A Comparative Analysis of Perceptions of Insecurity in Milan and Beijing Metro Stations

Abstract. Metro stations, as essential public spaces, not only serve as vital transportation hubs but also form part of the broader built environment that shapes people's perceptions of insecurity. An important concern for passengers in these environments is safety, particularly in underground public space where the design and organization of the physical surroundings play a crucial role. Despite various modern renovations in older metro stations, newer stations are generally perceived as safer. To understand this discrepancy, this research compares how visual factors in the built environments of old and new metro stations influence people's perceptions of insecurity. By examining two cities—Milan and Beijing, which follow distinct urban development models—this research also explores how differences in urbanization processes affect the contrast between old and new stations. This research introduces a novel methodology for analyzing underground public space by integrating 360-degree image capture, an enhanced semantic segmentation process, and predictive modeling using XGBoost and SHAP to reveal the complex relationships between these visual factors and safety perceptions. The results indicate that while factors like artificial light, floor, and the presence of people are significant across all stations, certain factors are particularly influential in specific contexts— for example, exposed pipes are more negatively associated with safety perception in Beijing's old stations, and platform doors have a strongly positive effect in Milan's new stations. The findings provide valuable insights for guiding the modernization of metro stations in the future, and offering an innovative approach to studying underground public space.

Keywords: Metro station; Underground public space; Perceptions of insecurity; Visual factors; Machine learning models; XGBoost

1. Introduction

Public transport plays a crucial role in achieving the Global Goals for Sustainable Development (Tiwari and Phillip, 2021). Efficient public transport systems contribute to sustainable urban development models by providing accessible and affordable mobility options, which are essential for fostering economic growth and promoting social inclusion (Birch, 2016). In cities, public transport infrastructure serves not only as a means of mobility but also as part of the built environment, functioning as public spaces where the physical design and environmental conditions impact social interactions. As part of the built environment, safety in these systems is a significant concern. Previous research highlights that safety perceptions strongly influence individuals' willingness to use public transport (Ceccato and Loukaitou-Sideris, 2021).

Unlike other public transport infrastructure, such as track and road systems, station facilities are highly enclosed environments with controlled access points and a significant concentration of passengers (Ding and Hou, 2022). This concentrated spatial creates specific environmental and psychological conditions that directly impact passengers' sense of safety (Ingvardson and Nielsen, 2022). Metro stations, in particular, due to their highly enclosed nature, dense population, and integral role in daily urban commuting, have drawn attention in safety perceptions studies. Research in Stockholm confirms that passengers generally feel more insecurity in metro stations than while traveling on trains (Ceccato, 2013). Interestingly, despite undergoing extensive renovations to align with modern standards, older metro stations are still commonly perceived as less safe than newer stations. This persistent perceptions raises the question: why do older stations, despite upgrades, continue to evoke stronger feelings of insecurity? This research aims to explore the visual factors contributing to the disparity in safety perceptions between old and new metro stations, with the goal of informing strategies to improve safety perceptions, especially in older stations.

A major challenge is defining "old" and "new" metro stations, as urban development models vary between cities. These differences lead to varying timelines and criteria for classifying stations. To address this issue, the research focuses on two cities, Beijing and Milan, which represent contrasting models of urban development, but share a similar starting point in the mid-to-late 20th century and have continued to expand.

Milan, a major European city with a long history of urbanization, has seen its metro system develop steadily over the past century (Foot et al., 2011). Urbanization in Europe dates back to around 700 BC and accelerated during the Industrial Revolution of the 18th and 19th centuries (Antrop, 2004). In contrast, Beijing has experienced rapid urbanization in recent decades (Shao et al., 2022). Beijing's population has surged to over 20 million in 2023 (Liu et al., 2018), while Milan's remained under 2 million. These distinct urbanization processes offer a valuable context for exploring the differences in safety perceptions between old and new metro stations across varying development models.

The historical trends of the construction of the metro systems in Milan and Beijing reflects these differences in urban development models. As shown in Figure 1, while Milan's metro system shows a steady growth over the entire period, the number of metro stations in Beijing remains very low and relatively flat from 1964 to around 2000. After 2000, the number of metro stations begins to increase rapidly, causing Beijing's metro network to now far exceed Milan's in terms of the number of stations in operation.

This research specifically focuses on comparing how the built environment in old and new stations contributes to people's perceptions of insecurity. Despite the extensive use of Street View imagery and semantic segmentation in studies of human perceptions of the built environment, a significant gap remains in the analysis of underground public spaces, particularly in public transportation systems. Most existing image datasets and segment technology are focused on above-ground environments and are therefore inadequate for capturing the unique characteristics of underground metro stations. As a result, current methodologies are insufficient for effectively analyzing the built environment in these spaces.

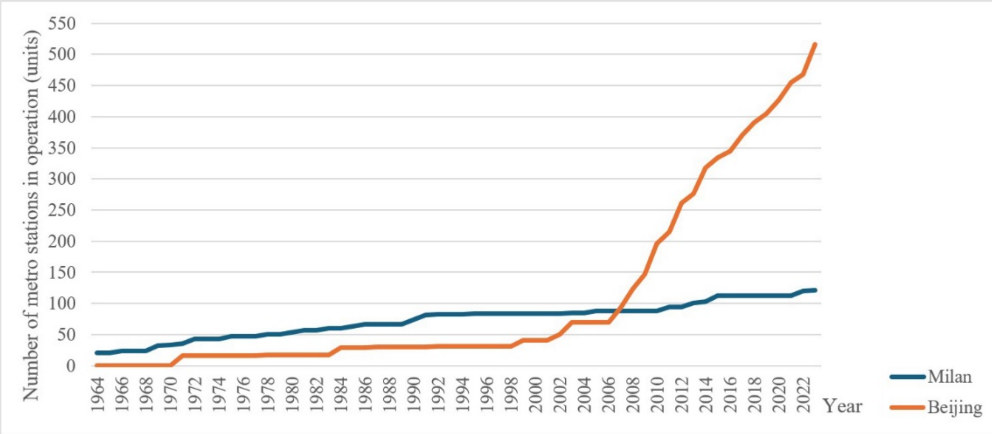


Figure 1. Comparative analysis of historical trends in metro stations in Milan and Beijing
Source: Authors' calculations based on data from The People's Government of Beijing Municipality (https://www.beijing.gov.cn/) and Milanese Transports Company JSC (https://www.atm.it/it/Pagine/default.aspx), 2024

To bridge this gap, this research proposes a novel framework to the analysis of underground public space. It employs a 360-degree mobile capture method and develops a unique dataset for segmenting and analyzing the interior factors of metro stations. Instead of traditional regression models, this research uses the XGBoost model to capture complex relationships between visual factors and safety perceptions, with SHAP explaining each feature's impact on safety perceptions.

By comparing the influence of visual factors associated with safety perceptions in old and new metro stations, this research aims to offer valuable insights into how the visual factors of built environment shapes safety perceptions. The findings will help inform targeted safety improvements and modernization efforts for older metro stations, while also presenting a new methodological approach for studying built environments in underground public space.

2. Literature Review

2.1 The relationship between built environment and people's insecurity

The notion of perceived insecurity encompasses a broad spectrum, spanning objective factors like regional crime and accident rates (Xu et al., 2018; Yue et al., 2022; Zhang et al., 2021), along with subjective factors such as personal evaluations (Dubey et al., 2016; Zhang et al., 2018). As with some other studies about people's perception, this research refers to a broader conception of perceived security intertwined with various human emotions (Cui et al., 2023; Dubey et al., 2016; Zhang et al., 2018).

The relationship between factors affecting people's insecurity and the built environment is a complex issue that has been explored in various studies. Firstly, the influence of the built environment on either promoting or preventing violent crime has been extensively studied (He et al., 2017). Research on the broken windows theory demonstrate that physical disorder is associated with crime in the environment (He et al., 2017). People, especially women, are afraid of potential crime, and certain environmental factors, such as poorly lit areas and isolated spaces, can increase their fear (Koskela and Pain, 2000). Secondly, the form of public spaces can exert a significant influence on people's perceptions of insecurity (Valentine, 1990). Removing barrier walls and eliminating graffiti as design interventions have an impact on safety perceptions (Navarrete-Hernandez et al., 2021). Thirdly, previous research has confirmed that the physical infrastructure in built environment could impact people's safety experience (Obiadi and Nzewi, 2018). Nesoff et al. (2018) created an inventory for pedestrian safety infrastructure to assess the pedestrian environment. Therefore, the built environment has a significant impact on people's feelings of insecurity (Sadeghi et al., 2023). Studying the correlation between people's insecurity and the built environment is valuable (Koskela and Pain, 2000).

In metro stations, which are unique public spaces within the urban built environment, safety perceptions are influenced by specific environmental. Unlike open public spaces, metro stations present unique challenges due to their enclosed nature and high-density usage (Ding and Hou, 2022). The spatial layout, visual accessibility, human scale, and the quality of station infrastructure are all crucial in shaping people's safety perceptions (Paydar et al., 2017; Sadeghi and Jangjoo, 2022; Sakip and Abdullah, 2012). Furthermore, security facilities such as surveillance cameras and security checks help enhance safety and reduce common concerns about enclosed transit spaces (Stjernborg, 2024). Analyzing the relationship between the built environment in metro stations and safety perceptions could provide valuable insights for developing safer and more user-friendly transit environments.

2.2 Components of perceptions of insecurity in built environment

Previous research has confirmed that various factors in the built environment influence people's safety perceptions when using public transport or being in public spaces, which have been identified and classified in Table 1. The authors' focus extends beyond public spaces within the metro system to include other areas such as train stations and urban streets, providing a comprehensive understanding of the factors influencing safety perceptions in public spaces.

Table 1
Components affecting people's safety perceptions in urban spaces

Space	Researcher	Influencing factors and classifications
Urban street	Park and Garcia, 2020	**Imageability** (historic buildings, identifiers, accessible open space) **Human scale** (street furniture, lights on buildings) **Transparency** (long sight line, first floor window, day & night activity) **Complexity** (pedestrian activity, business type & density, building color & design, outdoor dining, public art) **Tidiness** (ped. Infrastructure, traffic density & speed, noise & smells, sidewalk, litter & graffiti, landscape & vegetation)
	Cui et al., 2023	**Positive** (road, sidewalk, car, rail, person, skyscraper, fence, plam) **Negative** (house, truck, plant, bridge, path, sky, ground)

	Paydar et al., 2017	**Vitality through furniture arrangement** (few or lack of vacant houses, presence of large number/considerable number of benches along the walkways, presence of plazas with more number of people who sit there and along the walkways, often visiting the children activities in plazas and pathways) **Surveillance** (high possibility to be seen by others in the walkways, high possibility to be seen by others from their home, much number of people in the walkways, much number of shops beside the walkways, much traffic in the streets) **Diversity (**much presence of newness of the buildings and their façade, presence of much number of cars which are parked) **Signs of disorder** (much presence of walkways with good maintenance, little or lack of seeing the graffiti, little or lack of seeing the people who wander, little or lack of seeing the vacant lots, little or lack of seeing the litter on the floors) **Vegetation** (much presence of vegetation, much presence of trees)
All urban public spaces	Sadeghi and Jangjoo, 2022	**Accessibility and permeability** (access to public parking, access to public transportation, physical quality of sidewalks, access to the street through various routes) **Security** (legibility and clarity of urban space, visibility of space, night life of land uses and activities, the presence of other women in the urban space) **Land use and activity** (recreational and welfare land uses in space, ability to perform a variety of activities in space, mixed land uses, usability in different seasons and hours of the day) **Environmental and visual comfort** (the choice of sitting or walking in the shade or in the sun, environmental cleanliness and sanitation facilities avoiding all kinds of pollution, beautiful form and facade of buildings) **Facilities and services** (resting facilities in space, sufficient light and effective lighting, suitable urban furniture of the space, freedom to operate in space)
Metro stations	Ceccato et al., 2013	Crowded, lighting, distance from city centre, number of police stations, CCTVs, visibility, cross-sections / junctions / disruptions, escalators, social disorder, number of exits, seats/benches, foreign, number of platforms, pleasant atmosphere, number of ATMs within 100 m, dark corners, surveillance by others, presence of hiding places, physical deterioration, located underground, sunlight, social disorder, litter, graffiti, open exit, population density, owned housing, rain shield, net population
	Ceccato and Paz, 2017	Number of passengers / employee / security employees, Visibility from outdoors/indoors, Presence of bars / restaurants / shopping mall / ATM / bank / hiding corners/dark corners / cross section/junctions / litter/ physical deterioration / seats/ benches / civil guards / police
Railway Stations	Cozens et al., 2003	Lighting, surveillance camera, transparent shelter, staff, cleaner stations, longer trains, vegetation, underpasses, gangs, sign, frequent trains, help points, reliable service, coat of paint, clear information, access, links to community, open platform, people, spaces between seats, coordinated system, old buildings
	Coppola and Silvestri, 2021	**Main entrance** (greenery, road crossings, intermodal infrastructure, security personnel, surveillance cameras, commercial activities) **Internal lobby with the waiting room** (security personnel,

		surveillance cameras, commercial activities, artificial lighting, crowding) **Waiting areas near the platforms** (security personnel, surveillance cameras, commercial activities, crowding, tactile path and signage, decorum and maintenance)
All public transport	Stjernborg, 2024	**Desolate, drab, bare environments** (Lighting, seating, Distance between stops) **The social environment around stations and bus stops** (Intoxicated people, “Homeless” people, Men and young people) **Surveillance of stations and bus stops** (Guards and CCTV, Frequency of buses, Border controls, Police presence) **Surrounding transport environment at public transport nodes** (Buses, bicycles and cars with high speed, Buses, bicycles and cars which do not stop at pedestrian crossings)

Source: The authors, 2024

2.3 Prevailing research methodology

2.3.1 Data resources

Traditionally, evaluating urban perceptions has relied on conventional, time-consuming methods such as interviews and questionnaires (Yao et al., 2019; Dadvand et al., 2016; Paydar et al., 2017). However, the rapid expansion of geospatial big data from various sources, especially the abundance of large geo-tagged image datasets such as street-view and check-in images (Zhou et al., 2014), has opened new opportunities for evaluating human perceptions through visual surroundings. Street-view imagery has been used as a data source for research in urban perceptions (Guan et al., 2022). Meanwhile, semantic segmentation, a technique widely used across various fields involving image analysis, has been applied to areas such as Street View data to better understand perceptions of built environments. For example, Yao et al.(2019) employed street-view images and integrated semantic segmentation to develop a human-machine adversarial scoring framework. This approach incorporates deep learning and iterative feedback with recommendation scores, allowing for a rapid assessment of urban perceptions.

Despite the widespread use of Street View imagery and semantic segmentation for studying human perceptions, this approach is limited by the lack of image coverage for underground public spaces, particularly public transportation areas. Research on public transport safety perceptions has focused on methods such as surveys, interviews, virtual reality (VR), GIS, and crime data analysis. Cozens et al. (2003) utilizes VR technology to investigate passengers' perceptions of personal safety within and around the station environment. Ceccato et al. (2013) used crime and public disorder events across the entire Stockholm underground stations, employing GIS and spatial data analysis to assess the safety conditions. More recently, Joshi and Bailey (2023) used the story completion method to explore safety in Delhi metro system. Participants from diverse backgrounds in Delhi were given an online form with five fictional story beginnings to complete.

2.3.2 Analysis methodology

Traditional research methodologies, such as linear regression models, have been used to analyze the relationships between people's perceptions and built environment variables (Hong and Chen, 2014). However, these methods are limited in their ability to study complex relationships. Recent advances in machine learning and interpretation methods have significantly improved this capability (Rossetti et al., 2019). Machine learning techniques like eXtreme Gradient Boosting (XGBoost) not only excel at studying these complex relationships but also offer more reliable and accurate predictions, compared to traditional regression methods (Shehadeh et al., 2021; Xiao et al., 2021). Despite these advantages, the "black box" nature of machine learning models has faced criticism due to the lack of transparency in their decision-making processes.

Explainable Artificial Intelligence (XAI), a field within Artificial Intelligence (AI), aims to generate high-quality, interpretable, and intuitive explanations for AI models, providing a new pathway to enhance the interpretability of machine learning (Das and Rad, 2020). One notable XAI method is SHapley Additive exPlanations (SHAP), which uses Shapley values for model explainability, as introduced by Lundberg et al. (Lundberg and Lee, 2017), SHAP can identify variable importance in predictions and analyze non-linear relationships and interaction effects (Li et al., 2024).

2.4 Research gaps

2.4.1 Limitations in existing research frameworks

Despite the widespread use of Street View imagery and semantic segmentation in studies related to human perceptions of built environments, there remains a significant gap in the analysis of underground public spaces, particularly within public transportation systems. Most existing image datasets lack coverage of underground public spaces, and due to the unique environmental factors of metro stations, widely used semantic segmentation datasets, such as ADE20K, are unsuitable for accurately segmenting the distinct characteristics of underground metro stations. Consequently, current methodologies are insufficient for effectively analyzing the built environment within these spaces.

To bridge this gap, this research proposes a novel framework specifically to the analysis of underground public transportation environments. The authors employ a 360-degree mobile capture method and train a YOLOv8 model, a tool extensively used in semantic segmentation across various fields, to accurately segment and analyze the unique factors of underground metro stations. Rather than using traditional regression models to explore the relationship between visual factors and safety perceptions, the authors employ the XGBoost model to more accurately capture these complex relationships and use the SHAP package to explain the contribution of each factor to safety perception. Overall, this research proposes a novel methodology, offering a new approach for future studies on the built environment in underground public space.

2.4.2 Limitations in study area

Due to early construction, old metro systems are more likely to have potential safety risks (Bădău, 2022). Even if old metro systems have undergone many modernizations, it is widely recognized that the built environment in new stations generally provides a better experience than that in old stations. Therefore, studying the differences between old and new metro stations is meaningful and can provide strategies for future improvements. However, previous studies on safety perceptions in public transport have mainly concentrated on entire areas or specific areas within a city (Ceccato et al., 2013; Ceccato and Paz, 2017; Cozens et al., 2003a). These studies assessed all stations using a uniform standard, without considering that old and new stations vary significantly in terms of equipment, lighting, and spatial conditions, all of which directly influence passengers' sense of safety.

This research addresses this gap by comparing the factors in the built environment of new and old metro stations. By examining the differences between these two types of stations, this research highlights the evolving nature of public transport environments and their direct implications for safety. The findings can guide future upgrades for older stations and inform the planning and construction of new ones, addressing potential safety concerns more effectively.

3. Research Design and Methods

3.1 Research framework

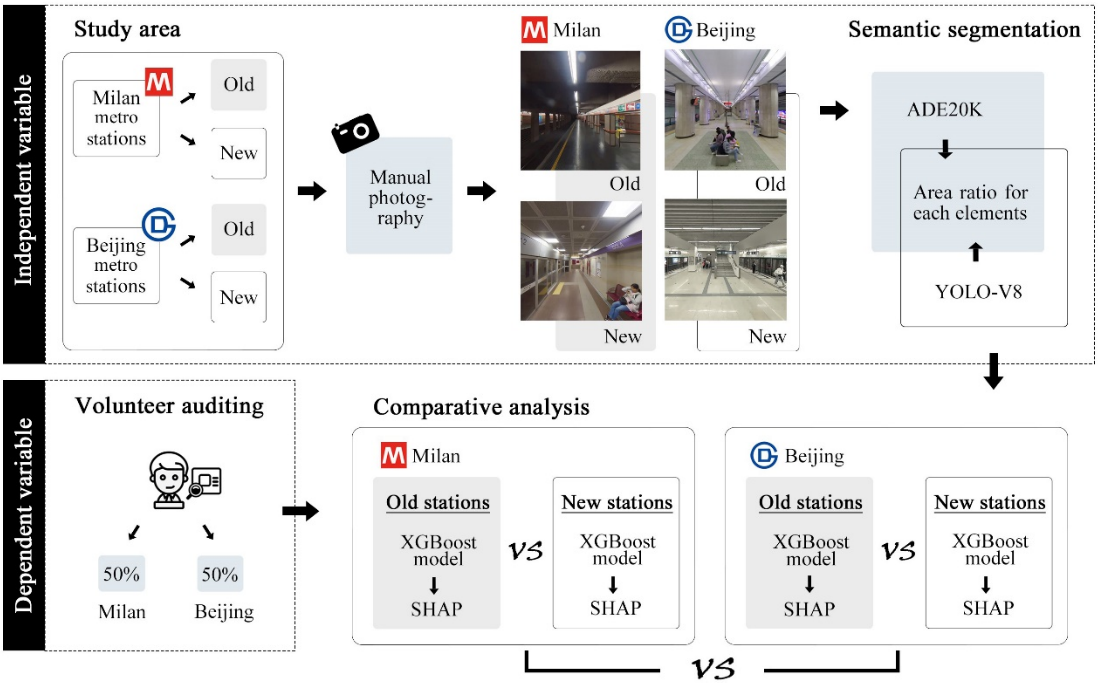


Figure 2. Research framework
Source: The authors, 2024

Figure 2 presents the framework developed to compare the visual factors in the built environment that contribute to feelings of insecurity when using the metro system in Milan and Beijing. The process begins with the identification of both old and new metro stations in these cities, followed by capturing images of these stations through manual photography. These images are then processed using semantic segmentation to calculate the area ratio for each element.

Due to the unique characteristics of metro station interiors, the accuracy of semantic segmentation is compromised when using open datasets like ADE20K. To address this, ADE20K is first used for initial segmentation. For factors that are either inaccurately segmented or unrecognized by ADE20K, the YOLOv8 model is employed for more precise segmentation. Meanwhile, a volunteer audit process is conducted to obtain perceptual safety perceptions scores for each image. Volunteers are evenly divided, with 50% originating from Milan and 50% from Beijing.

By integrating the data on area ratios of visual factors with perceptual scores from the volunteer audit, the XGBoost algorithm is utilized to establish predictive models for estimating perceptions scores. In both cities, XGBoost models are applied to both old and new stations, with the results interpreted using SHAP values to explain the contribution of each feature to the models' predictions. Within each city, the authors conduct a comparative analysis of old and new stations to identify potential improvements that could enhance safety. Finally, the authors broaden the analysis by comparing the metro systems of Milan and Beijing, offering insights into how development models influence the evolution and modernization of metro systems globally.

3.2 Study areas

3.2.1 Milan metro system

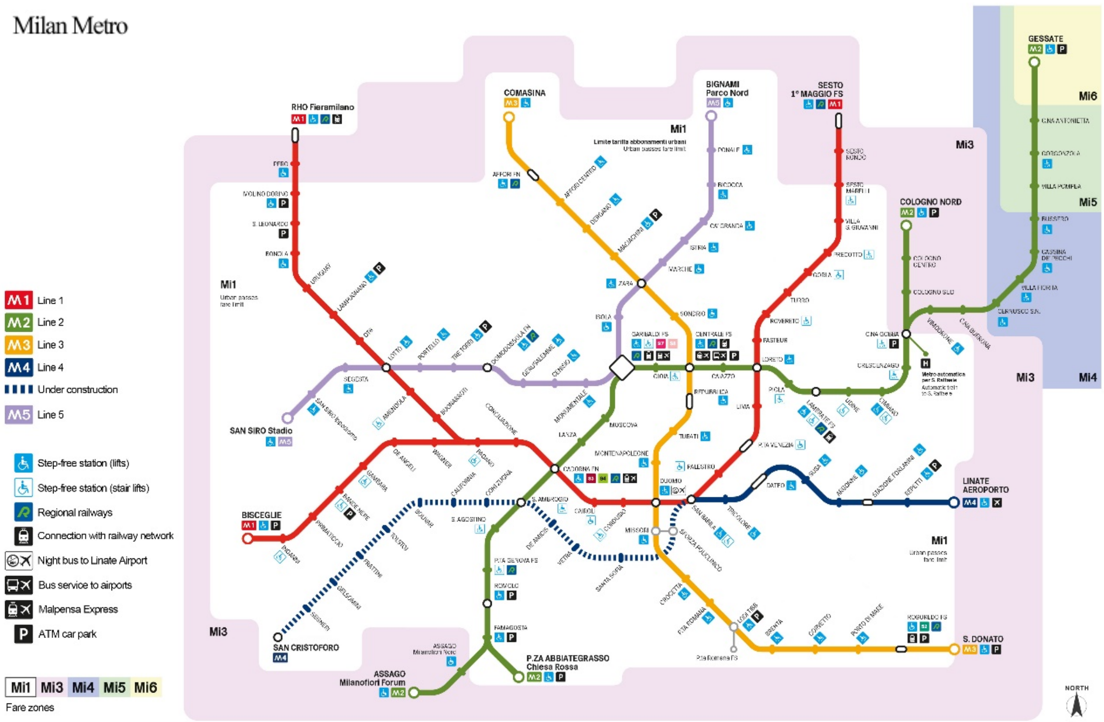


Figure 3. Milan metro map
Source: Milanese Transports Company JSC
(https://www.atm.it/it/Pagine/default.aspx), 2024

The Milan Metro system has steadily developed over a long period. Construction began in 1957, with the first line opening in 1964. Throughout the 1960s and 1970s, the network expanded to two lines. In 1990, the third line began opening in stages, with significant extensions completed in the early 2000s. The expansion continued into the 2010s, with Line 5 opening in multiple stages from 2013 to 2015 (ATM, 2019). After 2020, the first section of Line 4 was inaugurated in 2022, followed by an extension in 2023 (ATM Group, 2023). Currently, the system comprises five lines and a total of 113 stations, which includes 8 interchange stations (Figure 3).

3.2.2 Beijing metro system

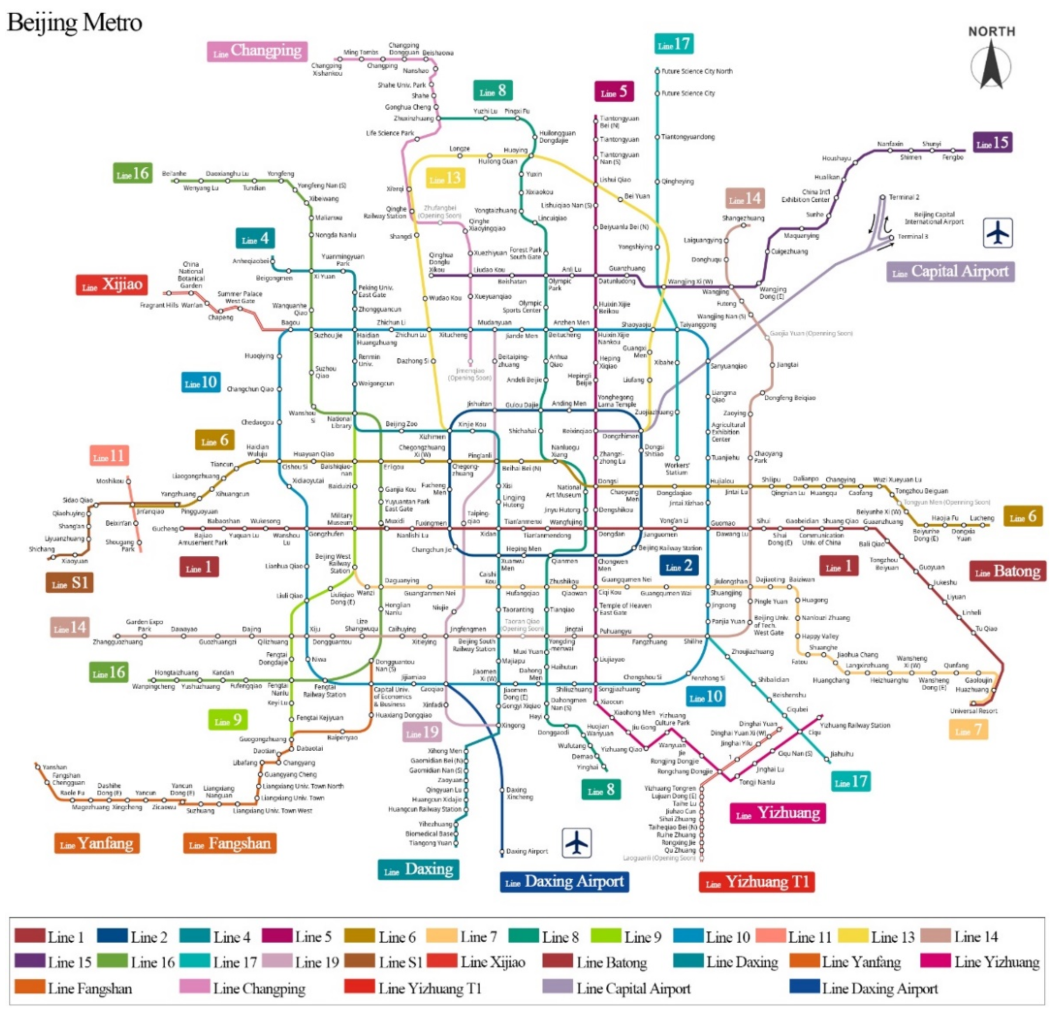


Figure 4. Beijing metro map

Source: Beijing Subway (https://www.bjsubway.com/), 2024

The Beijing Subway first opened in 1971 (Wang et al., 2009). Before 2000, the city's underground construction progressed slowly, with only two lines in operation (Guerra, 2015). However, after Beijing won the bid to host the 2008 Summer Olympics in 2001, the city accelerated its subway expansion plans. Six new metro lines were built and opened. The pace of construction further increased following the Olympics. By 2023, the Beijing Metro had grown to 27 lines, encompassing a total operational mileage of 836 kilometers and 490 stations, including 83 interchange stations (The People's Goveronment of Beijing Municipality, 2023). Figure 4 shows the latest metro map.

3.2.3 Criteria of station selection

In this research, specific metro stations in Milan and Beijing are selected and categorized into two groups: old stations and new stations. The following criteria are considered:

Firstly, the research focuses exclusively on underground stations, as their built environments differ from above-ground stations. Above-ground stations rely on natural light and often have better accessibility due to integration with existing infrastructure. Since lighting and accessibility significantly affect the sense of security, this research is confined to underground stations

Secondly, interchanges at intersections of different metro lines are treated as separate stations because they constructed in different periods. For example, in Milan, the San Babila interchange, where lines 1 and 4 intersect, is considered two distinct stations. The line 1 station is categorized as "old" due to its earlier construction, while the line 4 station is classified as "new" because it was built more recently.

Thirdly, the authors employ different criteria to define new and old metro stations in Milan and Beijing. Due to the distinct development models of the two cities, the conditions of their metro stations vary significantly. In Milan, the majority of metro stations are constructed before the 21st century, while Beijing's metro system experiences limited growth prior to the 21st century, followed by rapid expansion in recent decades.

Due to these differences, this research employs two distinct criteria. The first criterion is temporal, where stations built before 2000 are classified as old, and those constructed from 2014 to 2023 (the least decade) are considered new. The second criterion is based on a percentage-based approach, where a certain percentage of the earliest constructed stations are categorized as old, and the same percentage of the most recent stations as new. This dual approach enables a nuanced comparison, ensuring that the authors consider both the historical evolution and the pace of development in shaping the safety perceptions of metro users in these two cities.

3.2.4 Outcomes of station selection

The outcomes of the station selection under two criteria are presented in Table 2.

Table 2
Summary of selected old and new metro stations in Milan and Beijing

Criteria	City	Old stations	New stations
Criteria 1	Milan	Year: Before 2000 Sample Size: 42 (Out of 84 stations built before 2000, 50% are selected)	Year: 2014-2023 Sample Size: 20
	Beijing	Year: Before 2000 Sample Size: 40	Year: 2014-2023 Sample Size: 48 (Out of 240 stations built between 2014 and 2023, 20% are selected)

Criteria 2	Milan	Percentage: Earliest 30% (Year range: 1964-1972) Sample Size: 36	Percentage: Latest 30% (Year range: 2005-2023) Sample Size: 36
	Beijing	Percentage: Earliest 8% (Year range: 1971-2000) Sample Size: 40	Percentage: Latest 8% (Year range: 2021-2023) Sample Size: 40

Source: The authors, 2024

It is important to note that different percentages are employed to classify the old and new metro stations in two cities. As for criteria 1, 50% of Milan's old stations and 20% of Beijing's new stations are selected. This discrepancy arises due to the substantial number of metro stations constructed in Milan before 2000 and Beijing's rapid expansion after 2014. To balance the sample size across both cities, different percentages are utilized. For Criteria 2, 30% of Milan’s new stations, built since 2005, are selected to ensure a sufficient sample size. In contrast, only 8% of Beijing’s old stations (1971-2000) are included to avoid overlapping with the classification of new stations. This is necessary because all Beijing stations built between 2001 and 2006 are above ground, and selecting more than 8% would include post-2007 stations, conflicting with Milan’s classification.

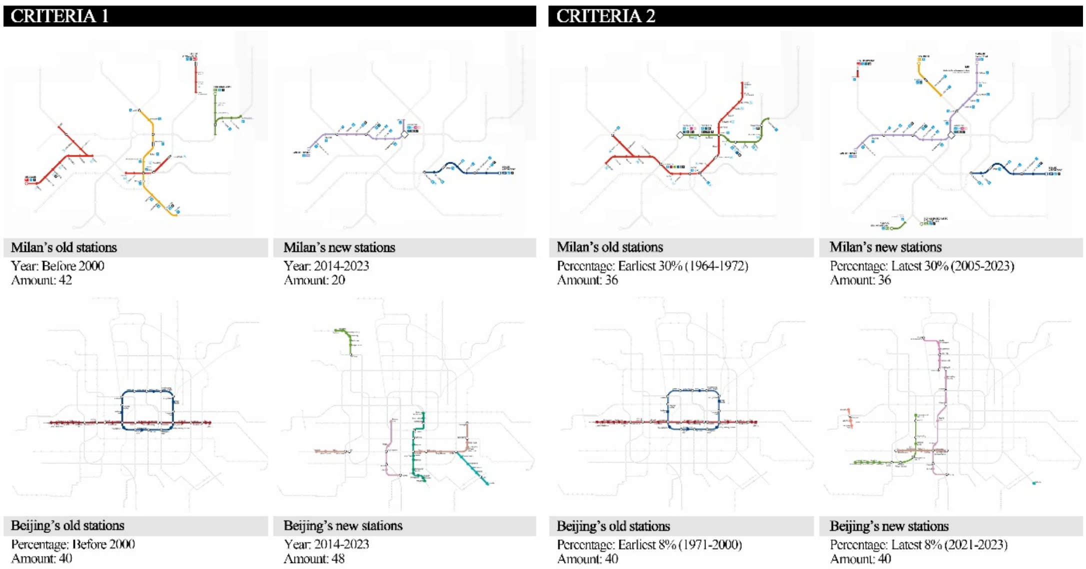


Figure 5. Geographical distribution of selected old and new metro stations in Milan and Beijing (both Criteria 1 and 2)

Source: The authors, 2024

Figure 5 presents a visual representation of the selected stations on the maps of Milan and Beijing. Notably, the selections of Beijing’s old stations are consistent under both Criteria 1 and Criteria 2, as both criteria limit the old stations in Beijing to those constructed before 2000.

3.3 Data collection

3.3.1 Image collection

The authors employ Insta360 X3 to manually capture 360-degree photographs of selected metro stations, focusing on key areas like security checks, exits, ticket gates, platforms, and interchange passages—all of them are critical areas for passengers. A total of 1186 images are collected, with 563 from Milan and 623 from Beijing. However, compared to the distorted equirectangular projection of a 360 panorama, perspective views are more familiar for most people to understand. Therefore, the authors transform the 360° panorama photographs into four 90-degree perspective photographs (Figure 6). After removing invalid images, the authors obtain 1324 images from Milan and 1194 images from Beijing.

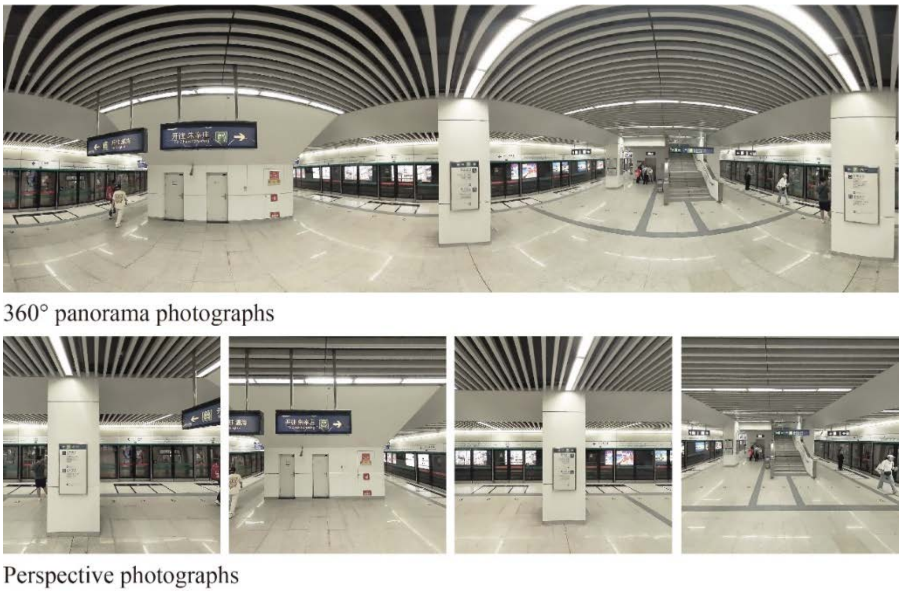


Figure 6. Transformation of 360° panorama photographs into 90-degree perspective views
Source: The authors, 2024

3.3.2 Identification and classification of principal component in metro stations

Building on previous research and considering the physical environment of metro stations in Milan and Beijing, the authors select identifiable visual factors as representative factors and categorize them into distinct groups (Table 3).

Table 3
Principal component analysis of built environment in metro stations

Classification	Factors	Segmentation	Sources
Visual accessibility	Light	Artificial light	(Coppola and Silvestri, 2021; Sadeghi and Jangjoo, 2022; Stjernborg, 2024)
	Enclosure	Floor	
		Wall	
		Column	
	Equipment	Platform doors	
		Stairs	
		Railing	
Surveillance	Passengers	People	(Ceccato and Paz, 2017; Coppola and Silvestri, 2021; Cui et al., 2023; Paydar et al., 2017; Sakip and Abdullah, 2012; Stjernborg, 2024)
	Monitor	Surveillance camera	
	Security	Security booth	
		Security check	
	Window	Visual window	
Vitality	Passengers	People	(Ceccato and Paz, 2017; Cui et al., 2023; Paydar et al., 2017; Stjernborg, 2024)
	Storefronts	Store / Vending machine	
	Signboards	Billboards / Signage	

	Graffiti	Graffiti	
Disorder	Broken	Broken pavement Broken ceilings Broken walls	(Ceccato and Paz, 2017; Park and Garcia, 2020; Paydar et al., 2017)
	Litter	Scattered litter	
	Equipment	Exposed pipes	
	Graffiti	Graffiti	

Source: The authors, 2024

It is important to note that due to the distinct characteristics of the built environments in Milan’s and Beijing’s metro stations, the presence and form of these components varies between the two cities. For instance, ‘Graffiti’ is only present in Milan’s stations, whereas ‘Security equipment’ is much more prevalent in Beijing’s stations. Milan's old station lacks platform gates, whereas both the old and new stations in Beijing are equipped with them. Moreover, even when some components can be found in both cities, they sometimes manifest in very different forms. For example, ‘Exposed pipes’ in Beijing are highly conspicuous and contribute to the visual clutter of the space, whereas in Milan, they are more likely to be concealed between beams or designed to resemble handrails. Figure 7 illustrates these differences, and the authors conjecture that the same component may have different impacts on safety perceptions in these two cities. Consequently, the authors decide to extract and analyze principal components from these two cities separately.

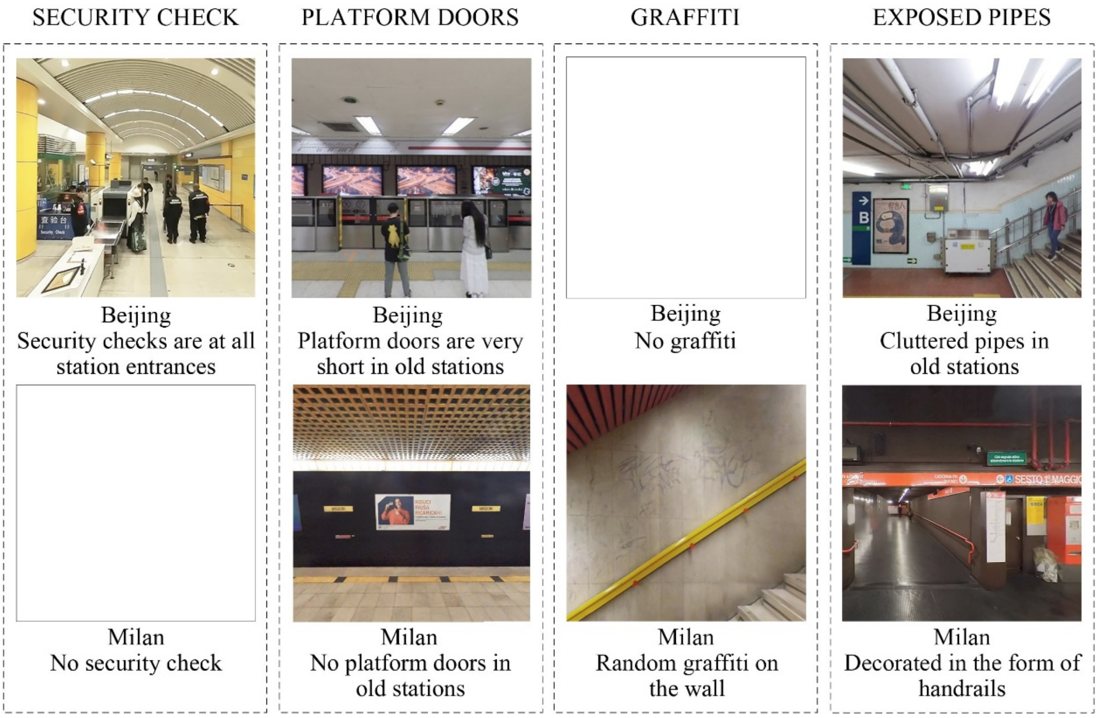


Figure 7. Comparison of the different built environment features in metro stations in Milan and Beijing

Source: The authors, 2024

3.3.3 Segmentation of principal component

To segment the principal components in metro stations from perspective photographs, the authors employ a fully connected network trained on the ADE20K dataset to semantically segment selected factors in each image. By counting the number of pixels in each segmentation mask, the network calculates the area ratio of each factor (Yao et al., 2019). This method has proven effective in identifying factors that affect people's perceptions of insecurity. For the features that show low accuracy or cannot be segmented with ADE20K, such as ‘Security equipment’, ‘Surveillance camera’, and ‘Exposed pipes’. The authors utilize the YOLOv8s-Seg network to perform segmentation of these features in this research.

The YOLO series is a deep-learning model for detecting objects (Yue et al., 2023). YOLOv8-Seg, an extension of YOLOv8 designed for segmentation, is widely used in fields like medical imaging, agriculture, and infrastructure maintenance(Yue et al., 2023; Wu et al., 2023; Chaoju and Jian, 2024). Due to its effectiveness, YOLOv8-Seg has also been applied in architectural and urban studies. For instance, Roudbari et al. (2023) used YOLOv8-seg to segment common household factors, Zhang et al. (2024) employed a YOLOv8-seg based model to detect and semantic several factors in urban street. This research selects YOLOv8-Seg to within metro stations, capitalizing on its demonstrated capabilities in similar contexts.

There are five different scale models of YOLOv8-Seg network: YOLOv8n-Seg, YOLOv8s-Seg, YOLOv8m-Seg, YOLOv8l-Seg, and YOLOv8x-Seg. Previous comparative studies of these models, such as those by Yue et al. (2023) and Paul et al. (2024), evaluated these models and identified YOLOv8s-Seg as the optimal choice for balancing segment $mAP_{@0.5}$ and model size. Therefore, this research employs the YOLOv8s-Seg model to segment visual factors in metro station.

The segmentation program is conducted separately for the Milan and Beijing datasets, encompassing four distinct phases. Initially, 1,000 images are randomly selected and split into training and validation datasets at an 8:2 ratio. Next, semantic segmentation labeling is meticulously performed on the selected features using specialized annotation software Lableme. Thirdly, the authors apply YOLOv8s-Seg, progressively increasing the number of images until the model achieves a Segment $mAP_{@0.5}$ of 70% on the validation set. Upon reaching this benchmark, the model is deployed to segment features in the remaining images.

However, using YOLOv8 requires a sufficiently large number of photos for effective training. For features with limited photos, YOLO training is not feasible, so the authors manually annotate them using Labelme. Table 4 shows the type of segmentation for different factors. Finally, the Openpyxl and Shapely libraries in Python are utilized to compute the area ratio of each feature in the images. This provides the area ratio of each feature across all images from both cities.

Table 4
Segmentation methods for different features in metro stations

ADE20K	YOLOv8s-Seg	Manual mode
Artificial light	Platform doors	Security check

Floor	Surveillance camera	Visual window
Wall	Security booth	Graffiti
Column	Store / Vending machine	Broken pavement
Stairs	Scattered litter	Broken ceilings
Railing		Broken walls
People		Exposed pipes
Billboards / Signage		

Source: The authors, 2024

3.3.4 The volunteers and auditing

Forty volunteers participate in auditing the images. The volunteers are long-term residents of Beijing or Milan who regularly use the metro system and are concerned about its safety risks. The demographics of these respondents are detailed in Table 5.

Table 5
Descriptive statistics for the volunteers

Variables	Proportion/Mean (SD)
Residence (%)	
Milan	50
Beijing	50
Gender (%)	
Male	42.50
Female	57.50
Age	25.31 (4.79)
Education (%)	
High school or below	12.5
College and above	87.5

Source: The authors, 2024

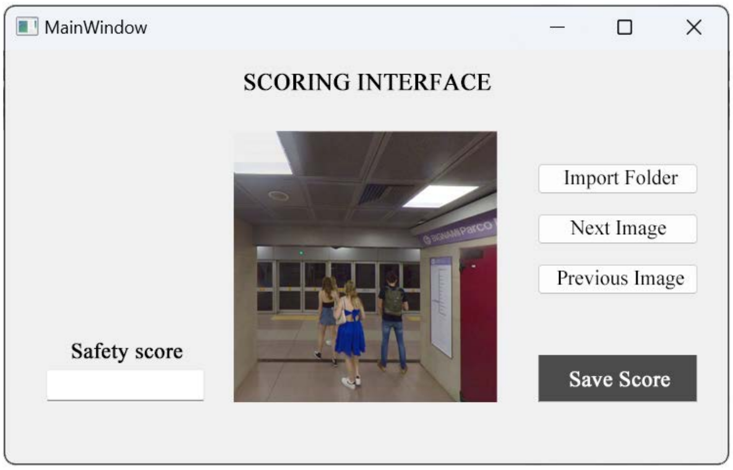


Figure 8. Volunteer auditing interfaces
Source: The authors, 2024

This research develops a Graphical User Interface for volunteer auditing, enabling volunteers to assign scores that assess safety perceptions (Figure 8). To ensure uniformity, all volunteers receive an operational manual outlining the criteria and score intervals. The intervals are defined as follows: below 40 denotes very insecurity, 40-60 signifies insecurity, 60-80 indicates security, and above 80 represents very security.

3.4 Research models

The authors define the area ratio of different factors as the independent variable and the volunteer auditing scores as the dependent variable, using the XGBoost method to build regression models and SHAP for interpretation. This research compares old and new metro stations in Milan and Beijing, developing four distinct models for each criteria: old and new stations in both cities: Milan's old stations, Milan's new stations, Beijing's old stations, Beijing's new stations.

3.4.1 Modelling approach

The XGBoost algorithm is an advanced machine learning library that enhances gradient boosting algorithms (Chen and Guestrin, 2016). This algorithm represents an advanced version of the ensemble learning method gradient boosted decision tree (GBDT). It addresses GBDT's limitations by incorporating a second-order Taylor expansion of the loss function and integrating L2 regularization. These enhancements help to reduce complexity and prevent overfitting (Zhou et al., 2022). The authors briefly presented its mathematics:

i. The authors investigate a dataset containing n sets of data: $D = \{(x_i, y_i): i = 1,2 \dots, n, x_i \in R^p, y_i \in R\}$, where x_i represents the area ratio of each feature, y_i represents the safety perceptions score given by volunteers. Assuming k decision tree functions f_k are trained, the final prediction for the k-th boost result $\widehat{y_i}$ is:

$$\widehat{y_i} = \sum_{k=1}^{K} f_k(x_i)\,, (f_k \in \mathcal{F}, \mathrm{i} \in \mathrm{n}) \tag{1}$$

where $\mathcal{F}$ is the ensemble space of the regression trees. The objective function is defined as follows:

$$obj = \sum_{i=1}^{n} \mathcal{L}(y_i, \widehat{y_i}) + \sum_{k=1}^{K} \Omega(f_k) \tag{2}$$

This objective function combines the loss function and the regularization term. The loss function measures the deviation between predicted values and true values. The regularization term $\Omega(f_k) = \gamma T + \frac{1}{2}\lambda\|\omega\|^2$ is used to control the complexity of the algorithm, where γ and λ are regularization coefficients, T is the number of tree leaf nodes, and ω represents the leaf weight values.

ii. The authors perform a second-order Taylor expansion for the objective function. After removing the constant term, the authors simplify and obtain the objective function:

$$obj = \sum_{i=1}^{n} \left[g_i f_k(x_i) + \frac{1}{2} h_i f_k^2(x_i) \right] + \Omega(f_k) \tag{3}$$

where g_i and h_i are known as the first-order gradient and second-order gradient for sample i.

iii. A leaf node score ω_j is introduced to represent the weight of the leaf node j. And $q(x)$ is the function that maps sample x to its corresponding leaf node index. Therefore, the authors could define $f_k(x_i)$, the sample set I_j, and $\Omega(f_k)$:

$$f_k(x_i) = \omega_{q(x_i)} \tag{4}$$

$$I_j = \{i | q(x_i) = j\} \tag{5}$$

$$\Omega(f_k) = \gamma T + \frac{1}{2} \lambda \sum_{i=1}^{T} w_i^2 \tag{6}$$

By converting the sample set into corresponding leaf nodes, the authors define: $G_j = \sum_{i \in I_j} g_i$, $H_j = \sum_{i \in I_j} h_i$. Then the authors obtain the new objective function:

$$obj = \sum_{j=1}^{T} \left[G_i \omega_j + \frac{1}{2} (H_j + \lambda) \omega_j^2 \right] + \gamma T \tag{7}$$

iv. When $\omega_j = -\frac{G_j}{(H_j + \lambda)}$, the authors obtain the minimum value of the objective function:

$$obj = -\frac{1}{2} \sum_{j=1}^{T} \frac{G_j^2}{(H_j + \lambda)} + \gamma T \tag{8}$$

v. To find the optimal tree structure, enumerating all possible structures is impractical due to the infinite number of possibilities. Instead, the authors use a greedy algorithm to simplify the objective function as follows:

$$obj = \frac{1}{2} \left[\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{G_L^2 + G_R^2}{H_L + H_R + \lambda} \right] - \gamma \tag{9}$$

Where L represents the left nodes after the split, R represents the right nodes after the split.

This research develops regression models using the XGBoost algorithm. The data is randomly split into 70% for the training set and 30% for the test set. To optimize the model's performance, the authors adjust the settings of hyperparameters, an approach that has been proven effective in previous studies (Ryu et al., 2020; Zhou et al., 2022). Table 6 displays the ranges of the hyperparameters.

Table 6
The optimal combination of hyperparameters in different models

Hyperparameters	Value range
n_estimators	[100, 200, 300, 400, 500, 600, 700]
max_depth	[4, 5, 6, 7, 8, 9, 10, 11, 12]
learning_rate	[0.005, 0.01, 0.02, 0.05, 0.1, 0.15]

min_child_weight	[1, 2, 3, 4, 5, 6, 7, 8, 9, 10]
subsample	[0.5, 0.6, 0.7, 0.8, 0.9, 1.0]
colsample_bytree	[0.5, 0.6, 0.7, 0.8, 0.9, 1.0]

Source: The authors, 2024

To systematically explore the optimal combination of these parameters, the authors employ GridSearchCV, utilizing five-fold cross-validation to evaluate model performance across different parameter sets. The iterative process is automatically terminated once the model evaluation score ceases to improve, leading to the identification of the optimal set of hyperparameters that minimizes the error rate. Subsequently, the authors train models using the optimal parameters. The final model performance is evaluated using several widely accepted metrics, including MAE, RMSE, MAPE, and R-squared. Table 7 shows the performance of all regression models.

Table 7
Model performances

Hyper-parameters		Criteria 1				Criteria 2			
		Milan		Beijing		Milan		Beijing	
		old	new	old	new	old	new	old	new
R-squared	Training set	0.95	0.92	0.88	0.93	0.93	0.96	0.88	0.97
	Test set	0.73	0.71	0.69	0.70	0.77	0.77	0.69	0.76
MAE	Training set	2.35	1.69	2.95	1.95	2.82	1.35	2.95	1.16
	Test set	5.24	3.39	4.82	4.10	5.13	4.04	4.82	3.31
RMSE	Training set	3.08	2.14	3.89	2.64	3.59	1.78	3.89	1.57
	Test set	7.00	4.06	6.30	5.46	6.71	5.15	6.30	4.61
MAPE	Training set	5.97	2.81	5.45	3.28	6.64	2.36	5.45	1.83
	Test set	12.26	5.59	9.63	6.83	11.83	7.80	9.63	5.33

Source: The authors, 2024

3.4.2 Model interpretation approach

SHAP is a Python package designed for model interpretation, capable of explaining the output of any machine learning model. Drawing from cooperative game theory, SHAP builds an additive explanatory model where all features are treated as "contributors" (Lundberg and Lee, 2017). This approach is capable of elaborating on the complicated nonlinear effects of variables on safety perceptions. The authors briefly introduce this interpretation approach:

i. For each predicted sample, the model generates a "SHAP value," representing the sum of the contributions from each feature. Shapley value $\phi_k(f, x)$ is defined as:

$$\phi_k(f, x) = \sum_{S \epsilon \mathcal{S}} \frac{1}{K!} \left[f_x \left(P_k^S \cup k - f_x(P_k^S) \right) \right] \quad (10)$$

where k represents a specific feature or variable for which the Shapley value is being calculated, $\mathcal{S}$ is the set of all variables, K is the number of variables. P_k^S

represents the set of variables that come before the variable k in the ordering S, x is the values of the explanatory variables.

ii. These values are then analyzed by determining the contribution of each feature to the overall predictions, so the single prediction $f(x)$ is explained by:

$$f(x) = \phi_0(f,x) + \sum_{k=1}^{K} \phi_k(f,x) \tag{11}$$

where $\phi_0(f,x)$ is the average value of overall predictions.

iii. The relative importance of each variable is quantified by defining the importance measure of variable k as:

$$I_k = \frac{1}{n}\sum_{i=1}^{n}\left|\phi_k^{(i)}\right| \tag{12}$$

where $\phi_k^{(i)}$ is the Shapley value of variable k for the i-th sample. This formula calculates the mean of the absolute Shapley values for feature k across all samples, reflecting the importance of that feature to the model's predictions.

This research employs the SHAP package to interpret the regression model constructed with XGBoost. SHAP values provide a theoretically sound and consistent measure of feature importance, allowing the authors to thoroughly understand how each visual factor in metro stations contributes to people's perceptions of insecurity.

4. Analysis and Results

4.1 Comparative analysis of old and new metro stations in Milan

4.1.1 Relative importance of vary visual factors

Figure 9 illustrates the relative importance of various visual factors in Milan’s old and new metro stations, evaluated under Criteria 1. The figure is divided into two panels: the upper panel represents the old stations, while the lower panel represents the new stations. Each panel features bar charts on the left, displaying the mean SHAP values in descending order, indicating the global importance of each feature. On the right, scatter plots illustrate the local effects of these features on specific samples.

SHAP values above zero indicate positive effects on safety perception, while values below zero reflect negative effects. The color gradient from blue to red represents the range from low to high feature values. The relationship between SHAP values and color allows for inferences about how feature levels influence safety perception. For instance, in new metro stations, red points for platform doors predominantly appear on the positive side, while blue points are mainly negative, indicating that higher feature values are positively associated with safety perception.

Figure 9 indicates that some visual factors are highly correlated with people’s safety perception. In old stations, the area ratio of artificial light is the most dominant explanatory variable, followed closely by people, store/vending machines, and floor, all of which are positively associated with safety perception. Conversely, walls, stairs, and scattered litter are negatively associated with safety perception. In new stations, platform doors emerge as the most significant factor, followed by people, floor, and artificial light, all of which are positively associated with safety.

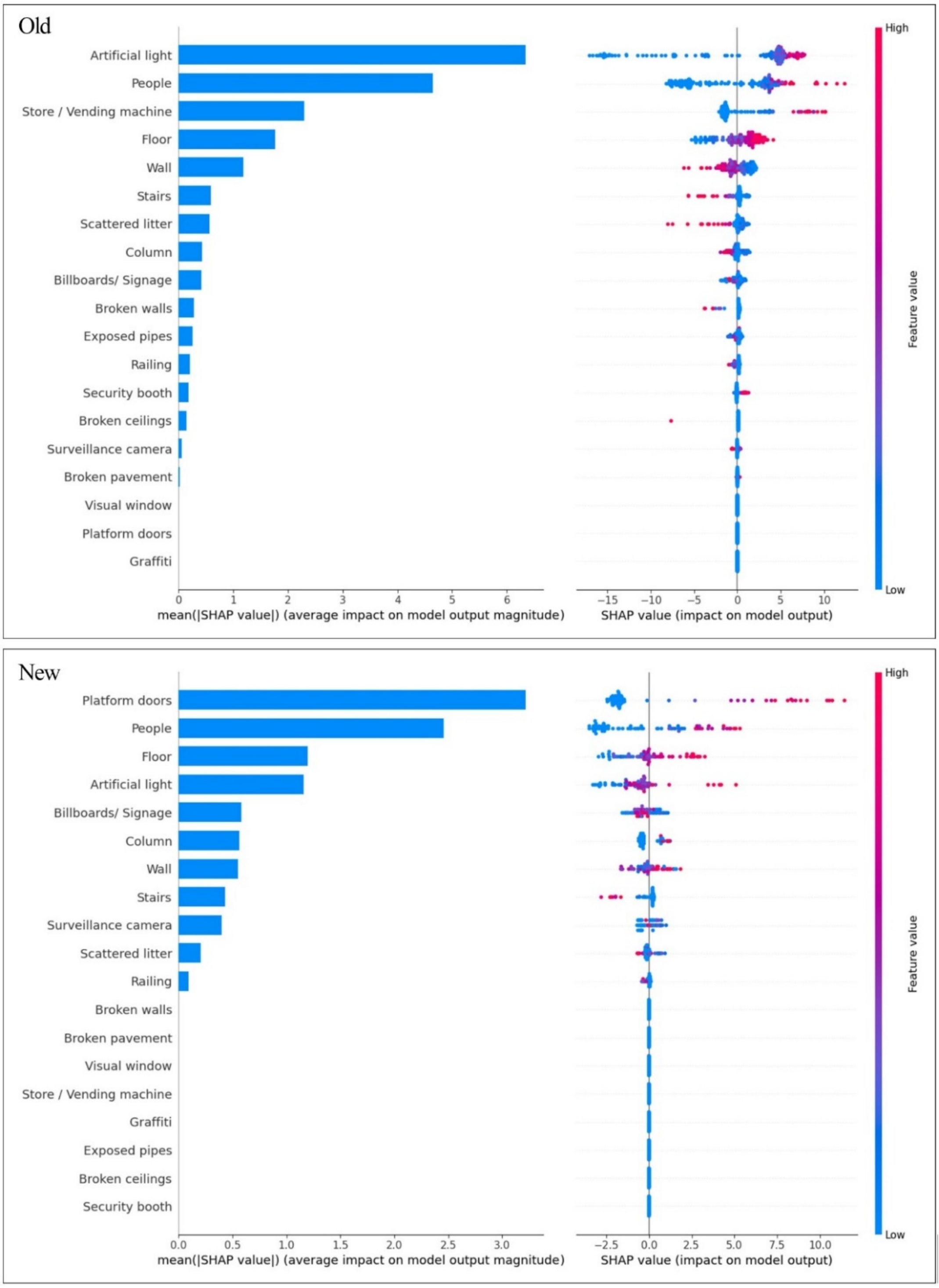


Figure 9. Comparison of relative importance of visual factors in Milan's old and new metro stations (Criteria 1)
Source: The authors, 2024

4.1.2 Nonlinear effects of vary visual factors

To further explore these relationships, the authors employ local dependence plots to analyze the nonlinear effects of various visual factors, highlighting distinct trends between the two types of stations (Figure 10).

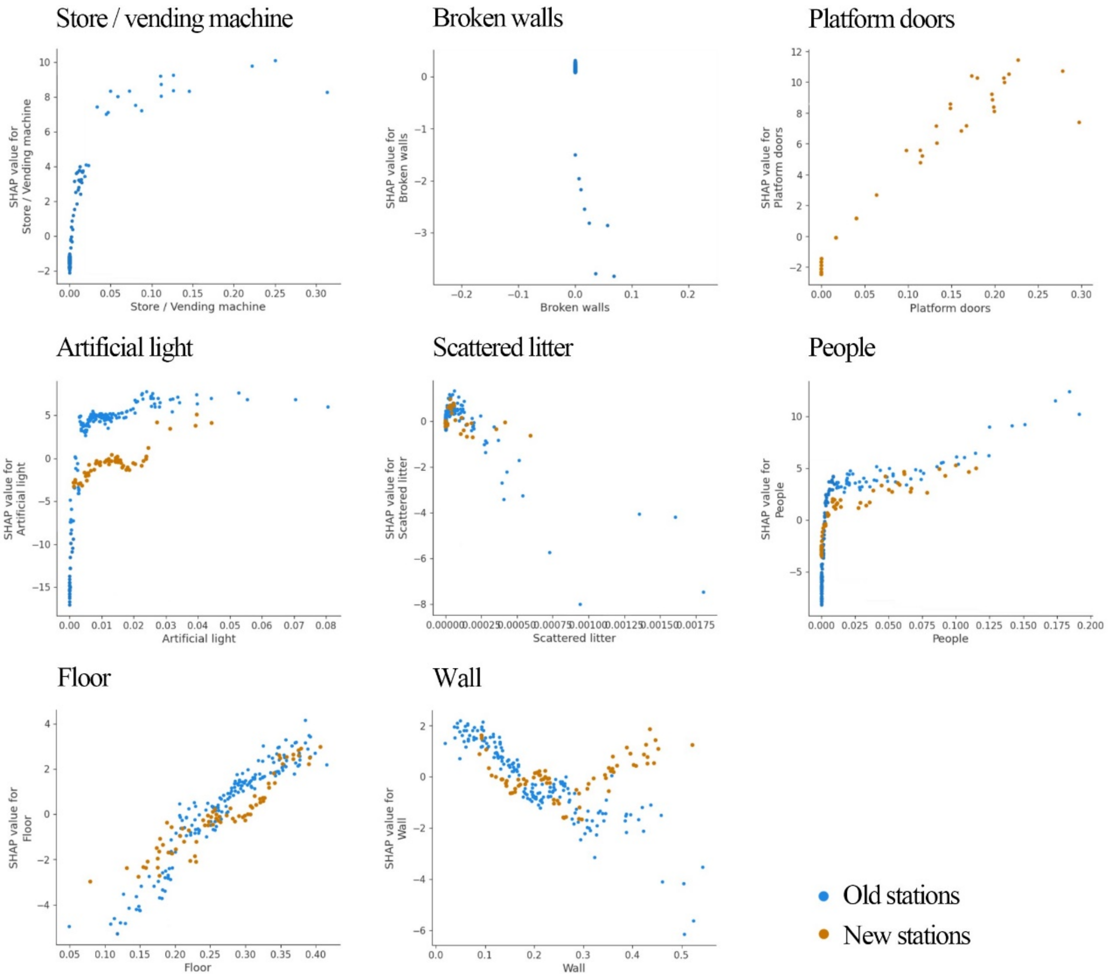


Figure 10. Comparison of local effects of variables on people's safety perception in Milan's old and new metro stations (Criteria 1)

Source: The authors, 2024

For variables predominantly associated with either older or newer stations, store and vending machines mainly exist in old stations, where the area ratio of stores/vending machines shows a positive linear association with safety perceptions when it ranges between 0-0.05. Beyond this range, the effect remains high but stable. Broken walls are also primarily present in old stations, with their local effect showing a sharp decline. In contrast, platform doors exist only in new stations. When the area ratio exceeds 0, their local effects present a positive and approximately linear association with safety.

For variables present in both old and new stations, there are significant differences in local effects. Artificial lighting in old stations shows a sharp rise in local effects as the area ratio increases from 0 to 0.03, jumping from −18 to 5 before leveling off. In new stations, however, the local effect increases more gradually, with no significant changes observed at low area ratios.

Similarly, the effect of scattered litter varies between old and new stations. In old stations, a small amount of litter (area ratio 0-0.00025) has little effect on safety, but beyond this threshold, the local effects sharply decrease. However, there are no significant changes in new stations. For other factors, in contrast, the presence of people shows a consistent rise in local effects in both old and new stations.

The comparison of floors and walls shows different results. In both old and new stations, the floor area is positively associated with safety. However, the wall area exhibits varying effects: in old stations, an increase in the wall area ratio leads to a sharp decline in safety, while in new stations, although there is an initial increase, beyond a certain point (area ratio >0.3), the trend reverses.

4.1.3 Comparative analysis for Criteria 1 and Criteria 2

Under Criteria 2 in the selection process for old and new stations, Figure 11 (upper) illustrates the relative importance of various visual factors in Milan's metro stations. Comparing Figures 9 and 11, the lower panels for new stations show significant variation.

In new stations, under Criteria 1, the presence of platform doors is the most dominant explanatory variable, while artificial lighting plays a less significant role in influencing safety perception. In contrast, under Criteria 2, artificial lighting becomes the most critical variable, particularly evident from the longer blue tail on the left side of the distribution, indicating a stronger negative association with safety perceptions in poorly lighting conditions.

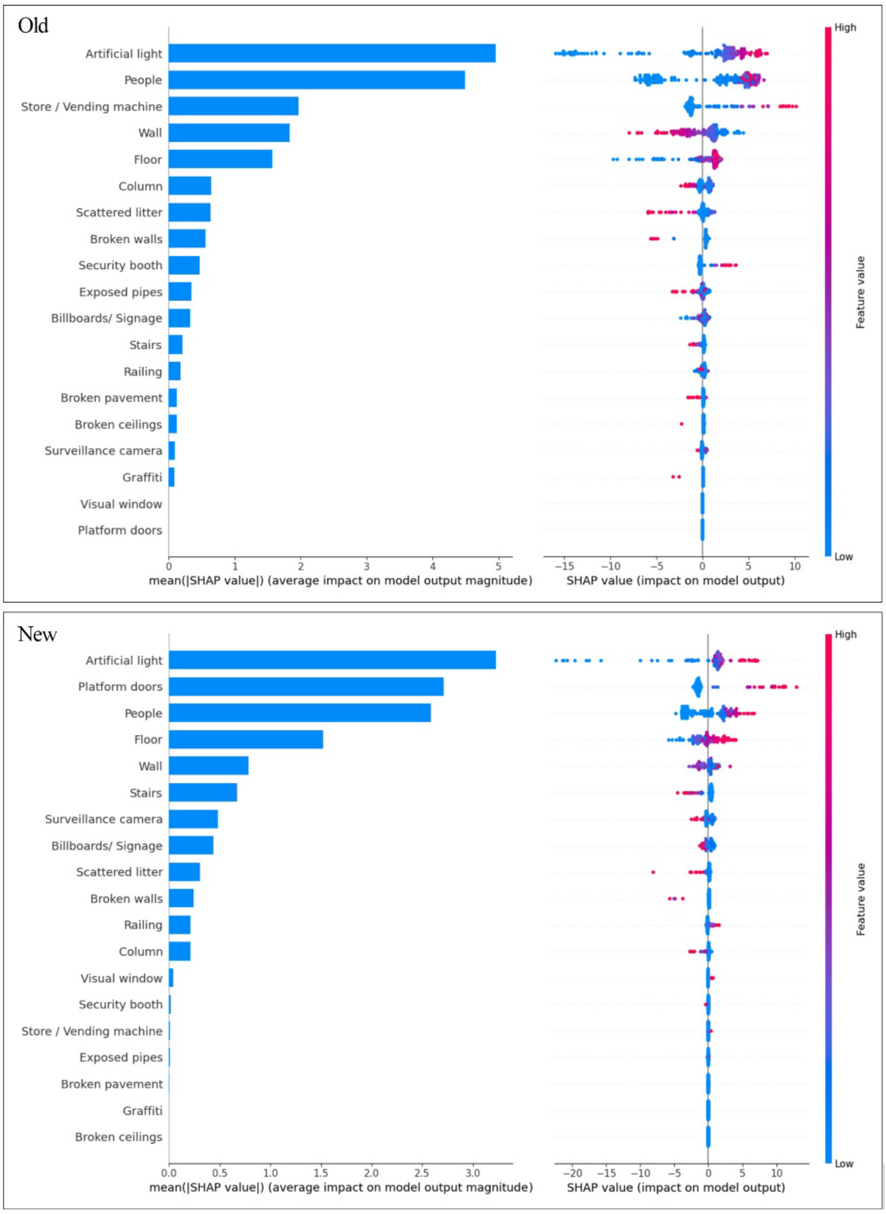
To further explore these findings, the authors use local dependence plots to compare the significance of artificial lighting and platform doors in new stations under Criteria 1 and Criteria 2 (Figure 11 lower). For platform doors, the growth rates of local effects are relatively consistent between different criteria. However, for artificial lighting, under Criteria 2, the local effects rise sharply than Criteria 1.

4.2 Comparative analysis of old and new metro stations in Beijing

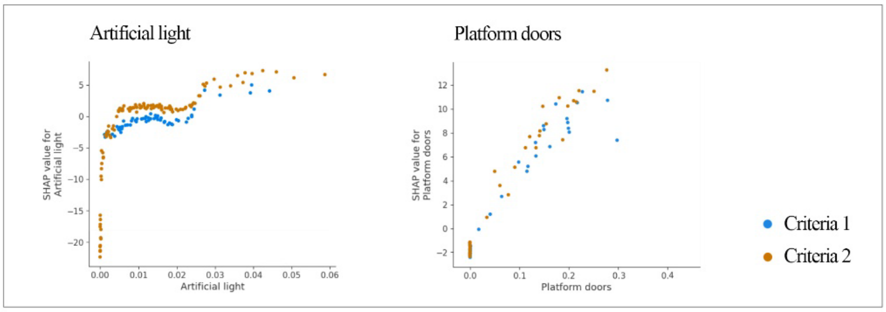
4.2.1 Relative importance of vary visual factors

Figure 12 illustrates the comparison of relative importance of visual factors in Beijing's old and new metro stations. In old stations, the most influential variable is exposed pipes, followed by the floor, people, and wall. Exposed pipes and walls generally exhibit a more significant negative association with safety perception, as indicated by the pronounced red tail on the left. Conversely, the floor and people show relatively positive associations. Additionally, artificial light, surveillance camera, and billboards/signage also positively correlated with safety perception.

In new stations, artificial light emerges as the most dominant factor, followed by the floor, people, and columns, all of which positively influence safety perception. Conversely, walls and stairs exhibit relatively negative associations. Furthermore, the presence of stores and vending machines also contributes positively to the safety perceptions .



Comparison of relative importance of visual factors in Milan's old and new metro stations (Criteria 2)



Comparison of local effects of variables on people's safety perception in Milan's new metro stations (Criteria 1 and Criteria 2)

Figure 11. Comparative analysis for Criteria 1 and Criteria 2 (in relation to Figure 9 & 10)

Source: The authors, 2024

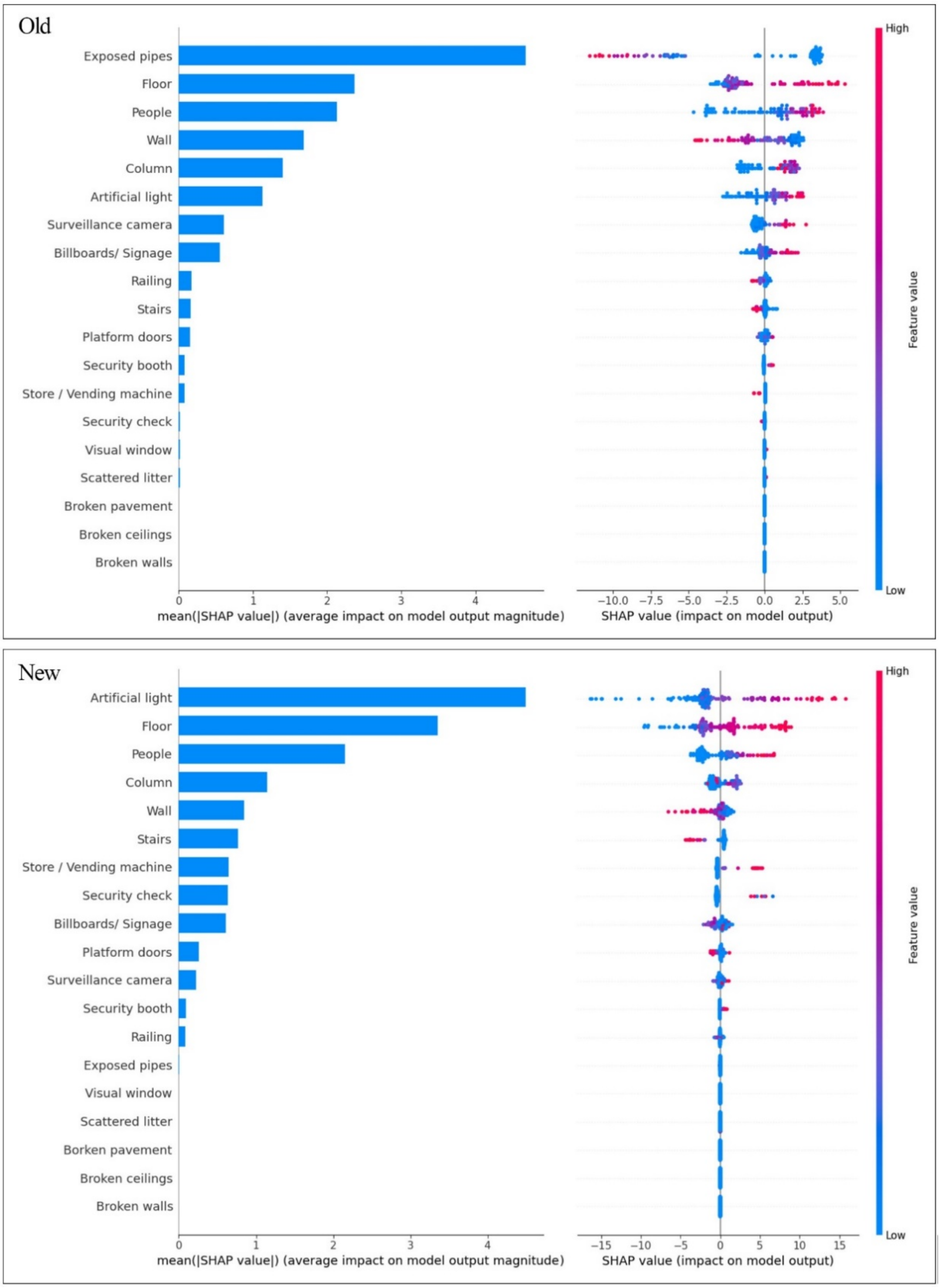


Figure 12. Comparison of relative importance of visual factors in Beijing's old and new metro stations (Criteria 1)

Source: The authors, 2024

4.2.2 Nonlinear effects of vary visual factors

To further explore these relationships, local dependence plots are used to analyse. As shown in the figure 13, all these variables are present in both new and old station.

The area ratio of artificial light and people are more positively associated with safety in new stations than old, with a particularly pronounced difference for artificial light. In new station, when area ratio of artificial light is between 0 and 0.1, the local effects increase sharply from -18 to -2. As the area ratio increases further, the effects stabilize and then continue to rise. However, in old stations, overall change is relatively smooth. For other element, surveillance cameras have a more pronounced positive impact on safety perceptions in old stations, while their effect is minor in new stations.

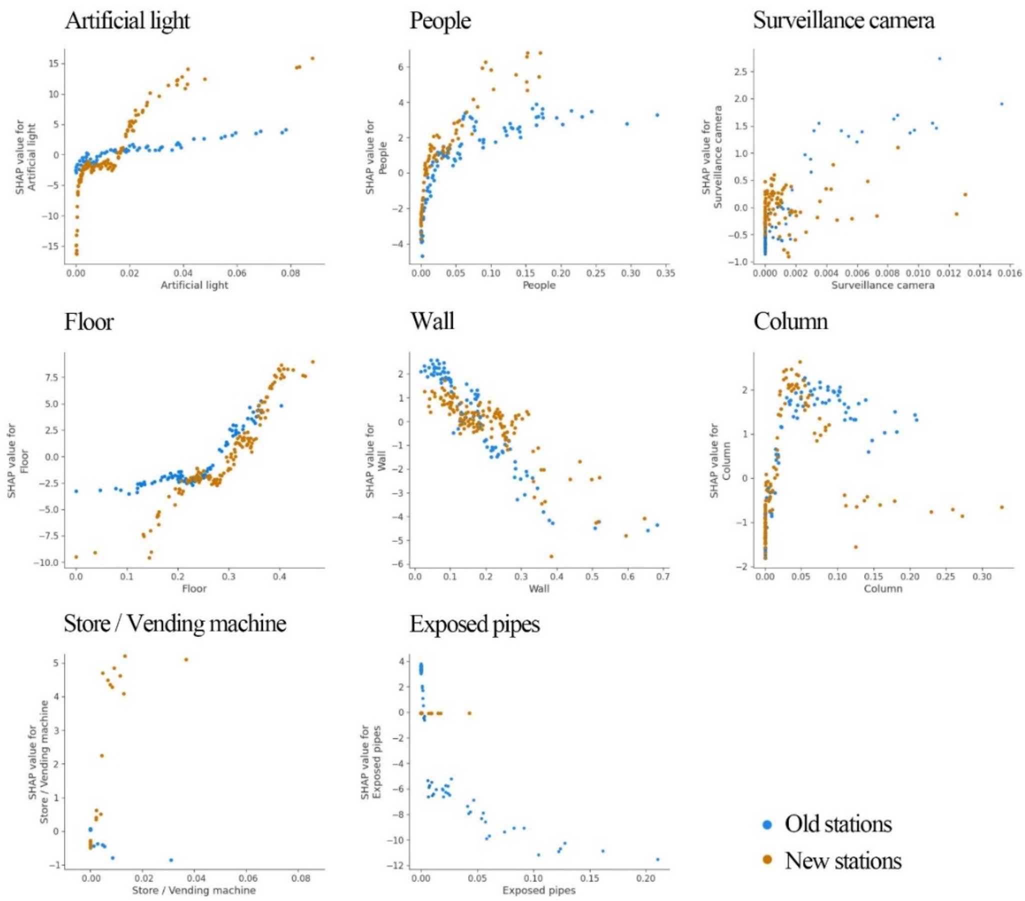
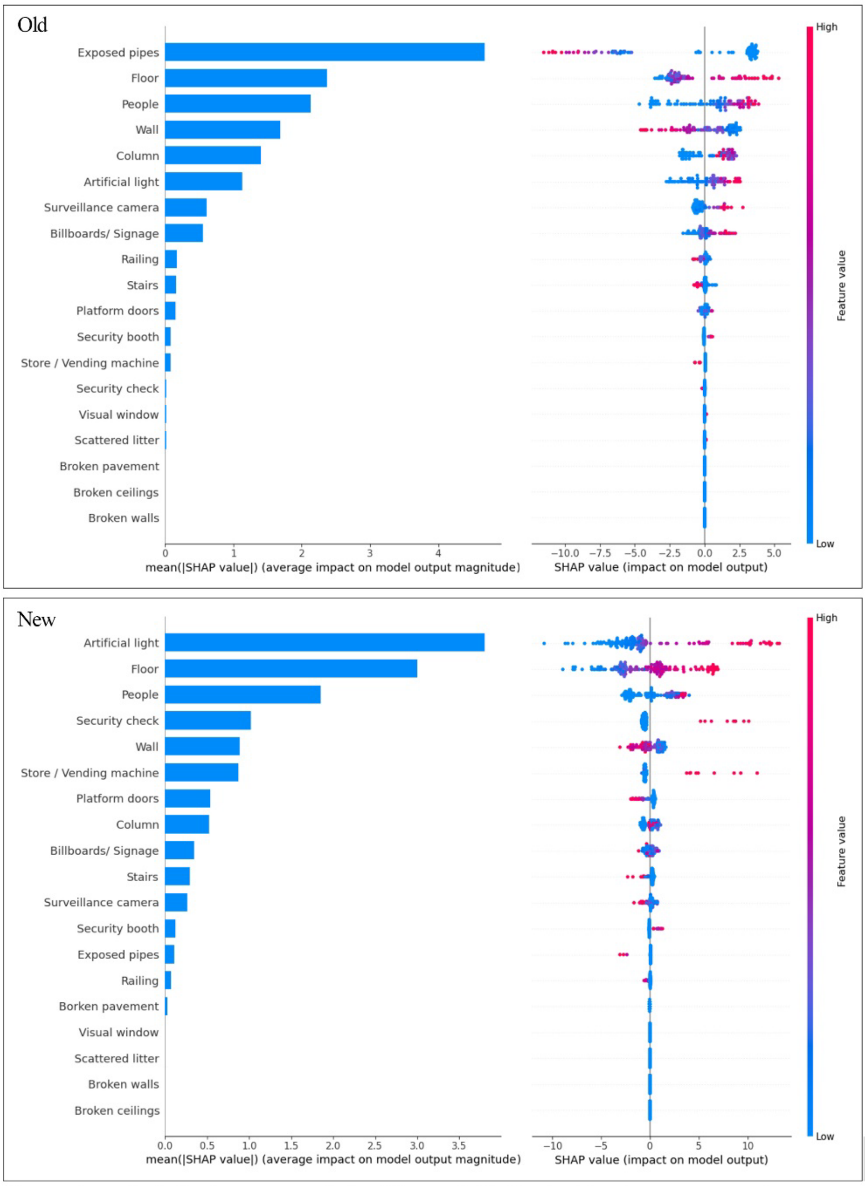


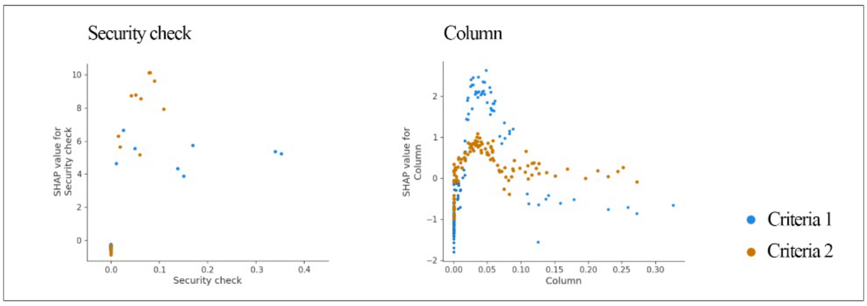
Figure 13. Comparison of local effects of variables on people’s safety perception in Beijing’s old and new metro stations (Criteria 1)
Source: The authors, 2024

Floors, walls, and columns, essential elements in any space, impact safety perceptions in distinct ways. For walls, both old and new stations display similar downward trends. However, floors show different effects: in new stations, they have a sharper positive impact, particularly when the area ratio is between 0.1 and 0.25. For columns they initially exhibit same positive effect in both old and new stations, but in new stations, this effect declines more sharply as the column area ratio increases.

For store/vending machine and exposed pipes, they present totally opposite trend. The local effects of stores/vending machines are generally higher in new stations compared to old ones, where the effects are negative and less varied. In contrast, exposed pipes, a dominant factor in old stations, exhibit more negative local effects that gradually worsen as the area of exposed pipes increases. In new stations, however, local effects remain stable around zero.



Comparison of relative importance of visual factors in Beijing's old and new metro stations (Criteria 2)



Comparison of local effects of variables on people's safety perception in Beijing's new metro stations (Criteria 1 and Criteria 2)

Figure 14. Comparative analysis for Criteria 1 and Criteria 2 (in relation to Figure 12 & 13)

Source: The authors, 2024

4.2.3 Comparative analysis for Criteria 1 and Criteria 2

Under Criteria 2 in the selection process for Beijing's metro stations, the relative importance of various visual factors in old and new stations is illustrated in Figure 14 (upper). Since the selection of old stations remains consistent between Criteria 1 and Criteria 2, the comparative analysis primarily focuses on new stations.

Artificial light, floor, and the presence of people are consistently the most significant factors influencing safety perceptions across both criteria. but security check and column show significant differences. To further examine these differences, local dependence plots are employed to compare the significance of security checks and columns in new stations under Criteria 1 and Criteria 2 (Figure 14 lower). For security checks, the local effects are much higher under Criteria 2, indicating a stronger positive association with safety perception. In contrast, for columns under Criteria 1, the local effects rise sharply to a distinct peak, then decline but mostly stay positive. Under Criteria 2, the local effects are lower and evenly spread.

5. Discussion of the Results

5.1 Milan's old and new metro stations

The comparative analysis of the relative importance and nonlinear effects of various visual factors under Criteria 1 reveals distinct similarities and differences between old and new metro stations. The authors discuss according to the classification of all visual factors:

1. Visual accessibility: Artificial light plays a crucial role in shaping safety perception, particularly in old stations. In old stations, when the area ratio of artificial light is low, the local effect on safety perceptions rapidly shifts from negative to positive. This indicates that inadequate lighting in older stations strongly contributes to feelings of insecurity. In contrast, new stations show a more gradual and steady increase in safety perceptions as lighting improves, reflecting the influence of modern, evenly distributed, and efficient lighting systems. However, the outdated or uneven lighting in older stations makes safety perceptions more sensitive to changes in artificial light coverage.

   For other factors, both old and new stations show a high correlation between floor area ratio and safety, suggesting that more open spaces enhance safety perception, but walls play a more significant role in old stations. The authors suspect that, this difference is due to the lower quality of walls in older stations, which may appear more dilapidated or cluttered compared to the cleaner walls in new stations.

2. Surveillance: People are consistently significant in association with safety perceptions in both old and new stations, regardless of the varying quality of the spaces

3. Vitality: Stores and vending machines, more common in old stations, are positively associated with safety. In contrast, platform doors, found exclusively in new stations, have a substantial impact on safety perception. The recent installation of platform doors on newer metro lines has likely increased public awareness of their safety benefits. Therefore, adding commercial facilities to new stations and installing platform doors in old stations are likely to significantly enhance the sense of safety.

4. Disorder: Scattered litter and broken walls have a more substantial negative impact on safety in old stations. Possibly due to the overall cleanliness and modern infrastructure, which mitigate the negative effects of litter and broken walls.

The comparative analysis of Criteria 1 and Criteria 2 highlights the significant role of artificial lighting in older stations. In new stations under Criteria 1, platform doors are the most influential factor affecting safety perception. However, in Criteria 2, artificial lighting takes precedence as the dominant factor. A closer examination of local effects under Criteria 2 reveals that when the area ratio of artificial lighting is low, safety perceptions drops sharply, indicating the significant negative influence of poor lighting conditions. In contrast, no such drastic changes are observed under Criteria 1, where safety perceptions remain more stable.

The differences can be attributed to the variations in infrastructure and technological advancements between the stations included under Criteria 1 and Criteria 2. The new stations selected under Criteria 1, as part of the latest metro lines (Lines 4 and 5), benefit from modern equipment such as advanced lighting systems and platform doors. In contrast, the stations under Criteria 2 include older lines (Lines 1, 2, and 3) with a broader range of construction dates, where infrastructure may not have been updated to the same standards. Many of these stations have not been upgraded to the same technological standards, the poor lighting conditions in some stations create a stronger negative association with safety perception.

5.2 Beijing's old and new metro stations

The comparative analysis of visual factors under Criteria 1 shows key similarities and differences in how they influence safety perceptions between old and new stations:

1. Visual accessibility: Artificial light plays a more significant role in enhancing the perceptions of safety in new stations. The comparison of local effects shows that in new stations, the impact on safety perceptions increases rapidly from negative to neutral under low lighting conditions, whereas old stations do not exhibit such significant changes. This finding is not surprising, as during the manual photography process, the authors observe that certain areas in new stations, especially interchange passages, have very poor lighting conditions, which inhibits safety perception.

Factors such as floors, walls, and columns demonstrate various association with safety in both old and new stations. A higher wall area ratio reduces perceived openness and visibility, contributing to a sense of insecurity. Floors, however, show a sharper positive impact on safety perceptions in new stations, probably because new stations tend to have higher floor heights, amplifying the effect of floor area on spatial perception. In contrast, columns initially contribute positively to safety in both old and new stations. However, in new stations, the local effect diminishes more sharply as the area ratio increases. This decline is likely due to modern station designs that prioritize open and spacious environments, and columns may be perceived as obstructions to movement or as visual barriers.

2. Surveillance: The presence of people remains a vital factor in both old and new stations. However, surveillance cameras are more positively associated with safety in old stations, likely because the built environment in old stations is generally worse than in new stations, making the presence of monitoring equipment more significant.
3. Vitality: The presence of stores and vending machines is positively associated with safety in new stations but has a slightly negative impact in old stations. This suggests that the facilities in old stations may contribute to a sense of insecurity, possibly due to inadequate maintenance or outdated design.
4. Disorder: In old stations, exposed pipes are the most dominant explanatory variable, which is negatively associated with people's safety perception. However, this association is minimal in new stations. It makes sense that the exposed pipes mainly exist in old stations, and some of them are highly conspicuous, contributing to the disorder of the space.

The comparative analysis of Criteria 1 and Criteria 2 highlight the significant role of security check and column. The nonlinear analysis indicates that security checks have a stronger positive impact on safety perceptions under Criteria 2, while columns show more positive effect under Criteria 1.

These differences can be attributed to the varying construction periods encompassed by Criteria 1 and Criteria 2. The new stations selected under Criteria 2 represent the latest metro lines in Beijing, constructed between 2021 and 2023, while the stations under Criteria 1 span a broader range of construction dates, from 2014 to 2023. Therefore, for security checks, the local effects are much higher under Criteria 2, probably due to the more advanced security equipment in the latest metro stations significantly enhances safety perception. For columns, consistent with previous assumptions, they are likely perceived less favorably in latest stations due to modern preferences for open spaces and clear visibility.

5.3 The comparison of the old and new metro stations in Milan and Beijing

The comparative analysis of the old and new metro stations in Milan and Beijing reveals both similarities and distinct differences in these two cities' metro systems:

1. Visual accessibility: Given that Milan's oldest metro stations are constructed earlier than Beijing's, it makes sense that they have poorer lighting due to aging infrastructure, which has been significantly associated with perceptions of insecurity. But newer stations have limited associated with safety perception. However, despite being more modern, some areas in Beijing's new stations, such as interchange passages, still suffer from poor lighting, which is significant associated with perceptions of insecurity. This indicates that Beijing's newer stations are still face challenges in providing consistent lighting across all spaces.

   Additionally, in both cities, the area ratio of floors is positively associated with safety perception, while walls show the opposite association. The negative association of walls is more pronounced in older stations, likely due to the deterioration of aged walls, which is strongly associated with perceptions of insecurity in both cities.

2. Surveillance: The presence of people consistently correlates with safety perceptions in both Milan and Beijing, regardless of the station's age. However, surveillance cameras play a more prominent role in Beijing.

3. Vitality: Platform doors show significant difference between two cities. In Beijing, platform doors are present in both old and new metro stations, and they have little influence on safety perception, probably because residents are already accustomed to their presence. In contrast, in Milan, platform doors only exist in new stations, where they have a strongly positive effect on safety. This comparison suggests that, since platform doors are a newer feature in Milan, their presence is significantly associated with enhanced safety.

   Additionally, stores and vending machines have varying effects on safety perceptions in the two cities. In Milan, these amenities are common in old stations and are positively associated with safety, while in Beijing, they are limited in amount and negatively impact safety in old stations. This contrast may be because, despite the less favorable environments in old stations in both cities, the numerous storefronts in Milan improve safety by enhancing visibility and oversight.

4. Disorder: factors are all negatively associated with safety perceptions in both cities, particularly in older stations. In Milan, a combination of broken walls, ceilings, pavement, and litter plays a role, whereas in Beijing, exposed pipes are the dominant factor. Therefore, addressing these specific factors of disorder—repairing broken structures in Milan and concealing exposed pipes in Beijing—are highly like to improve the safety perceptions in older metro stations in these cities.

The authors randomly selected some typical space from old and new stations in both cities to better understand the difference. Figure 15 presents localized explanations for the selected spaces. In Milan, the presence of platform doors, effective lighting, and clean floors contributes to a heightened sense of safety in the newest stations. In Beijing, exposed pipes are strongly associated with perceptions of insecurity in older stations, while certain areas in new stations receive lower safety prediction scores due to inadequate lighting.

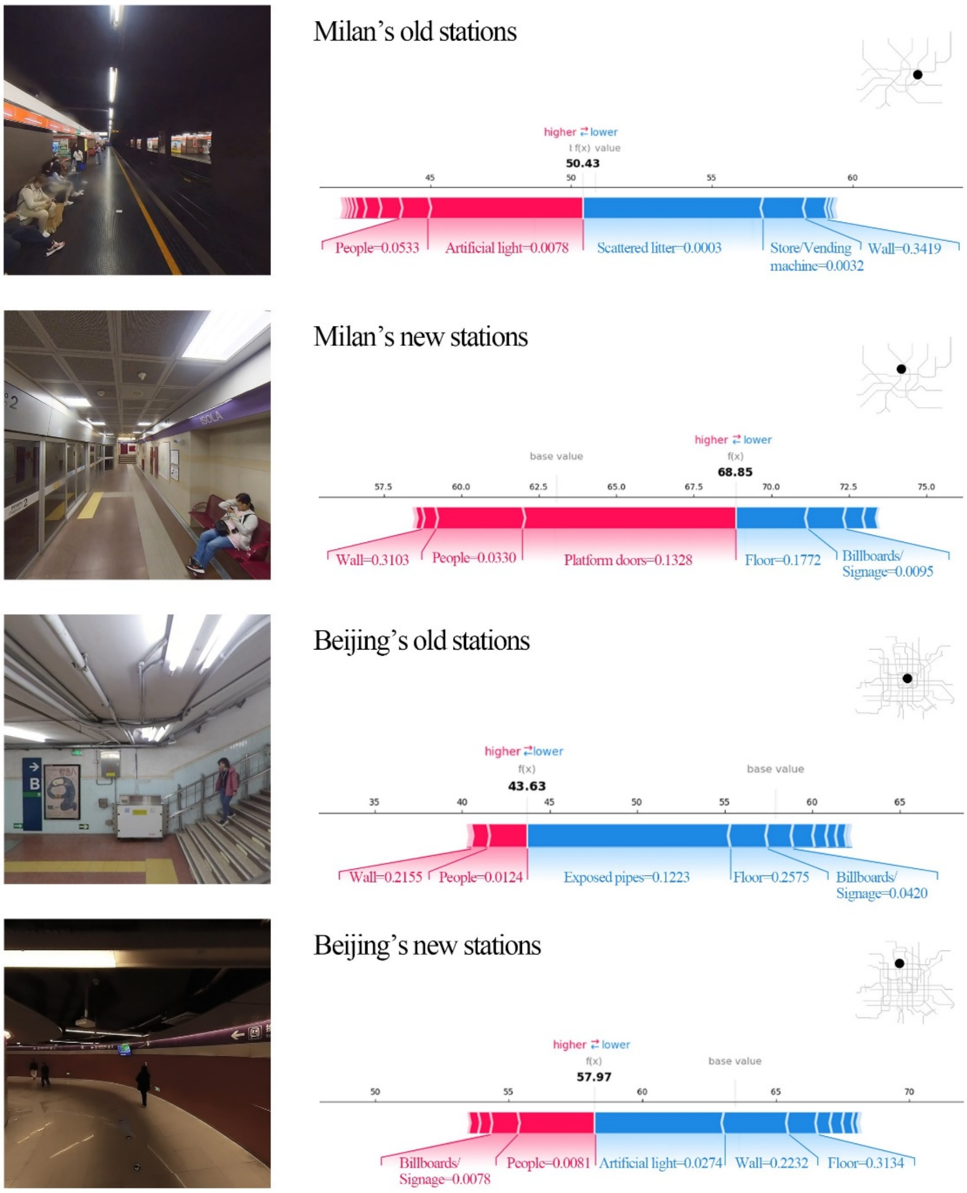


Figure 15. Localized safety perception analysis of typical spaces from old and new stations in both cities

Source: The authors, 2024

6. Conclusions and Recommendations

This research presents a comparative analysis of the built environments in old and new metro stations, focusing on visual factors that shape safety perceptions. Additionally, the authors extend the analysis by comparing the metro systems of Milan and Beijing, representing distinct urbanization processes. By examining how different urbanization models influence the evolution of metro systems, the authors provide a global perspective on the disparities between old and new stations. Using 360-degree mobile image capture, integrating the ADE20K dataset with YOLOv8 models for semantic segmentation, and developing GBDT models and interpreting them locally with SHAP, this research introduces a novel methodology for analyzing underground public spaces.

The findings indicate that in Milan, platform doors, which are only present in new stations, have a strong positive effect on people's safety perception. Similarly, stores and vending machines, which are more common in older stations, also contribute positively to safety. Therefore, in prospective renovations, setting platform doors in old stations and providing more commercial areas in new stations are highly likely to amplify people's sense of safety in stations. Additionally, in both old and new stations, artificial lighting, the presence of people, and open floor spaces positively influence safety, while walls, scattered litter, and broken structures have a negative impact. These negative effects are more pronounced in older stations, where local effects show a steeper decline. As a result, increasing lighting, improving aged floors, walls, and ceilings, and keeping stations clean will be more effective in older stations.

In Beijing, exposed pipes are strongly correlated with negative perceptions in older stations, while inadequate lighting negatively affects safety in some areas of newer stations. As such, effective strategies encompass concealing exposed pipes in the ceilings of older stations and ensuring consistent lighting throughout all new stations. Furthermore, while stores and vending machines are highly likely to increase safety in new stations, they have the opposite effect in older ones. Therefore, in future renovations of older stations, maintaining commercial facilities and harmonising them with their surroundings are highly likely to amplify people's sense of safety.

This research primarily focuses on identifying the differences between old and new metro stations. By integrating the findings, urban planners and designers can develop more effective strategies for modernizing metro stations and improving the safety of public transport systems. However, the research faces limitations. Firstly, it depends on a visual-based measurement method, which lacks the ability to capture more subjective experiences that contribute to people's perceptions of insecurity. Factors such as cultural context, sound conditions, and decoration styles, which play a crucial role in shaping individual safety perceptions in public space, are not fully considered within the current framework. Secondly, the research is limited to Milan and Beijing. Although these cities represent distinct urbanization processes and provide a useful basis for comparison, the findings may not be directly applicable to other cities with different metro system designs or urban development histories. Nevertheless, the results offer valuable insights that could inform similar studies in other regions, and the framework could be used to different metro systems and urban contexts in the future. Furthermore, the novel research framework developed in this research not only advances the analysis of underground public spaces but also provides a flexible tool for future research on various types of underground environments, such as pedestrian tunnels, parking garages, and large-scale transit hubs.

Data availability

Data for this research were obtained legally and can be requested from the corresponding author on reasonable request. However, the research team can only provide segmented data, as the original images contain identifiable facial information of passengers, which cannot be shared to protect privacy and portrait rights. Furthermore, due to the Beijing Metro's function as a civil defense infrastructure, the systematic public release of these images is restricted.

References

Antrop, M., 2004. Landscape change and the urbanization process in Europe. Landscape and Urban Planning, Development of European Landscapes 67, 9–26. https://doi.org/10.1016/S0169-2046(03)00026-4

ATM, 2019. «Atm, record di passeggeri e centomila multe in più». Lineadiretta ATM. URL https://medium.com/lineadiretta/atm-record-di-passeggeri-e-centomila-multe-in-pi%C3%B9-b3e218743f3d (accessed 6.20.24).

ATM Group, 2023. Integrated annual report.

Bădău, F., 2022. Railway interlockings – a review of the current state of railway safety technology in Europe. Promet - Traffic&Transportation 34, 443–454. https://doi.org/10.7307/ptt.v34i3.3992

Birch, E.L., 2016. The new urban agenda's road map for planning urban spatial development: Tangible, manageable and measurable. UN Chronicle 53, 8–11. https://doi.org/10.18356/48644704-en

Ceccato, V., 2013. Moving safely: crime and perceived safety in Stockholm's subway stations. Lexington Books.

Ceccato, V., Loukaitou-Sideris, A., 2021. Fear of sexual harassment and its impact on safety perceptions in transit environments: A global perspective. Violence Against Women. https://doi.org/10.1177/1077801221992874

Ceccato, V., Paz, Y., 2017. Crime in São Paulo's metro system: sexual crimes against women. Crime Prev Community Saf 19, 211–226. https://doi.org/10.1057/s41300-017-0027-2

Ceccato, V., Uittenbogaard, A., Bamzar, R., 2013. Security in Stockholm's underground stations: The importance of environmental attributes and context. Secur J 26, 33–59. https://doi.org/10.1057/sj.2011.32

Chaoju, Y., Jian, R., 2024. Road segmentation algorithm based on improved YOLOv8-seg, in: 2024 IEEE 4th International Conference on Electronic Technology, Communication and Information (ICETCI). Presented at the 2024 IEEE 4th International Conference on Electronic Technology, Communication and Information (ICETCI), pp. 1512–1516. https://doi.org/10.1109/ICETCI61221.2024.10594282

Chen, T., Guestrin, C., 2016. XGBoost: A scalable tree boosting system, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery

and Data Mining, KDD ’16. Association for Computing Machinery, New York, NY, USA, pp. 785–794. https://doi.org/10.1145/2939672.2939785
Coppola, P., Silvestri, F., 2021. Gender inequality in safety and security perceptions in railway stations. Sustainability 13, 4007. https://doi.org/10.3390/su13074007
Cozens, P., Neale, R., Whitaker, J., Hillier, D., 2003. Managing crime and the fear of crime at railway stations—a case study in South Wales (UK). International Journal of Transport Management 1, 121–132. https://doi.org/10.1016/j.ijtm.2003.10.001
Cui, Q., Gong, P., Yang, G., Zhang, S., Huang, Y., Shen, S., Wei, B., Chen, Y., 2023. Women-oriented evaluation of perceived safety of walking routes between home and mass transit: A case study and methodology test in Guangzhou. Buildings 13, 715. https://doi.org/10.3390/buildings13030715
Dadvand, P., Bartoll, X., Basagaña, X., Dalmau-Bueno, A., Martinez, D., Ambros, A., Cirach, M., Triguero-Mas, M., Gascon, M., Borrell, C., Nieuwenhuijsen, M.J., 2016. Green spaces and general health: roles of mental health status, social support, and physical activity. Environment International 91, 161–167. https://doi.org/10.1016/j.envint.2016.02.029
Das, A., Rad, P., 2020. Opportunities and challenges in Explainable Artificial Intelligence (XAI): A survey. https://doi.org/10.48550/arXiv.2006.11371
Ding, Y., Hou, Y., 2022. Analysis of indoor environment state characteristics of urban rail transit stations based on actual measurements in Chongqing, China. Energy and Buildings 277, 112544. https://doi.org/10.1016/j.enbuild.2022.112544
Dubey, A., Naik, N., Parikh, D., Raskar, R., Hidalgo, C.A., 2016. Deep learning the city: quantifying urban perception at a global scale, in: Leibe, B., Matas, J., Sebe, N., Welling, M. (Eds.), Computer Vision – ECCV 2016, Lecture Notes in Computer Science. Springer International Publishing, Cham, pp. 196–212. https://doi.org/10.1007/978-3-319-46448-0_12
Guan, F., Fang, Z., Wang, L., Zhang, X., Zhong, H., Huang, H., 2022. Modelling people’s perceived scene complexity of real-world environments using street-view panoramas and open geodata. ISPRS Journal of Photogrammetry and Remote Sensing 186, 315–331. https://doi.org/10.1016/j.isprsjprs.2022.02.012
Guerra, E., 2015. Suburban transit in Mexico city. ACCESS Magazine 1(46), 20–25.
He, L., Páez, A., Liu, D., 2017. Built environment and violent crime: An environmental audit approach using Google Street View. Computers, Environment and Urban Systems 66, 83–95. https://doi.org/10.1016/j.compenvurbsys.2017.08.001
Hong, J., Chen, C., 2014. The role of the built environment on perceived safety from crime and walking: examining direct and indirect impacts. Transportation 41, 1171–1185. https://doi.org/10.1007/s11116-014-9535-4
Ingvardson, J.B., Nielsen, O.A., 2022. The influence of vicinity to stations, station characteristics and perceived safety on public transport mode choice: a case study from Copenhagen. Public Transp 14, 459–480. https://doi.org/10.1007/s12469-021-00285-x
Joshi, S., Bailey, A., 2023. What happens next? Exploring women’s transport motility through the story completion method. Journal of Transport Geography 107, 103547. https://doi.org/10.1016/j.jtrangeo.2023.103547
Kokkola, M., Nikolaeva, A., Brömmelstroet, M. te, 2023. Missed connections? Everyday mobility experiences and the sociability of public transport in Amsterdam during COVID-19. Social & Cultural Geography.

Koskela, H., Pain, R., 2000. Revisiting fear and place: women's fear of attack and the built environment. Geoforum 31, 269–280. https://doi.org/10.1016/S0016-7185(99)00033-0
Li, J., Song, Z., Wen, J., Cai, C., Tang, P., 2024. Exploring nonlinear relationship between built environment and street vitality using machine learning. Presented at the CAADRIA.
Lundberg, S.M., Lee, S.-I., 2017. A Unified approach to interpreting model predictions, in: advances in neural information processing systems. Curran Associates, Inc.
Navarrete-Hernandez, P., Vetro, A., Concha, P., 2021. Building safer public spaces: Exploring gender difference in the perception of safety in public space through urban design interventions. Landscape and Urban Planning 214, 104180. https://doi.org/10.1016/j.landurbplan.2021.104180
Nesoff, E.D., Milam, A.J., Pollack, K.M., Curriero, F.C., Bowie, J.V., Gielen, A.C., Furr-Holden, D.M., 2018. Novel methods for environmental assessment of pedestrian injury: Creation and validation of the inventory for pedestrian safety infrastructure. J Urban Health 95, 208–221. https://doi.org/10.1007/s11524-017-0226-2
Obiadi, B.N., Nzewi, N.U., 2018. The architecture of the urban fronts, the case of urban experience and pressure on the infrastructure. Mgbakoigba: Journal of African Studies 7, 114–130.
Park, Y., Garcia, M., 2020. Pedestrian safety perception and urban street settings. International Journal of Sustainable Transportation 14, 860–871. https://doi.org/10.1080/15568318.2019.1641577
Paul, A., Machavaram, R., Ambuj, Kumar, D., Nagar, H., 2024. Smart solutions for capsicum Harvesting: Unleashing the power of YOLO for detection, segmentation, growth stage classification, Counting, and real-time mobile identification. Computers and Electronics in Agriculture 219, 108832. https://doi.org/10.1016/j.compag.2024.108832
Paydar, M., Kamani-Fard, A., Etminani-Ghasrodashti, R., 2017. Perceived security of women in relation to their path choice toward sustainable neighborhood in Santiago, Chile. Cities 60, 289–300. https://doi.org/10.1016/j.cities.2016.10.002
Rossetti, T., Lobel, H., Rocco, V., Hurtubia, R., 2019. Explaining subjective perceptions of public spaces as a function of the built environment: A massive data approach. Landscape and Urban Planning 181, 169–178. https://doi.org/10.1016/j.landurbplan.2018.09.020
Roudbari, A., Dalvand, M., Pasiar, N., Asadi, N., Yousefi, M., Masouleh, M.T., 2023. Autonomous vision-based robotic grasping of household objects: A practical case study, in: 2023 11th RSI International Conference on Robotics and Mechatronics (ICRoM). Presented at the 2023 11th RSI International Conference on Robotics and Mechatronics (ICRoM), pp. 918–924. https://doi.org/10.1109/ICRoM60803.2023.10412434
Ryu, S.-E., Shin, D.-H., Chung, K., 2020. Prediction model of dementia risk based on XGBoost using derived variable extraction and hyper parameter optimization. IEEE Access 8, 177708–177720. https://doi.org/10.1109/ACCESS.2020.3025553
Sadeghi, A.R., Baghi, E.S.M.S., Shams, F., Jangjoo, S., 2023. Women in a safe and healthy urban environment: environmental top priorities for the women's presence in urban public spaces. BMC Women's Health 23, 163. https://doi.org/10.1186/s12905-023-02281-8

Sadeghi, A.R., Jangjoo, S., 2022. Women's preferences and urban space: Relationship between built environment and women's presence in urban public spaces in Iran. Cities 126, 103694. https://doi.org/10.1016/j.cities.2022.103694

Sakip, S.R.M., Abdullah, A., 2012. Measuring crime prevention through environmental design in a gated residential area: A pilot survey. Procedia - Social and Behavioral Sciences, ASEAN Conference on Environment-Behaviour Studies (AcE-Bs), Riverside Majestic Hotel, 7- 8 July 2010, Kuching, Sarawak, Malaysia 42, 340–349. https://doi.org/10.1016/j.sbspro.2012.04.198

Shao, Q., Tao, R., Luca, M.M., 2022. The effect of urbanization on health care expenditure: Evidence from China. Front Public Health 10, 850872. https://doi.org/10.3389/fpubh.2022.850872

Shehadeh, A., Alshboul, O., Al Mamlook, R.E., Hamedat, O., 2021. Machine learning models for predicting the residual value of heavy construction equipment: An evaluation of modified decision tree, LightGBM, and XGBoost regression. Automation in Construction 129, 103827. https://doi.org/10.1016/j.autcon.2021.103827

Stjernborg, V., 2024. Triggers for feelings of insecurity and perceptions of safety in relation to public transport; the experiences of young and active travellers. Applied Mobilities 0, 1–21. https://doi.org/10.1080/23800127.2024.2318095

The People's Goveronment of Beijing Municipality, 2023. Three metro lines to open for trial operation on 30 December, Beijing's urban rail transit mileage jumps to first in China. (In Chinese) [WWW Document]. URL https://www.beijing.gov.cn/fuwu/bmfw/sy/jrts/202312/t20231230_3521334.html

Tiwari, G., Phillip, C., 2021. Development of public transport systems in small cities: A roadmap for achieving sustainable development goal indicator 11.2. IATSS Research 45, 31–38. https://doi.org/10.1016/j.iatssr.2021.02.002

Valentine, G., 1990. Women's fear and the design of public space. Built Environment (1978-) 16, 288–303.

Wang, Y., Hayashi, Y., Kato, H., 2009. Performance evaluation of subway lines with data envelopment analysis and geographic information systems, a case study of Beijing.

Wu, Y., Han, Q., Jin, Q., Li, J., Zhang, Y., 2023. LCA-YOLOv8-Seg: An improved lightweight YOLOv8-Seg for real-time pixel-level crack detection of dams and bridges. Applied Sciences 13, 10583. https://doi.org/10.3390/app131910583

Xiao, L., Lo, S., Liu, J., Zhou, J., Li, Q., 2021. Nonlinear and synergistic effects of TOD on urban vibrancy: Applying local explanations for gradient boosting decision tree. Sustainable Cities and Society 72, 103063. https://doi.org/10.1016/j.scs.2021.103063

Xu, S., Li, S., Wen, R., 2018. Sensing and detecting traffic events using geosocial media data: A review. Computers, Environment and Urban Systems 72, 146–160. https://doi.org/10.1016/j.compenvurbsys.2018.06.006

Yao, Y., Liang, Z., Yuan, Z., Liu, P., Bie, Y., Zhang, J., Wang, R., Wang, J., Guan, Q., 2019. A human-machine adversarial scoring framework for urban perception assessment using street-view images. International Journal of Geographical Information Science 33, 2363–2384. https://doi.org/10.1080/13658816.2019.1643024

Yue, H., Xie, H., Liu, L., Chen, J., 2022. Detecting people on the street and the streetscape physical environment from Baidu street view images and their

effects on community-level street crime in a Chinese city. IJGI 11, 151. https://doi.org/10.3390/ijgi11030151
Yue, X., Qi, K., Na, X., Zhang, Y., Liu, Y., Liu, C., 2023. Improved YOLOv8-Seg network for instance segmentation of healthy and diseased tomato plants in the growth stage. Agriculture 13, 1643. https://doi.org/10.3390/agriculture13081643
Zhang, F., Fan, Z., Kang, Y., Hu, Y., Ratti, C., 2021. "Perception bias": Deciphering a mismatch between urban crime and perception of safety. Landscape and Urban Planning 207, 104003. https://doi.org/10.1016/j.landurbplan.2020.104003
Zhang, F., Zhou, B., Liu, L., Liu, Y., Fung, H.H., Lin, H., Ratti, C., 2018. Measuring human perceptions of a large-scale urban region using machine learning. Landscape and Urban Planning 180, 148–160. https://doi.org/10.1016/j.landurbplan.2018.08.020
Zhang, X., Dong, T., Yan, L., Yang, Z., Zhang, J., 2024. LAtt-Yolov8-seg: Video real-time instance segmentation for urban street scenes based on focused linear attention mechanism, in: Proceedings of the International Conference on Computer Vision and Deep Learning, CVDL '24. Association for Computing Machinery, New York, NY, USA, pp. 1–5. https://doi.org/10.1145/3653804.3656278
Zhou, B., Liu, L., Oliva, A., Torralba, A., 2014. Recognizing city identity via attribute analysis of Geo-tagged images, in: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (Eds.), Computer Vision – ECCV 2014, Lecture Notes in Computer Science. Springer International Publishing, Cham, pp. 519–534. https://doi.org/10.1007/978-3-319-10578-9_34
Zhou, S., Liu, Z., Wang, M., Gan, W., Zhao, Z., Wu, Z., 2022. Impacts of building configurations on urban stormwater management at a block scale using XGBoost. Sustainable Cities and Society 87, 104235. https://doi.org/10.1016/j.scs.2022.104235