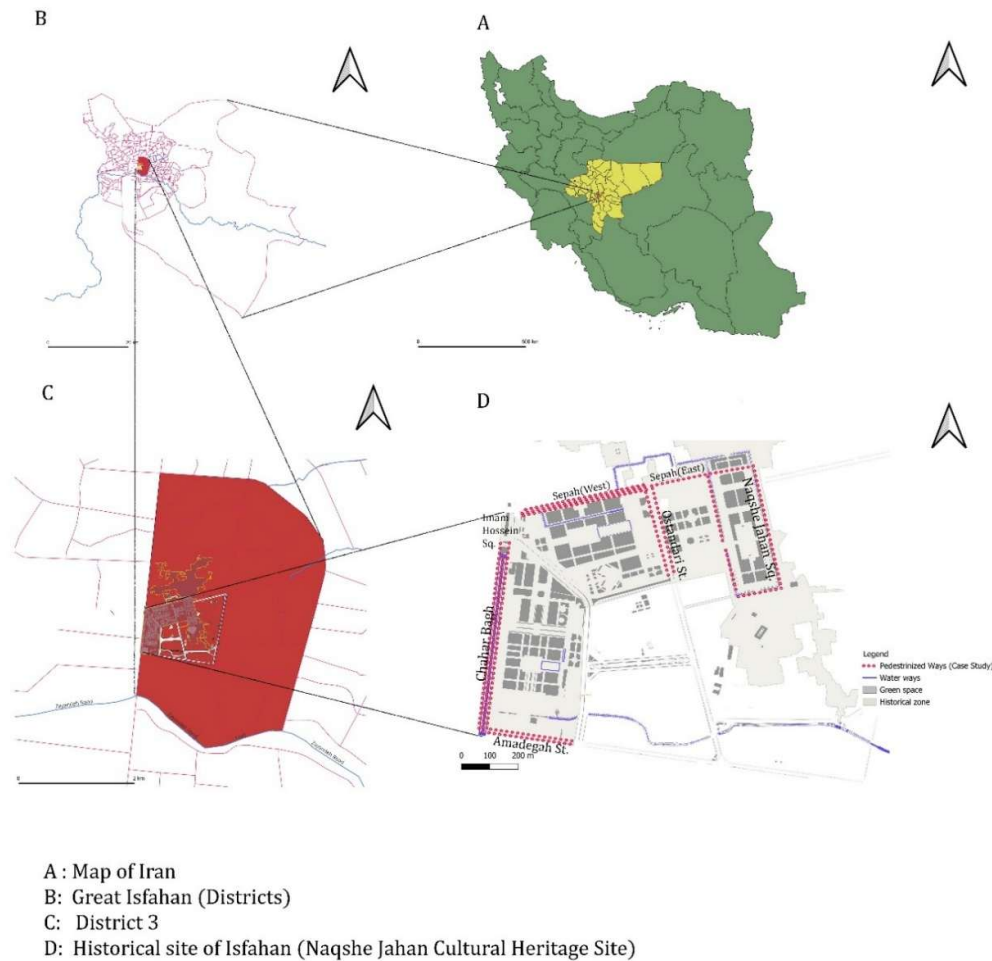


Figure 2. Case Study Map and Historical Area.

Recent municipal efforts have included relocating motorcycle parking and introducing new street furniture to promote walkability and encourage greater public use. Despite these improvements, several challenges persist. These factors make Isfahan an appropriate case for evaluating how semantic segmentation–derived perceptual indices can guide design improvements in heritage contexts.

3.1.1. Historical and Urban Context

Naqsh-e Jahan Square and Chahar Bagh Street are the main components of the movements constructed in Safavid Dynasty (Figure 3). The square of Naqsh-e Jahan remained in morphology the same even the functions of buildings and hubs (Figure 4). Just palaces around turned to museums. Chahar Bagh St. was originally 54.86 m wide, with eight rows of white poplar (*Populus alba*) and plane trees (*Platanus orientalis*). This street benefited five stone-lined water channels. The central lane of

Chahar Bagh was paved with stone (Figure 5) and flanked by gardens along both sides of the street, enclosed by screened walls that separated them from the street [36]. It is now 34 m. (Figure 6) and the area has undergone multiple physical and functional transformations over time. Despite these changes, the streets and the Square continue to embody layers of collective memory embedded in everyday social life. Beyond its historical significance, the recent pedestrianization has reshaped both the street and the garden within the Square, creating a renewed experience of movement along this urban axis. This transformation into a pedestrian-oriented corridor was intended to enhance urban livability and sociability. Nevertheless, empirical studies investigating the relationship between street width, spatial configuration, and quantitatively measured social elements remain limited, particularly in relation to perceived sense-of-place qualities.

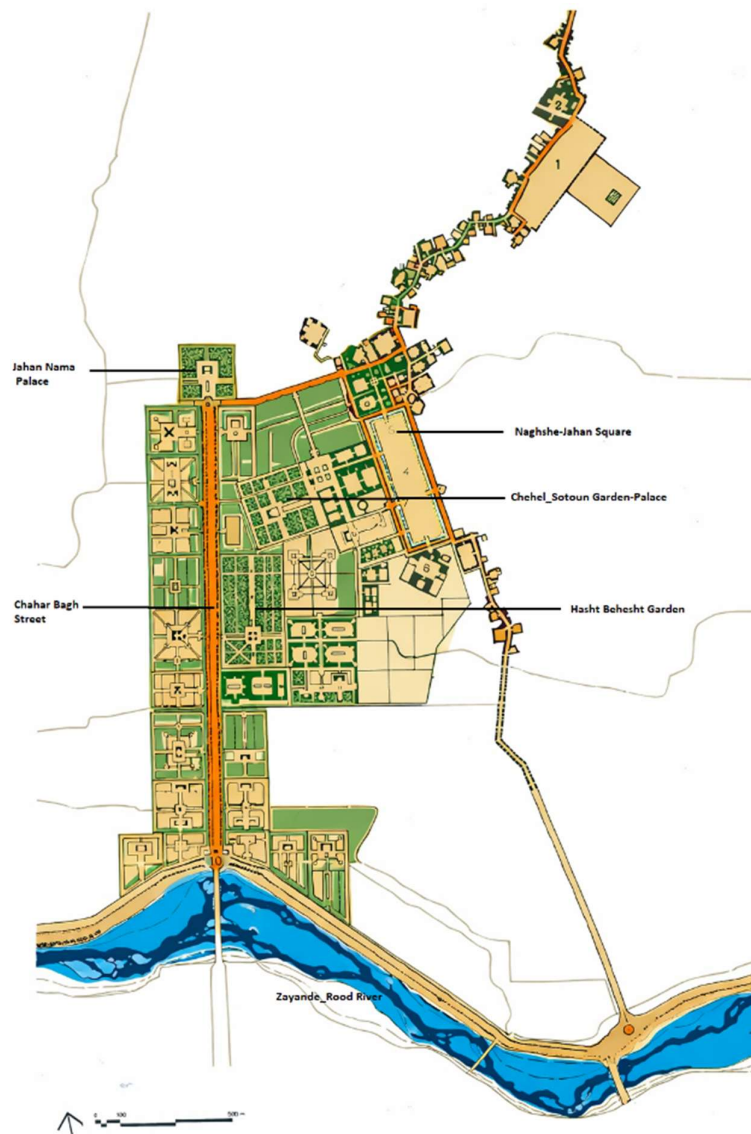


Figure 3. Plan of case study in origin shape (redrawn by the author) [36].

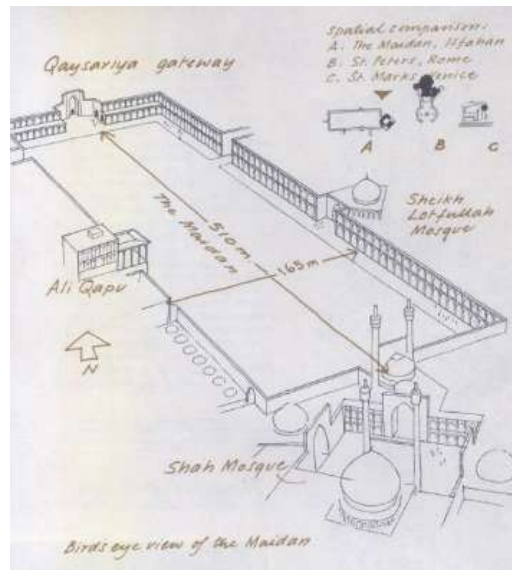


Figure 4. The birdview of Naqsh-e Jahan and its components [43].

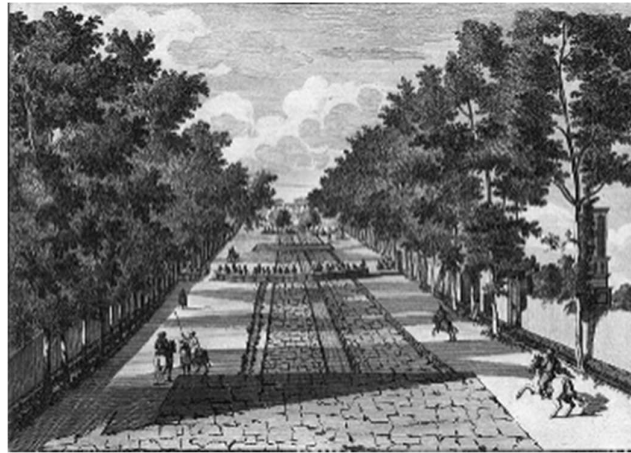


Figure 5. Chahar Bagh morphology in original and the garden palaces along the street [36].

3.1.2. Morphological Characteristics of the Case Studies

The selected case studies exhibit distinct morphological configurations in terms of street width and spatial layering, which are expected to influence pedestrian perception. Among them, Chahar Bagh Street represents a formally structured promenade characterized by a wide cross-section and a symmetrical spatial composition. Its layout consists of multiple pedestrian layers, including side sidewalks, tree-lined zones, central walkways, and a continuous water channel, which together create a strong sense of enclosure, rhythm, and visual continuity. In contrast, the other pedestrianized streets and the square display comparatively simpler and less hierarchical spatial arrangements, with narrower widths and without integrated water features. These morphological differences provide an essential spatial context for interpreting variations in perceived streetscape qualities across the case studies.

A detailed cross-sectional representation is provided for Chahar Bagh Street, as it constitutes the main experiential spine and uniquely integrates a central water channel and historically designed as a processional promenade. Other pedestrianized streets and the square are characterized by simpler and more heterogeneous spatial configurations; therefore, their morphological properties are represented through aggregated width measurements rather than detailed sections.

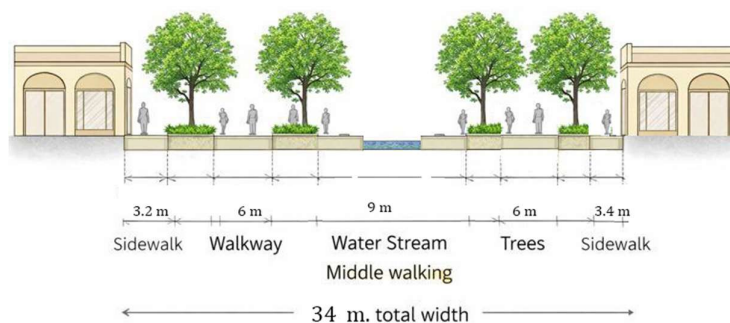


Figure 6. Cross-sectional profile of Chahar Bagh Street based on field measurements. Source: Authors (2025).

3.2. Methodology

This study analysed twelve pedestrianised street sections within the historic core of Isfahan, which had previously been assessed for subjective perception and syntactic integration. The methodology included four main phases: (i) collection of First-Person Pedestrian View (FPPV) images, (ii) semantic segmentation using a U Net model, (iii) calculation of perceptual scores, and (iv) evaluation of walking potential.

Street level images were captured every twenty metres during September and October 2024, resulting in a dataset of 1,026 annotated scenes. Because Google Street View is unavailable in Iran and existing datasets such as ADE20K lack contextual relevance, a new dataset was developed. Each image was manually annotated at the pixel level into predefined classes, including buildings, trees, sidewalks, and sky. Annotation was carried out by two trained researchers who followed a standardised labelling protocol to maintain class consistency. To improve reliability, a random subset of fifty images was cross checked by both annotators, and any discrepancies were resolved through discussion. The dataset was then divided into 784 training images (76.4 percent), 139 validation images (13.5 percent), and 103 test images (10.0 percent) using a stratified random procedure. This ensured reliable evaluation while maintaining adequate diversity in both the training and validation sets.

For model training, the images were subdivided into 1,026 annotated patches with a size of 512 by 768 pixels, which served as the final input samples. In this context, 1,026 refers to the total number of annotated image patches (training samples) generated from 195 original street scenes. The dataset split remained consistent at 784 training images (76.4 percent), 139 validation images (13.5 percent), and 103 test images (10.0 percent). Data augmentation was applied exclusively during the training phase of the semantic segmentation model to improve robustness. Pixel-based perceptual indicators were computed only for the original, non-augmented images to avoid duplication of visual content and ensure statistical independence of observations.

The U-Net model achieved a mean pixel accuracy of approximately 83 percent, with performance evaluated using Intersection over Union (IoU) and Dice coefficients [40]. The resulting semantic maps allowed a structured analysis of urban scenes and were aggregated to generate perceptual indices.

Following semantic segmentation, pixel-level outputs for each annotated image patch were converted into quantitative data for statistical analysis. For each semantic class, the proportion of pixels relative to the total image area was calculated, producing normalized pixel-based indicators representing visual elements such as greenness, enclosure, openness, and built form. These values

were first exported into structured spreadsheet files and subsequently organised into a relational SQLite database to enable efficient handling of multi-image and multi-class data.

Statistical analysis was conducted using Python 3.11, employing standard scientific libraries to compute descriptive statistics, including means and standard deviations, for each perceptual indicator across all street sections. Correlation analyses were performed to examine relationships between visual elements and derived perceptual qualities.

Pearson correlation was used to assess linear relationships between continuous variables derived from aggregated pixel-based proportions ($n = 41$), where approximate linearity was observed. Spearman's rank correlation was applied for analyses involving smaller sample sizes and potential non-normal distributions, particularly in examining the relationship between street width and perceptual indices ($n = 12$).

This database-driven workflow ensured reproducibility and enabled scalable analysis of streetscape visual composition while maintaining a transparent link between segmented imagery and perceptual indicators.

3.2.1. Selection of Perceptual Indices

Recent research has emphasized the importance of integrating multiple spatial and visual attributes when evaluating pedestrian environments, rather than assessing them in isolation. Foundational work by Ewing and Handy [37] identified key perceptual qualities such as imageability, enclosure, human scale, and complexity as essential components of streetscape assessment. Subsequent studies have extended this approach by incorporating quantitative measures derived from street-level imagery and spatial data, demonstrating that combined indicators provide a more robust and reliable representation of pedestrian perception [38,39]. This multidimensional perspective highlights the interaction between physical form and perceptual experience, supporting more comprehensive and scalable methods for assessing walkability and urban design quality.

Consistent with previous urban design studies [9,16], five perceptual indices were selected: imageability, enclosure, human scale, greenness, and walkability index. Complexity was excluded due to overlap with imageability. Raw segmentation outputs (pixel ratios of urban elements) were aggregated into these indices by logically grouping classes (e.g., façades, doors, and windows into imageability; tree canopy and vertical structures into enclosure).

This indicator system transforms pixel-level outputs into actionable urban design metrics, bridging quantitative deep learning results with perceptual quality evaluation. It allows for comparative assessment with subjective studies while maintaining interpretability for planning and renewal strategies.

A total of 195 street-level images were collected at regular 20-meter intervals throughout the study area, with one photograph captured at each sampling point. For model training, these images were augmented and divided into 1,026 annotated patches to improve segmentation accuracy. However, for the statistical analysis of perceptual indices (e.g., greenness, enclosure, imageability), calculations were based on the 195 original images. This ensured that each observation corresponded to an independent street segment, avoiding potential bias from multiple patches derived from the same scene.

The first Table summarises the operational definition of the selected perceptual indices and their corresponding computational formulations. Each index is derived from pixel-level proportions of semantically segmented streetscape elements and expressed as a normalized value at the street-segment level. Rather than presenting implementation-specific code, the table reports the conceptual structure and mathematical logic of each index to ensure transparency and reproducibility while remaining accessible to urban design and planning audiences. The formulations illustrate how low-level visual information is translated into higher-order perceptual constructs commonly used in walkability and streetscape assessment.

Table 1. A summary of using five indices for calculation of the perception scores.

Indices	Formula	Description
Imageability	$I_i = \frac{1}{n} \sum_{i=1}^n (Bn) + \frac{1}{n} \sum_{i=1}^n (TCn + TOn) + \frac{1}{n} \sum_{i=1}^n PDn + \frac{1}{n} \sum_{i=1}^n Wtn + \frac{1}{n} \sum_{i=1}^n PALn + \frac{1}{n} \sum_{i=1}^n Sn$	The sum indicates the total number of tree pixels in each image. Bn indicates the proportion of building pixels; TCn indicates the proportion of pixels of trees with columnar shape; TOn indicates the proportion of pixels of trees with oval shape; PDn indicates the proportion of pedestrian pixels; Wtn indicates the proportion of water pixels;
Enclosure	$E_i = \frac{\frac{1}{n} \sum_{i=1}^n Bn + \frac{1}{n} \sum_{i=1}^n (TCn + TOn) + \frac{1}{n} \sum_{i=1}^n Fn}{\frac{1}{n} \sum_{i=1}^n PVn + \frac{1}{n} \sum_{i=1}^n SKn}$	The sum indicates the total number of tree pixels in each image: Bn indicates the proportion of building pixels, TCn indicates the proportion of pixels of trees with columnar shape; TOn indicates the proportion of pixels of trees with oval shape; Fn indicates the proportion of fences and balustrade on the wall pixels; PVn indicates the proportion of pavement pixels; SKn indicates the proportion of sky pixels.
Human Scale	$HS_i = \frac{1}{n} \sum_{i=1}^n Sn + \frac{1}{n} \sum_{i=1}^n Pln + \frac{1}{n} \sum_{i=1}^n PDn + \frac{1}{n} \sum_{i=1}^n Fn$	Sn indicates the proportion of seats, symbols, windows pixels; Pln indicates the proportion of plant covers such as grass and flowers pixels; PDn indicates the proportion of pedestrian pixels. Fn indicates the proportion of fences and balustrade on the wall pixels.
Walking Index	$WS_i = \frac{\frac{1}{n} \sum_{i=1}^n PDn}{\frac{1}{n} \sum_{i=1}^n PVn}$	PDn indicates the proportion of pedestrian pixels, and PVn for pavement pixels
Greenness	$G_i = \frac{1}{n} \sum_{i=1}^n (TCn + TOn + PALn) + \frac{1}{n} \sum_{i=1}^n Pln$	TCn indicates the proportion of pixels of trees with columnar shape; TOn indicates the proportion of pixels of trees with oval shape; PALn indicates the proportion of palm trees pixels; Pln indicates the proportion of plant cover such as grass and flowers pixels

Imageability

Imageability refers to the extent to which a place is distinct, recognisable, and memorable [3,26]. It is largely influenced by the richness of architectural elements, street furniture, landmarks, and natural features such as trees and water [15,18]. In this study, the index includes buildings, historic structures, symbols, and seating elements, and extends previous formulations by explicitly incorporating trees and water because of their importance in shaping pedestrian perception.

Pixel-based segmentation results for these classes were combined to produce the imageability score, which represents the visual prominence of elements that contribute to cultural identity and aesthetic preference. For example, previous surveys have shown a preference for turquoise and light green water, illustrating how colour and materiality enhance the memorability of heritage environments. While earlier survey findings highlighted the community's sensitivity to the

preservation of historic buildings and greenery, this study translates such qualitative insights into a quantitative framework through the proposed formula.

Enclosure

Enclosure is a key spatial attribute that enhances walkability and comfort by shaping the degree to which pedestrians feel surrounded by vertical elements [24,31]. It is defined here as the ratio of built and vertical natural features (e.g., façades, walls, trees) to open horizontal space (e.g., sky, pavement).

In our approach, enclosure is calculated pixel-wise by using the visible proportion of buildings, walls, fences, and trees in the numerator, and the proportion of sky pixels in the denominator. This adapts Ewing's description of physical enclosure features [26], into a vision-based metric. The measure therefore reflects the perceived height-to-width relationship of urban canyons through visual rather than geometric parameters.

A higher enclosure score indicates stronger spatial definition and visual containment, which has been linked to perceptions of security, comfort, and belonging. Streets with continuous façades or tree rows typically achieve higher enclosure, contributing to a more compact and legible pedestrian environment.

Human Scale

Human scale is a fundamental principle of urban design that emphasises the relationship between the physical environment and the dimensions of the human body [10,26,31]. It describes the extent to which urban elements are designed to be perceptible, approachable, and meaningful from a pedestrian perspective. Streetscapes that maintain human scale provide details and features that can be experienced at eye level and walking speed, generally within the range of one to two metres.

In this study, human scale is defined through the presence of elements such as benches, signage, small vegetation including shrubs and flower beds, and windows at ground level. These components enhance legibility by breaking large building façades into finer details and by offering opportunities for direct human interaction. For example, benches and seating areas allow physical rest and social contact, signage and symbols support orientation and identity, and windows or doors at ground level encourage visual connection between private and public spaces.

Quantitatively, the pixel proportions of these elements were measured using semantic segmentation to estimate the density of human scale features across pedestrianised streets. This approach transforms an abstract design principle into a measurable indicator. A streetscape rich in human scale details is expected to enhance comfort, psychological security, and spatial familiarity, whereas environments lacking such elements may appear oversized, monotonous, or impersonal to pedestrians. By incorporating human scale into the perceptual indices, this study underscores its role in linking measurable spatial form with lived pedestrian experience.

Greenness

Greenness describes how much of the street scene is formed by trees and other natural features [15, 26, and 31]. Streets rich in greenery are often calmer, more inviting, and visually pleasing. In this study, trees were grouped by canopy form, some tall and columnar, others broad and oval, and palm trees were recorded separately to reflect the mix of species typical of Isfahan's streets. Smaller plants such as shrubs, hedges, flower boxes, and grass were also mapped to give a fuller picture of the green layer visible to pedestrians.

Looking at these vegetation types separately makes it possible to see how each one contributes to shade, enclosure, and the overall character of the street. Beyond the ecological value of vegetation, greenery in Iranian cities also carries cultural meaning, shaping comfort, identity, and a sense of belonging within the public realm.

Walkability Index

Instead of relying on the commercial Walk Score® metric, which measures accessibility to nearby amenities, this study defines a walkability index based on actual pedestrian activity. Building on earlier research that used syntactic integration and connectivity mapping, the analysis is extended here to include pixel-level pedestrian counts obtained from the 195 segmented images.

The index is calculated as the ratio between pedestrian pixels and pavement pixels, representing the intensity of pedestrian movement in relation to the available walking space. Higher index values do not directly indicate walkability, but rather show how effectively design factors such as accessibility, enclosure, greenery, and human scale encourage people to walk. This perspective links walkability with both spatial capacity and observed use, providing conceptual clarity and reducing the risk of misinterpretation.

3.3. Technical Information

Street View images (SVIs) provide valuable datasets for documenting and analysing detailed representations of urban environments. Through semantic segmentation, each pixel in an image is assigned to a specific category such as road, pedestrian, building, or greenery, allowing automated interpretation of street scenes. These techniques have become widely used in fields such as autonomous driving, robotics, geospatial analysis, and urban planning [40]. Fan et al. [41] discussed the potential bias of street view imagery in urban studies and proposed a computational method to estimate the proportion of environmental elements captured within SVIs, improving the reliability of streetscape assessments.

Open resources like Google Street View and large annotated datasets including Cityscapes and ADE20K have been central to the progress of segmentation research. The Cityscapes dataset, created for urban scene analysis, offers finely labelled imagery for understanding the structure of street environments [40]. ADE20K provides a broader range of indoor and outdoor scenes that support general-purpose image parsing. More recently, the Segment Anything Model (SAM), introduced by Meta AI, has demonstrated strong adaptability across segmentation tasks without the need for retraining for specific applications.

In this study, a U-Net model was used for the semantic segmentation of Isfahan's heritage streetscapes. U-Net is particularly effective when annotated data are limited because its encoder-decoder structure and skip connections preserve fine spatial details. The model was implemented using the Segmentation Models PyTorch library with MobileNetV2 as the encoder backbone, pre-trained on ImageNet. The encoder depth was set to five, and the decoder channel dimensions were 256, 128, 64, 32, and 16. Raw logits were generated with activation set to None, and Cross Entropy Loss was applied as the loss function. Model optimisation was conducted using the Adam optimiser with a dynamic learning rate scheduler, and early stopping was applied after twelve epochs without improvement in validation performance.

Performance Metrics

To evaluate the quality of the segmentation results, several quantitative metrics were employed. The training loss (`train_loss`) represented the pixel wise error during model fitting, whereas the validation loss (`val_loss`) measured the error on unseen data, providing an indication of generalisation and potential overfitting. The mean intersection over union (mIoU) was computed separately for the training and validation datasets and served as a more precise indicator of segmentation accuracy than raw classification accuracy. The accuracy metrics (`train_acc` and `val_acc`) reflected the proportion of correctly classified pixels, while the learning rate (`lrs`) represented the adaptive adjustment of the optimiser during training.

The model was trained for 32 epochs (Figure 8), and all performance metrics were recorded after each epoch. A consistent decline in both training and validation losses indicated stable learning behaviour and effective convergence of the model. Among all measures, the mIoU provided the most

reliable assessment of segmentation quality, capturing the balance between correctly and incorrectly classified pixels across classes.

Overall, this configuration was designed to achieve robust multi class segmentation under conditions of data scarcity. It produced stable and replicable perceptual indices, such as enclosure, imageability, human scale, and greenness that quantitatively describe the visual and spatial qualities of pedestrianised heritage streetscapes.

3.4. Measurement of Street Width as Physical Attribute

As one of the morphology attributes, street width was treated for each street rather than an image-specific variable. As multiple street view images were captured along the same street at approximately 30 m intervals, all images belonging to a given street segment were assigned very similar width value (Figure 7). To ensure scale consistency between morphological and perceptual variables, perception indices derived from semantic segmentation were aggregated at the street level using mean values prior to correlation analysis.

Street width and spatial layer dimensions were obtained through direct field measurements conducted along the selected pedestrianized streets and the square. On-site measurements were employed because available digital sources, such as Google Maps and OpenStreetMap, do not provide sufficient geometric precision for pedestrian-scale elements, particularly in heritage environments. In addition, several sections of the study area have undergone recent reconstruction and pedestrianization projects, resulting in discrepancies between current physical conditions and previously documented layouts.

Measurements were taken at representative segments along each street to capture the effective pedestrian space, including sidewalks, walkways, tree zones, and water channels where present. The recorded values were then averaged to represent the characteristic width of each case study and used as morphological indicators in the subsequent correlation analysis with perceptual variables.

To examine the relationship between street morphology and perceptual qualities, imageability scores derived from multiple street view images were aggregated at the street level. These aggregated values were then correlated with measured street widths using Spearman's rank correlation coefficient, given the small sample size ($n = 12$) and non-normal data distribution.

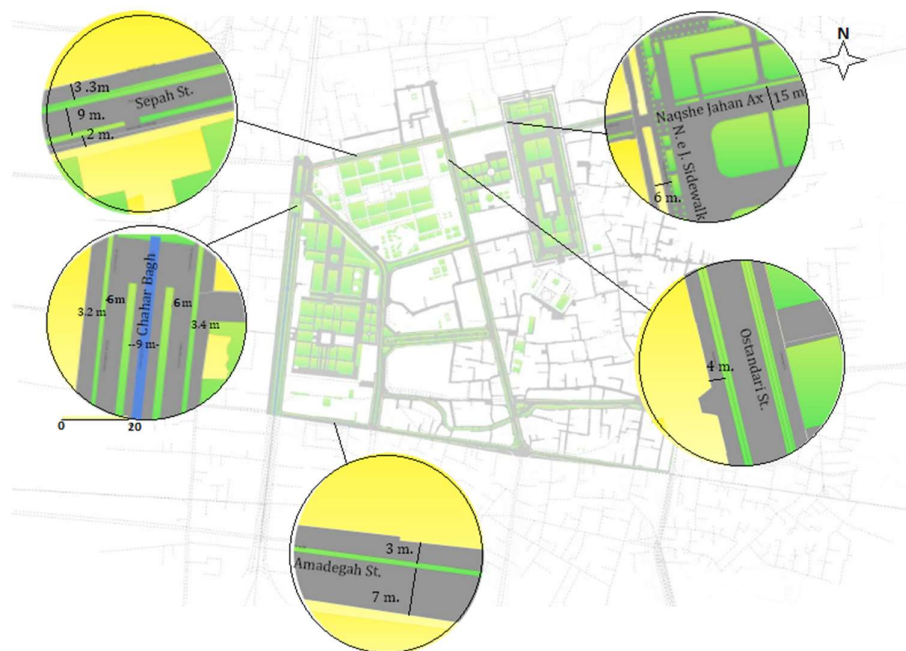


Figure 7. Comparative street width measurements across the selected case study. Source: Authors (2025).

4. Results

4.1. Performance of the U-Net Framework

The U Net architecture, which is widely recognised for its encoder decoder configuration and skip connections, was applied to the semantic segmentation of Isfahan's heritage streetscapes. In this framework, the encoder progressively reduced the spatial dimensions of the input images to extract hierarchical feature representations, while the decoder reconstructed the segmentation map through up sampling and detailed refinement. The skip connections between corresponding encoder and decoder layers helped preserve fine spatial information throughout the network. This structure allows U Net to maintain the spatial precision necessary for representing complex urban environments and makes it particularly suitable for applications involving limited annotated data.

The dataset consisted of 195 annotated training images covering twenty two semantic object classes. Annotation focused on pedestrian scale features, including façades, street furniture, vegetation, and human presence. Images were collected at different times of day to capture variations in lighting conditions and pedestrian density. High activity periods were represented by samples from 11:00 to 14:00 and from 17:00 to 19:00, whereas low activity periods were drawn from 09:00 to 11:00 and from 14:00 to 17:00. Pre-processing involved pixel normalisation, resizing image patches to 512 by 768 pixels, and extensive data augmentation through operations such as flipping, rotation, and scaling. These procedures, implemented using the Albumentations library, were intended to improve model generalisation and training stability.

Performance evaluation employed several standard segmentation metrics, including overall pixel accuracy, mean intersection over union (mIoU), class specific IoU, precision, recall, and the F1 score. A confusion matrix (Appendix Table 1) was also generated to identify common misclassifications, particularly among visually similar categories such as walls and building façades. The model achieved an overall accuracy of approximately 83 percent, with the highest IoU values observed for larger and more distinct classes such as sky, vegetation, and pavement. Lower IoU values were recorded for smaller or irregular categories such as windows. These outcomes are consistent with earlier segmentation studies, which report that class imbalance and fine scale variations often reduce predictive accuracy (Appendix Table 2).

While the U Net model effectively captured the dominant visual elements of the streetscape, challenges remained in areas characterised by low contrast, irregular boundaries, or overlapping features. These limitations highlight the need for larger annotated datasets and suggest that future research may benefit from hybrid or attention-based architectures to enhance model precision and robustness.

4.1.2. Class Merging and Optimization.

The dataset initially consisted of 34 annotated classes. During the analysis, it became evident that several categories overlapped or represented objects serving similar functions within the streetscape. To improve interpretability, these classes were progressively merged, resulting in 22 broader categories, as shown in Table 2. The consolidation process was guided by both visual similarity and the functional role of each element within its context. For example, brick wall with fence and metal fence were combined under wall or fence, while steps and path or sidewalk were grouped as pavement elements. Likewise, flower, flower box, and planter were merged under low vegetation cover, bulletin board and statue were combined as urban markers, and building wall together with building containing columns was simplified into building façade.

Following this reorganisation, the dataset comprised windows, doors, benches and food stands, walls and fences, steps and pavements, vegetation (including flowers, shrubs, grass, and trees), symbols and signs, human figures, water, domes, palm trees, sky, and vehicles. This refinement enhanced the internal consistency of the dataset and reduced confusion among visually similar classes, resulting in clearer annotations and more stable segmentation outcomes.

4.1.2. Hyper-Parameter Tuning

To enhance model performance, several hyperparameters and preprocessing steps were refined during the training process. The learning rate followed an adaptive schedule that supported steady convergence and reduced the likelihood of overfitting. The batch size was selected according to GPU memory capacity, maintaining a balance between computational efficiency and gradient stability. Because the dataset contained an uneven distribution of object classes, a weighted CrossEntropyLoss function was employed to account for this imbalance. The model was trained for a maximum of 32 epochs, and an early stopping criterion was applied when the validation loss showed no improvement across twelve consecutive epochs. This configuration ensured efficient use of computing resources while maintaining model robustness.

Before training, image data were normalised using the mean and standard deviation of the dataset. Augmentation techniques such as rotation, flipping, and rescaling were then applied to expand the diversity of training samples and improve generalisation. Each image and its corresponding segmentation mask were divided into smaller patches of 512 by 768 pixels, enabling the model to capture fine spatial details and adapt to variation in image composition.

Performance was assessed using pixel accuracy and mean Intersection over Union (mIoU). Throughout the training, the weighted cross entropy loss consistently decreased, and both training and validation curves indicated stable learning with limited overfitting, as illustrated in Figure 8. After about 25 to 30 epochs, accuracy levels for both sets converged, showing close agreement between predictions and the ground truth masks. The final model achieved an average pixel accuracy of approximately 0.83 and a mean IoU of 0.80 across the 22 visual classes (Figure 9). These results demonstrate that the U Net framework produced reliable and consistent segmentation outcomes under the applied configuration.

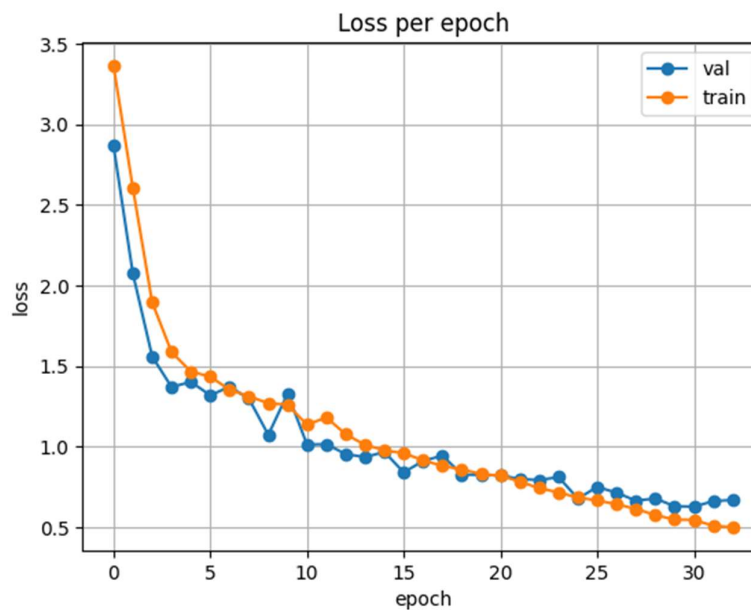


Figure 8. Loss metrics in model.

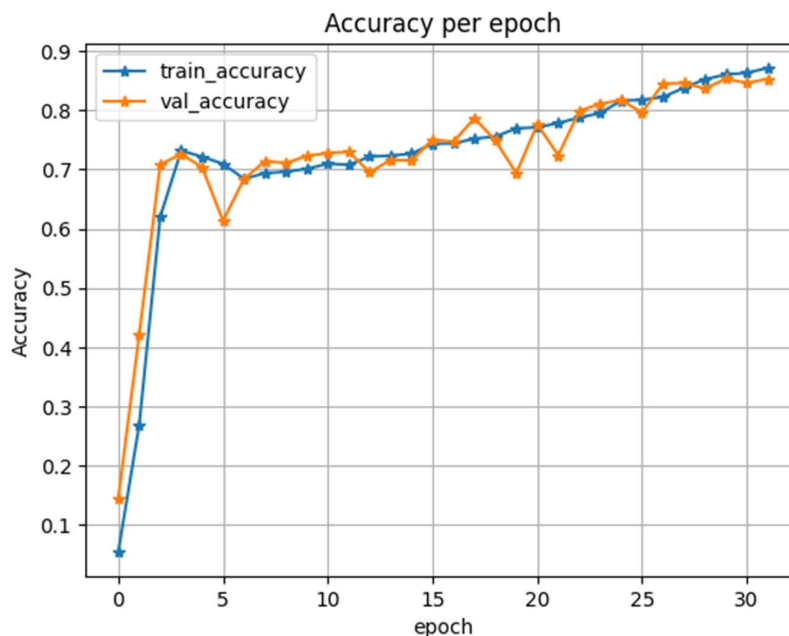


Figure 9. Accuracy of model.

4.2. Descriptive Statistics of the Segmented Results (First-Person Pedestrian Views)

The U Net model was used to perform pixel level segmentation on the collected street images, accurately identifying 22 distinct visual elements (Table 2). Descriptive statistics, including maximum, minimum, and standard deviation values, were calculated to describe the distribution of these elements and their relationship to the five selected walkability parameters. These parameters reflect how pedestrians perceive the morphology, identity, and overall quality of the streetscape. Among all detected elements, buildings, sky, and trees appeared most frequently, exerting a strong influence on imageability and enclosure and therefore shaping the overall walking experience.

Data from twelve pedestrian streets were processed, and the semantic segmentation results were compiled in Excel for subsequent analysis. Perception scores were computed according to the indices and formulas listed in Table 1. All perceptual indices were computed using custom Python scripts, with pixel statistics extracted from semantic segmentation outputs and stored in a SQLite database. Descriptive statistics and correlation analyses were performed using NumPy and Pandas libraries. The resulting values were visualised in Figures 5 and 6. Complementary pie charts illustrated the pixel distribution of key streetscape components. Each chart expanded radially in increments of ten percent, offering an intuitive view of the relative proportion of each visual element. Together, these graphics demonstrate that the five walkability parameters vary substantially between fully pedestrian streets and semi pedestrian pathways around Naqsh-e Jahan Square.

Table 2. Statistics of the segmented results from FPPV (%).

Indicators	Visual elements	Mean	Max	Min	S.D.
Classes in Dataset	1 Sky	.04	0.341271	0	.07
	2 Buildings (Bn)	.15	0.604509	0	.13
	3 Trees with columnar shape	.17	0.451022063	0	.11
	4 Trees with oval shape	.19	0.502243287	0.017691326	.13
	5 Palm trees	.007	0.214644	0	.04

6	Plant cover	.05	0.187119872	0	.03
7	Water	.013	0.251371126	0	.04
8	Wall (Bn)	.002	0.712366595	0	.06
9	Fence (Fn)	.01	0.249894023	0	.02
10	Symbol (Sn)	.002	0.46962293	0	.03
11	Sun shade (Sn)	.001	0.096984	0	.02
12	Pavement (PVn)	.27	0.483494471	0.028383866	.13
13	Person in Group (PDn)	.003	0.238764738	0	.03
14	Person Standing (PDn)	.03	0.426627789	0	.04
15	Person Sitting (PDn)	.003	0.091998322	0	.01
16	Shops Shutter (Sn)	.0055	0.516099	0	.03
17	Bench and Food stands (Sn)	.005	0.186044	0	.03
18	Window (Sn)	.0008	0.387737	0	.02
19	Door (Sn)	.007	0.15860	0	.03
20	Motorbike	.006	0.131518784	0	.03
21	Bike	.0005	0.204095	0	.01
22	Car	.003	0.029565	0	.03

Although dominant elements such as buildings, sky, and trees remained relatively consistent across locations, certain visual classes were confined to specific streets. For instance, water features were detected only in the central walkway of Chahar Bagh Street, while palm trees appeared solely along Amadegah Street. In several cases, walls took the place of buildings, particularly on the southern side of Sepah Street and along Ostandari Street; therefore, the class wall was substituted for building in the computation of perceptual indices. Windows were mainly identified on Sepah Street, whereas doors were more common in the Naqsh-e Jahan sidewalks. Both features were treated as indicators of human scale in the analysis. Pathways and sidewalks were combined into a single category, pavement, to ensure consistency in classification and measurement.

Table 3. Summary of dominant segmented classes across 12 pedestrianized streets.

Class / Element	Average Share (%)	Notes on Distribution
Buildings / Walls	~35–40%	Highest share across all streets; walls substituted for buildings in Sepah South and Ostandari.
Sky	~20–25%	Consistently visible in most panoramas, influencing enclosure.
Trees (Oval/Columnar)	~15–18%	Key contributor to enclosure and greenness; palm trees unique to Amadegah.
Pavement (Sidewalks/Pathways)	~10–12%	Unified class for consistency.
Windows / Doors	~5–7%	More visible in Sepah (windows) and Naqsh-e Jahan (doors).
Water	<3%	Only observed in the Chahar Bagh middle pathway.
Street Furniture / Symbols	<2%	Scattered elements, often associated with human-scale features.

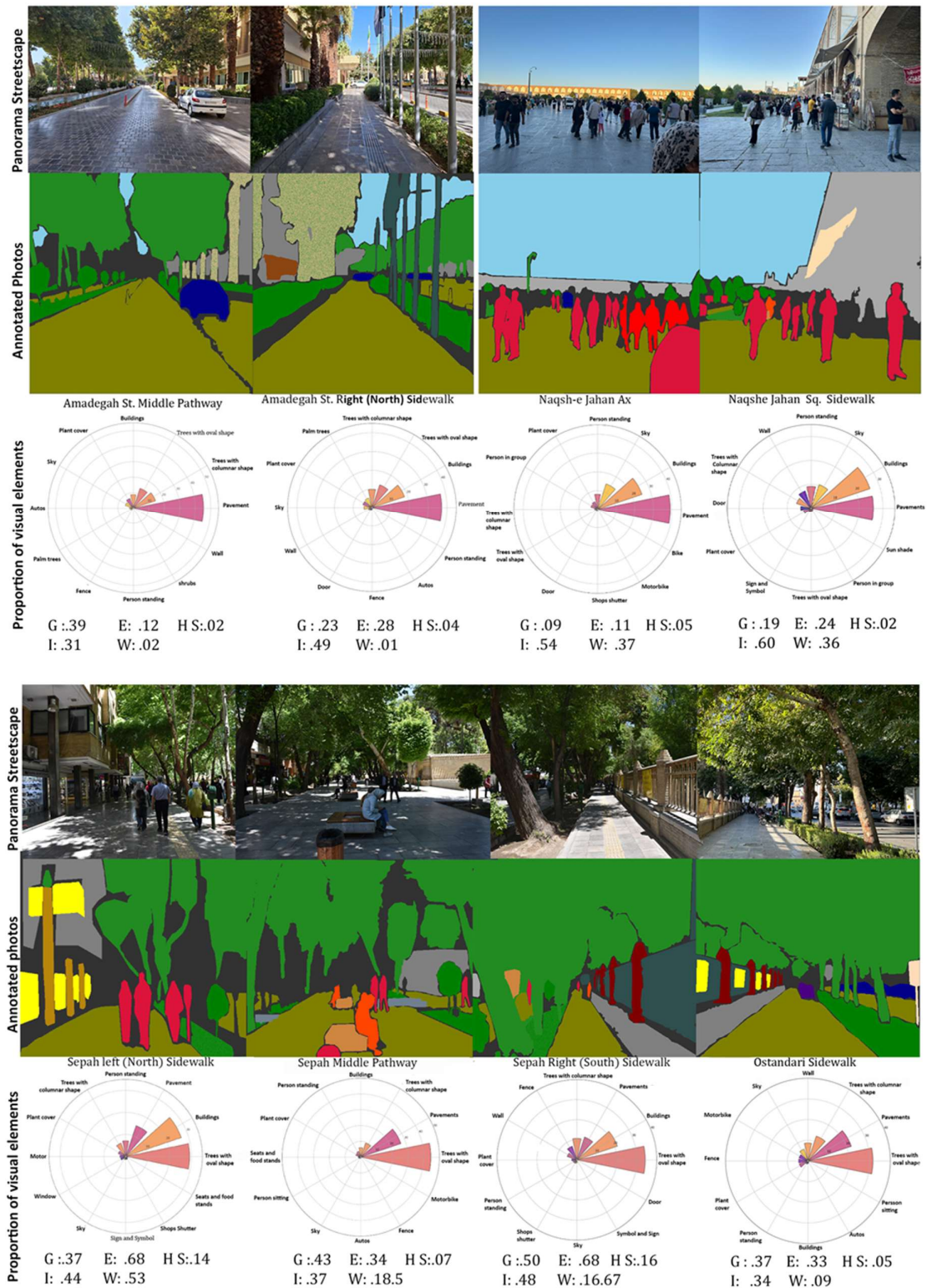


Figure 10. Streetscape and annotated photos of Sites, classified visual elements and perception scores (G: Greenness, E: Enclosure, H S: Human Scale, I: Imageability, W: Walkability index).

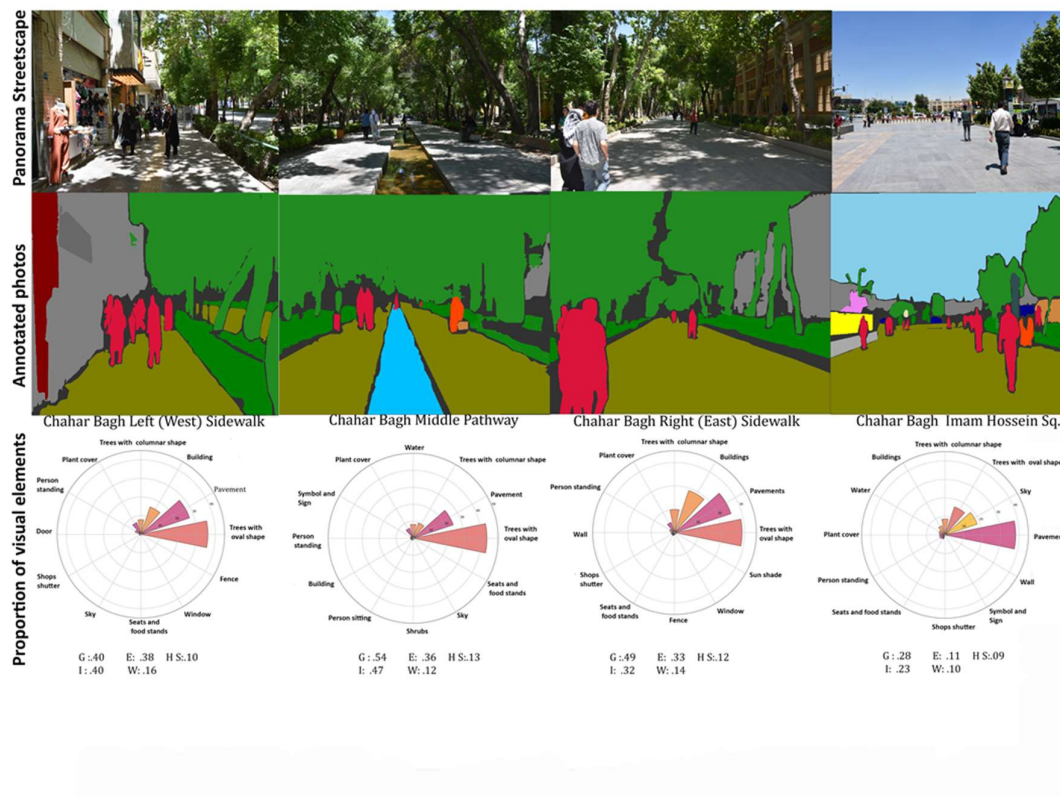


Figure 11. Streetscape and annotated photos of Sites, classified visual elements and perception scores (G: Greenness, E: Enclosure, H S: Human Scale, I: Imageability, W: Walking Index).

4.3. Correlation Analysis of Streetscape Perception Indices and Elements

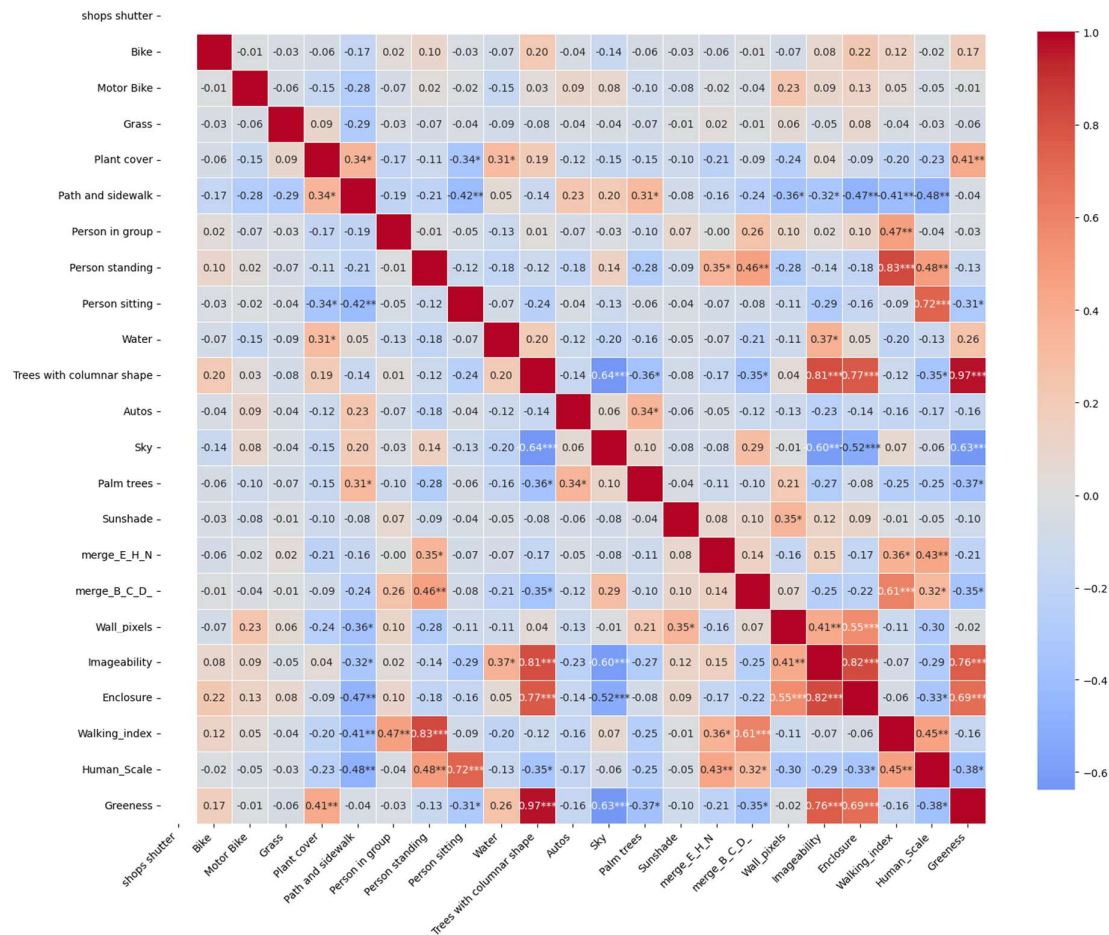
Normality of variables was assessed using the Shapiro–Wilk test. As several variables did not meet normality assumptions, Spearman’s rank correlation was also applied where appropriate to ensure robustness of the results.

Although the study area comprised twelve sampling sites, the statistical analysis was not limited to site-level averages. Instead, each annotated first-person pedestrian view (FPPV) was treated as an independent observation. To reduce local noise and enhance robustness, pixel-based indicators were spatially aggregated into 41 grouped means of 195 taken photos, each representing the mean composition of a small sequence of adjacent images. This approach provided a sufficient sample size for correlation analysis while preserving spatial variability along the streetscape.

It should be explained that 1026 photos are total images were used only for training the U-Net, not for perception analysis. If we included them in Excel, then it would artificially multiply the same perception and violate statistical independence. Data augmentation was applied exclusively during the training phase of the semantic segmentation model to improve robustness. Pixel-based perceptual indicators were computed only for the original, non-augmented images to avoid duplication of visual content and ensure statistical independence of observations.

The analysis was implemented in Python 3.11 using structured pixel-based data stored in a SQLite database, derived from one hundred and ninety-five annotated streetscapes. Given the aggregation level and sample size, moderate to high correlation coefficients indicate statistically meaningful relationships, while weak associations were interpreted cautiously. Classes exhibiting minimal variance across observations were excluded from analysis to ensure statistical validity.

Table 4. The correlation heatmap reveals several coherent clusters linking physical composition and perceptual outcomes (N = 41). The heatmap visualizes pairwise Pearson correlation coefficients (r), indicating the strength and direction of linear relationships. (Color intensity reflects correlation strength (dark orange = positive, blue = negative and yellow=weak). Statistical significance: * p < 0.05; ** p < 0.01; *** p < 0.001.



Strong relationships emerge among built elements, enclosure-related indicators, and pedestrian activity, while natural elements form a distinct cluster associated with greenness and visual comfort. These patterns indicate that streetscape perception results from interacting morphological components rather than isolated features.

Imageability demonstrates its strongest positive associations with enclosure ($r = 0.82$), trees with columnar shape ($r = 0.81$), and greenness ($r = 0.76$), indicating that visually memorable streets in the study area are primarily shaped by vertical green structure and spatial containment rather than by building mass alone. A moderate positive relationship is also observed with wall pixels ($r = 0.41$), suggesting that visible façade surfaces contribute to image formation, though less strongly than vegetation structure. Imageability is negatively associated with sky visibility ($r = -0.60$), confirming that excessive openness reduces perceived memorability.

The enclosure index shows strong and internally consistent relationships. It correlates positively with imageability ($r = 0.82$), trees with columnar shape ($r = 0.77$), wall pixels ($r = 0.55$), and greenness ($r = 0.69$), confirming that vertical elements and edge definition enhance spatial containment. Conversely, enclosure exhibits a negative correlation with sky visibility ($r = -0.52$) and a moderate negative relationship with path and sidewalk width ($r = -0.47$), indicating that increased openness and wider ground planes reduce perceived containment.

Human scale presents moderate but meaningful associations. It shows a positive relationship with person standing ($r = 0.48$) and the walking index ($r = 0.45$), suggesting that environments

perceived as human-scaled tend to coincide with active pedestrian presence near commercial frontages. Importantly, human scale demonstrates a moderate positive correlation with architectural detail clusters, particularly merge_E_H_N ($r = 0.43$), which includes elements such as benches and finer façade components, indicating that small-scale street furniture and edge articulation contribute to proportional perception. A moderate positive relationship is also observed with façade-related elements such as windows and doors (merge_B_C_D) ($r = 0.32$), suggesting that façade permeability and rhythmic openings reinforce the experiential sense of scale.

Conversely, human scale shows a negative association with enclosure ($r = -0.33$) and greenness ($r = -0.38$), implying that highly enclosed or heavily vegetated environments do not necessarily enhance perceived proportional intimacy.

Greenness shows a strong positive correlation with trees with columnar shape ($r = 0.97$), confirming that canopy vegetation is the dominant contributor to perceived greenery. It also correlates positively with imageability ($r = 0.76$) and enclosure ($r = 0.69$), suggesting that structured vegetation reinforces both memorability and spatial containment. Negative correlations with sky visibility ($r = -0.63$) and human scale ($r = -0.38$) indicate a trade-off between dense canopy coverage and openness or proportional perception.

The walking index is strongly associated with person standing ($r = 0.83$), and moderately associated with group presence ($r = 0.47$) and façade-related variables (merge_B_C_D) ($r = 0.61$), suggesting that pedestrian aggregation and spatial configuration, consistent with space syntax findings, contribute directly to walking activity. Interestingly, walkability shows weak or negligible direct correlations with imageability ($r = -0.07$) and enclosure ($r = -0.06$), indicating that pedestrian intensity is more strongly related to actual human presence and network structure than to perceptual qualities alone.

The moderate negative relationship between the walking index and path pixels ($r = -0.41$) suggests a spatial trade-off whereby larger exposed pavement areas correspond to lower pedestrian intensity, while higher walking index values reflect greater pedestrian concentration within more spatially defined corridors.

Overall, the correlation analysis reveals a structured pattern of interrelationships among streetscape components and perceptual indices. Strong positive associations—such as those between greenness and trees with columnar shape ($r = 0.97$), enclosure and imageability ($r = 0.82$), and enclosure and wall pixels ($r = 0.55$)—support the internal consistency of the segmentation-based methodology. At the same time, several negative relationships highlight inherent spatial trade-offs within urban form. Collectively, these findings suggest that walkable and visually coherent historic streets emerge from a calibrated balance between vertical definition, structured greenery, architectural articulation, and pedestrian activity, rather than from isolated spatial elements. To maintain clarity and avoid redundancy in the presentation of results, the detailed correlation coefficients are summarized in the heatmap (Table 4). While the full correlation matrix provides precise statistical values, the heatmap representation effectively conveys the direction, strength, and relative patterns of associations among streetscape elements and perceptual indices. This visual synthesis enables a more intuitive interpretation of complex interrelationships and supports comparison with spatial and perceptual patterns discussed in the results and discussion sections.

The heatmap visualizes pairwise Pearson correlation coefficients (r), indicating the strength and direction of linear relationships between pixel-based streetscape components and perceptual indices. For correlation analysis, segmentation outputs were aggregated into 41 spatial units to reduce pixel-level noise and avoid artificial inflation of observations, ensuring statistically meaningful relationships between morphological and perceptual variables.

Imageability exhibits strong positive correlations with structured vegetation ($r = 0.87$), greenness ($r = 0.84$), and enclosure ($r = 0.69$), alongside a strong negative association with sky visibility ($r = -0.66$). The near-zero relationship with wall pixels ($r = 0.06$) indicates that façade quantity alone does not generate memorability, supporting the interpretation of imageability as a compositional rather than purely architectural construct.

Greenness is strongly associated with tree presence ($r = 0.97$), confirming segmentation reliability. Moreover, its strong overlap with imageability and enclosure suggests that vegetation contributes simultaneously to environmental quality and spatial definition. Enclosure demonstrates positive correlations with both vegetation structure ($r = 0.77$) and wall pixels ($r = 0.55$), indicating that vertical boundaries—whether natural or built—collectively produce spatial containment. The Walking Index aligns most strongly with active human presence ($r = 0.83$) and enclosure ($r = 0.69$), suggesting that pedestrian intensity in the case study is socially and spatially conditioned rather than solely infrastructural. Human Scale exhibits comparatively moderate and diffuse correlations, reinforcing its perceptual complexity and supporting earlier claims that it does not follow a simple linear morphological pattern. All reported correlations above $|r| \geq 0.40$ are statistically significant (two-tailed, $p < 0.01$, $N = 41$).

4.4. Spatial Patterns of Aggregated Visual Perception Variables

Perception scores for the five indices were calculated for all twelve sampling sites. The pixel level segmentation results were then combined into spatial units and visualised through circle maps, as shown in Figure 12. These maps provide a clear and accessible way to interpret how visual elements and perceptual indices vary across the heritage streets. By examining these spatial patterns, it becomes possible to recognise whether a street is open and exposed to the sky or shaded by trees, to identify streets with stronger cultural or historical prominence, and to assess other spatial qualities such as pedestrian activity and enclosure.

The pedestrian density maps, which were generated from annotated first person pedestrian views, display a strong alignment with the integration and connectivity outcomes from the earlier space syntax analysis. In particular, the main axis of Naqsh-e Jahan Square and the northern section of Sepah Street show both high syntactic integration and dense pedestrian activity. This consistency supports the reliability of the image-based approach and demonstrates how it complements conventional syntactic techniques in evaluating spatial performance within heritage settings.

Streets that show high pedestrian potential, including Sepah Street and Naqsh-e Jahan Square, appear as key priorities for renovation and design enhancement. In Figure 12, these locations are represented in yellow and red, indicating higher walkability scores and highlighting the areas where improving spatial quality would likely have the most impact. It is important to note, however, that higher walkability indicates potential for pedestrian movement rather than a direct measure of visual or aesthetic quality.

The red marked segments correspond to areas with favourable visual conditions, such as a continuous building frontage that creates strong enclosure, tree canopies that offer shade, a pronounced human scale through façade articulation and pedestrian level features, and moderate pedestrian activity that contributes to a sense of liveliness. For instance, although Chahar Bagh Street is of major cultural importance, observations in 2022 showed that only one street café was present, limiting its ability to support a vibrant pedestrian culture.

Photographic documentation further clarifies these spatial differences. The left sidewalk of Sepah Street reflects many human scale qualities such as surface texture, building height, and façade rhythm, yet its narrow width restricts pedestrian movement. The right sidewalk across the street provides stronger enclosure due to continuous built edges but accommodates fewer people. In contrast, Naqsh e Jahan Square has greater openness and high sky visibility, conditions that attract more pedestrians even when the sense of enclosure is weaker.

Overall, the interaction between perceptual indices such as enclosure, human scale, and imageability defines distinct spatial experiences that cannot be explained by pedestrian density alone. The integration of pixel-based segmentation with spatial mapping allows for a more comprehensive understanding of how heritage streets perform both visually and functionally. From a design perspective, this approach offers practical guidance for improving heritage streetscapes by prioritising shaded, human scaled, and visually legible areas as the foundation for more comfortable and sustainable pedestrian environments.

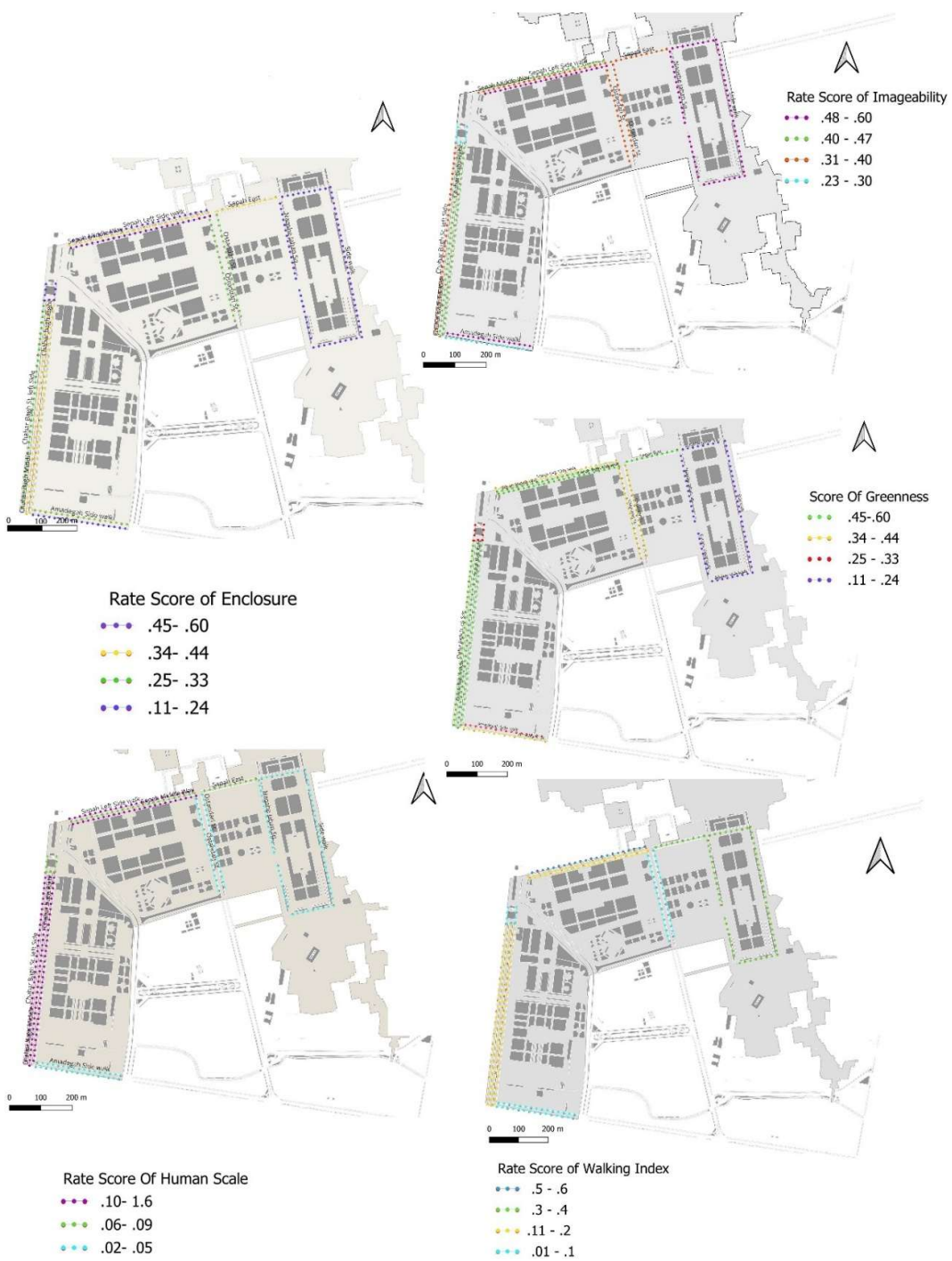


Figure 12. Perceptual indices in dotted maps in the study area.

4.5. Width of Street as a Morphological Attribute

Figure 13 illustrates the relationship between street width and aggregated perceptual indices across the 12 street segments. Spearman's rank correlation analysis indicated a moderate negative monotonic relationship between street width and imageability ($\rho = -0.466$); however, this association is not significant ($p = 0.13$).

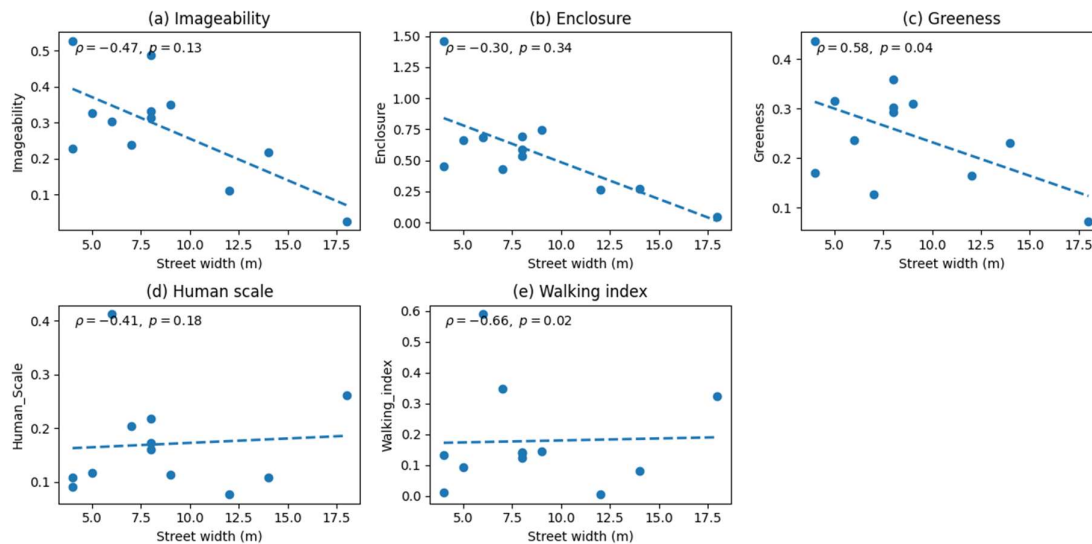


Figure 13. Relationships between street width and perceptual indices: (a) imageability, (b) enclosure, (c) greenness, (d) human scale, and (e) walking index. Points represent individual street segments. Dashed lines indicate linear trend lines added for visual guidance only. Statistical relationships were assessed using Spearman's rank correlation (ρ).

Street width indicated a negative tendency with enclosure ($\rho = -0.54$), suggesting that wider streets tend to be perceived as less enclosed, although this relationship is not significant ($p = 0.07$).

Spearman correlation analysis revealed no significant association between street width and perceived human scale ($\rho = 0.056$, $p = 0.862$). This suggests that human-scale perception is largely independent of absolute street width and is likely influenced by other spatial and visual factors such as enclosure, vertical elements, and façade articulation.

Spearman correlation analysis indicated a moderate negative association between street width and greenness ($\rho = -0.416$); however, this relationship was not statistically significant ($p = 0.178$). This suggests that street width as separate factor does not sufficiently explain variations in greenness across the studied streets.

Even when accounting for human presence and social activity, street width showed limited explanatory power. This suggests that pedestrian behavior in heritage streetscapes is shaped more by spatial quality, enclosure, and cultural affordances than by dimensional scale alone.

5. Discussion

5.1. Interpretation of Morphological–Perceptual Relationships in Relation to Street Structure

The results are better understood when interpreted through the structural configuration of individual streets rather than through isolated correlation coefficients. Semantic segmentation outputs were aggregated across grouped image sequences (41 spatial units), and mean pixel-based indicators were calculated for each street segment. The following discussion therefore links statistical relationships to the actual morphological character of the streets.

5.1.1. Structurally Enclosed and Architecturally Defined Streets

Sepah right sidewalk presents the highest Enclosure value, accompanied by relatively high Imageability (0.526) and moderate Greenness (0.436). This configuration reflects a strongly defined vertical edge condition, where architectural façades dominate the visual field. Generally, Imageability exhibited strong positive correlations with enclosure ($r = 0.82$, $p < 0.001$) and trees with columnar shape ($r = 0.81$, $p < 0.001$), indicating that spatial containment and vertically structured vegetation play a dominant role in shaping perceptual memorability. Despite its strong enclosure and imageability, the Walking Index (0.133) remains moderate, indicating that spatial containment indices not automatically generate pedestrian intensity. However, the Walking Index demonstrates a significant moderate correlation with Human Scale, suggesting that façade permeability—expressed through windows and doors—contributes to pedestrian activity. Street segments characterized by blank walls, such as the Chehelsotoon Garden edge, tend to show lower walking intensity, indicating that architectural openness supports street vitality even in contexts with strong cultural attachment. Similarly, Chahar Bagh Right Sidewalk and ChaharBagh, Middle Way, exhibit high Enclosure values (0.742 and 0.695 respectively) with moderate Imageability (0.349 and 0.487). Enclosure and columnar trees exhibited stronger associations with imageability than wall surfaces alone (Sepah Right, Chehel Sotoon Garden Wall), suggesting that vertical spatial definition and organized greenery play a more decisive role than façade repetition. However, their Walking Index values remain relatively low (0.145 and 0.124). These streets demonstrate that enclosure enhances visual coherence and memorability but does not independently drive pedestrian flow. These results totally align with syntax space results from these streets in previous study. This pattern reinforces the earlier correlation finding: enclosure is strongly associated with imageability ($r = 0.82$), yet only indirectly linked to walking activity.

5.1.2. Highly Walked but Weakly Enclosed Axial Spaces

A contrasting condition appears in Naqsh-e Jahan Ax and Naqsh-e Jahan Sidewalk.

- Naqsh-e Jahan Ax shows very low Enclosure (0.046) and very low Imageability (0.027), yet a high Walking Index (0.323).
- Naqsh-e Jahan Sidewalk exhibits moderate Enclosure (0.429) and low Imageability (0.239), but similarly elevated Walking Index (0.346).

These results confirm that pedestrian concentration in these axial segments is not driven by perceptual containment or visual classes, but rather by functional centrality and spatial integration. This finding aligns with the absence of a significant correlation between Imageability and Walking Index ($r = -0.07$, n.s.). In these cases, movement intensity is structurally driven by configurationally accessibility and socio-functional attraction rather than perceptual enclosure.

5.1.3. Human Scale as a Supportive Rather Than Dominant Factor

Human Scale values remain moderate across most streets. The highest value appears in Sepah Left Sidewalk (0.412), where Walking Index is also very high (0.588). This suggests that proportional articulation and façade detail like door and windows of shops may reinforce pedestrian comfort in already active streets.

However, other segments with moderate Human Scale (e.g., Naqsh-e Jahan Ax = 0.262 and Naqsh-e Jahan Sidewalk = 0.20; which pedestrian comfort for window shopping) do not exhibit correspondingly high enclosure or imageability. Furthermore, the absence of correlation between Human Scale and street width (Spearman $r = 0.056$, $p = 0.862$) confirms that proportional perception is not mechanically determined by geometric width alone. Instead, Human Scale appears to function as a perceptual moderator that enhances experiential comfort in active environments and acting as a primary generator of pedestrian flow. While historic façades, particularly in Naqsh-e Jahan Square, contribute to visual identity, it may have been an advantage to human scale classes which lead to higher human activity.

5.1.4. Vegetation Structure and Greenness

Greenness values are highest in Sepah Right Sidewalk (0.436) and relatively elevated in Chahar Bagh (0.359). These streets also show moderate to high enclosure levels, supporting the earlier correlation between Greenness and Enclosure ($r = 0.69$).

In contrast, Naqsh-e Jahan Ax = 0.072 and Naqsh-e Jahan Sidewalk = 0.125 show minimal greenness despite screening facades and high pedestrian intensity. This again reinforces the distinction between visual–environmental quality and socio-functional movement patterns. Vegetation contributes strongly to perceptual coherence and imageability, but not directly to pedestrian density.

5.1.5. Structural Differentiation of Street Types

Based on combined indices, three morphological street types emerge in layers of this historic context can be defined:

Type A – Visually Structured and Enclosed

(Sepah Right Sidewalk, Chahar Bagh Right Sidewalk, Chahar Bagh Middle)

High enclosure and imageability, moderate walking activity.

These streets prioritize visual containment and vertical articulation.

Type B – Functionally Active Axial Corridors

(Naqsh-e Jahan, Sepah Left Sidewalk)

High walking intensity, low enclosure and low imageability. Movement here is driven by accessibility and urban centrality.

Type C – Intermediate Transitional Streets

(Amadegah Right Sidewalk, Ostanandari St, Chahar Bagh Left Sidewalk)

Moderate enclosure, moderate greenness, moderate walking. These segments represent balanced but less specialized spatial conditions.

5.1.6. Structural Interpretation

When interpreted through actual street structure, the findings reveal a differentiated experiential logic:

- Walkability (Walking Index) operates primarily as a socio-functional condition driven by accessibility and pedestrian aggregation.
- Imageability operates as a spatial–visual condition shaped by enclosure and structured vegetation.
 - Enclosure functions as a structural mediator that enhances perceptual coherence but does not independently generate pedestrian flow.
 - Human Scale acts as a perceptual regulator that improves comfort within active streets but does not correlate directly with width or movement intensity.
 - Importantly, the highest pedestrian density (SepahL: 0.588) occurs in a segment combining moderate enclosure (0.686) with the highest Human Scale (0.412), suggesting that pedestrian vitality emerges when configurational accessibility aligns with perceptual proportionality—not from visual memorability alone.

5.1.7. Implications for Future Research

While the present study provides a structured analysis of morphological and perceptual indices within selected heritage streets, further research is required to expand both the spatial scope and methodological depth of the investigation. Future studies could incorporate additional case studies across a wider network of surrounding streets to enhance comparative robustness. Integrating large-scale spatial coverage through updated Google Maps and Google Street View (GSV) datasets would allow for broader geographic sampling and temporal updating of visual data.

At a larger scale, automated image segmentation using advanced deep learning architectures such as DeepLabv3, a well-established model in computer vision, could improve classification accuracy and enable more comprehensive pixel-level measurements across extensive urban areas. Such an approach would strengthen the scalability and reproducibility of morphological-perceptual analysis.

Moreover, expanding the research context beyond public streets to include semi-private environments—such as historic gardens—could provide valuable insight into the role of spatial configuration in shaping human activity under different levels of accessibility and enclosure. Comparative analysis between open street environments and semi-enclosed garden contexts would clarify how movement patterns and perceptual responses vary across urban typologies.

Finally, future investigations could incorporate behavioral observation methods to examine real-time reactions to streetscape elements, including pause duration, interaction patterns, and movement trajectories. Integrating behavioral metrics with spatial and perceptual indices would contribute to a more comprehensive understanding of how urban form influences human experience.

5.2. Contribution of the Study and Application of Pedestrian-Oriented Streetscapes

This research contributes to the field of urban design by linking computer vision techniques with morphological analysis, demonstrating how image segmentation can support the design and renewal of pedestrian environments. It provides a detailed examination of visual perception indices within the pedestrianization program of Isfahan's historic core and shows how technological methods can complement design thinking. In contrast with earlier approaches that depend largely on surveys or manual pedestrian counts, this work applies deep learning based semantic segmentation to produce objective and reproducible measures of spatial cohesion and accessibility.

The segmentation outputs were verified through manually annotated samples to ensure accuracy. The U-Net model achieved a mean Intersection over Union (mIoU) of 0.80 and comparable pixel level accuracy, confirming the reliability of the perceptual indices reported in earlier work. These results validate the approach as a credible means of quantifying spatial perception, offering a methodological bridge between data science and urban design practice.

The analysis highlights that perceptions of streetscapes vary significantly across streets serving different functions. The uneven distribution of the five indices reveals how walkability values can be influenced by both morphological form and functional diversity. This variability provides planners with practical guidance for prioritising urban renewal. Streetscapes that scored highly on imageability were generally characterised by continuous façades, historic architectural detail, and distinctive frontage design. When combined with human scale elements and tree canopy shading, these attributes strengthened enclosure and improved pedestrian comfort. At the same time, the findings indicate that functional aspects, particularly access to shops, services, and public transport, had a stronger influence on pedestrian density than visual quality alone [3].

A key strength of this study is its exclusive focus on pedestrianised streets rather than citywide averages. This narrower perspective makes the findings directly relevant to pedestrian oriented planning, where morphology-based evaluations can guide renewal and retrofitting efforts. When segmentation-based walkability maps were compared with space syntax integration maps, strong spatial correspondence was observed. Areas showing high pedestrian intensity, such as Sepah Street and Naqsh-e Jahan Square, coincided with regions of high syntactic integration. This agreement across analytical methods strengthens the credibility of the segmentation approach and demonstrates how visual perception indices can complement traditional spatial tools in assessing urban performance.

Beyond technical validation, this study extends the conceptual understanding of walkable heritage environments. Earlier research has often emphasised landmarks, memory, and reference frames in pedestrian wayfinding [44]. The present analysis advances this discussion by using deep learning to quantify how physical design features affect real time visibility and perception in historic contexts. The framework presented here offers planners a robust and adaptable tool for evidence-

based interventions, supporting the creation of streets that are both culturally sensitive and experientially rich.

5.3. Triangulation with Previous Empirical Findings (Integrated Triangular Conceptual Model)

This study extends prior research on Isfahan's pedestrianized heritage streets by integrating AI-based semantic segmentation with earlier questionnaire and space syntax analyses. Rather than simply confirming earlier findings, we present refined results and conceptually differentiate them within a unified analytical framework. Consistent with the space syntax analysis, accessibility remains the strongest predictor of pedestrian movement. However, the current findings clarify that imageability does not directly translate into walking intensity. Instead, imageability operates primarily as a spatial-visual dimension associated with memorability, while enclosure emerges as a structural mediator linking perceptual coherence and pedestrian continuity. The role of vegetation further illustrates this differentiation. Trees contribute positively to greenness, enclosure, and human scale, enhancing perceptual comfort. Yet their influence on imageability is context-dependent, as dense canopies may partially obscure architectural features. This nuance reinforces the importance of interpreting visual indices within a broader structural and cultural framework rather than as isolated predictors of movement.

Space syntax results identified integration and connectivity as the strongest predictors of pedestrian flow, a pattern consistent with the current model in which enclosure and social presence correlate more strongly with walking than imageability components. Furthermore, questionnaire findings demonstrated that wayfinding behavior was significantly influenced by sense of attachment and the perceived cultural-historic value of destinations. This affective dimension appears to operate independently of formal imageability metrics, suggesting that symbolic meaning and heritage recognition attract movement beyond purely morphological or visual descriptors. The convergence of computational perception in deep learning results by data analysis, behavioral surveys in objective method, and configurational analysis thus reinforces the robustness of the findings and reduces the likelihood that observed relationships are method-specific artifacts. The triangle logic model (Figure 15) clarifies that walkability, imageability, and attachment operate as related but distinct dimensions within historic urban environments.

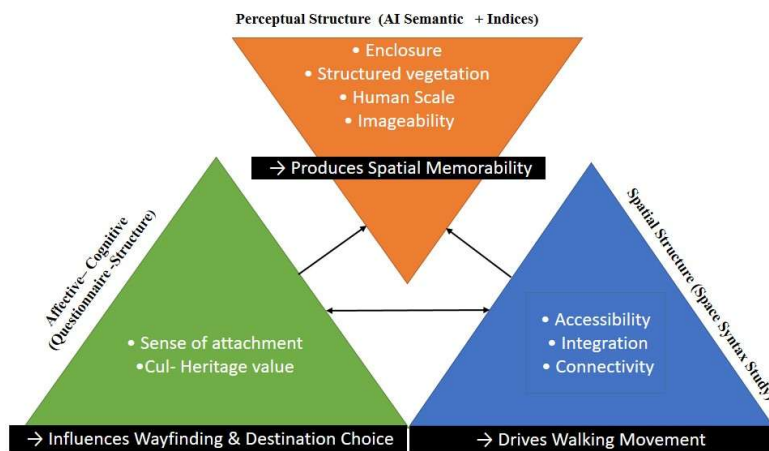


Figure 15. Integrated Triangular Conceptual Model.

Overall, the triangulation of configurational analysis, perceptual surveys, and AI-based visual metrics supports a multi-layered understanding of pedestrian experience in heritage streets in which behavioral dynamics cannot be reduced to visual form alone. Accessibility explains where people move, perceptual structure shapes how spaces are experienced and remembered, and affective attachment influences destination choice and shapes wayfinding. By distinguishing these interrelated yet independent dimensions, the study advances a more integrated model of heritage urbanism that bridges functionality, perception, and cultural meaning.

To empirically validate the triangulated framework, AI-derived perceptual metrics were compared with previously reported questionnaire-based assessments. Tables 4–6 summarize the comparative results for imageability, enclosure, and human scale, enabling evaluation of methodological convergence and construct stability across independent analytical techniques.

Table 5. Comparison of Imageability assessment between previous study and current study.

Aspect	Previous Study	Current Study (This Manuscript)
Concept & Method	Emphasized landmarks, architectural memorability, and street aesthetics as drivers of spatial imageability and walkability. Based on surveys, expert opinion, and field observation (subjective assessment).	Extends prior focus by quantifying visual components (buildings, trees, pedestrians) through semantic segmentation. Uses pixel-based ratios to construct an imageability index. Non-historic edges were down-weighted in formula.
Findings	Imageability linked to symbolic/memorable architecture and pedestrian density. Chahar Bagh benefited from shade and water, enhancing spatial appeal. Landmark buildings acted as termination/orientation points.	Imageability shows trade-offs with other elements; landmark buildings remain decisive (visual termination, contrast, orientation). Negative correlation with trees ($r = -0.325$) suggests their inclusion in the index requires contextual interpretation.
Innovation	Recognized weaker imageability in streets with plain façades (e.g., Chahar Bagh, Sepah)	Combines perception-based indices with pixel data, showing that façade richness, historic character, and functional trees enhance imageability in heritage-sensitive design. Future work can refine landmark detection in segmentation models.

Table 6. Comparison of Enclosure assessment between previous study and current study.

Aspect	Previous Study	Current Study (This Manuscript)
Concept & Method	Discussed enclosure elements only implicitly—through shading, tree canopies, and spatial continuity. No formal definition or measurement of enclosure; pedestrian surveys could not directly capture it.	Pixel-based segmentation quantifies proportions of buildings, trees, fences, pavement, walls, and sky to compute a measurable enclosure index.
Findings	Key excerpts point to enclosure-related elements: buildings define edges; fragmented façades reduce continuity; trees add comfort but without measurement.	In heritage streets, enclosure tends to be low where façades are discontinuous. Tree morphology (oval vs. columnar canopies) strongly shapes enclosure, while greater sky visibility signals weaker enclosure.
Innovation	Lacked quantification—treated enclosure as theoretical.	Proposes a refined enclosure metric as an image-based index that integrates vegetation, vertical density, and sky exposure. Links enclosure with walkability and street width, highlighting pedestrian-friendly design needs.

Table 7. Comparison of Human Scale assessment between previous study and current study

Aspect	Previous Study	Current Study (This Manuscript)
Concept & Method	Based on visual analysis, pedestrian surveys, and expert questionnaires. Attention given to benches, cafés, shopfronts, and street furniture.	Semantic segmentation of street images, focusing on spatial proportions and presence of human-scaled elements (façade continuity, windows, seating, shading).
Findings	Human scale enhanced by commercial frontage, urban furniture, shade, and active edges (cafés, vans). Elements at eye level encouraged lingering.	Dataset lacked consistent visual capture of furniture, cafés, signage. Windows often undetected due to absence in many images. Human scale therefore depended strongly on spatial framing and façade continuity.
Innovation	Addressed design factors qualitatively but without quantification.	Among the first attempts to quantify human scale pixel-wise. Shows its dependence on enclosure and continuous building façades, confirming links between street wall definition and pedestrian comfort.

As summarized in Table 7, the present study introduces a quantitative approach to assessing human scale through semantic segmentation. Whereas earlier research relied primarily on surveys and field observations, the current method operationalizes human-scale qualities through measurable spatial proportions and visual components.

The findings indicate that human scale is closely associated with enclosure and façade continuity, reinforcing the importance of defined street edges in shaping perceptual proportionality. However, its relationship with walking remains moderate rather than dominant, suggesting that human scale enhances experiential comfort without independently generating pedestrian movement.

By translating qualitative perceptions into pixel-level indicators, the study refines earlier interpretations while situating human scale within a broader structural framework that includes accessibility, enclosure, and affective attachment. This integration clarifies that human scale functions as a perceptual regulator within the pedestrian experience rather than as a primary determinant of urban vitality.

Table 8. Comparison of Walkability assessment between previous study and current study.

Aspect	Previous Study	Current Study (This Manuscript)
Concept & Method	Depth map-based space syntax (integration & connectivity) combined with field observations of pedestrian activity in heritage streets.	Advances walkability analysis by integrating semantic segmentation data of pedestrian activity (walking, sitting). Pixel-level element classification allows pavement coverage to be linked with pedestrian density and street width.
Findings	Integration directly influenced pedestrian flow: streets with higher connectivity supported more movement, regardless of aesthetic quality.	Walkability index showed a positive correlation with buildings ($r = 0.698$), suggesting urban form influences density. Discrepancies appeared where urban furniture and detailed façades were visually absent, lowering perceived walkability.

Aspect	Previous Study	Current Study (This Manuscript)
Innovation	Social behavior understood cognitively through integration maps, without direct visual metrics.	Treating walkability index as density of pedestrian activity, segmentation results can predict high-walkability zones with strong similarity to syntax analysis maps. This provides a visual validation of syntactic predictions.

The comparison between the previous research and the current study shows a methodological progression in the analysis of pedestrian behaviour. The present work builds upon earlier space syntax investigations by integrating pixel-level measures of pedestrian activity, shifting the focus from abstract integration values to directly observable spatial patterns. While earlier analyses identified connectivity as the main determinant of movement, the segmentation-based walkability index provides additional insight by illustrating how façade articulation, the presence of urban furniture, and visual richness influence walking potential. Together, these findings confirm the strength of connectivity-driven models while demonstrating the complementary value of visual perception metrics in guiding heritage-sensitive pedestrian design.

Table 9. Comparison of Greenness assessment between previous study and current study.

Aspect	Previous Study	Current Study (This Manuscript)
Method	Recognized tree cover, shading, and vegetation presence as contributors to comfort, walkability, and livability—particularly in Chahar Bagh. Field observations, questionnaires, and user perceptions described greenery's presence and priority.	Extends greenery analysis through pixel-based segmentation (trees, grass, planters), enabling quantification of vegetation in streetscape images. Allows direct measurement of relationships between greenness and other indices such as imageability and enclosure.
Findings	Trees and vegetation enhanced comfort and attractiveness; greenness was valued for shading and functionality but could also disrupt imageability or aesthetics. Favoured by pedestrians but not consistently measured. Suggested integrating ecological design into heritage planning.	Tree alignment and species strongly influence aesthetics and functionality. Trees moderately correlate with enclosure and human scale depending on form and placement. Tree-rich pedestrian streets (e.g., Chahar Bagh) showed higher perceived quality, while in Sepah St., trees combined with walls improved enclosure. Positive correlation observed between fences and greenness (Ma, 2021).
Innovation	Greenness acknowledged as beneficial but lacked a quantifiable framework; its role in cognitive perception was suggested but not measured.	Provides a quantifiable, image-based greenness index, differentiating tree types and their contribution to enclosure and imageability. Offers new insight into vegetation's functional role in shaping cultural perception, shading, and pedestrian wayfinding.

The comparison between the previous research and the current study reflects a methodological advancement in the analysis of pedestrian behaviour. While earlier space syntax investigations identified connectivity as the principal determinant of movement, the integration of pixel-level pedestrian activity and visual metrics allows behavioural patterns to be examined alongside perceptual structure. Rather than replacing connectivity-based models, the segmentation-derived indices complement them by revealing how enclosure, vegetation, and spatial definition contribute to the experiential dimension of walking.

In relation to greenery, the present approach extends earlier descriptive assessments by embedding vegetation within a multi-index analytical framework. Vegetation demonstrates strong associations with greenness, enclosure, and human scale, confirming its role in shaping spatial intimacy and perceptual comfort. At the same time, its influence on imageability is context-

dependent, as dense canopies may partially obscure architectural features. This dual effect underscores the importance of understanding greenery not merely as an environmental amenity, but as a structural component within the morphological and perceptual system of heritage streets.

At the methodological level, recent applications of semantic segmentation and street-view analysis have shown that pixel-based metrics can effectively quantify visual walkability and perceptual indices [15,16,44]. The findings of this study align with a growing body of international research linking spatial form, visual perception, and greenery to pedestrian experience. Prior studies have demonstrated that enclosure, façade continuity, and spatial rhythm contribute to perceived comfort and pedestrian presence, supporting the present observation that defined vertical edges and structured vegetation enhance experiential coherence [29]. Research on tree canopy and shading similarly confirms the positive relationship between greenery, thermal comfort, and perceived walkability [13,42].

Importantly, while international research often emphasizes visual predictors of walking, the present findings refine this perspective by demonstrating that, in a heritage context, accessibility remains the dominant structural driver of movement. Visual factors such as enclosure, vegetation, and façade articulation shape perceptual comfort and memorability but do not independently determine pedestrian flow. This distinction contributes to a more differentiated understanding of walkability across cultural and spatial settings.

5.4. Width Measurement Correlations

While statistical significance remains marginal, the direction and magnitude of the relationship indicate a systematic morphological tendency rather than random variation, particularly within historically structured pedestrian streets.

When examined in relation to enclosure, street width demonstrates a moderate negative association ($\rho = -0.54$, $p = 0.07$), indicating a tendency for narrower sections to exhibit stronger spatial definition. Although this relationship does not reach conventional levels of statistical significance, it reflects a morphological pattern consistent with the significant negative correlation observed between enclosure and path in the heatmap analysis. These findings suggest that reduced cross-sectional width is associated with increased vertical containment and decreased sky exposure.

A similar negative trend appears between width and imageability, implying that narrower heritage streets may intensify perceptual coherence through stronger enclosure and visual continuity. However, this relationship remains indirect. The results indicate that width influences perception primarily through its structural interaction with enclosure rather than as an independent determinant of imageability. Variability across the case-study streets and the presence of complementary spatial elements—such as vegetation and façade articulation—likely moderate this effect.

This divergence is further reflected in the walking index analysis. In heritage streets, higher walking index values tend to coincide with narrower widths, suggesting that spatial compression, enclosure, and pedestrian activity reinforce one another.

5.5. Design Strategies for Walkable Public Spaces in Heritage Sites

The findings of this study emphasize that morphological and perceptual qualities, particularly imageability, enclosure, and visual connectivity play a central role in shaping walkable environments within heritage contexts. Reconfiguring public spaces to prioritize these qualities can significantly enhance pedestrian experience while maintaining cultural authenticity. In Isfahan's pedestrianized streets, such as Sepah, Chahar Bagh, and Amadegah, reconfiguration has included the reduction or removal of vehicular access, the introduction of human-scale design elements, and the integration of culturally significant features into the public realm. These interventions have not only improved safety and accessibility but have also enriched the experiential quality of the streetscape by reinforcing enclosure, strengthening visual continuity, and supporting a distinctive spatial identity. For example, relocating motorcycles to designated parking lots and introducing additional benches

along Sepah Street would further enhance walkability by reducing conflicts with vehicles and encouraging lingering behavior.

Design strategies inspired by these typologies should focus on shaded pedestrian corridors, stronger perceptual enclosure through vegetation and architectural continuity, and visual coherence across façades and street furniture. Such interventions provide a dual benefit: preserving the cultural identity of heritage sites while aligning with contemporary objectives of sustainable mobility and public health. Urban designers and planners should therefore approach heritage streetscapes not only as functional spaces but also as perceptual landscapes, where spatial form and sensory experience work together to shape walkability and encourage social interaction.

6. Practical Design Guidelines:

1. **Enhance shading and enclosure** – Introduce continuous tree canopies, arcades, or colonnades that reduce sky openness and improve pedestrian comfort in hot climates.
2. **Strengthen human-scale features** – Integrate benches, shop front windows, signage, and localized cultural motifs at eye level to reinforce street vibrancy and encourage interaction.
3. **Balance heritage preservation with modern mobility** – Relocate conflicting modes (e.g., motorcycles, service vehicles) and create small rest areas, ensuring accessibility while preserving the cultural and visual identity of the street.

7. Conclusions

This study investigated pedestrian dynamics in historic streets by integrating established approaches, including space syntax modeling and questionnaire-based behavioral analysis, with current semantic segmentation analysis and morphological width measurements within a unified analytical framework. The findings demonstrate that pedestrian behavior in heritage environments is structured through three interacting yet analytically distinct subsystems: spatial configuration, perceptual structure, and affective–cognitive valuation, each conditioned by measurable geometric properties.

Configurational accessibility—expressed through integration and connectivity—remains the dominant structural predictor of pedestrian movement. However, correlation analysis indicates that street width and sky exposure ratios significantly influence enclosure values, thereby indirectly affecting walking continuity. Wider sections with greater visible sky tend to weaken spatial containment, whereas proportionally balanced width-to-height relationships enhance enclosure and stabilize pedestrian trajectories. Thus, width operates not as a direct driver of movement intensity but as a morphological moderator shaping perceptual structure.

AI-derived visual metrics further demonstrate that enclosure exhibits stronger alignment with walking continuity than imageability alone. Vegetation increases greenness and reinforces enclosure, yet its impact on imageability remains context-dependent due to façade visibility interactions. These findings confirm that perceptual indices are conditioned by geometric proportions and vertical boundary definition rather than by façade quantity in isolation.

Methodologically, the triangulation of space syntax metrics, width-based geometric indicators, survey data, and convolutional neural network–based segmentation enhances internal validity and reduces single-method bias. The Integrated Triangular Conceptual Model formalizes this multi-layered structure, positioning network accessibility as the structural driver of movement, width-conditioned enclosure as the perceptual stabilizer, and attachment as the cognitive selector of destinations.

Overall, the results indicate that pedestrian vitality in heritage streets depends on the interaction between network hierarchy, proportional street geometry, perceptual containment, and cultural meaning. By incorporating measurable width correlations into a computational–behavioral framework, the study advances a transferable and reproducible methodology for analyzing pedestrianized historic environments beyond the Isfahan case.

From a technical point of view, the segmentation model achieved classification accuracies above eighty five percent for most major visual categories, confirming its reliability as a tool for studying urban morphology in heritage settings. Future studies could extend this approach by using larger samples, a wider range of streets, and more complex statistical analyses to better understand the links among perception, spatial form, and behaviour. Comparative research in different climates and cultural contexts could also help identify patterns that support design decisions for pedestrian-oriented heritage areas.

Overall, the results show that every heritage street has its own character and spatial logic. The relationships among enclosure, greenery, and imageability should be viewed as context dependent rather than universal. The insights gained from this work can guide planners and designers in developing renewal strategies that respect cultural heritage while improving walkability and urban quality.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org.

Author Contributions: **Hessameddin Maniei:** Conceptualization, Methodology, Software, Visualization, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Final manuscript preparation. **Elham Mehrinejad Khotbehsara:** Review & Editing. **Dietwald Gruehn:** Supervision, Project administration, Methodology, Review & Editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: No new data involving human participants were collected for the current study. The research relied solely on previously obtained, non-personal, and anonymized data, which had already been ethically reviewed in the context of earlier related work. Therefore, ethical approval was not required at this stage. In accordance with institutional guidance (TU Dortmund University), the study falls under the category of research without own data collection, and no sensitive personal data (e.g., names, contact details, or identifiers) were gathered. The study aligns with the principles of the Declaration of Helsinki.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Morales-Flores, P.; Marmolejo-Duarte, C. Understanding the Relationship between Public Space Social Capital Expressions and Pedestrian-Oriented Design: A Case Study within the Superblock Action Plan in Barcelona. *Cities* 2025, 162, 105968. <https://doi.org/10.1016/j.cities.2025.105968>.
2. Cervero, R. Public Transport and Sustainable Urbanism: Global Lessons. In *Transit Oriented Development: Making It Happen*; Curtis, C.; Renne, J.L.; Bertolini, L., Eds.; Ashgate: Aldershot, UK, 2009; pp. 37–50.
3. Maniei, H.; Askarizad, R.; Pourzakarya, M.; Gruehn, D. The Influence of Urban Design Performance on Walkability in Cultural Heritage Sites of Isfahan, Iran. *Land* 2024, 13, 1523. <https://doi.org/10.3390/land13091523>.
4. Carmona, M.; Gabrieli, T.; Hickman, R.; Laopoulou, T.; Livingstone, N. Street Appeal: The Value of Street Improvements. *Prog. Plan.* 2018, 126, 1–51. <https://doi.org/10.1016/j.progress.2017.09.001>.
5. Qi, Z.; Li, J.; Yang, X.; et al. How the Characteristics of Street Color Affect Visitor Emotional Experience. *Comput. Urban Sci.* 2025, 5, 7. <https://doi.org/10.1007/s43762-025-00167-z>.
6. Harvey, C. *Measuring Streetscape Design for Livability Using Spatial Data and Methods*. Ph.D. Thesis, University of Vermont, Burlington, VT, USA, 2014.
7. Park, K.; Tian, G.; Larsen, S.S. Street Life and the Built Environment in an Auto-Oriented US Region. *Cities* 2019, 88, 243–251. <https://doi.org/10.1016/j.cities.2018.11.005>.
8. De Vos, J.; Lättman, K.; van der Vlugt, A.L.; Welsch, J.; Otsuka, N. Determinants and Effects of Perceived Walkability: A Literature Review, Conceptual Model and Research Agenda. *Transp. Rev.* 2022, 43, 303–324. <https://doi.org/10.1080/01441647.2022.2101072>.

9. Ewing, R.; Handy, S. Measuring the Unmeasurable: Urban Design Qualities Related to Walkability. *J. Urban Des.* 2009, 14, 65–84. <https://doi.org/10.1080/13574800802451155>.
10. Gehl, J. *Cities for People*; Island Press: Washington, DC, USA, 2013.
11. Elzeni, M.M.; ElMokadem, A.A.; Badawy, N.M. Impact of Urban Morphology on Pedestrians: A Review of Urban Approaches. *Cities* 2022, 129, 103840. <https://doi.org/10.1016/j.cities.2022.103840>.
12. Ibrahim, S.; Younes, A.; Abdel-Razek, S.A. Impact of Neighborhood Urban Morphologies on Walkability Using Spatial Multi-Criteria Analysis. *Urban Sci.* 2024, 8, 70. <https://doi.org/10.3390/urbansci8020070>.
13. Huang, X.; Zeng, L.; Liang, H.; et al. Comprehensive Walkability Assessment of Urban Pedestrian Environments Using Big Data and Deep Learning Techniques. *Sci. Rep.* 2024, 14, 26993. <https://doi.org/10.1038/s41598-024-78041-x>.
14. Song, L. The Street Space Planning and Design of Artificial Intelligence-Assisted Deep Learning Neural Network in the Internet of Things. *Heliyon* 2024, 10, e35031. <https://doi.org/10.1016/j.heliyon.2024.e35031>.
15. Ma, H.; Li, J.; Ye, X. Deep Learning Meets Urban Design: Assessing Streetscape Aesthetic and Design Quality through AI and Cluster Analysis. *Cities* 2025, 162, 105939. <https://doi.org/10.1016/j.cities.2025.105939>.
16. Nagata, S.; Nakaya, T.; Hanibuchi, T.; Amagasa, S.; Kikuchi, H.; Inoue, S. Objective Scoring of Streetscape Walkability Related to Leisure Walking: Statistical Modelling Approach with Semantic Segmentation of Google Street View Images. *Health Place* 2020, 66, 102428. <https://doi.org/10.1016/j.healthplace.2020.102428>.
17. Khotbehsara, E.M.; Yu, R.; Somasundaraswaran, K.; Askarizad, R.; Kolbe-Alexander, T. The Walkable Environment: A Systematic Review through the Lens of Space Syntax as an Integrated Approach. *Smart Sustain. Built Environ.* 2025. <https://doi.org/10.1108/SASBE-02-2024-0049>.
18. Ewing, R.; Cervero, R. Travel and the Built Environment: A Meta-Analysis. *J. Am. Plan. Assoc.* 2010, 76, 265–294. <https://doi.org/10.1080/01944361003766766>.
19. Hassan, G.F.; Rashed, R.; El Nagar, S.M. Regenerative Urban Heritage Model: Scoping Review of Paradigms' Progression. *Ain Shams Eng. J.* 2022, 13, 101652. <https://doi.org/10.1016/j.asej.2021.101652>.
20. Jacobs, J. *The Death and Life of Great American Cities*; Random House: New York, NY, USA, 1961.
21. Li, H.; Ikebe, K.; Kinoshita, T.; Chen, J.; Su, D.; Xie, J. How Heritage Promotes Social Cohesion: An Urban Survey from Nara City, Japan. *Cities* 2024, 149, 104985. <https://doi.org/10.1016/j.cities.2024.104985>.
22. Alnaim, M.M.; Albaqawy, G.; Bay, M.; Mesloub, A. The Impact of Generative Principles on the Traditional Islamic Built Environment: The Context of the Saudi Arabian Built Environment. *Ain Shams Eng. J.* 2023, 14, 101914. <https://doi.org/10.1016/j.asej.2022.101914>.
23. Wu, C.; Peng, N.; Ma, X.; Li, S.; Rao, J. Assessing Multi-Scale Visual Appearance Characteristics of Neighborhoods Using Geographically Weighted Principal Component Analysis in Shenzhen, China. *Comput. Environ. Urban Syst.* 2022, 92, 101732. <https://doi.org/10.1016/j.compenvurbsys.2022.101732>.
24. Zube, E.H. Cross-Disciplinary and Intermode Agreement in the Description and Evaluation of Landscape Resources. *Environ. Behav.* 1974, 6, 68–69. <https://doi.org/10.1177/001391657400600112>.
25. Kaplan, R.; Kaplan, S. *The Experience of Nature: A Psychological Perspective*; Cambridge University Press: Cambridge, UK, 1989.
26. Ewing, R.; Brownson, R.C.; Clemente, O.; Winston, E.; Handy, S. Identifying and Measuring Urban Design Qualities Related to Walkability. *J. Phys. Act. Health* 2006, 3, S223–S240. <https://doi.org/10.1123/jpah.3.s1.s223>.
27. Westerholt, R.; Acedo, A.; Naranjo-Zolotov, M. Exploring Sense of Place in Relation to Urban Facilities—Evidence from Lisbon. *Cities* 2022, 127, 103750. <https://doi.org/10.1016/j.cities.2022.103750>.
28. Burrage, H. Green Hubs, Social Inclusion and Community Engagement. *Proc. Inst. Civ. Eng. Munic. Eng.* 2011, 164, 167–174.
29. Ogawa, Y.; Oki, T.; Zhao, C.; Sekimoto, Y.; Shimizu, C. Evaluating the Subjective Perceptions of Streetscapes Using Street-View Images. *Landsc. Urban Plan.* 2024, 247, 105073. <https://doi.org/10.1016/j.landurbplan.2024.105073>.
30. Lynch, K. *The Image of the City*; MIT Press: Cambridge, MA, USA, 1960.

31. Ameli, S.H.; Hamidi, S.; Garfinkel-Castro, A.; Ewing, R. Do Better Urban Design Qualities Lead to More Walking in Salt Lake City, Utah? *J. Urban Des.* 2015, 20, 393–410. <https://doi.org/10.1080/13574809.2015.1041894>.
32. Yin, L.; Wang, Z. Measuring Visual Enclosure for Street Walkability Using Machine Learning Algorithms and Google Street View Imagery. *Appl. Geogr.* 2016, 76, 147–153. <https://doi.org/10.1016/j.apgeog.2016.09.024>.
33. Jeon, J.; Woo, A. Deep Learning Analysis of Street Panorama Images to Evaluate the Streetscape Walkability of Neighborhoods for Subsidized Families in Seoul, Korea. *Landsc. Urban Plan.* 2023, 230, 104631. <https://doi.org/10.1016/j.landurbplan.2022.104631>.
34. Hölscher, C.; Brösamle, M.; Vrachliotis, G. Challenges in Multilevel Wayfinding: A Case Study with the Space Syntax Technique. *Environ. Plan. B* 2012, 39, 63–82.
35. Fang, Y.-N.; Tian, J.; Namaiti, A.; Zhang, S.; Zeng, J.; Zhu, X. Visual Aesthetic Quality Assessment of the Streetscape from the Perspective of Landscape–Perception Coupling. *Environ. Impact Assess. Rev.* 2024, 106, 107535. <https://doi.org/10.1016/j.ear.2024.107535>.
36. Wilber, D.N. *Persian Gardens and Garden Pavilions*, 2nd ed.; Dumbarton Oaks: Washington, DC, USA, 1979.
37. Ewing, R.; Clemente, O. *Measuring Urban Design: Metrics for Livable Places*; Island Press: Washington, DC, USA, 2013.
38. Helbich, M.; Yao, Y.; Liu, Y.; Zhang, J.; Liu, P.; Wang, R. Using Deep Learning to Examine Street View Green and Blue Spaces and Their Associations with Mental Health. *Environ. Int.* 2019, 126, 136–145. <https://doi.org/10.1016/j.envint.2019.02.041>.
39. Mehrinejad Khotbehsara, E.; Somasundaraswaran, K.; Kolbe-Alexander, T.; Yu, R. The Influence of Spatial Configuration on Pedestrian Movement Behaviour in Commercial Streets of Low-Density Cities. *Ain Shams Eng. J.* 2025.
40. Dubey, A.; Naik, N.; Parikh, D.; Raskar, R.; Hidalgo, C.A. Deep Learning the City: Quantifying Urban Perception at a Global Scale. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2016.
41. Fan, Z.; Feng, C.; Biljecki, F. Coverage and Bias of Street View Imagery in Mapping the Urban Environment. *Comput. Environ. Urban Syst.* 2025, 117, 102253. <https://doi.org/10.1016/j.compenvurbsys.2025.102253>.
42. Zhou, B.; He, S.; Cai, Y.; Wang, M.; Su, S. Social Inequalities in Neighborhood Visual Walkability: Using Street View Imagery and Deep Learning. *Sustain. Cities Soc.* 2019, 50, 101605. <https://doi.org/10.1016/j.scs.2019.101605>.
43. Cantacuzino, S.; Bowne, K. Special Issue: Isfahan. *Archit. Rev.* 1976, 159, 255–321.
44. Nasar, J.L. *The Evaluative Image of the City*; Sage Publications: Thousand Oaks, CA, USA, 1998.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.