Applied Geography 63 (2015) 337–345

Contents lists available at ScienceDirect

Applied Geography

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

'Big data' for pedestrian volume: Exploring the use of Google Street View images for pedestrian counts

Li Yin [a,*], Qimin Cheng [b], Zhenxin Wang [c], Zhenfeng Shao [d]

[a] Department of Urban and Regional Planning, University at Buffalo, The State University of New York, Buffalo, NY 14214, USA
[b] Department of Electronics and Information Engineering, Huazhong University of Science & Technology, Wuhan 430074, China
[c] Center for Human-Engaged Computing, Kochi University of Technology, 185 Miyanokuchi, Tosayamada-Cho, Kami-Shi, Kochi 782-8502, Japan
[d] The State Key Laboratory of Information Engineering on Surveying Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China

ARTICLE INFO

Article history:
Received 9 March 2015
Received in revised form
19 July 2015
Accepted 19 July 2015
Available online 30 July 2015

Keywords:
'Big data'
Google Street View
Pedestrian count
Walkability

ABSTRACT

New sources of data such as 'big data' and computational analytics have stimulated innovative pedestrian oriented research. Current studies, however, are still limited and subjective with regard to the use of Google Street View and other online sources for environment audits or pedestrian counts because of the manual information extraction and compilation, especially for large areas. This study aims to provide future research an alternative method to conduct large scale data collection more consistently and objectively on pedestrian counts and possibly for environment audits and stimulate discussion of the use of 'big data' and recent computational advances for planning and design. We explore and report information needed to automatically download and assemble Google Street View images, as well as other image parameters for a wide range of analysis and visualization, and explore extracting pedestrian count data based on these images using machine vision and learning technology. The reliability tests results based on pedestrian information collected from over 200 street segments in Buffalo, NY, Washington, D.C., and Boston, MA respectively suggested that the image detection method used in this study are capable of determining the presence of pedestrian with a reasonable level of accuracy. The limitation and potential improvement of the proposed method is also discussed.

1. Introduction

"[B]ig data and computational sciences are changing how we can analyze and understand individual and collective human behavior" (Ruppert, 2013; p269). Planners and social scientists have been using survey, interview, and field work to collect empirical data for years (Ruppert, 2013). Recent studies suggested to attend to new forms of empirical data and conceptions (Lury and Wakeford, 2012; Ruppert, 2013). 'Big data' and computational analytics create important opportunities for interdisciplinary approach to study new phenomena or to study old questions with new data and insights (Arribas-Bel, 2014; Ruppert, 2013). They have a potential to help study social phenomena in ways never before imagined and possible (Arribas-Bel, 2014; Watts, 2007).

New data sources and technologies have simulated research on walkability (Lee & Talen, 2014). Responding to the growing demand for walkable and transit-oriented development in recent years, pedestrian activity has been used to study what it is about the built environment that gets people active such as walk and bike in their neighborhood (Ewing & Clemente, 2013). Pedestrian count is a quantitative measure of pedestrian volume to help evaluate pedestrian activity, walkability and how it correlates with land use, and other built environment characteristics (Ewing & Clemente, 2013; Hajrasouliha & Yin, 2014). The count data can also be used as baseline data to help inform planning and funding decisions. Even though some methods have been developed to estimate pedestrian volumes (Ercolano, Olson, & Spring, 1997; Landis, 1996), they are usually not designed for actual pedestrian counts and many of them do not have a fine-grained geographic scale (Scheneider et al., 2009). Collection of detailed information about non-motorized activity is insufficient and inefficient in many transportation and built environment studies, especially at a large scale.

A large amount of data on many aspects of human behavior are available on various websites nowadays (Arribas-Bel, 2014). Google Street View provides panoramic views along public streets in the U.S. and many countries around the world. It has recently been used

* Corresponding author.
E-mail address: liyin@buffalo.edu (L. Yin).

http://dx.doi.org/10.1016/j.apgeog.2015.07.010
0143-6228/

to help audit the built environments (Badland, Opit, Witten, Kearns, & Mavoa, 2010; Ben-Joseph et al., 2013; Clarke, Ailshire, Melendez, Bader, & Morenoff, 2010; Lee & Talen, 2014; Rundle, Bader, Richards, Neckerman, & Teitler, 2011) and for pedestrian counts (Ewing & Clemente, 2013). Current studies, however, are still limited and subjective with regard to the use of Google Street View for environment audits or pedestrian counts because of the manual information extraction and compilation, especially for large areas.

Building on recent studies with manual collection of pedestrian counts (Purciel et al., 2009; Ewing & Clemente, 2013), our study contributes in three ways. First, we propose an alternative method to collect pedestrian volume information more consistently and objectively and at a larger scale using automatic information extraction on Google Street View images. It can possibly be used for environment audits by automatically detecting benches, trees, building shapes, etc. Pedestrian volumes are found to be correlated with land use and development density, street network, and a sense of safety including crime and traffic (Chu, 2005; Hajrasouliha & Yin, 2014; Ozer & Kubat, 2007). Previous studies called for better data on pedestrian volumes and more effective methodologies for counting and modeling pedestrian volumes with existing data (Lee &Talen, 2014; Scheneider et al., 2009). Our method can help identify places with different pedestrian needs using readily available Google Street View data to prioritize investment in order to improve pedestrian environment with more safety, comfort and convenience for building walkable environment.

The second and third contributions lie in how this study can stimulate and push forward the interdisciplinary discussion of the use of online 'big data' and recent computational advances for planning and design. Google and other websites make a large amount of information available. However, many companies and websites, such as Google, target mainly for the third party web-based development. APIs were provided for third party websites to display Google maps and street views. Google Street View does not explicitly provide direct and enough information on their street view images, such as image parameters and how the images were stored and assembled for other purposes. Going through Google Street View and Google map websites and API codes, we report and describe related Google Street View information needed to automatically download and assemble images, as well as other image parameters for a wide range of analysis and visualization. The automatically downloaded images can be transformed into images for different research purposes, such as neighborhood audits. Finally, we present how we borrowed a tool developed in another discipline, in particular, from the most recent development in machine learning to help detect and extract pedestrian volume for design and planning of walkable environments.

2. Pedestrian activity, 'big data', and Google Street View images

Building walkable and healthy communities is a heightened and widespread interest in recent years among researchers and practitioners. Many studies are in an effort to improve the pedestrian environment and pedestrian has become the subject of increasing attention among planners, engineers and public health officials (Clifton et al., 2007; Ewing & Bartholomew, 2013). The built environment, and streets in particular as one important element, should be designed not only to enhance mobility choices but also to reinforce walkability, livability, and sustainability (Ewing & Clemente, 2013; Yin, 2014). How people use their built environment and how the built environment characteristics influence physical and psychological health for people of all ages has increasingly being studied in recent years (Frank, Sallis, Conway, Chapman, & Saelens, 2006; Forsyth, Schmitz, Hearst, & Oakes, 2008; Yin et al., 2013; Kim & Susilo, 2013). Pedestrian count has been used as an important measure for these studies.

Pedestrian count data has been traditionally collected along sampled streets through field work, self-reported survey, or automated counting. Automated counting technology for pedestrian is less developed even though it has been used for motor vehicles for many years. Most pedestrian counts are done manually. Like in-person audits for walkability, the significant limitation of current pedestrian count method is mainly on cost, time, data accuracy, and subjectivity (Badland et al., 2010; Ben-Joseph et al., 2013; Purciel et al., 2009; Rundle et al., 2011). The observation time, number of observers, and training sessions required can be substantial (Lee & Talen, 2014). There are data errors due to human mistakes in counting and data entry. Collection of such data is also less unfeasible with large and spatially dispersed samples (Purciel et al., 2009; Rundle et al., 2011). As suggested in Ewing and Clemente (2013), sample sizes are usually small when the process is manually done. In addition, Field work based on observations and self-reported surveys are more subjective than automatic counts using video-taping or sensors. The counting methods used vary in different studies (Rundle et al., 2011; Ewing & Clemente, 2013). Finally, pedestrian counts cannot be acquired easily as a secondary data source; in other words, they are usually collected and used by the same group without being verified or made available to the public. Using existing data can potentially help to increase accuracy and improve efficiency, as well as increase reuse of the data (Lee & Talen, 2014).

With the recent rapid development of internet and cloud computing, we are entering the era of 'big data' with the 'Internet of Things' and People (O'Leary, 2013). 'Internet of Things' is configured to include inputs from humans and many different things linked to the internet (O'Leary, 2013). 'Big data' was described as data gathered from different online sources by Goodchild (2013). Another definition of 'big data' is about the effort to make the rapidly expanding amount of digital information analyzable and "the actual use of that data as a means to improve productivity, generate and facilitate innovation and improve decision making" (O'Leary, 2013, p54).

Google Street View is a component of Google Map and Google Earth that serves millions of people daily with images captured in many cities in over 20 countries across four continents (Anguelov et al., 2010). It allows users to see panoramic images from points along public streets to replicate an eye-level experience and to virtually walk down the street (Ewing & Clemente, 2013; Ben-Joseph et al., 2013). Google Street View has provided an unprecedented source of visual information about our streets, for instance, pedestrians, trees, and building features. It has become a source of 'big data' and it is readily available to anyone with access to internet.

Google Street View "has rarely been utilized in published research" until recently (Ewing & Clemente, 2013, p85). A number of studies related to urban planning and public health that used Google Street View have been published since 2010 and they all found that Google Street View offered a reliable alternative for neighborhood audits associated with walking and cycling (Badland et al., 2010; Ewing & Clemente, 2013). Web-based tools such as Google Street View offer a more resource-efficient substitute for on-site audits to save time and cost by allowing for preliminary audits to be performed accurately from remote locations, and increasing the effectiveness of subsequent on-site visits (Badland et al., 2010; Ben-Joseph et al., 2013). However, current studies manually processed information from Google Street View. This is partly because Google Street View targets primarily third party web-based development with little and limited information available for other purposes; partly because recent tools developed in other disciplines to process such information have not been well applied by planners and social scientists.

Video-based and image-based human detection have had a wide range of applications in robotics, intelligent transportation and other fields for collision prediction, driver assistance, and demographic recognition etc. (Gallagher & Chen, 2009; Prioletti et al., 2013). The classic methods of pedestrian detection involve extracting the features of images first, and then applying classifiers such as Support Vector Machine, Adaboost, decision tree, etc. for feature classification (Benenson et al., 2014). One of the most widely used pedestrian detection algorithms was proposed by Dalal and Triggs (2005) and was characterized by the histogram based on the gradient direction: Histogram of Oriented Gradient (referred to as HOG) (Dalal & Triggs, 2005; Benenson et al., 2014). HOG describes the distribution of the intensity and direction of the gradient of the local image regions.

A large amount of data and computation are usually involved in a pedestrian detection process. The accuracy of recognition needs to be considered and balanced with processing time and file size. For a city like Buffalo, it would take more than 72 h to complete pedestrian detection for a total of 260,000 images to be processed, assuming one recognition algorithm can handle one frame per second. Dollar et al. (2014) studied the relationships of HOG at different scales in the image pyramid. In order to increase the computational efficiency, the Aggregated Channel Features (ACF) algorithm is designed to first estimate the effectiveness of the HOG at the large scale and then neglect the useless parts in small scale images. This method can detect pedestrians in 640*480 images with a reasonably low missing rate at 32 frames per second.

Using automatic pedestrian detecting methods developed in the machine learning field such as ACF, Google Street View images can potentially help identify number of pedestrians on a particular street objectively to generate patterns of walkability for the understanding of pedestrian activity across a city. Comparing with data collection from traditional survey or questionnaire, or manual information collection from Google Street View like Ewing & Clemente (2013), automatic information extraction based on Google Street View images is potentially more cost-effective and information may be updated more frequently once the algorithm of image download and assembling, and pedestrian detection is developed.

3. Method

This study uses City of Buffalo as a case study area to explore extracting pedestrian count data based on Google Street View images using machine vision and learning technology. We started with going through Google Map, Google Street View, and APIs to develop an algorithm for downloading images and transforming images for the automatic detection. A Matlab toolkit was used to detect pedestrian in the second step and the last step validated the detection results. These steps were reported in three subsections below.

3.1. Google Street View images: download and transformation

We found that Google Street View provides images in seven different resolution levels. At the resolution level zero, there is one image for the panorama of each shooting location. The file size of such an image is about 23 KB. There are two images for level one and eight images for level two respectively. The size of images is about 65 KB and 175 KB for these two levels respectively. At level three, number of images for each shooting location increases to 28 with a file size of about 611 KB. Level four and five have 91 and 338 images respectively. Every panorama from level 0 through 5 is composed of 512 * 512 tiles. The level 6 images do not have higher resolutions than level 5, but with a set of smaller tiles of 256 * 256 pixels.

Fig. 1 illustrated how we downloaded Google Street View images for one street block and transformed them into images for pedestrian detection and counting. Going through all related websites and APIs, we found information of these panoramic view images including the corresponding image ID, the IDs of adjacent shooting locations, shooting location and other shooting parameters, etc. As shown in the upper part of Fig. 1, one street block can have a number of shooting locations along the street, as marked by "x" in the figure. There are about 130,000 locations with panoramas from Google Street View in a city like Buffalo. If all panoramic images of a street block need to be downloaded, the coordinates of the starting point of the street block needs to be identified to find the first panorama and then move to the adjacent location along the street block to download the corresponding panoramic images until the whole street block is done. The middle section of Fig. 1 shows an example of downloaded panoramic images for one of the shooting locations (marked by the bolded x) on the example street. These panoramic images may give people an impression of street intersections because each side of the street was half of the panoramic circle with perspective views. Image one (marked both in the images in the upper right corner and in the second row) covers the lower side of the street and shows what one person sees from the bolded x with a vertical view range of 90° at the top, 0 in the middle, and minus 90° at the bottom, and a horizontal view range of 0° on the left, 90° in the middle, and 180° on the right. Image two covers the upper side of the street with the same vertical view range and a horizontal view range from 180° to 270 in the middle to 360°.

Panoramas such as those shown in the second row of Fig. 1 can be used for pedestrian detection directly; however, the detection results are less desirable because of the image distortion from perspective views. As shown in Fig. 1, the part with least distortion is in the center of image one and two. Thus, we drew squares in the center of each side of the panoramic circle to create images shown at the bottom section. The sizes of the squares were based on information about the average distance between each pair of contiguous shooting locations on a street block, and the mean horizontal viewing angle (α) that allows to see the whole street block averaged from each shooting position (see the illustration in the third section of Fig. 1). After the downloaded panorama being transformed to the street side view images as if somebody views from the street centerline for better pedestrian detection results, two images were created for each shooting locations, each of which can be assembled together with other images from the same street block to construct one image for each side of the street block. As a result, two images were created after assembling several images like image one and two on each side of the street for every street block.

3.2. Pedestrian detection

We followed the image recognition algorithm ACF proposed by Dollar, Appel, Belongie, and Perona (2014) for pedestrian detection. The training sample used in the method was collected by a camera in a car that travels on streets, which is similar to how Google Street View images were taken. Therefore, we used the trained model to detect pedestrians without having to manually annotate the pedestrians to re-train the model parameters.

Fig. 2 illustrates the steps of pedestrian detection and counting used in this study. The first step involved extracting the regions of possible pedestrians to narrow detection range, and to eliminate a large number of disturbing areas or background regions to achieve the primary detection. We then applied the ACF model to determine whether there is a pedestrian or not, followed by counting detected pedestrians. A Matlab toolbox developed by Dollar et al.

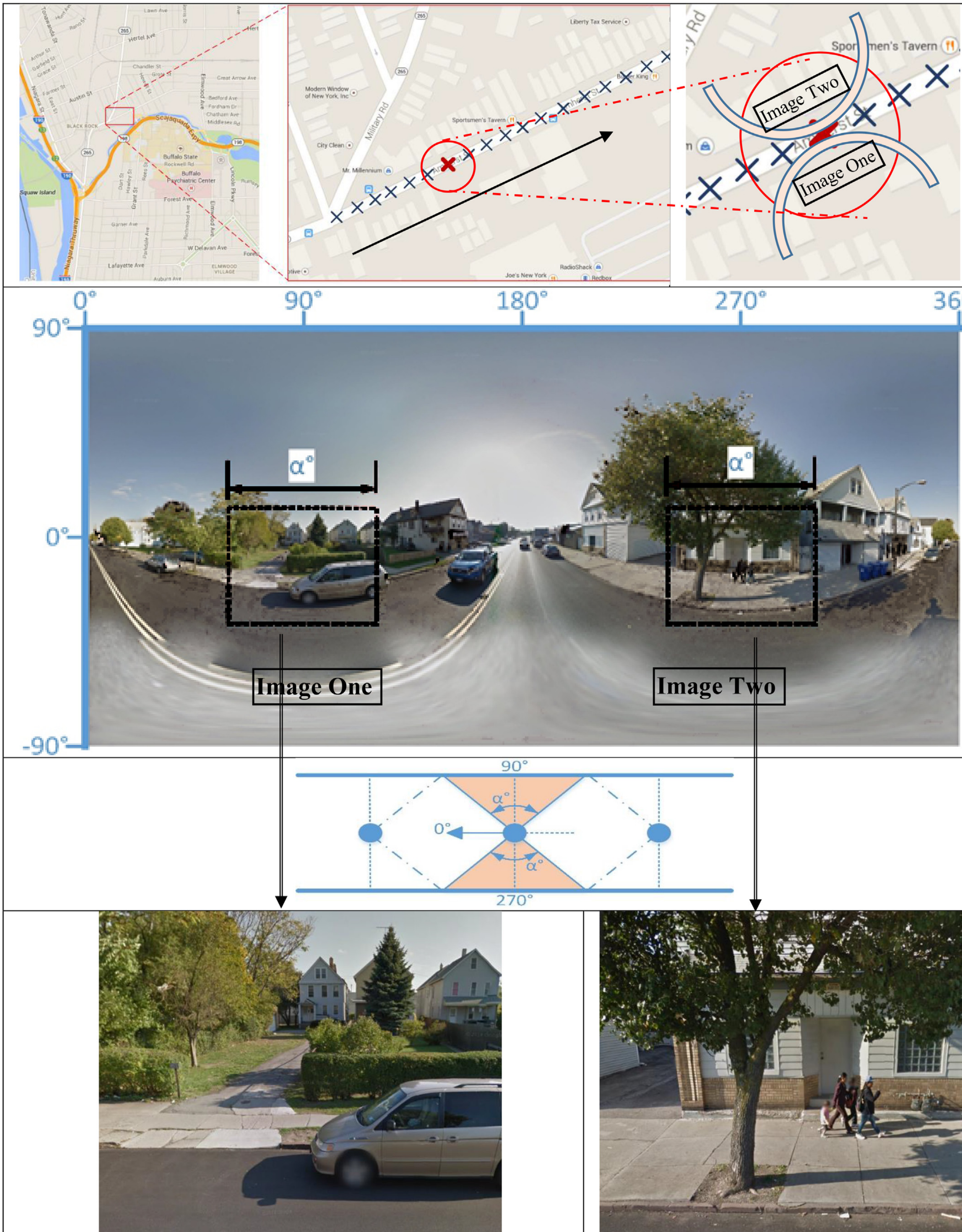

Fig. 1. Retrieving Google Street View images.

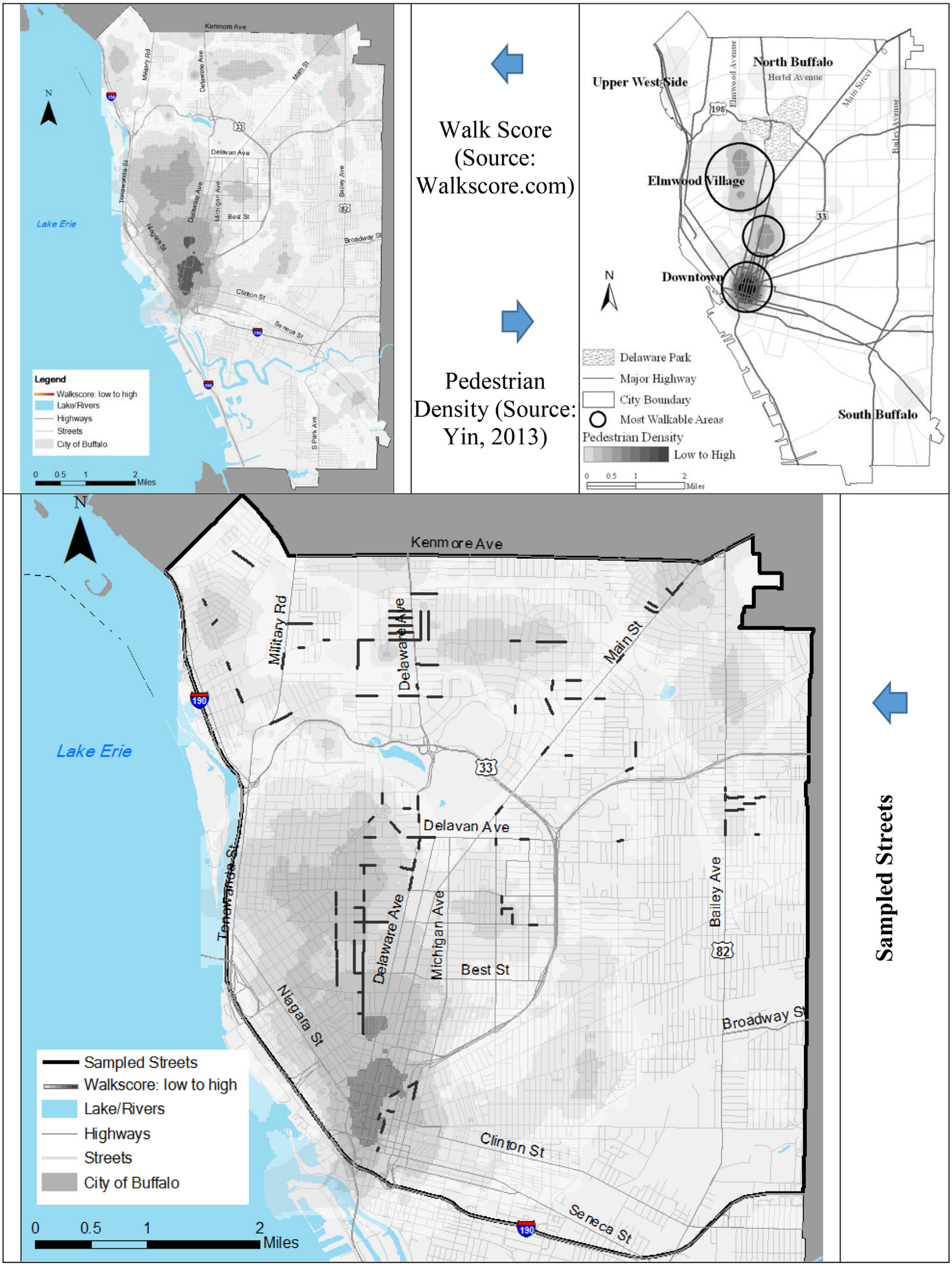

Fig. 2. Sampled streets.

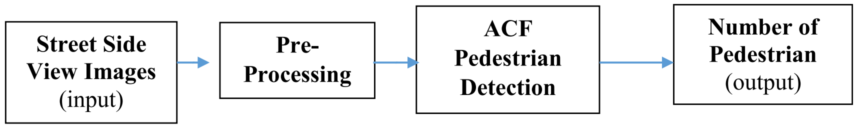

Fig. 3. Pedestrian counts extraction process.

(2014) and made available online (Dollar, 2015) was used to realize ACF. Dollar (2015), Dollar et al. (2014) provide detailed information on the detection method and the toolkit. This method was used to detect and count pedestrians for over 200 sampled street blocks in Buffalo, with a potential to apply to the entire city for the next step and other areas.

3.3. *Case study*

Fig. 3 shows the sampled streets in this study. The upper part of Fig. 3 has two maps showing the distribution of walk scores and pedestrian densities across the city. Walk score data was collected from WalkScore, a private company that provides walkability services to promote walkable neighborhoods, and pedestrian density was adopted from Yin (2013). Walk scores were calculated based on distance to the closest amenity such as parks, schools, etc. It "is an index that factors in several aspects of walkability such as accessibility and street network characteristics to offer an overall measure of the walkability of a location" (Arribas-Bel, 2014). It has been validated as a useful proxy for walkability (Lee & Talen, 2014; Carr, Dunsiger, & Marcus, 2011; Duncan, Aldstadt, Whalen, Melly, & Gortmaker, 2011). The darker the color is, the higher the walk score is (left image) and the higher the pedestrian density is (right image). These two maps show similar patterns on where pedestrians are expected in the city. Based on these two maps, street segments were sampled in the areas where there are certain levels of pedestrian volumes expected to exclude streets with very low or no possible pedestrian presence for this study. In other words, streets were sampled mostly in the areas with a color from light grey to black. Including many streets with no pedestrians may increase the matching rates between the automatic counts and counts used for validation tests but overestimate the level of reliability because of a high percentage of 0s in the sample.

We also focus on the part of the city north of interstate highway I190 partly due to the similar reason, and partly because this is where local researchers and practitioners are most interested for walkability and community and neighborhood development. A total of 212 street blocks in the City of Buffalo scattered across the focused area were sampled. Images for a total of 2661 shooting locations or panoramas were downloaded. The total file size of level two images is about 454 MB and 1.55 GB for level three images. Both resolution types were downloaded and used for pedestrian detection to test detection results by different image resolutions.

4. Validation

Several tests were performed on the reliability with respect to the automatic machine pedestrian counts from Google Street View images as an indicator of pedestrian activity. Three sets of data were used for the reliability and validation tests. One is pedestrian field counts data collected in 2013 by the graduate urban planning students at the University at Buffalo, The State University of New York. The second data were manual pedestrian counts based on Google Street View images collected by undergraduate students and randomly verified by graduate urban planning students. To test the machine count method and algorithm in other cities, Google pedestrian counts were collected for Washington, D.C. and Boston manually to be compared with the automatic machine counts for those cities. Over 200 randomly sampled street blocks were included for each of these cities. The manual counts were done by two graduated Masters of Urban Planning students independently for these two cities. The field pedestrian count data was collected by counting number of pedestrians on sampled street blocks scattered in the City of Buffalo at non-intersection locations, in a 15-min interval. Only pedestrians on the side where the data collectors were standing were counted. The counts based on Google Street View images were done for both sides of an entire street block.

The Google Street View images are static shots. Even though counts from the these static images will not match in absolute values with the 15-min field counts, our focus of measuring concordance between our automatic counts and other data sets is on the pattern match. As suggested by the literature, Cronbach's alpha can be used to test correlation between the Google machine counts, Google manual counts, and field counts (Ewing & Clemente, 2013).[1] Cronbach's alpha indicates whether different indictors or variables measure a single underlying construct and the degree to which they measure it. By relating values of different indicators to one another to test the degree of association, it can help to test whether results across different quantitative measurements on pedestrian volume are consistent.

5. Detection results and discussion

Fig. 4 illustrated how images (level 3) downloaded and processed were used for automatic pedestrian detection following steps in Fig. 2. Upper left corner is a street side view image used as the input for pedestrian detection. Upper right shows the results from pre-processing through image segmentation. At the end of this primary stage, background was removed, and some of the environmental interference was reduced to get a more effective area for the fine detection stage. This helps to reduce the false detection rate and unnecessary complex calculations at the fine detection stage to increase accuracy and processing speed. The image on the bottom shows the grey rectangles as outlined pedestrian areas for pedestrian counts. One rectangle represents one pedestrian. In this example, one pedestrian was not detected because it was partly obscured by the light pole. Fig. 4 shows a reasonably good pedestrian detection result with only one missing target. The missing error could happen on any street and may be offset when the algorithm is used for a large scale pedestrian count across a city.

This detection algorithm found a total of 59 pedestrians in the level 2 images, and 329 pedestrians in level 3 images for the sampled streets in Buffalo. Level 3 images with higher resolution, but also increased file sizes from 454 MB (level 2) to 1.55 GB (level 3) greatly helped to increase number of pedestrians detected. Through comparing ACF parameters and level 3 and 2 image

[1] Ewing & Clemente (2013) provides detailed information on using Cronbach's alpha to validate field counts against manual counts from different websites such as Google Street View.

Fig. 4. Results from ACF pedestrian detection.

parameters, we found that this is because the pedestrian resolution defined in level 3 images is closer to the training set developed with the ACF model than that in level 2.

The extraction algorithm ACF has some pre-set training parameters. One closely related parameter is the size or height for human, which defined the boxes that outlined pedestrians in Fig. 4. The default ACF parameters match level 3 images well. Level 2 images are half size of level 3 and level 4 is twice as much as level 3. None of them matches the parameters as well as the level 3 images. Therefore, images with higher or lower resolution do not generate results as well as level 3 images if the default training sets are used. That explains the lower matching rates for level 2 images. Level 4 images might perform better in a future study. However, in addition to process a much bigger dataset due to higher resolution, a re-training needs to be done to match the size of pedestrians before running ACF using level 4 images. A re-training requires collecting two more datasets: one for training and the other for test. There are also manual operations needed to adjust parameters. This study suggested that level 3 images should be used instead of level 2 or 4 or other levels in future research if the default parameters are used.

Table 1 shows the results from the Cronbach's alpha tests. Generally, the alpha value varies from 0 to 1. A Cronbach's alpha of 0.70 is acceptable, 0.80 is good, and 0.90 or higher is considered very good. The alpha for Buffalo Google machine counts and manual counts is 0.704, indicating acceptable internal consistency among these two counts. It means that a street with higher machine counts tended to have higher manual counts. The alpha for Washington, D.C. and Boston are slightly higher with a value of 0.724 and 0.780. This may be because of higher manual counting quality for these two cities. The alpha for Buffalo Google counts and field counts is 0.758, higher than the value for Google machine counts and manual counts. This may be because of the human mistakes during the counting process.

All alphas in Table 1 are at the acceptable level, with one close to the good level. Based on the acceptable level of agreement between Google machine counts and manual counts, and between machine counts and field counts, we can assume that the automatic machine counts based on Google images using the detection method in this paper are a reliable measure of pedestrian activity and can represent reasonably well how streets in Buffalo are really used by pedestrians.

6. Conclusion and limitations

In this study, we investigated the feasibility of using machine learning technology to detect and count pedestrians from Google Street View imagery. Pedestrian information on over 200 street segments in Buffalo, NY, Washington, D.C., and Boston, MA respectively was collected to help verify the detection and counting results. We found that Google Street View provides street images in seven different resolution levels, and suggested that level 3 images should be used in future research for pedestrian detection, if the default training set and parameters are used. We also suggested doing a re-training if higher resolution data are to be used so that the parameters of the pedestrians in ACF and the Google images match. In addition, we reported one way to transform panoramic images to side view images for higher matching rates. The automatic detection results were compared and tested to resemble the pedestrian counts collected by field

Table 1
Results of reliability tests.

	Cronbach's alpha values
Google (machine) vs. Google (manual): Buffalo	0.702
Google (machine) vs. Google (manual): Washington, D.C.	0.724
Google (machine) vs. Google (manual): Boston	0.780
Google (machine) vs. Field Counts	0.758

work and manual Google Street View Images counts. The results suggested that the image detection method used in this study is capable of determining the presence of pedestrian with a reasonable level of accuracy.

This method can be used by planners, engineers, public health professionals and others to get an objective and large scale reliable estimate of pedestrian volume and potentially conduct environment audits by automatically detecting benches, trees, building shapes, etc. This information can inform safety analysis, walkability analysis such as design of sidewalks and other pedestrian facilities in responding to anticipated pedestrian volumes, and modeling and prediction of pedestrian traffic and how it interacts with land use development and transit development, etc. The detection method used in this study has limitations but with a potential to be improved as discussed and outlined as follows.

To further explore the use of Google Street View images for automatic pedestrian counts and environment audits, future work can benefit from information and technology development in two perspectives as summarized below. The first one relies on information that Google may release in the future and second one on the recent technology development in machine learning.

1. Current Google Street View images as it is published and made available to the public have two aspects of limitations as follow for pedestrian detection and counting, and environment audits:
   (1) Metadata such as street-view image collection time is not complete. For example, image collection time is accurate to month. Weather and time are two important factors on walking and biking choices. The current incomplete metadata made it difficult to match information extracted from the images with the weather or time data to do more research on how weather and time influence pedestrian volume. The U.S. city wide weather forecasts have currently been accurate to the granularity of hours; therefore, if acquisition time granularity of street-view can be published accurate to hours, the validity of the data for research use will be improved.
   (2) The Google Street View data currently are un-measurable images. They are images acquired by vehicle platform with Charged Coupled Device (CCD) sensor from different viewpoints. Because images of different angles have different geometric distortions, it results in difficulties to be registered with vector data. As Google has not provided measurable parameters of the street-view currently, these images cannot be used for geometric measurements directly. If the internal and external orientation elements of the image are released in the future, the stereo image pair can be constructed. In that case, pedestrians can be distinguished from the poles, trash, roadside advertising board more accurately and effectively, and the pedestrian detection accuracy can be potentially higher. Measurable images can also be used to conducted 3D-GIS based calculation and analysis, such as proportion of sky for the research on walkability (Yin, 2014).
2. In terms of the automatic detection method, the following two ways might help to improve the efficiency and pedestrian detection rate for future work:
   (1) This paper used the existing trained model for pedestrian detection. A new training sample created by manually annotating pedestrians in side view images downloaded and processed as discussed above to retrain the model parameters might help to improve the detection rate, especially when level 4 or higher images are to be used.
   (2) Deep learning is part of the family of machine learning methods and it is the most promising method in image recognition (Bengio & Vincent, 2013). Because of the rapid development in computing power due to the Graphics Processing Unit (GPU) parallel processing, deep Learning (Bengio & Vincent, 2013) have recently made some breakthrough results in machine learning field, such as speech recognition (Hinton et al., 2012), image recognition (Krizhevsky, Sutskever, & Hinton, 2012), etc. Comparing with current models, deep neural networks "reduce the word error rate by about 30%" (Bengio & Vincent, 2013) in four main benchmarks of the speech recognition, which is closely related to image recognition. With regard to image recognition, almost all the competition teams in the Large Scale Visual Recognition Challenge 2014 (ILSVRC2014)[2] at ImageNet used a deep architecture.

Deep learning can potentially help to reduce the error rate of pedestrian detection in two aspects. First, deep neural networks have a potential to learn the feature of pedestrian in images because they have proven to successfully learn the feature of natural images (Bengio & Vincent, 2013). Second, since deep architecture allows to labeling and recognizing a scene (Farabet et al. 2013; Pinheiro & Collobert, 2014), we can label skies, roads, vehicles, and buildings, which can be used to reduce the areas of pedestrians, for pedestrians cannot appear on the building or upon a car.

References

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 (June).
Arribas-Bel, D. (2014). Accidental, open and everywhere: emerging data sources for the understanding of cities. *Applied Geography, 49*, 45–53.
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.
Ben_Joseph, E., Lee, J. S., Cromley, E. K., Laden, F., & Troped, P. J. (2013). Virtual and actual: relative accuracy of on-site and web-based instruments in auditing the environment for physical activity. *Health and Place, 19*, 138–150.
Benenson, R., Omran, M., Hosang, J., & Schiele, B. (2014). *Ten years of pedestrian detection, what have we learned? In Computer Vision-ECCV 2014 Workshops (September)* (pp. 613–627). Springer International Publishing.
Bengio, Y. A., & Vincent, C. P. (2013). Representation learning: a review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 35*(8), 1798–1828.
Carr, L. J., Dunsiger, S. I., & Marcus, B. H. (2011). Validation of walk score for estimating access to walkable amenities. *British journal of Sports Medicine, 45*(14),

[2] http://image-net.org/challenges/LSVRC/2014/results.

1144–1148.
Chu, S. C. H. (2005). When and why do people walk in the city: the influence of urban elements on time-pattern of pedestrian movement. In *Proceedings of the sixth international walk 21 conference* (pp. 1–13).
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., Smith, A. D. L., & Rodriguez, D. (2007). The development and testing of an audit for the pedestrian environment. *Landscape and Urban Planning, 80*, 95–110.
Dalal, N., & Triggs, B. (2005). Histograms of oriented gradients for human detection. In *Computer vision and pattern recognition, 2005, IEEE computer society conference* (vol 1, pp. 886–893).
Dollar, P. (2015). *Piotr's computer vision matlab toolbox*. Access in Feb. 2015 http://vision.ucsd.edu/~pdollar/toolbox/doc/index.html.
Dollar, P., Appel, R., Belongie, S., & Perona, P. (2014). Fast feature pyramids for object detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 36*(8), 1532–1545.
Duncan, D. T., Aldstadt, J., Whalen, J., Melly, S. J., & Gortmaker, S. L. (2011). Validation of walk score for estimating neighborhood walkability: an analysis of four US metropolitan areas. *International Journal of Environmental Research and Public Health, 8*(12), 4160–4179.
Ercolano, J., Olson, J., & Spring, D. (1997). *Sketch-plan method for estimating pedestrian traffic for central business districts and suburban growth corridors*. Transportation Research Record 1578 (pp. 38–47). Washington D.C: Transportation Research Board.
Ewing, R., & Bartholomew, K. (2013). *Pedestrian and transit-oriented design*. Washington, D.C.: Urban Land Institute and American Planning Association.
Ewing, R., & Clemente, O. (2013). *Measuring Urban Design: Metrics for Livable Places*. Washington D.C.: Island Press.
Farabet, C., Couprie, C., Najman, L., & LeCun, Y. (2013). Learning hierarchical features for scene labeling. *Pattern Analysis and Machine Intelligence, IEEE Transactions on, 35*(8), 1915–1929.
Forsyth, A., Schmitz, K., Hearst, M. O., & Oakes, J. M. (2008). Design and destinations: factors influencing walking and total physical activity. *Urban Studies, 45*(9), 1973–1996.
Frank, L., Sallis, J., Conway, T., Chapman, J., & Saelens, B. (2006). Many pathways from land use to health: associations between neighbourhood walkability and active transportation, body mass Index and air quality. *Journal of the American Planning Association, 72*(1), 75–87.
Gallagher, A., & Chen, T. (2009). Understanding images of groups of people. In *Proc. CVPR*.
Goodchild, F. M. (2013). The quality of big (geo) data. *Dialogues in Human Geography, 3*(3), 280–284.
Hajrasouliha, A., & Yin, L. (2014). *The impact of street network connectivity on pedestrian movement*. Urban Studies Online.
Hinton, G., Deng, L., Yu, D., Dahl, G. E., Mohamed, A. R., Jaitly, N., et al. (2012). Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups, Signal Processing Magazine,. *IEEE, 29*(6), 82–97.
Kim, N. S., & Susilo, Y. O. (2013). Comparison of pedestrian trip generation models. *Journal of Advanced Transportation, 47*, 399–412.
Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097–1105).
Landis, B. W. (1996). The bicycle system performance measures: the intersection Hazard and latent demand score models. *ITE Journal, 66*(2), 18–26.
Lee, S., & Talen, E. (2014). Measuring walkability: A note on auditing methods. *Journal of Urban Design, 19*(3), 368–388.
Lury, C., & Wakeford, N. (Eds.). (2012). *Inventive methods: the happening of the social*. London, UK: Routledge.
O'Leary, D. E. (2013). 'Big data', the 'internet of things' and the 'internet of signs. *Intelligent Systems in Account, Finance and Management, 20*, 53–65.
Ozer, O., & Kubat, A. S. (2007). Walking initiatives: a quantitative movement analysis. In *Proceedings of the sixth internal space syntax symposium*. Istanbul: ITU Faculty of Architecture.
Pinheiro, P., & Collobert, R. (2014). *Recurrent convolutional neural networks for scene labeling*. JMLR.
Prioletti, A., Møgelmose, A., Grisleri, P., Trivedi, M. M., Broggi, A., & Moeslund, T. B. (2013). Part-based pedestrian detection and feature-based tracking for driver assistance: real-time, robust algorithms, and evaluation. *Intelligent Transportation Systems, IEEE Transaction on, 14*(3), 1346–1359.
Purciel, M., Neckerman, K. M., Lovasi, G. S., Quinn, J. W., Weiss, C., Bader, M. D. M., et al. (2009). Creating and validating GIS measures of urban design for health research. *Journal of Environmental Psychology, 29*(4), 457–466.
Rundle, A. G., Bader, M. D. M., Richards, C. A., Neckerman, K. M., & Teitler, J. O. (2011). Using Google Street View to audit neighborhood environments. *American Journal of Preventive Medicine, 40*, 94–100.
Ruppert, E. (2013). Rethinking empirical social sciences. *Dialogues in Human Geography, 3*(3), 268–273.
Scheneider, R. J., Arnold, L. S., & Ragland, D. D. (2009). Pilot model for estimating pedestrian intersection crossing volumes. *Transportation Research Record: Journal of the Transportation Research Board, 2140*, 13–26.
Watts, D. J. (2007). A twenty-first century science. *Nature, 445*, 489.
Yin, L. (2013). Assessing walkability in the city of Buffalo: an application of agent-based simulation. *Journal of Urban Planning and Development, 139*(3), 166–175.
Yin, L. (2014). Review of Ewing, reid and Otto clemente. 2013. Measuring urban design: metrics for livable places. *Journal of Planning Literature, 29*(3), 273–274.
Yin, L., Raja, S., Li, X., Lai, Y., Epstein, L. H., & Roemmich, J. N. (2013). Neighborhood for playing: using GPS, GIS, and accelerometry to delineate areas within which youth are physically active. *Urban Studies, 50*(14), 2922–2939.