RESEARCH ARTICLE

Classifying crime places by neighborhood visual appearance and police geonarratives: a machine learning approach

Md Amiruzzaman[1] · Andrew Curtis[2] · Ye Zhao[1] · Suphanut Jamonnak[1] · Xinyue Ye[3]

Abstract

The complex interrelationship between the built environment and social problems is often described but frequently lacks the data and analytical framework to explore the potential of such a relationship in different applications. We address this gap using a machine learning (ML) approach to study whether street-level built environment visuals can be used to classify locations with high-crime and lower-crime activities. For training the ML model, spatialized expert narratives are used to label different locations. Semantic categories (e.g., road, sky, greenery, etc.) are extracted from Google Street View (GSV) images of those locations through a deep learning image segmentation algorithm. From these, local visual representatives are generated and used to train the classification model. The model is applied to two cities in the U.S. to predict the locations as being linked to high crime. Results show our model can predict high- and lower-crime areas with high accuracies (above 98% and 95% in first and second test cities, accordingly).

Keywords Geonarrative · Machine learning · Semantic segmentation · Street-view image analysis · Urban crime

Introduction

The nature of "place" plays a vital role when it comes to understanding the location and context for many social problems. For instance, opioid overdoses were a major topic of discussion prior to Covid-19. Now, as we move through the

✉ Md Amiruzzaman
   mamiruzz@kent.edu

1   Kent State University, Kent, USA

2   Case Western Reserve University, Cleveland, USA

3   Texas A & M University, College Station, USA

⚭ Springer

various stages of the epidemic, there is an indication that this situation has worsened [87]. This leads to questions such as why and where will this occur? Understanding such contextualized locations are vital to target intervention. The crime landscape can be classified in different ways, for example as a micro-environment or macro-environment [67, 86]. For example, a city, zipcode, and neighborhood areas can be defined as a *macro-environment*. In contrast, a *micro-environment* is more granular and can be defined as a place within any of those areas. Micro-environment crime location classification can be a challenging task due to a lack of data and associated restrictions such as confidentiality [40]. These spaces can be classified in different ways, one of the most obvious being its visual appearance [47, 76], the most famous of which is the theory of "broken windows" that continues to inform current research [85, 92]. However, an under-researched aspect of linking place-based visual imagery to crime involves AI, no significant study was found that tried to use both. In this study, we address this gap by classifying granular scale crime places based on potential connections to activities such as where drugs are purchased, where drug use occurs, and where overdoses will occur most frequently [25].

If successful, the identification of potential drug overdose locations might improve intervention, such as knowing where to place Project Dawn kits [71]. This same logic applies across a variety of other health and crime examples, for example knowing where people feel they are or are less safe [70]. While there has been considerable research on these topics generally this work takes place at a single location with a suggested transferability of findings (such as street lighting) to other locations [62]. Fewer studies have considered the transferability of these findings to different locations.

To achieve this, granular detailed primary data needs to be collected in the form of environmental audits or participant observations [31, 52]. To this end, this paper will leverage previously collected geonarratives to acquire fine-scale multi-time period contextualized data, an approach which has successfully been used to understand the heterogeneous variations in a variety of different environments [2, 21, 25, 26, 41, 49, 55]. Advancing this body of work, and addressing the topic of transferability of findings, this paper will present an automatic classification of contextualized locations deemed to be important to explain negative localized events and then transfer these findings to other test locations. More generally, a further contribution is that automation in crime place classification could provide faster and more accurate results while also reducing human overheads.

To do this we extend previous geonarrative research focused on crime landscapes with an AI-based Google Street View (GSV) image analysis to classify multiple urban places. The AI-based image segmentation tools were used in some social applications (see "Semantic segmentation and applications" section), but they were not explicitly applied to crime place classification. More specifically, our approach and contributions of this study are as follows:

- Locations are evaluated by local police officers who provide professional insights especially related to drug activities using a geonarrative approach. These geonarratives are processed to label specific places as high-crime or lower-crime.

- Instead of linking a described location to its exact image, a "fuzzy" classification occurs of the place using a group of images extracted from the environment surrounding the single place. In this way, a more transferable holistic representation of that type of location is acquired.
- Semantic segmentation based on a deep learning algorithm is used to extract semantic categories (sky, greenery, building, etc.) from these neighborhood images. A location visual representative is then computed to model the environmental features of the place. We study different ways to define the representative and identify the essential semantic categories that can lead to a good classification template.
- The location classification of high-crime and lower-crime areas is implemented by training a ML classification model with GSV images and geonarratives from several police officers patrolling the same set of neighborhoods. Multiple ML algorithms will be tested and compared, before the most successful is used to identify similar spaces in a different city where validation occurs using police report data.
- We further investigate the usability and limitation of the model by testing it across various other US cities with differing urban characters, using local crime indexes to gauge the performance of classification in each location.

Related work

Linking crime to detailed landscapes

Fear of crime is a product of actual and perceived threats, environmental and human based, and that can negatively impact the quality of life [13, 16, 77, 83]. Arguably being able to identify and understand the geographic nature of these fears and actual risks can lead to more effective intervention strategies. However, the required data and associated knowledge, at such fine sub-neighborhood scales are often hard to acquire [11, 61, 91]. For example, the risk of violence or where drug overdoses will occur is linked to a variety of different environmental factors, such as the quality of housing stock, local vegetation, lighting, open and dense spaces, and the interrelationship between all of these.

The local perception of what this mix means in terms of risk translates into how, where, and when people conduct their daily activity [32, 77, 78]. An alternative conceptualization is that this mix results in a landscape of actual and perceived criminal opportunity and victims [89]. While there is a rich literature that has delved into such interconnections [73], especially the importance of micro spaces [12], and patterns of opportunity and victims [18], less has been attempted in developing more transferable rules. Yet, given the challenge in finding detailed local data, alternative more ubiquitous solutions to gauge such localized risk is required.

To effectively achieve this, we also have to add spatial context; it is not enough to just find overlay associations of where crimes and environments intersect, but rather we need to know why they occur there. For example, while we may know on which street a rape or a drug overdose has occurred, it is far harder to understand that event

in terms of the knowledge that can be transferred elsewhere. Advances to more traditional crime data analysis include both big data [10, 30] and primary data solutions using new field methods. In this paper, we leverage aspects from both of these advances [74, 75].

Ground level observations and geonarratives

Advances in online spatialized ground-level imagery, for example, GSV and the advances in global positioning system (GPS) cameras have opened various possibilities for auditing within neighborhood environments for different time periods [19, 75, 79]. One frequently used source for these audits is GSV due to their ubiquitous nature [36, 46]. There are, however, limitations including varying time frames within the imagery, not having recent imagery, and geographic gaps in the collection [5, 27].

An advance on GSV as an audit tool has been putting similar technology in terms of GPS enabled cameras into the hands of local practitioners or researchers so that data can be collected for any space and any time period. Simply put, data can be collected in a more responsive way to the environment being studied—either filling in data gaps, capturing landscapes immediately after temporal inflection points (such as after a political or natural hazard externality), or to investigate changes over short (by month) or longer (by year) durations [20, 24].

A companion data collection is the spatial video geonarrative (SVG). Simply put, by adding an expert "witness," not only are images and coordinates collected, but so to their context [21–23]. This is vital as it not only improves official data with more depth but can be used to fill in the gaps caused when geographic (areas too dangerous to collect in) or institutional bias (not deemed important enough to collect) are at play. For example, an event such as a rape or overdose is more than just a point on a map. It is the location of a geographic story that involves a narrative of the victim, perpetrator, other actors, society, and the physical environment. These types of spatial [39] or "Go along" interviews have proven useful in adding depth for this and other topics notoriously missing or lacking richness in official data sources such as genocide spaces, homelessness, drug overdoses, and infectious disease spread [25, 26]. SVG is a qualitative GIS [43], and mixed-method [48, 80] that lends richness to more traditional spatial data and methods.

Indeed, the geonarrative not only provides an insightful commentary of objects and places in the environment but moving through that landscape also helps inspire that commentary [3, 7, 15, 31, 42, 52, 66]. Places that are identified in these narratives can then be mapped because of the associated coordinate information [2]. In this way, an alley is not only described but can be mapped - it is not just a linear object from another spatial data source, but a series of interconnected places where different but interlinked events occur.

SVG can be seen as part of the current theoretical shift to include behavior and physical environment at the micro-space scale to understand how and why events occur [8, 34, 72, 96]. More specifically these methods also collect and analyze data in such a way that interventions can be developed [2, 49]. The advance this paper

makes is combining the advances of both on-the-ground imagery availability with these contexts generating geonarratives in a machine learning environment to make these insights transferable to other locations based on the visual appearance of the landscape.

Semantic segmentation and applications

Our goal was to understand the difference between places in terms of the presence and combination of visible objects. For example, two places may differ in terms of the amount of greenery, building type, or quality of the building. A GSV image from a commercial area may show more buildings and less greenery compared to a residential block. In this study, we wanted to see if there were any differences between high-crime and lower-crime areas based on their semantic segmentation information (SSI) which is the extraction of objects using computer vision. To do this, AI-based models can be used to predict object types within an image and then provide associated and transferable labels [95].

More specifically, semantic segmentation methods label pixels in specific regions of an image for known objects, then scene parsing tools segment and label the whole image within semantic categories. Different deep learning methods have been successfully applied to achieve this including DeepLab [17]), SegNet [6], DPN [59], LRR [35], Piecewise [58], and PSPNet [95].

Other research has used SSI from images of urban environments to understand and visualize different patterns [19, 24, 57, 60, 69, 82]. For instance, Odgers et al. [69] investigated visual indicators of economic variation; more greenery was associated with higher median home prices. Similar findings were reported by Li et al. [57], while Ye, Zeng, Shen, Zhang, and Lu [94] quantitatively measured the perceptual-based visual quality of streets. We intend to extend these approaches to show how semantic categories (extracted from GSV images around known event locations) can also be used to classify *potential* crime activities in other locations.

Methodological framework

In order to develop an effective transferrable classification scheme, it is important to expand the area of interest beyond too specific an image. For example, while a single streetlight may be known locally as where violence occurs, it is important to capture the immediate surroundings of that location as it is not useful to identify all streetlights as being dangerous. Therefore, when classifying an object, it is not wise to decide about a class based on a single object [28, 50]. To achieve this goal, a more holistic approach is needed to summarize the area in terms of multiple spatial objects and their interconnection.

Figure 1 illustrates our approach for classifying places associated with crimes. First, the insights of police officers who patrol city streets on a daily basis are captured as geonarratives (Fig. 1a).
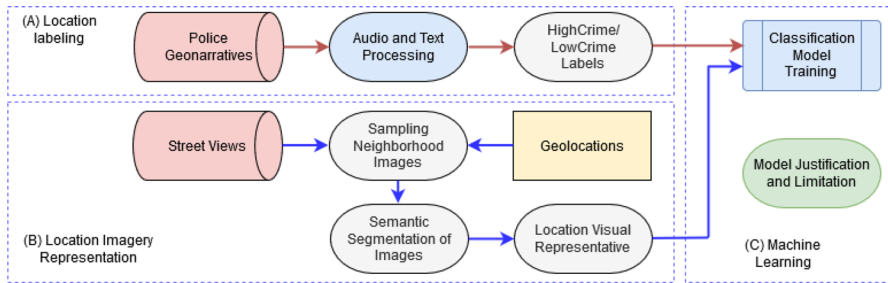
Fig. 1 Overview of the approach. **a** Showing labeling of geonarratives as high-crime and lower-crime, **b** Use SSI from GSV images to find representative, and **c** Use ML techniques to classify places and validate the findings

These narratives are then classified based on the keywords where police officer described a place as being problematic based on a serious crime or not. While future work can further work on the nuances required to tease out specific crime types, here, to prove the conceptual applicability, we reduce crime locations into this binary of higher or lower levels of crime activity. Second, for each of these location types, GSV images are sampled and extracted and then segmented into categories (e.g., road, sky, greenery, etc.) utilizing an AI based SSI extraction algorithm (Fig. 1b). The achieved semantic representations of the neighborhood images are used to compute location visual representatives, where important subsets of the semantic categories are studied and selected. Third, a ML classification model is established between the visual representatives and the location crime labels, which is tested using multiple ML algorithms (Fig. 1c). We implement this model with GSV and the geonarratives recorded by several police officers in the same city. We further apply this trained model to a different Midwest city where similar places are labeled using a geo-tagged police report dataset. Finally, the model is tested in different geographical areas in the U.S. to examine the usability and limitation of the model for different visual appearances.

Location labeling with police geonarratives

Multiple geonarratives were recorded on police rides for a single U.S. city with a population of about 200,000. The geonarrative data consists of over six months of conversation between the time of 8:00 am to 5:00 pm, and eight different police officers participated to describe crime places. The purpose of these rides was to collect insights regarding the link between the built environment and different types of crime. Explanations about data collection protocols have previously been described [22]. The audio narratives were transcribed into text files. From these narrative files, sentences mentioning specific locations with crimes related to drugs, robbery, theft, etc. were labeled.

Obviously, not all crimes have the same level of severity and we are following a similar classification to the FBI in terms of more severe (violent) and less severe

(property) crimes [9, 33]. For the purpose of this study, acts of violence are used to define high-crime areas and lower-crime areas (meaning crime still had occurred but was not of the highest concern) was matched with property crime events. If a sentence had a violence-related keyword, then the corresponding location was labeled as being a high-crime area. Similarly, the description of property crimes was labeled as signaling a lower-crime area. These locations provide PlaCes Of Interests (PCOIs) with high-crimes and lower-crimes. Moreover, we randomly sampled the city for PCOIs with lower-crime activities (i.e., places are not labeled as high-crime areas). Then, n PCOIs in the city, $P_1, P_2, \ldots, P_n$, are labeled as $P_i \in \{\text{HighCrime|LowerCrime}\}$, $i = 1, 2, \ldots, n$. In our experiment, we use $n = 400$. The details of crime location classification is described in "Crime location classification" section.

Location imagery representation

Place neighborhood sampling

To understand the proximate environment of a PCOI holistically, we focus on the visual appearance of that place as well its neighborhood. To capture the surrounding for a PCOI P_i, a circular neighborhood area is defined as $\Omega(P_i, R)$ with a radius R. The road network inside Ω is retrieved from OpenStreetMap (OSM). Then, m_i Neighborhood Sampling Points (NSP) S_i^j, $j = 1, 2, \ldots, m_i$, are uniformly sampled on these street segments, where S_i^j is γ meters apart from each other. Here, two parameters are specified related to the questions of spatial characteristics:

- R defines the neighborhood size: how big is the neighborhood whose visual appearance can indicate the crime tendency of a location?
- γ defines the sampling resolution: what is the appropriate number of street images needed to represent a neighborhood?

While the correct settings will likely vary by location and will require local expert insight, for this paper we used heuristics to evaluate a set of options to finally decided upon $R = 200$ m and $\gamma = 20$ m (see Fig. 2). The total number of sampling points m_i at each P_i varies in a range between [100, 300]. The total number of GSV images used in our classification is $2\Sigma_{i=1}^n m_i$ (2 for left-side and right-side street views) which in practice is about 200,000 images.

GSV image extraction

GSV provides panoramic street views of most U.S. locations. In order to extract images of actual buildings and landscapes (i.e., side-view), but not the road ahead (i.e., road-view), the heading of each street at each NSP was calculated. Since the default camera angle 0° is fixed to the north; three consecutive NSPs along the street
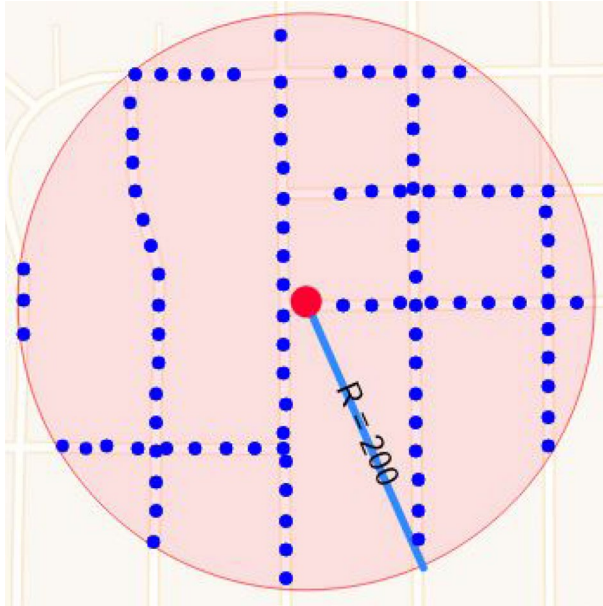
Fig. 2 Data collection procedure in an area of radius R. The red circle represents the target center of the location, and the sample points are shown by blue circles

are utilized: (lat_0, lng_0), (lat_1, lng_1), (lat_2, lng_2), with their latitudes and longitudes. The heading angle at (lat_1, lng_1) is computed as:

$$\theta = \text{atan2}(x, y) \tag{1}$$

where

$$x = \cos(lat_0) \times \sin(|lng_0 - lng_2|),$$
$$y = \cos(lat_0) \times \sin(lat_2)$$
$$- \sin(lat_0) \cos(lat_2) \times \cos(|lng_0 - lng_2|).$$

Based on θ, the side-view angles to the left and right sides are computed and used to retrieve the images from GSV.

Semantic image segmentation

A neural network-based semantic segmentation tool PSPnet [95] is used to extract the SSI from the images (see Fig. 3). Each image is represented by a 19-dimension vector of occupancy values of 19 different object categories (classes), namely, road, sidewalk, building, wall, fence, pole, traffic light, traffic sign, vegetation, terrain, sky, person, rider, car, truck, bus, train, motorcycle, bicycle.

To get the occupancy of an object in an image, we calculate the ratio of the total number of pixels representing the object to the total pixels in the image (see Eq. 2).

Fig. 3 Example of GSV and Semantic segmented images: (left) GSV image and its corresponding (right) semantic segmented image. In the semantic segmented image light-blue color represents as sky, green color as trees, light green color as grass, gray color as pole, dark-gray as building, pink color as sidewalk, blue color as cars, and purple color as road respectively

$$\text{Occupancy of an object}_i = \frac{\text{Pixel count of the object}_i}{\sum_{i=1}^{n} \text{Pixel count of the object}_i} \qquad (2)$$

So, essentially each image has the occupancy values calculated for 19 different categories, which forms the vector to represent a PCOI in this study. If any category is missing in an image, the corresponding value is zero in the vector. This process allows us to represent the significance of different categories of objects present in a scene.

Location visual representative

To train the classification model, a representative of P_i acquired from the neighborhood images of the labeled location is defined. To achieve this, two major questions need to be answered:

- How to extract P_i from the neighborhood image segmentation vectors?
- How to find essential semantic categories that improve the classification results?

Representative identification

In this section we present three different approaches to find representative vectors. First, we show use of Singular-Value Decomposition (SVD) method to find representative vectors (see "SVD method" section). Second, we show use of Principal Component Analysis (PCA) to obtain representative vectors (see "SVD method" section), and Third, we show use of Central Tendency Method to find representative vectors (see "SVD method" section). Details of each approach is presented below.

SVD method

Each PCOI P_i is represented by a matrix of semantic segmentation results, $A_{19 \times 2m_i}$, where m_i is the number of NSPs and 2 is for the left and right side images at each NSP. 19 is the number of semantic categories. Note that m_i is not a fixed number for each location. We apply multiple approaches to extract a good representative of the matrix, and then use it as the characteristic feature in ML classification.

First, A_i is factorized by a Singular-Value Decomposition (SVD) [54] as $A_i = U_{19 \times 19} \Sigma_{19 \times 2m_i} V_{2m_i \times 2m_i}$, where U and V are orthogonal matrices with orthonormal eigenvectors, and Σ is a diagonal matrix with eigenvalues. Then, top k largest values in Σ is selected to reduce dimensionality so that an approximation matrix is achieved:

$$\hat{A}_i = \hat{U}_{19 \times 19} \hat{\Sigma}_{19 \times k} \hat{V}_{k \times k}. \tag{3}$$

Here $\hat{A}_i$ is a $19 \times k$ matrix as the location visual representative of P_i. First, each P_i is represented by the same size matrix so it can be applied in classification. Second, different k values can be set to test the performance of classification. We test from $k = 20$ leading to a large matrix representative, to the smallest value $k = 1$ where $\hat{A}_i$ becomes a 19 dimensional representative vector. In our experiments, the vector representation creates better classification outcome.

PCA method

In data analysis and dimensionality reduction, Principal Component Analysis (PCA) is one of the popular methods. As done in "SVD method" section, in this section we used PCA to reduce dimensions row-wise and find centroid [44], then we found the vector that is closest to the centroid using Eq. 4, where, J is the minimum distance between jth centroid and one of its vectors.

$$J = \min(||v_i^{(j)} - c_j||^2) \tag{4}$$

where v_i is random vectors and c_j is the jth centroid of v_i vectors. Based on the minimum distance between the centroid and vectors associated with a place, we selected a vector that closely resembles the centroid and used that vector as a representative to classify places.

Central tendency method

In geography, statistical measures of central tendency (i.e., mean, and median) have been used in defining a representative location of a small-size areal distribution [38]. Visual appearances in a small neighborhood can be considered as a specific areal distribution.

The location representative at P_i follows the concept of central tendency. First, the image segmentation vectors at NSPs S_i^j are ordered by the distances (Euclidean

or cosine) to the vector at the original place P_i, and then a *median vector* is achieved. Second, a *mean vector* of these vectors is directly computed by averaging the 19 category dimensions. In our experiments, the mean vector leads to higher accuracy in classification than the median vector, and it has better computational efficiency.

Categorical subsets

Initially we extracted 19 semantic categories (i.e., road, sidewalk, building, wall, fence, pole, traffic light, traffic sign, vegetation, terrain, sky, person, rider, car, truck, bus, train, motorcycle, and bicycle) from GSV images and used the SSI to classify places. However, the classification model was complicated due to the 19 different independent variables. Also, we understand that using all 19 categories in the classification model may not be necessary. There may be some categories that do not contribute to classifying crime locations, so it is more efficient to identify a subset of more important classifiers. For example, greenery and open spaces can play a critical role in determining the safety of a neighborhood [56]. We investigate multiple combinations of the dimensions and compare their classification performance in a heuristic way. We find that six major categories out of 19 can achieve the same level of accuracy in crime location prediction. Please see more discussion in "Experiment results and discussion" section.

Crime location classification

We used sentiment analysis to determine high- and lower-crime area related sentences. Sentiment analysis is a process to identify positive and negative sentences using text-mining [14]. In this study, positive sentences are those that are not related to crime, and negative sentences are those related to either violent or property crime. The bag-of-words is a popular text-mining approach to understand the sentiment of a sentence [90].

In this study, the keywords, such as murder, robbery, gun, drugs, and assault (and their variations for example robberies) are used to identify negative sentences, and beautiful, amazing, happy, and family are used to identify positive sentences. A frequency count of positive and negative words was calculated to classify sentences from the geonarrative data. We checked each sentence manually for its rightful category. Because keywords alone do not fully capture an event, to increase accuracy, we manually analyzed those sentences so that they could be classified into high- or lower- crime categories. In the manual analysis, two researchers independently analyzed the result and then discussed on disputed categories and finalized the categorization after an agreement. We discarded neutral sentences (i.e., not related to places) from further analysis. From these, a classification model was trained and tested using a three-step approach to gauge effectiveness and limitations. The results are reported and discussed in "Experiment results and discussion" section.

Step 1: Several supervised ML algorithms including Logistic Regression, Support Vector Machine (SVM), Random Forest (FR), and Naive Bayes (NB) were

trained to recognize high-crime or lower-crime PCOIs in the imagery. About $n = 400$ PCOIs were labeled for the test city, of which 80% were used to train the model, and 20% to validate the classification results. In particular, three comparison experiments were performed with different model inputs:

- Using different location representatives as discussed in "SVD method" section, in order to identify a lower level of crime severity using only the visual characterization of the neighborhood.
- Using the neighborhood representatives versus using only the image segmentation vector at the exact location P_i, in order to justify our approach of using street-level appearance in a neighborhood.
- Using the full 19-dimension representative vectors versus using different combinations of the semantic dimensions, in order to find essential categories linked to linking to tendency of crimes and drug uses.

Step 2: The trained model is applied in another city, approximately 20 miles from the original test environment (i.e., City 1). To evaluate the classification accuracy, locations in this second city (i.e., City 2) are labeled as being high-crime/lower-crime from a police report dataset. The report included both crime and the location of the crime. Using the FBI crime severity, we labeled places as high-crime and lower-crime (Fig. 4).

Step 3: To assess global transferability the trained model is tested on a varying set of different urban environments from across the US. These locations are labelled based on their crime indexes and then used to evaluate the model's effectiveness as the region changes.

To verify our findings and model accuracy, we downloaded the Federal Bureau of Investigation's (FBI's) Uniform Crime Report (UCR) from their official website. Following the guidelines provided by Douglas, Burgess, Burgess, and Ressler [29], we grouped crime incidents information into two categories: violent crime and property crime. We used the UCR data to calculate the crime scores. In this study, we considered criminal activities such as, (a) violent crime (murder, rape, robbery, assault), and (b) property crime (burglary, theft, vehicle theft). We
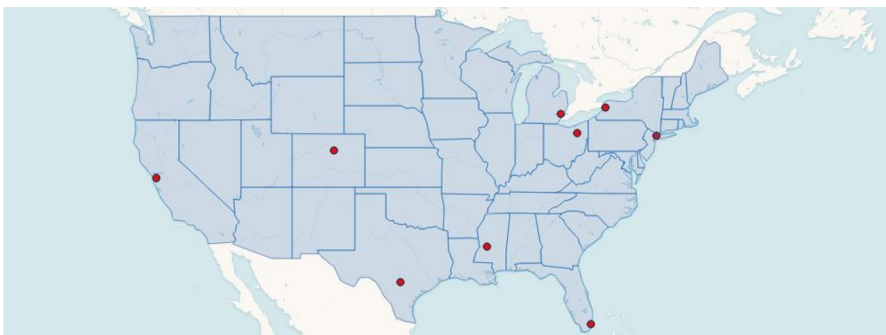

Fig. 4 Model accuracy analysis using different US zipcode locations

divided the number of criminal activities by their respective population to get the crime rate for each type of criminal activity separately, then we normalized the crime rate for 100 residents, this was done so that we can compare crime scores of neighborhoods. All crime types should not be considered the same based [9] so violent crime are weighted differently to property crime. We assigned these "seriousness weights" to the FBI UCR data, and noticed that the average value for violent crimes is three times that of property crimes. Hence, considering the nature and severity of the crime in the crime score calculation we multiplied violent crime by 0.75 and property crime by 0.25, i.e.,

$$\text{crime score} = ((\text{violent crime} \times 0.75) + (\text{property crime} \times .025)) \qquad (5)$$

In addition, we compared neighborhood crime scores to both the proximate neighborhood crime scores and the average national crime score. As a result, a higher crime score means a high-crime area and a lower crime score means a lower-crime area.

Experiment results and discussion

In this section, we present our experimental results: first, we show how our model classified high and lower crime areas within a city. Second, we show how our model performance was evaluated using police recorded crime data. Third, we show our proposed model performance in other geographical areas. Finally, we discuss the model's performance and limitations.
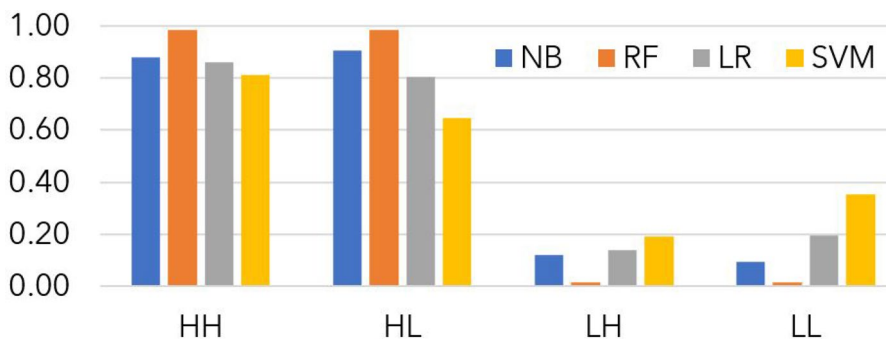

Fig. 5 Classification results in four cases (HH, HL, LH, LL) of $n = 400$ locations in City 1. Y-axis shows the classification rate (0-1) and different ML algorithms are shown in different colors

Classification performance in the test urban environment

Comparing different location visual representatives

Figure 5 shows the classification results for $n = 400$ locations where 200 are labelled as high-crime and another 200 as lower- crime. The mean vector is used and the rates of classification in four cases are shown: (1) HH: high-crime identified as high-crime; (2) HL: high-crime identified as lower-crime; (3) LH: lower-crime identified as high-crime; (4) LL: low- crime identified as lower-crime. We compute a classification accuracy by:

$$\text{Accuracy} = \frac{(HH + LL)}{(HH + HL + LH + LL)} \tag{6}$$

Table 1 reports the classification performance of different location visual representatives. In general, the accuracy of using the mean vector is the highest: LR (83.50%), SVM (72.75%), RF (98.75%), and NB (92.50%), and RF algorithm shows the best performance (so it is used as the default for the other experiments below). In contrast, the median vector only achieves a 46.25% accuracy with the RF algorithm. The reason for this is that the median vector only selects one NSP from the neighborhood which lacks the necessary representation. The accuracy of the SVD method increases from lower than 50% to above 80%, when k (i.e., row-dimension) decreases from 50 to 1. Also, representative vector obtained from PCA helped to achieve better classification accuracy than SVD vectors (see Table 1). The reason may be arguably explained as: Visual appearance in the neighborhood is an anisotropic geometric distribution with sporadic changes. Using a large k includes considerable variation which in turn negatively impacts the classification, while finding a few major components with a small k can remove such variations. In addition, the accuracy of the mean vector shows the best performance which indicates that the classification of social areal attributes

Table 1 Classification accuracy with different location visual representatives in percentage (%)

Method	Representative dimensions	LR	SVM	RF	NB
	20×19	45.50	46.25	48.50	47.50
	15×19	52.50	57.50	57.75	56.25
	10×19	56.50	61.75	62.25	62.50
SVD	5×19	65.25	66.50	77.75	73.50
	4×19	66.75	67.50	78.50	74.75
	3×19	68.25	68.75	79.50	75.25
	2×19	73.50	69.50	82.25	79.50
	1×19	75.25	70.50	84.50	82.25
PCA	1×19	80.75	71.25	87.25	85.75
Mean	1×19	83.50	72.75	98.75	92.50
Median	1×19	42.25	45.50	46.25	48.25

Fig. 6 GSV images are classified as being high-crime (top row) and lower-crime (bottom row). Images A and B are from two different locations, and they have similar visual appearances, however the addition of their neighborhood images (in the same row) can better predict their classes

Fig. 7 Semantic segmented images are classified as being high-crime (top row) and lower-crime (bottom row). Images A and B are from two different locations, and they have similar visual appearances, however the addition of their neighborhood images (in the same row) can better predict their classes

may respond better to an aggregated global representation rather than visual categories. We realize the danger in drawing such a conclusion from this initial work and we intend to further explore this finding.

Comparing with the semantic vector at the exact location (City 1)

When only two GSV images are extracted at P_i, the classification accuracy after training drops to below 50% with all four ML methods. The negative comparison to the 19-dimension mean vector validates our assumption that GSV images in a neighborhood can better predict crime tendency due to being less reliant on heavily weighting a single image. For example Figs. 6, and 7 shows two similar locations though the "context" of their surrounding neighborhood images results in a different classification.

Comparing subsets of semantic dimensions

In order to determine which of the 19 semantic categories best detects crime events, non-significant categories such as pole, traffic-light, traffic-sign, person, rider, truck, bus, train, motorcycle, bicycle are removed. While some of these may play a role in the crime "story" as extracted from the narratives, their general infrequency and therefore lack of pixel portions makes them unsuitable for classification.

In the end, six significant categories (i.e., road, sidewalk, building, fence, vegetation, sky), as shown in the second row of Table 2, reach a similar accuracy level when using all 19 categories.

These 6 dimensions are further explored with road, sidewalk, and building found to be the most important, while the other three can be used to improve the accuracy of the subset.

Model performance in City 2

The trained model with City 1 data is tested for a nearby and, therefore, similar City 2, with a population of about 15,000. A police report data set including geo-tagged crime information in four consecutive years was processed to extract high-crime locations with high activities of gunshot, robbery, drug arrests, etc. A study conducted by Andresen, Linning, and Malleson [4] used random samples to understand spatial concentrations and spatial stability of criminal event data at the micro-spatial unit. The authors mentioned that random sampling can help increase confidence in the results. Random lower-crime locations were similarly sampled in the city as well. With this dataset, the trained classification models are tested on $n = 135$ PCOIs using about 45,196 images. The model achieves good classification accuracy with RF at 95.55%. It shows that this model can be used in another but generally similar urban environment since City 2 is only 20 miles away from City 1. Also, this case study indicates that the model supports the local crime report from the police.

Table 2 Classification accuracy report with semantic categories in percentage (%)

Location representative vector semantic dimensions	Logistic regression	SVM	Random forest	Naïve Bayes
All 19 categories	83.50	72.75	98.75	92.50
Road, sidewalk, building, fence, vegetation, sky	83.25	72.50	98.45	92.35
Road, sidewalk, building, fence, sky	81.50	71.40	97.45	91.75
Road, sidewalk, building, fence, vegetation	78.80	70.25	97.25	91.20
Road, sidewalk, building, fence	76.50	68.25	96.75	89.80
Road, sidewalk, building, vegetation	74.75	67.50	95.45	87.90
Road, sidewalk, building	72.25	64.80	89.75	86.50
Road, building	66.40	63.24	76.50	74.30

Table 3 Classification accuracy report at different areas

Areas	ZRs	HH	HL	LH	LL	Accuracy (%)
NY(HC)	14203, 14210, 14204	97	78	22	3	87
NY(LC)	14209, 14222					
MI(HC)	48201, 48226	88	70	30	12	80
MI(LC)	44221					
CO(HC)	80211	72	78	22	28	75
CO(LC)	80209, 80210					
FL(HC)	33127, 33137	74	72	28	26	73
FL(LC)	33125, 33126					
MS(HC)	39209	69	68	32	31	68
MS(LC)	39211					
CA(HC)	94603	66	67	33	34	65
CA(LC)	94611, 94516					
TX(HC)	78227	61	65	35	39	63
TX(LC)	78258					
NYC(HC)	10453, 10454, 10455	66	71	29	34	68
NYC(LC)	10019, 10023, 10304					

RB is used. HC: High-Crime; LC: Low-Crime

Model performance across geographical regions

While being able to translate findings to similar local urban environments is useful, a test of true transferability is in how the model performs in geographically distinct regions (see Fig. 4). To answer this question, seven US states (Table 3) are selected in which to apply the model. First, a few zipcode regions (ZR) of these states are selected with high and lower crime occurrences based on the FBI Uniform Crime Report. As previously described, (a) violent crime (murder, rape, robbery, assault), and (b) property crime (burglary, theft, vehicle theft) are used to define high and lower crime locations.

Second, from those ZRs 200 PCOIs are sampled in each state, 100 each in high and lower ZRs. These PCOIs are randomly generated to find their accurate geolocations within the ZRs. Their neighborhood images are retrieved from GSV and semantic segmentation is applied to them (see Fig. 4).

Third, the trained model is tested by using the mean vector with six dimensions (i.e., road, sidewalk, building, fence, vegetation, sky) as the location representative vector to classify these PCOIs. Table 3 reports the classification results with RF algorithm in the four cases (HH, HL, LH, LL). The total accuracy is reported for these states ranging from high to low.

Finally, we wanted to see whether our model can classify places (i.e., zipcode areas) in other states than Ohio. To accomplish this task, we used FBI uniform crime report and selected high and low crimes zipcode areas. Also, in our study, we test the model in New York City, which has a markedly different urban landscape to most other US cities. Similarly, 100 high-crime PCOIs and 100

lower-crime PCOIs are selected and classified. This result is shown in the last row of Table 3.

Multi-dimension scaling

Multi-Dimension Scaling (MDS) is a method to convert high-dimensional data to a lower-dimension. In this study, we used MDS methods in "Location visual representative" section to reduce samples and find representative vectors. In this section, we convert the samples obtained in "Location visual representative" section to 2D form and plot them using scatterplots. In machine learning, researchers often use MDS techniques to separate high-dimensional data to reduced or low-dimensions [63, 64], such as Principal Component Analysis (PCA) and t-distributed Stochastic Neighbor Embedding (t-SNE).

In this study, we used both PCA and t-SNE to convert data to 2-dimensions (see Fig. 8). This allowed us to see how high- and lower-crime areas are visually separate from each other. Figure 8 shows most high- and lower-crime places are far from each other. However, a few of them are very close to each other. This helps to understand why the ML model failed to achieve 100% accuracy.

Discussion

It is widely accepted that different "local" or microenvironments are linked to, or are even predictive of crime events [37, 84]. In this paper, we have used machine learning approaches to see if it's possible to use street-level built environment imagery to classify those types of locations in an automated and geographically replicable manner. Evidence from Table 3 shows that this is indeed possible, with the model trained in City 1 achieving an accuracy of 87% and 80% in the other similar "regional" states of New York and Michigan. This is largely due to the similarities in their visual appearance. The model's accuracy decreases though with distance, as does also the visual appearance of sub-neighborhood spaces. In Colorado, Florida, and Missouri, for example, the accuracy falls to 65% to 75%. In California and Texas, the landscapes have even greater dissimilarity to Ohio, reflected in model performance drops below 65%. Again, this can be attributed to many of those micro space elements which have been linked to crime, such as different building types, sidewalk styles, openness, and vegetation types. This is not to say there isn't nuance
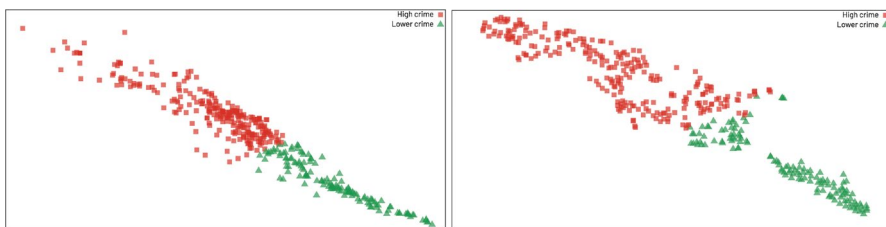

Fig. 8 Scatterplots (left) 2-components from PCA, (right) reduced dimension to 2D using t-SNE

within a straight distance decay effect of visual similarity; the classification accuracy for high-crime PCOIs in New York City was only 68%, since it has dense high-rise buildings and landscapes [1, 53].

The finding that the ecological connection between micro space and crime will vary geographically in terms of content is not surprising. This raises the question of how replicable are the classic crime-and-built environment research [12, 89] to other built environments in terms of replicating their specific detail using a machine learning approach. For example, how transferable are systematic observations of neighborhood spaces beyond their study space [80]? Likewise, can the results from other AI-enhanced single location studies find application beyond their study site [93]? This leads to other questions such as, where does the model accuracy change, meaning where are those boundaries of - regional difference? For example, the results for City 2 were acceptable. It could also be argued that the results for New York State and Michigan were acceptable, or at least the models could be tweaked with minimal local image training. Might it be possible, by understanding these boundaries, to create image libraries in order to tweak classification models regionally so that research in "City A" would need the "Region A" trained model supplemented with 20% additional training?

Even now, the models presented still achieved 60+% prediction for any test environment. Can these results be further mined to identify common location-neutral built environment characteristics and crime drivers? This will be explored in further research, where more specific crime types are investigated using this modeling approach.

Implications

Often local law-enforcement agencies help to classify places as high and low crime areas [88]. However, human-led classification of places may be biased, because of personal belief and misjudgment [68]. Our approach uses AI and computer vision to classify places, which has the advantage through machine learning of increased accuracy and bias-free results [81]. Evidence from our study suggests that among all the semantic categories, road, sidewalk, building, fence, vegetation, and sky are the major categories that can help to determine if a place can be labeled as high/lower crime. For example, the semantic category of fence was commonly found with high-crime areas, which has support in the crime literature by Kim [47] and then Rooney [76].

Likewise, our study also suggests areas with more vegetation have more positive associations and are more visually pleasing. According to Kuppinger [51], more green areas are attractive to home buyers and greenery is related to comfort, quality of lifestyle, and convenience. Conversely, crime tends to locate in less green areas. In our study, greenery was an important category that helped to separate high- and lower-crime areas (see Table 2). Similarly, a study conducted by Katyal [45] noticed that less building and openness of an area help identify crime areas. In other words, the density of the built environment is negatively correlated with high-crime areas. The results of our study indicate that the semantic

category building was one of the major predictors of crime classification (see Figs. 6 and 7). However, while the evidence from our study generally supports these studies in terms of buildings, openness, and crime, we also acknowledge that considerable complexity exists within these overall categories, and that the next steps are to further extract these details. For example, while vegetation in general is a positive association, we know of the research connection between different park types and crime, or the perception of crime [65]. Further image analysis could again consider such nuances in vegetative cover, or even the type of open space.

Conclusions and future work

This paper presents an ML approach to automatically identify the types of places linked to crime based on their visual characteristics, with thematic classification occurring through the mining of police officer geonarratives. By using this contextualized labeling of images, in addition to taking a more "complete" visual of the neighborhood by extracting images around the described location, predictive models were generated that could successfully identify crime environments in other cities beyond the point of data collection. In this way, potentially, model findings can be extrapolated where little local data exists. Even for more data-rich environments, this type of automatic classification approach could be more operational for more resource stretched police departments. A further benefit in model adoption would be the reduction in human-led classification biases.

By comparing these model outputs to different regions, it was found that a distance decay in model performance was evident, with neighboring (and therefore more similar) urban environments being better predicted. Future questions to be explored include, how to define regions based on model accuracy (and where additional training is needed), how the model performs for more specific crime types, whether it is possible to directly establish crime-environment patterns from the images using deep learning, instead of performing semantic segmentation first.

What we have shown is that it is possible to apply models and findings from more data and resource-rich environments to more challenging jurisdictions. Future work might show us where, for example, potential rape locations can be found in any urban environment using minimal additional model training. That type of tool could prove invaluable in getting ahead of, rather than just reporting about, where crimes are likely to occur.

Compliance with ethical standards

Conflicts of interest The authors declare that they have no conflict of interest.

References

1. Adjoian, T., Dannefer, R., & Farley, S. M. (2019). Density of outdoor advertising of consumable products in nyc by neighborhood poverty level. *BMC Public Health*, *19*(1), 1–9.

2. Ajayakumar, J., Curtis, A., Smith, S., & Curtis, J. (2019). The use of geonarratives to add context to fine scale geospatial research. *International Journal of Environmental Research and Public Health*, *16*(3), 515.

3. Anderson, J. (2004). Talking whilst walking: a geographical archaeology of knowledge. *Area*, *36*(3), 254–261.

4. Andresen, M. A., Linning, S. J., & Malleson, N. (2017). Crime at places and spatial concentrations: Exploring the spatial stability of property crime in Vancouver bc, 2003–2013. *Journal of Quantitative Criminology*, *33*(2), 255–275.

5. Bader, M. D., Mooney, S. J., Bennett, B., & Rundle, A. G. (2017). The promise, practicalities, and perils of virtually auditing neighborhoods using google street view. *The Annals of the American Academy of Political and Social Science*, *669*(1), 18–40.

6. Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *39*(12), 2481–2495.

7. Bell, S. L., Phoenix, C., Lovell, R., & Wheeler, B. W. (2015). Using gps and geo-narratives: A methodological approach for understanding and situating everyday green space encounters. *Area*, *47*(1), 88–96. https://doi.org/10.1111/area.12152.

8. Berke, E. M. (2010). Geographic information systems (gis): Recognizing the importance of place in primary care research and practice. *The Journal of the American Board of Family Medicine, 23*(1), 9–12. https://doi.org/10.3122/jabfm.2010.01.090119

9. Blumstein, A. (1974) Seriousness weights in an index of crime. American Sociological Review pp. 854–864

10. Bogomolov, A., Lepri, B., Staiano, J., Letouzé, E., Oliver, N., Pianesi, F., et al. (2015). Moves on the street: Classifying crime hotspots using aggregated anonymized data on people dynamics. *Big Data*, *3*(3), 148–158.

11. Boyd, S. J., Armstrong, K. M., Fang, L. J., Medoff, D. R., Dixon, L. B., & Gorelick, D. A. (2007). Use of a "microecologic technique" to study crime around substance abuse treatment centers. *Social Science Computer Review, 25*(2), 163–173.

12. Brantingham, P.L., & Brantingham, P.J. (1999). A theoretical model of crime hot spot generation. Studies on Crime & Crime Prevention

13. Browning, C. R., Cagney, K. A., & Iveniuk, J. (2012). Neighborhood stressors and cardiovascular health: Crime and c-reactive protein in dallas, usa. *Social Science & Medicine*, *75*(7), 1271–1279.

14. Cambria, E. (2016). Affective computing and sentiment analysis. *IEEE Intelligent Systems*, *31*(2), 102–107.

15. Carpiano, R. M. (2009). Come take a walk with me: The "go-along" interview as a novel method for studying the implications of place for health and well-being. *Health & Place, 15*(1), 263–272.

16. Chandola, T. (2001). The fear of crime and area differences in health. *Health & Place*, *7*(2), 105–116.

17. Chen, L. C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2017). Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *40*(4), 834–848.

18. Clarke, R. V., & Felson, M. (1998). Opportunity makes the thief: Practical theory for crime prevention. *Police Research Series*, *98*, 1–36.

19. Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., Franke, U., Roth, S., & Schiele, B. (2016) The cityscapes dataset for semantic urban scene understanding. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 3213–3223

20. Curtis, A., Blackburn, J. K., Widmer, J. M., & Morris, J. G., Jr. (2013). A ubiquitous method for street scale spatial data collection and analysis in challenging urban environments: mapping health risks using spatial video in haiti. *International Journal of Health Geographics, 12*(1), 21.

21. Curtis, A., Curtis, J. W., Ajayakumar, J., Jefferis, E., & Mitchell, S. (2019). Same space-different perspectives: Comparative analysis of geographic context through sketch maps and spatial video geonarratives. *International Journal of Geographical Information Science*, *33*(6), 1224–1250.

22. Curtis, A., Curtis, J. W., Porter, L. C., Jefferis, E., & Shook, E. (2016). Context and spatial nuance inside a neighborhood's drug hotspot: Implications for the crime-health nexus. *Annals of the American Association of Geographers*, *106*(4), 819–836.

23. Curtis, A., Curtis, J. W., Shook, E., Smith, S., Jefferis, E., Porter, L., et al. (2015). Spatial video geonarratives and health: Case studies in post-disaster recovery, crime, mosquito control and tuberculosis in the homeless. *International Journal of Health Geographics*, *14*(1), 1–15.

24. Curtis, A., & Fagan, W. F. (2013). Capturing damage assessment with a spatial video: An example of a building and street-scale analysis of tornado-related mortality in joplin, missouri, 2011. *Annals of the Association of American Geographers*, *103*(6), 1522–1538.

25. Curtis, A., Felix, C., Mitchell, S., Ajayakumar, J., & Kerndt, P. R. (2018). Contextualizing overdoses in los angeles's skid row between 2014 and 2016 by leveraging the spatial knowledge of the marginalized as a resource. *Annals of the American Association of Geographers*, *108*(6), 1521–1536.

26. Curtis, A., Tyner, J., Ajayakumar, J., Kimsroy, S., & Ly, K. C. (2019). Adding spatial context to the april 17, 1975 evacuation of phnom penh: how spatial video geonarratives can geographically enrich genocide testimony. *GeoHumanities*, *5*(2), 386–404.

27. Curtis, J. W., Curtis, A., Mapes, J., Szell, A. B., & Cinderich, A. (2013). Using google street view for systematic observation of the built environment: analysis of spatio-temporal instability of imagery dates. *International journal of health geographics*, *12*(1), 53.

28. Deng, Z., Zhu, X., Cheng, D., Zong, M., & Zhang, S. (2016). Efficient knn classification algorithm for big data. *Neurocomputing*, *195*, 143–148.

29. Douglas, J. E., Burgess, A. W., Burgess, A. G., & Ressler, R. K. (2013). *Crime classification manual: A standard system for investigating and classifying violent crime*. New York: Wiley.

30. Duan, L., Ye, X., Hu, T., & Zhu, X. (2017). Prediction of suspect location based on spatiotemporal semantics. *ISPRS International Journal of Geo-Information*, *6*(7), 185.

31. Evans, J., & Jones, P. (2011). The walking interview: Methodology, mobility and place. *Applied Geography*, *31*(2), 849–858.

32. Foster, S., Giles-Corti, B., & Knuiman, M. (2014). Does fear of crime discourage walkers? A social-ecological exploration of fear as a deterrent to walking. *Environment and Behavior*, *46*(6), 698–717.

33. Freisthler, B., Ponicki, W. R., Gaidus, A., & Gruenewald, P. J. (2016). A micro-temporal geospatial analysis of medical marijuana dispensaries and crime in long beach, California. *Addiction*, *111*(6), 1027–1035.

34. Garner, A. S., Shonkoff, J. P., Siegel, B. S., Dobbins, M. I., Earls, M. F., Garner, A. S., et al. (2012). Early childhood adversity, toxic stress, and the role of the pediatrician: Translating developmental science into lifelong health. *Pediatrics*, *129*(1), e224–e231. https://doi.org/10.1542/peds.2011-2662.

35. Ghiasi, G., & Fowlkes, C.C. (2016) Laplacian pyramid reconstruction and refinement for semantic segmentation. In: European conference on computer vision, pp. 519–534. Springer

36. Gong, F. Y., Zeng, Z. C., Zhang, F., Li, X., Ng, E., & Norford, L. K. (2018). Mapping sky, tree, and building view factors of street canyons in a high-density urban environment. *Building and Environment*, *134*, 155–167.

37. Groff, E. R., Weisburd, D., & Yang, S. M. (2010). Is it important to examine crime trends at a local "micro" level?: A longitudinal analysis of street to street variability in crime trajectories. *Journal of Quantitative Criminology, 26*(1), 7–32.

38. Hart, J. F. (1954). Central tendency in areal distributions. *Economic Geography*, *30*(1), 48–59.

39. Hawthorne, T. L., & Kwan, M. P. (2013). Exploring the unequal landscapes of healthcare accessibility in lower-income urban neighborhoods through qualitative inquiry. *Geoforum, 50,* 97–106. https://doi.org/10.1016/j.geoforum.2013.08.002

40. Hipp, J. R., Bates, C., Lichman, M., & Smyth, P. (2019). Using social media to measure temporal ambient population: Does it help explain local crime rates? *Justice Quarterly*, *36*(4), 718–748.

41. Jamonnak, S., Zhao, Y., Curtis, A., Al-Dohuki, S., Ye, X., Kamw, F., & Yang, J. (2020).Geovisuals: A visual analytics approach to leverage the potential of spatial videos and associated geonarratives. International Journal of Geographical Information Science pp. 1–21

42. Jones, P., Bunce, G., Evans, J., Gibbs, H., & Hein, J. R. (2008). Exploring space and place with walking interviews. *Journal of Research Practice*, *4*(2), D2–D2.

43. Jung, J. K., & Elwood, S. (2010). Extending the qualitative capabilities of gis: Computer-aided qualitative gis. *Transactions in GIS*, *14*(1), 63–87.

44. Kambhatla, N., & Leen, T. K. (1997). Dimension reduction by local principal component analysis. *Neural computation*, *9*(7), 1493–1516.

45. Katyal, N. K. (2002). Architecture as crime control. *The Yale Law Journal*, *111*(5), 1039–1139.

46. Kelly, C.M., Wilson, J.S., Baker, E.A., Miller, D.K., Schootman, M. (2013) Using google street view to audit the built environment: inter-rater reliability results. *Annals of Behavioral Medicine* **45**(suppl_1), S108–S112

47. Kim, S.K. (2006) The gated community: Residents' crime experience and perception of safety behind gates and fences in the urban area. Ph.D. thesis, Texas A&M University

48. Knigge, L., & Cope, M. (2006). Grounded visualization: Integrating the analysis of qualitative and quantitative data through grounded theory and visualization. *Environment and Planning A*, *38*(11), 2021–2037.

49. Krystosik, A. R., Curtis, A., Buritica, P., Ajayakumar, J., Squires, R., Dávalos, D., et al. (2017). Community context and sub-neighborhood scale detail to explain dengue, chikungunya and zika patterns in cali, colombia. *PLoS ONE*, *12*(8), e0181208.

50. Kuncheva, L. I., & Rodríguez, J. J. (2010). Classifier ensembles for FMRI data analysis: An experiment. *Magnetic Resonance Imaging*, *28*(4), 583–593.

51. Kuppinger, P. (2004). Exclusive greenery: New gated communities in cairo. *City & Society*, *16*(2), 35–61.

52. Kwan, M. P., & Ding, G. (2008). Geo-narrative: Extending geographic information systems for narrative analysis in qualitative and mixed-method research. *The Professional Geographer*, *60*(4), 443–465.

53. Lai, Y., & Kontokosta, C. E. (2018). Quantifying place: Analyzing the drivers of pedestrian activity in dense urban environments. *Landscape and Urban Planning*, *180*, 166–178.

54. Leskovec, J., Rajaraman, A., & Ullman, J. D. (2019). *Mining of massive data sets*. Cambridge: Cambridge University Press.

55. Lewis, P., Fotheringham, S., & Winstanley, A. (2011). Spatial video and GIS. *International Journal of Geographical Information Science*, *25*(5), 697–716.

56. Li, X., Zhang, C., & Li, W. (2015). Does the visibility of greenery increase perceived safety in urban areas? Evidence from the place pulse 1.0 dataset. *ISPRS International Journal of Geo-Information*, *4*(3), 1166–1183.

57. Li, X., Zhang, C., Li, W., Ricard, R., Meng, Q., & Zhang, W. (2015). Assessing street-level urban greenery using google street view and a modified green view index. *Urban Forestry & Urban Greening*, *14*(3), 675–685.

58. Lin, G., Shen, C., Van Den Hengel, A., & Reid, I. (2016) Efficient piecewise training of deep structured models for semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 3194–3203

59. Liu, Z., Li, X., Luo, P., Loy, C.C., & Tang, X. (2015) Semantic image segmentation via deep parsing network. In: Proceedings of the IEEE international conference on computer vision, pp. 1377–1385

60. Long, Y., & Liu, L. (2017). How green are the streets? an analysis for central areas of chinese cities using tencent street view. *PLoS one*, *12*(2), e0171110.

61. Lorenc, T., Clayton, S., Neary, D., Whitehead, M., Petticrew, M., Thomson, H., et al. (2012). Crime, fear of crime, environment, and mental health and wellbeing: Mapping review of theories and causal pathways. *Health & Place*, *18*(4), 757–765.

62. Loukaitou-Sideris, A. (1999). Hot spots of bus stop crime: The importance of environmental attributes. *Journal of the American Planning association*, *65*(4), 395–411.

63. Ma, J., & Yuan, Y. (2019). Dimension reduction of image deep feature using PCA. *Journal of Visual Communication and Image Representation*, *63*, 102578.

64. Van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-sne. *Journal of machine learning research, 9*, 11.

65. Marquet, O., Ogletree, S. S., Hipp, J. A., Suau, L. J., Horvath, C. B., Sinykin, A., & Floyd, M. F. (2020). Peer reviewed: Effects of crime type and location on park use behavior. *Preventing chronic disease, 17*, 54.

66. Miaux, S., Drouin, L., Morency, P., Paquin, S., Gauvin, L., & Jacquemin, C. (2010). Making the narrative walk-in-real-time methodology relevant for public health intervention: Towards an integrative approach. *Health & Place*, *16*(6), 1166–1173.

67. Nasar, J. L., & Fisher, B. (1993). 'hot spots' of fear and crime: A multi-method investigation. *Journal of Environmental Psychology, 13*(3), 187–206.

68. Nolan, J. J., III., McDevitt, J., Cronin, S., & Farrell, A. (2004). Learning to see hate crimes: A framework for understanding and clarifying ambiguities in bias crime classification. *Criminal Justice Studies, 17*(1), 91–105.

69. Odgers, C. L., Caspi, A., Bates, C. J., Sampson, R. J., & Moffitt, T. E. (2012). Systematic social observation of children's neighborhoods using google street view: A reliable and cost-effective method. *Journal of Child Psychology and Psychiatry*, *53*(10), 1009–1017.
70. Ogneva-Himmelberger, Y., Ross, L., Caywood, T., Khananayev, M., & Starr, C. (2019). Analyzing the relationship between perception of safety and reported crime in an urban neighborhood using GIS and sketch maps. *ISPRS International Journal of Geo-Information*, *8*(12), 531.
71. Ohio Department of Health: Project dawn (deaths avoided with naloxone) (2020). Retrived 30 Sept 2020 from https://odh.ohio.gov/wps/portal/gov/odh/know-our-programs/violence-injury- prevention -program/projectdawn.
72. Oliver, M. N. (2010). Mapping primary care: Putting our patients in context. *The Journal of the American Board of Family Medicine, 23*(1), 1–3. https://doi.org/10.3122/jabfm.2010.01.090249
73. Patterson, E. B. (1991). Poverty, income inequality, and community crime rates. *Criminology*, *29*(4), 755–776.
74. Porter, L. C., Curtis, A., Jefferis, E., & Mitchell, S. (2020). Where's the crime? Exploring divergences between call data and perceptions of local crime. *The British Journal of Criminology*, *60*(2), 444–467.
75. Porter, L. C., De Biasi, A., Mitchell, S., Curtis, A., & Jefferis, E. (2019). Understanding the criminogenic properties of vacant housing: A mixed methods approach. *Journal of Research in Crime and Delinquency*, *56*(3), 378–411.
76. Rooney, T. (2015). Higher stakes-the hidden risks of school security fences for children's learning environments. *Environmental Education Research*, *21*(6), 885–898.
77. Ross, C. E. (1993). Fear of victimization and health. *Journal of Quantitative Criminology*, *9*(2), 159–175.
78. Ross, C. E. (2000). Walking, exercising, and smoking: Does neighborhood matter? *Social Science & Medicine*, *51*(2), 265–274.
79. Rundle, A. G., Bader, M. D., Richards, C. A., Neckerman, K. M., & Teitler, J. O. (2011). Using google street view to audit neighborhood environments. *American Journal of Preventive Medicine*, *40*(1), 94–100.
80. Sampson, R. J., & Raudenbush, S. W. (2004). Seeing disorder: Neighborhood stigma and the social construction of "broken windows". *Social Psychology Quarterly, 67*(4), 319–342.
81. Seyfioğlu, M. S., Özbayoğlu, A. M., & Gürbüz, S. Z. (2018). Deep convolutional autoencoder for radar-based classification of similar aided and unaided human activities. *IEEE Transactions on Aerospace and Electronic Systems*, *54*(4), 1709–1723.
82. Shen, Q., Zeng, W., Ye, Y., Arisona, S. M., Schubiger, S., Burkhard, R., et al. (2017). Streetvizor: Visual exploration of human-scale urban forms based on street views. *IEEE Transactions on Visualization and Computer Graphics*, *24*(1), 1004–1013.
83. Sundquist, K., Theobald, H., Yang, M., Li, X., Johansson, S. E., & Sundquist, J. (2006). Neighborhood violent crime and unemployment increase the risk of coronary heart disease: A multilevel study in an urban setting. *Social Science & Medicine*, *62*(8), 2061–2071.
84. Taylor, R. B. (1997). Social order and disorder of street blocks and neighborhoods: Ecology, micro-ecology, and the systemic model of social disorganization. *Journal of Research in Crime and Delinquency*, *34*(1), 113–155.
85. Troxel, W. M., Haas, A., Ghosh-Dastidar, B., Holliday, S. B., Richardson, A. S., Schwartz, H., et al. (2020). Broken windows, broken ZZS: Poor housing and neighborhood conditions are associated with objective measures of sleep health. *Journal of Urban Health*, *97*(2), 230–238.
86. Visser, M., Scholte, M., & Scheepers, P. (2013). Fear of crime and feelings of unsafety in European countries: Macro and micro explanations in cross-national perspective. *The Sociological Quarterly*, *54*(2), 278–301.
87. Wakeman, S.E., Green, T.C., & Rich, J. (2020) An overdose surge will compound the covid-19 pandemic if urgent action is not taken. Nature Medicine pp. 1–2
88. Wang, X., & Brown, D. E. (2012). The spatio-temporal modeling for criminal incidents. *Security Informatics*, *1*(1), 2.
89. Weisburd, D., Groff, E. R., & Yang, S. M. (2014). The importance of both opportunity and social disorganization theory in a future research agenda to advance criminological theory and crime prevention at places. *Journal of Research in Crime and Delinquency*, *51*(4), 499–508.
90. Whitelaw, C., Garg, N., & Argamon, S. (2005) Using appraisal groups for sentiment analysis. In: Proceedings of the 14th ACM international conference on Information and knowledge management, pp. 625–631

91. Wiehe, S. E., Kwan, M. P., Wilson, J., & Fortenberry, J. D. (2013). Adolescent health-risk behavior and community disorder. *PloS ONE*, *8*(11), e77667.
92. Wilson, J. Q., & Kelling, G. L. (1982). Broken windows. *Atlantic monthly, 249*(3), 29–38.
93. Xia, Z., Stewart, K., & Fan, J. (2021). Incorporating space and time into random forest models for analyzing geospatial patterns of drug-related crime incidents in a major us metropolitan area. *Computers, Environment and Urban Systems*, *87*, 101599.
94. Ye, Y., Zeng, W., Shen, Q., Zhang, X., & Lu, Y. (2019). The visual quality of streets: A human-centred continuous measurement based on machine learning algorithms and street view images. *Environment and Planning B: Urban Analytics and City Science*, *46*(8), 1439–1457.
95. Zhao, H., Shi, J., Qi, X., Wang, X., & Jia, J. (2017) Pyramid scene parsing network. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2881–2890
96. Zonfrillo, M. R., Melzer-Lange, M., & Gittelman, M. A. (2014). A comprehensive approach to pediatric injury prevention in the emergency department. *Pediatric emergency care*, *30*(1), 56–62.