Article

Vision AI reveals waste-safety perception relationships in urban environments

Check for updates

Siwei Zhang, Xujing Yu & Jun Ma ✉

The perception of safety plays a crucial role in fostering sustainable urban communities, as it influences the willingness of citizens to engage in community activities and their residential choices. Understanding the factors associated with safety perception is, therefore, critical for community development. While previous studies focus on static built environmental features, this research leverages artificial intelligence (AI) to examine how street-level waste management shapes safety perceptions in metropolitan areas. Through computer vision and machine learning approaches, we quantify safety perception levels and identify various types of street-level waste in New York City. Our analysis reveals a strong negative relationship between uncontrolled waste (particularly widespread litter) and perceived safety, while properly managed waste shows weaker associations with safety perception. These findings demonstrate that dynamic environmental management factors, rather than just static infrastructure, are critically associated with urban safety perceptions. The study advances both theoretical understanding and practical strategies for enhancing urban safety perception through improved waste management services.

Visual safety perception in urban environments refers to how individuals evaluate the safety of their surroundings based on visual environmental cues, distinct from actual crime experiences[1]. This assessment focuses on perceived risk of criminal victimization (e.g., theft, assault, harassment), which is interpreted through environmental cues such as physical disorder, poor maintenance, and incivilities in the urban space. This distinct psychological construct differs from objective safety measures, as it specifically captures residents' interpretations of environmental cues that signal potential crime risk in their urban surroundings[2]. This psychological dimension of urban experience has gained increasing attention as research demonstrates that environmental impacts on mental well-being are as crucial as physical health outcomes in urban settings[3]. Moreover, the influence of safety perception transcends individual experiences to fundamentally shape community dynamics, including social cohesion, civic participation, and urban vitality, thus serving as a critical determinant of sustainable urban development. The significance of urban perception of safety is further emphasized in global frameworks, notably the United Nations Sustainable Development Goals (SDGs), particularly Goal 11: Sustainable Cities and Communities, which explicitly aims to "make cities and human settlements inclusive, safe, resilient, and sustainable[4]." Within this framework, ensuring safety and fostering a sense of safety are crucial to achieving urban inclusivity and sustainability[5].

Safety perception analysis from a human-centered perspective offers valuable insights into residents' subjective experiences, informing urban design and municipal management practices[6,7]. Previous studies have

extensively examined various factors that shape urban perception, with particular emphasis on static built environment characteristics, such as urban greenery[8], architectural design[9], public space configuration[10], and street layouts[11]. However, the potential impact of dynamic management practices, particularly ongoing maintenance and operational interventions, remains relatively unexplored. This research gap warrants attention, given that urban environments are complex adaptive systems that require systematic maintenance and management interventions to sustain their intended functions and qualities.

Among various dynamic urban management practices, street waste management serves as a fundamental indicator of urban service quality and operational effectiveness. Street waste manifests in two primary forms: controlled waste, which includes waste properly disposed of through formal collection systems and designated facilities, and uncontrolled waste, which encompasses unauthorized disposal, littering, and informal dumping practices[12,13]. The effective management of urban waste not only ensures daily urban operations but also contributes to the global sustainability agenda. Specifically, waste management is identified as a key target under SDG 11, particularly Target 11.6, which addresses the reduction of cities' adverse environmental impact through municipal waste management[14]. Furthermore, SDG 12 reinforces this priority by emphasizing waste reduction through prevention, recycling, and reuse[15]. These international frameworks establish waste management as a critical component in achieving sustainable urban development.

Department of Urban Planning and Design, The University of Hong Kong, Hong Kong SAR, China. ✉e-mail: junma@hku.hk

While urban waste management represents a visible and immediate measure of service quality that residents encounter daily, current literature exhibits some research gaps. First, there is insufficient research on the accurate detection and classification of street waste presence patterns, as well as a limited understanding of their spatial distribution in urban environments. Second, despite the prominence of waste management in urban operations, limited empirical evidence exists on how varying forms of street waste presence influence residents' perception of safety, further leaving unexplored whether such dynamic management practices have comparable importance with static built environment characteristics as demonstrated in previous studies.

In this study, we aim to explore how the presence of various forms of street waste influences perceived safety and analyze the relative importance of different factors in shaping safety perception. We selected New York City as our study area, considering its representativeness as a major global metropolis with diverse urban landscapes and complex waste management issues[16]. To systematically address these objectives, we integrate street view imagery (SVI), computer vision techniques, and machine learning approaches. Specifically, our analysis comprises three main components: (1) We employ computer vision algorithms to quantify and map perceived safety across the urban landscape, establishing baseline spatial patterns of safety perception. (2) We develop deep learning models for detecting and classifying different types of street waste accumulation, enabling the systematic mapping and analysis of their spatial distribution patterns. (3) We investigate the relationship between waste presence and safety perception through multiple analytical approaches, examining the relative contribution of dynamic management practices and static built environment characteristics. Through explainable machine learning techniques, we further analyze the importance and directional effects of these environmental and management factors in shaping urban safety perception.

This study advances both theoretical understanding and methodological approaches while offering practical implications for urban management. From an academic perspective, our research makes two main contributions. Theoretically, through analyzing the associations between different types of street waste and safety perception, our study explores how street management services relate to community safety perceptions. This extends existing frameworks beyond static built environment characteristics to encompass dynamic management factors. Methodologically, we develop a novel analytical framework that integrates visual AI, computer vision technology, and explainable machine learning with SVI data. Through analyzing the associations between different types of street waste and safety perception, our study explores how street management services relate to community safety perceptions.

In terms of practical implications, our findings identify deficiencies in current waste management practices and their negative impacts on safety perception, providing evidence-based guidance for policymakers to develop targeted management strategies. Furthermore, this research highlights a broader insight into urban governance: the creation of safe and sustainable communities depends not only on initial planning and construction but also on the effectiveness of long-term management practices. This emphasizes the crucial role of sustainable urban management in shaping urban experiences.

Results

This study explores the relationships between different categories of urban street waste and safety perception, while examining waste presence as an indicator of dynamic urban management effectiveness in relation to broader factors shaping urban safety perceptions.

Our analysis unfolds in four interconnected stages. First, we demonstrate the performance of our computer vision model for safety perception calculation (Table 1), followed by mapping the predicted perceptions across NYC (Fig. 1) and presenting relevant statistical analyses. Second, we examine various categories of street waste, encompassing both controlled and uncontrolled waste (Fig. 3) while validating our computer vision model's capability in waste identification (Table 2). The spatial distribution and statistical characteristics of waste presence across the city are then presented (Fig. 4). Third, we investigate the statistical relationships between waste presence and safety perception, examining the correlation patterns and magnitude across different waste types (Fig. 5). Finally, we explore the relative importance of waste presence as a contributing factor to safety perception, identifying the dominant waste types that influence perceived safety. To investigate these dominant factors, we employ multiple analytical methods in the final section, utilizing explainable machine learning techniques alongside Class Activation Mapping (CAM) visualization (Figs. 6 and 7).

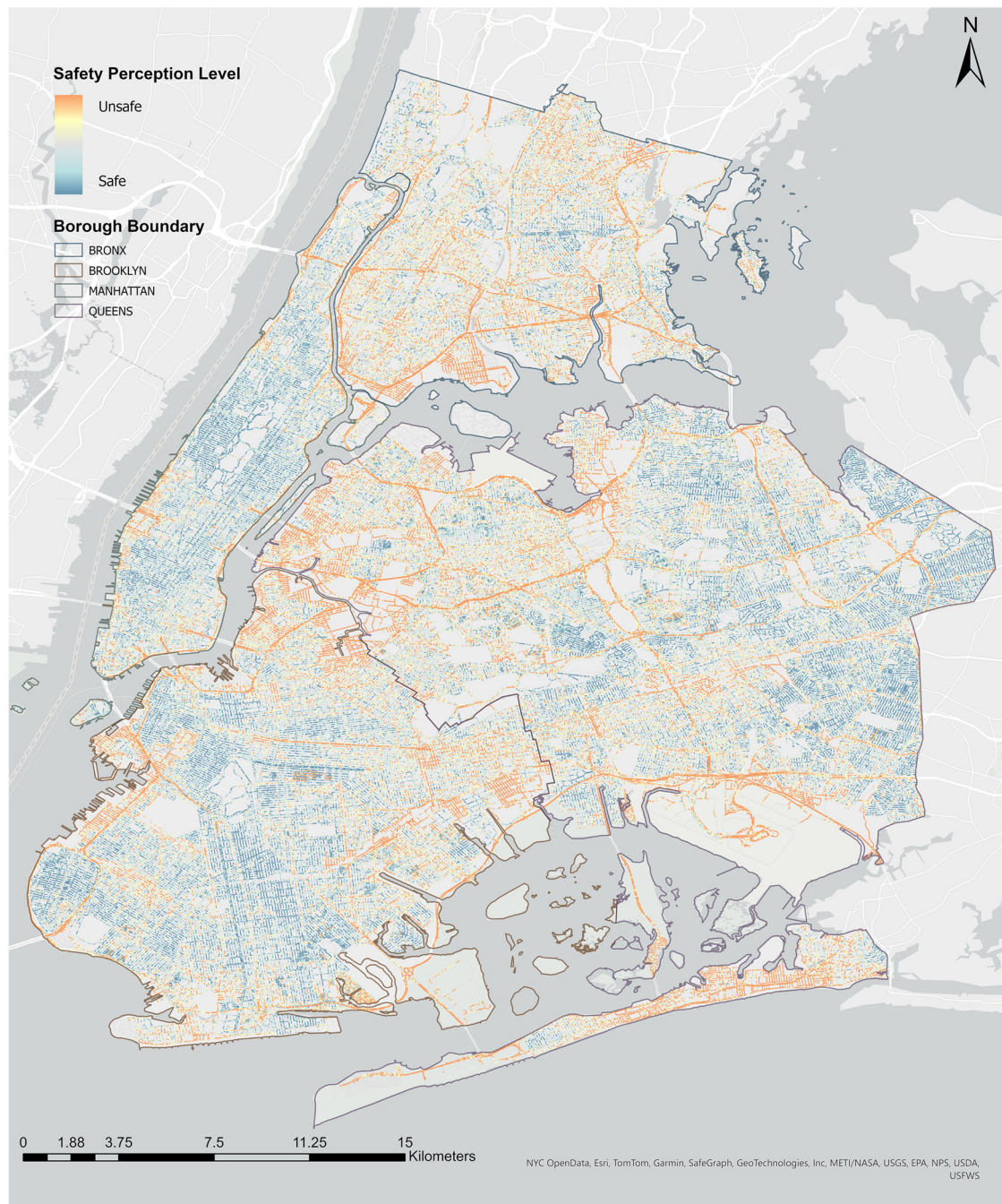### Urban safety perception: model performance, spatial distribution, and statistical analysis

Based on comprehensive experimental evaluations of four mainstream CNN architectures (Table 1), we selected ResNet-50 as our primary model architecture. ResNet-50 demonstrated superior overall performance with the highest accuracy (0.748) and consistently balanced metrics across safe and unsafe classifications (F1 scores of 0.746 and 0.750, respectively). While MobileNet-V2 achieved comparable accuracy (0.745), it showed notable disparities between safe and unsafe categories (precision: 0.719 vs 0.775; recall: 0.789 vs 0.702), indicating potential classification bias. EfficientNet-B0 and ShuffleNet-V2, despite their computational efficiency, exhibited lower overall accuracy (0.738 and 0.678) and less consistent performance across evaluation metrics. The balanced precision-recall trade-off and robust F1 scores of ResNet-50, combined with its well-established architecture in computer vision tasks, made it the optimal choice for our safety perception model.

The trained model was applied to infer safety perceptions from street-view images across the study area. We adopted a confidence-based scoring methodology, following established approaches in urban perception studies[17,18], to transform binary classifications into continuous safety scores. Specifically, locations classified as safe were assigned positive values while those classified as unsafe received negative values, with the magnitude proportional to the model's prediction confidence. This quantitative transformation enables more nuanced differentiation of safety perceptions, capturing subtle variations between locations that share the same binary classification but exhibit different degrees of perceived safety.

The spatial analysis of perceived safety across New York City reveals distinct geographical patterns. Our model-inferred safety perception scores, as visualized in Fig. 1, show a core-periphery pattern in each borough, particularly evident in Manhattan and Brooklyn, where central areas

Table 1 | Performance comparison of deep learning models for street-level safety perception assessment

Model	Precision		Recall		F1 score		Accuracy
	Safe	Unsafe	Safe	Unsafe	Safe	Unsafe	
ResNet-50	0.739	0.757	0.753	0.743	0.746	0.750	0.748
MobileNet-V2	0.719	0.775	0.789	0.702	0.753	0.737	0.745
EffcientNet-B0	0.723	0.754	0.757	0.720	0.740	0.736	0.738
ShuffleNet-V2	0.723	0.754	0.757	0.720	0.740	0.736	0.678

Fig. 1 | Spatial distribution of safety perception across New York City. The map shows perceived safety levels derived from street-level imagery analysis, revealing distinct patterns between central and peripheral areas. Safety scores are categorized from unsafe (orange) to safe (blue).

consistently exhibit higher safety perception compared to their peripheral counterparts. Notable concentrations of high safety perception are observed in Midtown Manhattan, central Brooklyn, and eastern Queens, as indicated by the pronounced blue regions in Fig. 1.

The relationship between safety perception and socioeconomic indicators (population density, education level, and income level, shown in Supplementary Figs. 2–4) exhibits complex spatial variations across boroughs. In Manhattan and Brooklyn, areas of high population density strongly correlate with elevated safety perception scores. However, this relationship does not persist uniformly across all boroughs. Eastern Queens presents a notable exception, displaying high safety perception scores despite relatively lower population density, suggesting that population density alone cannot fully explain the variations in perceived safety. Income

level distributions show strong spatial correspondence with safety perception patterns, most notably in Queens and the Bronx, where communities with higher income levels consistently report higher levels of perceived safety. Additionally, areas with higher educational attainment exhibit elevated safety perception scores across all boroughs, indicating a robust relationship between education level and perceived environmental safety.

Statistical analysis of the safety perception scores yields a mean value of −0.047, indicating a neutral perception baseline, with a standard deviation of 0.668. Based on these parameters, we established four distinct safety perception categories based on the statistical distribution of scores, using standard deviation (σ) thresholds from the mean ($\mu = -0.047$, $\sigma = 0.668$). Areas were classified as follows: low safety perception ($x < \mu - \sigma$, or $x < -0.715$), moderately low safety perception ($\mu - \sigma \leq x < \mu$, or

Table 2 | Performance metrics of waste classification models across different categories

	Precision	Recall	F1-score
(1) Controlled waste category: bagged waste			
Bagged waste	84.51%	75.23%	79.63%
Non-bagged waste	94.02%	96.13%	95.06%
Accuracy	92.01%		
(2) Uncontrolled waste category: widespread litter			
Widespread litter	82.53%	91.84%	86.92%
Non-widespread litter	97.42%	94.46%	95.92%
Accuray	93.17%		
(3) Uncontrolled waste category: uncontrolled litter dumpsite			
Litter dumpsite	38.42%	42.81%	40.52%
Non-litter dumpsite	94.87%	93.84%	94.35%
Accuracy	90.43%		
(4) Uncontrolled waste category: construction waste			
Construction waste	37.24%	43.52%	40.13%
Non-construction waste	98.53%	97.85%	98.19%
Accuracy	96.14%		

$-0.715 \leq x < -0.047$), moderately high safety perception ($\mu \leq x < \mu + \sigma$, or $-0.047 \leq x < 0.621$), and high safety perception ($x \geq \mu + \sigma$, or $x \geq 0.621$), where x represents the safety perception score. Representative street view images (SVIs) from each category are presented in Fig. 2.

Further analysis of SVI across these safety categories reveals distinct environmental characteristics. Areas classified as safe consistently exhibit well-maintained streetscapes featuring abundant greenery, organized parking infrastructure, and well-preserved building facades, as evidenced in Fig. 2. Conversely, areas categorized as unsafe frequently display vacant lots, active construction sites, and deteriorating infrastructure. The moderately safe category comprises areas with mixed urban features, characterized by intermediate levels of maintenance and organization that contribute to moderate safety perception scores.

Urban street waste: waste categories, model performance, and spatial distribution

In this study, we focus on street waste, defined as waste accumulation occurring on urban streets outside of designated waste containers. To minimize the impact of random noise on our research findings, we specifically concentrate on identifiable waste clusters while excluding randomly scattered single pieces of litter. Based on extensive observations of urban street conditions, we identified distinct patterns in waste manifestation across the urban landscape. These patterns vary significantly in their spatial distribution and formation mechanisms. Consequently, we categorized street waste into two primary classifications: controlled and uncontrolled waste.

Controlled waste refers to temporarily placed, properly contained waste (such as securely bagged garbage or systematically stacked recyclables) positioned at designated collection points along streets in accordance with municipal collection schedules and regulations (Fig. 3A). Uncontrolled waste refers to improperly disposed materials that deviate from municipal waste management guidelines, encompassing three distinct subtypes:

(1) Construction waste: this category includes improperly disposed construction materials, scattered demolition debris, and abandoned construction supplies that are not contained within designated storage areas or proper disposal containers. Such waste often accumulates beyond designated construction zones, encroaching upon public spaces and walkways, indicating inadequate construction site waste management practices and non-compliance with municipal waste disposal regulations. (Fig. 3B).

(2) Widespread litter: distinguished by dispersed debris covering significant surface areas without substantial vertical accumulation. While less severe than uncontrolled dumpsites, this category nevertheless indicates deficiencies in local waste management practices and represents a crucial category for urban waste management assessment (Fig. 3C).

(3) Uncontrolled litter dumpsites: characterized by substantial accumulations of mixed waste, predominantly composed of household litter but also containing other waste types. These sites typically indicate prolonged inadequacies in waste management practices (Fig. 3D).

Based on our waste categorization, we developed specialized deep learning models for each waste type. The waste identification models, implemented using the Swin Transformer architecture, demonstrated robust performance across all waste categories, validating our approach's effectiveness in real-world urban waste detection scenarios. The detailed performance metrics are presented in Table 2.

The model achieved notable accuracy across different waste categories, with particularly strong performance in controlled waste detection (92.01% accuracy for bagged waste) and widespread litter identification (93.17% accuracy). The latter demonstrated high precision in distinguishing between areas with and without widespread litter, effectively capturing varying degrees of litter presence. Overall model performance remained consistently strong across all waste categories, with accuracies ranging from 90.43 to 96.14%.

We observed relatively lower performance in detecting uncontrolled dumpsites and construction waste categories. This pattern largely reflects their natural occurrence patterns in urban environments, as these categories constitute relatively smaller proportions of our dataset (3.8% and 7.7% respectively). To address this class distribution characteristic, we implemented targeted data augmentation techniques, which helped maintain model performance while preserving the authentic representation of waste distribution patterns in urban settings. While acknowledging these performance limitations, we implemented rigorous quality control measures. During the inference phase, we conducted thorough manual verification of all positive detections to ensure classification accuracy. Although our approach may have resulted in conservative waste counts due to their inherent sparsity, all detected cases underwent careful manual verification to ensure high reliability. This verification process effectively minimizes potential classification bias, strengthening the robustness of our subsequent relationship analysis between waste presence and perceived safety, and providing a solid foundation for our analytical conclusions.

Utilizing our developed waste classification models, we systematically identified and mapped the spatial distribution of distinct waste categories across the study area, as visualized in Fig. 4. Our detection system revealed 2351 instances of bagged waste (controlled waste), 1771 cases of widespread litter, 614 uncontrolled litter dumpsites, and 358 locations with construction waste.

The spatial analysis revealed distinct distribution patterns between controlled and uncontrolled waste categories. Controlled waste, primarily represented by bagged waste, showed the highest concentration in Manhattan compared to other boroughs. In contrast, uncontrolled waste categories exhibited markedly different spatial patterns, with a notably lower presence in Manhattan and a significant concentration in the Rockaway Peninsula area of southern Queens. Among the uncontrolled waste categories, widespread litter emerged as the most prevalent issue, while construction waste showed a relatively modest presence throughout the city.

Relationship between safety perception and waste distribution

We analyzed the cumulative distribution of safety perception scores across different waste categories, as illustrated in Fig. 5. In our analysis, safety perception scores range from −1 to 1, where higher scores indicate greater perceived safety. The median score (at cumulative proportion = 0.5) indicates the safety perception value where 50% of locations within each category are concentrated, providing a representative measure of the overall

Fig. 2 | Representative street-level imagery showing the safety perception in New York City. The visual images typically range from areas perceived as unsafe (top row), often characterized by vacant lots and deteriorating infrastructure, through moderately unsafe and moderately safe neighborhoods (middle rows) generally showing mixed residential developments, to areas perceived as safe (bottom row), commonly distinguished by tree-lined streets and well-maintained residential environments.

safety perception level for that specific waste category. The analysis reveals distinct patterns in how safety perception varies between areas with and without waste presence, and among different waste types.
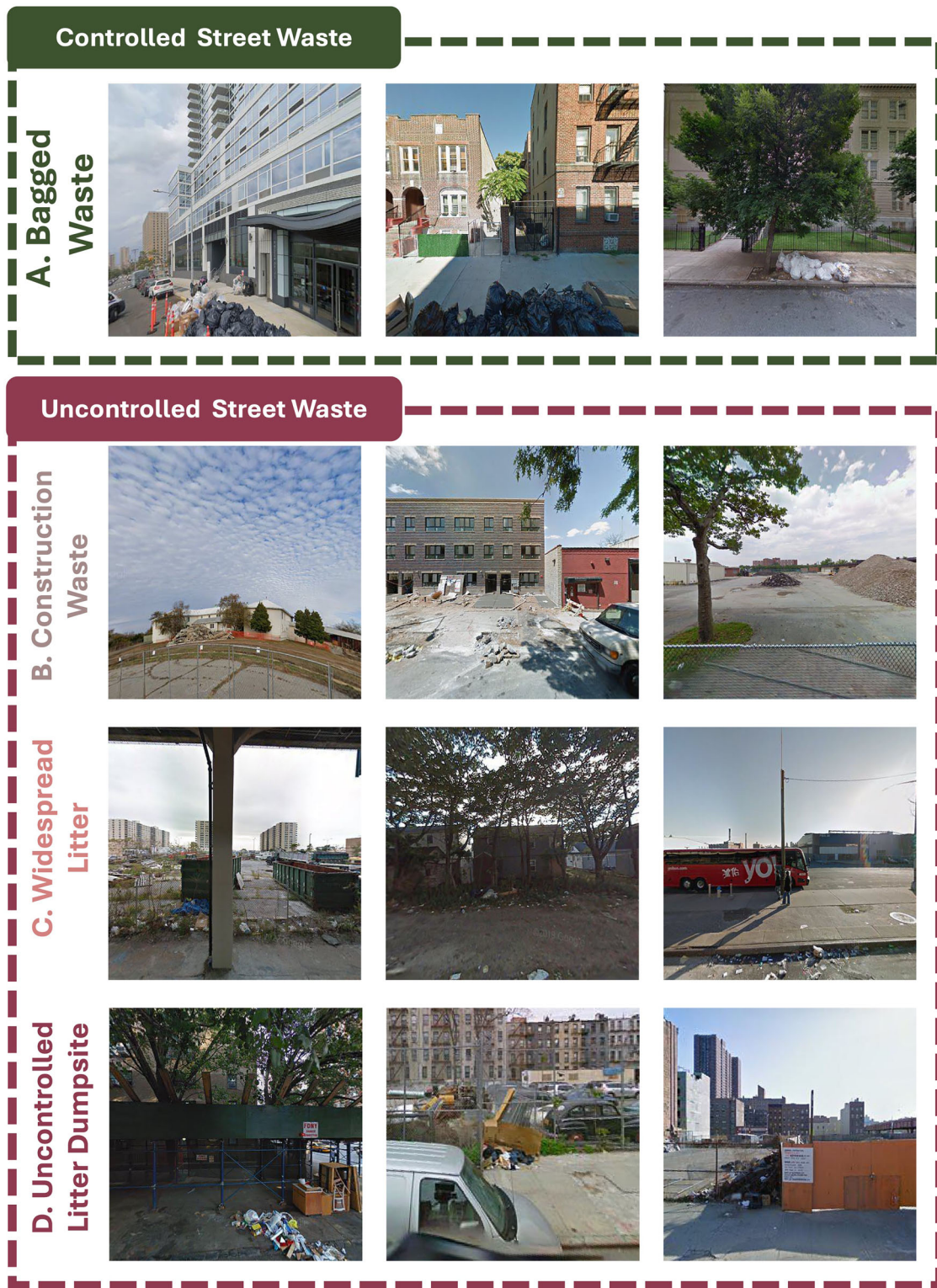
Out of the total 295,189 sampling points in NYC, only 4697 points (approximately 1.6%) contained waste in any form. The baseline distribution across all points in NYC (yellow line) shows that safety perception scores are approximately normally distributed, with a median score of −0.04, indicating a relatively neutral overall safety perception in the city. However, when examining areas where any type of waste is present (purple line), the distribution shifts notably toward lower safety scores, with the median dropping to −0.53, suggesting a substantial negative association between waste presence and perceived safety.

Further analysis of specific waste categories reveals a marked distinction between controlled and uncontrolled waste types. Controlled waste, represented by bagged waste (green line), shows a relatively modest negative association with safety perception, with a median score of −0.128. The gradual slope of its cumulative distribution curve indicates considerable variation in safety perceptions in areas with bagged waste, suggesting that its presence does not consistently correspond to negative safety perceptions.

In contrast, uncontrolled waste categories demonstrate a remarkably stronger negative relationship with perceived safety. Areas with construction waste, widespread litter, and uncontrolled litter dumpsites exhibit substantially lower median safety scores of −0.923, −0.921, and −0.896, respectively. The cumulative distribution curves for these uncontrolled waste categories display steeper slopes and closely aligned patterns, indicating a more consistent and pronounced negative relationship with safety perception. The similarity in both the median values and distribution patterns among uncontrolled waste types suggests that the presence of any form of uncontrolled waste corresponds strongly with reduced safety perception, regardless of the specific type.

Key factors shaping safety perception
To systematically investigate the underlying mechanisms driving these correlations, we employed two analytical methods to examine the relationships. For statistical analysis, we utilized explainable machine learning techniques to assess the relative importance and directional effects (positive or negative) of various environmental factors on safety perception. These factors encompass both static environmental characteristics, such as road
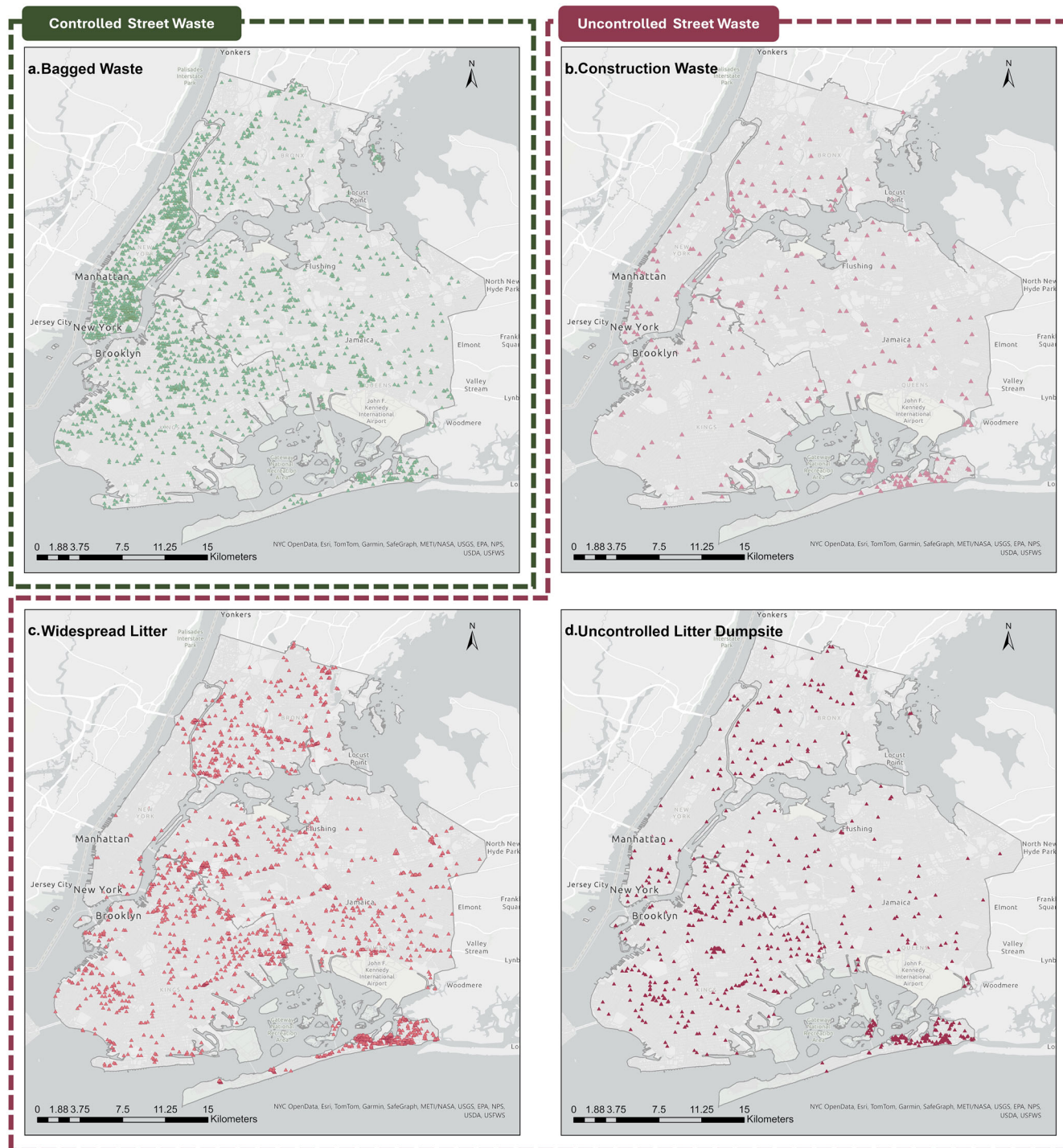
Fig. 3 | Representative street-level imagery depicting street waste morphologies in New York City. The images illustrate four distinct categories: **A** controlled waste areas, typically characterized by organized bagged waste awaiting collection; **B** construction waste zones, often marked by building materials and debris; **C** widespread litter areas, commonly showing scattered refuse across public spaces; and **D** uncontrolled dumping sites, frequently exhibiting accumulated waste in unauthorized locations. These patterns reflect varying levels of waste management practices and enforcement across urban neighborhoods.

and wall surface areas, and dynamic management indicators, such as waste presence, extracted from SVI. Additionally, we implemented CAM as a visual interpretation technique to identify and highlight specific regions within SVI that significantly influence the model's safety judgments. This visualization approach provides insights into the spatial attention patterns of the model, helping us understand how environmental features are weighted in the algorithmic assessment of safety perception.

To analyze the determinants of visual safety perception, we employed four regression models: Ordinary Least Squares (OLS), Random Forest, XGBoost, and Gradient Boosting Decision Tree (GBDT). These models

Fig. 4 | Spatial distribution of urban waste across New York City. The maps show four distinct waste categories: **a** controlled bagged waste; **b** construction waste sites; **c** widespread litter; and **d** uncontrolled dumping sites. Each point represents an observed instance identified through street-level imagery analysis.
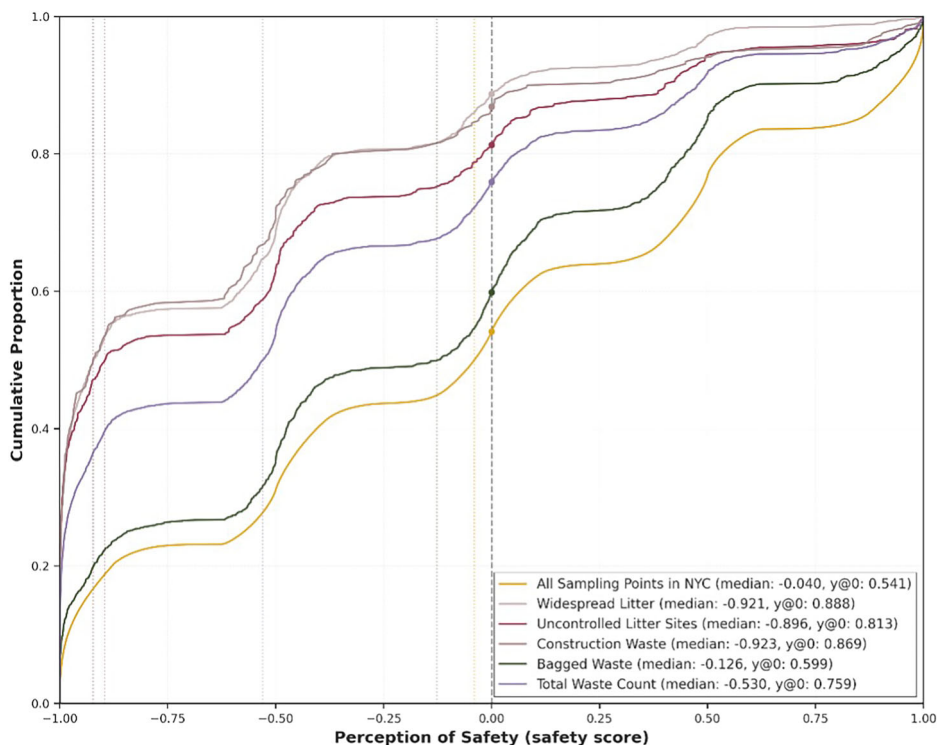
incorporated sociodemographic factors, visual environmental characteristics, and waste-related variables. To assess the specific impact of waste-related variables on model performance, we conducted parallel analyses with and without these variables for each model architecture. The comparative results are presented in Table 3.

The GBDT regression model demonstrated superior predictive performance, achieving the highest R^2 (69.47%) and lowest Mean Squared Error (0.101) among all tested models. Notably, the inclusion of waste-related variables consistently enhanced model performance across all architectures. Specifically, the GBDT model with waste-related variables

showed a 4.0% point improvement in R^2 compared to its counterpart without these variables (65.48%). This pattern of improvement was consistent across all models, with performance gains ranging from 3.5 to 6.0% points in R^2, underscoring the significant contribution of waste-related factors to safety perception prediction.

Based on these results, we selected the GBDT model for subsequent analysis. To further elucidate the complex relationships between environmental factors and safety perception, we employed SHapley Additive exPlanations (SHAP) value analysis. This approach enabled us to quantify both the relative importance and directional effects of individual

Fig. 5 | Cumulative distribution of safety perception scores across different waste categories in New York City. The plot shows the relationship between safety perception scores (−1 to 1) and cumulative proportions (0 to 1) for six waste-related categories. Vertical dashed line at $x = 0$ represents neutral safety perception, with negative values indicating perceived unsafe conditions and positive values indicating perceived safe conditions. Median values and intersection points at neutral perception (y@0) are shown for each category, demonstrating varying degrees of safety concerns associated with different waste types.

Table 3 | Comparison of model performance with and without waste-related variables

Model	Include waste-related variables	R^2	MSE
OLS Regression	Yes	58.51%	0.137
	No	52.54%	0.156
Random Forest Regression	Yes	68.09%	0.105
	No	64.03%	0.119
XGBoost Regression	Yes	69.11%	0.102
	No	65.65%	0.113
GBDT Regression	**Yes**	**69.47%**	**0.101**
	No	65.48%	0.114

environmental factors on safety perception, providing interpretable insights into the model's decision-making process.

The SHAP value analysis (Fig. 6) reveals the complex interplay between visual environmental features and sociodemographic characteristics in shaping visual safety perception. The results demonstrate diverse patterns of influence, varying in both magnitude and direction. Notably, while both physical environmental characteristics and socioeconomic factors significantly influence safety perception, their relative importance differs substantially.

Among the analyzed features, environmental elements demonstrate the strongest influence on safety perception. Sky visibility and tree coverage emerge as the most influential factors, exhibiting notable non-linear relationships with safety perception. Higher tree coverage is consistently associated with enhanced safety perception, suggesting that abundant urban vegetation contributes to a visual safety perception. Conversely, increased sky visibility correlates with decreased perceived safety, potentially indicating that more enclosed urban spaces, with limited sky exposure, are perceived as safer environments. This finding aligns with urban design theories about human-scale spaces and the role of natural surveillance in safety perception.

Notably, waste-related factors, as binary variables (0,1), demonstrate substantial negative impacts on safety perception. Both widespread litter and uncontrolled litter dumpsites show distinct binary distributions in their effects, confirming their categorical nature in the environment. Both widespread litter and uncontrolled litter dumpsites demonstrate strong negative relationships with safety perception. The magnitude of these effects positions waste-related factors among the top influential features, suggesting that waste management issues serve as powerful environmental cues for visual safety assessment. This finding emphasizes the critical role of municipal waste management in urban safety perception.

Built environment features and human activity indicators display moderate but consistent effects. Houses demonstrate positive associations with safety perception, while walls show a negative correlation. Notably, higher volumes of vehicles and pedestrians are associated with increased safety perception, suggesting that more developed urban environments with greater human activity generally foster stronger safety perception.
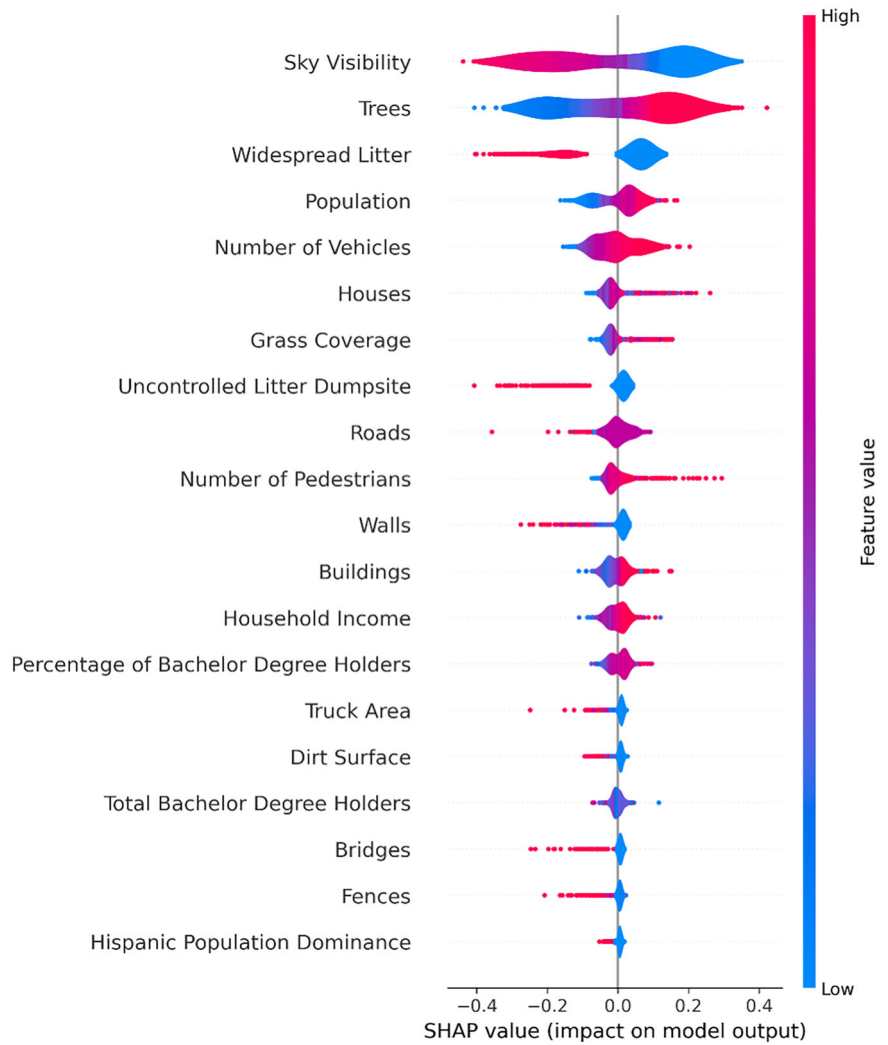
Regarding demographic and socioeconomic indicators, population density shows a positive correlation with safety perception, supporting the notion that more densely developed urban areas tend to evoke stronger perceptions of safety. Other socioeconomic indicators, including household income and educational attainment, show relatively modest influences on visual safety perception, suggesting their impact is less pronounced compared to physical environmental features.

These findings highlight the multifaceted nature of environmental safety perception, with particular emphasis on the significant impact of waste-related issues and urban vitality. The results suggest that urban visual safety enhancement strategies should prioritize effective waste management while maintaining active, well-populated spaces, alongside traditional urban design elements such as green space and built environment features.

To provide an interpretable visualization of our findings, we employed CAM to reveal the model's attention patterns in safety assessment. As illustrated in Fig. 7, the heat maps reveal regions of model attention through color intensity, with red areas indicating zones of highest attention. In images classified as unsafe, the model's attention predominantly

Fig. 6 | Feature importance analysis showing the impact of environmental elements on safety perception using SHAP values. The violin plots visualize both the distribution patterns and impact magnitude of each urban feature's contribution to safety perception. The width of each violin represents the density of observations at different contribution levels. Positive values (right side) indicate features that enhance perceived safety, while negative values (left side) represent features that diminish safety perception. The color gradient from blue to red corresponds to the magnitude of feature values, where blue indicates lower feature values and red signifies higher feature values.

concentrated on areas containing scattered litter, construction debris, and illegal dumping sites. This focused attention pattern suggests that the model identified these uncontrolled waste elements as key visual indicators of unsafe environments, aligning with human perceptual patterns.

In the analysis of scenes classified as safe, an intriguing pattern emerged: despite the presence of controlled waste (i.e., bagged waste) in some scenes, the model's attention primarily focused on surrounding architectural features rather than the waste itself. This selective attention suggests that controlled waste plays a less prominent role in safety perception, particularly when situated within well-maintained urban environments with established waste management systems.

These CAM-based visual interpretations provide strong supporting evidence for our findings: while uncontrolled waste serves as a primary visual indicator for unsafe environment perception, controlled waste demonstrates substantially less influence in shaping safety perceptions. This complementary relationship between the two analyses offers a more comprehensive understanding of how different waste management practices influence perceived safety in urban landscapes.

Discussion

Our study reveals three key findings that advance understanding of urban safety perception through the lens of environmental management. First, our analysis reveals strong associations between uncontrolled waste, particularly widespread litter, and perceived safety in urban environments, where areas with higher presence of litter correspond to lower safety perception scores. In contrast, controlled waste exhibits only a minimal influence on safety

perception, suggesting that properly managed waste systems have a limited negative impact on residents' perception of safety. Second, our analysis mapped the visual safety perception distribution, which reveals that safety perception is shaped by a complex interplay of spatial, social, and economic factors. Third, we identify distinct spatial patterns in waste distribution that reflect underlying urban socio-economic dynamics: controlled waste is predominantly concentrated in areas with established infrastructure and robust management systems, while uncontrolled waste is prevalent in peripheral locations with limited-service accessibility.

These findings provide novel insights into the relationship between environmental management and urban safety perception, particularly by quantifying the impacts of different forms of waste presence and accumulation in urban spaces. While previous research has predominantly focused on static physical features such as building design and street layout, our results demonstrate that dynamic environmental factors play a crucial role in shaping safety perceptions. Specifically, the stark contrast between controlled and uncontrolled waste's impact on safety perception (with bagged waste showing negligible impact while uncontrolled waste is strongly associated with negative safety perception) suggests that the manner of waste management, rather than the mere presence of waste, fundamentally influences community perceptions.

Our findings further reveal the dual importance of waste management and urban greening in shaping safety perceptions, with tree presence enhancing safety perception while uncontrolled waste diminishing it. To validate these statistical findings and explore potential interventions, we conducted simulation analyses using generative AI (detailed in

Fig. 7 | Visual attention analysis of urban safety perception using CAM heat maps. Street-level images and their corresponding attention maps illustrate how the model interprets environmental features associated with different safety perceptions. The upper panel shows scenes typically perceived as less safe, where attention often focuses on elements such as vacant lots, scattered debris, and deteriorating infrastructure. The lower panel displays environments generally perceived as safer, where the model typically highlights features such as well-maintained buildings, organized parking, and street trees. Heat map colors range from blue (low attention) to red (high attention), indicating regions most influential in the model's safety assessment.

Supplementary Note 3). These simulations demonstrate that combined interventions—removing uncontrolled waste and increasing tree coverage—can effectively transform environments with critically low safety perception scores into spaces with moderate safety ratings. This synergistic effect underscores the importance of implementing comprehensive environmental management strategies that address both waste control and urban greening simultaneously, rather than treating them as separate initiatives. Such integrated approaches offer promising pathways for improving perceived safety in urban environments, particularly in areas currently experiencing low safety perception scores.

This strong relationship between uncontrolled waste and visual safety perception aligns with insights from "broken windows" theory, which posits that visible signs of disorder can trigger a cascade of negative behavioral responses[19]. Our findings demonstrate that the presence of uncontrolled waste significantly correlates with decreased safety perception, suggesting its role as a visible indicator of environmental disorder. This relationship resonates with the theory's core premise about environmental cues and perception, though our study specifically focuses on the perceptual dimension rather than behavioral outcomes. While we cannot directly validate the full behavioral cascade proposed by the theory, our results highlight how visible environmental disorder, particularly uncontrolled waste, may

contribute to decreased perceptions of neighborhood safety, which could potentially create conditions conducive to further deterioration of community well-being.

Our findings have important implications for urban policy and management. First, our analysis reveals significant spatial disparities in waste management service distribution, with lower-income areas experiencing substantially reduced service coverage compared to more economically developed, densely populated districts. This service disparity significantly contributes to lower safety perceptions in marginal areas, suggesting an urgent need for more equitable distribution of urban environmental services. Second, while major infrastructure projects are important, our results reveal associations between dynamic management practices, particularly waste collection and vegetation maintenance, and urban safety perception. Our analysis shows correlations between well-maintained environments and higher safety perception scores, with waste-related features demonstrating significant relationships with how residents perceive safety in their neighborhoods. These findings contribute to our understanding of the relationship between environmental management practices and residents' perceptions of their urban environment.

Based on these findings, we recommend that city planners and policymakers: (1) implement systematic monitoring systems to identify and rapidly respond to uncontrolled waste hotspots, particularly in areas

characterized by persistent litter accumulation and uncontrolled domestic waste sites; (2) develop targeted programs to address service disparities; and (3) prioritize regular maintenance of green spaces and waste management services over large-scale infrastructure modifications. These strategies offer an efficient, equitable, and sustainable approach to enhancing urban safety perception while promoting environmental justice across all urban areas.

This study makes two contributions to urban research. First, we shift the paradigm from static infrastructure analysis to dynamic environmental management by demonstrating how day-to-day waste management practices fundamentally shape urban safety perceptions. This complementary perspective extends the traditional focus on fixed urban features and reveals that dynamic factors account for a substantial portion of variation in safety perception, addressing a critical gap in environmental psychology literature. Second, we establish an innovative methodological framework that leverages artificial intelligence for urban safety analysis. By integrating computer vision with explainable AI techniques (SHAP values and CAM), our approach provides comprehensive analytical insights into the relationship between environmental features and safety perception. This methodology not only offers a more nuanced understanding of urban safety dynamics but also demonstrates the potential of AI applications to transform urban research and practice.

Our study has one key limitation: the reliance on Google SVI provides only snapshots of urban conditions. While we ensured annual consistency in our analysis by using images from the same years, these single-time-point observations cannot fully capture the variations in waste distribution and their associations with safety perception across different times of day, weather conditions, and seasons.

Several temporal aspects warrant consideration. Our SVI collection period (2019–2021) coincided with the COVID-19 pandemic, which introduced unique temporal patterns. During this period, reduced human mobility likely altered waste generation and distribution patterns, while pandemic-related disruptions may have affected waste management services and collection frequencies. Seasonal variations could also influence waste visibility and detection. Winter snow coverage might partially obscure ground-level waste, especially widespread litter, while different seasonal lighting conditions and vegetation coverage could affect waste visibility. Furthermore, diurnal variations in waste patterns present another temporal consideration, as the timing of Street View image capture relative to scheduled waste collection could affect the observed presence of controlled waste, particularly bagged waste.

However, it's important to note that our methodology analyzes both waste presence and safety perception based on the same SVI. This simultaneous assessment means that while temporal variations might affect the absolute levels of observed waste or general safety perceptions, the relationship between these two factors within each image remains valid and reliable.

For future research aiming to capture more comprehensive temporal patterns of both waste distribution and safety perception, we suggest expanding beyond commercial SVI through a two-pronged approach. First, incorporating alternative data sources such as crowdsourced street-level imagery (e.g., Mapillary, OpenStreetCam) could provide broader temporal coverage and capture variations across different times of day, seasons, and urban conditions. Second, establishing focused monitoring programs in priority areas through systematic longitudinal data collection would enable a deeper understanding of persistent problem areas and the dynamic relationship between waste patterns and safety perception. This integrated approach would not only advance our understanding of urban safety dynamics but also support the development of more effective, evidence-based waste management strategies, contributing to the creation of sustainable and safe communities.

Beyond its immediate findings, our research has broader implications for urban sustainability and social equity. The relationship between environmental management and residents' management and residents' safety perception suggests that effective waste management strategies could simultaneously address multiple urban challenges, from how residents perceive their environment's safety to environmental quality. This understanding contributes to developing more sustainable and livable urban environments that foster community well-being and enhanced quality of life, particularly relevant as cities worldwide grapple with growing environmental and social challenges.

Methods
Study areas
New York City, one of the most densely populated and dynamic urban areas globally, is selected as the study site due to its diverse built environment, high pedestrian activity, and significant variation in neighborhood characteristics. To systematically examine the relationship between street waste distribution and safety perception, we focus on four boroughs: Manhattan, Brooklyn, Queens, and the Bronx. Staten Island is excluded from this study due to its lower population density, suburban land use patterns, and distinct waste management practices, which differ significantly from the denser and more urbanized boroughs[20].

The selected boroughs face substantial waste management challenges that may influence public perception of safety. Systematically, New York State generated 42.2 million tons of total waste in 2018, with New York City contributing 8.16 million tons that required processing or disposal[21]. At the operational level, service coverage disparities exist across neighborhoods, particularly in terms of collection frequency and street cleaning services, leading to uncontrolled waste issues and public complaints, with over 40,464,714 improper street waste-related complaints recorded in the city's 311 data from 2010 to 2025[22]. From a policy perspective, despite the Department of Sanitation of New York's expanding budget—which is projected to increase from $1.9 billion in the Fiscal 2025 Preliminary Financial Plan to $2 billion by the end of the plan period with an annual growth rate of 4.5%—the rising operational demands and cost pressures continue to necessitate strategic allocation of cleaning services[23]. These challenges underscore the importance of identifying critical locations for targeted intervention, particularly in areas where waste accumulation may impact perceived safety.
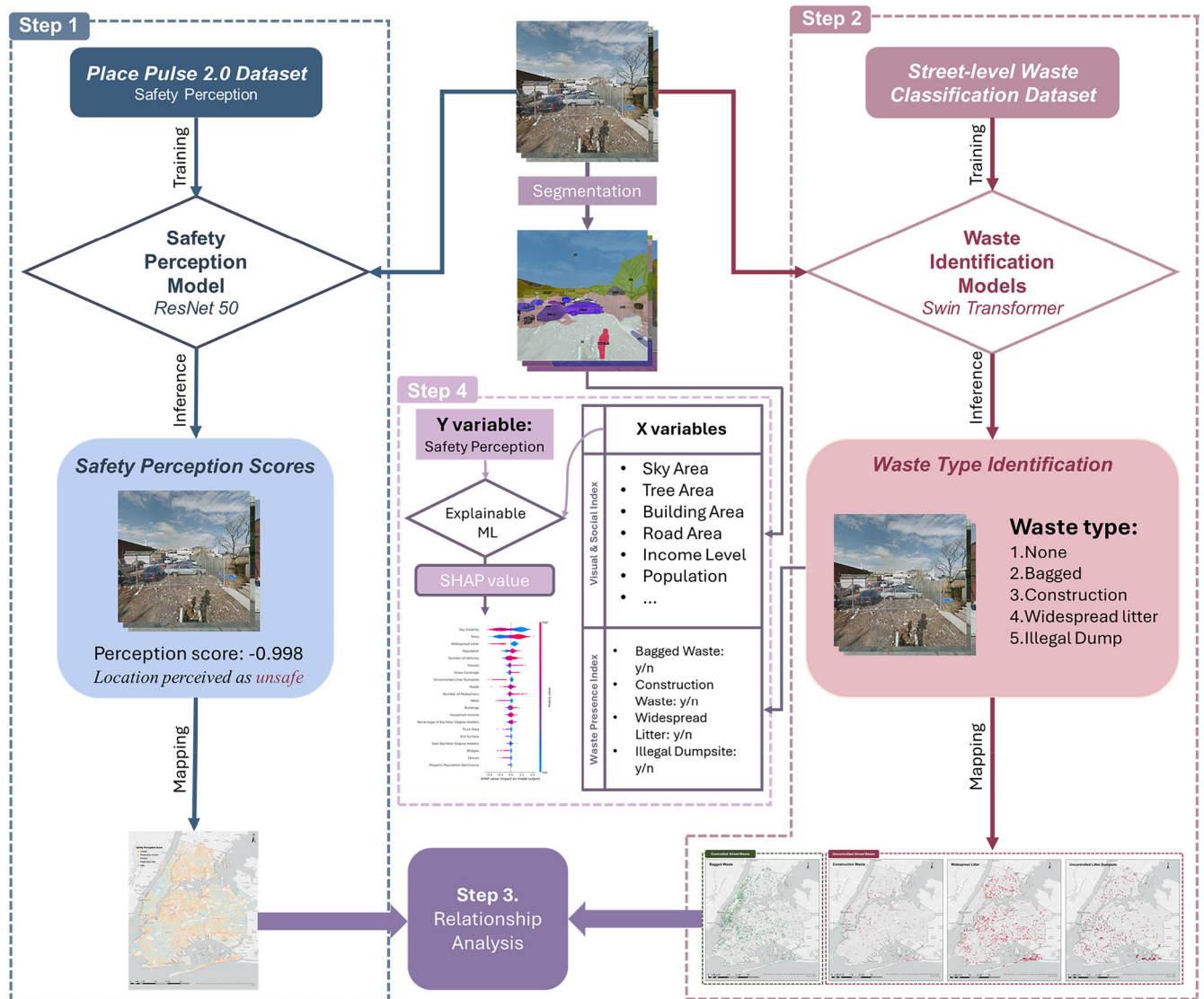
Research design
A comprehensive methodological framework was developed to address the problem, incorporating a systematic stepwise approach. As illustrated in Fig. 8, the methodology consists of four interconnected analytical phases:

Step 1. Safety perception distribution analysis: the Place Pulse 2.0 Dataset was utilized to train a deep learning model for safety perception score prediction. The trained model was subsequently employed to compute and spatially map safety perception scores across four boroughs of New York City.

Step 2. Waste location identification and mapping: using the Street-Level Waste Classification Dataset, we developed four specialized classification models to detect distinct waste categories: bagged waste, construction waste, accumulated litter dumpsites, and widespread litter. These models were subsequently deployed to analyze the comprehensive image collection spanning four boroughs, facilitating the spatial mapping of waste distribution. It is worth emphasizing that individual locations may exhibit multiple waste categories simultaneously.

Step 3. Correlation analysis: to examine the relationship between roadside waste presence and safety perception, cumulative proportion analyses were conducted. This analytical approach enabled the assessment of whether locations with waste presence exhibited disproportionately lower safety perception scores compared to the city-wide averages.

Step 4. Factor analysis: to investigate whether the presence of street-level waste significantly influences safety perception, we employed two complementary analytical approaches. For statistical assessment, we utilized explainable machine learning techniques to evaluate the relative importance of various factors and their positive or negative contributions to safety perception. The visualization analysis leveraged CAM to intuitively illustrate the key attention areas that the safety perception model focuses on when simulating human judgment processes.

Fig. 8 | Methodological framework for analyzing urban safety perception and waste patterns. The workflow comprises four main steps: Step 1 implements a ResNet-50-based safety perception model trained on Place Pulse 2.0 dataset to calculate safety perception score; Step 2 employs a Swin Transformer for street-level waste classification; Step 3 conducts spatial relationship analysis between safety and waste distributions; and Step 4 utilizes SHAP values to interpret feature importance.

Through the systematic implementation of these four analytical steps, this study establishes a comprehensive understanding of how controlled and uncontrolled roadside waste influences residents' safety perceptions in urban environments. This methodological framework enables a comprehensive interpretation of the relationship between waste presence and perceived safety.

Dataset

In this research, "Place Pulse 2.0 dataset", collected through an internet-based platform devised by MIT, was utilized to train the safety perception model. The platform displayed pairs of SVIs selected randomly from various international cities to the participants. These participants were prompted to either select one of two images or state if they viewed them as comparable, in response to evaluative queries such as "which place looks safer?" along with other attributes including "beautiful," "depressing," "lively," "wealthy," and "boring"[18,24]. The resulting dataset, which is downloadable from https://centerforcollectivelearning.org/urbanperception, encompasses the SVIs and a document detailing each image's ID, the posed question, and the participants' responses. This study specifically concentrated on analyzing safety perceptions derived from the query "which place looks safer?"

Following this focus, we developed a dedicated sub-database of SVIs for training safety perception prediction models. A comprehensive description of the sub-dataset and its characteristics is presented in Supplementary Note 1.

The road network data utilized in this study is sourced from the NYC Street Centerline dataset[25], which offers detailed and up-to-date information on street networks. To systematically sample the street environment, sampling points were generated along road segments at 50 m intervals, which has been widely adopted in street-level urban environment studies[18,26–28]. This sampling interval enables comprehensive coverage of the urban streetscape while maintaining computational efficiency. The temporal scope for SVI collection was set to 2019–2021. These spatial and temporal parameters were then utilized as input parameters for the Google Maps Platform API to retrieve street-level imagery at each sampling location (https://developers.google.com/maps). Four SVIs were captured at 90° horizontal intervals for comprehensive environmental coverage at each sampling point, with the vertical angle maintained at 0° to ensure a consistent viewing perspective. We used four directional images rather than panoramic views to avoid geometric distortions that could affect subsequent analyses, particularly in panoptic segmentation, where accurate area ratio calculations are

Table 4 | Distribution of waste categories in street-view images (*N* = 8142)

Waste category	Image count	Image percentage[a]
No waste	4071	50%
Controlled waste	1500	18.4%
Bagged waste	1500	18.4%
Uncontrolled waste	2714	33.3%
Construction waste	306	3.8%
Widespread litter	1912	23.5%
Uncontrolled litter dumpsite	629	7.7%
Total	8142	100%

[a]Total number of images = 8142. Each image can have multiple labels. Percentages are calculated based on the total number of images (*N* = 8142).

crucial. Additionally, this approach maintains consistency with the "Place Pulse 2.0" dataset's image perspective, which is essential for ensuring reliable perception evaluation results. This systematic sampling approach resulted in the successful collection of SVI for 295,189 sampling points in the study area, yielding a total of 1,180,756 images. This extensive image dataset was subsequently utilized for safety perception evaluation and roadside waste detection analyses.

To develop a comprehensive roadside waste classification model for New York City, we constructed a specialized street-view image dataset that distinguishes between controlled and uncontrolled waste patterns. This distinction is crucial for understanding spatial distribution patterns and supporting urban waste management strategies. While existing studies have developed street-view-based waste classification models, they primarily focus on developing regions with significantly different waste patterns from Western metropolitan areas. Given that urban waste manifestation patterns are heavily influenced by local waste management policies, cultural practices, and urban infrastructure, we recognized the need for a dataset specifically tailored to capture New York City's unique waste distribution characteristics.

Our dataset development followed a systematic two-phase approach. Initially, we collected 1,180,756 SVI across New York City. To efficiently identify waste-containing images from this extensive collection, we first employed an EfficientNet-based binary classification model pre-trained on the UrbanDumpSight dataset[12]. Despite this existing dataset's focus on Shenzhen dumpsites, the model's fundamental capability in detecting waste-related features proved effective for our initial screening. The preliminary model demonstrated robust performance metrics (precision: 0.93, recall: 0.97, F1-score: 0.95) on the UrbanDumpSight dataset and was optimized to prioritize recall, ensuring comprehensive capture of potential waste instances while accepting some false positives. This screening phase effectively reduced the candidate pool to 56,031 images (4.74% of the original dataset). This filtering significantly reduced the number of images requiring annotation while maintaining high sensitivity to waste-relevant scenes. The remaining filtered images then proceeded to detailed manual annotation.

The filtered images then underwent a detailed manual annotation process conducted by two trained annotators working independently. Our structured hierarchical annotation protocol involved a two-stage assessment: first verifying waste presence, then categorizing confirmed waste into four distinct categories—bagged waste (controlled waste), construction waste, widespread litter, and uncontrolled litter dumpsites. Multiple categories could be assigned when different types of waste co-existed in a single image. To ensure annotation reliability and consistency, we established a rigorous quality control procedure. The author reviewed all annotated data, focusing particularly on cases where the two annotators' assessments differed. When disagreements occurred, all team members jointly reviewed

these cases and discussed until reaching a consensus on the appropriate classification.

Through this systematic process, we identified 4071 images containing various types of waste. To create a balanced dataset, we randomly selected an equal number of verified waste-free street views, resulting in a final dataset of 8142 images. Some locations may contain multiple types of waste simultaneously, and in such cases, the corresponding images were labeled with all applicable waste categories. Table 4 illustrates the distribution of waste categories across the dataset, showing the frequency of different waste classifications. All annotations were systematically recorded in a structured database format across five binary classification fields: waste presence, bagged waste presence, construction waste presence, widespread litter presence, and uncontrolled litter dumpsite presence.

Safety perception analysis

This study implements a binary image classification model leveraging SVIs to quantify safety perception. The model generates confidence scores that are subsequently utilized to compute safety perception scores for each sampling point within the study area. The model development framework encompasses three key components: label classification schema in the training dataset, neural network architecture, and model evaluation, which are detailed in the subsequent sections.

Given that the original dataset comprises comparative assessments, a transformation methodology was implemented to convert the relative comparisons into absolute safety indices for individual images. This transformation enables each Street View Image to be assigned a standardized safety level score within the complete dataset. Following established methodologies from previous studies[17,18], we employed the Q-score metric to quantify absolute safety perception levels. The computational process consists of three main steps:

(1) Win-loss ratio computation. The initial step involves calculating the win ratios (W_i) and loss ratios (L_i) for each image i, expressed as:

$$W_i = \frac{w_i}{w_i + l_i + e_i}$$

$$L_i = \frac{l_i}{w_i + l_i + e_i} \tag{1}$$

Where w_i, l_i and e_i represent the number of instances where image i wins, loses, or equals its paired counterpart, respectively.

(2) Q-score calculation. The Q-score, normalized to a range of 0–10, is then computed using the following equation:

$$Q_i = \frac{10}{3}\left(W_i + \frac{1}{w_i}\sum_{j_1=1}^{w_i} W_{j1} - \frac{1}{l_i}\sum_{j_2=1}^{l_i} L_{j2} + 1\right) \tag{2}$$

Where j_1 and j_2 denote images lose to or win image i in the comparison.

(3) Binary classification threshold. To mitigate potential noise and subjective bias in human perception, we established robust classification thresholds for binary safety labels using the following criteria:

$$Q_{high} = \bar{Q} + \delta\sigma$$

$$Q_{low} = \bar{Q} - \delta\sigma$$

$$\text{Label} = \{\text{positive (safe), if } Q > Q_{high} \text{ negative (unsafe), if } Q < Q_{low}\} \tag{3}$$

To establish a robust classification framework, we implemented a thresholding mechanism using two boundaries: $\bar{Q} + \delta\sigma$ and $\bar{Q} - \delta\sigma$, where σ represents the standard deviation and δ controls the bandwidth of the gap between thresholds. This approach allows us to identify and remove noise samples lying between the thresholds, while retaining clear positive (labeled

as "1") and negative (labeled as "0") samples. To optimize the bandwidth parameter δ, we conducted extensive experiments examining the trade-off between noise reduction and maintaining sufficient training samples for model generalization[18].

We selected ResNet50 as our baseline model for parameter optimization experiments due to its proven stability and widespread adoption in computer vision tasks. To determine the optimal bandwidth parameter (δ), we conducted experiments with different δ values (0.5, 0.7, 1.0, 1.2, and 1.5) using a 7:3 train-test split ratio. As shown in Supplementary Fig. 1, there is a clear trade-off between sample size and model accuracy across different δ values. At $\delta = 0.5$, the dataset retained the largest number of samples (34,083 positive and 30,685 negative), but yielded lower accuracy. As δ increased, we observed a consistent decrease in sample size along with an improvement in model accuracy. The accuracy plateaued at $\delta = 1.2$ (with 11,004 positive and 11,371 negative samples) and showed only marginal improvement at $\delta = 1.5$, despite further reduction in sample size (8073 positive and 7673 negative samples).

Based on the experimental results (shown in Supplementary Fig. 1), we selected $\delta = 1.2$ as the optimal bandwidth parameter, which effectively balances noise minimization while preserving an adequate number of training samples to ensure model generalizability and performance. Following this process, the final dataset comprises 11,371 images classified as unsafe and 11,004 images classified as safe, providing a relatively balanced dataset for subsequent model development.

To develop an efficient and deployable safety perception model, we conducted a comparative study of four lightweight CNN architectures: ResNet50, MobileNetV2, EfficientNetB0, and ShuffleNetV2. The focus on lightweight architecture was motivated by the practical requirements of edge deployment, where computational resources and memory constraints are significant considerations. All models were pre-trained on ImageNet and fine-tuned on our safety perception dataset. The training process utilized the Adam optimizer with a learning rate of 0.0001 and a cross-entropy loss function. Models were trained for 20 epochs with a batch size of 32, using a 7:3 train-test split ratio to ensure robust evaluation of model performance. We modify the output layer into a binary classification head.

Accuracy is used as the evaluation metric, which is formulated as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

where TP, FP, FN, TN represent the number of true positive, false positive, false negative and true negative instances.

The trained model was deployed to evaluate safety perception across the study areas in New York City using the collected SVIs. For each location, safety perception was assessed through the following process: First, the model generates confidence scores for each SVI, indicating the probability of the image being classified as either safe or unsafe. To maintain directional consistency in the safety metrics, confidence scores for unsafe classifications were assigned negative values, while safe classifications retained positive values. The comprehensive safety perception score for each sampling point was computed by aggregating the model's predictions from all four cardinal directions where SVIs were captured. Specifically, the overall safety level (S) at location i can be expressed as:

$$S_i = \sum_{(j=1 to 4)} C_{ij} \quad (5)$$

where C_{ij} represents the confidence score of the j-th directional view at location i, with its sign determined by the classification outcome (positive for safe, negative for unsafe).

This aggregation method provides a holistic assessment of safety perception by considering the environmental characteristics from multiple viewpoints at each location, thereby capturing the complete visual context of the street environment.

Waste location identification and mapping

This section presents a framework for waste location identification and mapping, encompassing both the development of neural network architecture and subsequent inference processes. The proposed methodology addresses the challenges of accurate waste detection while ensuring practical applicability in real-world urban environments.

We implement a binary classification approach for waste detection, developing separate models for each of the four waste categories identified in the Street-Level Classification Dataset. The architecture employs Swin Transformer (base model, patch size 4, window size 7) as the backbone for image feature extraction and classification[29]. This architectural choice is motivated by the complex nature of waste presentation in urban environments, where traditional CNN models, despite their effectiveness in local feature detection, often fail to capture the contextual relationships between waste objects and their surroundings. The transformer-based approach enables more comprehensive environmental context modeling, facilitating more accurate waste classification decisions.

The Swin Transformer architecture, characterized by its hierarchical structure, efficiently models long-range dependencies while maintaining computational efficiency. We initialize the model using pretrained ImageNet weights and modify the final fully connected layer to accommodate binary classification tasks. To ensure consistent model performance, we implement a standardized preprocessing pipeline. The process begins with image resizing to 224×224 pixels, matching the Swin Transformer's input requirements. Subsequently, pixel normalization is performed using ImageNet mean and standard deviation values, facilitating improved model convergence.

The training process utilizes cross-entropy loss, optimized specifically for binary classification tasks. We employ the Adam optimizer with an initial learning rate of 0.0001, enabling adaptive parameter updates throughout the training process. To enhance convergence characteristics, a cosine learning rate scheduler is implemented, providing gradual learning rate reduction over the training period. Model selection is guided by accuracy metrics, with the highest-performing models retained for subsequent waste detection processes.

The inference and mapping phase implements a systematic approach to identify and geographically visualize waste presence across the study area. Each specialized waste classification model is applied to the image collection in the study area to detect potential waste instances. To ensure detection accuracy and mitigate the impact of misclassification on subsequent correlation analyses, we implemented a two-stage verification procedure. Initially, the trained models identify candidate images containing specific waste categories through positive predictions. These preliminary results then undergo manual verification to minimize false positives and enhance detection reliability. The geographical coordinates associated with each verified positive detection are extracted and recorded, enabling the spatial mapping of waste distribution across different categories.

Factor analysis

To investigate the relationship between waste presence and perceived safety in urban environments, we conduct a two-faceted analysis. The statistical analysis examines key factors shaping safety perception, with particular focus on identifying which waste categories most significantly influence perceived safety levels. Complementing this statistical approach, our visualization analysis employs advanced mapping techniques to illustrate the critical areas that inform safety perception decisions, providing interpretable evidence of how waste presence affects residents' sense of safety.

To systematically analyze the relationships between environmental features and safety perception in urban environments, we employed explainable machine learning as our primary statistical analysis approach. This approach enables the identification and quantification of how various environmental features influence residents' safety perceptions. Our analytical framework incorporates both static environmental elements, sociodemographic variables, and waste-related variables extracted from SVIs. The complete set of

Table 5 | Feature variables for urban safety perception analysis: built environmental indicators, waste indicators, and sociodemographic indicators

Category	Variables	Type
Built environmental view index-area	- Roads - Bridge - Fences - Pavements - Houses - Dirt - Skies - Buildings - Trees - Grasses - Walls - Wall bricks - Railroads - People	Continuous variable
Urban infrastructure and activity index-count	- People - Bicycles - Cars - Buses - Trucks - Traffic lights - Stop signs	Continuous variable
Sociodemographic index	- Population - Household income - POI density (100 m radius) - Number of residents with Bachelor's degrees - Percentage of residents with Bachelor's degrees	Continuous variable
	- Census blocks with Asian majority - Census blocks with Black majority - Census blocks with Hispanic majority - Census blocks with White majority	Dummy variable
Waste presence index	- Widespread litter - Uncontrolled litter dumpsites - Construction waste - Bagged waste	Dummy variable

feature variables utilized in the machine learning analysis is presented in Table 5. For feature extraction, we utilized a pre-trained Detectron2 panoptic segmentation model to identify and quantify static environmental elements from SVIs. A detailed description of the Detectorn2 panoptic model architecture and implementation is provided in Supplementary Note 2. The feature set was also supplemented with binary indicators representing the presence or absence of four distinct waste categories. These features constitute the input space for our analysis, with the binary safety perception score (safe/unsafe) serving as the target variable.

We evaluated the performance of four regression models: OLS, Random Forest, GBDT, and XGBoost. The model demonstrating superior performance was selected for subsequent variable interpretation. To interpret the model's decision-making process and quantify feature contributions, we utilized SHAP values[30]. This interpretability technique provides both global feature importance rankings and local explanations for individual predictions. SHAP values enable the determination of not only the magnitude of each feature's influence but also the directionality of their contributions to safety perception. This comprehensive analytical framework offers insights into how different environmental elements, including various waste types, collectively shape safety perception in urban spaces.

In addition to statistical modeling, we applied CAM as a visualization technique to interpret how the safety perception model simulates human judgment[31]. This visualization technique illuminates the areas within images that contribute most substantially to the classification decisions, providing insights into the model's attention mechanisms during the safety perception assessment process. CAM operates by highlighting regions in images that are most influential in the model's decision-making process. The technique extracts feature maps from the final convolutional layer and combines them with the weights from the fully connected layer to generate a heatmap. This heatmap effectively visualizes the model's focus, emphasizing regions that are particularly relevant to specific classification decisions.

In our implementation, the perception model utilizes the ResNet-50 architecture as the backbone. The feature maps from the final convolutional layer (layer 4) are selected for CAM generation, as this layer captures high-level semantic information crucial for identifying significant image regions. The process involves weighing these feature maps using the corresponding weights from the fully connected layer dedicated to safety perception classification. The CAM generation process encompasses several key computational steps. Initially, we extract the weights corresponding to the target class from the fully connected layer. These weights are then applied to the feature maps generated by layer 4 through a weighted summation process. The resulting activation map undergoes ReLU activation to eliminate negative values, ensuring focus on positively contributing features. The final visualization is achieved through normalization to a 0–1 range, followed by the application of a jet colormap for enhanced visual interpretation. The resulting heatmap is then resized to match the original image dimensions and overlaid on the input image, creating an interpretable visualization of the model's attention patterns.

This visualization approach enables comprehensive analysis of the model's decision-making process by revealing the global features and critical regions that influence safety perception assessment. The resulting visualizations provide valuable insights into the model's behavior and validate its ability to focus on relevant environmental features when evaluating urban safety.

Data availability
The urban perception data used in this study are available from the "Place Pulse 2.0" dataset (https://centerforcollectivelearning.org/urbanperception). Street-level imagery of New York City was obtained through the Google Maps Platform Street View Static API (https://developers.google.com/maps). The waste image classification dataset used in this study is available from the authors upon request.

References
1. Zhang, F., Fan, Z., Kang, Y., Hu, Y. & Ratti, C. "Perception bias": deciphering a mismatch between urban crime and perception of safety. *Landsc. Urban Plan*. **207**, 104003 (2021).
2. Ito, K., Kang, Y., Zhang, Y., Zhang, F. & Biljecki, F. Understanding urban perception with visual data: a systematic review. *Cities* **152**, 105169 (2024).
3. Lefosse, D. C. et al. Feeling nature: measuring perceptions of biophilia across global biomes using visual AI. *npj Urban Sustain.* **5**, 1–13 (2025).
4. Al-Zu'bi, M. & Radovic, V. *SDG11—Sustainable Cities and Communities: Towards Inclusive, Safe, and Resilient Settlements* (Emerald Publishing Limited, 2018).
5. Pascual, U. et al. Diverse values of nature for sustainability. *Nature* **620**, 813–823 (2023).
6. Kang, Y., Zhang, F., Gao, S., Peng, W. & Ratti, C. Human settlement value assessment from a place perspective: considering human dynamics and perceptions in house price modeling. *Cities* **118**, 103333 (2021).
7. Kang, Y., Kim, J., Park, J. & Lee, J. Assessment of perceived and physical walkability using street view images and deep learning technology. *ISPRS Int. J. Geo. Inf.* **12**, 186 (2023).
8. Jing, F. et al. Assessing the impact of street-view greenery on fear of neighborhood crime in Guangzhou, China. *Int. J. Environ. Res. Public. Health* **18**, 311 (2021).

9. Foroughi, M., de Andrade, B., Roders, A. P. & Wang, T. Public participation and consensus-building in urban planning from the lens of heritage planning: a systematic literature review. *Cities* **135**, 104235 (2023).

10. Ho, R. & Au, W. T. Scale development for environmental perception of public space. *Front. Psychol.* **11**, 596790 (2020).

11. He, Z., Wang, Z., Xie, Z., Wu, L. & Chen, Z. Multiscale analysis of the influence of street built environment on crime occurrence using street-view images. *Comput. Environ. Urban Syst.* **97**, 101865 (2022).

12. Zhang, S., Ma, J. & Jiang, F. Monitoring street-level improper dumpsites via a multi-modal and LLM-based framework. *Resour. Conserv. Recycl.* **218**, 108227 (2025).

13. Zhang, S. & Ma, J. CascadeDumpNet: enhancing open dumpsite detection through deep learning and AutoML integrated dual-stage approach using high-resolution satellite imagery. *Remote Sens. Environ.* **313**, 114349 (2024).

14. Lord, J. et al. Global potential for harvesting drinking water from air using solar energy. *Nature* **598**, 611–617 (2021).

15. Meara, J. G. et al. Global Surgery 2030: evidence and solutions for achieving health, welfare, and economic development. *Lancet* **386**, 569–624 (2015).

16. Kontokosta, C. E., Hong, B., Johnson, N. E. & Starobin, D. Using machine learning and small area estimation to predict building-level municipal solid waste generation in cities. *Comput. Environ. Urban Syst.* **70**, 151–162 (2018).

17. Yu, X., Ma, J., Tang, Y., Yang, T. & Jiang, F. Can we trust our eyes? Interpreting the misperception of road safety from street view images and deep learning. *Accid. Anal. Prev.* **197**, 107455 (2024).

18. Zhang, F. et al. Measuring human perceptions of a large-scale urban region using machine learning. *Landsc. Urban Plan.* **180**, 148–160 (2018).

19. Price, G. N. Broken windows and crime in development challenged urban areas: evidence from Jackson, Mississippi USA. *J. Dev. Areas* **50**, 209–220 (2016).

20. Papas, P. The crossroads of the middle colonies the people, society, and environment of Staten Island. in *That Ever Loyal Island: Staten Island and the American Revolution* 9–26 (New York University Press, 2009).

21. New York State Solid Waste Management Plan: Building the Circular Economy through Sustainable Materials Management. New York State Solid Waste Management Plan - NYSDEC https://dec.ny.gov/environmental-protection/waste-management/solid-waste-management-planning/nys (2023).

22. 311 Service Requests from 2010 to Present | NYC Open Data. https://data.cityofnewyork.us/Social-Services/311-Service-Requests-from-2010-to-Present/erm2-nwe9/about_data (2025).

23. Johnson, H. C., Dromm, H. D. & Reynoso, H. A. *THE COUNCIL OF THE CITY OF NEW YORK*. https://council.nyc.gov/budget/wp-content/uploads/sites/54/2021/05/DSNY.pdf (2025).

24. Dubey, A., Naik, N., Parikh, D., Raskar, R. & Hidalgo, C. A. Deep learning the city: quantifying urban perception at a global scale. in *Computer Vision—ECCV 2016* 196–212 (Springer, 2016).

25. NYC Street Centerline (CSCL) | NYC Open Data. https://data.cityofnewyork.us/City-Government/NYC-Street-Centerline-CSCL-/inkn-q76z/about_data (2025).

26. Huang, Y. et al. Comprehensive urban space representation with varying numbers of street-level images. *Comput. Environ. Urban Syst.* **106**, 102043 (2023).

27. Ma, X. et al. Measuring human perceptions of streetscapes to better inform urban renewal: a perspective of scene semantic parsing. *Cities* **110**, 103086 (2021).

28. Wang, L., Han, X., He, J. & Jung, T. Measuring residents' perceptions of city streets to inform better street planning through deep learning and space syntax. *ISPRS J. Photogramm. Remote Sens.* **190**, 215–230 (2022).

29. Liu, Z. et al. Swin transformer: hierarchical vision transformer using shifted windows. In *Proc. 2021 IEEE/CVF International Conference on Computer Vision (ICCV)* 9992–10002 (IEEE, 2021).

30. Lundberg, S. M. & Lee, S.-I. A unified approach to interpreting model predictions. in *Advances in Neural Information Processing Systems* 30 (Curran Associates, Inc., 2017).

31. Ayyar, M. P., Benois-Pineau, J. & Zemmari, A. A feature understanding method for explanation of image classification by convolutional neural networks. in *Explainable Deep Learning AI* (eds Benois-Pineau, J., Bourqui, R., Petkovic, D. & Quénot, G.) 79–96 (Academic Press, 2023).

Author contributions
S.Z. conceived the research idea, developed the computer vision models, conducted data analysis and visualization, and wrote the initial manuscript. X.Y. contributed to data collection and statistical analysis. J.M. supervised the research, provided critical feedback, refined the methodology, and revised the manuscript. All authors contributed to the interpretation of results, participated in manuscript revision, and approved the final version.

Competing interests
The authors declare no competing interests.

Declaration of generative AI and AI-assisted technologies in the writing process
During the preparation of this work, the author(s) used ChatGPT-4 to improve readability and language. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42949-025-00269-x.

Correspondence and requests for materials should be addressed to Jun Ma.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.