Modeling Perceived Street Safety from Street View Imagery: Global and Local Perspectives on Street Space Features and Crash Incidences

Janine A. Mendoza[1], Kim Paolo L. Satera[1], Karl Adrian P. Vergara[1]

[1]Department of Geodetic Engineering, University of the Philippines Diliman, Quezon City, Philippines –
{jamendoza16, klsatera, kpvergara}@up.edu.ph

Keywords: Street View Imagery, Safety Perception, Machine Learning, Gender, Roles in Traffic

Abstract

Road safety remains a global issue, with traffic crashes causing approximately 1.19 million deaths annually. While global datasets have been used to assess safety perception, limited studies examine how local perceptions vary and relate to street features and crash incidences. This study addresses that gap by modeling perceived street safety using Street View Imagery (SVI), comparing global (Place Pulse) and local (Quezon City, Philippines) perspectives. Models were developed to analyze how different populations perceive safety in relation to street space characteristics, and how these perceptions align spatially with crash incidences. A PSPNet-based semantic segmentation model extracted 28 features from SVIs which were modeled with perceptions from diverse groups defined by gender, road user role, and geographic context. Group-specific multiple linear regression models were developed to predict safety perception, and SHAPley Additive exPlanations (SHAP) interpreted feature influence. Results revealed natural and open-space elements like trees, sky, and sidewalk increased perceived safety, while vehicles, walls, and dense buildings reduced it. Local models outperformed global (R^2 up to 0.55 vs. ≈ 0.12), highlighting the value of localized, group-specific modeling. However, a perception–risk gap emerged: local respondents perceived crash-prone areas as safer than non-crash zones, whereas global perception aligned better with actual crash data. These findings emphasize that familiarity and environmental normalization can shape safety perception independently of real-world risk. The study highlights the limitations of relying solely on global data and advocates for equity-oriented urban design, prioritizing safe, inclusive, and context-aware public spaces that reflect lived realities of diverse communities.

1. Introduction

1.1 Background of the Study

Road safety continues to be a major global concern, especially in urban areas where streets make up 20–30% of city space and play a vital role in mobility, social interaction, economic activity, and safety. Perceptions of road safety emerge from a layered interaction of socio-economic status, enforcement strength, and cultural norms that differs markedly from one setting to another. A 175-country comparison, for instance, shows that crash fatality rates rise steeply as national income falls, while religion or culture alone has no consistent statistical link to those rates once traffic-law enforcement is taken into account (Magableh et al., 2021). Country-pair studies reinforce that message: Norwegians and Ghanaians, exposed to very different road environments, report contrasting attitudes toward speed, risk, and rule compliance, and these attitudes significantly predict self-reported collisions (Lund and Rundmo, 2009). In short, the same street geometry can “read” as safe in one context and hazardous in another, and global design ideals cannot simply be transplanted without local calibration.

Safety is not just a product of built environment itself but of how different groups interpret that built environment. These global–local contrasts set the stage for examining whether visual attributes that evoke feelings of security also coincide with recorded crash hotspots in a given city.

There's still a lack of research connecting these environments to how people perceive safety, and an important gap remains in understanding how safety perception aligns—or misaligns—with actual crash data. This is important because places that look or feel safe may not actually be so. Urban environments are visually complex, and it is often unclear whether individuals can accurately perceive potential road hazards (Yu et al., 2024).

In recent years, SVI has emerged as an effective tool in urban studies which provides ground-level visual data of street environments. It supports automated research assessments and enables quantification of the built environment features such as crosswalks, traffic signage, sidewalks, and smaller details like trash cans (Li et al., 2022). This technology facilitates a scalable and replicable means of examining the urban environment at the human-eye level.

This study leverages SVI and machine learning to model perceived street safety in Quezon City, Philippines, analyze the influence of streetscape features, and examine the alignment—or misalignment—between perception and recorded crash incidences from both global and local viewpoints.

1.2 Objectives

The primary objective of this study is to examine the potential of SVI in analyzing street space features that influence safety perception and to relate these perceptions to actual crash frequency in Quezon City. Specifically, the study seeks to develop safety perception models using street-level features extracted from SVI and to determine which elements most significantly impact predicted safety perception scores. It also evaluates the performance of various regression techniques and examines differences in perceived safety based on gender, road user type (driver vs. non-driver), and geographic context (local vs. global). Lastly, the study investigates the relationship between predicted safety scores and the locations of actual crash incident to inform data-driven road safety planning and policy development.

1.3 Scope and Limitations

Google Street View (GSV) images are the primary data source for extracting street-level features using a pre-trained PSPNet semantic segmentation model based on the ADE20K dataset. The model identifies 28 element classes: sky, tree, road, grass, earth, mountain, plant, house, sidewalk, streetlight, traffic light, fence, rail, signboard, pole, sidewalk, path, bridge, sunblind, wall, trashcan, skyscrapers, water, bus, truck, car, motorbike and buildings.

The data is limited by the quality and timeliness of the images, as image resolution, capture dates, and update frequency vary by location. In some areas, the available imagery may be several years old, which poses a temporal limitation: the visual conditions captured may no longer reflect the current conditions of street spaces. This may affect perception assessment, since respondents evaluate images based on past conditions that may not align with present realities. In addition, the crash data may be incomplete, lack precise location details, and have no available information on the level of accuracy. Furthermore, inter-rater consistency checks and sensitivity analyses for the survey participants were not conducted.

2. Literature Review

2.1 Machine Learning and Street View Imagery in Urban Safety Analysis

The integration of machine learning with Street View Imagery (SVI) has enabled researchers to quantify and analyze urban features influencing safety perception. Semantic segmentation models such as PSPNet have proved effective in extracting built environment elements such as roads, sidewalks, and traffic related infrastructure, with high accuracy. Studies like those of Cui et al. (2023) and Ma et al. (2021) demonstrate how segmented features from SVI can be used to evaluate urban qualities like greenness, enclosure, and walkability. These visual indices, when modeled using advanced regression algorithms and interpretability tools like SHAP, allow for robust predictions of perceived safety across diverse settings. Datasets such as MIT's Place Pulse 2.0 have further expanded the scale of analysis, offering perceptual scores for training machine learning capable of classifying safety across cities and countries.

2.2 Discrepancies Between Perceived Safety and Actual Crash Data

Despite continuous technological progress, literature consistently reveals a misalignment between perceived safety and actual crash data. Some streets may look unsafe but do not necessarily experience frequent accidents, while others that appear safer may have higher crash occurrences. This pattern has been noted in the works of Yu et al. (2024) and Hamim and Ukkusuri (2023), where areas with fewer recorded crashes often received lower safety perception scores. These mismatches call the need to reconcile perceptual biases with empirical crash data to inform effective urban safety interventions.

2.3 Contextual and Demographic Gaps in Existing Research

Gaps still persist in the current body of research as few studies contextualized findings within countries like the Philippines, where urban dynamics and transport conditions differ significantly from global contexts. Moreover, most perception studies rely on binary comparisons without unpacking why certain visual features evoke a sense of safety or danger. Demographic factors such as gender, age, and road usage frequency are often overlooked, despite their influence on safety perception (Cui et al, 2023). Addressing these gaps by conducting localized and demographically stratified analyses would enhance the reliability and inclusivity of safety perception models in urban planning.

3. Methods

3.1 Study Area

This study was conducted in Quezon City, a highly urbanized and landlocked city in Luzon, Philippines. With a total land area of about 171.71 square kilometers, Quezon City accounts for nearly one-fourth of the National Capital Region (NCR), making it the largest local government unit (LGU) in Greater Manila (Philippine News Agency, 2018).

Quezon City's diverse urban landscape—spanning residential, commercial, industrial, institutional, and open spaces—is supported by an extensive road network that connects key areas but also contributes to traffic and safety issues. In 2023, the Metro Manila Accident Recording and Analysis System (MMARAS) reported 32,387 road crashes in the city, the highest in the region. The leading fatal collision types were vehicle-pedestrian crashes (27.77% of fatalities), single-vehicle crashes, rear-end collisions, collisions with fixed objects, and side swipes. This highlights the urgent need for data-driven traffic safety measures.

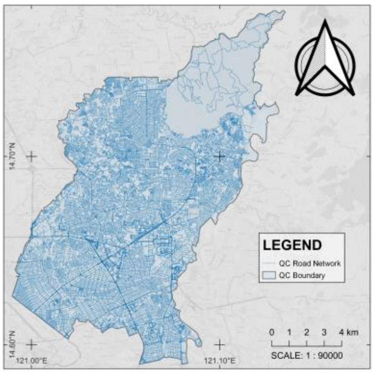

Figure 1. Quezon City Administrative Boundary and Road Network.

3.2 Data Used

The data used in this study are summarized in Table 1. Both local and global datasets were incorporated to explore differences in how people perceive safety in urban environments. The local dataset, focused on Quezon City, reflected the street conditions familiar to locals. In contrast, the global dataset, Place Pulse 2.0, contained a wide variety of streetscape images from different countries, evaluated by respondents from diverse cultural and geographic backgrounds.

The local dataset included 48,150 SVIs from Quezon City, collected at 50-meter intervals along the road network. To maintain consistency with the Place Pulse 2.0 dataset, the following parameters were applied for each request: image size of 400×300 pixels, a field of view (FOV) of 90 degrees, a pitch of 0 degrees, and a compass heading aligned with the street orientation. From this pool, a sample of 320 images was used in a local safety perception survey.

Administrative boundary, road network, and land use shapefiles supported spatial sampling and analysis. The global dataset, Place Pulse 2.0, comprised 84,899 images and 502,357 pairwise safety comparisons from respondents worldwide.

Data	Information	Data Source
Administrative Boundary	Study area shapefile	Humanitarian Data Exchange (HDX) sourced from NAMRIA
Road Network	Road links shapefile	Open Street Map (OSM)
Land Use Data	Classification of land uses shapefile	QC City Planning and Development Department
Local Dataset: Street View Images	Street space features	Google Street View API
Local Dataset: Safety Perception Data	Survey results of the local survey	Survey participants, Qualtrics Survey Platform
Global Dataset: Place Pulse 2.0	Street space features and survey results on perceptions related to safety	Massachusetts Institute of Technology (MIT)
Crash Accident Reports	Location of 2020 Traffic Accidents in Quezon City	SafeTravelPH Mobility Innovations Organization Inc.

Table 1. Data Used

3.3 Local Survey

The survey was crowdsourced online through Qualtrics survey platform and involved 266 Filipino participants who compared image pairs based on road safety. Respondents were stratified by gender (135 females, 131 males) and road user role (138 drivers, 128 non-drivers). This allows comparative analysis across these subgroups. Compared to prior crowdsourced studies, this sample size is relatively large: several studies in China and other countries relied on only 5 to 60 volunteers (Baran et al., 2018; Ma et al., 2021; Qiu et al., 2021; Wang et al., 2022). Moreover, Cui et al. (2023) demonstrated that reliable safety perception scores can be obtained even with a group of 60 volunteers (30 women and 30 men), achieving a strong correlation between predicted and evaluated scores ($R^2 = 0.664$, Pearson $r = 0.815$, $p < 0.01$). In this research, the use of 266 participants provides a stronger basis for subgroup analysis and enhances the reliability of the findings, although the results may still not fully generalize to the Quezon City population.

Survey design incorporated measures to reduce bias: SVIs were randomly paired and presented, crime-related considerations were explicitly excluded to focus on road safety, and each image was evaluated by at least five respondents to minimize individual bias.

3.4 Workflow

This study adopted a multi-stage methodology, as illustrated in Figure 2, comprising four main phases: Safety Perception Modeling, Feature Contribution Analysis, Spatial Analysis of Safety Perception, and Safety Perception–Crash Incidence Relationship Analysis.

3.5 Safety Perception Modeling

The methodology models perceived safety using street-level imagery by integrating semantic segmentation, pairwise comparison-based safety perception scoring, and regression modeling. Both global and local models were developed to capture demographic and geographic variation, which enables deeper insights into how the urban street environment influences subjective perceptions of safety.

3.5.1 Image Segmentation and Feature Extraction: To extract interpretable street-level features from the SVIs, this study employed semantic segmentation using a Pyramid Scene Parsing Network (PSPNet) model, pre-trained on the ADE20K dataset, assigning semantic labels to each pixel of an image. The PSPNet architecture integrates a ResNet-50-Dilated backbone with a Pyramid Pooling Module (PPM) and deep supervision, enabling both fine-grained local and contextual global feature extraction.

This implementation was adopted from the publicly available GitHub repository CSAILVision/semantic-segmentation-pytorch, developed by the Computer Vision group at MIT's Computer Science and Artificial Intelligence Laboratory (CSAIL). It was selected for its robust performance in parsing complex urban scenes, achieving a Mean IoU of 42.14 and Pixel Accuracy of 80.13% under standard benchmark conditions. Each image was segmented into 150 ADE20K classes, from which 28 elements were selected to be relevant in this study. Post-segmentation, the number of pixels per class was divided by the total pixel count (120,000) to compute the class-wise pixel proportion.

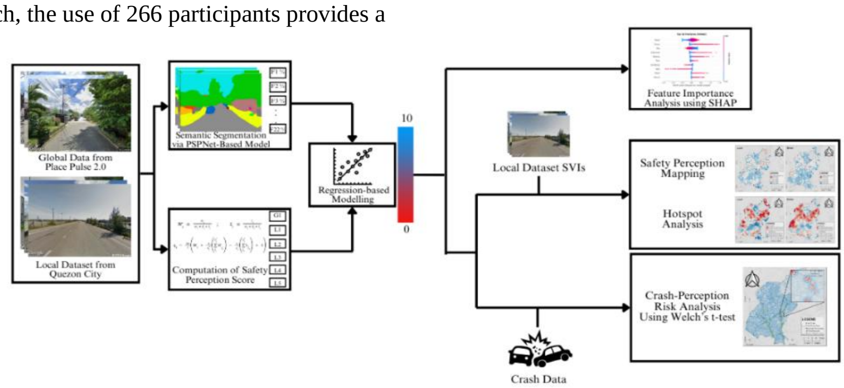

Figure 2. General Workflow.

3.5.2 Safety Perception Score Computation: Safety perception scores were computed from pairwise comparisons of images in both the local and global datasets. Following Moreno-Vera (2021), we adopted the concept of *"intensity of perception"*, defined as the proportion of times an image was chosen as safer.

For each image i, the positive rate W_i and negative rate L_i were calculated as:

$$W_i = \frac{w_i}{w_i + d_i + l_i} \quad ; \quad L_i = \frac{l_i}{w_i + d_i + l_i} \tag{1}$$

To compute the safety perception score (Q-score), the following formula was used:

$$q_i = \frac{10}{3}\left(W_i + \frac{1}{n_i^w}\left(\sum_{j_1}^{n_i^l} W_{j_1}\right) - \frac{1}{n_i^l}\left(\sum_{j_2}^{n_i^l} L_{j_2}\right) + 1\right) \tag{2}$$

(where, q_i: Q-score; n_i^w: total count of images that the image i wins against; n_i^l: total count of images that the image i loses to; j_1: denotes the images that lose against the image i; j_2: denotes the images that win against the image i)

Only images that were compared at least five times were assigned a safety perception score to ensure reliability and consistency (Yu et al., 2024). In the local dataset, each image appeared in at least 200 comparisons.

3.5.3 Safety Perception Model Development: Safety perception models were developed by linking segmented street features from SVIs with computed safety perception scores from pairwise comparisons. Two types of models were developed: (1) a global model trained on Place Pulse 2.0 dataset to capture generalized safety perceptions, and (2) local models trained on local dataset to reflect context-specific perceptions. Local subgroup models were developed for (3) male; (4) female; (5) driver; and (6) non-driver. These groupings based on gender (male and female), road user role (driver and non-driver), and geographic context (local and global) served as the basis for comparing how different populations perceive safety in relation to various street space characteristics. The study employed various regression models to predict safety perception scores: Adaptive Boost (ADAB), Bagging Regression (BR), Decision Tree (DT), Multiple Linear Regression (MLR), Random Forest (RF), Support Vector Machine (SVM), and Voting Selection (VS). Hyperparameter tuning was performed using Optuna to optimize the performance of each model.

3.6 Feature Contribution Analysis

SHAP (Shapley Additive Explanations) method was used to analyze the roles of various street features in shaping perceived safety. Shapley values were assigned to the selected classes to quantify their contributions to safety perception. Based on game theory, SHAP assigns Shapley values to features, indicating both the magnitude and direction (positive or negative) of their influence on individual predictions (Hao et al., 2024). To identify the top 10 features in each safety perception model, SHAP computed the mean absolute Shapley values, ranking features based on the strength of their impact, regardless of direction, on the model's predictions. This approach identifies which features have the most overall effect on the model's behavior.

3.7 Spatial Analysis of Safety Perception

To explore spatial patterns in perceived safety model predictions were aggregated using the H3 hexagonal hierarchical spatial index at a resolution level of 9 (approx. 0.105 km²), following Zhao et al. (2023). The H3 index was implemented using Python, wherein each SVI prediction was spatially assigned to its corresponding hexagonal grid cell. Each indicator value for a hexagon represents the mean predicted safety perception score of all SVIs located within that specific cell. This method allowed the aggregation of localized perceptions into a standardized spatial unit, providing a consistent basis for comparative spatial analysis between groups.

QGIS was used to visualize spatial distributions and reveal disparities in safety perceptions across participant groups (e.g., male vs. female, driver vs. non-driver). To detect statistically significant clusters of high or low perceived safety, Getis-Ord Gi* analysis was applied. These hotspots and coldspots highlighted areas for potential urban design interventions.

3.8 Safety Perception–Crash Incidence Relationship Analysis

To examine the relationship between perceived safety and actual road crash incidents, crash data with geographic coordinates from a one-year period were analyzed. A 50-meter buffer zone was generated around the estimated crash locations and classified as "crash-prone," while areas beyond this but within a 100-meter buffer zone were categorized as "non-crash-prone" (Hamim & Ukkusuri, 2023). Moreover, a Welch's t-test was employed to assess whether significant differences existed between the predicted safety perception scores of crash-prone and non-crash-prone areas across subgroups.

4. Results and Discussion

4.1 Semantic Feature Extraction from SVI

A total of 28 semantic features were extracted from Street View Imagery (SVI) using a PSPNet-based semantic segmentation model. These features served as inputs for modeling perceived safety scores across different population subgroups. To improve interpretability, semantically related features such as poles (combining poles, traffic lights, and streetlights), buildings (buildings and skyscrapers), and vehicles (cars, trucks, buses, motorbikes) were aggregated. Each feature represents the proportion of pixels occupied, reflecting its visual prominence within the streetscape.

Feature correlations were generally low, indicating minimal multicollinearity, supporting their use in regression analysis. Sample segmented images, as seen in Figure 3, demonstrated the model's ability to classify key street elements, with the most prominent features being road, buildings, sky, trees, and vehicles. Minor segmentation inaccuracies were noted but deemed acceptable given the study's focus on overall scene perception rather than pixel-perfect accuracy. Accuracy assessment was based on visual inspection due to the absence of ground truth annotations, aligning with the study's emphasis on perceived safety and supported by relevant literature on pre-trained model reliability. To provide quantitative context, the utilized PSPNet model which is commonly used for scene parsing has demonstrated benchmark performance of around 43% mean IoU on the ADE20K dataset (e.g. 43.3 % with Resnet-101, around 43.5% with Resnet-152. Qualitative inspection of the imagery of this study exposed that fine-grained features were more prone to misclassification, whereas dominant structural elements were reliably segmented.

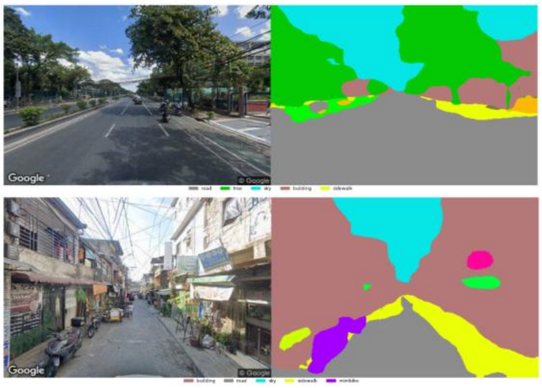

Figure 3. Sample segmented images.

4.2 Descriptive Analysis of Perceived Safety Scores

Perceived safety scores were analyzed using two data sources: a local survey and the global Place Pulse 2.0 dataset.

	Local Survey					Place Pulse 2.0
Metric	Male	Female	Driver	Non-Driver	Local	Global
Min	3.260	3.330	3.120	3.330	3.370	0.200
Max	5.620	5.830	5.680	5.900	5.610	9.630
Mean	4.626	4.618	4.583	4.662	4.627	4.726
SD	0.494	0.593	0.524	0.563	0.527	1.255

Table 2. Descriptive Statistics of Survey-Derived Safety Perception Scores

Results reveal that local scores ranged narrowly from about 3.1 to 5.9, with mean values around 4.6 and low variability across demographic subgroups, indicating relatively consistent perceptions. Slight differences emerged, with females rating safety marginally lower than males, and non-drivers perceiving higher safety than drivers, suggesting subtle influences of gender and mobility experience. In contrast, the global dataset showed a much wider range and higher variability in safety scores, reflecting the diverse cultural and environmental contexts represented. This contrast highlights the context-specific nature of safety perception, where local respondents draw on shared experiences, while global evaluations encompass varied, unfamiliar urban settings. The findings underscore the value of localized data in informing urban design and policy, as global datasets alone may not fully capture the nuanced perceptions of city residents.

4.3 Group-Based Variation in Safety Perception

Subgroup-level differences in perceived safety were examined across gender (male vs. female), role in traffic (driver vs. non-driver), and geographic context (local vs. global). Statistical tests on pre-modeling safety perception scores revealed significant variation across the two subgroups.

	Group 1	Group 2	Group 3
t-statistic	0.513	-5.078*	-3.337*
Cohen's d	0.029	-0.284	-0.103

Note: *p-value < 0.01

Table 3. Statistical Test Results

As shown in Table 3, no significant difference was found between male and female respondents ($p = 0.608$, $d = 0.029$), suggesting similar interpretations of safety. However, significant differences emerged between drivers and non-drivers ($p < 0.01$, $d = –0.284$), and between local and global respondents ($p = 0.001$, $d = –0.103$), with non-drivers and local participants rating scenes as safer. While gender differences were statistically insignificant, we still constructed separate models for each subgroup to account for potential differences in how visual features influence their perceptions.

4.4 Evaluation of Machine Learning Models for Safety Perception Score Prediction

Machine learning models were evaluated across demographic (gender, road user role) and spatial (local, global) subgroups (Figure 4). MLR consistently performed well, ranking highest in the local, male, and driver groups, with R^2 ranging from 0.125 (global) to 0.548 (local). MAE and RMSE were lowest in subgroup settings, confirming MLR's suitability for perception prediction in homogeneous groups.

Figure 4. Comparison of R^2 across all machine learning models and subgroups.

In contrast, Random Forest led the global model ($R^2 = 0.167$), reflecting its strength in handling heterogeneous data. However, overall performance in the global setting was weaker, likely due to perceptual and contextual variation. Local models performed better due to consistent lived experiences among respondents. Furthermore, to assess model performance, an 80:20 train-test split was used. The MLR model showed consistent R^2 values around 0.50 across most subgroups, indicating a reasonable fit and suggesting that the model captures a substantial portion of the variance in perceived safety without overfitting.

4.5 Influence of Street Space Features on Safety Perception

The MLR models provided a clear and interpretable way to predict safety perception, with moderate R^2 values (~0.5) indicating acceptable fit given the inherent complexity of human behavior and perception. However, MLR alone cannot fully capture the varying influence of street-level features across different subgroups. To deepen this understanding, SHAP analysis was used to quantify feature importance, revealing how each feature contributed positively or negatively to the predictions on an individual basis.

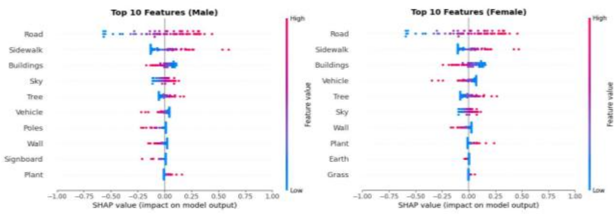

Figure 5. SHAP value of top features used in MLR (Group 1).

The SHAP analysis revealed that features such as roads, sidewalks, and trees consistently emerged as strong positive predictors of safety perception across all subgroups. However, group-specific variations were observed. Gender-based differences emerged, as seen in Figure 5, with males reacting more to infrastructure complexity and females more sensitive to vehicle presence and density which suggests that gender plays a key role in shaping how people perceive safety in urban environments.

In particular, women often perceived congested streets as less safe, suggesting that traffic conditions exert a stronger influence on their sense of safety. This indicates that measures such as traffic calming and pedestrian buffers are especially important for addressing female safety concerns. This divergence highlights the importance of incorporating diverse perspectives in safety modeling, as men may focus on navigational complexity while women prioritize environmental stressors like congestion and exposure to traffic. Such nuances point to the value of designing urban spaces that reduce visual and physical clutter while enhancing comfort and protection, especially for more vulnerable groups.

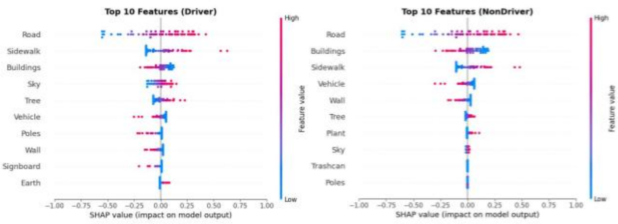

Figure 6. SHAP value of top features used in MLR (Group 2).

Differences also emerged based on road user role. As seen in Figure 6, Drivers tended to associate natural features with safety but viewed man-made elements such as poles and buildings as negative, while non-drivers were more sensitive to buildings and walls as indicators of lower safety. These patterns may stem from varying familiarity and exposure to risk, with drivers focusing more on navigational clarity and non-drivers on their immediate vulnerability as pedestrians. This suggests the importance of designing streets that balance mobility efficiency with pedestrian comfort and protection.

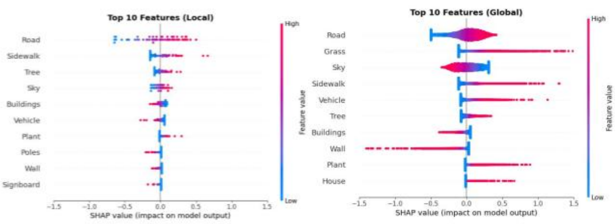

Figure 7. SHAP value of top features used in MLR (Group 3).

Comparing spatial contexts, as shown in Figure 7, local respondents emphasized real-world conditions such as vehicle density and walls lowering safety, while global respondents favored greener, more orderly environments. This reflects an idealized notion of safety often shaped by clean, well-maintained, and visually structured spaces—features commonly promoted in global urban design standards. While these visual cues suggest order and livability, they may overlook deeper, lived realities. The contrast highlights the importance of grounding perception models in local experience, especially in cities where surface-level aesthetics don't always match the on-the-ground sense of safety.

4.6 Safety Perception Score Hotspot Analysis

Spatial patterns of predicted safety perception scores across Quezon City were examined through hotspot analysis through Getis-Ord Gi* (see Figures 8-10), generating hexagonal grid maps by group and identifying significant hotspots and cold spots at varying confidence levels.

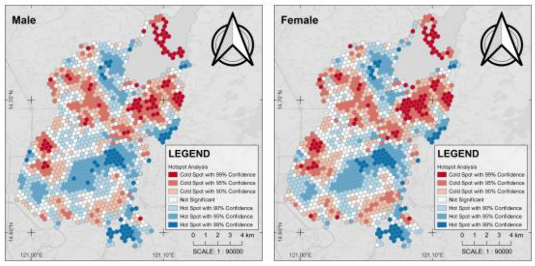

Figure 8. Quezon City Hotspot Map (Group 1).

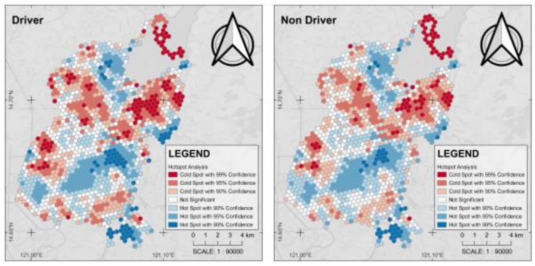

Figure 9. Quezon City Hotspot Map (Group 2).

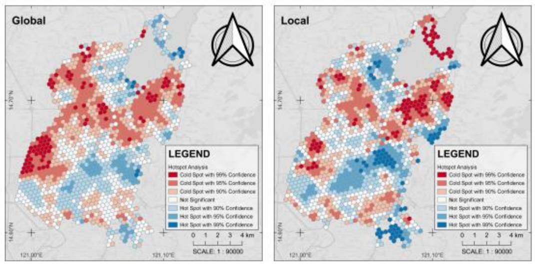

Figure 10. Quezon City Hotspot Map (Group 3).

Results showed strong spatial similarities between males and females (Figure 8), and between drivers and non-drivers (Figure 9), particularly highlighting the northeastern region as a consistent cold spot. This suggests shared spatial interpretations of safety within locally grounded experiences.

In contrast, perceptions diverged more between local and global (Figure 10): locals viewed southern areas as safer and northeastern areas as less safe, while global respondents perceived more widespread unsafety, especially in the northern and western zones, with notable contrasts in how specific regions were evaluated.

The hotspot analysis revealed that while local and global respondents shared some spatial patterns in perceived safety, global participants tended to view more areas, especially in the north as safe. This contrast suggests that locals, informed by lived experiences, may perceive risks not visible in images, such as crime or traffic issues. It highlights the need for localized models that reflect these ground-level realities. Areas in central and southern Quezon City, with better infrastructure like sidewalks, greenery, and wider roads, were generally seen as safer, while the less developed northern zones were rated lower, especially by locals.

4.7 Safety Perception–Crash Incidence Relationship Analysis

Metric		Group 1		Group 2		Group 3	
		Male	Female	Driver	Non-Driver	Local	Global
mean	crash	4.770	4.770	4.700	4.850	4.760	4.640
mean	non-crash	4.740	4.750	4.690	4.810	4.750	4.670
	t-statistic	-4.096**	-3.064**	-2.430*	-6.379**	-2.053*	5.267**

Note: *p-value < 0.05; **p-value < 0.01

Table 4. Summary findings of two-tailed Welch's t-test for crash-prone and non-crash-prone.

The relationship between perceived safety and actual crash incidence was assessed by comparing mean safety perception scores in crash-prone and non-crash-prone areas using Welch's two-tailed t-tests across demographic and spatial groups.

Results indicated a consistent, counterintuitive pattern among local groups—crash-prone areas were perceived as safer than non-crash-prone zones. This trend was observed across gender and road user roles. For instance, non-drivers perceived crash-prone areas with a higher mean score (4.85) compared to 4.81 in non-crash-prone areas, a difference that was highly significant ($t = -6.379$, $p < 0.01$). Similarly, the male group gave a mean score of 4.77 to crash-prone areas versus 4.74 in non-crash-prone areas ($t = -4.096$, $p < 0.01$), while female group also perceived crash-prone areas slightly higher at 4.77 compared to 4.75 ($t = -3.064$, $p < 0.01$). Local group overall followed the same pattern, with the contrast remaining statistically significant at level $\alpha = 0.05$. This consistent trend suggests that locals tended to misperceive crash-prone zones as safer.

In contrast, only the global group showed the expected pattern: higher safety perception scores in non-crash-prone areas (4.67 vs. 4.64), indicating better alignment with actual risk ($t = 5.267$, $p < 0.01$). This divergence points to a perception–risk gap with critical implications for urban safety assessment and policy. While visual assessments of safety are increasingly used to inform urban design, these findings suggest that perception may not be a reliable proxy for actual risk, particularly in familiar environments. Several behavioral theories provide insight into this discrepancy. The risk avoidance hypothesis suggests that people naturally avoid areas they perceive as unsafe, leading to lower exposure and, consequently, fewer recorded incidents in those areas, even if underlying risks remain (Cho et al., 2009). On the other hand, the caution enhancement hypothesis posits that individuals behave more cautiously in spaces they deem dangerous, which may reduce crash occurrences due to increased vigilance and risk-averse behavior (Edwards, 1999; Schneider et al., 2004; Dinh et al., 2020; Boua et al., 2022).

Familiarity with a place may also contribute to perceptual bias. As noted by Traunmueller et al. (2016), repeated exposure to a familiar environment can normalize risk, causing residents to overlook hazards that outsiders might notice immediately. This can explain why local respondents consistently perceived crash-prone zones as safer than they were. In contrast, global respondents, unfamiliar with local nuances, relied more heavily on visible cues and visual heuristics, resulting in assessments that more closely aligned with actual crash data. Another relevant lens is the risk compensation theory, which suggests that when individuals perceive an environment as safe, perhaps due to visible infrastructure or design interventions, they may behave more recklessly, inadvertently increasing risk (Levy & Miller, 2000; BMJ, 2002). This paradox may help explain why crash-prone areas can appear visually "safe" but still produce high incidence rates.

In addition to behavioral explanations, certain methodological limitations may have influenced the results. The use of static street view images captures only a snapshot in time, omitting important temporal and behavioral dynamics such as traffic flow, time-of-day effects, and pedestrian movement. Furthermore, semantic segmentation techniques used to classify street-level elements may misidentify subtle features that influence perceived safety, such as degraded signage or temporary obstructions. These technical constraints may partially explain mismatches between safety perception scores and actual crash records.

Synthesizing these findings reveals the complexity of interpreting perceived versus actual safety in urban environments. The consistent misalignment among local respondents suggests that perception-based measures alone may not be sufficient, as they may overlook critical safety risks. While aesthetics and streetscape design remain important, they must be complemented with behavioral insights and empirical crash data to ensure that interventions address not just how safe spaces look, but how safe they truly are.

5. Conclusions and Recommendations

This study compared both global and local perspectives on street safety and explored how these perceptions relate to street space features and actual crash incidence. Using semantic segmentation via PSPNet, localized safety perception surveys, and machine learning models, particularly Multiple Linear Regression, the research analyzed perceptions across subgroups defined by gender, road user roles, and geographic context.

Findings revealed significant differences between local and global perspectives. While the Place Pulse 2.0 data-trained global model showed only modest predictive performance ($R^2 \approx 0.12$), local models trained on local data performed significantly better (R^2 up to 0.55). This emphasizes the advantage of using localized, group-specific data when modeling safety perception.

Features like roads, sidewalks, and trees were generally seen as safer, while dense infrastructure like buildings and walls were linked to lower safety scores. SHAP analysis further revealed that each subgroup prioritized different cues, like drivers valuing natural openness, and female respondents being more sensitive to traffic density.

Moreover, local respondents who share common lived experiences perceive safety in ways that were more uniform but misaligned with actual crash data. Specifically, local participants perceived crash-prone areas as safer, a counterintuitive result suggesting that place familiarity and environmental normalization could have an impact. This perception–risk gap was consistently observed across local subgroups (e.g., male, female, driver, non-driver), indicating that local perceptions may not always be reliable indicators of actual safety. This gap highlights the importance of integrating both perceived and actual safety in urban planning.

The safety perception maps and visualizations generated through this research, enabled by street view imagery, can support more inclusive and equity-oriented urban design strategies. Street view imagery provided a scalable, high-resolution, and location-specific visual dataset that allowed for analysis of urban environments at eye level, closely reflecting how streets are experienced by different populations. For urban planners, this offers several practical advantages: it eliminates the need for costly and time-consuming field audits and enables broad geographic coverage across entire cities. More importantly, SVI captures the physical elements like sidewalk, presence of greenery, buildings, or visual clutter—that influence people's sense of safety. These tools provide spatial insights that let policymakers target design interventions precisely where they will make streets both objectively safer and subjectively more comfortable, especially for vulnerable or underserved communities. By linking perception-based safety maps with existing planning frameworks, these findings can be directly translated into targeted and context-sensitive interventions. In this way, locally calibrated models do not just point out where people feel unsafe but also help planners decide what kind of changes could make the streets safer and more welcoming.

To improve future studies on safety perception, researchers are encouraged to explore better-performing or locally trained semantic segmentation models that can more accurately identify subtle or context-specific features like informal paths or urban greenery. Expanding the set of variables, like socioeconomic indicators, green view indices, or population density, could also add depth to the analysis, especially for aspects not directly visible in street images. Considering more dynamic inputs such as traffic flow patterns or crowd-sourced mobility information, may likewise provide a fuller picture on how safety is shaped by movement and activity in the streets. Future work would also benefit from a more diverse and larger survey pool, involving participants from a wider range of backgrounds and locations. This would help capture a fuller range of lived experiences and make models more reflective of how different people truly experience safety. Factoring in other demographics such as income level or age may further enhance understanding of perception differences. By improving both the modeling techniques and the diversity of input data, future research can produce more grounded, context-sensitive, and socially relevant safety assessments.

References

Abramov, M., 2024: *Best Datasets for Training Semantic Segmentation Models*. Keymakr. https://keymakr.com/blog/best-datasets-for-training-semantic-segmentation-models/

Baran, P. K., Tabrizian, P., Zhai, Y., Smith, J. W., Floyd, M. F., 2018: An exploratory study of perceived safety in a neighborhood park using immersive virtual environments. *Urban Forestry & Urban Greening*, 35, 72–81. doi.org/10.1016/j.ufug.2018.08.009

BMJ 2002: Does risk homoeostasis theory have implications for road safety. 324(7346): 1149–52. PMCID: PMC1123100.

Cui, Q., Zhang, Y., Yang, G., Huang, Y., Chen, Y., 2023: Analysing gender differences in the perceived safety from street view imagery. *International Journal of Applied Earth Observation and Geoinformation,* 124, 103537. doi.org/10.1016/j.jag.2023.103537

Hamim, O. F., Ukkusuri, S. V., 2023: Towards safer streets: A framework for unveiling pedestrians' perceived road safety using street view imagery. *Accident Analysis & Prevention*, 195, 107400. doi.org/10.1016/j.aap.2023.107400

Levym, D. T., Miller, T., 2000: Review: Risk Compensation Literature — The Theory and Evidence. *Journal of Crash Prevention and Injury Control*, 2(1), 75–86. doi.org/10.1080/10286580008902554

Li, Y., Peng, L., Wu, C., Zhang, J., 2022: Street View Imagery (SVI) in the Built Environment: *A Theoretical and Systematic review. Buildings,* 12(8), 1167. doi.org/10.3390/buildings12081167

Lund, I. O., Rundmo, T., 2009: Cross-cultural comparisons of traffic safety, risk perception, attitudes and behaviour. *Safety Science*, 47(4), 547–553. doi.org/10.1016/j.ssci.2008.07.008

Ma, X., Ma, C., Wu, C., Xi, Y., Yang, R., Peng, N., Zhang, C., Ren, F., 2021: Measuring human perceptions of streetscapes to better inform urban renewal: A perspective of scene semantic parsing. *Cities,* 110, 103086. doi.org/10.1016/j.cities.2020.103086

Magableh, F., Matawah, J. A., Freeman, B., 2021: Investigating the impact of income, belief and culture on road safety. *American Journal of Traffic and Transportation Engineering*, *6*(3), 81. doi.org/10.11648/j.ajtte.20210603.13

Moreno-Vera, F., Lavi, B., Poco, J., 2021: Quantifying urban safety perception on street view images. *IEEE/WIC/ACM International Conference on Web Intelligence*. doi.org/10.1145/3486622.3493975

Qiu, W., Li, W., Liu, X., Huang, X., 2021: Subjectively measured streetscape perceptions to inform urban design strategies for Shanghai. *ISPRS International Journal of Geo-Information,* 10(8), 493. doi.org/10.3390/ijgi10080493

Traunmueller, M., Marshall, P., & Capra, L., 2016: ". . .when you're a Stranger": Evaluating
Safety Perceptions of (un)familiar Urban Places. Association for Computing Machinery. 71-77. doi.org/10.1145/2962735.2962761

Wang, L., Han, X., He, J., Jung, T., 2022: Measuring residents' perceptions of city streets to inform better street planning through deep learning and space syntax. *ISPRS Journal of Photogrammetry and Remote Sensing,* 190, 215–230. doi.org/10.1016/j.isprsjprs.2022.06.011

Yu, X., Ma, J., Tang, Y., Yang, T., Jiang, F., 2024: Can we trust our eyes? Interpreting the misperception of road safety from street view images and deep learning. *Accident Analysis & Prevention*, 197, 107455. doi.org/10.1016/j.aap.2023.107455

Zhao, H., Shi, J., Qi, X., Wang, X., Jia, J., 2017: Pyramid Scene Parsing Network. *ArXiv*. Cornell University. doi.org/10.1109/cvpr.2017.660