



An AI-based framework for studying visual diversity of urban neighborhoods and its relationship with socio-demographic variables

Md Amiruzzaman¹ · Ye Zhao² · Stefanie Amiruzzaman³ · Aryn C. Karpinski⁴ · Tsung Heng Wu²

Received: 13 June 2022 / Accepted: 29 November 2022 / Published online: 28 December 2022
© The Author(s), under exclusive licence to Springer Nature Singapore Pte Ltd. 2022

Abstract

This study presents a framework to study quantitatively geographical visual diversities of urban neighborhood from a large collection of street-view images using an Artificial Intelligence (AI)-based image segmentation technique. A variety of diversity indices are computed from the extracted visual semantics. They are utilized to discover the relationships between urban visual appearance and socio-demographic variables. This study also validates the reliability of the method with human evaluators. The methodology and results obtained from this study can potentially be used to study urban features, locate houses, establish services, and better operate municipalities.

Keywords Google street-view · Visual diversity · Urban neighborhood · Interrater reliability · Social phenomena · Semantic segmentation

Introduction

Geospatial visual appearance depends on many factors, such as built structures (roads, buildings, and sidewalks), greenery, and openness as well as the presence of different visual objects and their ratio in an environment [1, 2]. Visual appearance of built environments are inherently related to various socio-economic outcomes, such as population concentration, economic disparity, prevalence of crime, and pedestrian safety [1, 3–7].

✉ Md Amiruzzaman
mamiruzzaman@wcupa.edu

¹ Department of Computer Science, West Chester University, West Chester, PA, USA

² Department of Computer Science, Kent State University, Kent, OH, USA

³ Department of Languages and Cultures, West Chester University, West Chester, PA, USA

⁴ Research, Measurement & Statistics, Kent State University, Kent, OH, USA

In this study, we define the *geospatial visual diversity* as a criterion by which to understand the visual appearance of a geographic area, which is considered as an important component of environmental design [8]. This definition of geospatial visual diversity is guided by previous studies, such as [8, 9]. According to Stamps III [8], geospatial visual diversity is an important component of environmental design. It contributes to understanding scenic beauty and esthetically pleasing landscapes [9–11], comparing neighborhoods [1, 5, 12], and providing a subjective visual preference [13]. A few existing approaches [7, 9, 14, 15] quantified visual diversity with different metrics, such as entropy [8].

Recently, AI-based image segmentation tools were utilized in extracting semantic object information from street-view images, which eased the burden of data access and computing [1, 10, 12, 16–18]. In the existing work, however, the extracted semantics from street-view images was not employed in computing geographical visual diversity.

In this study, we compute and investigate a variety of visual diversity indices based on the AI tools and a large set of street-view images. Further, we compare multiple indices to see which index is more suitable and how the indices relate to multiple social phenomena.

This study aims to advance urban technology in four ways:

- (1) Presenting a computational framework of geographical visual diversity based on the semantic segmentation that extracts semantically segmented information from a large set of street-view images of a neighborhood;
- (2) Computing and comparing multiple types of visual diversity indices, including both single group diversity indices and multi-group diversity indices;
- (3) Validating the reliability of using the computed diversity indices through a study with human evaluators; and
- (4) Extracting social-demographic information including economy, population and crime metrics, and studying the correlations between the visual diversity indices with the socio-demographic variables.

The contribution of this research is twofold: (1) by measuring visual diversity from street-view images for urban studies; and (2) by recognizing implications for urban neighborhood planning based on visual diversity. Of specific value to the current research, this study demonstrates the process and value in examining can reveal relationships between street-level urban design qualities and property values.

Related work

The term *diversity* helps to quantify and compare social phenomena [19]. Measuring diversity is crucial in several disciplines [20], such as economics [21], ecology [22], urban planning [9], and social studies (e.g., culture). Diversity can also help to understand and assess the distribution of resources of an area, such as greenery,

water, wetlands, and land use [23–25]. It also relates to human movement and urbanization [26, 27] as well as home prices [9].

De Jonge [28] interviewed residents and found that most people prefer to live in areas with more visually diverse. Stamps III [8, 29–31] reported similar findings, indicating that when an urban area has more visual diversity, people find the area more appealing. Visual diversity can be related to and improve livability in urban planning [9, 32].

Many past studies have used Street-View images to explore information about the environment. For example, presence of openness [33], healthy areas [5], green areas [33, 34], crime-prone areas [1, 35], local businesses [36], land use and vacant areas [37], urban effect over time [38, 39], voting pattern analysis [40], COVID-19 affected areas [41], and landscape analysis [18, 39]. Most studies have used street-view images because it is readily available for almost all cities and saves researchers time to walk around and collect street-view images from different geographic locations.

The aforementioned studies competently focused on particular problems and used street-view images as their data source. Past studies that used street-view image had different research goals than this study. However, one notable study seems similar to the study. For example, Wen, Liu, and Wu [18] used an entropy weighted method to quantify ecological matrices. While Wen et. al.'s [18] study contributes to urban planning by focusing on ecological aspects, this current study concentrates on other social aspects and uses both multi-category and single-category diversity indices. Literature review indicates a gap between the neighborhood and its relationship with social-demographic variables. This study also presents its results in terms of social aspects, such as population, economics, and crime.

In the past, the visual content of a map, such as landscape and land cover type, served as primary sources of information that contributed to measuring visual diversity [42]. In a recent study, Zhang and Dong [9] used the horizontal green view index (HGVI) to measure the visual diversity of greenery from street-view images. In this study, we instead study a set of diversity indices from multiple visual categories computed from street-view images, and we further examine their relationships to social variables.

Urban data and AI-based street-view image segmentation

We collected two types of data: (1) socio-demographic information; and (2) street-view images from a U.S. metropolitan area. The details of our data collection procedure is explained in following sections. Please see sections "[Open street-view image data](#)" for socio-demographic data collection, "[Open street-view image data](#)" for street-view image data collection, and "[Reliability study with human evaluator](#)" for human evaluator data collection.

Open socio-demographic information

For the socio-demographic information of neighborhoods, the researchers used open data from Zillow. Zillow is an online real estate marketplace [43]. For crime-related information, open data from the FBI Uniform Crime Report were obtained [44]. In addition, other socio-demographic information, such as population size and population density per square mile, was gathered from open data at Area Vibes. Area Vibes is a website that measures various neighborhood population parameters [45, 46].

Open street-view image data

Using a neighborhood's boundary information downloaded from Zillow, the street network of the neighborhood was retrieved from OpenStreetMap [47]. Next, a large set of locations were sampled on the street network. In particular, each street is sampled with points with a 20-meter distance (see Fig. 1) so as to capture each block from the neighborhood [10]. A similar approach was applied by previous studies as well [1, 39].

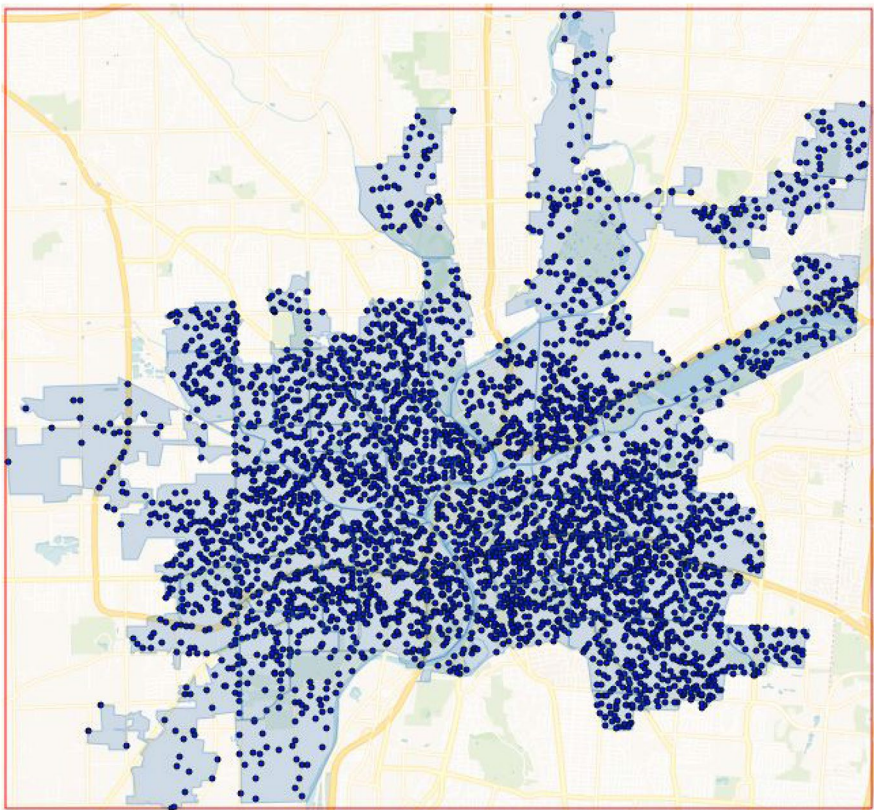


Fig. 1 A large set of sampling locations in a region

The blue dots in Fig. 1 represent the generated segmented geolocations, which are 20 m apart from each other.

Then, we used the segmented geolocations to download street-view images with the help of the Google Street-View (GSV) Application Programming Interface (API). We were only interested in the side-views (see Fig. 2) and ignored the front-view of roads. We ignored front and back views because those would mostly include roads, cars, and the sky, which could dominate the diversity computation (see Fig. 2); instead, we were mostly interested in scenic views. Scenic views primarily consist of the side-view of the street-view images [1].

To obtain side-views of the streets we computed the heading of the street and then added 90 degrees and 270 degrees, respectively (see Eq. 1). This idea was adopted from [1]:

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

where

$$\begin{aligned} x &= \{\cos(\text{lat}_0) \times \sin(\|\text{lat}_0 - \text{lng}_2\|)\}, \\ y &= \{\cos(\text{lat}_0) \times \sin(\text{lat}_2) \\ &\quad - \sin(\text{lat}_0) \cos(\text{lat}_2) \times \cos(\|\text{lng}_0 - \text{lng}_2\|)\}, \end{aligned} \quad (1)$$

where θ is the heading for geolocation $(\text{lat}_1, \text{lng}_1)$, using geolocation $(\text{lat}_0, \text{lng}_0)$ before the geolocation $(\text{lat}_1, \text{lng}_1)$ and after the geolocation $(\text{lat}_2, \text{lng}_2)$.

Semantic segmentation

From each street-view image, a deep learning-based semantic segmentation tool, PSP-net [48], extracted the visual information of each category from a total of 19 categories, namely road, sidewalk, building, wall, fence, pole, traffic light, traffic sign, vegetation, terrain, sky, person, rider, car, truck, bus, train, motorcycle, and bicycle. The segmented proportion for category i is computed as

$$c_i = \frac{\text{The number of pixels in category } i}{\text{Total number of pixels}}. \quad (2)$$



Fig. 2 Street-view images from four different sides: left view; front view; right view; and back view

The PSPnet model can reach 82.2% accuracy [48], as a state-of-the-art AI model at the time of this study. Figure 3 shows several examples of semantic segmentation results.

Visual diversity indices

The visual diversity was computed using a selection of indices in two categories: (1) multi-category indices; and (2) single-category indices. We computed multi-category indices to understand how different categories holistically affect the environment, while single-category indices explored individual category affects.

The PSPnet model provided the same 19 above-noted categories from each image, road, sidewalk, building, wall, fence, pole, traffic light, traffic sign, vegetation, terrain, sky, person, rider, car, truck, bus, train, motorcycle, and bicycle [10]. An initial qualitative review of these categories found that some were considered more “static” while others were more “transient.” Static categories consist of immobile objects, such as buildings, trees, sky, and fences. Transient categories include mobile objects, such as persons, cars, and trains.

We studied the 19 categories extracted from PSPNet in all images. Interestingly, we noticed that only five categories (i.e., road, building, vegetation, terrain, and sky) are normally distributed. The remainder of the categories are highly skewed and have high kurtosis values. Further analysis indicated that more than 97% images lacked presence of “transient” categories. Further, the presence of “transient” categories are more of time dependent. For example, we might see more presence of cars, bicycles, motorcycles during working hours, than weekends and after work hours. As such, we only considered these five categories in the diversity computation. For a given spatial unit (e.g., block, neighborhood, city), k street-view images in this unit were selected to compute the geospatial visual diversities shown below.

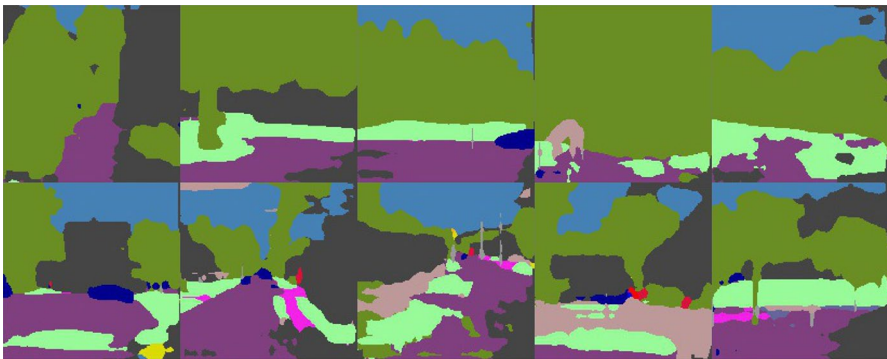


Fig. 3 Segmented images used in the content validation: (top) with low geospatial visual diversity and (bottom) with high geospatial visual diversity. Green: greenery such as trees and bushes. Dark gray: buildings. Light blue: sky or openness. Purple: roads and driveways. Light green: grass or vegetation. Beige: fences. Dark blue: cars. Red: person. Pink: sidewalk

Multicategory diversity indices

The following section describes the multi-category diversity indices used in this study.

Simpson index

The Simpson index [49] considers a sum of individual categories with respect to the sum of all categories:

$$D = 1 - \frac{\sum n_i \times (n_i - 1)}{N \times (N - 1)}, \quad (3)$$

where $n_i = \sum c_i^k$ for category $i \in [1, 5]$ in all k images and $N = \sum n_i$.

McIntosh index

The value of the McIntosh index [50] varies from 0 (no diversity) to 1 (extreme diversity):

$$M = 1 - \frac{N - \sqrt{\sum n_i^2}}{N - \sqrt{N}}, \quad (4)$$

Here n and N are the same as in the Simpson index.

Multiple-category entropy (MCE)

Considering the popularity of entropy as a diversity index (i.e., entropy of Shannon's entropy $H(p) = -\sum p_i \times \log_2(p_i)$, where p_i is the probability of each category), we extended it to multiple categories as presented in the following Eq. 5:

$$H = -\sum_{i=1}^n \sum_{j=1}^k (p_1, p_2, \dots, p_k) \log \left(\prod_{j=1}^k (p_1, p_2, \dots, p_k) \right), \quad (5)$$

where p is the probability of a single category computed by $p_i = \frac{c_i}{n_i}$.

Single-category indices

Entropy

One of the most-used diversity indexes is Shannon's entropy or entropy [51, 52]. A few recent studies also used entropy to compute diversity, such as [52, 53], and [18].

Entropy is computed as

$$H(p) = -\sum p_i \times \log_2(p_i), \quad (6)$$

where p is the same as in Equation 5. A comparable diversity index value obtained can be further obtained as

$$D = e^H, \quad (7)$$

where e is the Euler number (i.e., $e = (1 + 1/n)^n$) [54].

Horizontal view index

Li, Zhang, Li, Ricard, Meng, and Zhang [55] and Zhang and Dong [9] used a Horizontal View Green Index (HGVI) to measure the greenery of an area. To generalize this index, we call it Horizontal View Index (HVI), which is computed as

$$HVI = \frac{\sum n_i}{\sum N_i} \times 100. \quad (8)$$

Gini index

One of the most popular diversity indexes to understand economic diversity or income inequality is the Gini index [56]. The Gini index is frequently used to measure diversity in other fields, such as social and health [57]. Considering its widespread use, we used this index to understand individual category diversity. In this study, the Gini index was computed for each individual category in four steps: (1) sum all the proportion values of the category in all images; (2) sort the values in ascending order; (3) divide each value by the sum to get probabilities of each value; (4) compute the cumulative sum of all the probabilities.

Reliability study with human evaluator

It is important to validate the reliability of the proposed method (i.e., the consistency interpreting the definition of geospatial visual diversity and computational indices). Reliability analyses such as Inter-Evaluator Reliability or Inter-Rater Reliability (IRR), are helpful to measure the consistency among human ratings and computation results [58]. We conducted an IRR study with a group of five human evaluators. One of the biggest advantages of IRR is that it does not require a large pool of raters. Even in some cases, only raters can be sufficient to measure IRR [58, 59]. The definition of geospatial visual diversity, rating scales, and sample images were provided to the evaluators, who evaluated each image for its visual diversity.

We analyzed the semantic segmentation information of an urban neighborhood and found that the data was not normally distributed, which means that there were some outliers. To remove the outliers, we computed the z -scores of static categories, and we removed values ± 3 ($\alpha = .01$). Stamps III [8] showed that the visual diversity of individual images can be calculated using the entropy values. Thus, we computed entropy values for each image vector.

To find visually low and high diversity images, first, we computed the median from the entropy values and divided the dataset into four quartiles. As for the visually low diversity images, we selected five random images from the first quartile, and for the visually high diversity images, we selected five random images from the third quartile.

Second, ten representative images are used with a mix of low and high visual diversities as shown in Fig. 3. Then, the evaluators who were not aware of the actual calculated diversity rated each image for its level of visual diversity, using a 5-point Likert scale (i.e., 1 = Not Diverse and 5 = Extremely Diverse). The IRR of the evaluators were calculated by Intra-Class Correlation (ICC) [58] and Krippendorff's alpha [60], which were computed by the rating variance. Our resulting ICC was .836 for single measures and .953 for average measures. These ICC values indicated that the evaluators had a high degree of agreement and suggested that the evaluators similarly scored diversity in the images. The high ICCs also meant that a minimal amount of measurement error attributed to the evaluators, and thus power was not reduced. In addition, the obtained Krippendorff's Alpha was .775, which likewise indicated a moderate to high degree of agreement among evaluators [60]. In this study, SPSS version 24.0 was used to compute the ICC and Krippendorff's Alpha values. Finally, we performed correlation analysis between the average ratings and geospatial visual diversity indices. Before conducting the correlation analysis, we ran the assumptions of correlation and noticed that the data was normally distributed. Seeing this, we used Pearson correlation to analyze the relationship between average ratings and geospatial visual diversity indices. The correlations appear in Table 1.

The correlation between evaluators' average ratings and multiple category visual diversity index (e.g., Simpson index) was positive for both low and high diversity images (see Table 1). This indicates that the Simpson index could be helpful in assessing both low and high diverse images for a geospatial urban area. Despite this, the McIntosh index could be more appropriate for low diverse areas, whereas for high diverse areas, multi-category entropy might be the superior index to use.

For the single-category indices, entropy could be helpful to assess diversity for the building and greenery of a high diverse area. Similarly, the HVI index could be good for assessing a high diverse area in terms of building, greenery, and sky (see Table 1). Overall, the results here indicated that multiple category diversity indices show a stronger relationship than single category indices. This finding makes conceptual sense as different aspects of the geospatial area, and their proportions are typically considered together when appraising diversity, while single-category indices are better able to appraise which individual category impacts overall diversity.

Evidence from this study suggests that the validity for geospatial visual diversity computed from street-view images was high. It should be noted that the sample size was five, meaning that several correlation values might be high, but not significant. In this study, magnitudes and directions of the correlation values were the focus.

Correlation study between visual diversity and social variables

Street-view images totaling 351,246 in number were analyzed from 86 neighborhoods in a Midwest metropolitan area consisting of two major cities (City 1 and City 2). The sample size for the correlation analysis was computed using G*Power [61]. An a priori power analysis indicated that a total sample of 85 would be needed to detect medium effects with 80% power using an alpha of .05. The sample size for

Table 1 Correlation between average ratings and the computed geospatial visual diversity indices (upper diagonal correlation values are from high diversity images, and lower diagonal from low diversity images)

ht	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1 Avg. Rating														
2 Simpson	.789													
3 McIntosh	.923*	.600												
4 MCE	-.208	-.210	.137											
5 E Road	.129	.191	.402	.870										
6 E Building	-.181	-.688	.090	.372	.003									
7 E Green	.063	-.207	.406	.854	.655	.690								
8 E Terrain	-.049	.495	-.067	.419	.595	-.651	.030							
9 E Sky	.083	.116	.347	.897*	.806	.271	.871	.505						
10 HVI Road	-.079	.023	.214	.918*	.978**	.046	.647	.601	.791					
11 HVI Building	-.084	-.647	.103	.082	-.252	.954*	.476	-.844	.005	-.229				
12 HVI Green	.181	.317	-.190	-.975**	-.788	-.566	-.933*	-.213	-.869	-.831	-.300			
13 HVI Terrain	-.040	.463	-.017	.524	.652	-.534	.177	.987**	.631	.656	-.752	-.335		
14 HVI Sky	.168	.277	.378	.824	.792	.115	.780	.610	.984**	.757	-.140	-.770	.725	

* = $p < .05$, ** = $p < .01$, *** = $p < .001$

this study was 86 neighborhoods, enough to achieve 80% statistical power. Table 2 reports the result. Next, we present a few examples.

Relationship between geospatial visual diversity indices and population metrics

The results here indicated that there were mainly negative correlations between multi-category diversity indices and population metrics. For instance, Simpson index ($r_s = -.328$, $p < .05$ for City 1 and $r_s = -.442$, $p < .01$ for City 2, respectively) and McIntosh index ($r_s = -.390$, $p < .01$ and $r_s = -.423$, $p < .01$, respectively) have medium, negative correlation with total population. There were strong, negative correlations between MCE index ($r_s = -.502$, $p < .01$ and $r_s = -.513$, $p < .01$, respectively) and total population.

For diversity indices of individual visual category, the relationships depend on specific categories. For instance, high diversity of building and terrain often links to a large population. There were strong, negative correlations between total population and Gini index of building ($r_s = -0.727$, $p < .001$ and $r_s = -0.531$, $p < .01$, respectively), and the same occurred for the Gini index of terrain ($r_s = -0.706$, $p < .001$ and $r_s = -0.657$, $p < .001$, respectively). In addition, negative correlation was reflected between population density per square mile and Gini index of building ($r_s = -0.461$, $p < .01$ for City 1 and $r_s = -0.432$, $p < .01$ for City 2 respectively). In a similar manner, a negative correlation occurred between MCE index and population density per square mile ($r_s = -.309$, $p < .05$ and $r_s = -.420$, $p < .01$, respectively).

Relationship between geospatial visual diversity indices and economic indicators

For multi-category indices, negative correlations appeared between diversity and household income in a neighborhood. In some examples, this occurred between Simpson index and median household income ($r_s = -0.785$, $p < .001$ for City 1 and $r_s = -0.387$, $p < .05$ for City 2, respectively), and between McIntosh index and median household income ($r_s = -0.756$, $p < .001$ and $r_s = -0.349$, $p < .05$, respectively). For individual categories, positive correlations were detected between HVI of green and median household income ($r_s = 0.780$, $p < .001$ for City 1 and $r_s = 0.368$, $p < .05$ for City 2, respectively). Similarly, there were positive correlations between HVI index of green and median home value ($r_s = 0.663$, $p < .001$ and $r_s = 0.309$, $p < .05$, respectively). Greenery is often an indicator of a high-income neighborhood. Conversely, personal income has a negative correlation with diversity of building: HVI of building and median household income ($r_s = -0.551$, $p < .01$ and $r_s = -0.370$, $p < .05$, respectively).

Relationship between geospatial visual diversity indices and crime metrics

For multi-category indices, high diversity often indicates high crime activities. In the table, there were positive correlations between Simpson index and violent crime ($r_s = 0.721$, $p < .001$ for City 1 and $r_s = 0.323$, $p < .05$ for City 2, respectively), McIntosh index and violent crime ($r_s = 0.691$, $p < .001$ and $r_s = 0.337$, $p < .05$,

Table 2 Correlation between geospatial visual diversity and economic, population, and crime metrics (upper diagonal correlation values are from City 1, and lower diagonal from City 2)

Variable	1	2	3	4	5	6	7	8	9	10	11	12
1 Simpson		.984***	.489**	-.505**	.098	.943***	.253	-.943***	-.096	-.099	-.213	-.148
2 McIntosh	.970***		.482**	-.469**	.119	.950***	.272	-.926***	-.181	-.181	-.297	-.224
3 MCE	-.004	-.043		-.274	.533**	.583***	.789***	-.419*	-.404*	-.417*	-.467**	-.457**
4 Gini Road	-.704***	-.678***	.042		.224	-.349*	.061	.588***	-.440**	-.401*	-.332	-.376*
5 Gini Building	.085	.201	.030	.218		.231	.712***	-.095	-.602***	-.637***	-.595***	-.624***
6 Gini Green	.792***	.807***	.322*	-.388*	.314		.431*	-.873***	-.309	-.309	-.411*	-.359*
7 Gini Terrain	.406*	.449**	.548***	-.097	.549***	.714***		-.154	-.728***	-.751***	-.754***	-.788***
8 Gini Sky	-.807***	-.829***	-.108	.530**	-.163	-.843***	-.480**		-.020	-.004	.092	.034
9 Entropy Road	-.019	.169	.070	-.315	-.763***	-.241	-.464**	.078		.994***	.983***	.987***
10 Entropy Building	-.113	.259	.048	-.193	-.802***	-.328*	-.540***	.164	.981***		.985***	.991***
11 Entropy Green	-.328*	-.459**	.086	.054	-.685***	-.464**	-.544***	.309	.891***	.937***		.989***
12 Entropy Terrain	-.334*	-.458**	-.036	.041	-.693***	-.526**	-.663***	.327*	.876***	.924***	.979***	
13 Entropy Sky	-.057	.208	.185	-.103	-.688***	-.182	-.397*	.021	.939***	.949***	.927***	.897***
14 HVI Road	.779***	.816***	.064	-.856***	.174	.632***	.434**	-.613***	-.071	-.204	-.425**	-.433**
15 HVI Building	.761***	.684***	-.138	-.554***	-.246	.450**	.088	-.474**	.134	.099	-.141	-.138
16 HVI Green	-.980***	-.971***	.046	.690***	-.081	-.783***	-.347*	.814**	.046	.132	.360*	.355*
17 HVI Terrain	-.556***	-.586***	-.444**	.369*	-.303	-.722***	-.894***	.574***	.296	.392*	.472**	.578***
18 HVI Sky	.744***	.783***	.185	-.382*	.360*	.921***	.619***	-.921***	-.278	-.368*	-.479**	-.514**
19 Income	-.387*	-.349*	-.263	.053	-.387*	-.421**	-.439**	.311	.225	.238	.244	.303
20 Home Value	-.383*	-.346*	-.223	-.116	.021	-.256	-.114	.071	.101	-.134	-.099	-.037
21 Total Pop.	-.442**	-.423**	-.513**	.069	-.531**	-.554***	-.657***	.391*	.492**	.541***	.512**	.557***
22 Pop. Density	.177	.154	-.420**	-.125	-.432**	-.477**	-.615***	.248	.350*	.379*	.295	.354*
23 Violent Crime	.323*	.337*	.151	.084	.033	.328*	.215	-.155	.019	.041	.003	-.063
24 Property Crime	.324*	.339*	.128	.099	.063	.321*	.223	-.145	-.007	.014	-.024	-.089

Table 2 (continued)

Variable	13	14	15	16	17	18	19	20	21	22	23	24
1 Simpson	.161	.767***	.628***	-.987***	-.341*	.958***	-.785***	-.652***	-.328*	.102	.721***	.684***
2 McIntosh	.077	.785***	.598***	-.973***	-.335	.952***	-.756***	-.615***	-.390**	.084	.691***	.641***
3 MCE	-.295	.669***	.219	-.476**	-.793***	.511**	-.412*	-.358*	-.502**	-.309*	.373*	.352*
4 Gini Road	-.527**	-.661***	-.220	.476**	.163	-.530**	.194	.071	.178	.079	-.087	-.142
5 Gini Building	-.536**	.261	-.334	-.026	-.465**	.230	-.235	-.186	-.727***	-.461**	.175	.136
6 Gini Green	-.045	.759***	.590***	-.939***	-.465**	.926***	-.807***	-.709***	-.385*	.012	.751***	.707***
7 Gini Terrain	-.662***	.441**	.155	-.245	-.907***	.289	-.324	-.370*	-.706***	-.299	.328	.361*
8 Gini Sky	-.290	-.749***	-.461**	.916***	.248	-.970***	.698***	.527**	.128	.051	-.589***	-.583***
9 Entropy Road	.938***	-.198	-.086	.114	.491**	-.115	.215	.221	.722***	.032	-.231	-.199
10 Entropy Building	.934***	-.231	-.057	.111	.512**	-.132	.204	.213	.737***	.072	-.211	-.192
11 Entropy Green	.900***	-.313	-.170	.227	.544**	-.224	.283	.292	.736***	-.001	-.301	-.282
12 Entropy Terrain	.918***	-.261	-.134	.162	.573***	-.169	.249	.259	.731***	.043	-.261	-.249
13 Entropy Sky	-.004	-.004	.037	-.129	.424*	.166	.025	.062	.659***	.022	-.063	-.046
14 HVI Road	-.241	.497**	.336	-.737***	-.530**	.809***	-.459**	-.335	-.322	-.140	.359*	.372*
15 HVI Building	.050	.086	-.771***	-.772***	.346*	-.918***	.780***	.663***	.307*	-.151	-.737***	-.694***
16 HVI Green	.309	-.607***	-.284	.474**	.474**	-.341*	.346*	.438**	.465**	.160	-.376*	-.451**
17 HVI Terrain	-.203	.612***	.317*	-.745***	-.647***	-.745***	.758***	.596***	-.352*	-.093	.644***	.614***
18 HVI Sky	.139	-.177	-.370*	.368*	.307	.339*	.350*	.887***	.358*	-.102	-.940***	-.853***
19 Income	-.232	.075	-.226	.309*	.028	.318*	.350*	.289	.289	-.196	-.953***	-.930***
20 Home Value	.418**	-.272	-.264	.370*	.592***	-.502**	.563***	.221	.444**	-.261	-.228	.235
21 Total Pop.	.223	-.067	.083	.127	.438**	-.419**	.277	.098	.706***	-.122	.260	.926***
22 Pop. Density	.140	-.023	.310	-.280	-.176	.194	-.430**	-.924***	-.368*	-.122	.990***	
23 Violent Crime	.113	-.015	.304	-.282	-.169	.190	-.434**	-.910***	-.387*	-.122		

* = $p < .05$, ** = $p < .01$, *** = $p < .001$

respectively), Simpson index and property crime ($r_s = 0.684$, $p < .001$ and $r_s = 0.324$, $p < .05$, respectively), and McIntosh index and property crime ($r_s = 0.641$, $p < .001$ and $r_s = 0.339$, $p < .05$, respectively).

For single-category indices, it became evident that positive correlations exist between violent crime and Gini index of green ($r_s = 0.751$, $p < .001$ for City 1 and $r_s = 0.328$, $p < .05$ for City 2, respectively), and between property crime and Gini index of green ($r_s = 0.707$, $p < .001$ and $r_s = 0.321$, $p < .05$, respectively).

Discussion

One difference between our approach and those of other researchers was to understand a built environment using micro-level analysis, as we tried to capture every single block of a neighborhood using the street-view images. This process allowed us to gain information from every corner of a neighborhood and then compute the geospatial diversity. Moreover, we used different indices to understand the same information, which was an attempt to overcome underlying limitations and biases that each index could have. We also sought to use single categories to understand how each category relates to social phenomena. This was an attempt to discern the effect of individual categories and their relationships with different social aspects. The results of this study suggested that multi-category geospatial indices are more effective at explaining social phenomenon than single categories.

In this section, the results from the previous section are discussed in the order of the research questions. We ran two separate correlation analyses to see whether the correlation between visual diversity indices and social phenomenon varies for the two cities. Evidence obtained from the analysis suggests that the correlations between visual indices and social phenomena are somewhat similar but still different regardless of the cities.

Relationship between geospatial visual diversity indices and population metrics

In both cities, there were negative correlations between the Simpson, McIntosh, and MCE indices and total population (see Table 3, rows 1-3). Similarly, as total population and population density are related to where people live, Day [62] found that most residents preferred to have their homes in a less visually diverse area. Collis, Felton, and Graham [63] reported comparable findings.

There was a negative correlation between the MCE index and population density per square mile (see Table 3, row 16). This finding concurs with previous research indicating that low visually diverse areas are typically less crowded and contain fewer buildings, greenery, and sky [64]. In other words, the suburbs are considered less visually diverse and downtown urban areas are more visually diverse [37, 65]. In all, two of three multi-category diversity indices, Simpson and McIntosh diversity indices, correlated with the total population, but only one multi-category diversity index, MCE, correlated with population density per square mile.

The results indicated that there were positive correlations between Entropy indices of road, building, green, terrain, and sky and the total population in the neighborhoods of both cities (see Table 3, rows 8-12). As the research by Zhang and Dong [9] noted, people prefer to live in areas with more greenery and visible terrain. In addition, this finding was consistent with Noland [66], who noted that an increase in population demands an increase in vehicles and roads.

Negative correlations exist between Gini index of building, green, and terrain and the total population (see Table 3, rows 4-6). For instance, this indicates that high building diversity is related to a higher population density. This finding is supported by Gillis [67] as well as Ellis and Ramankutty [68], with each study asserting that the more variety of buildings in a given area are related to higher population density.

In addition, there was a negative correlation between HVI index of sky and the total population (see Table 3, row 15). Higher HVI of sky indicates more openness.

Table 3 Summary of the correlation table for diversity indices and population metrics.[†]

Row	Index	Variable	Direction	Sig
1	Simpson	Total population	–	M(C1)*, M(C2)**
2	McIntosh		–	M(C1)**, M(C2)**
3	MCE		–	S(C1)**, S(C2)**
4	Gini building		–	S(C1)***, S(C2)**
5	Gini green		–	M(C1)***, S(C2)***
6	Gini terrain		–	S(C1)***, S(C2)***
7	Gini sky		+	M(C2)*
8	Entropy road		+	S(C1)***, M(C2)**
9	Entropy building		+	S(C1)***, S(C2)***
10	Entropy green		+	S(C1)***, S(C2)**
11	Entropy terrain		+	S(C1)***, S(C2)***
12	Entropy sky		+	S(C1)***, M(C2)**
13	HVI green		+	M(C1)*, M(C2)*
14	HVI terrain		+	M(C1)**, S(C2)***
15	HVI sky		–	M(C1)*, S(C2)**
16	MCE	Population density	–	M(C1)*, M(C2)**
17	Gini building		–	M(C1)**, M(C2)**
18	Gini green		–	M(C2)**
19	Gini terrain		–	S(C2)***
20	Entropy road		+	M(C2)*
21	Entropy building		+	M(C2)*
22	Entropy terrain		+	M(C2)*
23	HVI building		+	S(C1)***
24	HVI terrain		+	M(C2)**
25	HVI sky		–	M(C2)**

[†] (C1 = city 1, C2 = City 2, S = Strong, M = Moderate, * = $p < .05$, ** = $p < .01$, *** = $p < .001$)

This explains that fewer people tend to live in open-sky areas in outer-city neighborhoods. In general, inner-city areas are covered with high-rise buildings or concrete jungles with a high population where high-rise buildings prevent fully viewing the sky from the street level [69].

Relationship between geospatial visual diversity indices and economic indicators

In both cities, there were negative correlations between the Simpson index and the McIntosh index and household income and median home value (see Table 4, rows 1-2, 13-14). That is, families/individuals with more household income as well as a higher home value tend to reside in less visually diverse areas (e.g., living in the outer-city or suburb areas), instead of more visually diverse areas (e.g., inner-city or downtown areas). These findings are supported by Howe, Bier, Allor, Finnerty, and Green [70], who found that most people want to live outside inner-city areas due to lower taxes and less crime. Despite these findings, the HVI of green and sky showed positive correlations with median household income (see Table 4, rows 10, 12) and median home value (Table 4, rows 20, 22). These findings are consistent with Kim and Kim [71], who showed that families/individuals with higher income and home values were more likely to live in green and open areas compared to those with lower incomes and home values.

Table 4 Summary of the correlation table for diversity indices and economic metrics

Row	Index	Variable	Direction	Sig
1	Simpson	Income	–	S(C1)***, M(C2)*
2	McIntosh		–	S(C1)***, M(C2)*
3	MCE		–	M(C1)*
4	Gini building		–	M(C2)*
5	Gini green		–	S(C1)***, M(C2)**
6	Gini terrain		–	M(C2)**
7	Gini sky		+	S(C1)***
8	HVI road		–	M(C1)**
9	HVI building		–	S(C1)***, M(C2)*
10	HVI green	+	S(C1)***, M(C2)*	
11	HVI terrain	+	M(C1)*	
12	HVI sky	+	S(C1)***, M(C2)*	
13	Simpson	Home Value	–	S(C1)***, M(C2)*
14	McIntosh		–	S(C1)***, M(C2)*
15	MCE		–	M(C1)*
16	Gini green		–	S(C1)***
17	Gini terrain		–	M(C1)*
18	Gini sky		+	S(C1)**
19	HVI building		–	S(C1)***
20	HVI green		+	S(C1)***, M(C2)*
21	HVI terrain		+	M(C1)**
22	HVI sky		+	S(C1)***, M(C2)*

* = $p < .05$, ** = $p < .01$, *** = $p < .001$)

In both cities, there was a negative correlation between HVI of building and median household income (see Table 4, row 9). A study by Ghose [72] suggesting that those with higher income want to reside in outer-city (suburb) areas with fewer high-rise buildings concurs with this study's results.

Relationship between geospatial visual diversity indices and crime metrics

There were positive correlations between the Simpson and McIntosh indices and both violent and property crime (see Table 5, rows 1–2, 11–12). A previous study reported that high visual inequality and crime were correlated [73]. Likewise, Lentz [74] asserted that the type of environment and crime are related.

In terms of single-category geospatial diversity indices and crime metrics, the results here suggested that there were positive correlations between Gini index of green and both violent and property crime (see Table 5, rows 4, 14). Notably, a higher value of Gini green indicates heterogeneity of greenery. In other studies, however, Kuo [75, 76] and Kuo and Sullivan [77] did not support this assertion. Rather, Kuo [76] explained that green areas decrease crime since paved areas with no vegetation are often seen as “no man’s lands.” In general, empty or “no man’s land” areas have less presence of residents or witnesses and increase crime activities, making criminals feel that they are less likely to be caught.

Table 5 Summary of the correlation table for diversity indices and crime metrics

Row	Index	Variable	Direction	Sig
1	Simpson	Violent crime	+	S(C1)***, M(C2)*
2	McIntosh		+	S(C1)***, M(C2)*
3	MCE		+	M(C1)*
4	Gini green	Property crime	+	S(C1)***, M(C2)*
5	Gini sky		–	S(C1)***
6	HVI road		+	M(C1)*
7	HVI building		+	S(C1)***
8	HVI green		–	S(C1)***
9	HVI terrain		–	M(C1)*
10	HVI sky		+	S(C1)***
11	Simpson	Property crime	+	S(C1)***, M(C2)*
12	McIntosh		+	S(C1)***, M(C2)*
13	MCE		+	M(C1)*
14	Gini green		+	S(C1)***, M(C2)*
15	Gini terrain		+	M(C1)*
16	Gini sky		–	S(C1)***
17	HVI road		+	M(C1)*
18	HVI building		+	S(C1)***
19	HVI green		–	S(C1)***
20	HVI terrain		–	M(C1)**
21	HVI sky		+	S(C1)***

* = $p < .05$, ** = $p < .01$, *** = $p < .001$

Often, studies containing crime statistics characterized empty areas as crime hot spots [78].

Limitations and future directions

The computation of diversity and the relationship between diversity and social variables are based on the primary results in the selected metropolitan area in the Midwest. We recognize that specific relationships between visual diversity indices and social factors might become different due to the variation of geographical regions or even different countries. Despite this, we contend that this study presents a useful methodology and substantive results in computing and linking the diversities with social outcome, which can be leveraged by urban researchers, residents, the workforce (i.e., businesses), and administrators.

For the validity and reliability analyses, only five evaluators were asked to participate. This smaller sample size was bolstered by the number of pictures rated (i.e., ten) and the rating scale that was used (i.e., 5-point Likert scale). Overall, the Inter-Rater Reliability (IRR) component in this study was exploratory in nature, and a total of five evaluators was considered adequate. Future studies should include more evaluators in order to approximate the magnitude of the reliability coefficient with more precision. Second, for the validity and reliability analysis, GSV images ($n = 10$) were randomly chosen from approximately 53,000 images from an urban neighborhood. In addition, only lower and higher visually diverse images were selected. Future studies should consider selecting more images and from a range of neighborhoods with levels of visually diverse images in more than two categories.

Although a total of 351,246 street-view images were collected in this study, the neighborhoods sample remained limited. More samples of neighborhoods should be selected in future research to verify whether obtained results of this study have any biases related to the sample size. Future studies should also explore how relationships might change or remain the same across more states or countries. In the end, limited sample size can dictate the kinds of analyses conducted. For instance, this study relied primarily on correlational analyses. A larger sample size and variety of geolocations (i.e., cities and neighborhoods from different states or countries) could be used in future research to develop more complex statistical models.

Another limitation of this study is that the accuracy of the geospatial visual diversity indices depends on the precision of the deep learning model. While PSP-Net achieves about 82.2% accuracy with cityscape images [48], future AI studies could improve the deep learning model, which could also enhance this method in the accuracy and reliability of the diversities and the correlation coefficients.

Conclusion

The work here presents a computing framework of geospatial visual diversity based on AI-based tools from street-view images. It shows that diversity indices can be helpful in understanding the built and natural environment as well as the social dynamics of an urban neighborhood. Still, correlation analysis does not imply causality or inference. Nevertheless, the results presented in this study can be used to understand the influence of visual diversity for cities or neighborhoods. This study indicated that multiple category geospatial indices could be more effective in explaining social phenomena in urban neighborhoods than in single-category indices. This approach can potentially be used by city administrators, policymakers, and urban planners for their work in urban and community study and improvement.

We considered three aspects of the social phenomenon: the economy; population; and crime. We considered three social aspects together because earlier studies found they are related, so it was important for the researchers to consider them together. This study used street-view images to capture neighborhood scenes and the semantic segmentation method to extract visual objects information, enabling the computation of geospatial visual diversity. This was an attempt to incorporate computer programs to understand geospatial visual diversity and automate the computational process. This approach can be employed by city administrators, policymakers, and environmental designers to understand geospatial visual diversity without leaving their offices. Our approach could potentially save time and cost to aid in better understanding a built environment.

Acknowledgements This work was supported by the U.S. National Science Foundation under Grant 1739491. Md Amiruzzaman was also supported by West Chester University Faculty Startup Grant.

Data availability statement Data sharing is not applicable to this article as datasets used in this study came from Google Street-View (GSV) images and other publicly available datasets. Google does allow sharing of GSV images publicly, anyone can download the GSV images using GSV API (e.g., [https://maps.googleapis.com/maps/api/streetview?size=400x400 &location=47.5763831,-122.4211769&fov=80 &heading=70&pitch=0 &key=YOUR_API_KEY &signature=YOUR_SIGNATURE](https://maps.googleapis.com/maps/api/streetview?size=400x400&location=47.5763831,-122.4211769&fov=80&heading=70&pitch=0&key=YOUR_API_KEY&signature=YOUR_SIGNATURE)). Statistical data used in this study are available at <https://www.zillow.com/howto/api/APIOverview.htm>, and at <https://www.niche.com/about/data/>.

Declarations

Conflict of interest There are no conflicts of interest.

References

1. Amiruzzaman, M., Curtis, A., Zhao, Y., Jamonnak, S., & Ye, X. (2021). Classifying crime places by neighborhood visual appearance and police geonarratives: A machine learning approach. *Journal of Computational Social Science*, 4(2), 813–837. <https://doi.org/10.1007/s42001-021-00107-x>
2. Yatmo, Y. A. (2008). Street vendors as 'out of place' urban elements. *Journal of Urban Design*, 13(3), 387–402.
3. Sampson, R. J., & Raudenbush, S. W. (1999). Systematic social observation of public spaces: A new look at disorder in urban neighborhoods. *American Journal of Sociology*, 105(3), 603–651.

4. Skogan, W. G. (1992). *Disorder and Decline: Crime and the Spiral of Decay in American Neighborhoods*. Berkeley and Los Angeles: Univ of California Press.
5. Rundle, A. G., Bader, M. D., Richards, C. A., Neckerman, K. M., & Teitler, J. O. (2011). Using google street view to audit neighborhood environments. *American Journal of Preventive Medicine*, *40*(1), 94–100.
6. Qian, X., Lu, X., Han, J., Du, B., & Li, X. (2017). On combining social media and spatial technology for poi cognition and image localization. *Proceedings of the IEEE*, *105*(10), 1937–1952.
7. Wang, W., Yang, J., & You, X. (2018). Combining elasticfusion with pspnet for rgb-d based indoor semantic mapping. In: 2018 Chinese Automation Congress (CAC), pp. 2996–3001 . IEEE
8. Stamps, A. E., III. (2003). Advances in visual diversity and entropy. *Environment and Planning B*, *30*(3), 449–463.
9. Zhang, Y., & Dong, R. (2018). Impacts of street-visible greenery on housing prices: Evidence from a hedonic price model and a massive street view image dataset in beijing. *ISPRS International Journal of Geo-Information*, *7*(3), 104.
10. Shen, Q., Zeng, W., Ye, Y., Arisona, S. M., Schubiger, S., Burkhard, R., & Qu, H. (2017). Streetvizor: Visual exploration of human-scale urban forms based on street views. *IEEE Transactions on Visualization and Computer Graphics*, *24*(1), 1004–1013.
11. Lothian, A. (2017). *The Science of Scenery*. South Carolina: CreateSpace Independent Publishing Platform.
12. Amiruzzaman, M. (2021) Studying geospatial urban visual appearance and diversity to understand social phenomena. PhD thesis, Kent State University
13. Dronova, I. (2017). Environmental heterogeneity as a bridge between ecosystem service and visual quality objectives in management, planning and design. *Landscape and Urban Planning*, *163*, 90–106.
14. Nasar, J. L., & Hong, X. (1999). Visual preferences in urban signscapes. *Environment and Behavior*, *31*(5), 671–691.
15. Vanegas, C. A., Aliaga, D. G., Wonka, P., Müller, P., Waddell, P., & Watson, B. (2010). Modelling the appearance and behaviour of urban spaces. *Computer Graphics Forum*, *29*(1), 25–42. Wiley Online Library.
16. Gong, F.-Y., Zeng, Z.-C., Zhang, F., Li, X., Ng, E., & Norford, L. K. (2018). Mapping sky, tree, and building view factors of street canyons in a high-density urban environment. *Building and Environment*, *134*, 155–167.
17. Ye, Y., Zeng, W., Shen, Q., Zhang, X., & Lu, Y. (2019). The visual quality of streets: A human-centred continuous measurement based on machine learning algorithms and street view images. *Environment and Planning B*, *46*(8), 1439–1457.
18. Wen, D., Liu, M., & Yu, Z. (2022). Quantifying ecological landscape quality of urban street by open street view images: A case study of xiamen island, china. *Remote Sensing*, *14*(14), 3360. <https://doi.org/10.3390/rs14143360>
19. Patton, D.R. (1975). A diversity index for quantifying habitat “edge”. *Wildlife Society Bulletin* (1973-2006) *3*(4), 171–173
20. Junge, K. (1994). Diversity of ideas about diversity measurement. *Scandinavian Journal of Psychology*, *35*(1), 16–26.
21. Dissart, J. C. (2003). Regional economic diversity and regional economic stability: Research results and agenda. *International Regional Science Review*, *26*(4), 423–446.
22. Chapman, S. K., & Koch, G. W. (2007). What type of diversity yields synergy during mixed litter decomposition in a natural forest ecosystem? *Plant and Soil*, *299*(1), 153–162.
23. Bolnick, D. I., & Ballare, K. M. (2020). Resource diversity promotes among-individual diet variation, but not genomic diversity, in lake stickleback. *Ecology Letters*, *23*(3), 495–505.
24. Daly, A. J., Baetens, J. M., & De Baets, B. (2018). Ecological diversity: Measuring the unmeasurable. *Mathematics*, *6*(7), 119.
25. Orubebe, B.B. (2020). In: Yahyah, H., Ginzky, H., Kasimbazi, E., Kibugi, R., Ruppel, O.C. (eds.) *Soil Governance and Sustainable Land Use System in Nigeria: The Paradox of Inequalities, Natural Resource Conflict and Ecological Diversity in a Federal System*, pp. 157–180. Springer, Cham . https://doi.org/10.1007/978-3-030-36004-7_9.
26. Kelly, R. P., O’Donnell, J. L., Lowell, N. C., Shelton, A. O., Samhoury, J. F., Hennessey, S. M., et al. (2016). Genetic signatures of ecological diversity along an urbanization gradient. *Peer J*, *4*, 2444.
27. Yeh, C.-T., & Huang, S.-L. (2009). Investigating spatiotemporal patterns of landscape diversity in response to urbanization. *Landscape and Urban Planning*, *93*(3–4), 151–162.

28. de Jonge, D. (1986). On the appreciation of visual diversity in housing environments. *The Netherlands Journal of Housing and Environmental Research*, 1(4), 299–304.
29. Stamps, A. E., III. (2002). Entropy, visual diversity, and preference. *The Journal of General Psychology*, 129(3), 300–320.
30. Stamps III, A.E. (2004) Entropy and visual diversity in the environment. *Journal of Architectural and Planning Research*, 239–256
31. Stamps, A. E., III. (2012). A walk down the block: Spatial and temporal parameters of aesthetic judgments about ordinary streetscapes. *Perceptual and Motor Skills*, 114(2), 553–562.
32. Adams, D., Tiesdell, S., & White, J. T. (2013). Smart parcelization and place diversity: Reconciling real estate and urban design priorities. *Journal of Urban Design*, 18(4), 459–477.
33. Lu, Y. (2019). Using google street view to investigate the association between street greenery and physical activity. *Landscape and Urban Planning*, 191, 103435.
34. Li, X., Zhang, C., Li, W., Ricard, R., Meng, Q., & Zhang, W. (2015). Assessing street-level urban greenery using google street view and a modified green view index. *Urban Forestry & Urban Greening*, 14(3), 675–685.
35. Vandeviver, C. (2014). Applying google maps and google street view in criminological research. *Crime Science*, 3(1), 1–16.
36. Anguelov, D., Dulong, C., Filip, D., Frueh, C., Lafon, S., Lyon, R., Ogale, A., Vincent, L., & Weaver, J. (2010). Google street view: Capturing the world at street level. *Computer*, 43(6), 32–38.
37. Li, H., Peng, J., Yanxu, L., & Yi'na, H. (2017). Urbanization impact on landscape patterns in beijing city, china: A spatial heterogeneity perspective. *Ecological Indicators*, 82, 50–60.
38. Li, Y., Peng, L., Wu, C., & Zhang, J. (2022). Street view imagery (svi) in the built environment: A theoretical and systematic review. *Buildings*, 12(8), 1167.
39. Hipp, J. R., Lee, S., Ki, D., & Kim, J. H. (2022). Measuring the built environment with google street view and machine learning: Consequences for crime on street segments. *Journal of Quantitative Criminology*, 38(3), 537–565.
40. Gebru, T., Krause, J., Wang, Y., Chen, D., Deng, J., Aiden, E. L., & Fei-Fei, L. (2017). Using deep learning and google street view to estimate the demographic makeup of neighborhoods across the united states. *Proceedings of the National Academy of Sciences*, 114(50), 13108–13113.
41. Nguyen, Q. C., Huang, Y., Kumar, A., Duan, H., Keralis, J. M., Dwivedi, P., et al. (2020). Using 164 million google street view images to derive built environment predictors of covid-19 cases. *International Journal of Environmental Research and Public Health*, 17(17), 6359.
42. Gaspar, J., Fidalgo, B., Miller, D., Pinto, L., & Salas, R. (2010). Visibility analysis and visual diversity assessment in rural landscapes. In: Proceedings of the IUFRO Landscape Ecology Working Group International Conference, pp. 486–490 . Instituto Politécnico de Bragança, Bragança, Portugal
43. Boeing, G. (2019). Street network models and measures for every us city, county, urbanized area, census tract, and zillow-defined neighborhood. *Urban Science*, 3(1), 28.
44. FBI: Federal Bureau of Investigation Uniform Crime Report. <https://www.fbi.gov/services/cjis/ucr/>. Accessed: 28 Nov 2019 (2019)
45. Broxterman, D. A., & Kuang, C. (2019). A revealed preference index of urban amenities: Using travel demand as a proxy. *Journal of Regional Science*, 59(3), 508–537.
46. Perry, T. S. (2018). What's the best city for software engineers?: Hint: It's not san jose or san francisco-[spectral lines]. *IEEE Spectrum*, 55(8), 5–5.
47. Haklay, M., & Weber, P. (2008). Openstreetmap: User-generated street maps. *IEEE Pervasive Computing*, 7(4), 12–18.
48. Zhao, H., Shi, J., Qi, X., Wang, X., & Jia, J. (2017). Pyramid scene parsing network. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2881–2890
49. McDonald, D. G., & Dimmick, J. (2003). The conceptualization and measurement of diversity. *Communication Research*, 30(1), 60–79.
50. McIntosh, R. P. (1967). An index of diversity and the relation of certain concepts to diversity. *Ecology*, 48(3), 392–404. <https://doi.org/10.2307/1932674>.
51. Rex, M. A. (1973). Deep-sea species diversity: Decreased gastropod diversity at abyssal depths. *Science*, 181(4104), 1051–1053.

52. Verma, D., Jana, A., & Ramamritham, K. (2020). Predicting human perception of the urban environment in a spatiotemporal urban setting using locally acquired street view images and audio clips. *Building and Environment*, 186, 107340. <https://doi.org/10.1016/j.buildenv.2020.107340>
53. Wu, T.H., Zhao, Y., & Amiruzzaman, M. (2020). Interactive Visualization of AI-based Speech Recognition Texts. In: Turkay, C., & Vrotsou, K. (eds.) EuroVis Workshop on Visual Analytics (EuroVA). The Eurographics Association. <https://doi.org/10.2312/eurova.20201091>
54. Rajaram, R., Castellani, B., & Wilson, A. (2017). Advancing shannon entropy for measuring diversity in systems. *Complexity*, 2017, 8715605. <https://doi.org/10.1155/2017/8715605>
55. Li, X., Zhang, C., Li, W., Kuzovkina, Y. A., & Weiner, D. (2015). Who lives in greener neighborhoods? the distribution of street greenery and its association with residents' socioeconomic conditions in hartford, connecticut, usa. *Urban Forestry & Urban Greening*, 14(4), 751–759.
56. Heilig, G. K. (2006). Many Chinas? The economic diversity of China's provinces. *Population and Development Review*, 32(1), 147–161.
57. Liao, T. F. (2006). Measuring and analyzing class inequality with the gini index informed by model-based clustering. *Sociological Methodology*, 36(1), 201–224.
58. Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology*, 8(1), 23.
59. Shou, Y., Sellbom, M., & Chen, H.-F. (2022). 4.02 - fundamentals of measurement in clinical psychology. In: Asmundson, G.J.G. (ed.) *Comprehensive Clinical Psychology* (Second Edition), Second edition edn., pp. 13–35. Elsevier, Oxford . <https://doi.org/10.1016/B978-0-12-818697-8.00110-2>. <https://www.sciencedirect.com/science/article/pii/B9780128186978001102>
60. Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, 1(1), 77–89.
61. Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191.
62. Day, K. (2003). New urbanism and the challenges of designing for diversity. *Journal of Planning Education and research*, 23(1), 83–95.
63. Collis, C., Felton, E., & Graham, P. (2010). Beyond the inner city: Real and imagined places in creative place policy and practice. *The Information Society*, 26(2), 104–112.
64. MacLean, S., & Moore, D. (2014). 'hyped up': Assemblages of alcohol, excitement and violence for outer-suburban young adults in the inner-city at night. *International Journal of Drug Policy*, 25(3), 378–385.
65. Ley, D. (1986). Alternative explanations for inner-city gentrification: A Canadian assessment. *Annals of the Association of American Geographers*, 76(4), 521–535.
66. Noland, R. B. (2001). Relationships between highway capacity and induced vehicle travel. *Transportation Research Part A*, 35(1), 47–72.
67. Gillis, A. R. (1974). Population density and social pathology: The case of building type, social allowance and juvenile delinquency. *Social Forces*, 53(2), 306–314.
68. Ellis, E. C., & Ramankutty, N. (2008). Putting people in the map: Anthropogenic biomes of the world. *Frontiers in Ecology and the Environment*, 6(8), 439–447.
69. Asgarzadeh, M., Koga, T., Hirate, K., Farvid, M., & Lusk, A. (2014). Investigating oppressiveness and spaciousness in relation to building, trees, sky and ground surface: A study in Tokyo. *Landscape and Urban Planning*, 131, 36–41.
70. Howe, S. R., Bier, T., Allor, D., Finnerty, T., & Green, P. (1998). The shrinking central city amidst growing suburbs: Case studies of Ohio's inelastic cities. *Urban Geography*, 19(8), 714–734.
71. Kim, Y.-J., & Kim, E. J. (2020). Neighborhood greenery as a predictor of outdoor crimes between low and high-income neighborhoods. *International Journal of Environmental Research and Public Health*, 17(5), 1470.
72. Ghose, R. (2004). Big sky or big sprawl? Rural gentrification and the changing cultural landscape of Missoula, Montana. *Urban Geography*, 25(6), 528–549.
73. Kelly, M. (2000). Inequality and crime. *Review of Economics and Statistics*, 82(4), 530–539.
74. Lentz, T. S. (2018). Crime diversity: Reexamining crime richness across spatial scales. *Journal of Contemporary Criminal Justice*, 34(3), 312–335.
75. Kuo, F. E., & Sullivan, W. C. (2001). Aggression and violence in the inner city: Effects of environment via mental fatigue. *Environment and Behavior*, 33(4), 543–571.

76. Kuo, F. E. (2003). Social aspects of urban forestry: The role of arboriculture in a healthy social ecology. *Journal of Arboriculture*, 29(3), 148–155.
77. Kuo, F. E. (2001). Coping with poverty: Impacts of environment and attention in the inner city. *Environment and Behavior*, 33(1), 5–34.
78. Loukaitou-Sideris, A. (1999). Hot spots of bus stop crime: The importance of environmental attributes. *Journal of the American Planning association*, 65(4), 395–411.

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.