urban science

Article

Analyzing Urban Crime Through Street View Imagery: Insights from Urban Micro Built Environment and Perceptions

Devin Yongzhao Wu [1,2] **and Jue Wang** [1,2,*]

[1] Department of Geography, Geomatics and Environment, University of Toronto Mississauga, Mississauga, ON L5L 1C6, Canada; yongzhao.wu@mail.utoronto.ca
[2] Department of Geography and Planning, University of Toronto, Toronto, ON M5S 1A1, Canada
* Correspondence: gis.wang@utoronto.ca

Abstract: Understanding the relationship between urban crime and the built environment is crucial for developing effective crime prevention strategies, particularly in the context of rapid urban development and city planning. As cities grow, urbanization leads to environments that either promote or inhibit criminal activity, making it essential to explore the interactions between urban design and crime. This study investigates the impact of micro built environment (MBE) elements and place perceptions on crime occurrences in Toronto using street view imagery (SVI) data and machine learning models. We used logistic regression models and an XGBoost (Version 1.7.5) classifier to assess the significance of MBE and perception variables in classifying crime and non-crime intersections. Our findings reveal that intersections with criminal activity tend to be related to more mobility-related features, such as roads and vehicles, and fewer natural elements, such as vegetation. The "beautiful" and "depressing" perceptions emerged as the most significant variables in explaining crime events, surpassing the commonly studied "safety" perception. The XGBoost model achieved 86% accuracy, indicating that MBE and perception variables are strong predictors of crime risk. These findings suggest that enhancing vegetation and improving street aesthetics could serve as effective crime prevention measures in urban environments. However, limitations include the general nature of the perception model and the reliance on aggregated crime data. Future research should incorporate local perceptions and fine-scale crime data to provide more tailored insights for urban planning and crime prevention

Keywords: micro built environment; environment perception; street view imagery; urban crime

Citation: Wu, D.Y.; Wang, J. Analyzing Urban Crime Through Street View Imagery: Insights from Urban Micro Built Environment and Perceptions. *Urban Sci.* **2024**, *8*, 247. https://doi.org/10.3390/urbansci8040247

Academic Editor: Guoray Cai

Received: 23 October 2024
Revised: 4 December 2024
Accepted: 5 December 2024
Published: 7 December 2024

1. Introduction

Crime poses a worldwide challenge, fostering social unrest and hindering urban development. It disrupts the day-to-day life of individuals and, meanwhile, presents a major barrier to achieving long-term city growth and stability. As cities expand and evolve, they face the complex interplay between urbanization, infrastructural development, and rising crime rates. Rapid urban development, when poorly planned or inadequately managed, can create environments where crime thrives, exacerbating social inequalities and straining public resources. This connection between crime and urban development underscores the need for more nuanced analyses of the built environment's role in shaping criminal activity.

As per the police-reported crime statistics in Canada, there has been a rising trend in the crime severity index since 2014 [1]. Research has confirmed that crime negatively affects urban public safety and welfare, with significant short- and long-term impacts on economic growth and public health [2–4]. Understanding the spatial and temporal patterns of crime is crucial for developing effective prevention strategies and mitigating its potential adverse effects on society.

Traditionally, studies have used built environment data, such as census information, to explore the spatial patterns of crime and its relationship with socio-economic factors.

However, this approach often fails to capture the finer details of the built environment around crime hotspots, as it relies on aggregated data from broader geographic units like census blocks. For example, a significant association between the percentage of households having porches, the average number of building stories, and crime was reported in the previous study [5]. The built environment data used in that research was at the census block level, which does not provide detailed information, such as the micro built environment (MBE), about the specific locations of crimes. MBE elements refer to components of human-made urban environment in a fine-scale setting, particularly those found on streets and encountered by people daily, such as trees, traffic lights, and vehicles. Unlike conventional built environment data that represents an entire administrative area and assumes uniformity in crime characteristics across that area, MBE elements provide detailed insights specific to each crime location, acknowledging the unique features present at each site.

The MBE plays an important role when studying crime events as it shapes human perception and impacts people's decisions. The importance of studying the relationship between MBE elements and crime is emphasized by the crime prevention through environmental design (CPTED) theory [6]. CPTED suggests that certain MBE elements are more influential in promoting or inhibiting crime. For example, a study in Portland found that larger trees convey the impression of better maintenance and heightened risk for offenders, thus reducing crime in the vicinity [7]. Other MBE elements, such as lighting, building height, and vegetation, have also been linked to crime in multiple studies [8–11]. Although many MBEs have been examined, few studies have focused on how human perceptions of MBEs relate to crime.

Perceptions were, arguably, closely linked to the built environment. In Jane Jacobs's theory, urban designs that encourage active street life and natural surveillance, such as features like mixed-use buildings, short blocks, and pedestrian-friendly layouts, enable residents and passersby to keep "eyes on the street," thereby deterring crime [12]. Similarly, Oscar Newman's theory relates safety perception with the built environment by promoting design elements like territorial boundaries, natural surveillance, and controlled access, which foster residents' sense of ownership and control over their spaces, thereby discouraging criminal activity and enhancing feelings of safety [13]. Both scholars emphasized the importance of street environment and street layout, which can be substantially proxied via street environment derived from SVI, such as MBEs. Additionally, previous urban studies have found associations between perceived characteristics derived from SVI and MBEs, such as safety, liveliness, and wealth, with outcomes like housing prices and restaurant reviews [14–16]. Researchers have also established connections between perceived safety and crime; for example, neighborhood disorders, such as broken windows, are associated with reduced perceptions of safety and higher crime rates [17,18].

Despite attempts to investigate associations among MBE elements, perceptions, and crime events, many existing studies are limited by small sample sizes and a lack of diversity in neighborhood characteristics [19]. Neighborhoods with different characteristics, such as population density and ethnic diversity, were found to have different crime patterns [20]. Such lack of diversity was due to the high cost of manually collecting and quantifying MBEs and perceptions in the city, as there can be tens of thousands of unique crime locations in a city. Such limitation may lead to inconsistent associations between perceived safety and crime events [5,19,21]. The perceptions were found to be largely impacted by the environmental design; specifically, proper streetscape design was suggested as one of the key methods to reduce the incidence of crime, according to CPTED principles [22]. Perceived safety has been largely explored in previous research, but other perceptions, such as perceived beauty, may also be associated with crime events as perceptions shape one's behavior. Studying how various perceptions are linked to crime incidence can deepen our understanding of the influence of environmental design on crime as suggested by CPTED. This involves examining how perception, shaped by a combination of factors including the MBE, impacts crime rates. To fill the research gap, this study implements a city-wide analysis that captures the MBE and perception characteristics on streets across the city of Toronto to further investigate the associations among MBEs, perception, and crime events.

To capture the characteristics of the street environment within a city, crowdsourced data that covers the street environment with high resolution and stable data quality is needed to make the large-scale analysis possible. Street view imagery (SVI) is crowdsourced data with high-resolution street appearance images similar to what pedestrians see while walking on the street and has been utilized extensively in recent urban studies. With SVI, researchers can evaluate walkability and quantify physical objects, such as trees and noise barriers, in urban areas [23,24]. Researchers found that SVI can identify physical MBE elements as a percentage of SVI objects on the street [25]. SVI provides opportunities for studies to quantify MBEs and perceptions across the city with the high quality and consistency of SVI [26,27]. Being able to quantify MBE elements and perceptions with well-trained models could minimize the bias and cost compared to manually labeled MBEs and perceptions, enabling big-data crime analysis on a larger scale.

This study employs deep learning models to extract MBE factors and estimate human perceptions from SVI around reported crime intersections in Toronto. Binomial logistic regression models were used to investigate the associations among different MBE and perception variables and crime events. The findings reveal that street environmental characteristics, including MBE elements and perceptions, are significantly associated with crime events. These insights can guide efforts to improve urban built environments, helping to mitigate rising crime trends and promote a more sustainable urban environment.

2. Materials and Methods

This study investigates street-level crime events and their association with the surrounding urban built environment in the city of Toronto. The analysis workflow is illustrated in Figure 1. Crime data from 2018, containing street addresses of reported incidents, were obtained from the Toronto Police Open Data [28]. To protect privacy, these addresses were aggregated to the nearest road intersection. The SVI for each road intersection was collected via online street view providers and further used to quantify MBE elements and perceptions by machine learning models. MBEs were extracted from the SVI into 19 types of commonly seen street elements, such as trees, roads, and cars, using the PSPNet model trained on the Cityscape dataset [29,30]. Perceptions were quantified on a scale from 0 to 1, utilizing a machine learning model trained on the Place Pulse 2.0 dataset [31,32]. Meanwhile, census variables at the dissemination area (DA) level, were collected as the control variables. The DA is the smallest census area in Canada, where each DA contains 400 to 700 residences. The average DA size is 0.18 square kilometers in Toronto. The control variables were selected based on suggestions from previous empirical studies [33,34]. To analyze the relationship between crime, SVI-derived MBEs, and perceptions, three analytical models were employed. First, a regression model was used to examine the associations among MBEs, perceptions, and crime events. Second, an XGBoost classifier was deployed to predict the likelihood of crime at all intersections in Toronto.

2.1. Data Collection

The crime data from Major Crime Indicators were obtained from the Toronto Police Open Data, containing 242,879 records of reported crimes from 2014 to 2022 [28]. For this study, crime data from 2018 were selected as the sampling data, as they provide a representative snapshot for analysis. This dataset includes information on premises type, offense type, and the geolocations where the crimes occurred. To protect the privacy of the individuals involved, the geolocations were adjusted to the nearest street intersections as provided in the official open data.

SVI was collected at all three and four-way intersections in Toronto (N = 20,040) to investigate the MBEs close to the crime scenes. For each intersection, SVI was captured from four angles (0, 90, 180, and 270 degrees relative to true north) to obtain a comprehensive 360-degree view and cover the surrounding MBEs. Since this study focuses on the relationship between MBEs and perceptions derived from SVI, and street crime, only crime events with a premises type classified as "outside" (indicating street-level occurrences) were included. After filtering, 2854 intersections with reported crimes were identified. To address potential data imbalance issues during model training, an equal number (2854) of non-crime intersections were randomly selected from all intersections in Toronto. The distribution of these intersections is illustrated in Figure 2.

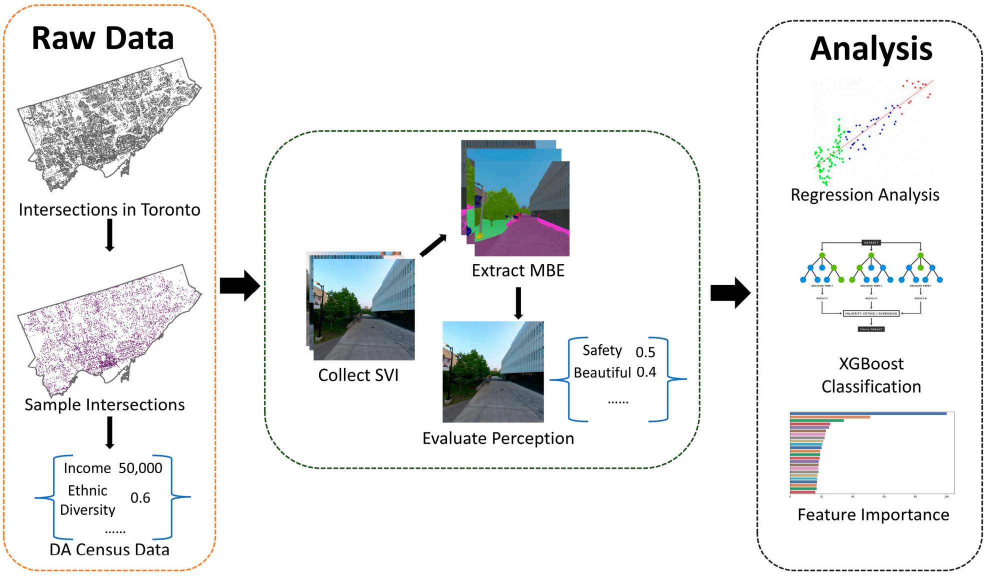

Figure 1. The workflow of research: (1) collect raw intersection and census data; (2) sample SVI in each of the intersections; and (3) analyze the characteristics of crime/non-crime intersections.

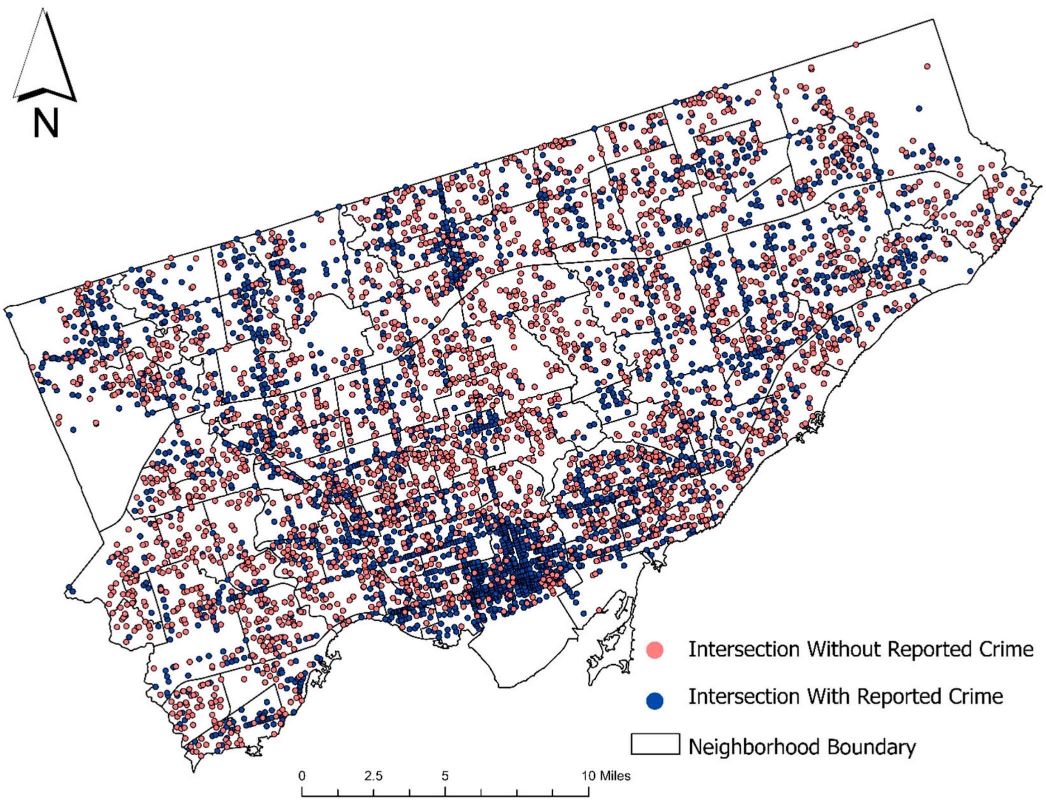

Figure 2. Road intersections that have reported crime records (in red) and no reported crime record (in blue).

Previous studies have found that street crime can be associated with factors such as social disadvantage, ethnic group composition, economic activity, and race [35–37]. Thus, in addition to street-view data, population, income, low-income percentage, and ethnic diversity were extracted from the most recent census data, the Toronto Census Data 2016 [38] at the DA level and used as control variables. The control variables were mapped to every intersection in the DA.

2.2. Micro Built Environment Extraction

MBEs were extracted from SVI using a deep-learning model, specifically, the Pyramid Scene Parsing Network (PSPNet). PSPNet is a deep convolutional neural network that employs convolutional blocks of sizes 1, 2, 3, and 6 to capture high-level details at different scales in the images, combining these details to segment MBEs [30].

The network was pre-trained on the Cityscapes dataset, which contains 25,000 annotated images focusing on urban street scenes [29]. This pre-trained PSPNet model was obtained from GluonCV [39], which provides implementations of top-performing computer vision models. The model's performance was evaluated using pixel accuracy and mean intersection over union (mIoU). Pixel accuracy measures the percentage of correctly classified pixels, while mIoU assesses the overlap between ground truth labels and prediction labels. The PSPNet model achieved 96.4% pixel accuracy and a 79.9% mIoU.

The Cityscapes dataset includes 19 object classes within 7 categorical street classes. Although another dataset, ADE20K offers 150 object classes (including both indoor and street objects) [40], its best model achieves only a 57% mIoU, significantly lower than Cityscapes. Since this study focuses solely on street MBEs, using ADE20K could introduce noise by misclassifying street objects as indoor elements. Therefore, the Cityscapes dataset was more suitable for this study's requirements. By deploying the trained PSPNet model with the Cityscapes dataset, 19 types of MBE are extracted from the SVIs. These 19 types of MBEs represent physical built environment elements commonly observed on the street, such as terrain, poles, and fences. A more detailed description and list of the MBE variables are listed in Appendix B. The extracted MBEs are quantified with values ranging from 0 to 1, indicating the percentage of pixels corresponding to each MBE type within an SVI. For example, a value of 0.1 for the "fence" MBE means that fences occupy 10 percent of the pixels in the image.

2.3. Perception Score

The perception score (Q-score) was derived from the Place Pulse 2.0 dataset, which contains 1,170,000 pairwise comparisons of SVI. Each comparison addressed questions such as "Which place looks safer?" or "Which place looks more beautiful?" across six perception variables: safety, beauty, liveliness, boredom, depression, and wealth [31]. Participants selected the image that best matched their perception without receiving predefined definitions of these variables. The pairwise comparisons were then transformed into a score scale ranging from 0 to 1 using the strength-of-schedule algorithm. For example, a safety score of 0 indicates that a street view feels very unsafe, while a score of 1 suggests that it feels very safe. The strength-of-schedule algorithm, often used in sporting competitions to evaluate team strength based on their matchups, was employed to compute the transformed Q-score. Detailed equations for this transformation can be found in Appendix A.

2.4. Classifying Crime Intersections Using Statistical and Machine Learning Models

Using the data obtained, three logistic regression models were built to explore the impact of MBE and place perception on street crime by classifying whether an intersection had a crime record. Additionally, six logistic regression models were constructed to assess the significance and performance of each perception variable in explaining crime events. An XGBoost model was also developed using parameters developed in Appendix C to classify crime and non-crime intersections using all variables, with its prediction accuracy validated through five-fold cross-validation. The XGBoost model is a powerful and scalable machine learning model that builds ensembled decision trees to iteratively improve accuracy. XGBoost is capable of identifying and modeling the non-linear relationships between variables. To identify the factors most relevant for crime intersection classification, feature importance was determined using data impurity-based metrics from the machine learning model. This model was further employed to predict the probability of an intersection being a crime location, with this probability serving as the crime risk indicator for each intersection.

The control variables included the neighborhood's low-income population, young population, unemployment rate, and ethnic origin. The low-income and young populations were normalized as percentages relative to the total population in the DA. Ethnic diversity was assessed using the Hirschman–Herfindahl Index in Equation (1), in which the higher the diversity, the higher the index score. In the equation, A represents the percentage of the Asian population, LA represents the percentage of the Latin American population, AA represents the percentage of the African American population, NV represents the percentage of non-visible minority populations, and O represents others [41]. The control variables used in this study were found to be related to crime events [42–46].

$$Ethnic\ Diversity = 1 - \frac{A^2 + LA^2 + AA^2 + NV^2 + O^2}{(A + LA + AA + NV + O)^2} \tag{1}$$

All the variables used in this study are described in Appendix B. By combining the control variables with MBE and place perception scores separately, this study compared the impact of different variables on street crime activities. To train the model, we checked the multicollinearity with variance inflation factors (VIF) and the correlation of different variables. If the VIF exceeds 10, the model may have potential multicollinearity and certain variables should be excluded from the model. Three models were constructed to investigate the associations between different variables and crime events. Model 1 included only control variables as a baseline, Model 2 incorporated both control variables and MBE data, and Model 3 included control variables along with perception scores. Additionally, six logistic regression models were built to evaluate the performance and significance of each perception variable in combination with the control variables.

3. Results

3.1. Machine Learning Based Feature Importance

The percentage difference between intersections with and without criminal activity is concluded in Figure 3, which indicates that intersections with criminal activity tend to have more buildings and roads but significantly fewer natural elements, including vegetation and terrain. The fitted XGBoost model achieved 86% accuracy when classifying crime and non-crime intersections using five-fold cross-validation. This observation aligns with the result obtained from the XGBoost model's feature importance analysis. In particular, the MBE variables exhibited high feature importance when classifying crime and non-crime intersections, as demonstrated in Figure 4. For instance, traffic lights, buildings, and vegetation rank as the second, fourth, and fifth most important variables.

3.2. Relationship Between MBEs and Street Crime Events

Logistic regression model results for control variables only (Model 1) and combining control variables and MBE variables (Model 2) indicate that the MBE variables greatly contribute to classifying crime and non-crime intersection. This is evidenced by Model 2 having a much lower AIC than Model 1, as shown in Table 1. Mobility-related variables, such as roads, sidewalks, and vehicles, demonstrate a significant positive association with intersections that experience crime events. The positive significance of buildings aligns with the findings in Figure 2, where a higher concentration of buildings at intersections is associated with an increase in the risk of nearby criminal activities. Although vegetation and terrain are both natural elements on the street, the association with crime events came from different directions, as terrain was identified to be positively associated with crime, and vegetation was identified to be negatively associated with crime.

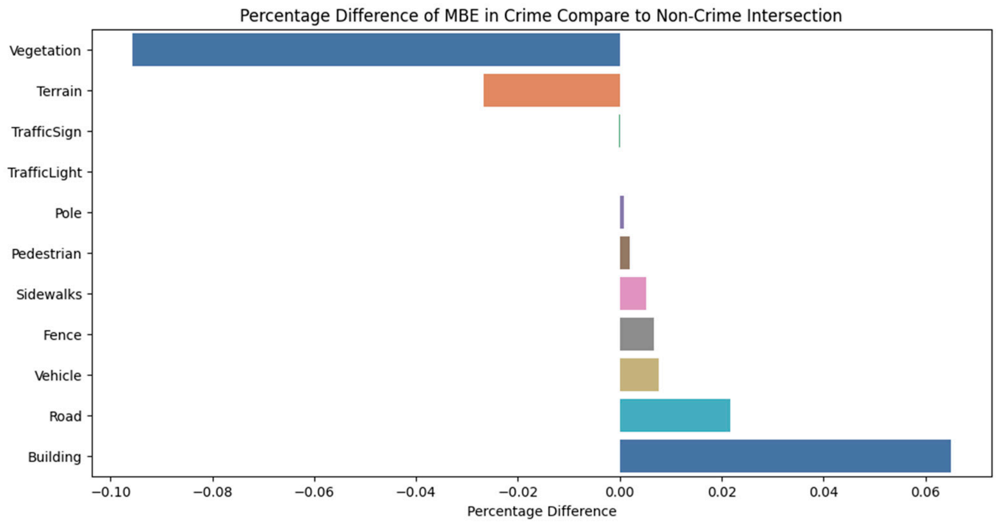

Figure 3. The MBE difference between intersections with crime records and intersections without crime records in 2018. (A positive percentage for an MBE indicates more appearances in crime intersections, and a negative percentage indicates fewer appearances).

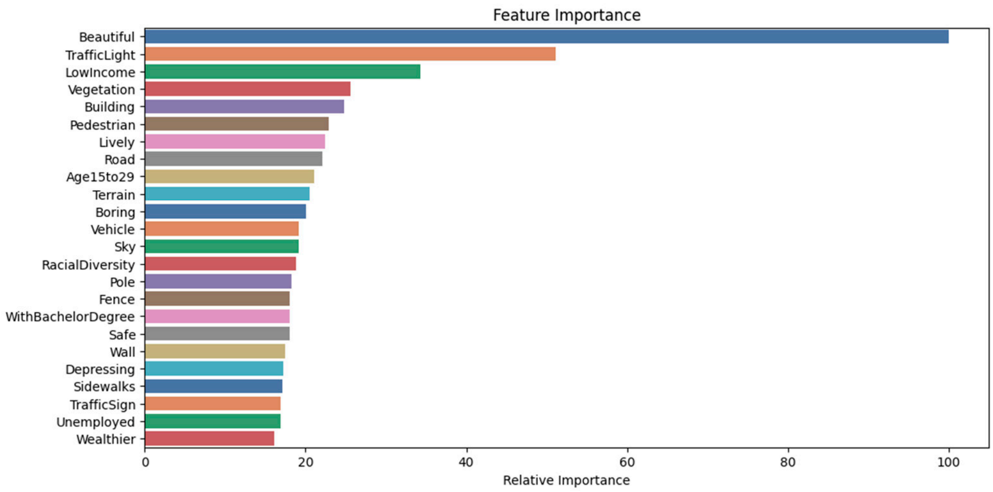

Figure 4. The importance of each feature when classifying crime/non-crime intersections.

3.3. Relationship Between Perception and Street Crime Events

All perception-related variables exhibit significant associations, except for the variable "lively", as shown in Table 2. These findings indicate that individuals typically perceived lower safety, wealth, beauty, and higher depressing and boring elements in the intersections with crime records compared to other intersections. Overall, for models with perception variables (Models 1–6), it is noteworthy that the model containing the beautiful perception (Model 4) has the smallest AIC values, which means the beautiful perception contributes the most to explaining crime events among all other perceptions in Toronto. Additionally, "beautiful" also has the highest feature importance, compared to other perception variables, according to Figure 4.

Table 1. Logistic regression model results for control variables only (Model 1) and combining control variables and MBE variables (Model 2).

	Model 1	**Model 2**
(Intercept)	−2.34 ***	−5.86 ***
Unemployed	−4.77 ***	−1.94
Low Income	5.26 ***	3.71 ***
Racial Diversity	1.45 ***	0.95 ***
Age 15 to 29	3.37 ***	1.99 ***
With Bachelor Degree	1.12 ***	0.09
Pedestrian		19.49 **
Road		10.16 ***
Sidewalks		8.67 ***
Vehicle		10.29 ***
Building		3.52 ***
Fence		8.42 ***
Pole		49.29 ***
Terrain		2.49 *
Traffic Light		106.64
Traffic Sign		−16.09
Vegetation		−0.96 *
AIC	7249	6561.7

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

Table 2. Logistic regression model results for safety perception (Model 1), wealthy perception (Model 2), depressing perception (Model 3), beautiful perception (Model 4), boring perception (Model 5), and lively perception (Model 6). (The control variables are consistent across all models, so they are omitted from the table to save space.)

	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 6**
(Intercept)	−0.64 ***	−1.07 ***	−3.31 ***	−0.60 ***	−3.07 ***	−2.32 **
Control variables	-	-	-	-	-	-
Safety	−2.37 ***					
Wealthy		−2.55 ***				
Depressing			3.11 ***			
Beautiful				−2.32 ***		
Boring					1.20 ***	
Lively						−0.07
AIC	6821	7018	6780	6739	7222	7251

*** $p < 0.001$; ** $p < 0.01$;

4. Discussion

The research findings indicate a significant negative association between safety perception and crime events, along with a significant positive association between MBE variables and crime events, which aligns with prior research [47–49]. This study advances the field by emphasizing that MBE variables offer a more effective explanation for crime events than perception variables. By integrating fine-scale data on urban microenvironments, this study offers compelling evidence that elements of the built environment play a crucial role in shaping crime patterns, potentially even more than subjective perceptions of safety. Unlike prior studies, which often focused on aggregated data, this research dives deeper into how specific MBE factors at crime locations contribute to criminal activity. Furthermore, this study not only examines how individual MBE elements, such as roads, vehicles, and fences, are related to crime, but also explores how perceptions modeled by the MBE are associated with crime. Notably, perception variables, though less effective than the MBE, still explain crime better than using control variables alone. Additionally, this study introduces the novel finding that perceptions of beauty and depression provide a more compelling explanation for crime events than the commonly examined safety perception. This finding provides insights for understanding how the aesthetic and emotional aspects of urban design influence criminal activity, an area that previous studies have largely overlooked.

The positive significance observed in all mobility-related elements, including roads, vehicles, sidewalks, and crime events, provides clear evidence that locations that provide greater accessibility were more likely to experience crime events [50]. This highlights a critical challenge in urban development: while mobility features are essential for ensuring efficient access and movement within cities, they may also inadvertently create opportunities for crime. However, the relationship between fences and poles and crime events varies across studies, potentially due to regional and cultural differences [51]. In North America, fences often serve as a physical barrier to safeguard residents' privacy and security for their properties. The fences can also serve as part of the architectural decorations for buildings, especially for old buildings that are concentrated in Toronto's downtown area [52]. In this study, the fence and poles were positively associated with crime events. This finding implies that elements such as roads, vehicles, sidewalks, fences, and poles can play important roles in influencing crime events, suggesting that areas with higher accessibility and mobility features might also be more susceptible to criminal activities.

These insights carry significant implications for urban planning and community safety. Urban planners can balance accessibility with safety considerations when designing public spaces. For instance, while ensuring efficient mobility through well-structured roads and sidewalks, targeted interventions such as improved lighting, surveillance, and strategic placement of fences can enhance safety. Moreover, fences and poles might be used more thoughtfully, with attention to their placement and design to avoid creating hiding spots or obstructing sightlines, which could help reduce crime opportunities. Incorporating these considerations into urban design could promote safer neighborhoods while still maintaining accessibility and aesthetic appeal.

Previous research suggests that more vegetation could mitigate crimes [4,53,54]. This connection between vegetation and crime was also revealed in the study region. The XGBoost model demonstrated that vegetation was one of the most important variables among other MBE variables in classifying crime and non-crime locations. The significant association between vegetation elements and crime events also appears in the logistic regression model. Furthermore, the association between vegetation and crime events in street intersections aligns with other studies, indicating that vegetation in street intersections is similarly as connected to crime events as vegetation in the middle of street segments. The result of this study suggests that improving vegetation near street intersections can also be an effective way to promote crime prevention, as with the effect of natural elements in the middle of street segments.

The statistically negative association of safety perception with crime events matches previous findings that indicate the connection between safety perception and crime events. In the context of Toronto's crime events, this study suggests the beautiful and depressing perception better explains crime. Due to limited data collection resources, previous research concentrated more on the relationship with the safety perception and did not pay enough attention to the other perceptions. Both safety and beautiful perception variables showed significance in the regression model and exhibited the highest performance among the perception regression models. Consequently, there is a need for more research to examine the relationship between these perceptions and crime as it is unclear what accounts for the better performance of beautiful and depressing.

Previous studies discussed how aesthetics related to crime incidence, where the existence of elements that represent neighborhood disorder, such as broken windows, leads to a negative influence on people's perception of safety, and further impacts street aesthetics [55]. The findings of mobility-related features and perceptions of beauty and depression are closely aligned with the safety perception theories introduced by Jane Jacobs and Oscar Newman [12,13]. Both of their theories suggest that the built environment's appearance can strongly influence crime perception and security. Jacobs's emphasis on vibrant, mixed-use areas with continuous human activity relates to how visually appealing and well-maintained areas attract more “eyes on the street,” deterring crime. Similarly, Newman’s Defensible Space Theory suggests that well-kept environments signal ownership and territoriality, while neglected or visually depressing spaces may invite crime by signaling vulnerability. Both theories, thus, support the idea that beauty and upkeep in the built environment contribute to a lower crime perception, resonating with our study’s findings on the role of visual aesthetics in crime modeling.

Street aesthetics, which is quantified as a beautiful perception in this study, can be a more informative perception than safety. A safe place might also be beautiful, as such, safety perception can be brought by the existence of specific elements instead of the overall street environment. For instance, the presence of a police station can make people feel safer than without the presence of a police officer, disregarding the orientation of elements on the street. However, the presence of elements that relate to street aesthetics, such as a modern building or art crafts, will not have as much impact as the police station on safety perception. Similarly, depressing perception has similar attributes, which can be more related to the overall environment design than the presence of a single element, which aligns with CPTED theory [22]. Studies emphasized people’s avoidance behavior in depressing places, where people try to avoid staying in disorderly places [56]. In contrast, people tend to stay in beautiful places like parks [57].

There are some limitations in this study that could be addressed in a future study. The findings may be limited by the data accuracy from the data collection stage. The crime locations were offset to the nearest intersections, which affected the number of roads and decreased natural elements in the SVI, potentially resulting in over- or underestimation of crime risk in certain areas. In the meantime, the crime data used in this study contain all types of crime that happen in outdoor conditions, such as robbery, theft, and assault. It is worth noticing that the connection between crime and the built environment varies greatly with different types of offenses. This study only focuses on the association between general crime and the built environment in various scales without categorizing crime types. Future studies may explore how the influence of factors found significant in this study, such as beautiful perception, changes across different types of crime. Future studies could utilize fine-scale crime data to mitigate such issues and validate the importance of perceived beauty and depression in crime risk models. The predicted perception represents general perceptions for people worldwide rather than just Toronto residents because the perception prediction model was trained using data collected from participants across the globe. Locals in the city have more knowledge about the characteristics of places, so their perceptions may differ from those predicted. These perception differences could be influenced by the different demographic backgrounds. For example, previous research indicated that low-income populations usually exhibit a lower perception of safety than high-income populations [58,59]. Future studies could investigate the connection between beautiful and depressing perceptions and criminal events. The significant association found in the study has only been tested in the City of Toronto, and the significance level might change if transferred to another city. Additionally, since the definitions of the different perceptions were not provided during the Place Pulse data collection, the participants responded to the perception survey based on their own interpretations. For example, future research could explore the connection between the beautiful/depressing perception and demographic or geographical characteristics. Further studies could validate how perception varies across people with different demographic and social backgrounds.

5. Conclusions

This study benefits urban sustainability by addressing the connection between social unrest and urban environmental factors through a data-driven crime analysis. We investigate the association between MBE elements, place perceptions, and street crime events using data that were extracted from SVI by machine learning models. From the SVI, we confirmed the strong positive significant association between mobility-related elements and crime events. For the perception variables, we found a negatively significant association between safety and crime events, which further provides evidence for existing research regarding the connection between fear of crime and crime events. Among all perception variables, we found that the beautiful and depressing perceptions, compared to the safety perception, could better explain the crime events in the study region. The association between perception variables and crime events suggests that perceptions beyond safety may also be related to crime events, which have yet to be investigated. Comparing the model with MBE elements and the model with perception variables, the model with MBE elements has the lowest AIC value. This finding indicates that the MBE can better explain crime events than the estimated place perception.

The findings of this study provide valuable insights regarding how the shaping of the urban MBEs and place perceptions could potentially affect street crimes. Results recommend that shaping the street environment to a more beautiful and safer sense could potentially reduce crime nearby and promote more sustainable urban development in the long term. The outcomes also provide practical interest to city policymakers who are interested in improving public security and building a sustainable urban environment.

Author Contributions: Conceptualization, D.Y.W. and J.W.; methodology, D.Y.W.; formal analysis, D.Y.W.; data curation, D.Y.W.; writing—original draft preparation, D.Y.W. and J.W.; writing—review and editing, D.Y.W. and J.W.; visualization, D.Y.W. and J.W.; supervision, J.W.; project administration, J.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Computational and Quantitative Social Sciences Grant [DSI-CQSSY3R1P08] from the Data Sciences Institute at the University of Toronto.

Data Availability Statement: These data were derived from the following resources available in the public domain: Toronto Police Open Data (https://data.torontopolice.on.ca/pages/open-data, accessed on 15 April 2024).

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

The strength-of-schedule algorithm estimated the Q-score by the win rate and loss rate for each image with the pairwise comparisons resulting from the following equations [60]. The equation transforms each image's win rate, loss rate, and win rate and loss rate to images that are compared with this image to provide a perception score for the image.

$$Award_i^{(v)} = \frac{1}{p_i^{(v)}} \sum_{j=1}^{k_1} \frac{p_i^{(v)}}{p_i^{(v)} + n_i^{(v)} + e_i^{(v)}} \tag{A1}$$

$$Penalty_i^{(v)} = \frac{1}{n_i^{(v)}} \sum_{j=1}^{k_2} \frac{n_i^{(v)}}{p_i^{(v)} + n_i^{(v)} + e_i^{(v)}} \tag{A2}$$

$$Q_i^{(v)} = \frac{10}{3} \left(\frac{p_i^{(v)}}{p_i^{(v)} + n_i^{(v)} + e_i^{(v)}} + Award_i^{(v)} - Penalty_i^{(v)} + 1 \right) \tag{A3}$$

In the Equations (A1) to (A3), $p_i^{(v)}, n_i^{(v)}$, and $e_i^{(v)}$ represent the number of times image i has been selected as the winner, loser, or equal, respectively, in the survey comparisons for perception v; k_1 and k_2 represent the number of times image i wins and loses a comparison; *Award* is the average win rate where the image i won the comparison; and *Penalty* is the average loss rate where the image i lost the comparison. The final Q-score starts with the win rate of image i, then adds the *Award* and subtracts the *Penalty*. By doing so, the final Q-score is based on the win rate and loss rate of the image compared with image j.

The transformed Q-score is not perfect, and bias could exist in the Q-score. Further reduction of the bias and noise in the Q-score is needed. As 1,169,078 survey results in 110,988 places formed the Q-score, there could be images that have only been compared once, which could lead to a biased perception score. To avoid this issue, only images that have more than $S_i^{(v)}$ votes in total were used. Experiments showed that with the increase of $S_i^{(v)}$, the base model would have a more remarkable performance.

Furthermore, researchers also introduce reducing noise by classifying the data into positive and negative examples with the following equations. The equation transforms the perception score to positive and negative by removing the scores close to the mean value. In Equation (A4), $\mu^{(v)}$ is the mean Q-score and $\sigma^{(v)}$ is the standard deviation in perception v. δ will be a tunable parameter to adjust how strictly we want to reduce the noise data.

$$y_i^{(v)} = \begin{cases} -1 \; if \; Q_i^{(v)} < \mu^{(v)} - \delta\sigma^{(v)} \\ 1 \; if \; Q_i^{(v)} > \mu^{(v)} + \delta\sigma^{(v)} \end{cases} \tag{A4}$$

A Support Vector Classifier (SVC) was used with the classified data to classify positive and negative perceptions. Equation (A5) demonstrates how SVC classifies data, where x is the MBE element extracted from SVI, $y \in \{-1, 1\}$, which represents positive/negative perceptions, and K is the kernel function. Training the SVC for every combination, we found that $S_i^{(v)} = 8$, $\delta = 1.8$, $K = RBF$ gave the highest accuracy, 79.44%, among all other settings in safety perception. The model has a similar performance with other perceptions as well.

$$f(x,\ y, K) = sgn\left(\sum_{i=0}^{N} y_i K(x) + b\right) \tag{A5}$$

Appendix B

Table A1. Description of all variables used in this study.

Variable	Description	Unit
Road	Percentage of roads in SVI	Percentage (%)
Sidewalk	Percentage of sidewalks in SVI	Percentage (%)
Building	Percentage of buildings in SVI	Percentage (%)
Wall	Percentage of walls in SVI	Percentage (%)
Fence	Percentage of fences in SVI	Percentage (%)
Pole	Percentage of poles in SVI	Percentage (%)
Traffic Light	Percentage of traffic lights in SVI	Percentage (%)
Traffic Sign	Percentage of traffic signs in SVI	Percentage (%)
Vegetation	Percentage of vegetation (vertical greenness, e.g., trees and hedge) in SVI	Percentage (%)
Terrain	Percentage of terrain (horizontal greenness, e.g., grass) in SVI	Percentage (%)
Sky	Percentage of the sky in SVI	Percentage (%)
Person	Percentage of people in SVI	Percentage (%)
Rider	Percentage of riders in SVI	Percentage (%)
Car	Percentage of cars in SVI	Percentage (%)
Truck	Percentage of trucks in SVI	Percentage (%)
Bus	Percentage of buses in SVI	Percentage (%)
Train	Percentage of trains in SVI	Percentage (%)
Motorcycle	Percentage of motorcycles in SVI	Percentage (%)
Bicycle	Percentage of bicycles in SVI	Percentage (%)
Safety	Modeled safety perception score	Index, 0–1
Wealthy	Modeled wealthy perception score	Index, 0–1
Depressing	Modeled depressing perception score	Index, 0–1
Beautiful	Modeled beautiful perception score	Index, 0–1
Boring	Modeled boring perception score	Index, 0–1
Lively	Modeled lively perception score	Index, 0–1
Unemployed	Percentage of unemployed population in the 2016 census	Percentage (%)
Low Income	Percentage of low-income population in the 2016 census	Percentage (%)
Racial Diversity	The racial diversity index calculated with the Hirschman-Herfindahl Index	Index, 0–1
Age 15 to 29	Percentage of the population between 15 to 29 years old in 2016 census	Percentage (%)
With Bachelor Degree	Percentage of the population with a bachelor degree or higher in the 2016 census	Percentage (%)

Appendix C

Table A2. Parameter chosen for XGBoost classifier.

Parameter	Description	Value
Learning Rate	Controls the step size at each iteration while moving toward a minimum in the loss function.	0.3
N_Estimators	The number of trees or boosting rounds the model will build.	100
Max Depth	Limits the maximum depth of each individual tree in the ensemble.	6
Minimum Child Weight	Controls the minimum number of instances needed in a leaf node for a split to happen.	1
Gamma	Specifies the minimum loss reduction needed for XGBoost to make a split in a tree.	0
Subsample	The fraction of samples used to build each individual tree.	1

References

1. Moreau, G.; Canadian Centre for Justice; Community Safety Statistics. Police-Reported Crime Statistics in Canada. 2021. Available online: https://www150.statcan.gc.ca/n1/pub/85-002-x/2022001/article/00013-eng.htm (accessed on 31 January 2023).
2. Baranyi, G.; Di Marco, M.H.; Russ, T.C.; Dibben, C.; Pearce, J. The Impact of Neighbourhood Crime on Mental Health: A Systematic Review and Meta-Analysis. *Soc. Sci. Med.* **2021**, *282*, 114106. [CrossRef] [PubMed]
3. Detotto, C.; Otranto, E. Does Crime Affect Economic Growth?: Does Crime Affect Economic Growth? *Kyklos* **2010**, *63*, 330–345. [CrossRef]
4. Estrada, M.A.R.; Ndoma, I. How Crime Affects Economic Performance: The Case of Guatemala. *J. Policy Model.* **2014**, *36*, 867–882. [CrossRef]
5. Schweitzer, J.H.; Kim, J.W.; Mackin, J.R. The Impact of the Built Environment on Crime and Fear of Crime in Urban Neighborhoods. *J. Urban Technol.* **1999**, *6*, 59–73. [CrossRef]
6. Lynch, K. *The Image of the City*; MIT Press: Cambridge, MA, USA, 1964; ISBN 0-262-62001-4.
7. Donovan, G.H.; Prestemon, J.P. The Effect of Trees on Crime in Portland, Oregon. *Environ. Behav.* **2012**, *44*, 3–30. [CrossRef]
8. Zhou, H.; Liu, L.; Lan, M.; Yang, B.; Wang, Z. Assessing the Impact of Nightlight Gradients on Street Robbery and Burglary in Cincinnati of Ohio State, USA. *Remote Sens.* **2019**, *11*, 1958. [CrossRef]
9. Lee, I.; Jung, S.; Lee, J.; Macdonald, E. Street Crime Prediction Model Based on the Physical Characteristics of a Streetscape: Analysis of Streets in Low-Rise Housing Areas in South Korea. *Environ. Plan. B Urban Anal. City Sci.* **2019**, *46*, 862–879. [CrossRef]
10. Wolfe, M.K.; Mennis, J. Does Vegetation Encourage or Suppress Urban Crime? Evidence from Philadelphia, PA. *Landsc. Urban Plan.* **2012**, *108*, 112–122. [CrossRef]
11. Wang, R.; Cleland, C.L.; Weir, R.; McManus, S.; Martire, A.; Grekousis, G.; Bryan, D.; Hunter, R.F. Rethinking the Association between Green Space and Crime Using Spatial Quantile Regression Modelling: Do Vegetation Type, Crime Type, and Crime Rates Matter? *Urban For. Urban Green.* **2024**, *101*, 128523. [CrossRef]
12. Jane, J. *The Death and Life of Great American Cities*; Random House: New York, NY, USA, 1961; ISBN 0-679-74195-X.
13. Newman, O. *Creating Defensible Space*; DIANE Publishing: Darby, PA, USA, 1997; ISBN 978-0-7881-4528-5.
14. Qiu, W.; Zhang, Z.; Liu, X.; Li, W.; Li, X.; Xu, X.; Huang, X. Subjective or Objective Measures of Street Environment, Which Are More Effective in Explaining Housing Prices? *Landsc. Urban Plan.* **2022**, *221*, 104358. [CrossRef]
15. Zhang, F.; Zu, J.; Hu, M.; Zhu, D.; Kang, Y.; Gao, S.; Zhang, Y.; Huang, Z. Uncovering Inconspicuous Places Using Social Media Check-Ins and Street View Images. *Comput. Environ. Urban Syst.* **2020**, *81*, 101478. [CrossRef]
16. Kruse, J.; Kang, Y.; Liu, Y.-N.; Zhang, F.; Gao, S. Places for Play: Understanding Human Perception of Playability in Cities Using Street View Images and Deep Learning. *Comput. Environ. Urban Syst.* **2021**, *90*, 101693. [CrossRef]
17. 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.* **2021**, *207*, 104003. [CrossRef]
18. Cinar, E.A.; Cubukcu, E. The Influence of Micro Scale Environmental Characteristics on Crime and Fear. *Procedia Soc. Behav. Sci.* **2012**, *35*, 83–88. [CrossRef]
19. Ogneva-Himmelberger, Y.; Ross, L.; Caywood, T.; Khananayev, M.; Starr, C. Analyzing the Relationship between Perception of Safety and Reported Crime in an Urban Neighborhood Using GIS and Sketch Maps. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 531. [CrossRef]
20. Hipp, J.R.; Wo, J.C.; Kim, Y.-A. Studying Neighborhood Crime across Different Macro Spatial Scales: The Case of Robbery in 4 Cities. *Soc. Sci. Res.* **2017**, *68*, 15–29. [CrossRef] [PubMed]
21. Pánek, J.; Ivan, I.; Macková, L. Comparing Residents' Fear of Crime with Recorded Crime Data—Case Study of Ostrava, Czech Republic. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 401. [CrossRef]
22. Casteel, C.; Peek-Asa, C. Effectiveness of Crime Prevention through Environmental Design (CPTED) in Reducing Robberies. *Am. J. Prev. Med.* **2000**, *18*, 99–115. [CrossRef]
23. Yin, L.; Wang, Z. Measuring Visual Enclosure for Street Walkability: Using Machine Learning Algorithms and Google Street View Imagery. *Appl. Geogr.* **2016**, *76*, 147–153. [CrossRef]

24. Zhang, K.; Qian, Z.; Yang, Y.; Chen, M.; Zhong, T.; Zhu, R.; Lv, G.; Yan, J. Using Street View Images to Identify Road Noise Barriers with Ensemble Classification Model and Geospatial Analysis. *Sustain. Cities Soc.* **2021**, *78*, 103598. [CrossRef]
25. Chudyk, A.M.; Winters, M.; Gorman, E.; McKay, H.A.; Ashe, M.C. Agreement between Virtual and In-the-Field Environment Audits of Assisted Living Sites. *J. Aging Phys. Act.* **2014**, *22*, 414–420. [CrossRef] [PubMed]
26. Chen, X.; Li, G.; Mehmood, M.S.; Jin, A.; Du, M.; Xue, Y. Using Street View Images to Examine the Impact of Built Environment on Street Property Crimes in the Old District of CA City, China. *Discret. Dyn. Nat. Soc.* **2023**, *2023*, e1470452. [CrossRef]
27. Huang, J.; Fei, T.; Kang, Y.; Li, J.; Liu, Z.; Wu, G. Estimating Urban Noise along Road Network from Street View Imagery. *Int. J. Geogr. Inf. Sci.* **2023**, *38*, 128–155. [CrossRef]
28. Major Crime Indicators. Available online: https://data.torontopolice.on.ca/pages/major-crime-indicators (accessed on 15 April 2024).
29. Cordts, M.; Omran, M.; Ramos, S.; Rehfeld, T.; Enzweiler, M.; Benenson, R.; Franke, U.; Roth, S.; Schiele, B. The Cityscapes Dataset for Semantic Urban Scene Understanding. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 3213–3223.
30. Zhao, H.; Shi, J.; Qi, X.; Wang, X.; Jia, J. Pyramid Scene Parsing Network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 2881–2890.
31. Dubey, A.; Naik, N.; Parikh, D.; Raskar, R.; Hidalgo, C.A. Deep Learning the City: Quantifying Urban Perception at a Global Scale. In Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 11–14 October 2016; Springer: Cham, Swizerland, 2016; pp. 196–212.
32. Urban Perception (Place Pulse, Streetscore, Streetchange). Available online: https://centerforcollectivelearning.org/urbanperception (accessed on 15 April 2024).
33. Zhou, H.; Liu, L.; Lan, M.; Zhu, W.; Song, G.; Jing, F.; Zhong, Y.; Su, Z.; Gu, X. Using Google Street View Imagery to Capture Micro Built Environment Characteristics in Drug Places, Compared with Street Robbery. *Comput. Environ. Urban Syst.* **2021**, *88*, 101631. [CrossRef]
34. Dugato, M. Understanding the Multi-Level Interactions Between Physical Environment and Neighborhood Characteristics in Assessing Vulnerability to Crime in Micro-Places. *Crime Delinq.* **2022**, *68*, 1923–1947. [CrossRef]
35. Bursik, R.J., Jr. Social Disorganization and Theories of Crime and Delinquency: Problems and Prospects. *Criminology* **1988**, *26*, 519–552. [CrossRef]
36. Lan, M.; Liu, L.; Eck, J.E. A Spatial Analytical Approach to Assess the Impact of a Casino on Crime: An Example of JACK Casino in Downtown Cincinnati. *Cities* **2021**, *111*, 103003. [CrossRef]
37. Shaw, C.R.; McKay, H.D. *Juvenile Delinquency and Urban Areas: A Study of Rates of Delinquency in Relation to Differential Characteristics of Local Communities in American Cities (1969)*; Routledge: London, UK, 2010; ISBN 0429245874.
38. Toronto Open Data. Available online: https://www.toronto.ca/city-government/data-research-maps/open-data/ (accessed on 15 April 2024).
39. Guo, J.; He, H.; He, T.; Lausen, L.; Li, M.; Lin, H.; Shi, X.; Wang, C.; Xie, J.; Zha, S. GluonCV and GluonNLP: Deep Learning in Computer Vision and Natural Language Processing. *J. Mach. Learn. Res.* **2020**, *21*, 1–7.
40. Zhou, B.; Zhao, H.; Puig, X.; Fidler, S.; Barriuso, A.; Torralba, A. Scene Parsing through Ade20k Dataset. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 633–641.
41. Hirschman, A.O. *National Power and the Structure of Foreign Trade*; Univ of California Press: Berkeley, CA, USA, 1980; Volume 105, ISBN 0-520-04082-1.
42. Browning, C.R.; Byron, R.A.; Calder, C.A.; Krivo, L.J.; Kwan, M.-P.; Lee, J.-Y.; Peterson, R.D. Commercial Density, Residential Concentration, and Crime: Land Use Patterns and Violence in Neighborhood Context. *J. Res. Crime Delinq.* **2010**, *47*, 329–357. [CrossRef]
43. Chiricos, T.G. Rates of Crime and Unemployment: An Analysis of Aggregate Research Evidence. *Soc. Probl.* **1987**, *34*, 187–212. [CrossRef]
44. Lochner, L. EDUCATION, WORK, AND CRIME: A HUMAN CAPITAL APPROACH*. *Int Econ. Rev* **2004**, *45*, 811–843. [CrossRef]
45. Papps, K.; Winkelmann, R. Unemployment and Crime: New Evidence for an Old Question. *N. Z. Econ. Pap.* **2000**, *34*, 53–71. [CrossRef]
46. Stucky, T.D.; Ottensmann, J.R. LAND USE AND VIOLENT CRIME*. *Criminology* **2009**, *47*, 1223–1264. [CrossRef]
47. Anderson, J.M.; MacDonald, J.M.; Bluthenthal, R.; Ashwood, J.S. Reducing Crime by Shaping the Built Environment with Zoning: An Empirical Study of Los Angeles. *Univ. Pa. Law Rev.* **2013**, *161*, 699–756. [CrossRef]
48. Chalfin, A.; Hansen, B.; Lerner, J.; Parker, L. Reducing Crime through Environmental Design: Evidence from a Randomized Experiment of Street Lighting in New York City. *J. Quant. Criminol.* **2021**, *38*, 127–157. [CrossRef]
49. Warr, M. Fear of Crime in the United States: Avenues for Research and Policy. *Crim. Justice* **2000**, *4*, 451–489.
50. Rhodes, W.M.; Conly, C. Crime and Mobility: An Empirical Study. In *Principles of Geographical Offender Profiling*; Routledge: London, UK, 2017; pp. 143–164. ISBN 1315246082.
51. Hipp, J.R.; Lee, S.; Ki, D.; Kim, J.H. Measuring the Built Environment with Google Street View and Machine Learning: Consequences for Crime on Street Segments. *J. Quant. Criminol.* **2021**, *38*, 537–565. [CrossRef]
52. Fumia, D. Divides, High Rise and Boundaries: A Study of Toronto's Downtown East Side Neighbourhood. *Ethnologies* **2010**, *32*, 257–289. [CrossRef]

53. Burley, B.A. Green Infrastructure and Violence: Do New Street Trees Mitigate Violent Crime? *Health Place* **2018**, *54*, 43–49.
54. Jing, F.; Liu, L.; Zhou, S.; Song, J.; Wang, L.; Zhou, H.; Wang, Y.; Ma, R. Assessing the Impact of Street-View Greenery on Fear of Neighborhood Crime in Guangzhou, China. *Int. J. Environ. Res. Public Health* **2021**, *18*, 311. [CrossRef]
55. Millie, A. Aesthetics and Crime. In *Philosophical Criminology*; Policy Press: Bristol, UK, 2016; pp. 51–68. ISBN 978-1-4473-2373-0.
56. Zhou, H.; Liu, L.; Wang, J.; Wilson, K.; Lan, M.; Gu, X. A Multiscale Assessment of the Impact of Perceived Safety from Street View Imagery on Street Crime. *Ann. Am. Assoc. Geogr.* **2024**, *114*, 69–90. [CrossRef]
57. Zhou, H.; Wang, J.; Wilson, K. Impacts of Perceived Safety and Beauty of Park Environments on Time Spent in Parks: Examining the Potential of Street View Imagery and Phone-Based GPS Data. *Int. J. Appl. Earth Obs. Geoinf.* **2022**, *115*, 103078. [CrossRef]
58. Lauritsen, J.L. The Social Ecology of Violent Victimization: Individual and Contextual Effects in the NCVS. *J. Quant. Criminol.* **2001**, *17*, 3–32. [CrossRef]
59. Will, J.A.; McGrath, J.H. Crime, Neighborhood Perceptions, and the Underclass: The Relationship between Fear of Crime and Class Position. *J. Crim. Justice* **1995**, *23*, 163–176. [CrossRef]
60. Zhang, F.; Zhou, B.; Liu, L.; Liu, Y.; Fung, H.H.; Lin, H.; Ratti, C. Measuring Human Perceptions of a Large-Scale Urban Region Using Machine Learning. *Landsc. Urban Plan.* **2018**, *180*, 148–160. [CrossRef]