AGILE: GIScience Series, 7, 33, 2026. https://doi.org/10.5194/agile-giss-7-33-2026
Proceedings of the 29th AGILE Conference on Geographic Information Science, 16–19 June 2026.
Eds.: Evelyn Uuemaa, Alexander Kmoch, Holger Virro, and Simon Scheider
This contribution underwent peer review based on a full paper submission.

Measuring Urban Green Space Vitality through Multi-Source Visual and Textual Data: Integrating Social Media and Street-Level Imagery in London

Fangyuan Li[1]†, Rui Wang[2]†, Yijing Li[2], and Zahratu Shabrina[1]

[1] Department of Geography, King's College London, London, United Kingdom
[2] Department of Informatics, King's College London, London, United Kingdom

† These authors contributed equally to this work.
Correspondence: Yijing Li (yijing.li@kcl.ac.uk), Zahratu Shabrina (zara.shabrina@kcl.ac.uk)

Abstract. Urban green space (UGS) vitality reflects actual public engagement rather than mere spatial provision. However, limited research has examined how micro-scale environmental characteristics shape participation intensity. This study investigates the relationship between street-level environmental features and green space vitality using cross-platform social media data in London. Geotagged Flickr and Instagram posts were used as proxies for usage intensity. Environmental attributes were extracted from Google Street View imagery through semantic segmentation, including vegetation coverage, walkability, spatial enclosure, and facility elements. Ordinary Least Squares models were applied to assess associations between environmental variables and social media density. Results indicate that structural and facility-related features show stronger relationships with participation intensity than simple vegetation measures.

Submission Type. Model, Analysis

BoK Concepts. [GS4-3b] Citizens and volunteered geographic information, [IP3] Image understanding

Keywords. Urban Green Space; Social Media; Street-View Imagery; Semantic Segmentation; Scene Recognition

1 Introduction

Since 2008, the majority of the global population has lived in urban areas, and by 2050 this proportion is projected to reach 70-80%. Rapid urban expansion and increasing density have intensified concerns about residents' physical and mental health. In high-density environments, urban green spaces (UGS) are widely recognised for influencing health and quality of life through environmental and behavioural mechanisms.

Green spaces can mitigate air pollution, urban heat island effects, and traffic noise, thereby improving local environmental conditions. They also provide accessible settings for walking and informal recreation, supporting higher levels of physical activity. Exposure to natural visual stimuli has been shown to reduce physiological stress and promote attention restoration. In addition, as shared public spaces, green areas facilitate social interaction and strengthen social cohesion.

However, growing evidence suggests that not only the presence but also the vitality of green spaces is crucial for human well-being (Mouratidis & Poortinga, 2020). As Jacobs (1961) argued, urban vitality refers to the extent to which public spaces are actively used and animated by everyday life, rather than merely their physical provision. From this perspective, green space availability does not necessarily translate into actual use. Evaluating UGS solely in terms of physical supply or accessibility captures potential opportunities but cannot fully explain real patterns of engagement and social vibrancy. Greater attention is therefore needed to understand how spatial characteristics and accessibility are transformed into everyday use.

Previous research on actual green space use has primarily relied on traditional survey-based approaches, such as visitor counts at park entrances or questionnaire surveys, to measure recreational demand and examine preferences for specific facilities (Cohen et al., 2010; Dallimer et al., 2014; Eagles, 2014). However, these methods are often constrained by limited personnel capacity, high costs when conducted regularly or across large multi-jurisdictional areas, and sometimes limited spatial precision (Cessford & Muhar, 2003; Freeman, 2014).

In recent years, social media data have increasingly become an important source for studying green space use. These data offer several advantages, including high temporal resolution, low acquisition cost, and the ability to complement traditional survey-based methods. Early studies mainly focused on national park contexts, where platforms such as Flickr, Instagram, and Twitter were used to analyse visitation patterns in lakes and national parks. These studies demonstrated that social media data can serve as effective proxies for visitation intensity (Sessions et al., 2016; Tenkanen et al., 2017; Keeler et al., 2015).

Subsequently, this line of research has been extended to urban park settings. Existing studies suggest that social media data can also act as reliable proxies for urban park visitation and serve as a valuable source for understanding user experiences (Hampstead et al., 2018). As social media content is generated through individuals' spontaneous expressions, shared images and texts can reflect how people interact with urban green spaces (Dietz et al., 2025).

Building on these data, researchers can not only identify popular green space destinations but also examine recreational activities and their spatial patterns within green spaces (Chen et al., 2024).

In addition, Street View Imagery (SVI) has recently emerged as an important data source for assessing urban green spaces. Using computer vision techniques, SVI enables the extraction of environmental features and provides fine-scale information of green spaces. Compared with traditional methods, it offers a scalable and consistent approach for analysing UGS across cities.

This study aims to address several limitations in existing literature.

First, although previous research has used geotagged social media data to examine green space use, cross-platform studies remain relatively limited. As a result, it is difficult to fully capture differences in user preferences across platforms and the potential biases they introduce. To address this gap, this study conducts a comparative analysis of Flickr and Instagram, examining differences in user behavior and expression within green spaces. This approach improves the robustness of urban vitality measurement by accounting for platform-specific characteristics.

Second, while prior studies have explored factors influencing park use, the integration of multi-source data remains limited. Few studies have combined micro-scale environmental features extracted from SVI, socio-demographic characteristics, and social media behavioural data within a unified framework. This makes it difficult to comprehensively understand the relationship between environmental conditions and actual usage behaviour. This study develops an analytical framework linking objective environmental conditions with subjective behavioural responses, thereby providing a more comprehensive understanding of how objective and perceptual factors jointly shape urban green space vitality.

London is a highly dense global city, known for its complex urban structure, diverse land use, and extensive network of urban green spaces. Urban green spaces play a central role in its spatial structure and are crucial for improving environmental quality and promoting public health (Greater London Authority, 2023; Mouratidis & Poortinga, 2020).

This study focuses on the Greater London Area (GLA) and develops an integrated analytical framework that combines cross-platform social media data with scene semantic analysis to characterise the vitality of urban green space use. Specifically, Flickr images are first analysed using scene recognition and semantic clustering to extract activity-related scenes from user-generated content, thereby providing a systematic representation of activity structures within green spaces.

Furthermore, Google Street View imagery within green space boundaries is used to extract key built-environment features through semantic segmentation, and regression analysis is used to examine the relationship between these features and green space use intensity derived from Flickr and Instagram. This framework provides quantitative evidence for identifying key environmental factors that promote green space use and offers methodological support for evidence-based urban green space planning.

2 Methodologies

2.1 Social Media Data

Flickr photo metadata was collected, yielding 271,954 photographs taken within green space boundaries between 2014 and 2025. To characterise behavioural patterns across green space environments, we identified the activity contexts represented in these images. Scene recognition models classify images at the level of overall environment (e.g. 'urban park' or 'waterfront'), making them more suitable for inferring activity settings.

Accordingly, we applied a ResNet50 model to perform automated scene classification. The model achieved approximately 54% Top 1 accuracy and over 85% Top-5 accuracy on the Places365 validation set (Zhou et al.,

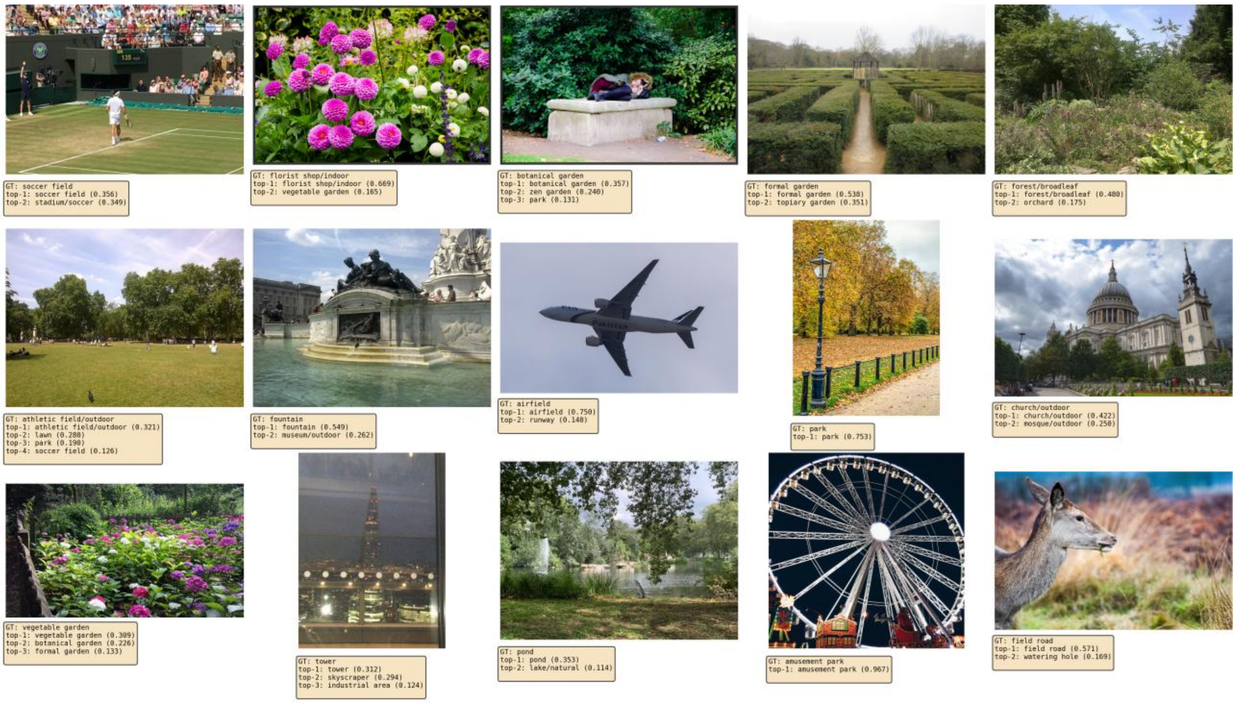


Figure 1. Flickr scene recognition results - example images.

2017), indicating robust scene-level performance. The resulting classifications enabled systematic identification of environmental contexts associated with green space activities (Fig. 1).

2.2 Street View-Based Measurement of the Built Environment

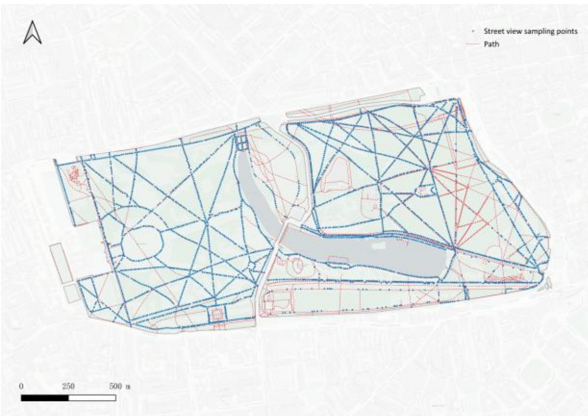


Figure 2. Illustration of Google Street View Sampling within Green Spaces (Hyde Park). Blue points represent sampling locations, and red lines indicate roads within the park.

For street-view images, we collected imagery located within the boundaries of open green spaces in London. Based on this dataset, we constructed street-view sampling points at 10-metre intervals along the OpenStreetMap (OSM) street network. Fig. 2 presents an example of the sampling points in Hyde Park. In total, we obtained 63,385 sampling points across green spaces in London, and for each point, we retrieved four directional street-view images. This design also helps reduce potential bias caused by visibility differences across environments.

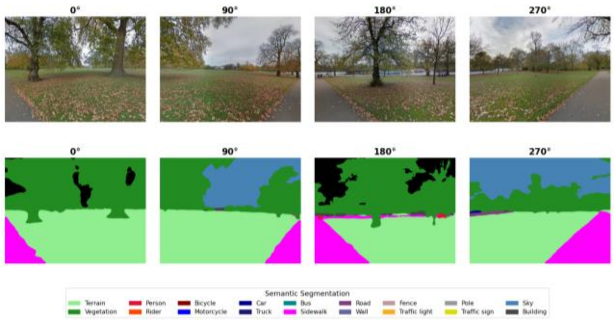


Figure 3. Segmentation Results at a Single Street View Sampling Point within a Green Space.

After obtaining the street-view images, we applied the Mask2Former panoptic segmentation model pre-trained on the ADE20K dataset to perform semantic segmentation, and the segmentation results are shown in Fig. 3. Mask2Former achieved approximately 57% mean Intersection over Union (mIoU) on the ADE20K semantic segmentation benchmark (Cheng et al., 2022). This model enables pixel-level semantic segmentation, enabling us to calculate the visible proportions of different environmental elements, such as trees, roads, and buildings. The calculation is expressed in Equation (1):

$$P_k = \frac{1}{N}\sum_{i=1}^{N} I(y_i = k), \quad k \in \mathcal{K} \qquad (1)$$

Where N represents the total number of pixels in the image, y_i denotes the class label of pixel i, and $\mathcal{K}$ represents the set of non-countable environmental elements.

2.3 Socio-economic Context and Control Variables

Urban green space visitation is influenced by broader socio-demographic structures (Jacobs, 1961; Gehl, 2010; Lefebvre, 1991). Therefore, to reduce potential bias in estimating the relationship between street-view environmental elements and green space vitality, this study incorporated socio-economic indicators derived from the 2021 UK Census at the LSOA (Lower Layer Super Output Area) level as control variables. These included age structure (proportions aged 0–19 and 60+), unemployment rate, annual total income, population density, and transport accessibility indicators (average Public Transport Accessibility Level (PTAL)). These variables were subsequently aggregated to the green space level to align with the analytical framework of the regression models (Fig. 4).

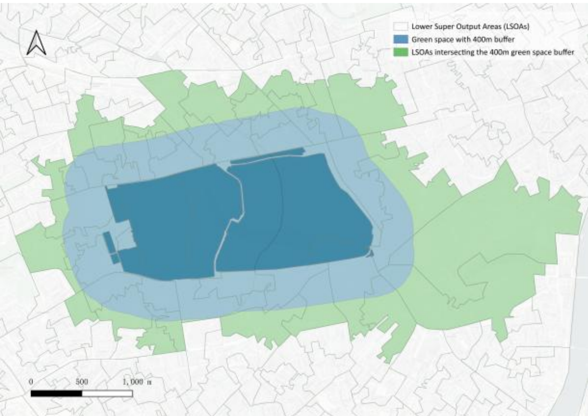


Figure 4. Area-weighted aggregation of LSOA-level socio-economic variables to the green space scale. A 400 m buffer (light blue) was generated around the green space (dark blue) and intersected with surrounding LSOAs (green). Socio-economic variables were aggregated based on the proportion of overlapping areas.

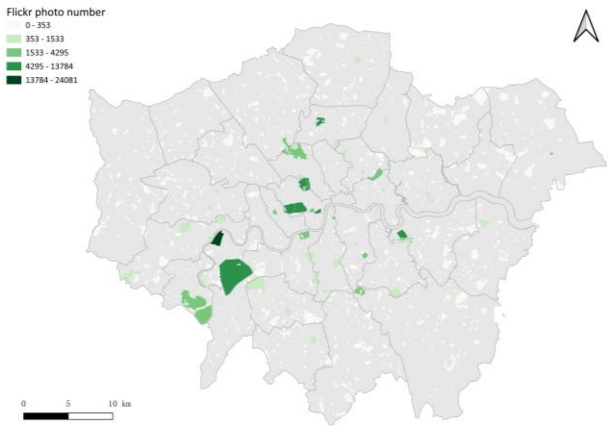
2.4 Data and Software Availability

The workflow of this study was implemented in Python. GeoPandas was used for spatial data processing, and PyTorch-based frameworks were employed for deep learning tasks.

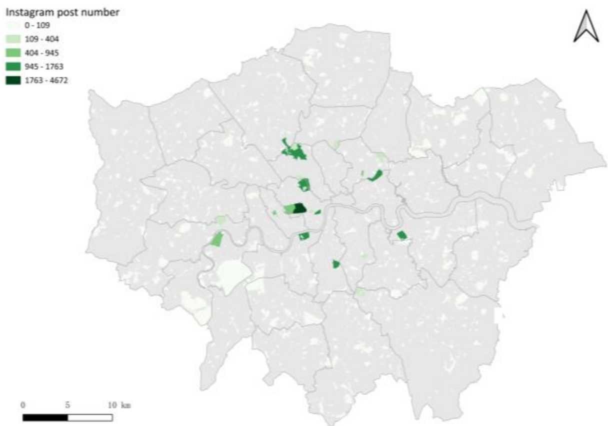
Green space boundaries were obtained from the Ordnance Survey (OS) Open Greenspace dataset, which is publicly available via the UK government open data portal. Socio-economic variables were derived from the 2021 UK Census and can be accessed through the Office for National Statistics (ONS). Flickr photo metadata were collected using the official Flickr API. Due to data access and privacy restrictions, the raw social media data cannot be publicly shared but can be reproduced through the API using equivalent spatial and temporal queries.

Google Street View (GSV) images were retrieved via the Google Street View Static API (https://developers.google.com/maps/documentation/streetview/overview), which requires an API key and is subject to usage limits defined by the Google Maps Platform.

For image analysis, a ResNet50 model pre-trained on the Places365 dataset was used for scene classification, and the Mask2Former model pre-trained on ADE20K was applied for panoptic segmentation. The implementation was based on PyTorch (version 2.0.1) with CUDA support.



(a)



(b)

Figure 4. Geographic distribution of Flickr photo number and Instagram post number. (a) Flickr; (b) Instagram.

3 Results

3.1 Spatiotemporal Differences in Green Space Popularity: Flickr vs Instagram

A mapping analysis of geotagged social media data in London's public green spaces reveal clear spatial and temporal heterogeneity between Flickr and Instagram. The Flickr dataset (2014-2025) shows broad spatial coverage across the city, with high-frequency observations concentrated not only in major Royal Parks in central London but also in secondary green spaces and peripheral nature reserves (Fig. 4(a)). In contrast, the Instagram dataset (January-June 2025) displays a markedly higher degree of spatial concentration. Activity points are clustered in a limited number of central green spaces, particularly in Zones 1 and 2, such as Hyde Park and Regent's Park, forming discrete high-density patches on the map (Fig. 4(b)).

3.2 Flickr-Based Scene Recognition and Clustering of Green Space Contexts

Figure 5. Semantic clustering results of Flickr photo activity scenes. Word clouds of six Flickr photo activity scene categories, with word size indicating frequency.

Based on the clustering results, Flickr photographs were grouped into six clusters. To aid interpretation of the activity contexts within each cluster, the associated scene categories were visualised as word clouds (Fig. 5).

Based on the semantic and spatial characteristics of dominant scenes in each word cloud, the clusters were labelled as follows: C0 Gardens & Natural Greenery; C1 Buildings & Indoor Spaces; C2 Water Features & Aquatic; C3 Public Venues & Religious Sites; C4 Towers & Urban Structures; and C5 Outdoor Activities & Sports. This classification enables multidimensional quantification of functional diversity within London's green spaces and supports further analysis of spatial preferences and behavioural patterns across different landscape settings.

3.3 Factors Associated with Green Space Popularity: A Comparative Regression Analysis of Flickr and Instagram

Table 1 presents the OLS regression results examining green space visitation in relation to demographic characteristics and street-view environmental elements. The Instagram model shows higher explanatory power ($R^2 = 0.623$) than the Flickr model ($R^2 = 0.420$).

To validate the appropriateness of the OLS regression, we performed Moran's I tests on the model residuals to check for potential spatial autocorrelation. Spatial weight matrices were constructed based on distance thresholds. The results indicated no significant spatial autocorrelation in the OLS residuals for either dataset (Flickr: $p = 0.365$; Instagram: $p = 0.280$). This demonstrates that the OLS models sufficiently accounted for the spatial structure, satisfying the assumption of residual independence.

Table 1. OLS regression model performance for Flickr and Instagram.

Metric	Flickr	Instagram
R^2	0.42	0.62
Adjusted R^2	0.41	0.61
AIC	-7670.91	-5146.12
Number of observations	1478	559
Significant variables ($p < 0.05$)	7	7

Fig. 6 and Table A1 report the OLS regression results incorporating demographic variables, green space characteristics, and street-view environmental indicators. Most population structure and socio-economic variables are not statistically significant in either platform model. The only exception is public transport accessibility (average Public Transport Accessibility Level, PTAL), which is positively and significantly associated with visitation in the Flickr model.

In contrast, the configuration of significant street-view environmental variables differs markedly between platforms. In the Flickr model, variables reflecting spatial structure and facility configuration - such as the number of sculptures, proportion of sidewalks, number of streetlights, wall coverage ratio, and water coverage ratio - are positively associated with visitation density. In the Instagram model, significant variables are more concentrated on perceptible landscape and usage features, including the number of sculptures and fountains, and the proportions of grass areas. Notably, in the Instagram model, the number of benches and the proportions of sky and tree canopy coverage show significant negative associations.

Overall, the results indicate that micro-scale environmental attributes captured from street-view

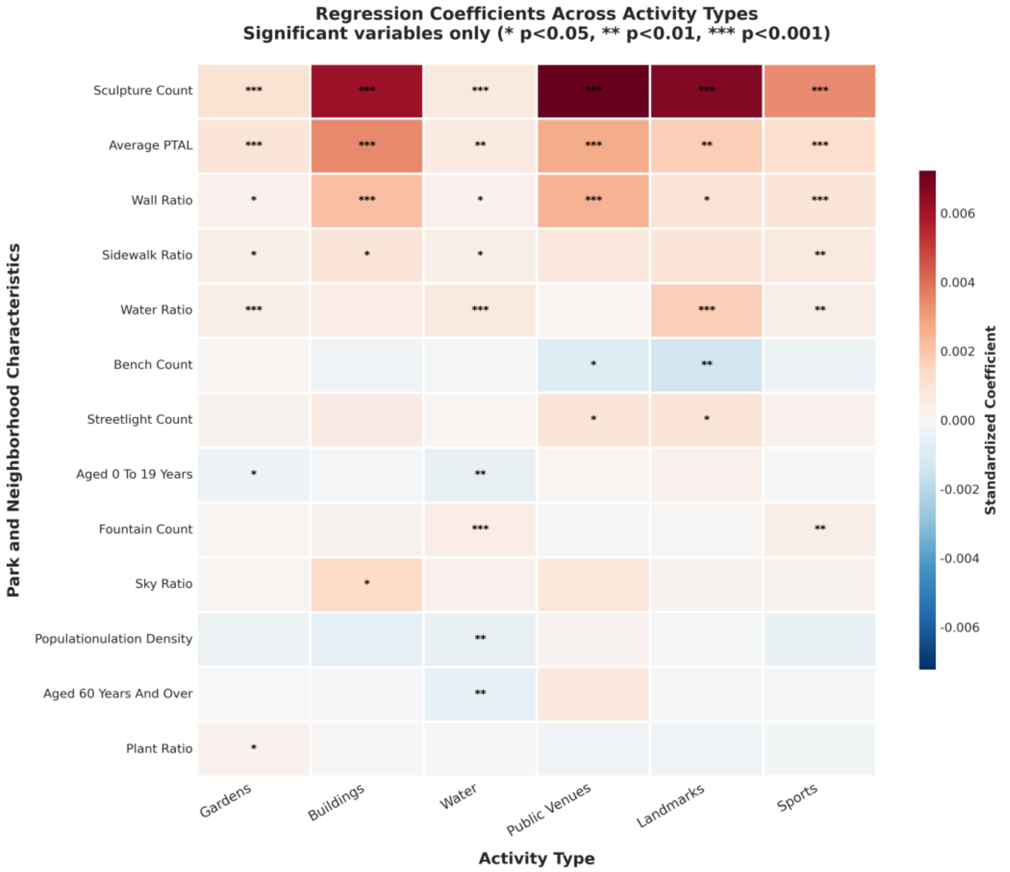


Figure 7. Heatmap of standardized OLS regression coefficients across six Flickr activity contexts (*p < 0.05, **p < 0.01, ***p < 0.001).

imagery play a more prominent role in shaping green space usage intensity than macro-level demographic factors or park size.

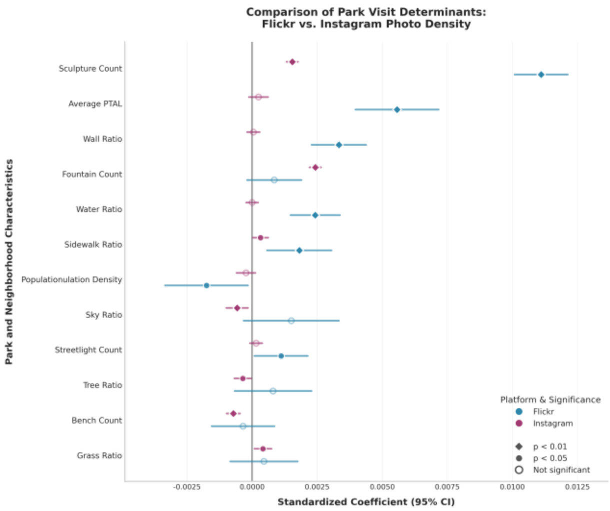


Figure 6. Standardized OLS coefficients for Flickr and Instagram models. Points show coefficient estimates with ±1.96 standard errors. Filled symbols indicate statistically significant effects ($p < 0.05$).

3.4 Factors Associated with Green Space Popularity Across Different Activity Contexts: A Comparative Regression Analysis

For different green space activity clusters, the results demonstrate clear variation in model performance (see Table A2). Clusters linked to built environments and public spaces (C1, C3, C4, and C5) display higher R^2 values, suggesting stronger explanatory power, whereas clusters dominated by natural landscapes (C0 and C2) exhibit comparatively lower model fit.

Fig. 7 presents a heatmap of regression coefficients showing the relationships between park visitation density and demographic and street-view environmental variables across the six activity contexts (C0-C5). Overall, sculpture count exhibits a significant positive association across all activity types and is the only environmental variable consistently significant in all six clusters. Average PTAL also shows a significant positive relationship in several categories, with particularly strong effects in Buildings & Indoor Spaces (C1) and Public Venues & Religious Sites (C3).

When examining the six Flickr photo activity contexts, scenes characterised by built environments and public spaces show stronger associations with structural and facility-related elements. Wall ratio is significantly and positively associated with visitation density in Buildings & Indoor Spaces (C1) and Public Venues & Religious Sites (C3). Streetlight count also shows a significant positive association in Public Venues & Religious Sites (C3) and Towers & Urban Structures (C4), although the effect size is modest. In contrast, bench count is significantly negatively associated with visitation density in these categories, with the strongest negative effect observed in Towers & Urban Structures (C4).

Activity contexts related to outdoor or natural environments display a different pattern, showing greater sensitivity to linear spatial elements and water-related features. Sidewalk ratio is positively associated with visitation density in C0, C1, C2, and C5, though the overall effect remains small. In water-related activity types (C0, C2, and C5), water ratio shows a significant positive association with visitation density. Similarly, fountain count is significantly positively associated in Water Features & Aquatic (C2) and Outdoor Activities & Sports (C5).

Notably, apart from plant ratio, which is positively associated with visitation density in Gardens & Natural Greenery (C0), typical natural green space indicators such as tree canopy coverage and grass ratio do not show significant associations with Flickr-based park visitation density.

4 Conclusion & Limitations

This study integrates multi-source social media data, street-view environmental indicators, and socio-economic variables within a unified analytical framework to examine how environmental characteristics influence the actual use of urban green spaces and the activities taking place within them.

We find that different social media platforms capture distinct behavioural dimensions. Flickr exhibits a broader spatial coverage, extending beyond central and iconic parks to include peripheral green spaces and nature reserves. This pattern may reflect Flickr users' preference for natural elements, indicating a greater tendency to document landscape quality and ecological value, including less prominent locations (Hausmann et al., 2018; Moreno et al., 2020). In contrast, Instagram activity is largely concentrated in a limited number of highly recognisable and symbolic central green spaces. This clustering pattern is likely related to platform-specific dissemination mechanisms, whereby visually appealing landmark spaces are more frequently photographed and shared, thereby reinforcing spatial hotspots (Boy & Uitermark, 2017).

The regression results further highlight differences in how environmental and socio-demographic factors relate to the activities observed on each platform. Although both models identify broadly similar significant factors, the Instagram model demonstrates higher explanatory power, indicating a stronger association between recorded activities and observable environmental characteristics. This may be because Instagram data are primarily drawn from more recent periods, meaning that the captured usage patterns are temporally closer to the street-view conditions and thus better reflect current use. In contrast, the decade-long Flickr dataset is more likely to represent the long-term accumulated patterns of green space use.

At the activity level, regression analyses across different scene types show that park environmental characteristics have stronger explanatory power for activities associated with built-environment features. These activities typically occur in green spaces with clearer functions and more defined spatial structures, such as areas containing buildings, plazas, or sports facilities. The coefficient heatmap further indicates that variables such as the number of sculptures, the proportion of walls, and the density of street lighting are positively associated with these scenes. Such elements tend to enhance the organisation and legibility of space: spatial enclosure helps define activity boundaries, while sculptures act as visual focal points that encourage people to stay, thereby increasing crowd concentration and activity intensity (Wu et al., 2023; Chen et al., 2024).

In contrast, park environmental characteristics show relatively weaker explanatory power for activities related to natural landscapes and water environments. These activities are more likely to reflect aesthetic experiences and psychological restoration, and their occurrence depends more on individual perception and external conditions, such as weather and seasonality. As a result, they are more difficult to explain consistently using environmental variables. Nevertheless, some structural features still play a role. For example, the proportion of pathways is one of the most strongly associated variables in nature- and water-related scenes. However, simple vegetation indicators, such as tree canopy coverage and grassland proportion, show relatively limited influence. This finding is consistent with Maas et al. (2008), who suggest that the mere presence of green space is not the primary driver of visitation; instead, landscape quality and walkability are more important determinants.

This study has several limitations that should be acknowledged. First, although previous research has shown that social media data are highly correlated with park visitation, they cannot be considered an exact proxy of actual use. Posting behaviour is influenced by multiple factors, including platform-specific user preferences, the availability of facilities, and aesthetic qualities, which may introduce bias when reflecting real usage patterns. In addition, social media users do not fully represent the general population, which may affect the generalisability of the findings.

Second, the relatively short temporal coverage of the Instagram dataset may emphasise recent trends and concentrated activity patterns, while being less capable of capturing long-term and stable patterns of green space use. Third, although the Places365 model performs well in recognising broad scene categories, the micro-scale or abstract compositions often found in Flickr images may introduce some uncertainty into the classification results.

Future research could combine social media data with on-site visitor statistics to further evaluate their validity as proxies for visitation. Incorporating longer-term datasets would also help capture more stable patterns of green space use. In addition, the adoption of more advanced visual recognition methods could improve the accuracy of environmental feature extraction and reduce potential misclassification.

Declaration of Generative AI in writing

The authors used Generative AI tools solely for language editing and improving the manuscript's grammar and structure; all scientific content, data analysis, and conclusions were developed independently by the authors without AI assistance.

Author Contribution.

Fangyuan Li: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing - original draft; Rui Wang: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing - original draft; Yijing Li: Conceptualization, Writing - review and editing, Supervision; Zahratu Shabrina: Conceptualization, Writing - review and editing, Supervision.

Acknowledgements.

This research was supported by the China Scholarship Council (CSC) from the Ministry of Education of P.R. China (CSC202408060082 & CSC202408060192).

References

Barros, C., Moya-Gómez, B., and García-Palomares, J. C.: Identifying temporal patterns of visitors to national parks through geotagged photographs, Sustainability, 11, 6983, 2019.

Boy, J. D. and Uitermark, J.: Reassembling the city through Instagram, Transactions of the Institute of British Geographers, 42, 612–624, 2017.

Cessford, G. and Muhar, A.: Monitoring options for visitor numbers in national parks and natural areas, Journal for Nature Conservation, 11, 240–250, 2003.

Chen, M., Cai, Y., Guo, S., Sun, R., Song, Y., and Shen, X.: Evaluating implied urban nature vitality in San Francisco: An interdisciplinary approach combining census data, street view images, and social media analysis, Urban forestry & urban greening, 95, 128289, 2024.

Cheng, B., Misra, I., Schwing, A. G., Kirillov, A., and Girdhar, R.: Masked-attention mask transformer for universal image segmentation, 1290–1299, 2022.

Cohen, D. A., Han, B., Derose, K. P., Williamson, S., Marsh, T., Rudick, J., and McKenzie, T. L.: The paradox of parks in low-income areas: Park use and perceived threats, American Journal of Preventive Medicine, 38, S423–S429, 2010.

Dallimer, M., Irvine, K. N., Anne, Davies, Z. G., Rouquette, J. R., Maltby, L. L., Warren, P. H., Armsworth, P. R., and Gaston, K. J.: Biodiversity and the feel-good factor: Understanding associations between self-reported human well-being and species richness, BioScience, 62, 47–55, 2012.

Dunkel, A.: Visualizing the perceived environment using crowdsourced photo geodata, Landscape and urban planning, 142, 173–186, 2015.

Eagles, P. F.: Research priorities in park tourism, Journal of Sustainable Tourism, 22, 528–549, 2014.

Gehl, J.: Cities for people, Island Press, 2010.

Giles-Corti, B., Bull, F., Knuiman, M., McCormack, G., Niel, V., Timperio, A., Christian, H., Foster, S., Divitini, M., and Middleton, N.: The influence of urban design on neighbourhood walking following residential relocation: longitudinal results from the RESIDE study, Social science & medicine, 77, 20–30, 2013.

Hamstead, Zoé A, Fisher, D., Ilieva, Rositsa T, Wood, S. A., McPhearson, T., and Kremer, P.: Geolocated social media as a rapid indicator of park visitation and equitable park access, Computers, Environment and Urban Systems, 72, 38–50, 2018.

Hausmann, A., Toivonen, T., Slotow, R., Tenkanen, H., Moilanen, A., Heikinheimo, V., and Enrico, D. M.: Social media data can be used to understand tourists' preferences for nature-based experiences in protected areas, Conservation Letters, 11, e12343, 2018.

Heikinheimo, V., Tenkanen, H., Bergroth, C., Järv, O., Hiippala, T., and Toivonen, T.: Understanding the use of urban green spaces from user-generated geographic information, Landscape and urban planning, 201, 103845, 2020.

Jacobs, J.: The death and life of great american cities, Random House, 1961.

Keeler, B. L., Wood, S. A., Polasky, S., Kling, C. L., Filstrup, C. T., and Downing, J. A.: Recreational demand for clean water: evidence from geotagged photographs by visitors to lakes, Frontiers in Ecology and the Environment, 13, 76–81, https://doi.org/10.1890/140124, 2015.

Lefebvre, H.: The production of space, Blackwell, 1991.

Maas, J., Verheij, R. A., Spreeuwenberg, P., and Groenewegen, P. P.: Physical activity as a possible mechanism behind the relationship between green space and health: a multilevel analysis, BMC public health, 8, 206, 2008.

Moreno-Llorca, R., Méndez, P. F., Ros-Candeira, A., Alcaraz-Segura, D., Santamaría, L., Ramos-Ridao, Ángel Fermín, Revilla, E., Bonet-García, F. J., and Vaz, A. S.: Evaluating tourist profiles and nature-based experiences in Biosphere Reserves using Flickr: Matches and mismatches between online social surveys and photo content analysis, Science of the Total Environment, 737, 140067, 2020.

Mouratidis, K. and Poortinga, W.: Built environment, urban vitality and social cohesion: Do vibrant neighborhoods foster strong communities?, Landscape and Urban Planning, 204, 103951, 2020.

Myrick, F. A.: The measurement of environmental and resource values, Resources for the Future. Washington DC, 2003.

Tenkanen, H., Enrico, D. M., Heikinheimo, V., Hausmann, A., Herbst, M., Kajala, L., and Toivonen, T.: Instagram, Flickr, or Twitter: Assessing the usability of social media data for visitor monitoring in protected areas, Scientific reports, 7, 17615, 2017.

Ulrich, R. S.: View through a window may influence recovery from surgery, Science, 224, 420–421, https://doi.org/10.1126/science.6143402, 1984.

Wartmann, F., Baer, M., Hegetschweiler, K., Fischer, C., Hunziker, M., and Purves, R.: Assessing the potential of social media for estimating recreational use of urban and peri-urban forests, Urban Forestry & Urban Greening, 64, 127261, 2021.

Wu, C., Ye, Y., Gao, F., and Ye, X.: Using street view images to examine the association between human perceptions of locale and urban vitality in Shenzhen, China, Sustainable Cities and Society, 88, 104291, 2023.

Zhou, B., Lapedriza, A., Khosla, A., Oliva, A., and Torralba, A.: Places: A 10 million image database for scene recognition, IEEE transactions on pattern analysis and machine intelligence, 40, 1452–1464, 2017.

Appendix

Table A1. OLS regression coefficients for Flickr and Instagram

Variable	Flickr β	Instagram β
Demographic & Park-related variables		
Aged 0–19	-0.001	0.000
Aged 60+	-0.000	-0.000
Average PTAL	0.006***	0.000
Park area (m²)	0.000	0.000
Population density	-0.002*	-0.000
Annual income (£)	-0.000	0.000
Unemployment rate	-0.001	0.000
Streetscape variables		
Bench count	-0.000	-0.001***
Earth ratio	0.000	0.000
Fence ratio	0.000	0.000
Fountain count	0.001	0.002***
Grass ratio	0.000	0.000*
Mountain ratio	-0.000	-0.000
Path ratio	-0.001	-0.000
Plant ratio	0.000	-0.000
Sculpture count	0.011***	0.002***
Sidewalk ratio	0.002**	0.000*
Sky ratio	0.002	-0.001**
Streetlight count	0.001*	0.000
Tree ratio	0.001	-0.000*
Wall ratio	0.003***	0.000
Water ratio	0.002***	0.000

Note: *p < 0.05; **p < 0.01; ***p < 0.001.

Table A2. OLS Regression Models Performance Across Green Space Activity Clusters

Cluster ID	Cluster Name	No. of Parks	R²
0	Gardens & Natural Greenery	1,160	0.214
1	Buildings & Indoor Spaces	1,135	0.436
2	Water Features & Aquatic	844	0.274

Cluster ID	Cluster Name	No. of Parks	R^2
3	Public Venues & Religious Sites	981	0.427
4	Vertical Landmarks & Urban Plazas	714	0.416
5	Outdoor Activities & Sports	1,090	0.398