Journal of Quantitative Criminology (2021) 37:361–391
https://doi.org/10.1007/s10940-021-09500-1

ORIGINAL PAPER

Explaining Crime Diversity with Google Street View

Samira Khorshidi[1] · Jeremy Carter[2] · George Mohler[1] · George Tita[3]

Accepted: 19 February 2021 / Published online: 15 March 2021

Abstract

Objectives Crime diversity is a measure of the variety of criminal offenses in a local environment, similar to ecological diversity. While crime diversity distributions have been explained via neutral models, to date the environmental and social mechanisms behind crime diversity have not been investigated. Building on recent work demonstrating that crime rates can be inferred from street level imagery with neural network computer vision models, in this paper we consider the task of inferring crime diversity through street level imagery.

Methods We use the Google Vision API, a deep learning image tagging service, to extract objects from sampled Google Street View (GSV) images in each census block of Los Angeles. For each census block we then compute indices for (1) object diversity, (2) diversity related to commonly employed census variables, and (3) crime diversity from reports provided by the Los Angeles Police Department. We then build ordinary least squares and geographically weighted regression models to explain crime diversity as a function of environmental diversity, population diversity, and population size.

Results We show that crime diversity arises via a combination of environmental diversity (as measured through street view object diversity), household diversity (as measured through the census), and population size. Population size and area of the census block both lend credence to the neutral model proposed by Brantingham for crime diversity. However, environmental and demographic diversity combined play an equally important role in explaining variation in crime diversity.

Conclusions Our study has two primary implications for research on crime and place. First, Google Street View (via the Google Vision API) can provide important, cost-effective empirical insights to best understand distinct geographic environments of crime. Second, environmental diversity, as measured by image tagging in GSV, was observed to be more predictive of crime diversity (variety of crime types) than commonly used census measures.

Keywords Crime diversity · Google street view · Geographically weighted regression · Computervision

✉ George Mohler
gmohler@iupui.edu

Extended author information available on the last page of the article

Introduction

Environmental and opportunity theories of crime emphasize the role of place in offender decision-making. Criminals interpret both general and specific cues that facilitate or hinder offending in a given place. From a measurement perspective, several environmental characteristics may be more pronounced and easily quantifiable, such as lighting, liquor establishments, or roadway accessibility. More challenging to measure are the subtle environmental cues. Signs of physical incivilities, such as graffiti or trash on sidewalks, opportunities such as parked vehicles, and possible guardians in the form of security systems, fences, or pedestrian traffic are each difficult to capture in examination of crime occurrence. To this end, the present study examines crime diversity as a measure of interest at place. Crime diversity (Brantingham 2016) refers to the variety of crimes that occur in a spatial area of observation. Environmental criminology (Brantingham and Brantingham 1981) posits that the crime occurrence in a given area is related to the natural and built environment. Thus, environmental diversity (variety of environmental characteristics) should give way to a range of criminal offending. Brantingham (2016) introduces crime diversity through the lens of a neutral model, where crime occurrence is assumed random and independent, leading to a law-like dependence of crime diversity as a function of the area under observation. This is to say that crime events share no relationship (for example, residential burglary and drug offending in the same place may not be related) and specific offending cues are likely randomly distributed across geographies. Thus, observations of crime diversity (and to a similar extent, crime patterns) are a derivative of environmental cues present across a geography.

Our study seeks to make two primary contributions to the crime at place literature broadly, and more specifically to the call of this special issue in identifying innovative methods and data to progress this body of literature. First, we demonstrate how Google Vision API - a deep learning API that tags features contained within Google Street View (GSV) images can be leveraged to glean place-based data that can be used to quantify environmental characteristics. Though previous studies have employed GSV data to explore crime at place (Langton and Steenbeek 2017), these examinations have relied upon qualitative coding procedures consistent with systematic social observation. Such processes are time consuming and subject to inter-rater reliability issues. As we discuss in our limitations below, GSV data via the Google Vision API are not without limitations as well, though they present an innovative source of rich place-based data that captures traits of urban environments. These data hold substantive promise for scholars examining environmental and opportunity theories of crime. Second, we build upon the emerging notion of crime diversity through an alternative approach using geographically weighted regression to predict crime diversity as a function of environmental diversity, social diversity, demographics, and population size. We leverage a unique data collection tool, the Google Vision API from Google Street View (GSV) to capture image content and quantify place-based characteristics at the street-level. In doing so we are able to explain a majority of the variation of crime diversity and estimate the effects of each of these explanatory variables. Our results indicate that population size (representative of Brantingham's neutral model), environmental diversity (as measured via Google street view images), household diversity, and median household income (which is negatively correlated) are the most important predictors of crime diversity. For diagnostic purposes of demonstrating the utility of GSV for crime at place research, we conduct separate analyses of environmental diversity as predictors of both violent crime and property crime. Results mirror those of our primary crime diversity models.

To be clear, the primary focus of the present study is to demonstrate the utility of the Google Vision API as a tool that can quantify characteristics of urban landscapes. Such metrics enable scholars to measure environmental attributes that are central to several theories of offending. The present study focuses on the diversity of both crime and environments, as we believe this approach best demonstrates the broad application of Google Vision API to further scholarly examinations of crime and place. Though our statistical findings are interesting and align with previous crime diversity at place studies, the substantive contribution of this study is that a host of place-based observations can be quantified through a mechanism that is much more accessible and cost-effective than traditional approaches. Given the nascency of Google Vision API data in criminological studies, we then identify and discuss several limitations that in turn help to chart a path forward for future research.

Offending, Environment, and the Utility of Google Street View

Explanations of crime at place hinge on several theoretical perspectives; each reliant on different sources of data and scales of measurement to appropriately test underlying assumptions. The discussion of theoretical frameworks that follows not only underpins the importance of place in studies of crime, but also contextualizes why Google Street View and Google Vision may be functional tools for crime and place research. As we near the 100th anniversary of the pioneering work of the Chicago School (Park et al. 1925; Shaw and McKay 1942; Thomas and Znaniecki 1918; Thrasher 1927), much of what we know about crime at places materialized from social disorganization that links crime with traits of neighborhoods and their residents (Sampson and Groves 1989). Emerging from this ecological perspective has been the growing recognition that places are quite nuanced in their capacity for crime. Indeed, much exploration about the causes and correlates of crime have shifted from the macro- and meso-levels to micro-levels of place (Pratt and Cullen 2005). Opportunity theories emphasize the spatiotemporal dimensions of crime; more specifically, how offenders interpret criminal opportunities and identify crime targets in the context of the social and physical environment (Eck and Weisburd 2015). Whether examining disorder or serious crimes, routine actives theory, environmental criminology and other placed-based theories have shifted the focus from ”neighborhoods” to street segments or particular addresses.

Rational choice theory contends that offenders engage in a rudimentary decision-making process that weighs perceived costs and benefits of crime commission (Cornish and Clarke 1986). Through a rather semi-conscious process, offenders evaluate criminal opportunities in the context of their environment. Offenders weigh desired rewards, likelihood of success, and likelihood of apprehension in this rational process. Thus, environmental cues in a given geography may trigger a rational decision to commit crime whereas others may have the opposite effect and result in crime deterrence. Situational crime prevention is built upon this rational foundation and leverages the reality that offenders will interpret their environment and criminal opportunity (Clarke 1997). For example, street lighting and presence of closed-circuit television signal to offenders an increased difficulty of committing an offense and likelihood of apprehension (Welsh and Farrington 2004). Conversely, the presence of graffiti and availability of surface objects to vandalize in the same location that also lacks effective mechanisms to control offending send signals of opportunity to offenders motivated by this crime type (Smith 2003).

Routine activity theory complements the rational choice perspective of offending. Under the routine activity framework, offenders become aware of opportunities to offend throughout their daily routines because of the spatio-temporal convergence of common situational elements necessary for a criminal offense (Cohen and Felson 1979). The presence of both a motivated offender and a suitable target, in the absence of a capable guardian, are likely to result in a criminal offense. Specific targets must also be available for certain crime types. Vehicles must be present for motor vehicle theft or residential homes for residential burglary. As these crime ingredients must converge in space and time for crime to occur, this presents opportunities for police and scholars to isolate specific dimensions of crime to identify primary catalysts of offending. A problem-oriented policing strategy known as the problem analysis crime triangle (Clarke and Eck 2014) focuses on these elements of crime and integrates an additional layer referred to as controllers who can exert crime control influence; which include managers of places, guardians of targets, and handlers of offenders. This crime analysis method enables criminal events to be examined in context of these elements to inform effective interventions. Given the empirical reality that crime concentrates in micro-places (Weisburd 2015), and that places offer more tangible crime prevention benefits as compared to individual offenders (Eck and Weisburd 2015), a focus on the place ingredient of crime has become increasingly fruitful. For example, a single street segment may exhibit high volumes of multiple crime types. Under a routine activity and problem analysis approach, multiple targets and offenders may be identified however the place dimension remains stable across offenders and criminal events. Further analyses may reveal that this street segment has a troublesome apartment complex, a bar with repeat calls for service, and poor lighting conditions. An effective path forward would then be to work with place managers as the controllers of this criminogenic area to generate interventions.

Crime pattern theory, or environmental criminology, leverages propositions from routine activity and rational choice theories to explain the spatiotemporal distribution of crime across various places. Crime pattern theorists emphasize the characteristics of places to explain crime occurrence. Here, the primary catalyst of crime can be attributed to the built environment. Risky facilities, roadway access, and geographic proximity of suitable targets relative to an offender’s residence (to name a few) are present in certain areas that will generate more offending opportunities than places absent these characteristics (Brantingham and Brantingham 2013). Offenders become aware of criminal opportunities through interpersonal communications (Lantz and Ruback 2017) or observations during their daily routines (Felson 2006). Characteristics of the environment heavily influence the decision-making process to offend. For example, and offender may learn of a neighborhood where homes are frequently unoccupied during certain times of day, is easily accessed by major street thoroughfares, and lack home security systems. It is also possible that an offender will personally observe these same characteristics while walking home from work, and make note of this opportunity to offend at a later time.

Routine activities and environmental characteristics coalesce to form what the Brantinghams (Brantingham and Brantingham 1981, 1995) coined as the environmental backcloth of places. Such a backcloth is comprised of features that attract and/or generate crime opportunities and rely upon effective place managers (Eck 1995). Recent advancements in crime forecasting have demonstrated the importance of the environmental backcloth, or risk factors, that identify highly criminogenic places (Kennedy et al. 2011). Onset of criminal offending at places may be described differently from a crime pattern or routine activity perspective. Routine activity theorists contend crime is a function of people in a given place. Controllers (e.g. place managers, guardians, and handlers) are believed to be effective at preventing crime in areas with lower crime rates, though less effective in high crime places. For crime pattern theorists, crime occurrence is a function of how criminal opportunities are identified and accessed by offenders (Braga and Weisburd 2010) placing a dominant role on environmental factors of a given place to explain offending. In a study of Chicago hot spots, drug dealers and robbers appeared to focus their offending in areas with specific business types, such as check-cashing and liquor stores. Subsequent interviews with drug dealers and robbers as part of this same study revealed offenders strategically selected places with these specific businesses (Jean 2008). Thus from a crime pattern perspective, offenders were drawn to these places in Chicago due to specific environmental cues. Several studies have demonstrated the importance of environment-type in explanations of crime (Kinney et al. 2008) such as the proportion of residential, commercial, and industrial buildings within geographic areas highly correlates with property crime (Boessen and Hipp 2015) and violent crime (Stucky and Ottensmann 2009) while physical urban layouts, such as streets and building placement, significantly influence overall patterns of crime (Brantingham and Brantingham 1993).

It is worth repeating that these characteristics of places reflect environmental cues interpreted by offenders through a rational decision-making process (Brantingham and Brantingham 1981; Eck and Weisburd 2015). Akin to the study of specialized offending at the individual-level (Osgood and Schreck 2007; Piquero et al. 1999), criminologists have sought to better understand specific mechanisms that facilitate this environment-crime relationship. This becomes conceptually and methodologically challenging as environmental cues may be general and/or specific to crime (Weisburd et al. 1992). General cues are requisite for crime occurrence (Brantingham and Brantingham 1993) while specific cues reflect facilitating characteristics for particular types of crime (Clarke and Cornish 1985). Though empirical assessments that point to tangible answers to the general versus specific correlates of crime are sparse, research suggests both mechanisms are salient to explaining crime at place (Weisburd et al. 1992; Quick et al. 2018).

Salient to the current discussion is the recognition that offending environments have multiple layers. From a built environment perspective, crime pattern theory (Brantingham and Brantingham 1993) suggests that street layout and building types, for example, define an environment and can help explain crime occurrence. A more particular, and difficult to quantify, aspect of environment are the more-subtle attributes of physical places. Things such as graffiti on buildings, trash or debris, pedestrian traffic, and presence of vehicles can be signals of disorder or opportunity. Such attributes can distinguish environmental differences among otherwise similar places as determined by traditional measures of environment. For example, two areas may exhibit similar building composition for land use and street layout, however one may have no physical signs of disorder while the other is rife with graffiti, litter, and abandoned vehicles. It is reasonable to assume the environment challenged with physical disorder would exhibit different crime patterns. This notion is consistent with broken windows theory - that physical incivilities signal to offenders an open environment for crime and disorder (Wilson and Kelling 1982). Controlled experiments have shown that graffiti on walls increases the propensity of individuals to litter on the same street (Keizer et al. 2008), and that geographic hot spots of crime exhibit similarly high concentrations of disorder (Weisburd 2015).

It is also the case that specific opportunities or targets must be present for specific crime types, yet similar targets and opportunities are not evenly victimized. Why some houses are burgled and others are not in the same neighborhood is difficult to explain (Palmer et al. 2002; Taylor and Gottfredson 1986; Wright and Decker 1996; Wright et al. 1995), as is higher burglary risk levels for homes located along the same side of the street in close geographic proximity to a previously victimized homes (Bowers and Johnson 2005; Vandeviver and Bernasco 2019). To date, much of the spatiotemporal modeling of crime at place has leaned upon community and structural indicators as valid empirical estimates that capture more granular characteristics of environments are more difficult to operationalize and quantify (Taylor 2015). Valid, reliable, and cost-effective means of quantifying the environmental backcloth of urban areas in which crime occurs at place are sorely needed to progress towards more effective crime prevention programs (Johnson 2010).

In his study of crime concentration in Boston across varying spatial scales, OBrien (2018) lends further support to the function between geographic size and crime occurrence and acknowledges the lack of attention given to explaining this offending cues-crime diversity issue: He notes that “some places have more crime and disorder simply by virtue of having more places where such events can occur. This should not be an entirely surprising finding, but it highlights the fact that this consideration has so rarely been addressed in the criminology of place literature” (OBrien 2018) (p.23). Despite rapid growth of crime and place research, such a void in this literature is largely a function of methodological challenges. Indeed, a consistent theme across environmental criminology and crime at place literatures is the lingering question of how can we better measure environment? To date, scholars have largely leveraged administrative data sources (such as census, 311 calls, and commercial databases) across various spatial scales to test the relationship between community factors and crime. While metrics such as land use and business types have been employed to examine the built environment to quantify risk of crime (Caplan et al. 2011; Kennedy et al. 2011), they likely mask unique characteristics of offending environments only observable at the street level.

Groff and colleagues (Groff et al. 2010) studied street to street crime trends to identify variability in Seattle. Their primary finding was that crime patterns differed across several adjacent street segments, and many street segments had significantly different crime patterns than their surrounding area. They provide a robust discussion of the community and environmental theories at play in micro places in discussing their results, largely acknowledging that the current state of science in crime and place research is unable to pinpoint direct environmental correlates of crime at localized scales. Groff et al. (2010) (p.26) urge scholars to pursue innovative ways to measure local environments: “Only by conducting research that uses street blocks as the unit of analysis and by thoroughly describing their characteristics will we be able to shed additional light on these questions. Future work should utilize prospective data collection to take advantage of the more robust information systems available today”. In a step to meet this call, the present study draws upon Google Vision API, a deep learning image tagging service, to extract objects from randomly sampled Google Street View images in each census block of Los Angeles.

Crime Diversity and Environment

Image tagging via Google Street View holds promise to identify a broad range of environmental characteristics, which in turn may signal offending cues related to a broad range of offenses in a given geography. Thus, our decision to focus on crime diversity is motivated by the desire to capture both general and specific cues of offending as they relate to several offense types. Crime diversity reflects the variation in observed crime types in a given geography. In the context of the aforementioned theories of crime, community schools of thought and general offending cues from environmental perspectives would suggest high crime diversity be observed in places. As social disorganization, concentrated disadvantage, and general offending cues are present, so too would all types of crime. Conversely, if specific offending cues are present, and/or the built environment lends itself to certain offenses, then there should be less crime diversity as certain crime types contingent upon the environment will dominate the area. For example, residential burglary cannot occur in a commercial area or park in the absence of residential housing and vehicles must be present for car thefts to occur.

Crime diversity is salient to crime at place lines of inquiry, and subsequent police interventions. Areas rich in crime diversity suggest the presence of variable crime generators and attractors. Such places are also likely to exhibit ineffective guardians, place managers, and handlers (Braga and Clarke 2014). Places plagued by an array of crime types pose difficult challenges to police. Place-based policing strategies tailored to combat multiple crime types are less effective than tailored approaches (Clarke 1997), investigatory (Malm et al. 2005) and proactive (Haberman 2017) resources are likely to be strained by multiple crimes, and community dynamics to institute social controls are likely to be less effective (Boessen and Hipp 2015). As noted by Brantingham (2016), both the volume of crime and array of crime types are equally important to communities.

Brantingham (2016), drawing from theories of ecology, conceptualized crime diversity as crime richness and crime evenness. The former refers to the simple occurrence of different crime types within a sample, while the latter speaks to the distribution frequency of observed crime types. Using crime classifications from Los Angeles, CA, he examines crime diversity as a function of geographic scale. His study adopted a neutral model in that crime types were assumed to be independently distributed across space, when conditioned on the number of crimes in an area. Under the neutral model, the distribution of crime diversity as a function of the area of observation and/or local crime rate is explained by a multinomial distribution of crime types conditionally independent of a locally dependent size parameter. In short, he found that larger geographic areas demonstrated higher crime richness (crime diversity) as predicted by the neutral model.

Scholars are most likely to take aim at Brantingham's (Brantingham 2016) central finding, that crime is a function of random processes as opposed to a linkage with certain offending cues present in a geography. Such a proposition challenges widely held beliefs among environmental criminologists and the dependencies between offenders and their environment. However, the neutral model simply predicts what the distribution and variation of crime diversity should be as a function of geographic area; the neutral model does not attempt to explain the observed variation. Lentz (2018) replicated Brantingham's (Brantingham 2016) study using publicly available data from Los Angeles in addition to data from St. Louis. His findings affirmed Brantingham's original observations, that crime diversity was indeed a function of geographic size. However, Lentz (2018) offers two interrelated observations that are central to the present study. First, he argued a more appropriate geographic sampling method would be to randomly select geographies from across a city (as opposed to random sampling by crime incident as was done by Brantingham). Lentz (2018) observed that levels of crime diversity varied more dramatically using a random sample of places, suggesting areas of a given city are likely to experience much different crime diversity dependent upon unique characteristics of the area sampled. Despite this observed variation of crime diversity, the primary finding held true that crime diversity exhibits a non-linear relationship with geographic size. Second, and related to the previous observation, Lentz (2018) suggests the randomness of neutral models may mask the reality that crime events are not random, but specific environmental offending cues are random.

Taken together, these observations suggest the importance of capturing more subtle attributes of geographies—the aforementioned indicators of guardians, physical incivilities, and opportunity that may, or may not, be present across geographies. Both studies suggest future research should attempt to identify methodological approaches to better understand the complex measurement issue of crime diversity and general versus specific cues of offending. These issues speak to the present study as both authors explain the need for a more contextual assessment of city environments that help to explain crime diversity.

Google Street View as a Valid Source of Place-Based Data

The use of innovative technologies as sources of data has slowly emerged in criminological research. Recently, scholars have conducted systematic social observation-style analyses leveraging video footage from CCTV cameras (Sytsma and Piza 2018) and police body cameras (Willits and Makin 2018). Despite rather well-documented uses in the areas of geographic and other social sciences, criminology has been slow to utilize Google Street View (GSV) as a valid source of data. Vandeviver (2014) provides a review of GSV applications to criminological research, noting most of which focus on GSV as a supplemental tool for understanding offending choices and environmental composition as opposed to a primary source of data collection. This is somewhat surprising given the exponential growth of interest in place-based crime research and reliance upon innovative mapping techniques (Kindynis 2014). Use of GSV has been proven to demonstrate measurement reliability (Clarke et al. 2010; Odgers et al. 2012; Rundle et al. 2011) and be a cost-effective and less time-intensive (Badland et al. 2010; Ben-Joseph et al. 2013; Kelly et al. 2012; Taylor et al. 2011) method to visually audit geographic places of interest.

A few notable contributions are pertinent to the present study. To estimate the risk of residential burglary in The Hague, Langton and Steenbeek (2017) merged burglary addresses from police with image-coded data of these same addresses using Google Street View to identify environmental traits of homes and their immediate surroundings. Their findings revealed that burglarized homes were more accessible due to a lack of gates, fences, and proximity to open or public areas. Homes that were more secluded from public view provided less surveillance and increased burglary likelihood. Such environmental factors that are indeed captured using the GSV Vision API data tool. Odgers and colleagues (Odgers et al. 2012) used a virtual systematic social observation (SSO) of 2,232 youth facing environmental risk across 1,012 neighborhoods in England and Wales. Their study coded physical disorder, physical decay, danger, green space, and child-safe streets. Observed SSO metrics were cross-referenced with census and study participant self-report data. In summary, their study concluded that "Findings from this study support the use of Google Street View as a reliable and cost effective tool for gathering information about local neighborhoods. Acceptable levels of inter-rater agreement were documented for the majority of the virtual-SSO items and scales, providing evidence that both negative and positive neighborhood features can be reliably coded within a virtual context" (Odgers et al. 2012) (p.1015). Similarly, He and colleagues (He et al. 2017) leveraged GSV to code physical disorder, pride in land ownership, and defensible space within each census block group of Columbus, OH to predict violent crime. Their findings are consistent with previous research using available census data at the block-group level, while noting that GSV serves as a unique source of quantitative data to capture the nuances of built environments and neighborhood characteristics that best inform crime correlates at the street-level.

Perhaps Robert Sampson (Sampson 2013) best captured this approach in his 2012 presidential address to the American Society of Criminology. When discussing how important visual assessments of the social environment are to understanding a variety of crime correlates, he draws specific attention to GSV as a legitimate source of environmental data: "Google Street View can be and is being used systematically to code a variety of urban street scenes, right down to the level of specific places" (Sampson 2013) (p.9).

Methodology

Our overall methodology follows the work by Graif and Sampson (2009), where homicide rate variation is explained through neighborhood diversity. Graif and Sampson model homicide rates as a function of language diversity (as measured via a diversity index) and other spatial covariates. In that work a geographically weighted regression (Anselin 2009) is compared to ordinary least squares regression, with the former yielding improved goodness of fit.

Here our goal is to explain the variation in crime diversity across geographic areas as a function of environmental diversity measured through Google street view, along with other spatial demographic covariates and derived diversity measures. For this purpose we use the Shannon Diversity Index (Shannon 1948), which is used in ecology to measure the uncertainty in predicting the species of an individual randomly drawn from the population in a given spatial area.

Let $i = 1, ..., R$ index each of R crime types and let $p_{i,k}$ denote the proportion of crime in spatial unit k given by crime type i. The Shannon Diversity Index H_k in spatial area k is given by,

Property	88%
Residential Area	86%
Tree	78%
Neighbourhood	74%
Architecture	73%
Home	72%
House	68%
Building	66%
Road	65%
Real Estate	65%
Street	65%
City	60%
Landscape	59%
Thoroughfare	58%
Asphalt	57%
Urban Design	56%
Intersection	55%
Road Trip	52%

Fig. 1 Example street view image and objects detected by the Google Vision API (along with object confidence score)

$$H_k = -\sum_{i=1}^{R} p_{i,k} \ln p_{i,k}. \tag{1}$$

Here H_k is maximized (greatest diversity) if all crime types are equally represented, $p_{i,k} = 1/R$, and is minimized (least diversity) if only one crime type is observed, $p_{i,k} = 1$ for some i. Each term in Equation 1 is defined as zero if $p_{i,k} = 0$. The spatial unit of analysis we consider in this paper is the census block level.

To measure environmental diversity, we again use Eq. 1 to estimate the diversity of objects in street view images. We first use the Google street view API to obtain the street view image at each node in the street network of Los Angeles containing at least one crime. We next use the Google Vision API to generate a list of objects present in the image. The Google Vision API utilizes a deep neural network for image object detection and returns a list of objects along with confidence scores (see Fig. 1). In one study evaluating the Google Vision API, the service was shown to be around 75% accurate [68] at identifying objects in an image, whereas in another study using the service for street sign detection, the GSV API was shown to be above 95% accurate (Campbell et al. 2019).

With this list of objects for each image in a census block, we then compute the Shannon diversity for image objects in the census block aggregated across all images in the census block. In Fig. 2, we provide examples of images in a high and low diversity census block along with the Shannon diversity index for crime and image object diversity. We note that synonyms can potentially arise in the GSV image object labels, for example home and dwelling, and that one could consider collapsing such synonyms into more general categories. However, we follow the work of Lentz (2018), Brantingham (2016) and leave the label categories disaggregated to the finest level. In the appendix we explore the sensitivity of the diversity index to high and low levels of categorization and find that the index is highly correlated between the two.

Fig. 2 Top row: high diversity census block with object diversity index of 0.79 and crime diversity index of 2.58. The area contains a variety of housing, commercial buildings, and vehicle types and subsequently contains a variety of property and violent crime. Bottom row: low diversity industrial census block with object diversity index of 0.62 and crime diversity index of 0.69. The area consists of mostly commercial buildings and the only crimes reported in the time period of observation were fraud and vandalism

We estimate both an ordinary least squares regression and, as is done in Graif and Sampson (2009), we also estimate a geographically weighted regression model (using the R package "spgwr" (Bivand et al. 2017)). We fit a third "spatial simultaneous autoregressive errors" model (errorsarlm R package), but it did not outperform the GWR in terms of AIC so we did not include it in the paper. The purpose of considering multiple model specifications is to check for the robustness of our results against different modeling assumptions. Because our goal is explaining crime diversity, rather than predictive accuracy, we include all relevant predictors to establish error bars and significance of the coefficients.

In the GWR model, crime diversity, y_k, in census block k is given by,

$$y_k = a_{k,0} + \sum_{j=1}^{M} a_{k,j} x_{k,j} + \epsilon_k \tag{2}$$

where $x_{k,j}$ is the value of the jth explanatory variable in census block k and $a_{k,j}$ is a spatially varying coefficient. The coefficient vector at location k is estimated via the equation,

$$\vec{a}_k = (\vec{x}^t W_k \vec{x})^{-1} \vec{x}^t W_k \vec{y}. \tag{3}$$

Here W_k is a diagonal matrix containing the weights of other census blocks relative to block k using a Gaussian kernel with bandwidth selected via cross-validation. When W_k is given by the identity matrix then gwr reduces to ordinary least squares regression, however as the bandwidth of the Gaussian kernel reduces in size then the coefficients become more localized in space.

Data

We use three types of data in this study: 1) crime incident data for Los Angeles in 2013, 2) Google street view object counts in census blocks and 3) census data defined at the block group level. In Table 1 we display the crime categories present in the Los Angeles crime data set (those containing at least 50 incidents) and in Table 2 we state the mean and standard deviation of the variables across census blocks. The list of crime categories includes both property and violent crime, ranging in severity from misdemeanors to felonies. When including census variables we followed previous work such as Graif and Sampson (2009) and the census variable list includes commonly used spatial predictors of crime such as poverty, income, unemployment, etc.

As discussed in the previous section, we obtained Google street view images using the Google street view API at locations on the street network containing at least one crime incident in the Los Angeles data set. This resulted in 123,912 unique images. We then used the Google Vision API to extract a list of objects detected in each image. Finally we aggregated these objects at the census block group level in order to compute a Google Street View (GSV) object diversity index (see Table 3 for a list of detect objects and their frequencies).

The census variables used can be found in Tables 5, 6, 7, 8, 9, and 10 in the results section. We analyze a wide range of theoretically relevant census variables commonly used to explain crime variation, though we limit the number of predictors to less than 1% of the total number of observations to prevent over-fitting (Harrell 2015). These measures are largely drawn from the social disorganization and concentrated disadvantage perspectives (Sampson and Groves 1989). Though social disorganization as a

Table 1 Crime categories and frequencies for Los Angeles 2013 crime data

Category	Code	Frequency
Homicide	110	250
Robbery	210	6859
Robbery - attempt	220	976
Agg. assault	230	6995
ADW against police officer	231	126
Child neglect	237	509
Shots fired inhabited dwel.	251	243
Burglary	310	14231
Burglary - attempt	320	1206
Burg. from vehicle	330	15445
Grand theft from vehicle	331	2539
Grand theft	341	6609
Grand shoplifting	343	165
Theft from person	350	1429
Pursesnatch	351	95
Identity theft	354	11,709
Burg. from vehicle - attempt	410	264
Petty theft from vehicle	420	7604
Theft from vehicle - attempt	421	81
Prop. missing (grand) recovered vehicle	430	55
Prop. missing (petty) recovered vehicle	431	80
Driving w/o owner consent	433	87
False imprisonment	434	81
Resisting arrests	437	338
Petty theft	440	15,913
Theft - attempt	441	198
Petty shoplifting	442	3637
Bicycle stolen	480	884
Stolen vehicle	510	14127
Stolen vehicle - attempt	520	320
Vehicle recovered	521	12,294
Battery on police officer	623	322
Battery (Misdemeanor)	624	18,882
Assault (Misdemeanor)	625	94
Throwing object moving vehicle	647	109
Arson	648	317
Forged document	649	2609
Credit card fraud (grand)	653	121
Credit card fraud (petty)	660	99
Computer crime	661	88
BUNCO (grand theft)	662	401
BUNCO (petty theft)	664	326
Embezzlement (Grand)	668	726
Vandalism (Grand)	740	8907
Vandalism (Petty)	745	9091
Shots fired	753	324

Table 1 (continued)

Category	Code	Frequency
Bomb threat	755	109
Brandish weapon	761	710
Stalking	763	180
Violation court order	900	1881
Violation restraining order	901	1747
Violation temp. restrain. order	902	90
Kidnapping	910	241
Kidnapping - attempt	920	80
Child stealing (custody dispute)	922	95
Threatening phone call	928	283
Criminal threats	930	5589
Prowler	933	140
Extortion	940	100
Cruelty to animals	943	110
Other crime	946	1929
Defrauding innkeeper	951	227
Obscene phone call	956	1866
Death report	979	130

salient community explanation of crime is generally agreed upon (Lowenkamp et al. 2003), specific measures used to operationalize this construct are less than prescriptive (Kubrin and Weitzer 2003; Veysey and Messner 1999). Which specific measures pertaining to race, ethnicity, poverty, housing population and type, education, age, sex, and unemployment should be indexed to reflect social disorganization varies across studies and available data samples. To best accommodate the range of census variables most commonly used to capture social disorganization, we err on the side of overly specified models. We believe this approach is best suited to demonstrate the explanatory power of GSV as compared to commonly used census measures, as well as several census-based demographic diversity indices (similar to the work in Graif and Sampson (2009)). For all diversity indices we use the Shannon Index found in the above methodology section. We include language heterogeneity or "diversity" as an extension of social disorganization and collective efficacy as differential opportunity theory (Cloward and Ohlin 2013) posits social ties among residents in areas of concentrated disadvantage affects crime occurrence and crime type as well as community mechanisms of social control (Bursik and Grasmick 1999) - an interaction observed in Chicago neighborhoods among social networks and crime in communities (Schreck et al. 2009). This approach is also consistent with previous research that demonstrates the importance of cultural disorganization within socially disorganized communities (Warner 2003).

In the following analysis, we restrict our attention to census block groups with at least 50 crime incidents in the observation period. This is to remove blocks with low volume of crimes where the diversity index will have high variance (and also the index is biased towards lower diversity). We used the bootstrap to simulate blocks with different crime volumes (and same diversity index) to empirically set the threshold at 50. In

Table 2 Mean and standard deviation of dependent and target variables for regression (all crime index target variable)

Name	Mean	SD
GSV image index	2.63	0.39
Household heterogeneity	2.02	0.23
Language heterogeneity	1.2	0.31
Race heterogeneity	1.16	0.26
Per capita income	31,998.14	24,864.97
Population	296.85	329.67
Census block land size	557,623.21	964,587.99
Population male	833.62	415.74
Population female	850.88	443.12
Median age	37.25	7.48
Hispanic or latino ethnicity	805.05	692.46
Poverty undetermined	1640.79	783.09
Median household Income	62,938.95	34,905.62
Labor force unemployed	83.72	70.85
Vacant housing units	35.28	40.76
Powerty status uder 50	141.88	167.39
Poverty status .50–.99	190.26	203.63
Poverty status 1.00–1.24	97.27	105.49
Poverty status 1.25–1.49	99.79	108.29
Poverty status 1.50–1.84	119.08	120.84
Poverty status 1.84–1.99	47.92	62.93
No schooling completed educ	35.36	38.78
Regular high school dipl educ	208.14	132.91
GED or alternative dipl educ	19.1	21.76
Less than 1 year college educ	45.19	38.27
Some college more than 1 year educ	165.15	103.64
Associates degree	69.82	54.82
Bachelors degree	244.68	210.62
Masters degre	77.46	84.92
Professional degree	30.54	53.73
White alone	922.7	581.21
African American alone	125.32	180.7
American Indian and Alaska Native	9.62	23.07
Asian alone	189.88	251.41
Native Hawaiian Pacific Islander	2.72	10.67
Other race alone	373.57	404.03
All crime index	1.58	0.75

Fig. 3 we plot the Shannon diversity index vs. number of crimes for two census block groups for the bootstrap simulations.

Table 3 Objects with frequency and frequencies for Los Angeles

Object	Frequency	Object	Frequency
Road	14,388	Car	12,551
Property	9855	Residential area	9753
Lane	9158	Mode of transport	8681
Asphalt	7999	Tree	7562
Sky	7168	Neighbourhood	6401
Transport	6254	Vehicle	5703
Town	5157	Motor vehicle	5143
Infrastructure	4396	Family car	4284
Plant	4228	Luxury vehicle	3555
Real estate	3100	Road trip	2398
Area	2397	Home	2337
Land lot	2019	House	1944
Street	1626	Suburb	1536
Land vehicle	1369	Road surface	1120
Woody plant	1118	Metropolitan area	887
landscape	862	Building	827
Highway	785	Public space	645
Estate	622	Automotive exterior	589
Path	579	Vegetation	439
Arecales	391	Wall	391
Grass	387	Facade	375
Walkway	364	Fence	313
Urban area	303	Palm tree	299
Traffic light	292	City	291
Roof	282	Signaling device	272
Yard	266	Shrub	254
Condominium	231	Thoroughfare	230
Flora	216	Cottage	211
Street light	193	Outdoor structure	184
Biome	177	Sedan	173
Mid size car	173	Parking lot	172
Plant community	171	Ecosystem	161
Parking	153	Automotive design	150
Structure	145	Architecture	137
Intersection	136	Window	132
Advertising	132	Cloud	123
Commercial vehicle	119	Villa	118
Recreation	117	Geological phenomeno	100
Village	100	Downtown	100

Results

Our goal is to explain variation in crime diversity, as measured via the Shannon Index, as a function of street view diversity and demographic features of the environment (and derived diversity measures). For this purpose we use ordinary least squares regression and geographically weighted regression and, in Table 4, we present the R^2 and AIC values of each model. For each type of regression we model all crime diversity (all crimes aggregated), property crime diversity and violent crime diversity.

In Table 5, we present the OLS coefficients for all crime diversity along with their standard errors and p-values. The explanatory variable with the largest estimated coefficient is population; Google Street View diversity has the third largest coefficient in magnitude (.172, 95% CI: [.149, .195]) behind population and median income. All variables have been standardized to mean zero and variance 1, so population having a coefficient of .46 means that the crime diversity index is predicted to increase by that amount when the population size increases by one standard deviation. Other statistically significant predictors include median household income, household heterogeneity (diversity), and area of the census block. Population size and area of the census block both lend credence to the neutral model proposed by Brantingham for crime diversity (Brantingham 2016). However, environmental and demographic diversity combined play an equally important role in explaining variation in crime diversity. Median household income is also an important, negatively correlated predictor, indicating that lower income levels in a geographic area lead to higher crime diversity. However, such a relationship could arise simply from lower income levels being correlated with higher crime rates, which the neutral model predicts to be associated with higher crime diversity.

The geographically weighted regression has a similar set of important explanatory variables in the case of all crime diversity compared to the OLS model. Population again has the largest global coefficient, with income, GSV diversity, household diversity, median age all with large, statistically significant global coefficients. In Tables 5, 6, 7, 8, 9, and 10 we display the model coefficients for OLS and GWR regressions on property crime diversity and violent crime diversity. Here we see that the coefficients of GSV diversity are still second (next to population) in the OLS models, but are less significant in the case of the GWR models. Thus, when accounting for geographic variation of the model (GWR), GSV diversity is predicting variations in crime diversity arising due to variations across property and violent crime, but less so when restricting to property or violent crime.

While our study here focuses on modeling crime diversity, we note that street view images may also be used to model crime volume in regressions with spatial covariates. In Table 11 we show results for spatial (linear) regressions estimated similarly to those above, but with the response variable being crime volume instead of diversity. We note that the street view diversity index is a statistically significant variable, though its importance is less than in the case of modeling crime diversity. For researchers interested in modeling

Table 4 Comparison of linear and gwr models

	All linear	Al gwr	Prop. linear	Prop. gwr	Viol. linear	Viol. gwr
R^2	0.426	0.974	0.381	0.970	0.375	0.976
AIC	10,578.9	− 3234.4	10,727.2	− 2585.0	8983.9	− 2922.8

Table 5 All crime diversity linear model

Coef.	Est.	Std. Err.	*t* val	*p* val	Sig.
Population	0.46	0.0143	32.264	< 2e-16	***
Median Household Income	− 0.28	0.0239	− 11.731	< 2e-16	***
GSV Image index	0.172	0.0119	14.485	< 2e-16	***
Household heterogeneity	0.154	0.0188	8.196	3.19E-16	***
Census block land size	0.121	0.0149	8.144	4.90E-16	***
Median age	− 0.148	0.0183	− 8.08	8.20E-16	***
Language heterogeneity	− 0.0703	0.0166	− 4.239	2.29E-05	***
No schooling completed Educ	0.0607	0.0154	3.938	8.34E-05	***
Poverty status 1.00–1.24	− 0.0543	0.0165	− 3.299	0.000977	***
Some college more than 1 Year Educ	0.0658	0.0208	3.159	0.001596	**
Labor force unemployed	− 0.0607	0.0195	− 3.108	0.001895	**
Professional degree	0.0656	0.023	2.848	0.004423	**
Poverty status 1.25–1.49	0.0448	0.0164	2.74	0.006166	**
Poverty status 1.84–1.99	0.0335	0.0128	2.616	0.008924	**
African American alone	0.106	0.0433	2.436	0.014884	*
Vacant housing units	0.0289	0.0144	2.006	0.044864	*
Bachelors degree	0.0625	0.0334	1.873	0.061093	.
Racial heterogeneity	−0.0409	0.0218	−1.871	0.06137	.
Less than 1 year college Educ	0.0263	0.0153	1.72	0.085511	.
Associates degree	−0.0256	0.0161	−1.585	0.113051	
Population female	−0.161	0.102	−1.582	0.11375	
Poverty status .50–.99	0.026	0.0217	1.201	0.229856	
Native Hawaiian Pacific Islander	0.014	0.0119	1.182	0.237298	
Per capita income	0.0277	0.0248	1.12	0.262719	
Asian alone	0.0497	0.0593	0.839	0.40154	
Other race alone	0.0722	0.0927	0.778	0.43657	
American Indian and Alaska Native	− 0.0093	0.0132	− 0.707	0.47968	
Poverty status under .50	0.00954	0.0194	0.493	0.622227	
White alone	− 0.0628	0.132	−0.475	0.634479	
Regular high school Dipl Educ	0.0109	0.0246	0.445	0.656398	
Masters degree	− 0.0108	0.0251	− 0.431	0.666549	
Hispanic or Latino Ethnicity	− 0.0215	0.0517	− 0.416	0.677185	
Population male	0.027	0.095	0.284	0.776147	
Poverty undetermined	0.0168	0.0698	0.241	0.809786	
GED or alternative Dipl Educ	0.00191	0.0134	0.142	0.886919	
Poverty status 1.50–1.84	− 0.000831	0.0172	− 0.048	0.961529	

Table 6 All crime diversity gwr model

Coef.	Min.	1st Q.	Median	3rd Q.	Max.	Global
Population	− 1.27	0.261	0.459	0.683	1.75	0.4598
Median household income	− 3.8	− 0.466	−0.116	0.203	4.42	−0.2803
GSV image index	− 0.528	− 0.0226	0.135	0.318	1.34	0.1722
Population female	− 14.6	− 0.819	0.537	1.83	20.5	− 0.1607
Household heterogeneity	− 3.48	− 0.219	0.0407	0.308	1.97	0.1539
Median age	− 1.55	− 0.365	− 0.153	0.0629	1.48	− 0.1476
Census block land size	− 4.59	− 0.6	0.00436	0.362	6.84	0.1211
African American alone	− 5.82	− 0.606	− 0.106	0.609	5.18	0.1055
Other race alone	− 10.5	− 1.06	− 0.269	0.665	11.6	0.0721
Language Heterogeneity	− 1.28	− 0.124	0.0237	0.249	1.85	− 0.0703
Some college more than 1 year educ	− 1.16	− 0.175	0.0619	0.334	1.92	0.0658
Professional degree	− 7.6	− 0.648	− 0.265	0.124	9.11	0.0656
White alone	− 14.9	− 1.82	− 0.694	0.684	18.5	− 0.0628
Bachelors degree	− 3.21	− 0.33	0.103	0.502	4.52	0.0625
Labor force unemployed	− 3.24	− 0.169	0.00194	0.224	2.32	− 0.0607
No schooling completed educ	− 1.26	− 0.147	0.0382	0.178	2.34	0.0607
Poverty status 1.00–1.24	− 0.807	− 0.204	− 0.0211	0.137	1.74	− 0.0543
Asian alone	− 11.3	− 0.766	− 0.104	0.501	5.49	0.0497
Poverty Status 1.25-1.49	− 3.44	− 0.196	− 0.0465	0.125	1.43	0.0448
Racial heterogeneity	− 2.91	− 0.239	0.0157	0.263	3.49	− 0.0409
Poverty status 1.85–1.99	− 1.17	− 0.169	− 0.000403	0.171	1.28	0.0335
Vacant housing units	− 2	− 0.0537	0.113	0.268	1.07	0.0289
Per capita income	− 7.13	− 0.577	0.107	0.551	5.11	0.0277
Population male	− 13	− 0.582	0.892	2.41	22.2	0.027
Less than 1 year college educ	− 3.66	− 0.272	− 0.068	0.19	1.83	0.0263
Poverty status .50–.99	− 2.05	− 0.358	− 0.121	0.117	2.01	0.026
Associates degree	− 1.54	− 0.211	−0.0216	0.13	1.63	− 0.0256
Hispanic or latino ethnicity	− 16.1	− 1.04	− 0.0233	0.956	6.58	− 0.0215
Poverty undetermined	− 21	− 2.89	− 0.273	1.35	31.6	0.0168
Native Hawaiian Pacific Islander	− 3.9	− 0.0667	0.0583	0.183	2.72	0.014
Regular high school dipl educ	− 1.95	− 0.441	−0.132	0.16	1.68	0.0109
Masters degree	− 1.83	− 0.228	0.121	0.524	3.98	−0.0108
Poverty status under .50	− 3.2	− 0.377	− 0.0656	0.207	1.33	0.0095
American Indian and Alaska native	− 0.872	− 0.171	0.00325	0.102	1.32	− 0.0093
GED or alternative dipl educ	− 1.33	− 0.232	− 0.0679	0.0693	0.811	0.0019
Poverty status 1.50–1.84	− 2.85	− 0.231	− 0.042	0.176	1.36	− 8e-04

crime volume, individual object counts may be better features (such as using the density of cars, houses, etc. as individual predictors).

Table 7 Property crime diversity linear model

Coef.	Est.	Std. Err.	t val	p val	Sig.
Population	0.483	0.0149	32.317	< 2e-16	***
Household heterogeneity	0.22	0.0196	11.274	< 2e-16	***
Median household income	− 0.18	0.0251	−7.155	9.75E-13	***
Median age	− 0.16	0.0193	− 8.292	< 2e-16	***
GSV image index	0.151	0.0125	12.07	< 2e-16	***
Census block land size	0.138	0.0156	8.849	< 2e-16	***
Language heterogeneity	− 0.0914	0.0175	− 5.23	1.77E-07	***
Some college more than 1 year educ	0.0911	0.0218	4.175	3.04E-05	***
Labor force unemployed	− 0.0814	0.0207	− 3.935	8.44E-05	***
Bachelors degree educ	0.119	0.035	3.396	0.000689	***
Poverty 1.25–1.49	0.05	0.0172	2.911	0.003617	**
Poverty status 1.00–1.24	− 0.047	0.0173	− 2.714	0.006681	**
Poverty undetermined	− 0.174	0.0733	− 2.38	0.017359	*
Associates Degree Educ	− 0.036	0.0169	− 2.133	0.03301	*
Population male	0.207	0.0997	2.075	0.038008	*
Vacant housing units	0.0297	0.0152	1.959	0.050137	.
Less than 1 Year College Educ	0.0308	0.016	1.921	0.054837	.
Poverty status 1.50–1.84	0.0334	0.0181	1.847	0.064885	.
Hispanic or Latino ethnicity	− 0.0996	0.0547	− 1.822	0.068574	.
Poverty status below .50	0.0308	0.0204	1.507	0.131928	
American Indian and Alaska Native	− 0.0203	0.0139	− 1.469	0.14196	
Poverty status .50–.99	0.0277	0.0227	1.218	0.223419	
White alone	− 0.14	0.138	− 1.015	0.310371	
No schooling completed Educ	0.016	0.0162	0.986	0.324113	
Regular high school Dipl Educ	0.0254	0.0257	0.986	0.324161	
Poverty status 1.84–1.99	0.0123	0.0134	0.917	0.359075	
Asian alone	−0.0551	0.0622	− 0.885	0.376289	
African American alone	0.0299	0.0455	0.658	0.510768	
Per capita income	− 0.0144	0.026	−0.555	0.579175	
Professional degree Educ	0.0121	0.0243	0.498	0.61835	
Native Hawaiian Pacific Islander	0.00421	0.0124	0.338	0.735292	
GED or alternative Dipl Educ	− 0.00405	0.0141	− 0.287	0.773912	
Racial heterogeneity	− 0.00504	0.0231	− 0.219	0.826974	
Population female	0.0101	0.106	0.095	0.924518	
Masters degree	− 0.000641	0.0264	− 0.024	0.980595	
Other race alone	− 0.000348	0.0974	− 0.004	0.99715	

Discussion

Several implications from the present study can be gleaned for research, and how place-based crime analysis may inform police interventions. Findings presented here indicate two primary implications for research. First, Google Street View (via the Google Vision API) can provide important, cost-effective empirical insights to best understand distinct

Table 8 Property crime diversity gwr model

Coef.	Min.	1st Q.	Median	3rd Q.	Max.	Global
Population	– 0.8981	0.306	0.4781	0.6474	1.417	0.4827
Household heterogeneity	– 2.656	– 0.3255	0.04398	0.2932	1.734	0.2204
Population male	– 15.53	– 0.8446	0.3235	1.724	22.93	0.207
Median household income	– 4.715	– 0.4662	– 0.02169	0.4019	2.489	– 0.1798
Poverty undetermined	– 18.8	– 2.142	– 0.5072	1.073	30.58	– 0.1744
Median age	– 1.411	– 0.3472	– 0.1669	0.08906	2.364	– 0.1596
GSV image index	– 0.8308	– 0.07034	0.1135	0.3496	1.163	0.1508
White alone	– 25.15	– 1.616	– 0.2665	1.135	29.25	– 0.1402
Census block land size	– 5.486	– 0.4833	0.03034	0.6415	7.746	0.1383
Bachelors degree	– 3.011	– 0.2361	0.4447	0.9402	6.624	0.119
Hispanic or Latino ethnicity	– 17.01	– 0.9097	0.1713	1.275	5.129	– 0.0996
Language heterogeneity	– 1.976	– 0.1775	0.0289	0.254	1.741	– 0.0914
Some college more than 1 year educ	– 1.349	– 0.144	0.1217	0.3838	2.701	0.0911
Labor force unemployed	– 2.265	– 0.06334	0.08595	0.2844	2.292	– 0.0814
Asian alone	– 13.74	– 0.8331	– 0.1624	0.4474	10.26	– 0.0551
Poverty status 1.25–1.49	– 1.165	– .1774	0.07297	0.2876	1.631	0.05
Poverty status 1.00–1.24	– 0.5411	– 0.1469	– 0.005384	0.1621	1.928	– 0.047
Associates degree	– 2.387	– 0.2	0.0129	0.2103	1.112	– 0.036
Poverty status 1.50–1.84	– 1.038	– 0.2433	0.001527	0.2132	1.366	0.0334
Poverty status under .50	– 1.283	– 0.3027	0.03996	0.3228	2.07	0.0308
Less than 1 year college educ	– 1.171	– 0.3033	– 0.05743	0.2322	1.205	0.0308
African American Alone	– 7.23	– 0.6503	– 0.08319	0.4918	8.14	0.0299
Vacant housing units	– 2.411	– 0.07539	0.09061	0.2691	1.837	0.0297
Poverty status .50–.99	– 1.703	– 0.3745	– 0.1075	0.1199	1.768	0.0277
Regular high school dipl educ	– 2.444	– 0.3602	0.05613	0.2851	1.295	0.0254
American Indian and Alaska Native	– 1.459	– 0.276	– 0.002181	0.1569	1.181	– 0.0203
No schooling completed educ	– 1.288	– 0.2331	– 0.06057	0.1252	1.274	0.0159
Per Capita income	– 7.231	– 0.7712	0.01281	0.7158	8.072	– 0.0144
Poverty status 1.85–1.99	– 0.921	– 0.1967	– 0.04175	0.1503	1.303	0.0123
Professional degree	– 8.18	– 0.6628	– 0.1126	0.3612	13.49	0.0121
Population female	– 21.01	– 1.165	0.2229	1.512	22.73	0.0101
Racial heterogeneity	– 3.464	– 0.262	– 0.02348	0.278	1.912	– 0.005
Native Hawaiian Pacific Islander	– 2.201	– 0.08311	0.04267	0.1774	2.577	0.0042
GED or alternative dipl educ	– 1.214	– 0.2189	– 0.04321	0.1215	0.7168	– 0.004
Masters degree	– 2.51	– 0.2339	0.1232	0.542	4.04	– 6e-04
Other race alone	– 16.68	– 1.087	– 0.03781	0.8486	18.28	– 3e-04

geographic environments of crime. Methods such as systematic social observations of places as well as qualitative data-coding by scholars examining images or video require significant resources. This is also true of access to commercially held databases in the United States. Though granular point-level data on environmental characteristics is more readily available in Europe and other places, such data in the United States requires scholars to pay premium purchase prices for access. Indeed, GSV is an additional methodological tool that can be

Table 9 Violent crime diversity linear model

Coef.	Est.	Std. Err.	t val	p val	Sig.
Population	0.382	0.0176	21.777	<2e-16	***
GSV image index	0.128	0.014	9.133	<2e-16	***
Median age	− 0.181	0.0217	− 8.346	<2e-16	***
Associates degree	0.149	0.0187	7.97	2.09E-15	***
No schooling completed educ	0.141	0.018	7.863	4.87E-15	***
Median household income	− 0.168	0.0273	− 6.153	8.41E-10	***
Language heterogeneity	− 0.127	0.0209	− 6.093	1.22E-09	***
Labor force unemployed	− 0.0972	0.0215	− .52	6.38E-06	***
Regular high school dipl educ	0.125	0.0288	4.334	1.50E-05	***
Professional degree	0.108	0.027	4.01	6.20E-05	***
Poverty status 1.25–1.49	− 0.0598	0.0191	− 3.135	0.00173	**
Poverty status under .50	0.0718	0.0231	3.101	0.00194	**
Poverty status .50–.99	0.0738	0.0254	2.903	0.00372	**
Bachelors degree	− 0.11	0.0405	− 2.709	0.00677	**
Census block land size	0.0395	0.0171	2.304	0.02128	*
Population male	− 0.243	0.114	− 2.131	0.03313	*
Vacant housing units	0.0362	0.0171	2.121	0.03401	*
Native Hawaiian Pacific Islander	0.0288	0.0141	2.049	0.04056	*
Asian alone	0.12	0.0595	2.017	0.04378	*
Hispanic or Latino Ethnicity	− 0.12	0.0622	− 1.934	0.05323	.
Racial heterogeneity	0.0454	0.0247	1.84	0.06581	.
Poverty status 1.85–1.99	0.0252	0.0148	1.706	0.08812	.
African American Alone	0.0926	0.0546	1.696	0.08995	.
Population female	− 0.183	0.117	− 1.56	0.11877	
Other race alone	0.178	0.114	1.556	0.11979	
GED or alternative dipl educ	0.0227	0.0155	1.461	0.1441	
Poverty status 1.50–1.84	− 0.0294	0.0212	− 1.39	0.1645	
Less than 1 year college educ	0.0239	0.0176	1.358	0.17442	
Per capita income	− 0.0307	0.0272	− 1.13	0.25867	
Household heterogeneity	0.0264	0.0235	1.12	0.26277	
Poverty undetermined	0.1	0.0896	1.12	0.26294	
American Indian and Alaska Native	0.012	0.0151	0.793	0.42805	
Masters degree	0.0215	0.0291	0.74	0.45953	
Some college more than 1 year Educ	− 0.0171	0.0243	− 0.704	0.48174	
Poverty status 1.00–1.24	0.012	0.0193	0.622	0.53401	
White alone	0.0356	0.154	0.231	0.81706	

utilized to quantify the environmental backcloth where crime occurs. Scholars seeking to best understand crime at place could utilize GSV as a supplemental source of data. Second, environment diversity, as measured by image tagging in GSV, was observed to be more predictive of crime diversity (variety of crime types) than commonly used census measures.

The observed explanatory power of our GSV index lends support to the central role of environmental cues of offending. Supplemental analyses of specific image tags, or what

Table 10 Violent crime diversity gwr model

Coef.	Min.	1st Q.	Median	3rd Q.	Max.	Global
Population	− 0.8091	0.2571	0.4417	0.6182	1.788	0.3821
Population male	− 7.129	− 0.9223	0.5318	2.864	36.73	− 0.2427
Population female	− 9.438	− 0.3429	0.9869	2.73	32.64	− 0.1832
Median age	− 2.606	− 0.3392	− 0.09489	0.1008	1.329	− 0.1809
Other race alone	− 10.04	− 1.642	− 0.3149	0.6741	5.611	0.1775
Median household income	− 2.747	− 0.2951	0.0562	0.3646	1.698	− 0.168
Associates degree	− 1.302	− 0.1176	0.09898	0.2807	1.899	0.1492
No schooling completed educ	− 1.296	− 0.1262	0.04969	0.1968	1.009	0.1413
GSV image index	− 0.7754	− 0.1017	0.1176	0.2831	1.01	0.128
Language heterogeneity	− 0.7788	− 0.0989	0.09304	0.3035	1.581	− 0.1271
Regular high school dipl educ	− 3.115	− 0.4047	− 0.08068	0.1947	3.198	0.1248
Hispanic or Latino ethnicity	− 23.72	− 1.225	− 0.371	0.3133	15.77	− 0.1203
Asian alone	− 11.09	− 1.007	− 0.2899	0.3934	2.792	0.1201
Bachelors degree	− 4.288	− 1.178	− 0.5341	− 0.0103	4.737	− 0.1097
Professional degree	− 6.926	− 0.3441	0.05258	0.6845	7.049	0.1082
Poverty undetermined	− 27.58	− 2.338	− 0.2021	1.039	10.96	0.1004
Labor Force unemployed	− 1.215	− 0.1892	0.004301	0.2666	1.585	− 0.0972
African American alone	− 8.643	− 0.9648	− 0.3758	0.2837	5.287	0.0926
Poverty status .50–.99	− 2.97	− 0.2092	0.05316	0.3516	1.724	0.0738
Poverty status under .50	− 1.802	− 0.2409	0.02635	0.3519	2.868	0.0718
Poverty status 1.25–1.49	− 0.9304	− 0.3035	− 0.1065	0.1846	1.318	− 0.0598
Racial Heterogeneity	− 2.635	− 0.2552	− 0.02935	0.3326	1.766	0.0454
Census block land size	− 7.515	− 0.6778	0.05457	0.6541	6.558	0.0395
Vacant housing units	− 3.324	− 0.1142	0.0657	0.2608	1.537	0.0362
White Alone	− 15.37	− 2.224	− 0.5699	0.5618	7.33	0.0356
Per capita income	− 5.104	− 0.5847	0.01716	0.3028	6.77	− 0.0307
Poverty status 1.50–1.84	− 2.054	− 0.239	− 0.006698	0.2159	2.082	− 0.0294
Native Hawaiian Pacific Islander	− 2.299	− 0.1504	− 0.001367	0.1182	1.564	0.0288
Household heterogeneity	− 1.311	− 0.2777	− 0.02028	0.3642	3.38	0.0264
Poverty status 1.85–1.99	− 1.047	− 0.1854	− 0.03362	0.08083	1.906	0.0252
Less than 1 year college educ	− 2.231	− 0.4402	− 0.2426	0.07021	0.862	0.0239
GED or alternative dipl educ	− 1.306	− 0.1819	0.05808	0.184	2.159	0.0227
Masters degree	− 3.225	− 0.5078	− 0.08279	0.2885	5.156	0.0215
Some college more than 1 year educ	− 2.942	− 0.2876	0.03525	0.3583	1.999	− 0.0171
Poverty status 1.00–1.24	− 1.142	− 0.2394	− 0.009009	0.1521	2.042	0.012
American Indian and Alaska Native	− 1.624	− 0.2269	− 0.06771	0.03893	0.8692	0.012

is being observed in an image, reveals features consistent with geographic traits posited in opportunity and ecological perspectives of offending. Google Vision image tags capture crime targets such as vehicles, luxury vehicles, homes, vehicles parked on the street, and vehicles located in parking lots. Guardian aspects of targets are also present, including features like walls, fences, street lights, and gates (one could also contend that the presence of vehicles on the roadway is also a form of guardianship). Physical incivilities, such as graffiti

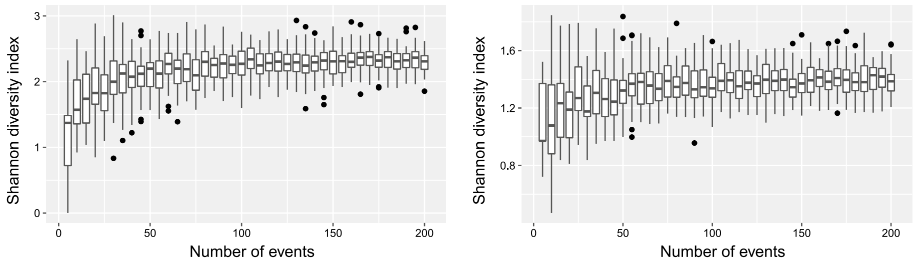

Fig. 3 Shannon diversity index vs. number of crimes for two census block groups (box intervals determined using bootstrap resampling)

Table 11 Predicting crime volume instead of diversity

Response var.	Best predictor	Coef	Stand. err.	Google div. coef	Stand. err.
All crime	Population	0.045	0.001	0.0094	0.001
Property crime	Population	0.043	0.002	0.0083	0.001
Violent crime	Population	0.039	0.002	0.0083	0.001

and trash, are also tagged in place-based images. Features of the environmental backcloth can also be captured, including bars, restaurants, paths, walkways, roadway thoroughfare, river, train tracks, and suburban versus urban. Beyond routine activity and crime pattern theories and related to social disorganization, parks, sports courts, places of worship, and voting centers are tagged in image features commonly associated with neighborhood capacity to increase collective efficacy. Relatedly, Jones and Pridemore (2019) recently articulated a multi-level theory of routine activities and social disorganization to explain crime concentration. Drawing on violent and property crime in Lexington, Kentucky they observed that micro-level social disorganization impacts crime differently at the street-segment level as compared to neighborhood-level disorganization. They go on to note that (p. 22) “The precise mechanisms that distinguish localized forms of social disorganization from neighborhood forms of social disorganization have not been fully considered in the empirical literature, and more research is required to determine the precise reasons for these independent effects”. Notably, their index of local social disorganization included disorder, mixed land used, and suburbanization at the micro-level each of which is captured via Google Vision image tags.

This proves advantageous for crime and place scholars as built environment data, such as land use metrics and business types, are commonly housed by commercial data vendors (such as InfoBase or InfoGroup) or require researchers to merge varying data sets and thereby introducing other sets of challenges. Moreover, GSV data to capture built and environmental features of place enable researchers to examine these features at more granular and localized levels. For instance, most readily available data regarding land use are limited to the block-group level whereas GSV can be drilled down to specific places. GSV images also provide rich content through the Google Vision image tag that can be quantified, such as graffiti or trees, while the images themselves can be analyzed by researchers to code qualitative distinctions akin to a systematic social observation methodology (Odgers et al. 2012). In addition, GSV boasts significant geographic coverage in the U.S., Europe, Australia, and many other countries. This is especially true for urban environments where the majority of crime and place research occurs.

As can be reviewed on Google Street View's "Explore" website, coverage of many suburban areas is robust. Rural areas are less represented and are often limited to major roadway thoroughfares, though planned coverage schedules indicate rural areas will be increasingly captured in the near future. Thus, GSV data may be more accessible to scholars seeking to explore crime and place phenomenon in areas where census-type data (community, built, or environmental) may not be available. We also note that GSV data can be used at micro-geographical units of observation where census data is not defined. For example, the correlation coefficient between GSV diversity and crime diversity increases from .2 to .4 when going from census blocks to 100mx100m grid cells. This is one direction for future research.

Methods used in the present study also have implications for a growing body of spatiotemporal analyses of crime. Repeat and near-repeat phenomena of crime (Bernasco 2008; Short et al. 2009) suggests offenders target and re-offend based on two hypotheses. The boost hypothesis suggests offenders operate among co-offending networks and communicate target suitability to one another (Bowers and Johnson 2004). Grounded in environmental explanations of crime, the flag hypothesis contends that repeat and near-repeat offenses occur as a result of environmental determinants conducive to certain offenses. For example, Townsley et al. (2003) observed the occurrence of near-repeat residential burglary among suburbs with homogeneous houses as opposed to more mixed-land use areas. Put simply, offenders diagnose environments and identify areas ripe for certain offenses, such as a poorly lit parking lot that contains numerous sought-after vehicles with easy street access (Piza and Carter 2018). Unfortunately and as discussed previously, environmental cues observed by offenders are difficulty to quantify when explaining spatiotemporal patterns of crime. Use of GSV may lend additional perspectives to inform the boost versus flag hypotheses debate as scholars would be able to empirically estimate environmental distinctions frequently observed in areas with repeat and near-repeat offenses.

Population and census block land size served as the primary catalysts for crime diversity, lending further support to observations by Brantingham (2016) and Lentz (2018) that crime diversity is indeed a function of geographic size. More specifically, these previous studies suggest crime diversity is a function of specific offending cues and geographic capacity to present such cues to offenders. Our results mirror these conclusions in that as population increases so too does environmental diversity and crime diversity. More specifically, the variety of observed environment diversity as measured by GSV may reflect the variability of cues that either facilitate or inhibit offending at places. Thus, if crime diversity is a function of geographic size, what are the implications for effective intervention?

To begin, police agencies seeking to engage in crime prevention strategies can tailor interventions to areas based upon their population or geographic size, or based on environment diversity indicators (either through GSV or patrol officer observations/intuition). Areas of high population or size may benefit from general prevention approaches proven effective in place-based policing studies, such as random proactive patrols (Koper 1995) or increased officer activity to increase perceptions of being caught (Ratcliffe et al. 2011). In this sense, high environmental and crime diversity may be approached through the general-crime perspective. Such areas are also likely to benefit from improved collective efficacy among residents to buttress informal social controls through community-building programs (Schreck et al. 2009; Sharkey et al. 2017). For areas with reduced environmental and crime diversity, officers can engage in problem-oriented approaches proven to generate notable crime reduction benefits in hot spots (Braga et al. 2014), as well as situational crime prevention techniques (Clarke 1995). This is not to suggest these interventions are mutually exclusive, however police agencies should understand the likelihood of crime-specific interventions may reduce certain crime types, but not generate overall significant crime reductions in a given area (Quick et al. 2018).

Haberman (2017) speaks to this complex intervention scenario in his study of overlapping crime hot spots in Philadelphia. In short, his analysis demonstrated that hot spots are indeed crime-specific and demonstrate minimal overlap or correlation across geographic occurrence of crime type. If the end-goal of a police agency is to reduce overall crime, an agency should target areas that are both high in crime and environmental diversity as well as population. Unfortunately, such an approach would be resource-intensive as officers would be expected to engage in higher volumes of patrols and activities across larger geographic spans. Conversely, agencies seeking to target specific crime problems could identify areas of interest and diagnose environmental factors likely serving as catalysts in such areas to engage in problem-solving, situational crime prevention, and crime prevention through environmental design strategies. As noted by Sampson and colleagues (Sampson et al. 1997), tangible crime reduction efforts are more likely to occur when the focus is on changing places and not people. Thus, the challenge remains as to what characteristics about places need to be changed? Certainly this will vary from place to place, but perhaps Google Vision API is another tool in the toolbox of criminologists and police departments to refine measurement of places and their environmental characteristics to identify such potential changes.

Limitations

This study is not without limitations. We focus on crime diversity and environmental diversity as a means to demonstrate the utility of GSV as a unique and cost-effective means of examining crime at place. Our decision to focus on crime diversity is an attempt to best capture the occurrence of crime at place. Scholars have debated the appropriate outcome of interest for crime at places studies (Jones and Pridemore 2019) and we believe crime diversity reflects the reality of crime occurrence and multi-crime challenges police face at places. Relatedly, our crime diversity measure is limited to the offenses captured by the Los Angeles Police Department, and as measured via their criminal code processing. Though there exist several hundred of such offenses, the maximum number of unique offenses is governed by the number of offense types reported. Future research may seek to refine such offense codes and either collapse or expand upon existing offense types. For example, some crime categories such as fraud and identity theft may not be strongly tied to environmental features observable through street view. Our decision to include fraud offense types was motivated by the reality that they are indeed a criminal offense and speak to the diversity of crime in a given environment. How the environment may, or may not, affect these offense types is beyond the scope of the present study. It may be plausible, given the literature on crime specialization, progressing, and sequencing, that offenders involved in certain crimes may also become involved in other related crimes - and the approach presented in the current study may be a pathway to discovery of these subtle relationships. As environmental diversity computed with street view does not explain all the variation in crime diversity, part of the remaining variation may be due to the presence of these crimes that are not tied to street-observable geography. How to deal with these variables, perhaps by systematically removing them or estimating crime diversity on subsets of cohesive crime topic groups (Kuang et al. 2017), will be a focus of future research.

Furthermore, environmental diversity is a function of the variability captured by Google Vision image tagging within GSV images. In sum, 408 different image tags were captured via Google Vision. Though several tags are quite unique, such as graffiti or iron rod fencing, others are vague or synonymous with other tags. For example, tags such as parking are vague and could reference a large surface parking lot or more finite curbside parking spots. Image tags such as home, house, dwelling, and residential building are rather synonymous. However, a review of sampled images and their associated tags reveal unique differences that may have implications for the nuances of place-based research. For example, an image tag such as ”fence” often refers to a commercial or industrial fence made of iron and of tall height. Conversely, a ”home fencing” tag reflects a fence that is wood and shorter. These similar yet different image tags also speak to the environmental context of a given fence and other objects in the image. ”Home fencing” is a fence in front of a residential home as opposed to a fence surrounding a commercial building. Similarly, different tags that capture residences may appear synonymous but reflect differences between stand alone single-family homes and connected row houses. The GSV Vision API attempts to distinguish between these similar but different structures - each of which have implications for defining offending environments. In fact, the classification is deterministic using a neural network where low level features (image texture, color, edges, and so forth) are then mapped to high level features, resulting in a prediction. Nonetheless, given the limited use of GSV Vision API in criminology, we urge future scholary inquiries to examine this issue in detail to provide an empirical explanation, or road map, as to what features in a given image generate specific tag identities resulting in distinguishable features. Emerging methods and advanced qualitative analytic programs, such as word2vec (Rong 2014), or other rich data mining programs may be fruitful in this endeavor. Our decision to leverage raw image tags in the current study was to maximize transparency with our analyses and enable replication while avoiding subjective or arbitrary collapsing of image tag categories.

Tagged images from Google Vision data face temporal challenges when merging with crime and other sources of data. Images captured by GSV are cross-sectional and represent a given place, or environment, at a single point in time. Though image updating does occur, it does not happen on a frequent basis. Thus, appropriately merging GSV image data with crime data along temporal consistencies presents a challenge that must be navigated. In some cases images used in this study were from 2015, whereas the crime incident data is from 2013 (hence the focus of this paper is on inferring correlations between variables, rather than causality). Google's Street view ”Explore” page provides image collection logs that may be leveraged to align GSV data with sources of data. This issue is less of a concern for image tags such as building, roadways, or urban versus rural as such environmental features are likely to be stagnant. However, indicators of physical incivility such as trash and graffiti as well as the presence of pedestrians or vehicles that may serve as targets or guardians are subject to change over time. Such temporal limitations of GSV data have been well documented in previous studies ([93], Charreire et al. 2014; Langton and Steenbeek 2017; Rundle et al. 2011). It should be noted that many commercial sources of data may suffer from similar time-dependent concerns. For example, official geocoded business records report a cross-sectional snapshot of business types in a given place. If a business were to close and become vacant, or a new business open in the same location, these data would suffer similar shortcomings. Despite these limitations, the present study has demonstrated that GSV can be leveraged as a viable means to capture environmental characteristics of places at the street-level, that such environmental characteristics significantly predict crime diversity consistent with previous research, and that such findings are consistent when examining both property and violent crime at place. These findings represent substantive promise, and challenges, for scholars seeking to leverage innovative methods to examine crime at place in the future.

Appendix: Handling Synonyms in GSV Object Detection

Here we explore the sensitivity of the diversity index with regards to expanding or collapsing GSV categories. Though some image tags may appear to be synonyms, many are in fact distinct and provide a richer understanding of what is contained within a given image. For example, a family home and apartment are clearly different, so too are row houses that are physically connected as compared to standalone single family homes. While these may indeed be homes or residences that could be combined into a single category, they are unique from one another and may have implications for how offending environments are interpreted. We queried a sample of images with similar tags (or potential synonyms) and reviewed these images in an attempt to decipher differences. For example, it appears yard tends to be the backyard of a residence whereas lawn is the front lawn. Certainly these may be synonymous, but they speak to the view of a given location and how an offender may interpret. Other examples include fence which appear to be metal or commercial fencing and home fencing which appear to be smaller wood fences. They are also distinguished by other characteristics in a given image. A residential home with a fence will be tagged with home fencing whereas a commercial building with a fence will be tagged as fence.

The GSV Vision API attempts to distinguish between these similar but different structures. In fact, the classification is deterministic using a neural network where low level features (image texture, color, edges, etc.) are then mapped to high level features, and finally a prediction. However, certain classes may be similar, for example dwelling and house (though there is something in the image that is causing dwelling to be selected, so the images are somehow distinguishable). To assess the potential sensitivity of the environmental diversity index to combining synonyms into more general categories, we run the following simulation. We simulate 1000 census blocks as follows. In each block we first draw a dirichlet distribution for 20 synthetic categories. We then sample M crimes (where M is uniform from 1 to 1000) according to the dirichlet distribution and allocate them to categories. Next we simulate adding synonym noise by distributing each event to one of 5 random synonyms. We finally compute the diversity index of each of these distributions.

In Fig. 4 we see that even when adding synonym noise, the diversity index is highly correlated (.919) with the original distribution. It is also worth pointing out that both Lentz (2018) and Brantingham (2016) encounter this same issue. In both papers, crime categories themselves are somewhat arbitrary, and different types of clustering could be possible. However they both use the raw categories to compute crime diversity, which is simpler, more reproducible, and less subjective. We prefer to take this approach here, both with the crime diversity index (to match more closely with Lentz (2018), Brantingham (2016)) and the GSV index. We believe this approach is most appropriate given the limited use of GSV Vision API in criminology for purposes of future replication. This approach also avoids the subjective and arbitrary process of assigning image tags to collapsed categories.

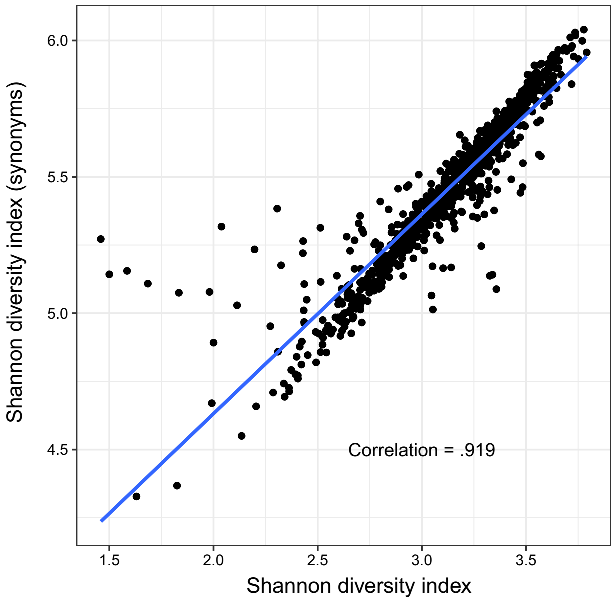

Fig. 4 Shannon diversity index of data with simulated synonyms vs Shannon diversity index of collapsed categories

References

Anselin L (2009) Spatial regression. SAGE Handbook Spatial Anal 1:255–276

Badland HM, Opit S, Witten K, Kearns RA, Mavoa S (2010) Can virtual streetscape audits reliably replace physical streetscape audits? J Urban Health 87(6):1007–1016

Ben-Joseph E, Lee JS, Cromley EK, Laden F, Troped PJ (2013) Virtual and actual: relative accuracy of on-site and web-based instruments in auditing the environment for physical activity. Health Place 19:138–150

Bernasco W (2008) Them again? same-offender involvement in repeat and near repeat burglaries. Europ J Criminol 5(4):411–431

Bivand R, Yu D, Nakaya T, Garcia-Lopez M-A, Bivand MR (2017) Package spgwr, R software package

Bloch S. An on-the-ground challenge to uses of spatial big data in assessing neighborhood character, Geographical Review

Boessen A, Hipp JR (2015) Close-ups and the scale of ecology: land uses and the geography of social context and crime. Criminology 53(3):399–426

Bowers KJ, Johnson SD (2004) Who commits near repeats? a test of the boost explanation. Western Criminol Rev 5:3

Bowers KJ, Johnson SD (2005) Domestic burglary repeats and space-time clusters: the dimensions of risk. Europ J Criminol 2(1):67–92

Braga AA, Clarke RV (2014) Explaining high-risk concentrations of crime in the city: social disorganization, crime opportunities, and important next steps. J Res Crime Delinquency 51(4):480–498

Braga AA, Weisburd D (2010) Policing problem places: crime hot spots and effective prevention. Oxford University Press, Oxford

Braga AA, Papachristos AV, Hureau DM (2014) The effects of hot spots policing on crime: an updated systematic review and meta-analysis. Justice Quarterly 31(4):633–663

Brantingham PJ (2016) Crime diversity. Criminology 54(4):553–586

Brantingham PJ, Brantingham PL et al (1981) Environmental criminology. Sage Publications, Beverly Hills

Brantingham PL, Brantingham PJ (1993) Environment, routine and situation: toward a pattern theory of crime. Adv Criminol Theory 5(2):259–94

Brantingham P, Brantingham P (1995) Criminality of place. Europ J Crim Policy Res 3(3):5–26

Brantingham P, Brantingham P (2013) Crime pattern theory, in Environmental criminology and crime analysis. Willan, pp 100–116
Bursik RJ Jr, Grasmick HG et al (1999) Neighborhoods & crime. Lexington Books
Campbell A, Both A, Sun QC (2019) Detecting and mapping traffic signs from google street view images using deep learning and gis. Comput Environ Urban Syst 77:101350
Caplan JM, Kennedy LW, Miller J (2011) Risk terrain modeling: brokering criminological theory and gis methods for crime forecasting. Justice Quarterly 28(2):360–381
Charreire H, Mackenbach JD, Ouasti M, Lakerveld J, Compernolle S, Ben-Rebah M, McKee M, Brug J, Rutter H, Oppert J-M (2014) Using remote sensing to define environmental characteristics related to physical activity and dietary behaviours: a systematic review (the spotlight project). Health Place 25:1–9
Clarke RV (1995) Situational crime prevention. Crime Justice 19:91–150
Clarke RVG (1997) Situational crime prevention. Criminal Justice Press, Monsey
Clarke RV, Cornish DB (1985) Modeling offenders' decisions: a framework for research and policy. Crime Justice 6:147–185
Clarke P, Ailshire J, Melendez R, Bader M, Morenoff J (2010) Using google earth to conduct a neighborhood audit: reliability of a virtual audit instrument. Health Place 16(6):1224–1229
Clarke R, Eck JE (2014) Become a problem-solving crime analyst. Willan
Cloward RA, Ohlin LE (2013) Delinquency and opportunity: a study of delinquent gangs. Routledge, Milton Park
Cohen LE, Felson M (1979) Social change and crime rate trends: a routine activity approach. American sociological review, pp 588–608
Cornish DB, Clarke RV (1986) The reasoning criminal: rational choice perspectives on offending. Transaction Publishers
Eck JE (1995) A general model of the geography of illicit retail marketplaces. Crime Place 4:67–93
Eck J, Weisburd DL (2015) Crime places in crime theory, Crime and place: Crime prevention studies, vol 4
Felson M (2006) Crime and nature. Sage, Thousand Oaks
Graif C, Sampson RJ (2009) Spatial heterogeneity in the effects of immigration and diversity on neighborhood homicide rates. Homicide Stud 13(3):242–260
Groff ER, 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. J Quantitat Criminol 26(1):7–32
Haberman CP (2017) Overlapping hot spots? examination of the spatial heterogeneity of hot spots of different crime types. Criminol Public Policy 16(2):633–660
Harrell FE Jr (2015) Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis. Springer, Berlin
He L, Páez A, Liu D (2017) Built environment and violent crime: an environmental audit approach using google street view. Comput Environ Urban Syst 66:83–95
Jean PKS (2008) Pockets of crime: broken windows, collective efficacy, and the criminal point of view. University of Chicago Press, Chicago
Johnson SD (2010) A brief history of the analysis of crime concentration. Europ J Appl Math 21(4–5):349–370
Jones RW, Pridemore WA (2019) Toward an integrated multilevel theory of crime at place: Routine activities, social disorganization, and the law of crime concentration. J Quantitat Criminol 35(3):543–572
Keizer K, Lindenberg S, Steg L (2008) The spreading of disorder. Science 322(5908):1681–1685
Kelly CM, Wilson JS, Baker EA, Miller DK, Schootman M (2012) Using google street view to audit the built environment: inter-rater reliability results. Ann Behav Med 45(1):S108–S112
Kennedy LW, Caplan JM, Piza E (2011) Risk clusters, hotspots, and spatial intelligence: risk terrain modeling as an algorithm for police resource allocation strategies. J Quantitat Criminol 27(3):339–362
Kindynis T (2014) Ripping up the map: criminology and cartography reconsidered. British J Criminol 54(2):222–243
Kinney JB, Brantingham PL, Wuschke K, Kirk MG, Brantingham PJ (2008) Crime attractors, generators and detractors: land use and urban crime opportunities. Built Environ 34(1):62–74
Koper CS (1995) Just enough police presence: reducing crime and disorderly behavior by optimizing patrol time in crime hot spots. Justice Quarterly 12(4):649–672
Kuang D, Brantingham PJ, Bertozzi AL (2017) Crime topic modeling. Crime Sci 6(1):12
Kubrin CE, Weitzer R (2003) New directions in social disorganization theory. J Res Crime Delinquency 40(4):374–402

Langton SH, Steenbeek W (2017) Residential burglary target selection: an analysis at the property-level using google street view. Appl Geogr 86:292–299
Lantz B, Ruback RB (2017) A networked boost: burglary co-offending and repeat victimization using a network approach. Crime Delinquency 63(9):1066–1090
Lentz TS (2018) Crime diversity: reexamining crime richness across spatial scales. J Contemp Criminal Justice 34(3):312–335
Lopes DPF, Neves AJ. A practical study about the google vision api
Lowenkamp CT, Cullen FT, Pratt TC (2003) Replicating sampson and groves's test of social disorganization theory: revisiting a criminological classic. J Res Crime Delinquency 40(4):351–373
Malm A, Pollard N, Brantingham P, Tinsley P, Plecas D, Brantingham P, Cohen I, Kinney B (2005) A 30 year analysis of police service delivery and costing. Int Centre Urban Res Stud (ICURS)
OBrien DT (2018) The action is everywhere, but greater at more localized spatial scales: Comparing concentrations of crime across addresses, streets, and neighborhoods. J Res Crime Delinquency, p 0022427818806040
Odgers CL, Caspi A, Bates CJ, Sampson RJ, Moffitt TE (2012) Systematic social observation of children's neighborhoods using google street view: a reliable and cost-effective method. J Child Psychol Psychiatry 53(10):1009–1017
Osgood DW, Schreck CJ (2007) A new method for studying the extent, stability, and predictors of individual specialization in violence. Criminology 45(2):273–312
Palmer EJ, Holmes A, Hollin CR (2002) Investigating burglars' decisions: factors influencing target choice, method of entry, reasons for offending, repeat victimisation of a property and victim awareness. Security J 15(1):7–18
Park RE, Burgess EW, McKenzie R (1925) The growth of the city: an introduction to a research project, The city, pp 47–62
Piquero A, Oster RP, Mazerolle P, Brame R, Dean CW (1999) Onset age and offense specialization. J Res Crime Delinquency 36(3):275–299
Piza EL, Carter JG (2018) Predicting initiator and near repeat events in spatiotemporal crime patterns: an analysis of residential burglary and motor vehicle theft. Justice Quarterly 35(5):842–870
Pratt TC, Cullen FT (2005) Assessing macro-level predictors and theories of crime: a meta-analysis. Crime Justice 32:373–450
Quick M, Li G, Brunton-Smith I (2018) Crime-general and crime-specific spatial patterns: a multivariate spatial analysis of four crime types at the small-area scale. J Criminal Justice 58:22–32
Ratcliffe JH, Taniguchi T, Groff ER, Wood JD (2011) The philadelphia foot patrol experiment: a randomized controlled trial of police patrol effectiveness in violent crime hotspots. Criminology 49(3):795–831
Rong X (2014) word2vec parameter learning explained. arXiv:1411.2738
Rundle AG, Bader MD, Richards CA, Neckerman KM, Teitler JO (2011) Using google street view to audit neighborhood environments. Am J Prevent Med 40(1):94–100
Sampson RJ (2013) The place of context: a theory and strategy for criminology's hard problems. Criminology 51(1):1–31
Sampson RJ, Groves WB (1989) Community structure and crime: testing social-disorganization theory. Am J Sociol 94(4):774–802
Sampson RJ, Raudenbush SW, Earls F (1997) Neighborhoods and violent crime: a multilevel study of collective efficacy. Science 277(5328):918–924
Schreck CJ, McGloin JM, Kirk DS (2009) On the origins of the violent neighborhood: a study of the nature and predictors of crime-type differentiation across chicago neighborhoods. Justice Quarterly 26(4):771–794
Shannon CE (1948) A mathematical theory of communication. Bell Syst Tech J 27(3):379–423
Sharkey P, Torrats-Espinosa G, Takyar D (2017) Community and the crime decline: the causal effect of local nonprofits on violent crime. Am Sociol Rev 82(6):1214–1240
Shaw CR, McKay HD (1942) Juvenile Delinquency and urban areas: a study of rates in relation to differential characteristics of local communities in American life. University of Chicago Press, Chicago
Short MB, Dorsogna MR, Brantingham PJ, Tita GE (2009) Measuring and modeling repeat and near-repeat burglary effects. J Quantitat Criminol 25(3):325–339
Smith MJ (2003) Exploring target attractiveness in vandalism: an experimental approach. Crime Prevent Stud 16:197–236
Stucky TD, Ottensmann JR (2009) Land use and violent crime. Criminology 47(4):1223–1264
Sytsma VA, Piza EL (2018) Script analysis of open-air drug selling: a systematic social observation of cctv footage. J Res Crime Delinquency 55(1):78–102

Taylor RB (2015) Community criminology: Fundamentals of spatial and temporal scaling, ecological indicators, and selectivity bias. NYU Press
Taylor RB, Gottfredson S (1986) Environmental design, crime, and prevention: an examination of community dynamics. Crime Justice 8:387–416
Taylor BT, Fernando P, Bauman AE, Williamson A, Craig JC, Redman S (2011) Measuring the quality of public open space using google earth. Am J Prevent Med 40(2):105–112
Thomas WI, Znaniecki F (1918) The polish peasant in Europe and America: monograph of an immigrant group, vol 2. University of Chicago Press, Chicago
Thrasher FM (1927) The gang: a study of 1313 gangs in Chicago. University of Chicago Press, Chicago
Townsley M, Homel R, Chaseling J (2003) Infectious burglaries a test of the near repeat hypothesis. Br. J. Criminol. 43(3):615–633
Vandeviver C (2014) Applying google maps and google street view in criminological research. Crime Sci 3(1):13
Vandeviver C, Bernasco W (2019) location, location, location: Effects of neighborhood and house attributes on burglars target selection. J Quantitat Criminol, pp 1–43
Veysey BM, Messner SF (1999) Further testing of social disorganization theory: an elaboration of sampson and groves's community structure and crime. J Res Crime Delinquency 36(2):156–174
Warner BD (2003) The role of attenuated culture in social disorganization theory. Criminology 41(1):73–98
Weisburd D (2015) The law of crime concentration and the criminology of place. Criminology 53(2):133–157
Weisburd D, Maher L, Sherman L, Buerger M, Cohn E, Petrisino A et al (1992) Contrasting crime general and crime specific theory: the case of hot spots of crime. Adv Criminol Theory 4(1):45–69
Welsh BC, Farrington DP (2004) Surveillance for crime prevention in public space: results and policy choices in britain and america. Criminol Public Policy 3(3):497–526
Willits DW, Makin DA (2018) Show me what happened: analyzing use of force through analysis of body-worn camera footage. J Res Crime Delinquency 55(1):51–77
Wilson JQ, Kelling GL (1982) Broken windows. Atlantic monthly 249(3):29–38
Wright RT, Decker SH (1996) Burglars on the job: Streetlife and residential break-ins. UPNE
Wright R, Logie RH, Decker SH (1995) Criminal expertise and offender decision making: an experimental study of the target selection process in residential burglary. J Res Crime Delinquency 32(1):39–53

Authors and Affiliations

Samira Khorshidi[1] · Jeremy Carter[2] · George Mohler[1] · George Tita[3]

[1] Department of Computer and Information Science, Indiana University - Purdue University, Indianapolis, USA

[2] O'Neill School of Public and Environmental Affairs, Indiana University - Purdue University, Indianapolis, USA

[3] Department of Criminology, Law and Society, University of California, Irvine, USA