ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume X-4/W7-2025
9th International Conference on Smart Data and Smart Cities (SDSC), 2–5 September 2025, Kashiwa, Japan

How Do Weather and Time of Day Affect Street Impression?

Keiki Harada[1], Takuya Oki[1]

[1] School of Environment and Society, Institute of Science Tokyo, 2-12-1 Ookayama, Meguro, Tokyo 152-8550, Japan - harada.k.15fd@m.isct.ac.jp, oki.t.e60f@m.isct.ac.jp

Keywords: Streetscape, Impression evaluation, Web questionnaire, Deep learning, Crowdsourcing, Regression analysis.

Abstract

Recent advancements in machine learning have enabled the modeling of people's perceptions of urban streets using large-scale image datasets such as Google Street View. However, such datasets are typically limited to images captured under clear daytime conditions, which constrain their ability to represent diverse environmental conditions in urban settings. This study aimed to quantitatively examine how weather and time of day influence the way people perceive streetscapes. Street images were collected by the author using a bicycle-mounted camera in four districts of Setagaya Ward, Tokyo, under various weather and lighting conditions. A web-based survey was conducted, in which participants evaluated these images along multiple dimensions, and a predictive model of street impressions was developed using the responses. It was found that: 1) the model identified streets where certain impression scores tended to be higher under nighttime or rainy conditions; 2) a regression-based factor analysis revealed that visual elements, weather, and time of day contributed significantly to impression evaluations; and 3) the characteristics of positively perceived streets varied depending on the environmental context. These findings provide a basis for considering the environmental conditions in streetscape evaluation studies and support context-aware evaluation and planning of streetscapes that reflect local environmental and social needs.

1. Introduction

1.1 Background

Conventional urban design has prioritized automobile convenience, but resultantly, problems such as traffic congestion, environmental pollution, and reduced pedestrian space have become apparent. By contrast, urban design that focuses on pedestrians and bicyclists has been attracting attention in recent years. During the said period, there has been a marked shift in the field of urban planning from automobile-centered design to human-centered design. This transformation was intended to create sustainable and livable urban environments, and many cities have introduced measures to prioritize pedestrians and bicyclists.

For instance, in Copenhagen, Denmark, the conversion of Strøget in the city center into a pedestrian-only street in 1962 is a particularly well-known example of successful urban development. Additionally, since the 1970s, Amsterdam has promoted a transportation policy that prioritizes bicycles and has developed approximately 767 km of bicycle paths. Consequently, bicycles have become the primary means of transportation for citizens, with bicycle use having increased by 40% since the 1990s. Bicycles are also an important element in the development of people-centered cities. In Japan, there is concern regarding the decline of urban vitality amid a declining and aging population. In this context, as one of the efforts to improve cities' attractiveness and create liveliness in town centers, the creation of *"Comfortable and Walkable"* towns is being promoted (Ministry of Land, Infrastructure, Transport and Tourism, 2021).

While physical design and policy have gained considerable attention, less focus has been placed on how external conditions—such as weather and time of day — affect the use and perception of urban spaces. However, these factors can influence people's behavior and experiences in cities and should therefore be carefully considered in the design and planning of urban spaces and streets.

1.2 Related Studies

Some studies from related fields have highlighted the importance of environmental conditions. For example, in mobility research,

	Google Street View		Mapillary		Ours	
Spatial Coverage	✓	Wide	×	Biased and not wide	○	Not wide but can access narrow streets
Temporal Coverage	×	Only daytime	✓	Flexible	✓	Flexible
Update Frequency	×	Every few years (Depends on areas)	✓	Flexible	✓	Flexible
Image Resolution	✓	High	×	Varied	✓	High and stable
Capture Interval	✓	More than 10 m	×	Depends on users	✓	Highly flexible
Shooting Height	×	2.1 m height	○	Varied	✓	1.5 m height
Cost/Effort	✓	Low for users	✓	Low for contributors	×	Moderate to high initially
Legends: Positive: ✓, Negative: ×, Neutral: ○						

Table 1. Comparison of Image Collection Method

https://doi.org/10.5194/isprs-annals-X-4-W7-2025-43-2025

Tin Tin et al. (2012) examined the impact of weather and seasonal changes on bicycle use in Auckland, New Zealand. In the field of environmental psychology, Huang and Wang (2018) showed that people’s emotional and perceptual responses to the same urban landscape can vary between day and night. These findings suggest that factors related to time and weather are essential in understanding how urban spaces function in practice.

In recent years, large-scale street image datasets such as Google Street View (GSV) and Mapillary have become widely available. This has led to a surge in research actively analyzing impressions of urban landscapes using these data sources (Biljecki and Ito, 2021). Such platforms are expected to be applicable in various fields, including urban planning. Advances in computer vision and machine learning are facilitating the evolution of techniques for the quantitative analysis of street images (Ogawa et al., 2024; Ye et al., 2019).

Most existing streetscape image datasets have several limitations. Table 1 compares the characteristics of the three street image collection methods in terms of the spatial coverage, temporal coverage, and viewpoint height.

First, most image datasets are biased toward images captured during sunny daylight hours. The question arises as to how people's impressions of streets change with the weather and time of day. These factors have not been adequately addressed in existing datasets and models.

The height and location from which the street images are captured are also important for planning human viewpoints. However, there is a lack of image datasets captured at an accurate pedestrian eye level or at an angle of view close to it. In GSV, the camera is mounted on top of the car, and the captured image only provides a view of the street from a position higher than the human viewpoint. In addition, because the camera is positioned near the center of the road, it deviates from the viewpoints of pedestrians and bicyclists, who usually pass close to the sides of the road. Ito et al. (2024) constructed an adversarial generative network technique that transforms street-view images into the viewpoints of pedestrians and bicyclists; however, the images may differ from those taken from the viewpoints of pedestrians or bicyclists.

Additionally, crowdsourcing platforms, such as Mapillary, may lead to variations in data consistency and accuracy owing to the dependence on the level of user participation and the quality of the equipment used to capture the data. This may not be conducive to a comprehensive analysis that considers various situations. With regard to time of day, several examples can be found that deal with this as a research topic. The use of nighttime street images to analyze urban lighting and activity has been explored (Fan and Biljecki, 2024). Studies have also generated nighttime images of a street from daytime street images by training pairs of daytime and nighttime street images (Ye et al., 2024). However, these generated nighttime images may differ from the actual street conditions at night. Moreover, to the best of our knowledge, no studies have analyzed weather in street space, and there remains room for research on how both weather and time of day affect people's impressions of the street environment.

1.3 Objectives

Based on previous background and related research, this study presents the following research questions:

RQ1: Can street image data collection methods that use bicycles compensate for the shortcomings of previous datasets and construct a new street image dataset that is closer to a pedestrian's perspective?
RQ2: Given that street conditions change from moment to moment and should not be uniquely defined, how do differences in time of day and weather conditions affect impressions of the same street?

This study seeks to achieve two main objectives:

1. To overcome the limitations of existing datasets that focus on daytime images by developing a bicycle-based street image collection method to acquire street images of the same street in multiple weather conditions and time periods.

2. To quantitatively analyze the effects of time of day and weather conditions on the evaluation of street impressions.

We propose a new method for collecting street-image data using bicycles; utilizing bicycles helps obtain street images more efficiently than pedestrians and closer to the pedestrian's line of sight. Additionally, because it is easy to capture pictures, various scenes can be considered. In this study, the same street was driven several times under different weather conditions and at different times of the day, while changes in impression evaluation on the same street were analyzed.

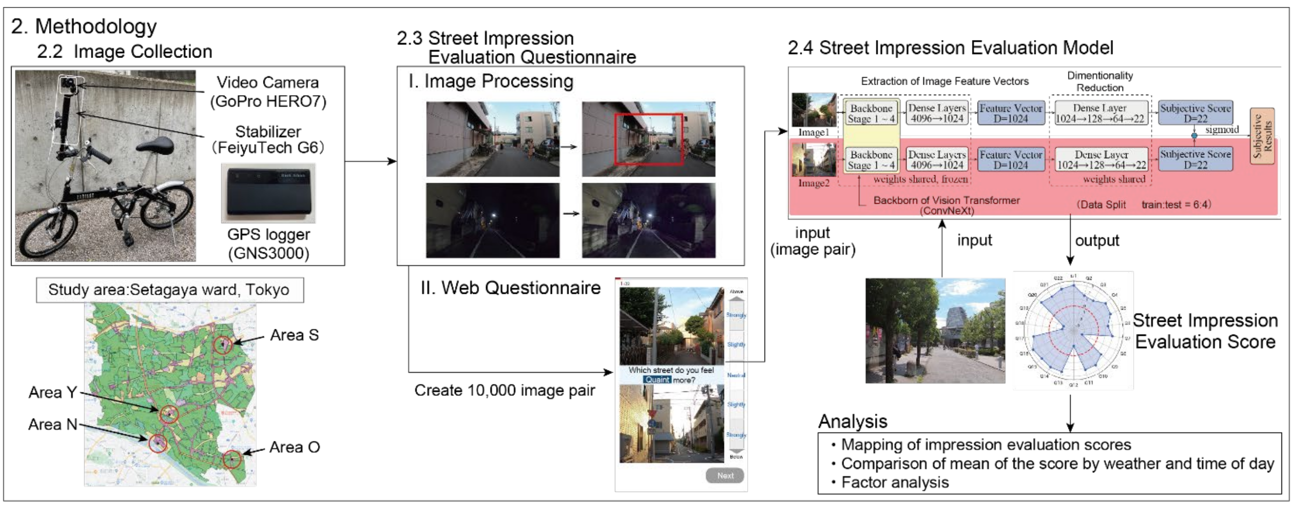

Figure 1. Research Framework

2. Methodology

2.1 Research Framework

Figure 1 illustrates the overall framework of this study, from data collection and the street impression evaluation questionnaire, to the training and evaluation of the model using the results of the questionnaire. Street images were collected using a bicycle-mounted camera under various weather and temporal conditions. The images were analyzed using a deep learning model, and subjective impressions were collected through a pairwise comparison survey. Each step of this framework is explained in the following sections.

2.2 Image Collection Method

In this study, the streetscape images were collected by one of the authors. A bicycle equipped with a video camera and GPS logger was used to record the streets. Figure 2 shows the equipment used for image collection, and Table 2 lists the details of the image collection method.

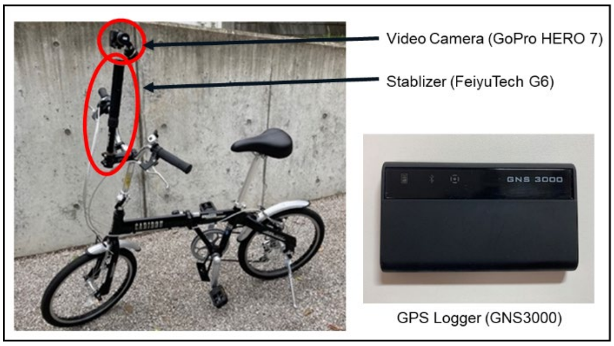

Figure 2. Equipment Used for Street Filming

Period	From 11th of August 2022 to 7th of October 2022
Shooting Area	Four areas in Setagaya ward, Tokyo, Japan
Shooting Position	Drove on the left side of the roadway or on the sidewalk Camera height: 1.5 m above ground level
Information of Camera	Mode: Wide-angle mode Resolution: 1920 x 1080 [pixel]
Weather	Sunny, Cloudy, Rainy
Time of the Day	Midday, Evening, Night

Table 2. Details of Image Collection

Four areas were selected based on "land-use zones," with each area reflecting different zoning characteristics within urban environments: Area O is predominantly residential; Area S includes a mix of commercial and residential zones; Area N is primarily commercial; and Area Y comprises residential and quasi-industrial zones. In Japan, there exist 13 types of zoning districts, including residential, commercial, and industrial zones, that define the general framework of land use in urban areas, and the types of buildings that can be built are defined according to the purpose of each district (Ministry of Land, Infrastructure, Transport and Tourism, 2003). Of the 13 types, 10 zoning districts exist in Setagaya Ward, Tokyo, and all are included in the selected areas. The details of these zones are shown in Figure 3.

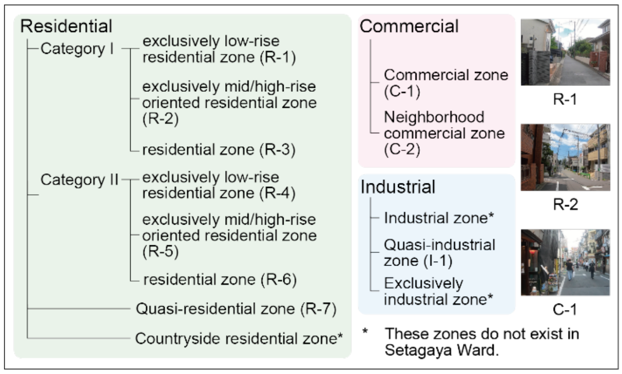

Figure 3. Overview of Land-Use Zones (Ministry of Land, Infrastructure, Transport and Tourism, 2003)

Street videos were recorded in each area for nine scenes under three weather conditions and three time periods.

2.3 Street Impression Evaluation Questionnaire

A web-based impression evaluation questionnaire was conducted using street images to obtain information on the subjective impressions of streetscapes. This survey was administered by a private research company and included a diverse age group of respondents from across Japan, which was considered appropriate for the study's objectives. Participants were compensated based on the number of image pairs they evaluated. Responses completed too quickly or containing uniform evaluation scores were excluded from compensation to ensure data quality. In addition, the survey was conducted with the approval of the university's research ethics review committee to which the authors are affiliated (Approval No.: 2022212).

The questionnaire method presented two images, as shown in Figure 4, and asked respondents to rate their impressions of each image on a 5-point scale for the 22 questions presented in Table 3.

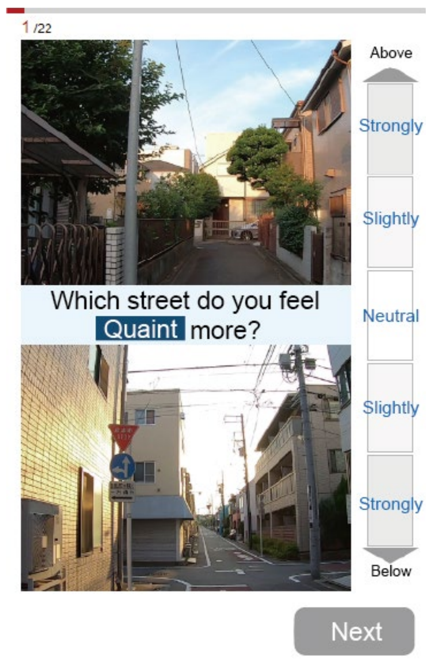

Figure 4. Example of the Questionnaire Form Screen

https://doi.org/10.5194/isprs-annals-X-4-W7-2025-43-2025

Q1	Open	Q12	Lived in feel
Q2	Familiar	Q13	Cozy
Q3	Lively	Q14	Walkable
Q4	Comfortable	Q15	Clean
Q5	Greenery	Q16	Beautiful
Q6	Calm	Q17	Wealthy
Q7	Bright	Q18	Boring
Q8	Old-fashioned	Q19	Anxious
Q9	Quaint	Q20	Favorite
Q10	Safe	Q21	Interesting
Q11	Neat	Q22	Willingness to pass

Table 3. Street Impression Evaluation Items

When selecting the 22 evaluation items, this paper considered the need to ensure generalizability and validity while minimizing the burden on survey respondents. A comprehensive review of previous studies on street impression evaluation was conducted, and frequently targeted items were extracted. Many items aligned with those used by Oki and Kizawa (2021), but items "Q9. Quaint," "Q14. Walkable," and "Q19. Anxious" were unique to this paper. No specific explanations were provided for the 22 items in the questionnaire, and respondents answered based on their own interpretations.

From the 60,041 images, we extracted 1,546 images and constructed 10,000 image pairs using the procedure shown in Figure 5. These pairs were used in the street impression evaluation questionnaire, with ten respondents assigned to each pair based on budgetary constraints. To ensure a diverse range of responses, each image pair was evaluated by two respondents from each of the five age groups (20s, 30s, 40s, 50s, and 60s), resulting in ten respondents per pair —one from each segment. These responses were then used for majority voting to evaluate image pairs, as described in Section 2.4. Since different users evaluated each image pair, comparisons between individual evaluators were not a concern. The total number of responses was balanced across age groups, with 20,000 responses collected per group.

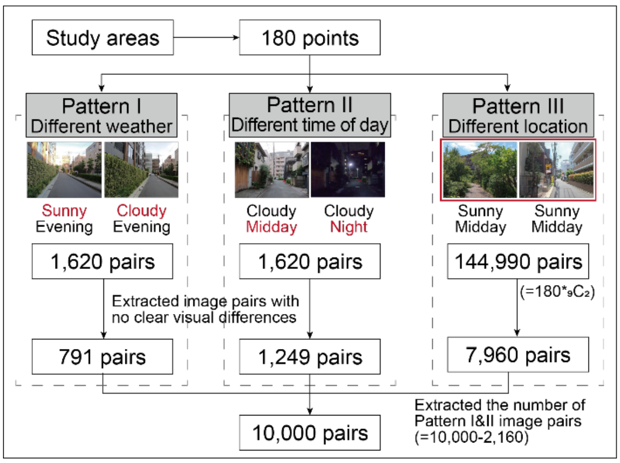

Figure 5. Selection Flow of the Pattern and the Number of Image Pairs Used for Questionnaire

Subsequently, image processing was performed as shown in Figure 6. Street images were cropped to remove distortions caused by shooting in wide-angle mode and to simulate a focal length of approximately 27.7 mm based on the diagonal crop factor applied to the original 16 mm images. Blurring was applied to the faces of passersby and car license plates. For nighttime images, we additionally corrected the brightness using contrast-limited adaptive histogram equalization (CLAHE) (Zuiderveld, 1994) to prevent the images from being too dark for the objects in the questionnaire to be identified.

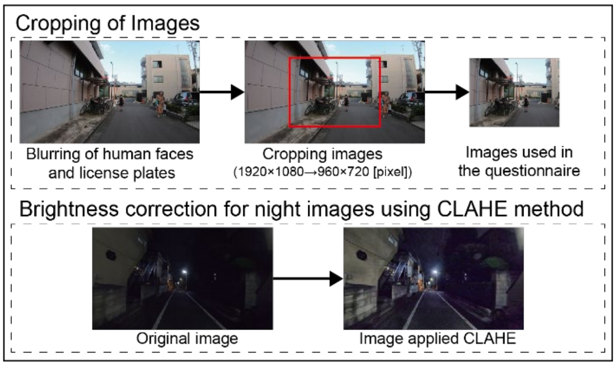

Figure 6. How to Trim Streetscape Images and Correct the Brightness of the Nighttime Images

2.4 Street Impression Evaluation Model

A street impression evaluation model was developed using the questionnaire results as training data. Figure 7 shows the structure of the model. This model is based on that proposed by Ogawa et al. (2024). This model enabled the simultaneous estimation of 22 street impression evaluation scores by inputting any two street images. The 10,000 image pairs used in the survey were divided into training data (60%), validation data (20%), and test data (20%) for model training.

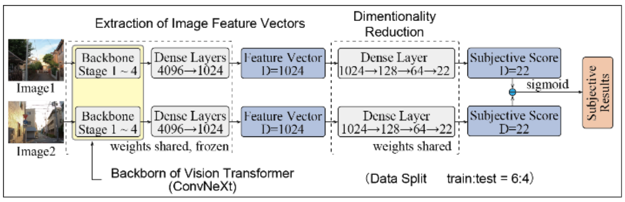

Figure 7. Structure of Street Impression Evaluation Model

First, feature vectors were calculated to quantitatively evaluate street images. In this study, ConvNeXt (Liu et al., 2022) was used as the backbone to extract 1,024-dimensional feature vectors from the images. Subsequently, dimensionality reduction was applied, and a sigmoid function was used to calculate provisional probabilities indicating which of the two images, Image1 or Image2, receives a higher evaluation for each of the 22 items.

The obtained probability values were compared with the survey results (a binary variable where the label was set to 1 if Image1 was rated higher by the majority of 10 respondents, and set to 0 if Image2 was rated higher). The model was trained to minimize the loss function (binary cross-entropy) using the Adam optimizer. The loss function was defined in Equation (1):

$$C_{ij} = -\bar{P}_{ij}\log P_{ij} - (1-\bar{P}_{ij})\log(1-P_{ij}) \qquad (1)$$

where $\bar{P}_{ij}$ = probability that the impression evaluation will be $i > j$ based on the questionnaire
P_{ij} = probability by applying sigmoid function to the difference between i and j

The "Subjective Score" calculated by the trained model, standardized based on the training images, was utilized in this study as the "Street Impression Evaluation Score."

2.5 Factor Analysis Method

It is difficult to quantitatively analyze which street components influence impression evaluation if the analysis is based solely on street impression evaluation scores. Therefore, factor analysis was conducted to determine the relationship between street impression evaluation scores and street components, thus enabling a statistical understanding of the street components that contribute to impression evaluation.

2.5.1 Semantic Segmentation: Semantic segmentation is a deep-learning task that makes sense of landscape elements in street images on a pixel-by-pixel basis. Semantic segmentation was performed using the Mask2Former model (Cheng et al., 2022), which was pretrained on the Mapillary Vistas Dataset (Neuhold et al., 2017).

2.5.2 LASSO (Least Absolute Shrinkage and Selection Operator) Regression Analysis: LASSO regression analysis (Tibshirani, 1996) was conducted using the following as explanatory variables: land-use zones, weather, time of day (dummy variables for each), and the percentage of street components in the image obtained by semantic segmentation; additionally, the street impression evaluation score was employed as the objective variable. LASSO regression analysis is an approach whereby the absolute value of the partial regression coefficient is used as the penalty term to prevent overlearning owing to a large coefficient in the linear regression. The parameters were determined to minimize the objective function E_{LASSO} in Equation (2) for each evaluation item. Specifically, the regularization parameter λ was determined through grid search using 5-fold cross-validation, with an exponential step of 0.1 over a logarithmic scale ranging from 10^{-6} to 1.

$$E_{LASSO} = \frac{1}{2}\sum_{1\le i\le n}\left\{y_i - \left(\sum_{1\le j\le m} a_j x_{ij} + b\right)\right\}^2 + \lambda \sum_{1\le j\le m} |a_j|, \quad (2)$$

where n = number of images used for factor analysis
m = number of explanatory variables
y_i = streetscape impression evaluation score for the i-th image
x_{ij} = value of the j-th explanatory variable for the i-th image
a_j = regression coefficient for the j-th variable
b = intercept
λ = hyperparameter

In this paper, interactions between explanatory variables were not considered to avoid the overcomplexity of the model.

3. Results

3.1 Evaluation of the Street Impression Evaluation Model

Table 4 presents the number of respondents by attribute. As a result of the questionnaire, 16,046 unique respondents representing a wide age range were recruited. The relatively small number of male respondents in their 20s and female respondents in their 60s indicates that the number of pairs or responses per person was high for these attributes. Since this study did not focus on comparisons between genders or age groups, such bias was unlikely to significantly affect the results or the overall significance of this research. The average number of pairs of responses per respondent was 6.2 (maximum, 90; minimum, 1).

Age	20s	30s	40s	50s	60s	Total
Male	480	1,368	2,203	2,219	1,472	7,742
Female	1,431	1,964	2,312	2,055	542	8,304
Total	1,911	3,332	4,515	4,274	2,014	16,046

Table 4. Number of Respondents by Gender and Age Group

Figure 8 shows the F-values of the street impression evaluation model. The average F-value for the test data is 86.0 %.

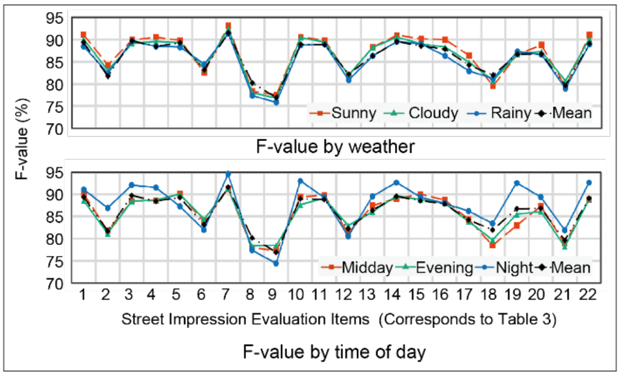

Figure 8. F-value of the Street Impression Evaluation Model

3.2 Examples of the Estimated Impression Evaluation Score

Figure 9 presents examples of the impression evaluation scores estimated using this model. The upper two images show differences in the street impression evaluation scores for many of the impression evaluation items, with the sunny daytime image having the highest impression evaluation score. Conversely, the impression evaluation scores for the lower two images show that the impression evaluation score is higher for rainy night in “Q21. Interesting.”

3.3 Comparison of Ratings based on Weather and Time of Day Combinations

The difference between the average score of street impressions in other weather and time periods was calculated for each zoning district based on the average score of street impressions in sunny weather and daytime. The results are shown in Figure 10. In Category 1 Exclusively Low-Rise Residential Zone, the scores for sunny and daytime were basically high, but in Commercial Zone, the scores for “Q3. Lively,” “Q9. Quaint,” and “Q21. Interesting” were higher at night, and “Q6. Calm” was higher in the rain.

3.4 Mapping of the Impression Evaluation Scores

All street images were input into the trained street impression evaluation model to predict the street impression evaluation score for the entire route. Figure 11 illustrates the results plotted on a map for the two street impression evaluation items, two regions, and two scenes for each item.

The mapping of “Q9. Quaint” shows that the circled area had a higher rating in sunny/midday, which is consistent with streets where “Q5. Greenery” was highly rated, but “Q3. Lively” was rated relatively low. However, the spatial distribution of these streets during the rainy and nighttime hours showed a decrease in

https://doi.org/10.5194/isprs-annals-X-4-W7-2025-43-2025

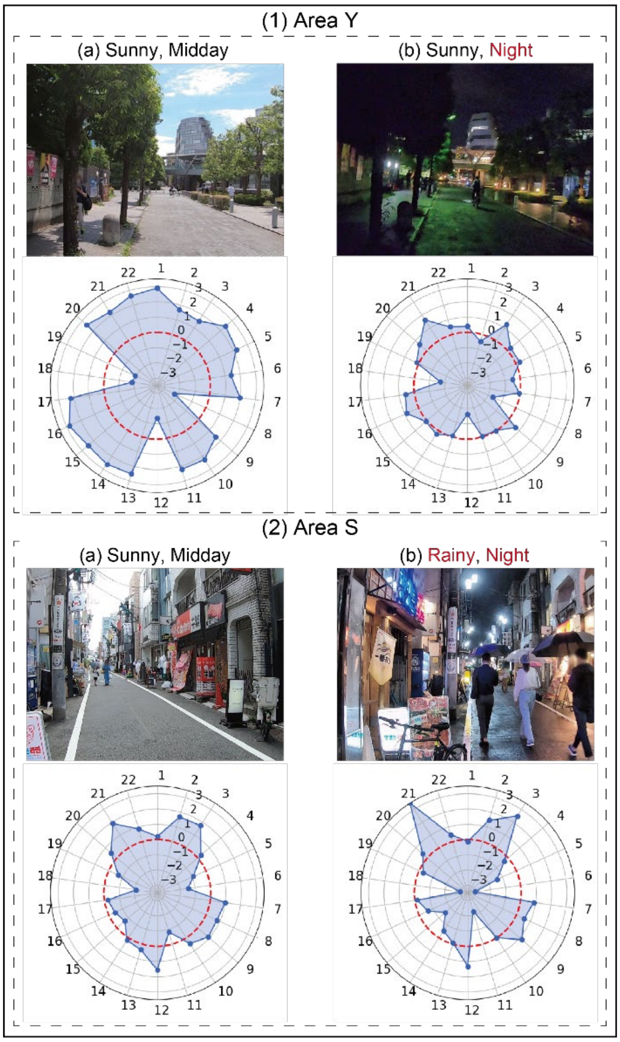

Figure 9. Examples of Impression Evaluation Scores Predicted by Trained Model (Comparison of the Same Point)

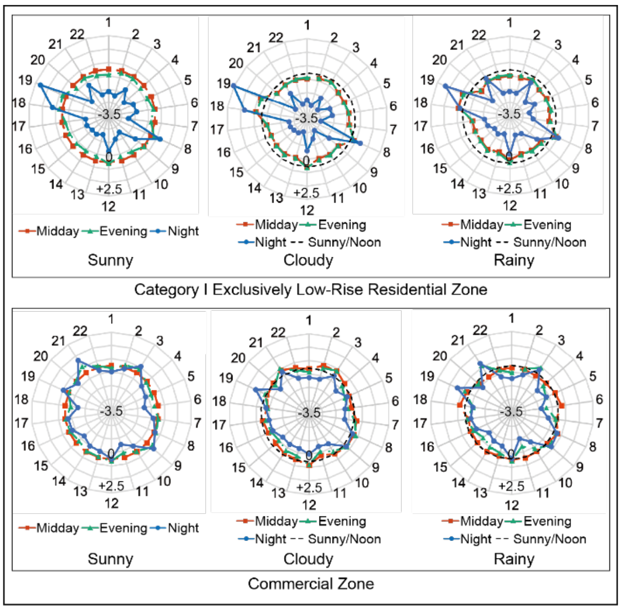

Figure 10. Difference in Street Impression Evaluation Scores (mean) based on Sunny and Daytime

the evaluation of the said streets and an increase in the evaluation of the streets around commercial areas.

The mapping of “Q19. Anxiety” indicates that the sense of anxiety increased during the night. Streets with relatively low anxiety levels at night were those in front of train stations, shopping streets, and streets with many cars. Additionally, although numerous points on a given street may be rated as highly anxiety-provoking, specific points along the same street were observed where the level of perceived uneasiness was notably lower.

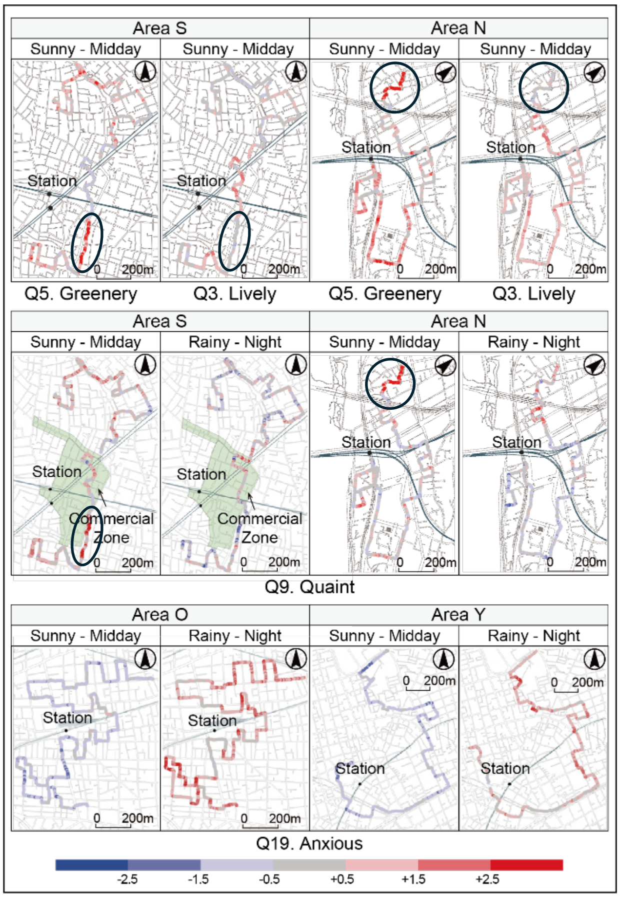

Figure 11. Mapping of the Street Impression Evaluation Scores

3.5 Relationship between Street Impression Scores and Streetscape Components

Table 5 shows the results of the LASSO regression analysis, which was conducted to analyze the relationship between streetscape impression evaluation scores and streetscape components. For each of the 22 impression evaluation items, the coefficients of determination and top five variables with the highest absolute values among the standard partial regression coefficients are shown.

The main components of the street, such as “Vegetation,” “Buildings,” or “Sidewalk,” appeared in all items, but “Night” and “Rainy” also ranked high in some items. For example, Night (-0.449) and Rainy (-0.129) negatively impacted “Q21. Interesting,” thus showing that “Banners” (i.e., signboards, etc.) and “Person” had a positive influence. “Building” and “Vegetation” positively influenced the evaluation in “Q10. Safe.” The influence of the surrounding environment was relatively small in “Q21. Anxious.”

Q1. Open		Q2. Familiar		Q3. Lively		Q4. Comfortable		Q5. Greenery		Q6. Calm	
R2 0.412	λ 3.2×10^{-4}	R2 0.367	λ 5.0×10^{-3}	R2 0.325	λ 2.5×10^{-3}	R2 0.416	λ 1.3×10^{-4}	R2 0.582	λ 5.0×10^{-4}	R2 0.432	λ 1.3×10^{-4}
Variable	Coef.	Variable	Coef.	Variable	Coef.	Variable	Coef.	Variable	Coef.	Variable	Coef.
Night	-0.437	Night	-0.427	Night	-0.214	Night	-0.613	Vegetation	0.613	Vegetation	0.733
Road	0.330	Fence	-0.159	Banner	0.181	Vegetation	0.410	Night	-0.294	Building	0.484
Sidewalk	0.320	Rainy	-0.114	Sidewalk	0.170	Road	0.358	Rainy	-0.072	Road	0.445
Sky	0.231	Banner	0.106	R-1*	-0.135	Sidewalk	0.347	R-1*	-0.063	Night	-0.322
R-1*	-0.225	Bridge	-0.100	Wall	-0.123	Building	0.303	Cloudy	-0.061	Sky	0.313
Q7. Bright		Q8. Old-fashoned		Q9. Quaint		Q10. Safe		Q11. Neat		Q12. Lived in feel	
R2 0.408	λ 5.0×10^{-3}	R2 0.295	λ 4.0×10^{-3}	R2 0.301	λ 5.0×10^{-3}	R2 0.423	λ 1.3×10^{-4}	R2 0.435	λ 2.0×10^{-5}	R2 0.467	λ 5.0×10^{-3}
Variable	Coef.	Variable	Coef.	Variable	Coef.	Variable	Coef.	Variable	Coef.	Variable	Coef.
Night	-0.497	Sidewalk	0.246	Vegetation	0.420	Night	-0.524	Road	0.567	Night	-0.207
Sidewalk	0.180	R-1*	0.222	Fence	-0.119	Sidewalk	0.318	Vegetation	0.549	Vegetation	-0.207
Road	0.105	Wall	0.148	Bridge	-0.114	Road	0.305	Night	-0.464	Building	0.200
Rainy	-0.099	Night	0.146	Banner	0.114	Vegetation	0.273	Building	0.449	R-1*	0.172
Wall	-0.098	R-6*	-0.122	Car	-0.109	Building	0.239	Sidewalk	0.447	R-6*	-0.141
Q13. Cozy		Q14. Walkable		Q15. Clean		Q16. Beautiful		Q17. Wealthy		Q18. Boring	
R2 0.401	λ 1.3×10^{-4}	R2 0.413	λ 1.3×10^{-4}	R2 0.400	λ 1.3×10^{-4}	R2 0.370	λ 1.0×10^{-6}	R2 0.303	λ 3.2×10^{-5}	R2 0.317	λ 4.0×10^{-3}
Variable	Coef.	Variable	Coef.	Variable	Coef.	Variable	Coef.	Variable	Coef.	Variable	Coef.
Vegetation	0.508	Night	-0.492	Vegetation	0.641	Vegetation	0.829	Vegetation	0.696	Banner	-0.213
Night	-0.483	Road	0.365	Building	0.523	Building	0.545	Building	0.596	Sidewalk	-0.188
Road	0.373	Sidewalk	0.363	Road	0.507	Road	0.480	Road	0.437	Person	-0.132
Building	0.359	Vegetation	0.348	Night	-0.460	Sidewalk	0.451	Sidewalk	0.417	Fence	0.131
Sidewalk	0.330	Building	0.245	Sidewalk	-0.437	Sky	0.387	Sky	0.381	Wall	0.121
Q19. Anxious		Q20. Favorite		Q21. Interesting		Q22. Willingness to pass					
R2 0.424	λ 2.0×10^{-3}	R2 0.364	λ 5.0×10^{-3}	R2 0.338	λ 4.0×10^{-3}	R2 0.389	λ 1.3×10^{-3}				
Variable	Coef.	Variable	Coef.	Variable	Coef.	Variable	Coef.				
Night	0.525	Vegetation	0.541	Banner	0.260	Night	-0.449				
Sidewalk	-0.184	Night	-0.404	Vegetation	0.136	Sidewalk	0.235				
Road	-0.121	Building	0.357	Person	0.122	Road	0.162				
R-1*	0.112	Road	0.357	Wall	-0.122	R-1*	-0.146				
Rainy	0.099	Sidewalk	0.355	Sky	-0.099	Rainy	-0.129				

(red) : Positive standardized regression coefficient
(blue) : Negative standardized regression coefficient
Street images extracted in Chapter 2.2 were divided into training and test datasets with a 0.75:0.25 split for model training.
(The coefficients shown in this table are derived from the test dataset)
* These abbreviations correspond to Figure 3.

Table 5. Standardized Partial Regression Coefficients for Each Impression Evaluation Item (Top 5 Variables and Coefficients of Determination)

4. Discussion

The results of the comparison of the average impression evaluation scores for each land-use zone suggest how street impression evaluation is characterized in each land-use zone. Notably, in commercial areas, the Q21. Interesting" was higher at night than during the day—a trend which differed from that of other impression evaluation items.

The predicted street impression rating scores were plotted for each area and compared. A comparison of streets rated "Q9. Quaint" suggests that, during relatively bright hours, people tend to find streets with a calm atmosphere, rich planting, and natural quaintness. However, in the spatial distribution of rain and nighttime, the evaluation of this item increased on streets around commercial areas. From these results, it can be inferred that the streets that people perceive may differ depending on the time of the day. This indicates the importance of considering different times of day and weather conditions.

We also conducted a factor analysis to analyze the relationship between street impression evaluation scores and elements of the streetscape. Even though "Building" and "Vegetation" positively influenced the evaluation in "Q10. Safe," they had little influence on the evaluation in "Q19. Anxious." This difference suggests that the psychological factors of darkness may have a greater impact on the evaluation.

We conducted a comparative analysis of weather and time of day. However, in terms of scene diversity, there are also differences between weekdays and holidays. Additionally, this study did not consider seasonal variations or the cultural characteristics specific to Tokyo. In particular, evaluations of street images taken in winter or spring, or during specific flowering seasons, may differ from the results of this study. Continued data collection will make it possible to understand the differences in impressions of a wider variety of scenes.

Some streets consistently receive particularly high or low impression scores; however, the underlying reasons for these evaluations remain unclear. Further investigation is required to clarify these patterns. As a next step, we plan to conduct a detailed factor analysis that incorporates the location, form, and quality of individual street components. This approach aims to identify interpretable indicators that influence street impressions and to support evidence-based urban design and decision-making. In future work, we will also consider including interaction terms in the regression model for factor analysis. For example, greenery and weather are interrelated: on sunny days, greenery appears vivid and gives an impression of "comfort" and "openness," whereas on cloudy days, colors appear darker and may be perceived negatively. Additionally, although the images were cropped to simulate a focal length of approximately 30 mm, it remains uncertain whether this representation truly reflects human visual experience. Verifying this correspondence is an important subject for future studies.

In the future, by integrating a range of technologies, such as computer vision, real-time geospatial analysis, sensor-based environmental monitoring, and machine learning, it may be possible to quantify impression evaluations of streets at near real-time speeds. For example, when combined with human flow data, such systems can reveal the relationship between street congestion and impression evaluations, helping to identify the types of congestion that are perceived as unpleasant. Furthermore, from a crime prevention perspective, it may become possible to assess perceived street safety alongside real-time human mobility, enabling the flexible and responsive rerouting of pedestrians or vehicular traffic.

Additionally, the imaging method used in this study was more labor-intensive than the GSV and other datasets. However, anyone with a camera attached to their bicycle can use this method to acquire street images. As the said method becomes more widely used, it will be possible to obtain datasets from a wider variety of areas under a wider variety of conditions, thus enabling a more comprehensive survey.

5. Conclusions

In this study, impression evaluation questionnaires were administered using street images captured in a unique way under various weather conditions, at different times of the day, and in multiple regions. Based on the street impression evaluation model trained using questionnaire results, the effects of weather and time of day on street impression evaluation were quantitatively verified. To understand the underlying factors shaping people's impressions of streetscapes, we conducted a factor analysis using LASSO regression. This approach allowed us to explore various streetscape components quantitatively. As the collection of street images via bicycles progresses using the methods proposed in this study, we can proceed with the evaluation of street impressions using existing deep learning models. Overall, the study has quantitatively shown that impressions on a street are not uniquely determined by a single condition, but rather vary depending on weather conditions and time of day. These changes vary by scene and land use type, highlighting the context-dependent nature of street perception.

Acknowledgements

This research was supported by JSPS KAKENHI Grant Number JP 22K04490 and the FY 2024 Urban Regeneration Research Grant from the Organization for Promoting Urban Development. We would like to express our gratitude to Rakuten Group, Inc. for their efforts in conducting the survey, and to all those who responded to the survey.

References

Biljecki, F., Ito, K., 2021. Street view imagery in urban analytics and GIS: A review. *Landsc. Urban Plan.*, 215, 104217. doi.org/10.1016/j.landurbplan.2021.104217.

Cheng, B., Misra, I., Schwing, A.G., Kirillov, A., Girdhar, R., 2022. Masked-Attention Mask Transformer for Universal Image Segmentation. Presented at the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 1290–1299.

Fan, Z., Biljecki, F., 2024. Nighttime street view imagery: A new perspective for sensing urban lighting landscape. *Sustain. Cit. Soc.*, 116, 105862. doi.org/10.1016/j.scs.2024.105862.

Huang, W. J., & Wang, P. 2018. “All that's best of dark and bright”: Day and night perceptions of Hong Kong cityscape. Tourism Management, 66, 274-286.

Ito, K., Quintana, M., Han, X., Zimmermann, R., Biljecki, F., 2024. Translating street view imagery to correct perspectives to enhance bikeability and walkability studies. *Int. J. Geogr. Inf. Sci.*, 38, 2514–2544. doi.org/10.1080/13658816.2024.2391969.

Liu, Z., Mao, H., Wu, C.-Y., Feichtenhofer, C., Darrell, T., Xie, S., 2022. A ConvNet for the 2020s. Presented at the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 11976–11986.

Ministry of Land, Infrastructure, Transport and Tourism, 2003. Introduction of Urban Land Use Planning System in Japan. Available at: https://www.mlit.go.jp/common/001050453.pdf (accessed on 11 April 2025).

Ministry of Land, Infrastructure, Transport and Tourism, 2021. Summary of the White Paper on Land, Infrastructure, Transport and Tourism in Japan. Available at: https://www.mlit.go.jp/en/statistics/content/001442571.pdf (accessed on 11 April 2025).

Neuhold, G., Ollmann, T., Rota Bulo, S., Kontschieder, P., 2017. The Mapillary Vistas Dataset for Semantic Understanding of Street Scenes. Presented at the Proceedings of the IEEE International Conference on Computer Vision, 4990–4999.

Ogawa, Y., Oki, T., Zhao, C., Sekimoto, Y., Shimizu, C., 2024. Evaluating the subjective perceptions of streetscapes using street-view images. *Landsc. Urban Plan.*, 247, 105073. doi.org/10.1016/j.landurbplan.2024.105073.

Oki, T., Kizawa, S. 2021. Evaluating visual impressions based on gaze analysis and deep learning: A case study of attractiveness evaluation of streets in densely built-up wooden residential area. *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.,* XLIII-B3-2021*,* 887-894. doi.org/10.5194/isprs-archives-XLIII-B3-2021-887-2021

Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Series B Methodol.*, 58(1), 267–288. Available at: http://www.jstor.org/stable/2346178 (accessed on 7 April 2025).

Tin Tin, S., Woodward, A., Robinson, E., & Ameratunga, S. 2012. Temporal, seasonal and weather effects on cycle volume: an ecological study. Environmental Health, 11, 1-9.

Ye, X., Wang, Y., Dai, J., Qiu, W., 2024. Generated Nighttime Street View Image to Inform Perceived Safety Divergence Between Day and Night in High Density Cities: A Case Study in Hong Kong. doi.org/10.2139/ssrn.4979919.

Ye, Y., Zeng, W., Shen, Q., Zhang, X., Lu, Y., 2019. The visual quality of streets: A human-centred continuous measurement based on machine learning algorithms and street view images. *Environ. Plan. B Urban Anal. City Sci.*, 46, 1439–1457. doi.org/10.1177/2399808319828734.

Zuiderveld, K. J. 1994. Contrast limited adaptive histogram equalization. *Gems*, 4(1), 474-485.

Appendix

The image data presented in this paper are available on request from the authors due to privacy restrictions.