Proceedings of the 4th International Conference on Functional Materials and Civil Engineering
DOI: 10.54254/2755-2721/2026.37113

Intelligent Assessment of Living Street Spatial Quality in Older Residential Communities Based on Streetscape Semantic Segmentation

Ying Jiang

Department of Architecture, Southeast University, Nanjing, China
jiangying02442@outlook.com

Abstract. Old residential neighborhoods include living streets that serve pedestrians, parking, accessibility, informal gathering, and maintenance of the neighborhood. Traditional approaches to evaluate such streets are based on field observations and expert ratings, limiting the possibility to compare different places. In this paper, an intelligent approach is developed to assess the living street spatial quality through the use of Baidu Street View imagery and semantic segmentation. Xuhui District, Shanghai, is the studied area with 1,200 directional street-view images taken in 2024 and analyzed through a fine-tuned DeepLabV3+ model. From the semantic output of pixels, six indices are calculated: greenness visibility index, pedestrian space ratio, traffic interference, enclosure index, facilities support, and interface openness. Indices are standardized and aggregated through entropy weighting method, with visualizations created in ArcGIS Pro. The model attains 0.786 ± 0.041 Intersection-over-Union (IoU) score and 0.842 ± 0.036 macro-F1 score. Scores show clear spatial distinctions between internal streets, boundary streets, and entrance interfaces.

Keywords: streetscape semantic segmentation, older residential communities, living street, spatial quality assessment, urban renewal

1. Introduction

Older residential communities are a major focus of urban regeneration because they were built under outdated standards for mobility, infrastructure, and public space. In dense areas, living streets function not only as transport corridors but also as everyday spaces for walking, resting, shopping, socializing, and managing conflicts among pedestrians, bicycles, vehicles, and service activities. Their quality should therefore be assessed through spatial composition rather than road width or traffic efficiency alone.

Street view imagery provides an eye-level representation of the built environment and has been widely used with computer vision to evaluate urban vitality, public space, walkability, and environmental change [1-6]. Semantic segmentation is particularly useful because it converts visible streetscape elements into measurable pixel-level indicators.

DOI: 10.54254/2755-2721/2026.37113

Compared with formal roads and commercial streets, older residential communities remain understudied despite their narrow, fragmented, and visually complex environments. This study develops an intelligent assessment framework combining Baidu Street View, DeepLabV3+ semantic segmentation, six indicators, entropy-weighted scoring, and GIS visualization. The framework identifies deficiencies in greenery, pedestrian space, traffic conditions, facilities, and interface openness to support evidence-based community renewal.

2. Literature review

2.1. Street spatial quality assessment

Street spatial quality assessment has increasingly shifted from vehicle-oriented measures toward human-centered evaluation of walking environments. Key dimensions include visual comfort, accessibility, safety, enclosure, greenery, and support for public activity [7]. These factors are especially important in older residential communities, where streets serve diverse users simultaneously. A street may remain accessible to vehicles yet perform poorly for pedestrians when sidewalks are blocked or interfaces are closed. Eye-level street imagery therefore provides a practical basis for linking visible spatial features with residents' everyday experience.

2.2. Streetscape semantic segmentation

Semantic segmentation classifies streetscape pixels into categories such as roads, sidewalks, buildings, vegetation, vehicles, walls, fences, entrances, and signs. Compared with object detection, it is better suited to spatial-quality assessment because it measures visual proportions rather than object presence alone [8]. DeepLabV3+ is particularly suitable for older community streets because its multi-scale context extraction and boundary recovery improve recognition in dense scenes with overlapping sidewalks, vehicles, walls, entrances, and street furniture.

2.3. Intelligent assessment in community renewal

Intelligent assessment converts visual evidence into comparable indicators, improving the coverage, consistency, and spatial precision of community renewal analysis. GIS visualization further helps identify low-quality streets and prioritize interventions. However, street-view data may contain temporal and coverage bias, requiring control of sampling interval, image direction, and quality [9]. In practice, indicators can directly guide renewal actions, including planting, sidewalk repair, parking control, facility improvement, and entrance or frontage activation [10].

3. Method

3.1. Data source and study area

This paper deals with the older residential communities found in the Xuhui District of Shanghai. These communities mainly consist of the residential blocks built before 2000, which have dense street networks, aging infrastructures, mixed interface boundaries, and constant conflicts between pedestrians and vehicles. In this study, the streets included are those within the communities, boundary living streets, and short connecting streets that provide direct pedestrian connections and services.

DOI: 10.54254/2755-2721/2026.37113

The main sources of imagery data are the Baidu Street View images collected in 2024. First, the correction of the street network is done using ArcGIS Pro 3.2 software and then sampling points are set up at a distance of 50 meters apart. Then at each point, four directional images are taken at 0°, 90°, 180°, and 270° angles. After removing the blurred images, obscured images, nighttime images, and repetitive images, 1,200 valid images remain. Every image is associated with its longitude, latitude, direction, community name, street type, segment ID, and collection state.

3.2. Semantic segmentation model

The semantic segmentation model used is DeepLabV3+ with ResNet-101 backbone, which was implemented in Python 3.10, PyTorch 2.1, CUDA 11.8, OpenCV 4.8, MMSegmentation 1.2, NumPy, pandas, and Labelme 5.3. In the experiment, twelve classes were chosen: sky, building, road, sidewalk, vegetation, vehicle, pedestrian, wall, fence, entrance, street object, and sign. Local data with 300 images were collected and manually annotated to include such features as informal parking, narrow sidewalks, small gates, walls, and street objects of different kinds of service.

The images from the dataset are divided into 210 images for training, 45 for validation, and 45 for testing. The images are rescaled to have dimensions of 1024 × 512 pixels. Augmentation includes horizontal flips, brightness adjustment, scale adjustment within the interval of 0.75-1.25, and random cropping. Training is performed during 80 epochs with the batch size of 4, using the AdamW optimizer, initial learning rate equal to 1.0×10^{-4}, cosine decay, and weight decay of 0.01. To solve the problem of imbalanced data, cross-entropy loss is combined with Dice loss.

3.3. Spatial quality indicator system

The measurement system converts the segmentation mask into six quantitative variables. The green view is calculated based on the vegetation pixels; pedestrian spaces based on sidewalk pixels; traffic disturbance based on road and car pixels; enclosure based on building, wall, and fence pixels; support facilities based on street object and sign pixels; and interface openness based on entrance pixels and active frontage elements. Four directions of images at one sampling point are taken and averaged, and the resulting values are then summed up for the whole street segment. The standardized indicator matrix is calculated as shown in Equation (1):

$$z_{ij} = \begin{cases} \frac{x_{ij}-\min(x_j)}{\max(x_j)-\min(x_j)+\varepsilon}, j \in P \\ \frac{\max(x_j)-x_{ij}}{\max(x_j)-\min(x_j)+\varepsilon}, j \in N \end{cases} \tag{1}$$

where x_{ij} is the raw value of indicator j for segment i, P represents positive indicators, N represents negative indicators, and ε prevents division by zero. Enclosure is transformed according to its distance from the empirical median before normalization because both excessive openness and excessive closure may reduce spatial quality.

The final score is generated through entropy weighting as shown in Equation (2)-(5):

$$p_{ij} = \frac{z_{ij}}{\sum_{i=1}^{m} z_{ij}+\varepsilon} \tag{2}$$

DOI: 10.54254/2755-2721/2026.37113

$$e_j = -\frac{1}{\ln m} \sum_{i=1}^{m} p_{ij} \ln (p_{ij} + \varepsilon) \tag{3}$$

$$w_j = \frac{1-e_j}{\sum_{i=1}^{n}(1-e_j)} \tag{4}$$

$$S_i = \sum_{j=1}^{n} w_j z_{ij} \tag{5}$$

where w_j is the entropy weight and S_i is the final living street spatial quality score.

The six indicators were derived from semantic pixel counts. Greenness visibility measured vegetation pixels, pedestrian space ratio measured sidewalk pixels, and traffic interference measured vehicle presence and was treated negatively. Enclosure combined building, wall, and fence pixels, with quality defined by deviation from the sample median to account for both under- and over-enclosure. Facility support represented street-object and sign pixels, while interface openness measured entrance visibility. Four directional images were averaged at each point, then aggregated by street segment to reduce image-number bias.

4. Experimental process

4.1. Image collection and spatial matching

OpenStreetMap road data are imported into ArcGIS Pro and corrected against Baidu Map basemaps. Internal streets, boundary streets, and short access roads are retained, while unrelated roads are removed. Each segment receives a unique ID, and 50-meter sampling points are generated along the corrected network. At intersections, points are adjusted to avoid duplicate panoramas. Baidu Map API downloads four directional images for each point, and the metadata file records coordinates, viewing direction, road name, community name, segment ID, and download status.

4.2. Image preprocessing and segmentation inference

All valid images are resized to 1024 × 512 pixels and converted to RGB format. OpenCV is used for blur detection, exposure screening, format conversion, and duplicate detection. Images with severe occlusion, night conditions, or abnormal brightness are excluded or replaced by the nearest valid panorama within a 15-meter buffer. The fine-tuned model produces semantic masks, NumPy counts class pixels, and four directional results are averaged to reduce viewpoint bias. A 30-point field-photo set checks whether computed grades match visible conditions.

4.3. Indicator calculation and quality grading

After positive and negative indicators were normalized, the entropy method assigned larger weights to indicators with greater differentiation among street segments. For each indicator, the normalized proportion of every segment was calculated first, followed by information entropy, difference coefficient, and final weight. The spatial quality score was obtained by multiplying each normalized indicator by its entropy weight and summing the six weighted values. Therefore, a higher score

DOI: 10.54254/2755-2721/2026.37113

indicates a segment with stronger greenery, pedestrian space, facility support, and interface openness, lower traffic interference, and enclosure closer to the sample median. Natural breaks classification was applied to the final score by minimizing within-grade variance and maximizing between-grade variance, producing five data-based quality grades.

5. Results

5.1. Semantic segmentation accuracy of living street elements

The fine-tuned model performed consistently across major street elements. Compared with the original baseline, mean IoU increased by +0.092 ± 0.018*** and macro-F1 increased by +0.074 ± 0.016*** (see Table 1). Vegetation, road, building, and sidewalk were recognized most reliably, while entrance and sign classes were more affected by occlusion and shadow.

Table 1. Semantic segmentation performance on the test set

Class	IoU	Precision	Recall	F1-score
Vegetation	0.842 ± 0.029***	0.894 ± 0.024***	0.871 ± 0.031***	0.882 ± 0.026***
Sidewalk	0.781 ± 0.046***	0.846 ± 0.038***	0.817 ± 0.044***	0.831 ± 0.041***
Road	0.856 ± 0.025***	0.912 ± 0.019***	0.889 ± 0.027***	0.900 ± 0.022***
Vehicle	0.764 ± 0.052***	0.833 ± 0.047***	0.801 ± 0.050***	0.817 ± 0.048***
Wall/Fence	0.719 ± 0.057**	0.806 ± 0.049**	0.756 ± 0.055**	0.780 ± 0.052**
Entrance/Sign	0.652 ± 0.064**	0.741 ± 0.058**	0.706 ± 0.061**	0.723 ± 0.059**
Macro mean	0.786 ± 0.041***	0.839 ± 0.035***	0.813 ± 0.039***	0.842 ± 0.036***

5.2. Spatial distribution of living street quality grades

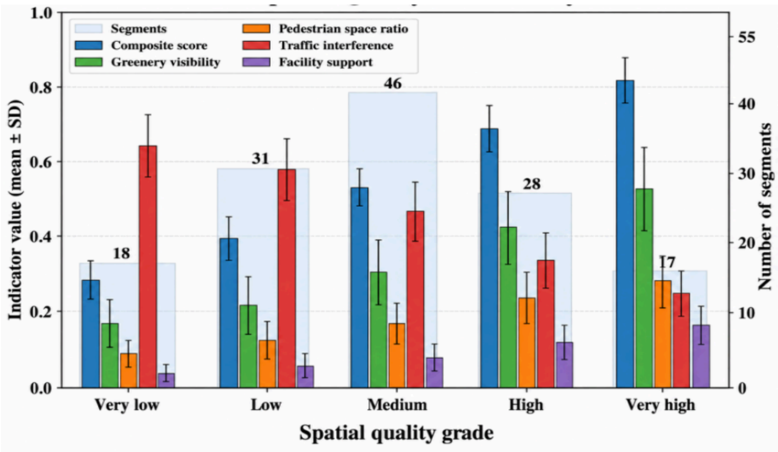

Figure 1. Spatial quality indicators by grade

The grade comparison reveals a consistent indicator combination rather than a single visible defect. From very low to very high quality, pedestrian space, greenery visibility, facility support, and interface openness increased, while traffic interference declined. High-grade entrance and mixed-service streets therefore received higher scores because several positive indicators improved

DOI: 10.54254/2755-2721/2026.37113

simultaneously. Low-grade internal streets combined sidewalk compression, vehicle occupation, weak facilities, and closed interfaces. The entropy-weighted result identifies how these conditions accumulate within the same segment, which cannot be established reliably through a general visual judgment alone (see Figure 1).

5.3. Identification of typical spatial quality problems

Three key problems are identified. First, informal parking compresses pedestrian space, with low-grade streets showing a sidewalk ratio of 0.128 ± 0.046 and high traffic interference. Second, enclosed streets show strong wall/fence exposure and low interface openness, indicating limited active frontage. Third, facility-deficient streets lack sufficient street objects and signage despite available pedestrian space. Accordingly, renewal should prioritize parking control and sidewalk recovery, frontage and entrance activation, and targeted improvements in seating, lighting, and signage.

6. Conclusion

This study develops an intelligent framework for assessing living street quality in older residential communities using semantic segmentation. Based on 1,200 Baidu Street View images from Shanghai's Xuhui District, DeepLabV3+ extracts six spatial indicators, while entropy-weighted scoring and GIS visualization identify deficiencies in greenery, pedestrian space, traffic conditions, interfaces, and facilities. The framework improves assessment consistency, coverage, and diagnostic precision, supporting targeted renewal. Its main limitation is the temporal and perceptual constraints of street-view imagery. Future work should integrate multi-period images, field observations, resident surveys, and post-renewal evaluation.

References

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

[2] Wang, M., & Vermeulen, F. (2021). Life between buildings from a street view image: What do big data analytics reveal about neighbourhood organisational vitality? *Urban Studies, 58*(15), 3118–3139.

[3] Chen, S., & Biljecki, F. (2023). Automatic assessment of public open spaces using street view imagery. *Cities, 137*, 104329.

[4] Fan, Z., Zhang, F., Loo, B. P. Y., & Ratti, C. (2023). Urban visual intelligence: Uncovering hidden city profiles with street view images. *Proceedings of the National Academy of Sciences of the United States of America, 120*, e2220417120.

[5] Li, Y., Yabuki, N., & Fukuda, T. (2022). Measuring visual walkability perception using panoramic street view images, virtual reality, and deep learning. *Sustainable Cities and Society, 86*, 104140.

[6] Zhou, Y., Thill, J.-C., Liu, X., Zhong, C., & Tu, W. (2025). Measuring and understanding changes in the physical built environment of cities with street view images. *Urban Informatics, 4*, 3.

[7] Guan, F., Fang, Z., Wang, L., Zhang, X., Zhong, H., & Huang, H. (2022). Modelling people's perceived scene complexity of real-world environments using street-view panoramas and open geodata. *ISPRS Journal of Photogrammetry and Remote Sensing, 186*, 315–331.

[8] Zhang, F., Salazar-Miranda, A., Duarte, F., Vale, L., Hack, G., Chen, M., Liu, Y., Batty, M., & Ratti, C. (2024). Urban visual intelligence: Studying cities with artificial intelligence and street-level imagery. *Annals of the American Association of Geographers, 114*(5), 876–897.

[9] Fan, Z., Feng, C.-C., & Biljecki, F. (2025). Coverage and bias of street view imagery in mapping the urban environment. *Computers, Environment and Urban Systems, 117*, 102253.

[10] Lu, Y., Ferranti, E. J. S., Chapman, L., & Pfrang, C. (2023). Assessing urban greenery by harvesting street view data: A review. *Urban Forestry & Urban Greening, 83*, 127917.