Accident Analysis and Prevention 146 (2020) 105716

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

An examination of the intersection environment associated with perceived crash risk among school-aged children: using street-level imagery and computer vision

Jae-Hong Kwon[a], Gi-Hyoug Cho[b,*]

[a] School of Urban and Environmental Engineering, Ulsan National Institute of Science and Technology, South Korea
[b] School of Urban and Environmental Engineering, Ulsan National Institute of Science and Technology, 50 UNIST-gil, Uljugun, Ulsan, 44949, South Korea

A B S T R A C T

While computer vision techniques and big data of street-level imagery are getting increasing attention, a "black-box" model of deep learning hinders the active application of these techniques to the field of traffic safety research. To address this issue, we presented a semantic scene labeling approach that leverages wide-coverage street-level imagery for the purpose of exploring the association between built environment characteristics and perceived crash risk at 533 intersections. The environmental attributes were measured at eye-level using scene segmentation and object detection algorithms, and they were classified as one of four intersection typologies using the k-means clustering method. Data on perceived crash risk were collected from a questionnaire conducted on 799 children 10 to 12 years old. Our results showed that environmental features derived from deep learning algorithms were significantly associated with perceived crash risk among school-aged children. The results have revealed that some of the intersection characteristics including the proportional area of sky and roadway were significantly associated with the perceived crash risk among school-aged children. In particular, road width had dominant influence on risk perception. The findings provide information useful to providing appropriate and proactive interventions that may reduce the risk of crashes at intersections.

1. Introduction

Computer vision is a field of study in which researchers develop techniques to help computers understand the content of digital images such as photographs and videos (Brownlee, 2019). Computer vision algorithms and applications have been adopted in various fields, such as image recognition (Farabet et al. 2012; Tompson et al. 2014) and object detection (Girshick et al. 2014; Sermanet et al., 2013). One of the elements essential to developing autonomous driving, for instance, is computer vision (Pohlen et al. 2017).

Meanwhile, several studies have applied computer vision techniques to an examination of traffic safety. Dubey et al. (2016) and Ordonez and Berg (2014) predicted the perceived safety of streetscapes using convolution neural network (CNN) and support vector machine (SVM) classifier trained on street-level imagery in order to visualize the spatial distribution of perceived safety at a city scale. Yet, the use of current computer vision methods has a clear limitation to investigating the complicated relationship between the built environment and perceived safety (Dubey et al., 2016). This is largely because the computer vision model makes it difficult to find the safety determinants of the built environment's characteristics. The reliance upon "black-box models," in which the rationale for the generated outputs is inscrutable (Cabitza et al. 2017), is another factor that reduces researchers' ability to understand the relationship between the characteristics of the built environment and perceptual outcomes. Opaque models like CNN, therefore, cannot serve as a theoretical tool for understanding why a neural network operates effectively (Alemi et al. 2016). Therefore, the black-box nature of neural networks is a barrier to their adoption in research applications where interpretability is paramount, like traffic safety (Shrikumar et al. 2017).

One of the ways to overcome the limitations of current computer vision methods is to use a semantic scene labeling method. Recently, there has been growing interest in using semantic scene labeling approach in the computer's vision (L.-C. Chen et al., 2014; Long et al., 2015; Noh et al., 2015; Shrikumar et al., 2017). In contrast to the conventional scene recognition approach, which aims to determine the overall scene category by placing emphasis on understanding its global properties (Zhou et al. 2014), the scene labeling approach seeks to identify the individual components of a whole scene as well as the complex relationships among the semantic entities usually found in street scenes, such as pedestrians, cars, road, or sidewalks (Cordts et al., 2016).

Furthermore, applying the scene labeling approach to street-level imagery allows for an analysis of the built environment characteristics measured at the eye-level. Eye-level measurements may more accurately reflect a person's actual perception of the attributes of a

* Corresponding author.
E-mail addresses: kjh623@unist.ac.kr (J.-H. Kwon), gicho@unist.ac.kr (G.-H. Cho).

https://doi.org/10.1016/j.aap.2020.105716
Received 31 October 2019; Received in revised form 31 July 2020; Accepted 3 August 2020
Available online 19 August 2020
0001-4575/

particular built environment (Jiang et al., 2015). Given the difficulty of scalably measuring the built environment at the street level, however, few studies on perceived safety have accounted for eye-level urban characteristics. Combining street-level imagery and the scene labeling approach may overcome this limitation by acquiring data on what people typically experience and perceive at the eye-level on the ground (Li et al., 2015).

The ability of scene labeling methods to analyze the urban built environment at the eye-level also provides the advantage of measuring the visibility of pedestrians and drivers, which is one of the important factors in urban traffic safety (Dadashova et al., 2016; Hills, 1980; Retting et al., 2003; Schofer et al., 1995). For instance, Meir et al. (2013) suggested that pedestrians' traffic crashes tend to take place at intersections or curves where visibility is obstructed by the road's curvature, fixed objects, parked vehicles, darkness, or other visibility limitations because these visibility obstructions interfere with the pedestrian's ability to detect oncoming vehicles while also obstructing the motor vehicle drivers' vision, preventing drivers from noticing pedestrians, who may be masked by obstructions (Aoki & Moore, 1996; Petch & Henson, 2000). Thus, improving pedestrian and driver visibility at intersections is essential both for enhancing perceptions of safety and reducing potential crash risks (Lee et al., 2016).

The present study examined the association between urban environmental attributes derived from scene labeling methods (i.e., scene segmentation and object detection algorithms) and school-aged children's perceptions of safety, the latter of which provides information on the potential risks of pedestrian crashes (Lee et al., 2016; Meir et al. 2015). First, we sought to overcome the technical limitations of current computer vision methods by applying the scene labeling approach, which allows us to analyze complex relationships among the urban characteristics of built environments. A second aim of this study was to inform public policy for reducing the crash risk of children in school zones through the use of this novel method for deriving accurate measures of intersection characteristics at the eye-level, which offers a more precise understanding of perceived safety.

2. Related works

2.1. Measuring urban attributes from street-view imagery

Audit instruments are commonly used to collect data on the built environment characteristics, including sidewalks conditions, bicycling facilities, building characteristics, and traffic conditions (Boarnet et al., 2006; Brownson et al., 2004; Cunningham et al., 2005; Pikora et al., 2003; Suminski et al., 2008). Since collecting data using audit tools typically requires in-person observations (Ewing, R et al., 2006), it is labor intensive and time consuming compared to using other types of GIS, aerial photos, or videotape data (Brownson et al., 2009). Therefore, in-person field audits can be logistically challenging when the study covers large or geographically dispersed areas or when data collection at multiple time points is necessary to supporting a longitudinal comparison (Wilson et al., 2012).

New technologies using omnidirectional imagery, such as GSV, offer a viable alternative to field audits since they expand the geographic and temporal scope of vision (Kelly et al., 2012). Street-view imagery can reduce the resources required for conducting audits because they are collected by virtually "driving" through a community. Most of this collected imagery is freely accessible in more than 20 countries (e.g., in Australia, Asia, Europe, North America, and Central America; Anguelov et al., 2010). Given these advantages, many urban studies researchers have used street-view imagery to audit such environmental attributes as public open space (Edwards et al., 2013; Taylor et al., 2011), neighborhood environment (Griew et al., 2013; Li et al., 2015; Odgers et al., 2012), and built environment (Kelly et al., 2012; Wilson et al., 2012).

Several studies have consistently reported high levels of agreement between street-view imagery and in-person measures (Badland et al., 2010; Clarke et al., 2010; Rundle et al., 2011). Kelly et al. (2012) suggested that street-view imagery audits have high inter-rater reliability for most items, including land use, public transportation, and street characteristics, but that characteristics with relatively lower reliability (e.g., on-street parking, tree shade on streets, sidewalk width, and curb cuts) should be included in datasets with caution. Several other studies have shown that audits using street-view images were effective and reliable in the place of field audits (Griew et al., 2013; Li et al., 2015; Wilson et al., 2012).

2.2. Applying computer vision algorithms to urban images

When the street-view imagery is combined with computer vision techniques, a large-scale automatic evaluation of various urban environment is possible (Liu et al., 2017). Li et al. (2017a,b) noted that machine learning techniques make it possible to extract additional geospatial information and more complex metrics from big image datasets. Previous studies on the automatic classification of architectural styles have demonstrated the possibility of using street-view images and computer vision algorithms to capture the urban characteristic of built environments. For instance, Doersch et al. (2013), Doersch et al. (2012) and Doersch et al. (2015) discovered geographically distinguishable patches between different cities from GSV images by employing a discriminative clustering algorithm. That method requires representing the patches by using a HOG + color descriptor (Dalal & Triggs, 2005) and an iterative training linear support vector machine (SVM) detector (Singh et al., 2012). Other work has parsed scene images into constituent objects such as windows, domes, doors, walls, roofs, arches, and balconies (Berg et al., 2007; Goel et al., 2012; Shalunts et al., 2011; Simon et al., 2012; Weissenberg et al., 2013).

The relevant research to the present study are those that focus on understanding how humans perceive urban scenes (Dubey et al., 2016; Naik et al., 2014; Ordonez & Berg, 2014; Porzi et al., 2015). Most of these studies have aimed to provide reliable methods for completing technical tasks like automatic scene classification and recognition, which are usually analyzed using street-view images and crowd-sourced data collected from volunteer participants. For instance, previous studies have employed such methods as SVM and CNN trained using street-view images and crowd-sourced safety perception scores of urban images such as Place Pulse 1.0 (Salesses et al., 2013). The algorithms employed in that research can accurately predict the safety score of streetscapes for which no evaluative data are available. Naik et al. (2014) and Ordonez and Berg (2014) validated such a prediction model by computing the correlations between predicted safety scores and local income and crime statistics. Their study proposed a tentative boundary for the validity of their predictors.

2.3. Risk perception

In the field of traffic safety, risk perception has been used to examine associations between the behaviors of road users and perceptions of safety (Diógenes et al., 2010). The major advantage of relating risk perception to traffic safety is that doing so provides information on potentially hazardous situations, which makes it possible to assess road safety even when no crashes have yet occurred (Diógenes et al., 2010; Meir et al., 2015). Studies on pedestrian crash risk using only recorded crash data commonly have difficulty assessing road safety because reliable data on pedestrian crashes are relatively scarce. In such cases, perception information can be used as an indicator of potential crash risks at locations where actual crashes have not occurred (Schneider et al., 2004). Cho et al. (2009), meanwhile, argued that while police reported crash risk data do not provide complete information on traffic safety, individuals' perceptions of crash risk provide additional information useful for proactively identifying potential risk locations or risk factors.

Understanding the ways that pedestrians perceive crash risk could also provide information about pedestrian behavior that could be used to guide the development of proactive countermeasures to reduce the number and severity of road crashes (Diógenes et al., 2010; Lee et al., 2016; Oltedal et al., 2004). Individuals adapt their behavior and make decisions that help them to avoid unwanted outcomes based on their perceptions of risk (Sjöberg et al., 2004). In traffic safety situations, people are more likely to change their attitudes and behaviors to avoid exposure to areas with a high potential risk when they perceive themselves as vulnerable to a crash risk (Rundmo, 1999; Ulleberg & Rundmo, 2003). Therefore, precautions or proactive countermeasures to prevent future pedestrian crashes are desirable when the perception of a crash risk is high (Lee et al., 2016; Schneider et al., 2004).

Critiques of using risk perception techniques to measure traffic safety are related to the fact that people can both underestimate and overestimate traffic dangers since risk perception is a subjective assessment of the probability of an unwanted event occurring (Sjöberg et al., 2004). In response to these criticisms of risk perception, the literature examining the extent to which the public perceives traffic risk correctly has shown a positive correlation between statistically estimated risk and perceived risk with different modes of transport, finding that the public's risk perception is broadly correct (Elvik & Bjørnskau, 2005), Rafaely et al. (2006) examined differences in the perception of traffic risk between older and younger adults and found that individual younger and older adults provided relatively accurate average risk estimates for their respective groups.

Used perceived crash risk among school-aged children as an indicator of potential crash risk, we hypothesized that perceived crash risk can be used to inform proactive interventions that are likely to improve road safety.

3. Methodology

3.1. Study area

The study site for this research was the city of Ulsan, one of the largest industrial complexes in Korea. We chose eight elementary school zones in the Dong district, where the number of police-reported pedestrian crashes between 2010 and 2014 was 792 (the highest level of crash occurrences in Ulsan; Fig. 1). The land use within the selected study sites was mainly residential, but mixed commercial and industrial areas are included. The number of enrolled students during the 2015 survey period was 4,130 (Lee et al., 2016).

3.2. Outcome: perceived safety

To measure perceived safety within the school zone areas, we conducted a risk-perception survey in July of 2015. The participants in the survey were 10-to-12 year-old students enrolled in eight elementary schools. The number of respondents was originally 1,263. The survey questionnaire gathered data on demographic features, the mode of travel to school, and the children's perceived crash risk along their walking routes to school. Among 1,263 students, 1,042 students completed all of the response on the questionnaire. Excluding the 243 respondents who used motorized travel for commuting, 799 students who commuted to school by walking were included as the final participants.

The survey participants reported perceiving 2,621 locations in the study area as having a high crash risk. Since our research was focused on intersections, the reported locations needed to be spatially matched to a corresponding intersection. A road intersection was defined as having a 15 m (50 ft) buffer zone from the crossing point of two road centerlines (Schneider et al., 2010). The number of risk points located at the 546 locations that were intersections was used as the crash risk score for each intersection (for more information about data collection, see Lee et al., 2016).

Fig. 1. The selected study sites in Ulsan, Korea

3.3. Control variable: risk exposure

Risk exposure is one of the most important factors to consider when determining the number of pedestrian crashes, thus it needs to be accurately accounted for when analyzing risk factors to pedestrian safety. In particular, elementary school students are likely to evaluate a specific intersection along their regular walking routes to school as places of high crash risk. Although it is desirable to use pedestrian counts for each intersection to create a measure of risk exposure (Clifton & Kreamer-Fults, 2007), count data for the study areas were not available, and taking manual counts at each intersection was not feasible given the number of intersections included. Consistent with the methodology used in a previous study (Lee et al., 2016), the present work used a proxy measure to represent risk exposure, and the number of students crossing at each intersection was derived from the survey data. Participants were asked to represent their regular walking routes to school on a local map. Based on that information, we counted the number of individual crossings at each intersection chosen for the route and used that number as one of the exposure variables. To calculate population and street density, population and street centerlines in the GIS from the Ulsan Statistical Information Service (2010) and Statistical Geographic Information Service (SGIS) in Korea were used.

3.4. Independent variables

3.4.1. Environmental attributes

Environmental attributes were derived from street-view imagery (map.naver.com) of each intersection using scene labeling algorithms that included scene segmentations and object detection. Fig. 2 shows exemplary images of the two algorithms. We used a pre-trained scene segmentation model (Xiao et al., 2018; Zhou et al., 2016, 2017) and object detection model (Redmon et al., 2016; Redmon & Farhadi, 2017). The variables derived from the scene segmentation model included area proportions of the sky, green space, road, and sidewalk. The area proportion of each environmental element in the image was calculated by counting the number of pixels in the segmentation mask (Wang et al., 2019; Yao et al., 2019; Zhang et al., 2018), a variable that Zhang et al. (2018) showed is significantly associated with various perceived qualities of urban environments. The variables measured by the object detection model were number of pedestrians, and number of vehicles.

Similar to GSV imagery, the street-view imagery we used, taken since 2007, provided real street images (Park et al., 2012). The street view imagery used in the present paper was taken in October 2017, which was the latest version of imagery we could download in July 2018. To download the targeted street-view imagery, we computed the latitudes and longitudes of the 546 intersections examined using ArcGIS 10.4. Based on the GPS coordinates of each intersection, we captured the 2,184 street-view images in four directions at 546 intersections. The pitch parameter (i.e., the up and down angle of the camera) and the FOV parameter (horizontal field view of the image) were set to 0 and 100. Among the 2,184 images, we manually excluded indoor images. Finally, 2,132 images at 533 intersections were used for the next step of the analysis.

The engineering modifications used to reduce pedestrian-vehicle crashes can be classified into three approaches: improvement to visibility, managing vehicle speed, and the separation of pedestrians from vehicles (Retting et al., 2003). Similarly, three types of environmental attributes were defined: 1) visibility of pedestrian/drivers; 2) traffic flow; and 3) separation of pedestrians from vehicles.

The visibility of pedestrians and drivers was measured using the proportional area of sky and green. A higher proportion of sky indicates an increase in the FOV of pedestrians and drivers, while a higher proportion of green tends to decrease FOV (Fig.2). From the street-view perspective, a small proportion of sky and a large proportion of green frequently limits the sight distance of drivers and pedestrians. To quantify the traffic flow at the intersections, we used the proportional area of the road derived from the scene segmentation model and the number of pedestrians and vehicles identified from the object detection model. The previous literature reported that the width of lanes (Noland & Oh, 2004; Ukkusuri et al., 2012) and higher traffic volumes (Chen et al., 2009; Harwood et al., 2008; Ma et al., 2010) are associated with a greater crash risk. Conversely, the separation of pedestrians from vehicles (Berhanu, 2004) can reduce vehicle-pedestrian crash rates. The separation of pedestrians from vehicles was represented using the proportional area of sidewalks, which were derived from the scene segmentation model.

3.4.2. Intersection Typology

Measurement of the built environment can produce various indicators, each representing independent characteristics of the built environment. In reality, however, these characteristics are not actually independent but are strongly associated each other. For instance, an intersection with more openness or a higher proportion of sky is likely to have a lower building density, and average building height is generally higher where roads are wider. Because of the strong correlation between various the dimensions of built environment, many urban spaces can be classified as a specific type of urban forms based on their similarities (Song & Knaap, 2007). In addition, we assumed that human perception of built environments tends to be more comprehensive or holistic than analytical. In other words, a typical street view can be classified into several groups, and children's perceptions of unique road environments may have typological patterns related to the environments' features.

To understand distinguishable variations in the built environment between groups, we used an empirical cluster analysis, which is a method of combining observations into groups based on their similarities within a set of predetermined characteristics (Song & Knaap, 2007; Wilks, 2011). In our study, a K-means cluster analysis was used to classify all 533 intersections into different types on the basis of their similarities in the values of the five factors derived from our scene segmentation algorithm: the proportional area of 1) building, 2) sky, 3) green space, 4) road, and 5) sidewalk. We used a scree plot by searching for a kink in the curve generated from the within sum of squares (WSS) and its logarithm [log(WSS)], and a η^2 coefficient, which is similar to R^2, and a proportional reduction of the err (PRE) coefficient to choose the optimal number of clusters for the final model (Schwarz, 2008). The best clustering solution, based on the interpretability of the results and associated cluster statistics, was found to be a four-cluster solution. STATA software (version 13.0) was used for the cluster analysis. (Fig. 3)

The values of the cluster centroids for each of the four intersection types are presented in Fig. 4 and Table 1.. The centroid of a cluster are representative of a typical set of intersection features (Song & Knaap, 2007), indicating the characteristics of each intersection type. Using the centroid values, each cluster was defined as four types of intersections: wide road with the presence of sidewalks (type 1), wide road with roadside trees (type 2), wide road with clear visibility (type 3), and narrow alley with dense buildings (type 4).

3.5. Statistical analysis

Two negative binomial (NB) regression analyses were conducted to examine the effects of environment attributes and their composition on perceived safety. The standard NB model, as compared to the Poisson regression model, is the most widely applied in traffic safety studies because it can account for the over-dispersion and between-location heterogeneity of crash variations (Bowman et al., 1995; Ukkusuri et al., 2012). The standard NB models have the following probabilistic structure: perceived crash risk Y_i, at an intersection i, when conditional on its mean θ_i is Poisson distributed and independent over all intersections.

$$Y_i \mid \theta_i \sim Poisson(\theta_i)$$

The mean of the Poisson is represented as,

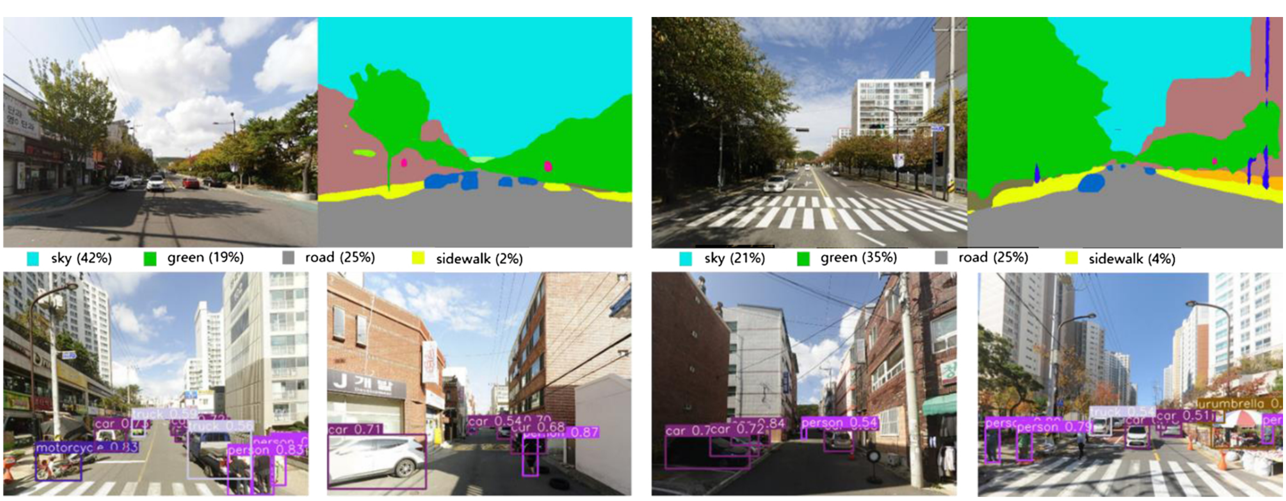

Fig. 2. Examples of scene segmentation algorithm (top), and object detection (bottom)

$\theta_i = f(X_i;\beta)\exp(\varepsilon_i)$

where the $\exp(\varepsilon_i)$ is the multiplicative random effect of the model, following a Gamma distribution, $f(X_i;\beta)$ is a function of the covariates, X_i is a vector of the environmental attributes, and β is the vector of unknown coefficients (Zou et al., 2015). We created two models for defining the environmental attributes. The first investigated associations between independent built environment features and perceived crash risk; the second focused on associations between intersection typology and perceived crash risk. STATA software (version 13.0) was used to estimate the NB regression models.

4. Results

4.1. Model performance

To test the validity of the applied computer vision algorithms, we estimated model performance using a random sample of 20% of the street-view imagery (434/2132) as our validation set. For a quantitative evaluation of the algorithm, we compared the ground truth with the results obtained from the scene segmentation algorithm and then calculated the resulting score. The model performance was evaluated using the following equations,

$$\text{Accuracy}=\frac{TP+TN}{TP+TN+FP+FN},$$

$$\text{Precision}=\frac{TP}{TP+FP},$$

$$\text{Recall}=\frac{TP}{TP+FN},$$

Where TP, TN, FP and FN stand for the number of true positive, true negative, false positive and false negative pixels in a confusion matrix. Precision indicates the probability that a selected positive pixel is correct, and recall refers to the percentage of relevant pixel that were selected. These two measures for information retrieval effectiveness are most frequent and basic approaches (Manning et al., 2010). To test the validity of our object detection model, we compared the ground truth and environmental attributes detected by the object detection algorithm and then calculated the average precision (AP) and mean average precision (mAP), for a quantitative evaluation of the algorithm. The AP and mAP for a object class can be obtained based on precision and recall domains, defined as follows,

$$\text{AP}=\frac{1}{11}\sum\nolimits_{p\in MP}(P),$$

$$\text{mAP}=\frac{\sum_{c\in S} AP_c}{|S|},$$

where MP is the maximum precision value from 0 to 1 step by 0.1 in the recall domain, and AP is the means of MaxPrecision in the 11 points. The mAP denotes the means of all object classes'AP, where S is a class set. In addition, 11-point metric and IoU $>\ =0.5$ are applied in object detection algorithm based on the standard VOC2007 (Everingham et al., 2010). The ground truth was manually annotated using Labelme (github.com/wkentaro/labelme), a polygonal image annotation tool written in Python. Fig. 5 shows an example of a ground truth image generated by the Labelme software.

Table 2 shows the estimated performance of the two algorithms. The results indicate that the scene segmentation model performed well for classifying variables that occupy a large proportion of area in the street-view imagery [e.g., road (precision = 86.26%, recall = 97.14%), and sky (precision = 96.19%, recall = 98.25%)], but that it tends to

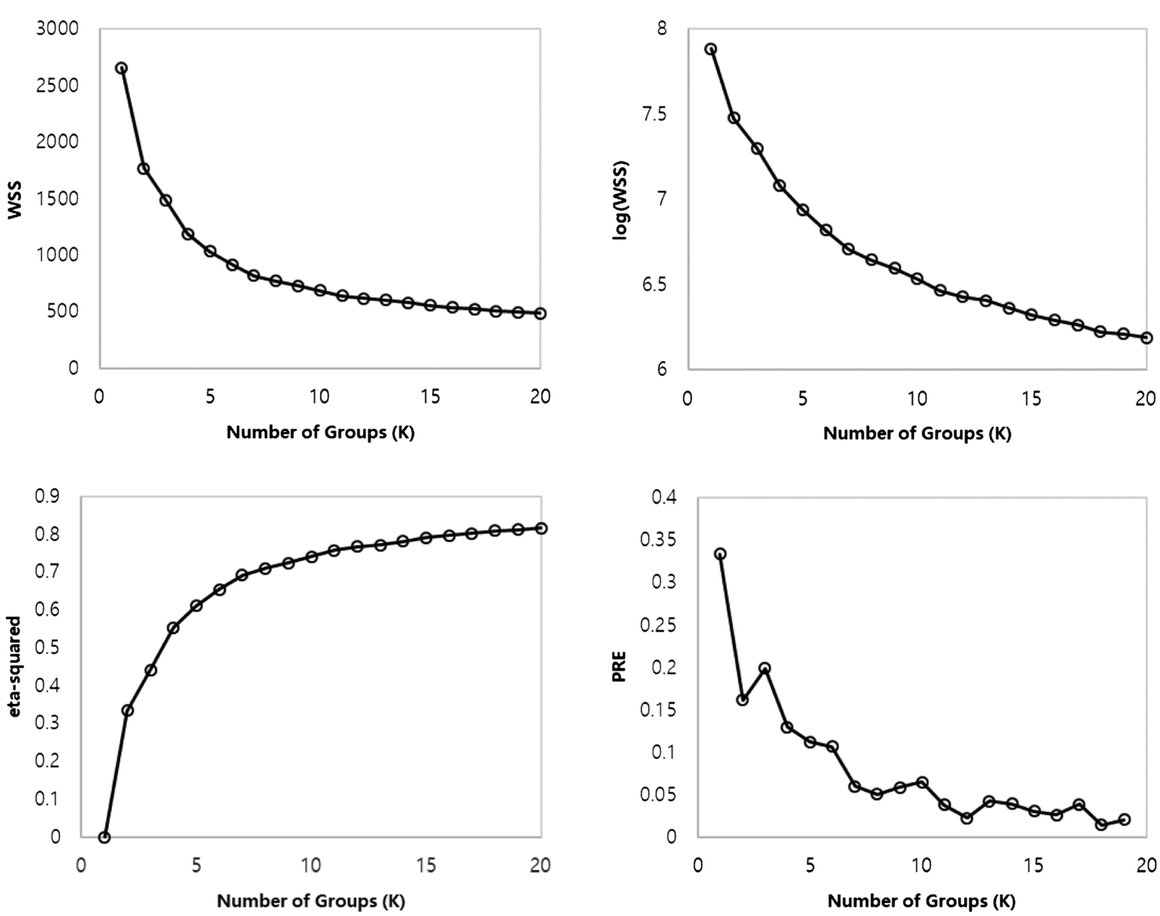

Fig. 3. The criterions for detecting the optimal number of clusters

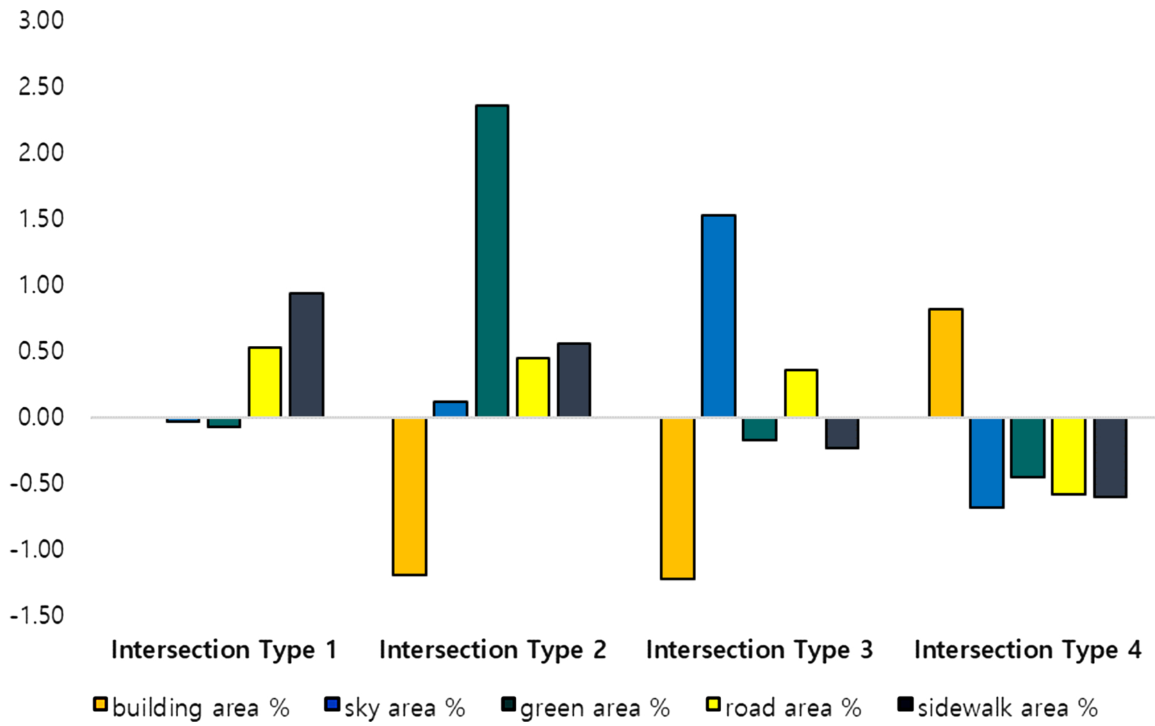

Fig. 4. Cluster centroid values for each of the intersection type

Table 1
Comparing intersections in five dimensions based on cluster centroid values.

Dimensions	Type1	Type2	Type3	Type4
Proportional area of building	0	−1.191	−1.215	0.8106
Proportional area of sky	−0.032	0.117	1.529	−0.679
Proportional area of green	−0.069	2.355	−0.173	−0.445
Proportional area of roadway	0.528	0.454	0.366	−0.581
Proportional area of sidewalk	0.944	0.563	−0.224	−0.594
Counts	139	56	103	235

produce relatively low performance for classifying sidewalk (precision = 46.96%, recall = 61.15%) and green area (precision = 90.25%, recall = 72.21%). The average proportion of green and sidewalk in the street-view images was 5% and 3%, respectively. The object detection model had a moderate level of performance at detecting pedestrians (AP = 53.8%), vehicles (AP = 65.5%) with 59.6% mAP. Although the overall accuracy is lower than the results of the best exemplary study (Ju et al., 2019), it is an acceptable level of accuracy when we consider relatively low resolution of street view images and different context of study sites from original training dataset (Redmon & Farhadi, 2018; Zhang et al., 2016).

4.2. Descriptive analysis

Table 3 provides descriptive statistics for the variables used in our analysis. On average, the perceived crash risk in an intersection was 3.09 (*SD* = 6.35). As Fig. 6 shows, each type of intersection was evenly distributed in study site. The mean value for daily crossings at the intersection was 4.75 (*SD* = 9.08). The maximum number of student crossing was 66, and 215 intersections were not selected for the routes to school. Not surprisingly, perceived crash risk and the number of crossings at an intersection were strongly correlated ($\rho = 0.836$). The mean values of the proportional area of sky, green, road, and sidewalk at each intersection was 28% (*SD* = 11%), 5% (*SD* = 5%), 14% (*SD* = 5%), and 3% (*SD* = 2%), respectively. Fig. 7 shows the intersections where the street-view imagery had the mean value of each of the variables at the study sites.

4.3. Intersection characteristics associated with perceived safety

Table 4 presents the parameter estimates and p values of the two NB models for perceived crash risk. The over-dispersion parameter estimation α was significantly higher than zero in both NB model 1 (95% CI (0.88, 1.37)) and NB model 2 (95% CI (0.89, 1.38)). In both models, all variables representing risk exposure had a significantly positive association with perceived crash risk. A higher number of students crossing, a higher population density, and a higher street density were all positively related to a higher perceived crash risk.

In NB model 1, the proportional area of sky significantly decreased perceived crash risk ($p = 0.024$). A higher proportion of sky indicates more visual “openness” on the street and more visible area for drivers and pedestrians. For that reason, the negative relationship between proportion of sky and perceived crash risk implies that the improvement in visibility at an intersection could be one of the factors that reduces perceived crash risk. As Lee et al. (2016) found, although the direct relationship between lower visibility and actual crash risk remains unclear, improving pedestrian visibility at an intersection apparently enhances feelings of safety and could change pedestrian behavior.

Meanwhile, the results of the NB model 1 showed that the proportional area of road has a significantly positive relationship with perceived crash risk ($p = 0.051$). The literature has consistently reported that the width of lanes and roads is one of the important factors that increases crash risk (Lee et al., 2016; Ukkusuri et al., 2012). Indeed, wider lanes mean higher vehicle speeds, which implies a greater likelihood of crashes involving pedestrians as well as more serious pedestrian injuries (Aarts & Van Schagen, 2006). Number of pedestrians was positively associated with the level of perceived crash risk ($p = 0.014$), perhaps because more pedestrians at the intersections hinders the ability of school-aged children to cross. The number of pedestrian variable may indicate crowdedness at the intersection, but it is worth noting that the number of pedestrians was derived from street-view images taken at a specific time.

The NB model 2 showed influence of street type of perceived crash risk. The students tended to perceive a lower risk at the intersections with wide roads and clear visibility (type 3) than at intersections with

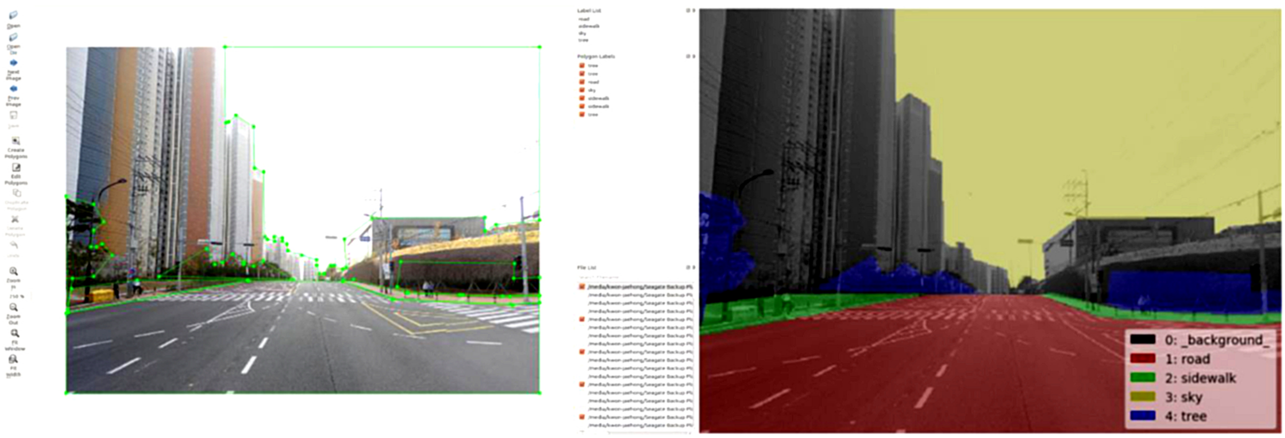

Fig. 5. A ground truth image (right) generated by Labelme software (left)

Table 2
Model performance in scene segmentation and object detection methods

Variables	Accuracy	Precision	Recall	mAP	AP
Overall	98.95%	89.87%	93.84%	-	
Road	97.48%	86.26%	97.14%		
Sky	98.39%	96.19%	98.25%		
Sidewalk	97.49%	46.96%	61.15%		
Green	98.43%	90.25%	72.21%		
Pedestrian	-			59.64%	53.76%
Vehicle					65.53%

* IoU > = 0.5

wide road and sidewalks (type 1; $p =$ 0.048). As shown in Fig. 8 and Fig. 10, a clear distinction between the two types of intersections was their releative visual openness (e.g., proportional area of sky) and the presence of sidewalks. It seems manifest that increasing the visbility of pedestrians could reduce perceived crash risk. The risk perception at the reference intersections was not statistically different from that at the intersections with wide roads and large roadside trees. Roadside trees may enhance the aesthetics of the street view (Fig. 9), but they do not alter the level of perceived safey. Indeed, in the traffic safey context, roadside trees are a visual and physical barriers (Budzynski et al., 2016). A comparison between intersection type 1 and 4 (Fig. 11) showed interesting results. School-aged children tended to report lower levels of perceived crash risk in Type 4 intersections(narrow alleys and dense development) than the reference intersenctions ($p =$ 0.086). The presence of sidewalks and visual openness may reduce perceived crash risk, but the influence of wide roads on perceived crash risk may override the influence of these factors.

5. Discussion and Conclusion

In this study, we applied a scene labeling technique for measuring built environment attributes at the eye-level. While investigating associations between urban environmental elements and perceived safety, we attempted to overcome the technical limitations of current computer vision techniques. Despite the success of deep neural networks (DNNs) in the field of computer vision (Krizhevsky et al., 2012; Roth et al., 2016; Szegedy et al., 2015), no comprehensive theoretical understanding of optimization processes or their inner organization (Shwartz-Ziv & Tishby, 2017) has yet been provided. Thus, neural networks are often criticized for being used as mysterious “Black Boxes” (Alain & Bengio, 2016). To address this issue, the current study presented processes for analyzing the relationships between urban environmental elements derived from deep learning algorithms and perceived safety.

Measuring the design characteristics of the built environment at the eye-level provided a better understanding of perceived safety. People rarely experience landscapes from a high altitude (like remotely sensed imagery), but they instead perceive the walking environment at eye-level (like street-view imagery; Jiang et al., 2017). Our analytical approach using street-view images and a scene labeling technique allows

Table 3
Description and descriptive statatistics of model variables

Variables		Description	Obs	Mean	Std. Dev.	Min.	Max.
Outcome	Perceived Risk	Sum of perceived risk point at the intersection	533	3.09	6.35	0.00	52.00
Risk exposure	Crossing	Estimated number of daily student crossings at the intersection	533	4.75	9.08	0.00	66.00
	Population Density	Population density of 2010	533	18.74	13.57	0.10	100.20
	Street Density	Street density of census block	533	55.84	27.16	0.38	125.50
Visibility	Sky	Proportion of pixels occupied by sky on street-view image at the intersection	533	0.28	0.11	0.01	0.60
	Green	Proportion of pixels occupied by green on street-view image at the intersection	533	0.05	0.06	0.00	0.41
Traffic Flow	Road	Proportion of pixels occupied by road on street-view image at the intersection	533	0.14	0.05	0.04	0.27
	Pedestrian	Number of pedestrians on street-view images at the intersection	533	0.27	0.43	0.00	4.00
	Vehicle	Number of vehicles on street-view images at the intersection	533	2.81	1.44	0.00	8.25
Pedestrian Separation	Sidewalk	Proportion of pixels occupied by sidewalk on street-view image at the intersection	533	0.03	0.02	0.00	0.09
Intersection Type	Intersection Type 1	Wide road with sidewalks	139				
	Intersection Type 2	Wide road with roadside trees	56				
	Intersection Type 3	Wide road with a clear visibility	103				
	Intersection Type 4	Narrow alley with high building density	235				

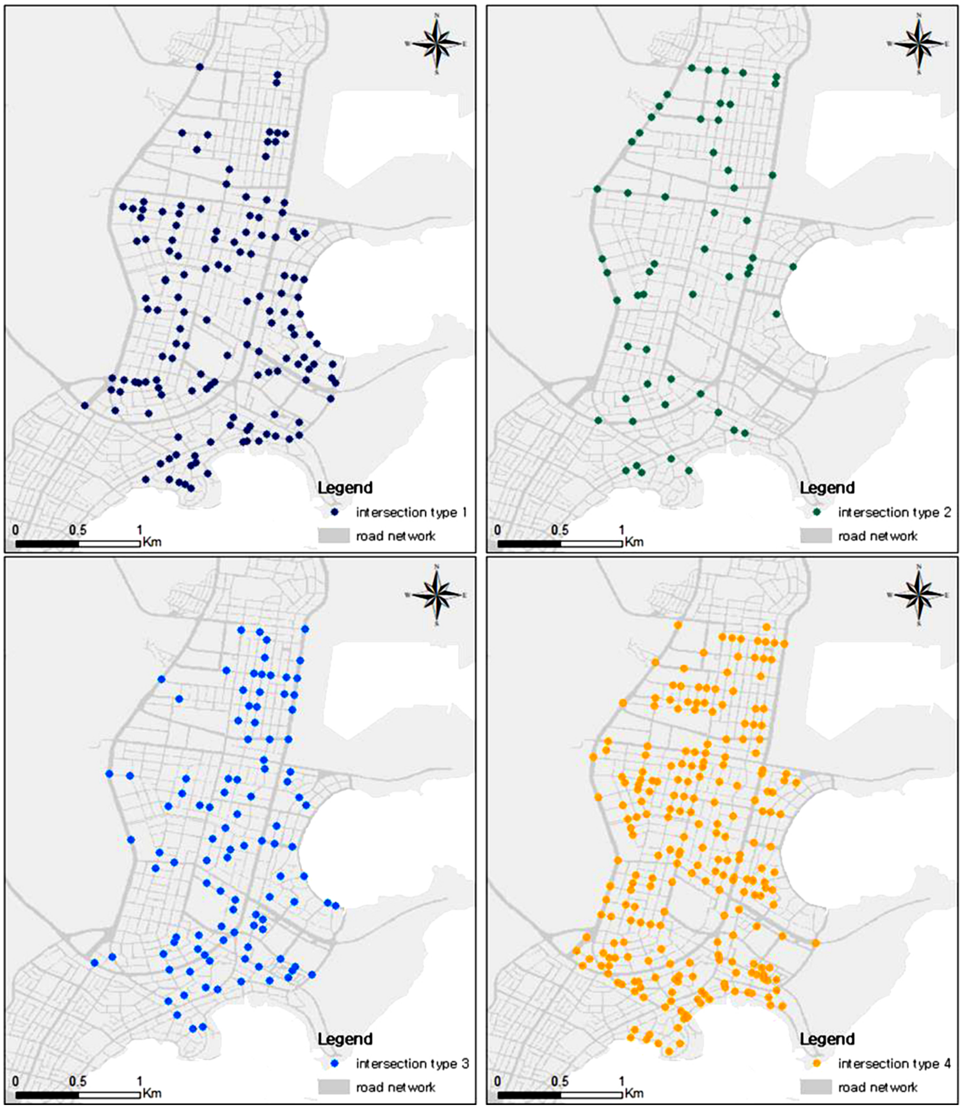

Fig. 6. Spatial distribution of four types of intersections

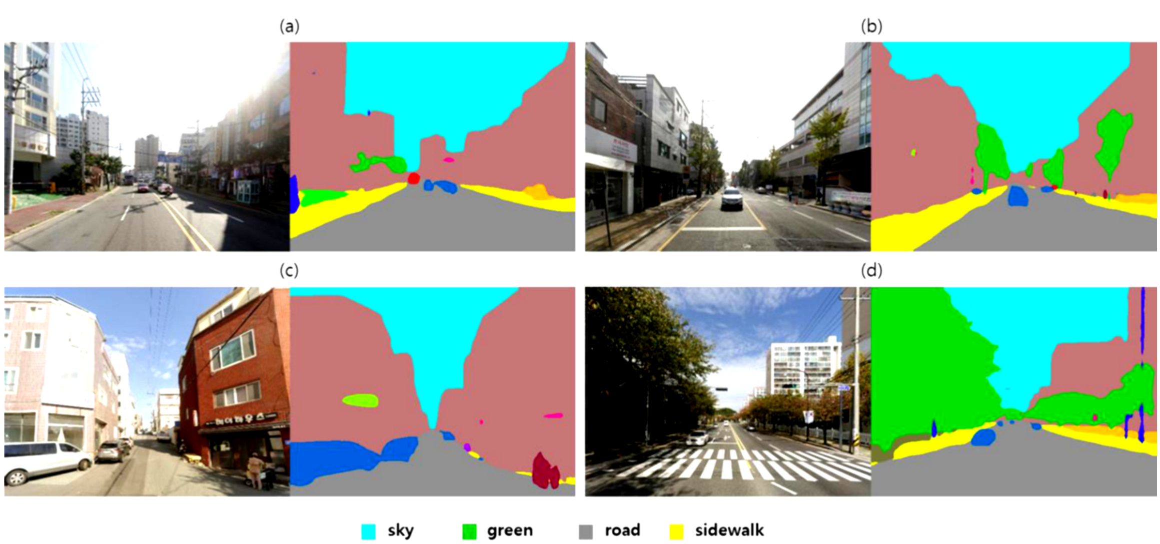

Fig. 7. Intersections where street-view imagery has mean value of proportion of (a) sky, (b) green, (c) road, (d) sidewalk

Table 4
Results of two standard NB models for percived crash risk at intersections

Variables		Model 1		Model 2	
		Coeff.	P-value	Coeff.	P-value
Visibility	Sky	−1.325	0.038		
	Green	1.280	0.203		
Traffic Flow	Road	2.789	0.051		
	Pedestrian	0.280	0.014	0.309	0.007
	Vehicle	0.014	0.764	0.017	0.695
Separation of Pedestrian	Sidewalk	0.795	0.803		
Risk Exposure	Crossing	0.106	0.000	0.107	0.000
	Population Density	0.009	0.126	0.011	0.069
	Street Density	0.010	0.000	0.010	0.001
Intersection Type	Intersection Type 1			(ref.)	
	Intersection Type 2			0.007	0.975
	Intersection Type 3			−0.356	0.048
	Intersection Type 4			−0.247	0.086
Number of Observations		533		533	
Dispersion Parameter		1.093		1.199	
AIC		1915.147		1916.201	
Log likelihood		−946.573		−948.101	

for the measurement of urban characteristics as perceived at the eye-level in an automated and scalable way.

Our model examining associations between the built environment and perceived safety at intersections presented the crash risk factors at intersections for school-aged children. The main findings of this study are consistent with the body of the literature investigating the environmental correlates of pedestrian crashes. Lower pedestrian visibility, wider roads, a higher number of student crossings, a higher population density, and higher street density were all significant and were all positively associated with a higher perceived crash risk at the intersections studied.

Previous research has shown that perceived crash risk can be an indicator of crash risk and a tool useful for providing guidelines for proactive interventions (Cho et al., 2009; Diógenes et al., 2010; Schneider et al., 2004). Intersections with lower visibility increased perceived crash risk, which implies that taking engineering measures to increase the visibility and conspicuity of pedestrians is necessary. Lee et al. (2016) suggested that imposing a building setback regulation or creating open spaces near intersections might be effective countermeasures that would increase visibility for both pedestrians and drivers and, therefore, reduce potential crash risks for children. In addition, parking restrictions (Agran et al., 1996), diagonal parking as a replacement for parallel parking (Retting et al., 2003), and relocating bus stops to the far side of intersections (Berger & Knoblauch, 1975) can enhance the visibility of both pedestrians and drivers.

Although narrow alleys and dense buildings reduce levels of visibility, the level of perceived crash risk at this type of intersections (type 4) was slightly lower than the risk at wide roads (type 1). The difficulty of crossing wide roads might be bigger concerns than their ability to visually detect incoming/outgoing traffic at intersections. Empirical studies of pedestrian crashes have commonly reported that increasing the number of lanes and road widths would increase the severity of vehicle-pedestrian crashes as well as crash rates (Ewing & Dumbaugh, 2009; Klaitman et al., 2018; Manuel et al., 2014; Pour et al., 2017; Yanmaz-Tuzel & Ozbay, 2010). We emphasize that more attention needs to be paid to school zones since child pedestrians need more time crossing wider roads (Wann et al., 2011). Thus, installing pedestrian crossing facilities and road medians on wider roads could be an effective means of improving pedestrian safety at wider road crossings (Pour et al., 2017). Other countermeasures for reducing crash risk include speed management, such as lane narrowing, and multi-way stop signs in residential neighborhoods with large numbers of children (Retting et al., 2003).

The technical limitations of this study need to be noted. The performance of the proposed algorithms is relatively accurate, but they less precise with respect to detecting small features such as greenery and

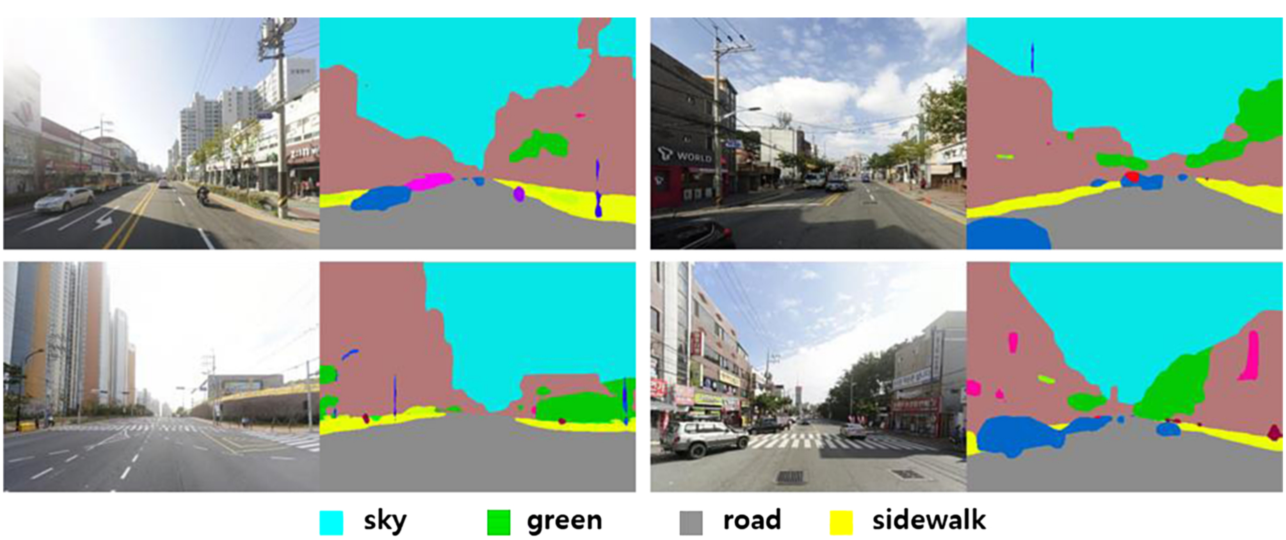

Fig. 8. Interesection Type 1 (wide road and presence of sidewalk)

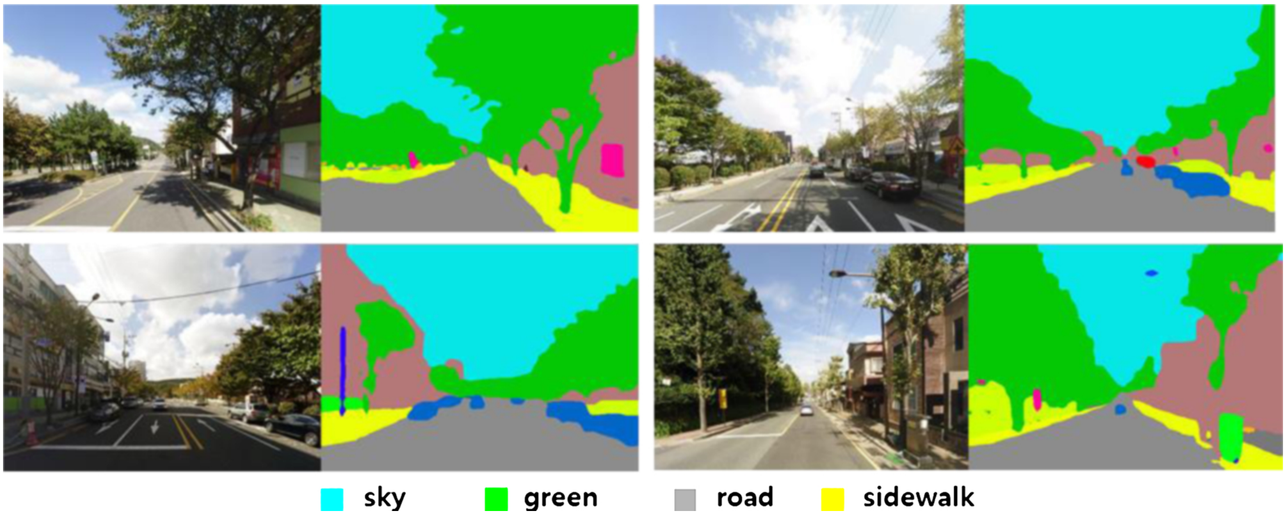

Fig. 9. Intersection Type 2 (wide road and large roadside trees)

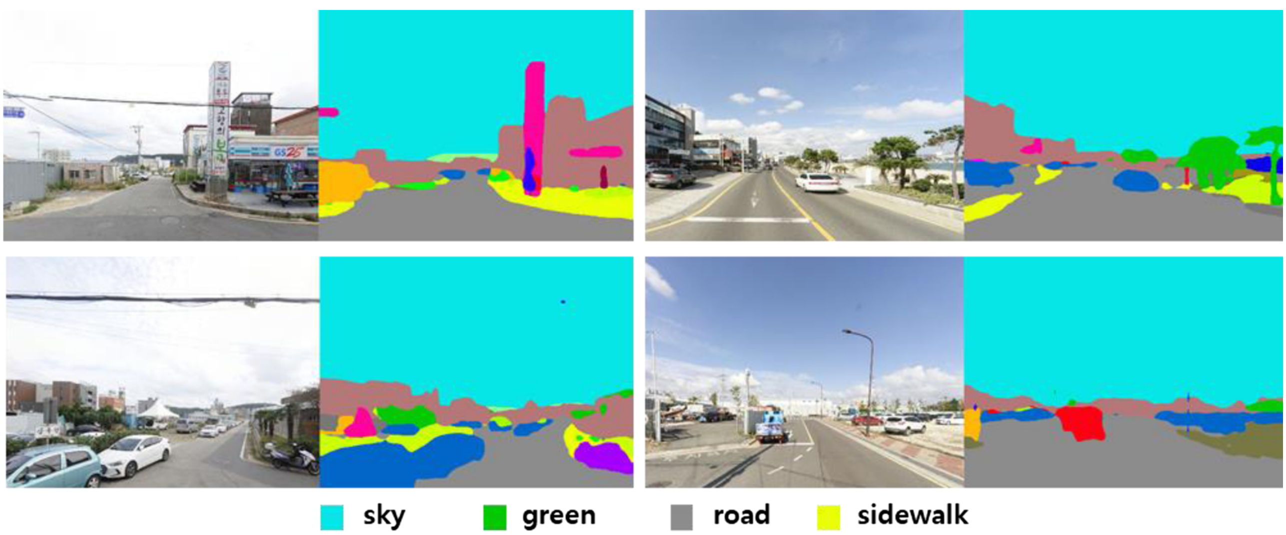

Fig. 10. Intersection Type 3 (wide road with clear visibility)

Fig. 11. Intersection Type 4 (narrow alley with dense buildings)

sidewalks. Relatively low precision of sidewalk detection may explain why the proportion of sidewalk was not significantly associated with perceived crash risk. Previous studies have indicated that number of transit stops (Pulugurtha & Sambhara, 2011; Ukkusuri et al., 2012; Yu & Zhu, 2016), lighting design (Peña-García et al., 2015; Siddiqui et al., 2006), and traffic-calming measures (Jones et al., 2005; Rothman et al., 2014) were all associated with an increase in pedestrian crashes. However, these features could not be successfully detected with the algorithm that we used. Future research is needed to uncover the influence of more refined measures of the built environment on perceived crash risk.

One of the common concerns for scene segmentation approach is measurement validity. We measured visual information of three dimensional built environment using the proportion of sky, green and road in street-view images. The present study interpreted that a larger proportion of sky implied visual openness of street canyon (Li & Ratti, 2019) or higher visible area from pedestrian's perspective. (Yin & Wang, 2016). Similarly, the proportion of green was understood as obstruction of street greenery (Li et al., 2017a,b), and the proportion of pixels occupied by road implied width of road (Ye et al. 2019). It is important to note that what we attempted to measure is visually perceived built environment and it may not exactly identical with objective measurement for the built environment. For instance, if a street-view image has higher proportion of pixels occupied by road, pedestrian may perceive the street wider than the objective measurement of width. Since the main outcome of our study is perceived crash risk, we assumed that measurements of visually perceived environment might be more appropriate measures to examine their influence on the outcome.

Although street-view imagery is a promising data source for urban studies, some limitations remain related to its use for cartographic purposes. First, the approximately two-year time gap between street-view imagery and present environment conditions may reduce the accuracy of intersection investigations (Lee et al., 2016). For instance, some neighborhood features are inherently unstable over time, including seasonal or weather-related fluctuations of greenery, number of pedestrians, and parked or moving vehicles. Second, street-view images were not collected for scientific purposes. Therefore, street-view imagery might have geometrical distortions that influence the accuracy of our analysis (Li et al., 2017a,b).

These shortcomings aside, this study demonstrated a method for capturing built environment measurements at the eye-level and their influence on perceived crash risk using street-view imagery and scene labeling techniques. To the best of our knowledge, no other study has measured the attributes of the built environment at the eye-level in an automated and scalable way. The scene labeling method used presented the possibility of opening a black box of deep learning, therefore overcoming the technical limitations of current computer vision. Moreover, this approach provides a better understanding of the appearance of urban variations on the ground and how they are perceived. Ultimately, this approach might provide appropriate and proactive interventions that reduce the potential crash risk at intersections.

CRediT authorship contribution statement

Jae-Hong Kwon: Conceptualization, Methodology, Writing - original draft, Formal analysis, Investigation, Software, Visualization. **Gi-Hyoug Cho:** Conceptualization, Methodology, Writing - review & editing, Supervision, Project administration, Funding acquisition, Resources.

Acknowledgements

This research was supported by a grant (19CTAP-C151930-01) from the Technology Advancement Research Program of the Korea Ministry of Land, Infrastructure and Transport

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.aap.2020.105716.

References

Aarts, L., Van Schagen, I., 2006. Driving speed and the risk of road crashes: A review. Accident Analysis & Prevention 38 (2), 215–224.
Agran, P.F., Winn, D.G., Anderson, C.L., Tran, C., Del Valle, C.P., 1996. The role of the physical and traffic environment in child pedestrian injuries. Pediatrics 98 (6), 1096–1103.
Alain, G., Bengio, Y., 2016. Understanding intermediate layers using linear classifier probes. arXiv preprint arXiv 1610.01644.
Alemi, A.A., Fischer, I., Dillon, J.V., Murphy, K., 2016. Deep variational information bottleneck. arXiv preprint arXiv 1612.00410.
Anguelov, D., Dulong, C., Filip, D., Frueh, C., Lafon, S., Lyon, R., et al., 2010. Google street view: Capturing the world at street level. Computer 43 (6), 32–38.
Aoki, M., Moore, L., 1996. KIDSAFE: A young pedestrian safety study. ITE journal 66, 36–45.
Badland, H.M., Opit, S., Witten, K., Kearns, R.A., Mavoa, S., 2010. Can virtual streetscape audits reliably replace physical streetscape audits? Journal of Urban Health 87 (6), 1007–1016.
Berg, A.C., Grabler, F., Malik, J., 2007. Parsing images of architectural scenes. Paper presented at the 2007 IEEE 11th International Conference on Computer Vision.
Berger, W.G., Knoblauch, R.L., 1975. Urban pedestrian accident countermeasures experimental evaluation Vol. 1 BioTechnology.
Berhanu, G., 2004. Models relating traffic safety with road environment and traffic flows on arterial roads in Addis Ababa. Accident Analysis & Prevention 36 (5), 697–704.
Boarnet, M.G., Day, K., Alfonzo, M., Forsyth, A., Oakes, M., 2006. The Irvine–Minnesota inventory to measure built environments: reliability tests. American journal of preventive medicine 30 (2), 153–159 e143.
Bowman, B.L., Vecellio, R.L., Miao, J., 1995. Vehicle and pedestrian accident models for median locations. Journal of transportation engineering 121 (6), 531–537.
Brownlee, J., 2019. Deep Learning for Computer Vision: Image Classification, Object Detection, and Face Recognition in Python. Machine Learning Mastery.
Brownson, R.C., Hoehner, C.M., Brennan, L.K., Cook, R.A., Elliott, M.B., McMullen, K.M., 2004. Reliability of 2 instruments for auditing the environment for physical activity. Journal of Physical Activity and Health 1 (3), 191–208.
Brownson, R.C., Hoehner, C.M., Day, K., Forsyth, A., Sallis, J.F., 2009. Measuring the built environment for physical activity: state of the science. American journal of preventive medicine 36 (4), S99–S123 e112.
Budzynski, M., Jamroz, K., Jelinski, L., Antoniuk, M., 2016. Why are trees still such a major hazard to drivers in Poland? Transportation research procedia 14, 4150–4159.
Cabitza, F., Rasoini, R., Gensini, G.F., 2017. Unintended consequences of machine learning in medicine. Jama 318 (6), 517–518.
Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L., 2014. Semantic image segmentation with deep convolutional nets and fully connected crfs. arXiv preprint arXiv 1412.7062.
Chen, Y., Meng, H., Wang, Z., 2009. Safety improvement practice for vulnerable road users in Beijing junctions. Retrieved from. .
Cho, G., Rodríguez, D.A., Khattak, A.J., 2009. The role of the built environment in explaining relationships between perceived and actual pedestrian and bicyclist safety. Accident Analysis & Prevention 41 (4), 692–702.
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.
Clifton, K.J., Kreamer-Fults, K., 2007. An examination of the environmental attributes associated with pedestrian–vehicular crashes near public schools. Accident Analysis & Prevention 39 (4), 708–715.
Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., et al., 2016. The cityscapes dataset for semantic urban scene understanding. Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.
Cunningham, G.O., Michael, Y.L., Farquhar, S.A., Lapidus, J., 2005. Developing a reliable senior walking environmental assessment tool. American journal of preventive medicine 29 (3), 215–217.
Dadashova, B., Ramírez, B.A., McWilliams, J.M., Izquierdo, F.A., 2016. The identification of patterns of interurban road accident frequency and severity using road geometry and traffic indicators. Transportation research procedia 14, 4122–4129.
Dalal, N., Triggs, B., 2005. Histograms of oriented gradients for human detection.
Diógenes, M.C., Lindau, L.A., de Transportes–LASTRAN, L. d. S, 2010. RISK PERCEPTION OF PEDESTRIANS AT MIDBLOCK CROSSINGS IN BRAZIL. Paper presented at the Proceedings of the 12th World Conference on Transport Research (WCTR).
Doersch, C., Gupta, A., Efros, A.A., 2013. Mid-level visual element discovery as discriminative mode seeking. Paper presented at the Advances in neural information processing systems.
Doersch, C., Singh, S., Gupta, A., Sivic, J., Efros, A., 2012. What makes paris look like paris?.
Doersch, C., Singh, S., Gupta, A., Sivic, J., Efros, A.A., 2015. What makes Paris look like Paris? Communications of the ACM 58 (12), 103–110.
Dubey, A., Naik, N., Parikh, D., Raskar, R., Hidalgo, C.A., 2016. Deep learning the city: Quantifying urban perception at a global scale. Paper presented at the European conference on computer vision.
Edwards, N., Hooper, P., Trapp, G.S., Bull, F., Boruff, B., Giles-Corti, B., 2013.

Development of a public open space desktop auditing tool (POSDAT): a remote sensing approach. Applied Geography 38, 22–30.
Elvik, R., Bjørnskau, T., 2005. How accurately does the public perceive differences in transport risks?: An exploratory analysis of scales representing perceived risk. Accident Analysis & Prevention 37 (6), 1005–1011.
Everingham, M., Van Gool, L., Williams, C.K., Winn, J., Zisserman, A., 2010. The pascal visual object classes (voc) challenge. International journal of computer vision 88 (2), 303–338.
Ewing, R., Dumbaugh, E., 2009. The built environment and traffic safety: a review of empirical evidence. Journal of Planning Literature 23 (4), 347–367.
Farabet, C., Couprie, C., Najman, L., LeCun, Y., 2012. Learning hierarchical features for scene labeling. IEEE transactions on pattern analysis and machine intelligence 35 (8), 1915–1929.
Girshick, R., Donahue, J., Darrell, T., Malik, J., 2014. Rich feature hierarchies for accurate object detection and semantic segmentation. Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.
Goel, A., Juneja, M., Jawahar, C., 2012. Are buildings only instances?: exploration in architectural style categories. Paper presented at the Proceedings of the Eighth Indian Conference on Computer Vision, Graphics and Image Processing.
Griew, P., Hillsdon, M., Foster, C., Coombes, E., Jones, A., Wilkinson, P., 2013. Developing and testing a street audit tool using Google Street View to measure environmental supportiveness for physical activity. International Journal of Behavioral Nutrition and Physical Activity 10 (1), 103.
Harwood, D.W., Bauer, K.M., Richard, K.R., Gilmore, D.K., Graham, J.L., Potts, I.B., et al., 2008. Pedestrian safety prediction methodology. Retrieved from. .
Hills, B.L., 1980. Vision, visibility, and perception in driving. Perception 9 (2), 183–216.
Jiang, B., Deal, B., Pan, H., Larsen, L., Hsieh, C.-H., Chang, C.-Y., Sullivan, W.C., 2017. Remotely-sensed imagery vs. eye-level photography: Evaluating associations among measurements of tree cover density. Landscape and Urban Planning 157, 270–281.
Jiang, B., Larsen, L., Deal, B., Sullivan, W.C., 2015. A dose–response curve describing the relationship between tree cover density and landscape preference. Landscape and Urban Planning 139, 16–25.
Jones, S.J., Lyons, R.A., John, A., Palmer, S., 2005. Traffic calming policy can reduce inequalities in child pedestrian injuries: database study. Injury Prevention 11 (3), 152–156.
Ju, M., Luo, H., Wang, Z., Hui, B., Chang, Z., 2019. The application of improved YOLO V3 in multi-scale target detection. Applied Sciences 9 (18), 3775.
Kelly, C.M., Wilson, J.S., Baker, E.A., Miller, D.K., Schootman, M., 2012. Using Google Street View to audit the built environment: inter-rater reliability results. Annals of Behavioral Medicine 45 (suppl_1), S108–S112.
Klaitman, S.S., Solomonov, E., Yaloz, A., Biswas, S., 2018. The Incidence of Road Traffic Crashes Among Young People Aged 15–20 Years: Differences in Behavior, Lifestyle and Sociodemographic Indices in the Galilee and the Golan. Frontiers in public health 6.
Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. Paper presented at the Advances in neural information processing systems.
Lee, G., Park, Y., Kim, J., Cho, G.-H., 2016. Association between intersection characteristics and perceived crash risk among school-aged children. Accident Analysis & Prevention 97, 111–121.
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.
Li, X., Ratti, C., Seiferling, I., 2017a. Mapping urban landscapes along streets using google street view. Paper presented at the International cartographic conference.
Li, X., Ratti, C., Seiferling, I., 2017b. Mapping urban landscapes along streets using google street view. Paper presented at *the International cartographic conference*.
Li, X., Ratti, C., 2019. Using Google street view for street-level urban form analysis, a case study in Cambridge, Massachusetts. The Mathematics of Urban Morphology. Springer, pp. 457–470.
Liu, L., Silva, E.A., Wu, C., Wang, H., 2017. A machine learning-based method for the large-scale evaluation of the qualities of the urban environment. Computers, Environment and Urban Systems 65, 113–125.
Long, J., Shelhamer, E., Darrell, T., 2015. Fully convolutional networks for semantic segmentation. Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.
Ma, W.J., Nie, S.P., Xu, H.F., Xu, Y.J., Zhang, Y.R., 2010. Socioeconomic status and the occurrence of non-fatal child pedestrian injury: results from a cross-sectional survey. Safety science 48 (6), 823–828.
Manning, C., Raghavan, P., Schütze, H., 2010. Introduction to information retrieval. Natural Language Engineering 16 (1), 100–103.
Manuel, A., El-Basyouny, K., Islam, M.T., 2014. Investigating the safety effects of road width on urban collector roadways. Safety science 62, 305–311.
Meir, A., Oron-Gilad, T., Parmet, Y., 2015. Can child-pedestrians' hazard perception skills be enhanced? Accident Analysis & Prevention 83, 101–110.
Meir, A., Parmet, Y., Oron-Gilad, T., 2013. Towards understanding child-pedestrians' hazard perception abilities in a mixed reality dynamic environment. Transportation research part F: traffic psychology and behaviour 20, 90–107.
Naik, N., Philipoom, J., Raskar, R., Hidalgo, C., 2014. Streetscore-predicting the perceived safety of one million streetscapes. Paper presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops.
Noh, H., Hong, S., Han, B., 2015. Learning deconvolution network for semantic segmentation. Paper presented at the Proceedings of the IEEE international conference on computer vision.
Noland, R.B., Oh, L., 2004. The effect of infrastructure and demographic change on traffic-related fatalities and crashes: a case study of Illinois county-level data. Accident Analysis & Prevention 36 (4), 525–532.
Odgers, C.L., Caspi, A., Bates, C.J., Sampson, R.J., Moffitt, T.E., 2012. Systematic social observation of children's neighborhoods using Google Street View: a reliable and cost-effective method. Journal of Child Psychology and Psychiatry 53 (10), 1009–1017.
Oltedal, S., Moen, B.-E., Klempe, H., Rundmo, T., 2004. Explaining risk perception: An evaluation of cultural theory. Trondheim: Norwegian University of Science and Technology 85 (1-33), 86.
Ordonez, V., Berg, T.L., 2014. Learning high-level judgments of urban perception. Paper presented at the European conference on computer vision.
Park, E., Kim, K.J., Jin, D., del Pobil, A.P., 2012. Towards a successful mobile map service: an empirical examination of technology acceptance model. Paper presented at the International conference on networked digital technologies.
Peña-García, A., Hurtado, A., Aguilar-Luzón, M., 2015. Impact of public lighting on pedestrians' perception of safety and well-being. Safety science 78, 142–148.
Petch, R., Henson, R., 2000. Child road safety in the urban environment. Journal of Transport Geography 8 (3), 197–211.
Pikora, T., Giles-Corti, B., Bull, F., Jamrozik, K., Donovan, R., 2003. Developing a framework for assessment of the environmental determinants of walking and cycling. Social science & medicine 56 (8), 1693–1703.
Pohlen, T., Hermans, A., Mathias, M., Leibe, B., 2017. Full-resolution residual networks for semantic segmentation in street scenes. Paper presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.
Porzi, L., Rota Bulò, S., Lepri, B., Ricci, E., 2015. Predicting and understanding urban perception with convolutional neural networks. Paper presented at the Proceedings of the 23rd ACM international conference on Multimedia.
Pour, A.T., Moridpour, S., Tay, R., Rajabifard, A., 2017. Neighborhood influences on vehicle-pedestrian crash severity. Journal of Urban Health 94 (6), 855–868.
Pulugurtha, S.S., Sambhara, V.R., 2011. Pedestrian crash estimation models for signalized intersections. Accident Analysis & Prevention 43 (1), 439–446.
Rafaely, V., Meyer, J., Zilberman-Sandler, I., Viener, S., 2006. Perception of traffic risks for older and younger adults. Accident Analysis & Prevention 38 (6), 1231–1236.
Redmon, J., Divvala, S., Girshick, R., Farhadi, A., 2016. You only look once: Unified, real-time object detection. Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.
Redmon, J., Farhadi, A., 2017. YOLO9000: better, faster, stronger. Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.
Redmon, J., Farhadi, A., 2018. Yolov3: An incremental improvement. arXiv preprint arXiv 1804.02767.
Retting, R.A., Ferguson, S.A., McCartt, A.T., 2003. A review of evidence-based traffic engineering measures designed to reduce pedestrian–motor vehicle crashes. American journal of public health 93 (9), 1456–1463.
Roth, H., Lu, L., Liu, J., Yao, J., Seff, A., Cherry, K., et al., 2016. Improving Computer-Aided Detection Using Convolutional Neural Networks and Random View Aggregation. IEEE transactions on medical imaging 35 (5), 1170–1181.
Rothman, L., Macarthur, C., To, T., Buliung, R., Howard, A., 2014. Motor vehicle-pedestrian collisions and walking to school: the role of the built environment. Pediatrics 133 (5), 776–784.
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.
Rundmo, T., 1999. Perceived risk, health and consumer behaviour. Journal of Risk Research 2 (3), 187–200.
Salesses, P., Schechtner, K., Hidalgo, C.A., 2013. The collaborative image of the city: mapping the inequality of urban perception. PloS one 8 (7), e68400.
Schneider, R.J., Diogenes, M.C., Arnold, L.S., Attaset, V., Griswold, J., Ragland, D.R., 2010. Association between roadway intersection characteristics and pedestrian crash risk in Alameda County. California. Transportation Research Record 2198 (1), 41–51.
Schneider, R.J., Ryznar, R.M., Khattak, A.J., 2004. An accident waiting to happen: a spatial approach to proactive pedestrian planning. Accident Analysis & Prevention 36 (2), 193–211.
Schofer, J.L., Christoffel, K.K., Donovan, M., Lavigne, J.V., Tanz, R.R., Wills, K.E., 1995. Child pedestrian injury taxonomy based on visibility and action. Accident Analysis & Prevention 27 (3), 317–333.
Schwarz, A., 2008. *Lokale Scoring-Modelle*: Eul.
Sermanet, P., Eigen, D., Zhang, X., Mathieu, M., Fergus, R., LeCun, Y., 2013. Overfeat: Integrated recognition, localization and detection using convolutional networks. arXiv preprint arXiv 1312.6229.
Shalunts, G., Haxhimusa, Y., Sablatnig, R., 2011. Architectural style classification of building facade windows. Paper presented at the International Symposium on Visual Computing.
Shrikumar, A., Greenside, P., Kundaje, A., 2017. Learning important features through propagating activation differences. Paper presented at the Proceedings of the 34th International Conference on Machine Learning-Volume 70.
Shwartz-Ziv, R., Tishby, N., 2017. Opening the black box of deep neural networks via information. arXiv preprint arXiv 1703.00810.
Siddiqui, N.A., Chu, X., Guttenplan, M., 2006. Crossing locations, light conditions, and pedestrian injury severity. Transportation Research Record 1982 (1), 141–149.
Simon, L., Teboul, O., Koutsourakis, P., Van Gool, L., Paragios, N., 2012. Parameter-free/pareto-driven procedural 3d reconstruction of buildings from ground-level sequences. Paper presented at the 2012 IEEE Conference on Computer Vision and Pattern Recognition.
Singh, S., Gupta, A., Efros, A.A., 2012. Unsupervised discovery of mid-level discriminative patches. Paper presented at the European Conference on Computer Vision.
Sjöberg, L., Moen, B.-E., Rundmo, T., 2004. Explaining risk perception. An evaluation of the psychometric paradigm in risk perception research. Rotunde publikasjoner

Rotunde 84, 55–76.
Song, Y., Knaap, G.-J., 2007. Quantitative classification of neighbourhoods: The neighbourhoods of new single-family homes in the Portland Metropolitan Area. Journal of Urban Design 12 (1), 1–24.
Suminski, R.R., Heinrich, K.M., Poston, W.S., Hyder, M., Pyle, S., 2008. Characteristics of urban sidewalks/streets and objectively measured physical activity. Journal of Urban Health 85 (2), 178–190.
Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., et al., 2015. Going deeper with convolutions. Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.
Taylor, B.T., Fernando, P., Bauman, A.E., Williamson, A., Craig, J.C., Redman, S., 2011. Measuring the quality of public open space using Google Earth. American journal of preventive medicine 40 (2), 105–112.
Tompson, J.J., Jain, A., LeCun, Y., Bregler, C., 2014. Joint training of a convolutional network and a graphical model for human pose estimation. Paper presented at the Advances in neural information processing systems.
Ukkusuri, S., Miranda-Moreno, L.F., Ramadurai, G., Isa-Tavarez, J., 2012. The role of built environment on pedestrian crash frequency. Safety science 50 (4), 1141–1151.
Ulleberg, P., Rundmo, T., 2003. Personality, attitudes and risk perception as predictors of risky driving behaviour among young drivers. Safety science 41 (5), 427–443.
Wang, R., Liu, Y., Lu, Y., Zhang, J., Liu, P., Yao, Y., Grekousis, G., 2019. Perceptions of built environment and health outcomes for older Chinese in Beijing: A big data approach with street view images and deep learning technique. Computers, Environment and Urban Systems 78, 101386.
Wann, J.P., Poulter, D.R., Purcell, C., 2011. Reduced sensitivity to visual looming inflates the risk posed by speeding vehicles when children try to cross the road. Psychological science 22 (4), 429–434.
Weissenberg, J., Riemenschneider, H., Prasad, M., Van Gool, L., 2013. Is there a procedural logic to architecture? Paper presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.
Wilks, D.S., 2011. Cluster Analysis. Vol. 100. International Geophysics. pp. 603–616.
Wilson, J.S., Kelly, C.M., Schootman, M., Baker, E.A., Banerjee, A., Clennin, M., Miller, D.K., 2012. Assessing the built environment using omnidirectional imagery. American journal of preventive medicine 42 (2), 193–199.
Xiao, T., Liu, Y., Zhou, B., Jiang, Y., Sun, J., 2018. Unified perceptual parsing for scene understanding. Paper presented at the Proceedings of the European Conference on Computer Vision (ECCV).
Yanmaz-Tuzel, O., Ozbay, K., 2010. A comparative Full Bayesian before-and-after analysis and application to urban road safety countermeasures in New Jersey. Accident Analysis & Prevention 42 (6), 2099–2107.
Yao, Y., Liang, Z., Yuan, Z., Liu, P., Bie, Y., Zhang, J., et al., 2019. A human-machine adversarial scoring framework for urban perception assessment using street-view images. International Journal of Geographical Information Science 1–22.
Ye, Y., Zeng, W., Shen, Q., Zhang, X., Lu, Y., 2019. The visual quality of streets: A human-centred continuous measurement based on machine learning algorithms and street view images. Environment and Planning B: Urban Analytics and City Science 46 (8), 1439–1457.
Yin, L., Wang, Z., 2016. Measuring visual enclosure for street walkability: Using machine learning algorithms and Google Street View imagery. Applied Geography 76, 147–153.
Yu, C.-Y., Zhu, X., 2016. Planning for safe schools: impacts of school siting and surrounding environments on traffic safety. Journal of Planning Education and Research 36 (4), 476–486.
Zhang, F., Zhou, B., Liu, L., Liu, Y., Fung, H.H., Lin, H., Ratti, C., 2018. Measuring human perceptions of a large-scale urban region using machine learning. Landscape and Urban Planning 180, 148–160.
Zhang, L., Lin, L., Liang, X., He, K., 2016. Is faster R-CNN doing well for pedestrian detection? October In: European conference on computer vision. Springer, Cham. pp. 443–457.
Zhou, B., Lapedriza, A., Xiao, J., Torralba, A., Oliva, A., 2014. Learning deep features for scene recognition using places database. Paper presented at the Advances in neural information processing systems.
Zhou, B., Zhao, H., Puig, X., Fidler, S., Barriuso, A., Torralba, A., 2016. Semantic Understanding of Scenes through the ADE20K Dataset. arXiv preprint arXiv 1608.05442.
Zhou, B., Zhao, H., Puig, X., Fidler, S., Barriuso, A., Torralba, A., 2017. Scene parsing through ade20k dataset. Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.
Zou, Y., Wu, L., Lord, D., 2015. Modeling over-dispersed crash data with a long tail: examining the accuracy of the dispersion parameter in negative binomial models. Analytic Methods in Accident Research 5, 1–16.