



## Automatic concrete sidewalk deficiency detection and mapping with deep learning

Yuhan Jiang<sup>a</sup>, Sisi Han<sup>b</sup>, Dapeng Li<sup>c</sup>, Yong Bai<sup>d</sup>, Mingzhu Wang<sup>e,\*</sup>

<sup>a</sup> Department of Built Environment, North Carolina A&T State University, Greensboro, NC 27411, United States

<sup>b</sup> Department of Civil, Construction and Environmental Engineering, Marquette University, Milwaukee, WI 53233, United States

<sup>c</sup> Department of Geography and Geospatial Sciences, South Dakota State University, Brookings, SD 57007, United States

<sup>d</sup> Department of Civil, Construction and Environmental Engineering, Marquette University, Milwaukee, WI 53233, United States

<sup>e</sup> School of Architecture, Building and Civil Engineering, Loughborough University, Loughborough, United Kingdom



### ARTICLE INFO

#### Keywords:

Computer vision  
Deep learning  
Semantic segmentation  
Concrete joint detection  
Sidewalk deficiency  
Point cloud

### ABSTRACT

Vertical displacement is a common concrete slab sidewalk deficiency, which may cause trip hazards and reduce wheelchair accessibility. This paper presents an automatic approach for trip hazard detection and mapping based on deep learning. A low-cost mobile LiDAR scanner was used to obtain full-width as-is conditions of sidewalks, after which a method was developed to convert the scanned 3D point clouds into 2D RGB orthoimages and elevation images. Then, a deep learning-based model was developed for pixelwise segmentation of concrete slab joints. Algorithms were developed to extract different types of joints of straight and curved sidewalks from the segmented images. Vertical displacement was evaluated by measuring elevation differences of adjacent concrete slab edges parallel to the boundaries of joints, based on which potential trip hazards were identified. In the end, the detected trip hazards and normal sidewalk joints were geo-visualized with specific information on Web GIS. Experiments demonstrated the proposed approach performed well for segmenting joints from images, with a highest segmentation IoU (Intersection over Union) of 0.88, and achieved similar results compared with manual assessment for detecting and mapping trip hazards but with a higher efficiency. The developed approach is cost- and time-effective, which is expected to enhance sidewalk assessment and improve sidewalk safety for the general public.

### 1. Introduction

Public sidewalks are essential infrastructures in cities to provide convenience for urban life. Deficiencies of sidewalks will lead to inconveniences, disruptions and potential hazards to residents. Hence, it is important to monitor and evaluate sidewalk condition such as to take necessary maintenance measures to ensure the normal functionality of sidewalks. To ensure public sidewalks remain in good conditions, local governments usually have their own sidewalk program to assist private property owners (who are the maintaining authority of the sidewalk adjacent to their property) with concrete slab evaluation and defect correction. The typical traditional approach for sidewalk surveying is using smart-level and measuring tools, e.g. tapes, to manually take slope readings and evaluating the compliance with related regulations. However, such manual surveying method takes a long time to assess

overall conditions of sidewalks, for example, the City of Middleton and the Village of Shorewood both require eight years to go through each neighborhood ([City of Middleton, 2021b](#); [Village of Shorewood, 2021](#)).

Particularly, vertical displacement, also known as vertical fault, is a common concrete slab sidewalk deficiency, which may cause tripping hazards and reduce wheelchair accessibility ([CCRPC, 2016](#); [City of Sioux Falls, 2017](#); [Yates, Fouts, Sehgal, & Mcren, 2017](#)). To be compliant with Americans with Disabilities Act (ADA), any vertical displacement of new sidewalk concrete slabs (at joints) must be less than 13 mm (1/2 in.) ([Department of Justice, 2010](#)). Local governments have different criterions (e.g., the listed criterions in [Table 1](#)) considering the repair cost and budgets to decide the maintenance actions to take. In general, grinding should be performed to correct the trip hazard, when a joint or crack has a vertical displacement between 13 mm (1/2 in.) and 3.81 cm (1½ inches); otherwise, replacement would be the best method to

\* Corresponding author.

E-mail addresses: [jyh100@live.com](mailto:jyh100@live.com) (Y. Jiang), [sisi.han@marquette.edu](mailto:sisi.han@marquette.edu) (S. Han), [dapeng.li@sdstate.edu](mailto:dapeng.li@sdstate.edu) (D. Li), [yong.bai@marquette.edu](mailto:yong.bai@marquette.edu) (Y. Bai), [M.Wang@lboro.ac.uk](mailto:M.Wang@lboro.ac.uk) (M. Wang).

**Table 1**

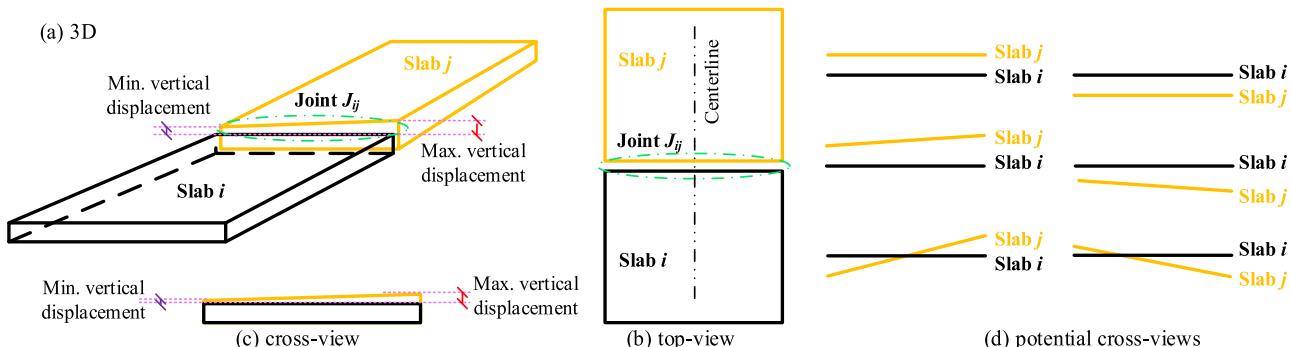
Vertical displacement (changes in level) regulations.

Sources	Criterion of vertical displacement (changes in level)	Comments
2010 ADA Standards for Accessible Design (Department of Justice, 2010, 2010)	<i>Section 303 Changes in Level</i> Changes in level of 6.4 mm (1/4 in.) high maximum shall be permitted to be vertical. Changes in level between 6.4 mm (1/4 in.) high minimum and 13 mm (1/2 in.) high maximum shall be beveled with a slope not steeper than 1:2. Changes in level greater than 13 mm (1/2 in.) high shall be ramped.	For design and construction
(Proposed) Public Rights-of-Way Accessibility Guidelines (Access Board, 2013)	<i>R302.7.2 Vertical Surface Discontinuities</i> Vertical surface discontinuities shall be 13 mm (1/2 in.) maximum. Vertical surface discontinuities between 6.4 mm (1/4 in.) and 13 mm (1/2 in.) shall be beveled with a slope not steeper than 50 percent. The bevel shall be applied across the entire vertical surface discontinuity.	
Sidewalk Deficiencies Examples, City of Middleton, WI (City of Middleton, 2021a)	At a crack or joint, has a vertical displacement greater than 13 mm (1/2 in.) height but less than 38.1 mm (1 ½ inches), the available repair options are: Saw Cut high edge, Mud Jack low edge, or Replacement.	For maintenance and replacement
Policy No. 36, Concrete Replacement Criteria, Village of Shorewood, WI (Board, 2019)	All concrete slabs that have cracks (including control joints) with differential settlement of greater than 19.05 mm (3/4 in.) which may cause tripping. A concrete slab is eligible for grinding to correct a trip hazard(s) if the vertical displacement is between 19.05 and 38.1 mm (3/4 and 1 ½ inches). This criterion does not require replacement of all concrete sidewalk that does not meet ADA requirements for change in level of 6.4 mm (1/4 in.) due to the significant amount of sidewalk replacement which would be required to meet these criteria.	For maintenance and replacement

mitigate trip hazards on public sidewalks (City of Middleton, 2021a; Board, 2019). Obviously, manually evaluating sidewalk condition based on such detailed criterions is very time-consuming and labour intensive. In addition, there is a lack of comprehensive and updated database of sidewalk features and condition. Some cities spent several months in a one-time Sidewalk Network Inventory and Assessment project to create a comprehensive database of sidewalk network features within the cities' urbanized areas (CCRPC, 2016; City of Sioux Falls, 2017). However, as time passed, these databases become too outdated to reflect the sidewalks' as-is conditions.

Therefore, automated sidewalk surveying methods are desired to improve the surveying efficiency and alleviate manual workloads. Previous studies have proposed using the Ultra-Light Inertial Profiler (ULIP) (Cole, 2013), which is a Segway-based sensor and acquisition system. In addition, ULIPr (which is the RoLine 1130 laser line scan sensor version of ULIP) was designed to capture a 3D representation of the travel surface (Starodub Inc, 2009a). Nevertheless, both ULIP and ULIPr have limited coverage of the sidewalk, making them likely to miss the vertical displacement between the sidewalk slabs. Pose estimation sensors such as IMU are also utilized to measure the condition of sidewalks, which however also suffer from the same problem of limited coverage of the object (Kim & Ahn, 2016; Kim, Ahn, & Yang, 2016).

To address limitations of previous studies and automatically detect and geo-visualize sidewalk trip hazards, this study proposed a cost-and time-effective approach that integrates mobile devices, deep learning, and geographic information systems (GIS) in the scanning phase, data processing phase, and trip hazard mapping phase. In the scanning phase, the as-is conditions of sidewalk sections are scanned using mobile devices, in which both low-cost LiDAR (light detection and ranging) scanner (which determines distances by targeting a surface with a laser and measuring the time for the reflected light to return to the receiver) (Wikipedia, 2022a) and SfM (structure from motion) photogrammetry (which estimates 3D structures from 2D image sequences) (Wikipedia, 2022b) are supported. In the processing phase, a method is developed to convert the obtained point cloud data to feature images, after which a deep learning-based segmentation model is adopted for precise concrete slab joint detection, which showed good performances in building and infrastructure defect detection (Jiang, Han, & Bai, 2021a). In addition, this study proposed a concrete slab joint extraction and vertical displacement measurement algorithm to extract joints from the segmented images of both straight and curved sidewalks with straight and oblique joints. After that, the elevation differences of adjacent slab edges are measured, while the detected trip hazards are marked at the higher edges with wavy lines and the calculated displacement values. In the mapping phase, all sidewalk concrete slab joints are mapped in the Web GIS platform with measured vertical displacement values and attached annotated images. All joints are classified as *trip hazard* and *normal* as well. Furthermore, comprehensive experiments were conducted to evaluate the developed sidewalk trip hazard detection and geo-visualization approach in three communities, including University

**Fig. 1.** Sketch of sidewalk concrete slabs at joint  $J_{ij}$ .

**Table 2**

Comparison of sidewalk surveying approaches.

Performance	Most common engineering method (CCRPC, 2016; City of Sioux Falls, 2017)	ULIP/ ULIPr (Starodub Inc, 2009a, 2009b; Yates, Fouts, Sehgal, & Mcren, 2017)	Pose estimation sensors (Frackelton et al., 2013; Kim & Ahn, 2016; Kim et al., 2016)	Camera and SfM photogrammetry (Tested in this paper)	Low-cost LiDAR scanner (Proposed in this paper, Fig. 2)
Field equipment	Smart-level (i.e., smart slope meter), and Measuring wheel (i.e., measuring tape), Table computer (e.g., iPad, Google Nexus), and Esri ArcGIS Collector application	ULIP is a Segway-based sensor and acquisition system. ULIPr is a RoLine 1130 laser line scan sensor version of ULIP	Smart phone (or table computer) with an inertial measurement unit (IMU) sensor (e.g., accelerometer and gyroscope) and a global positioning system (GPS)	Smart phone (or digital camera)	iPad Pro (iPhone Pro, or other table computers and smart phones with LiDAR sensor or depth camera)
Raw data and coverage	Points with measured information	Profile of the travel path. Approximate 100 mm width 3D representation of the travel surface	Signal vector magnitude (SVM) of the walked path	Numbers of highly overlapped images of the scanned full-width sidewalk surfaces	Dense 3D point clouds of the scanned full-width sidewalk surfaces
Limitation	Labor-intensive, inaccurate, time-consuming.	Limited coverage, which means a high chance of skipping maximum displacement caused by sedimentation	Limited coverage, which means a high chance of skipping maximum displacement caused by sedimentation	It requires a long image processing and 3D reconstruction time to get point clouds of the scanned concrete slabs. Resolutions are limited at 1 cm/pixel.	

of Wisconsin-Milwaukee (UWM), Shorewood Village, and South Dakota State University (SDSU).

The remaining of the paper is organized as follows. Section 2 reviews related works in this area. Section 3 introduces the proposed approach for detecting and mapping trip hazards. Section 4 presents the experiments and results, followed by Conclusion in Section 5.

## 2. Related works

### 2.1. Sidewalk surveying and assessment

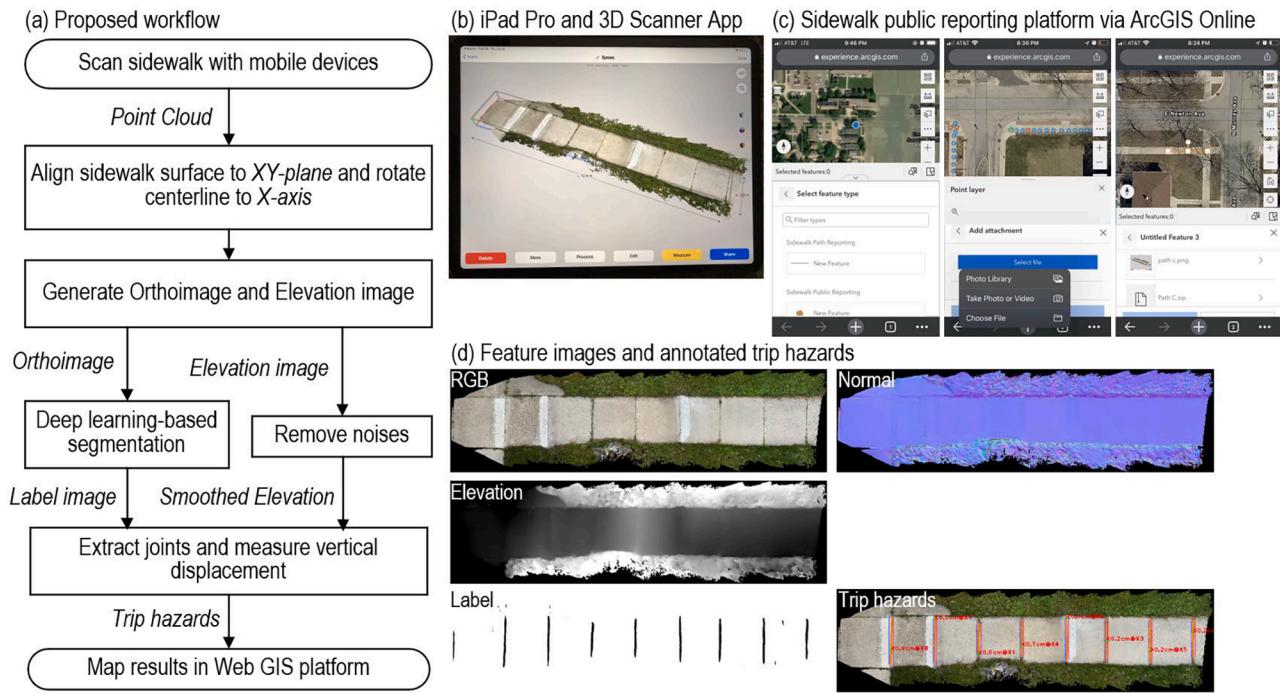
Currently, the most common engineering approach is using smart-level and measuring tape to manually take slope readings. This inspector-led practice is very labor-intensive and time-consuming, taking six times longer than the Ultra-Light Inertial Profiler (ULIP)-based approach (Cole, 2013). ULIP is a Segway-based sensor and acquisition system, which can assess the sidewalk by recording a profile of the travel path on the sidewalk. Meanwhile, ULIPr is a RoLine 1130 laser line scan sensor version of ULIP. However, even the most powerful ULIPr designed to capture a 3D representation of the travel surface (Starodub Inc, 2009a) only has a narrow coverage (approximately 100 mm) in width, which means a high chance of skipping maximum displacement caused by sedimentation. In addition, the pose estimation sensors-based approaches (Kim & Ahn, 2016; Kim et al., 2016) were also used for assessing sidewalks. They generate signal vector magnitude of the walked path using smart phone (or table computer) with an inertial measurement unit (IMU) sensor (e.g., accelerometer and gyroscope) and a global positioning system (GPS). Nevertheless, such method also suffers from the problem of limited coverage of the sidewalk,. As shown in Fig. 1, sedimentation of concrete slabs may cause different vertical displacement along the joint of two adjacent concrete slabs. Thus, the walked path- and travel surface-based approaches would skip the maximum vertical displacement when it is located at the end of a joint like the bottom four conditions shown in Fig. 1(d). To evaluate the elevation of the sidewalk in full-width coverage, 3D dense point clouds can potentially provide more comprehensive and accurate information. One method for sidewalk scanning is directly obtaining point cloud using laser scanning (LiDAR) while another is reconstructing 3D point cloud from images or videos using 3D reconstruction approaches such as

SfM photogrammetry. Both methods have been previously applied in roadway pavement evaluation (Edmondson et al., 2019; Gézero & Antunes, 2019; Li, Cheng, Kwan, Tong, & Tian, 2019; Roberts, Inzerillo, & Di Mino, 2020), indicating their potential applicability for sidewalk surveying as well. Table 2 summarizes the existing and potential approaches for sidewalk condition surveying.

However, SfM photogrammetry requires much more time in sidewalk scanning and point cloud acquisition due to two main reasons: (a) highly overlapped high-resolution images are essential raw data for SfM, which needs the camera to remain a small distance away from sidewalk surfaces, and slowly move over multiple paths to guarantee overlaps and full coverage; (b) image processing and 3D reconstruction are slowed down due to the number of images, for example, cloud processing one hundred images needs approximately 50 to 70 min by Autodesk ReCap Photo, which may need additional time for waiting in the queue before the processing (Jiang & Bai, 2021). In contrast, LiDAR can obtain dense point clouds for the scanned object immediately. It is the most effective technique for capturing 3D reality data if its price drops down. Fortunately, LiDAR sensors are built-in components for some relatively low-cost mobile devices such as iPad Pro and iPhone Pro. Thus, this study utilized and tested the low-cost LiDAR scanner for sidewalk as-is condition scanning while the SfM photogrammetry was tested for sidewalk scanning with a mobile phone without LiDAR sensor as well.

### 2.2. Infrastructure defect detection and assessment using deep learning

The sidewalk (concrete slabs) surfaces (see Fig. 1) are similar to the roadway pavement surfaces, both of which are relatively flat planes. It is feasible to simplify (convert) a 3D point cloud to a 3D image to represent the elevation feature of the sidewalk concrete slab surface. Similar idea has been applied in previous studies, such as the surface height plot (Edmondson et al., 2019), depth map (Roberts et al., 2020), range image (Zhou & Song, 2020a, Zhou and Song, 2020b) and elevation map (Jiang and Bai, 2021, 2020a, 2020b). In addition, the corresponding 2D image, e.g., top-view and drone photogrammetric orthophoto, can provide the spectral features (Red, Green, Blue) at the same pixel coordinates (Dadrasjavan, Zarrinpanjeh, Ameri, Engineering, & Branch, 2019; Jiang, Bai, & Han, 2020; Li et al., 2019). With the 2D or 3D images, deep learning approaches such as convolutional neural networks (CNNs)-



**Fig. 2.** The proposed automated trip hazard detection approach. (Map Data © 2021 Esri).

**Table 3**  
Parameters of scanning.

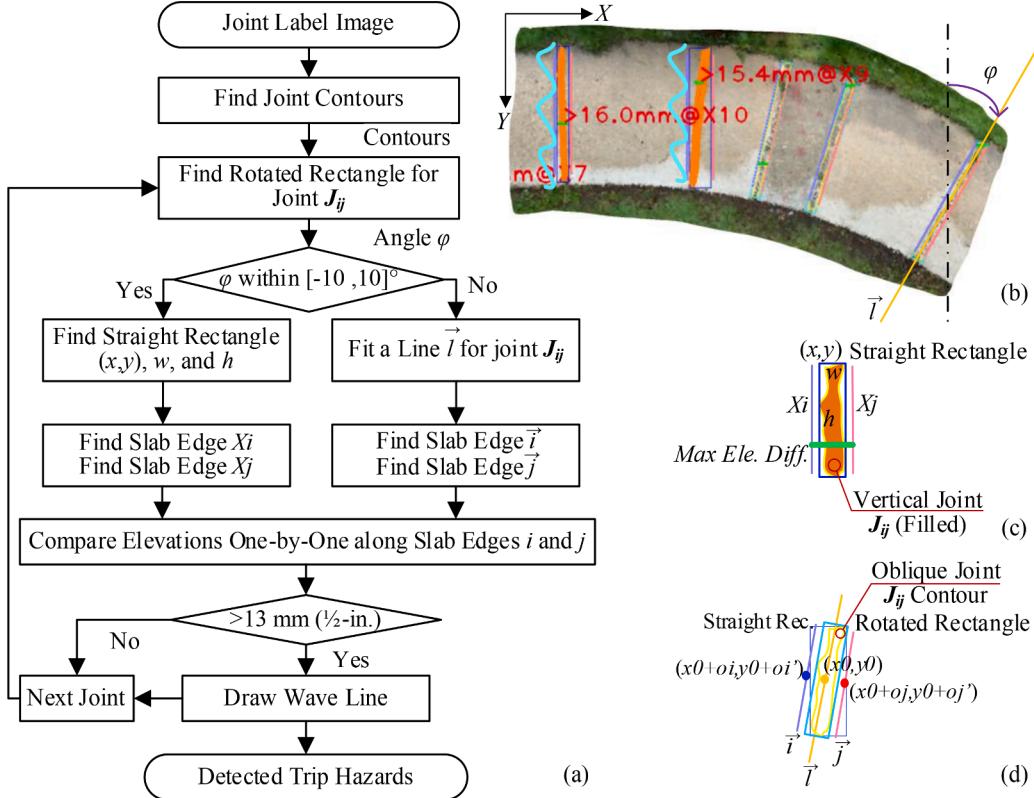
Parameters	Options	Instructions by ( <a href="#">Labs, 2021</a> )	This paper
Confidence	[Low, Medium, High]	Options for thresholding the data coming in from the sensor. [High] only keeps the best quality data, but reduces the amount of data available.	Medium
Range	0.3 m to 5.0 m	Discards LiDAR data after a certain distance. Limiting range reduces scan size and increases accuracy.	3.0 m
Resolution	5 mm to 20 mm	Lower values mean higher resolution, but also limited scan size.	10 mm and 5 mm
Masking	[None, Object, Person]	This feature masks LiDAR data based on the type of object in view. [Object] attempts to isolate prominent objects in view. For best results, keep the object fully in view, and use on a simple background without other clutter in view.	None
Point cloud export	[Low, Medium, High]	Allow the model to be shared as USDZ, GLB, GLTF, OBJ, DAE, STL, Point Cloud, or to Sketchfab.	High density LAS file.

based classification models (Ali, Valappil, Kareem, John, & Al Jassmi, 2019; Cheng & Wang, 2018; Fan et al., 2019; Jiang, Han, & Bai, 2021b; Maniat, 2019; Protopapadakis, Voulodimos, Doulamis, Doulamis, & Stathaki, 2019; Tan, Cai, Li, Chen, & Wang, 2021; Yang, Shi, Chen, & Lin, 2020; Zhou & Song, 2020a) and fully convolutional networks (FCNs)-based segmentation methods (Alipour, Harris, & Miller, 2019; Augustaukas & Lipnickas, 2019; Dung & Anh, 2019; Ji, Xue, Wang, Luo, & Xue, 2020; Jiang et al., 2021a; Jiang et al., 2021b; Liu, Cao, Wang, & Wang, 2019; Song, Jia, Zhu, Jia, & Gao, 2020; Zou et al., 2019) have

been used for roadway pavement distress detection. The successful applications of deep learning-based distress detection in previous studies making them potential methods for detecting sidewalk concrete slab joints and deficiencies as well.

To evaluate defects and structure condition, one method is to precisely extract boundaries of target objects (e.g., cracks). Previous studies adopted the deep learning based image semantic segmentation models such as FCN (Alipour et al., 2019; Shelhamer, Long, & Darrell, 2017), U-Net (Liu et al., 2019; Ronneberger, Fischer, & Brox, 2015), SegNet (Badrinarayanan, Kendall, & Cipolla, 2017; Kearney, Coops, Sethi, & Stenhouse, 2020), PSPNet (Zhao, Shi, Qi, Wang, & Jia, 2017) and DeepLabv3+ (Chen, Zhu, Papandreou, Schroff, & Adam, 2018; Ji et al., 2020) for crack detection on bridge and pavement surface. Some studies also developed specific deep learning models such as DeepCrack (Zou et al., 2019), CrackSeg (Song, Jia, Zhu, Jia, & Gao, 2020) and DilatSeg-CRF (Wang & Cheng, 2020) for crack detection with 2D images. The comparisons showed the performances of the above models were not significantly different in the CrackDataset (Song et al., 2020). However, the U-Net can reach a higher accuracy with fewer training data sets of images and labels (Liu et al., 2019; Zhang, Zhang, & Cheng, 2020), and has a good performance in thin cracks detection like in (Augustaukas & Lipnickas, 2019; Jiang et al., 2021b, 2021a; Majidifard, Adu-Gyamfi, & Buttlar, 2020). Hence, U-Net was used as an example segmentation model in this study.

In concrete slabs, temperature changes can result in concrete expanding or shrinking, thus joints are commonly designed and created by forming, tooling, sawing, and placing joint formers to prevent cracks when the concrete shrinks (Palmer, 2020; Rodriguez, 2021). Sidewalk concrete slab joints can be classified as contraction (control) joint, isolation (expansion) joint, and construction joint. They are normally in a straight-line shape, as shown in Fig. 1(b), ranging from 3 mm to 20 mm with different creating methods. As joints have similar features as



**Fig. 3.** The proposed algorithm for joint extraction and vertical displacement measurement.

cracks, this research explored deep learning-based segmentation model (e.g. U-Net) for sidewalk joint detection and segmentation.

### 3. The proposed automated trip hazard detection approach

To improve the efficiency in sidewalk assessment, this study proposed an approach of automatic trip hazard detection and mapping. Fig. 2(a) presents the procedure of the proposed approach. There are four main steps involved in the whole process, including (1) sidewalk scanning using mobile devices and feature image generation, (2) deep learning-based image segmentation, (3) joint extraction and vertical displacement measurement, and (4) trip hazard mapping and geovisualization in the Web GIS platform. Details of each step are introduced in the following.

#### 3.1. Sidewalk scanning and feature image generation

Mobile devices such as iPad Pro (or iPhone Pro) with LiDAR sensor are used for scanning sidewalks in this study to obtain the point cloud data. 3D scanning tools such as the 3D Scanner App (Labs, 2021) is used to acquire the 3D point cloud of the sidewalk in real-time. Table 3 lists the parameters of the high-resolution mode of 3D Scanner App used in this study. After scanning, the data will be processed to generate the textured point cloud in the 3D Scanner App, see Fig. 2(b). Then, the point cloud can be exported as a LAS file.

Moreover, this study proposed a sidewalk public reporting platform via a Web GIS system (i.e. ArcGIS online) for property owners or concerned citizens to report sidewalk trip hazards with their surveyed results. Users can place start and end points, paths, and regions to mark the

scanned sidewalks in the reporting platform, where location service is enabled to fast and accurately find the user's current location as shown in the left screenshot of Fig. 2(c). The scanned files can be uploaded to the reporting platform as shown in the right two screenshots of Fig. 2(c). As a result, the local governments can easily review the reported cases and assign teams to inspect or repair the concrete slabs.

Following that, the authors proposed a *pointcloud2orthoimage* algorithm and developed the tool (code and demo in (Jiang, 2022)) to automatically convert a sidewalk point cloud to feature images, e.g., orthoimage and elevation image, see Fig. 2(d). The algorithm has the following key processes: (a) Find a plane in the point cloud (in general, the plane contains most of the sidewalk surface points); (b) Calculate the rotation angle between the normal and Z-axis of the plane, and use the angle to rotate the point cloud such as to align the sidewalk surface with the XY-plane; (c) Find an oriented bounding box for the point cloud (with a rotation matrix  $R$ ), and use the inverse matrix  $R^{-1}$  to rotate the point cloud and make the sidewalk centerline codirect to the X-axis; and, (d) Translate the point cloud and make its center close to (0,0,0) if necessary. The developed tool utilized the plane segmentation and oriented bounding box functions in Open3D to obtain geometry information (Open3D, 2020). The generated point cloud feature images have unlimited size, which means any scanned long sidewalk can be presented in a continuous high-resolution image. Meanwhile, the differently sized RGB feature images also have the same geospatial resolution, known as ground sampling distance (GSD).

#### 3.2. Deep Learning-based image segmentation

After generating sidewalk feature images, a deep learning model was

**Table 4**

Parameters of the proposed algorithm for joint extraction and vertical displacement measurement.

Parameter	Definition	Comments
$(X, Y)$	Pixel coordinate, origin is top-left corner of a feature image.	Horizontal axis $X$ has positive direction in the right. Vertical axis $Y$ has positive in downward.
$\varphi$	Angle of joint $J_{ij}$ direction, positive in clockwise	An approximate straight joint $J_{ij}$ has $-10^\circ \leq \varphi \leq 10^\circ$ ; An oblique joint $J_{ij}$ has $\varphi < -10^\circ$ or $\varphi > 10^\circ$
$(x, y)$	Coordinate of top-left corner of a straight bounding box.	A straight bounding box that contains all pixels for a joint $J_{ij}$ .
$w, h$	Width and height of a straight bounding box	
$\{X_i, X_j\}, \{Y_i, Y_j\}$	Edges of left and right concrete slabs $i$ and $j$ around a joint $J_{ij}$	Left slab edge $\begin{cases} X_i = x - 3 \\ y + 10 < Y < y + h - 10 \end{cases}$ Right slab edge $\begin{cases} X_j = x + w + 3 \\ y + 10 < Y < y + h - 10 \end{cases}$ where 3 is used to offset the bounding box edges; and 10 is used to offset the sidewalk edges to avoid measuring vegetation.
$\vec{l}$	Line of an oblique joint $J_{ij}$ , which has the equation $(Y - y_0)/(X - x_0) = V_y/V_x$	Joint $J_{ij}$ has a center $(x_0, y_0)$ , and a normalized vector $(V_x, V_y)$ , which is the normalized vector collinear to the line $\vec{l}$ .
$\vec{i}$	Line of left slab edge, which is parallel to $\vec{l}$	Left slab edge has a center $(x_0 + o_i, y_0 + o_i')$ , right slab edge has a center $(x_0 + o_j, y_0 + o_j')$ , and $X$ has a range of $(-L/2, L/2)$ , where $o_i, o_i', o_j, o_j'$ are small values used to offset slab edges from the joint fitted line $\vec{l}$ , and $L = \sqrt{b^2 + w^2}$ .
$\vec{j}$	Line of right slab edge, which is parallel to $\vec{l}$	

proposed for pixelwise segmentation of sidewalk joints in the image. During model training and prediction stages, the disassembling and assembling algorithm presented in (Jiang et al., 2021b) was used because the dimension of a feature image may be too large to be processed by a workstation. Specifically, the algorithm first disassembled a large-resolution feature image into multiple small-patches with a dimension of  $128 \times 128$ -pixel, each of which overlaps 50% with the adjacent small-patches in both width and height directions. Then, the disassembled small-patches were fed into the deep learning segmentation model rather than directly using the large-resolution input, and correspondingly small-patch outputs would be generated by the model with segmented labels. In the end, the small-patch outputs were assembled to produce a segmented image with the same dimension as the original input image.

After segmentation, the 1-channel segmented label image has a pixel value range of 0 to 255 which is obtained by multiplying 255 with the results of Sigmoid (i.e. the activation function in the end layer of the segmentation model), which are in the range of 0 to 1. Then, any pixel with a value less than the threshold (which is  $255/2 = 127$ ) was updated to 0 to indicate the joint, otherwise, replaced with 255 to represent non-joint objects. As a result, the trained model can be utilized for label image generation (i.e. segmentation) for large-resolution feature image inputs as shown in Fig. 2(d).

### 3.3. Joint extraction and vertical displacement measurement

Typically, sidewalk concrete slabs are in a rectangular shape, and joints are perpendicular to the sidewalk centerline as shown in Fig. 1(b). Once the sidewalk path is rotated into the horizontal direction, the joints are in vertical directions in the feature image as shown in Fig. 2(d). In addition, the special case of a curved sidewalk is shown in Fig. 3(b), where joints are perpendicular to the centerline, while oblique joints have an angle  $\varphi$  to the vertical direction. An example of vertical

(straight) joint and oblique joint is illustrated in Fig. 3(c) and Fig. 3(d) respectively. To accurately measure the vertical displacement between adjacent concrete slabs, this study proposed a joint extraction and vertical displacement measurement algorithm to process both vertical (straight) joints and oblique joints. Details of the algorithm are shown in Fig. 3(a).

Table 4 summarizes the parameters of the proposed algorithm, which uses a joint label image as input. Since joints are scatter instances in the label image, edges of joints are easy to determine as contours. Then, the ensuing steps and geometry information as follows are used to process a joint  $J_{ij}$ .

(1) Find a rotated bounding rectangle for the extracted joint contour  $J_{ij}$ , where a center of  $(x_0, y_0)$  and an angle  $\varphi$  (compared to the vertical direction) are returned (see Fig. 3 (b)).

(2a) If angle  $\varphi$  falls in a range of  $[-10^\circ, 10^\circ]$ , the joint is classified as a vertical joint. Find a straight bounding rectangle for joint  $J_{ij}$ , which has the top-left corner  $(x, y)$ , width  $w$  and height  $h$  (see Fig. 3 (c)).

(2b) If the absolute value of angle  $\varphi$  is larger than  $10^\circ$ , the joint is classified as an oblique joint. Fit a line  $\vec{l}$  for the joint  $J_{ij}$ , and also find a straight bounding rectangle for joint  $J_{ij}$  (see Fig. 3 (d)). The fitted straight line would pass through the joint, and have the minimum sum of distances to all points of the joint.

(3a) For an approximate vertical joint  $J_{ij}$ , the concrete slab edges  $i$  and  $j$  are set as two-line segments with an offset to the straight rectangle (offset = 3-pixel) and have the pixel length of  $h=20$  (see Fig. 3 (c)).

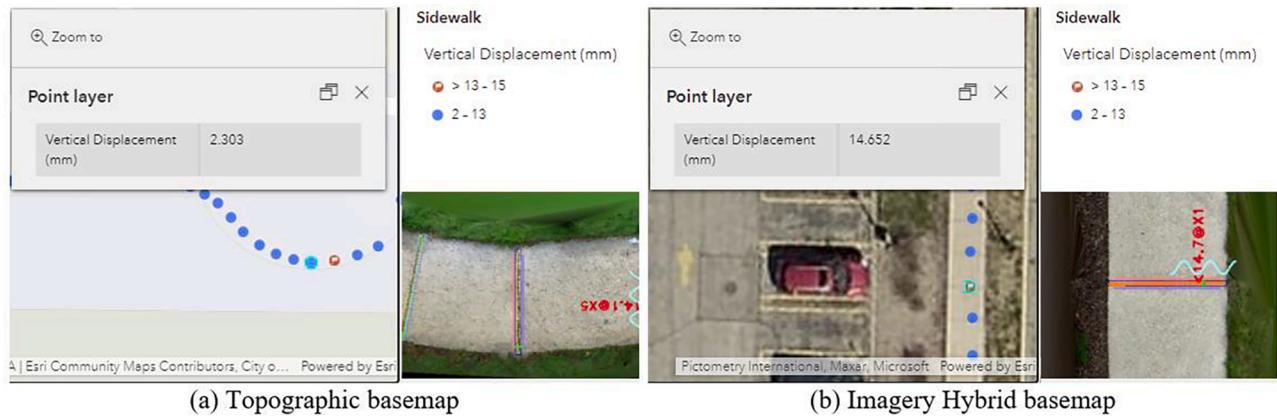
(3b) For an oblique joint  $J_{ij}$ , with a large  $\varphi$ , the concrete slab edges  $\vec{i}$  and  $\vec{j}$  are codirected to the line  $\vec{l}$ , and have the middle point  $(x_0 + o_i, y_0 + o_i')$  and  $(x_0 + o_j, y_0 + o_j')$ , respectively (see Fig. 3 (d)), where  $[o_i, o_i', o_j, o_j']$  are small values designed to offset slab edges from the fitted line of the oblique joint  $J_{ij}$ . For example, assume  $(V_x, V_y)$  is the normalized vector collinear to the line  $\vec{l}$ , if  $V_y/V_x$  greater than 0, set them as [-6, 3, 6, -3], or if  $V_y/V_x$  less than 0 (in Fig. 3 (d)), set them as [-6, -3, 6, 3].

(4) For a given  $(X, Y)$ , calculate elevation difference of the corresponding points on the two slab edges. Elevations of points can be obtained from the corresponding elevation image based on pixel coordinates of points. If the maximum elevation difference exceeds the vertical displacement criterion of 13 mm (1/2 in.), then fill joint  $J_{ij}$  (like Fig. 3 (c)) and annotate the maximum elevation difference at the corresponding location with a numerical value and a wavy line to create the annotation of trip hazards in the image like Fig. 3 (b).

Additionally, to automatically execute the proposed algorithm, image processing techniques are utilized to extract all individual concrete slab joints in a pixelwise segmented label image, as well as the contour features of straight bounding rectangle and rotated rectangle. Following that, the geometry information (e.g. center point, width, height, etc.) of joints is obtained by fitting functions (e.g. “fitting a line” (OpenCV, 2021a)) based on the extracted contours.

### 3.4. Trip hazards mapping and geo-visualization

To better visualize and manage the potential trip hazards, a method was proposed to map the detection results (joints and vertical displacements) to a Web GIS platform. The method was based on the image with annotated trip hazards and GPS coordinates of the scanning starting and ending point. If the sidewalk path is arbitrarily scanned, there are two scenarios for the sidewalk, including: Scenario 1, sidewalk (nearly) along the west-east direction, which has a longitude difference i.e.  $|Longitude_{end} - Longitude_{start}|$  larger than the latitude difference i.e.  $|Latitude_{end} - Latitude_{start}|$ ; and, Scenario 2, sidewalk (nearly) along the south-north direction, which has a latitude difference larger than the longitude difference. In addition, by considering orders of the scanning start and end points, the two scenarios can be classified into four cases, including: Case1, scan in West-East direction; Case 2, scan in East-West



**Fig. 4.** Examples of the mapped joint and trip hazards with annotated image segments. (Map Data © 2021 Esri).



**Fig. 5.** Experimental sidewalk paths. (Map Data © 2021 Google.).

direction; Case 3, scan in South-North direction; and Case 4, scan in North-South direction. The corresponding case is determined by comparing the longitude and latitude GPS coordinates of the start and

For Case 2 (East-West), the GPS coordinates of a joint  $J_{ij}$  can be determined with Eq. (1b),

$$\begin{cases} \text{Longitude}_{J_{ij}} = \left\{ \begin{array}{l} \text{Longitude}_{start} + (\text{Longitude}_{end} - \text{Longitude}_{start}) \times \text{ratio}_x \\ \text{Latitude}_{start} + (\text{Latitude}_{end} - \text{Latitude}_{start}) \times (\text{ratio}_x - 0.5 + \text{ratio}_y) \end{array} \right. \end{cases} \quad (1b)$$

end points. Then, GPS coordinates of trip hazards and sidewalk joints are determined using equations as follows.

For Case 1 (West-East), the GPS coordinates of a joint  $J_{ij}$  can be determined via Eq. (1a),

$$\begin{cases} \text{Longitude}_{J_{ij}} = \left\{ \begin{array}{l} \text{Longitude}_{start} + (\text{Longitude}_{end} - \text{Longitude}_{start}) \times \text{ratio}_x \\ \text{Latitude}_{start} + (\text{Latitude}_{end} - \text{Latitude}_{start}) \times (\text{ratio}_x - 0.5 + 1 - \text{ratio}_y) \end{array} \right. \end{cases} \quad (1a)$$

For Case 3 (South-North), the GPS coordinates of a joint  $J_{ij}$  can be determined with Eq. (1c),

**Table 5**  
Training and validation data for the segmentation model.

Dataset	RGB	RGB + Normal
Training	72,725	56,055
Validation	8,081	6,229
Total	80,806	62,284

where,  $\begin{cases} ratio_x = \frac{x0/ImageWidth}{1e} \\ ratio_y = \frac{y0/ImageHeight}{1e} \end{cases}$  (1e), and  $(x0, y0)$  are pixel coordinates of the middle points of joint  $J_{ij}$  (see Fig. 3(d)).

After obtaining the coordinates of trip hazards and sidewalk joints, Web GIS platform is utilized to geo-visualize the sidewalk assessment results, in which sidewalk concrete slab joints are mapped in a point layer with the longitudes and latitudes GPS coordinates of joints that are determined by Eqs (1a, b, c, and d). Point objects are classified as trip

$$\begin{cases} Longitude_{J_{ij}} = \left\{ \begin{array}{l} Longitude_{start} + (Longitude_{end} - Longitude_{start}) \times (ratio_x - 0.5 + ratio_y) \\ Latitude_{start} + (Latitude_{end} - Latitude_{start}) \times ratio_x \end{array} \right. \\ \end{cases} \quad (1c)$$

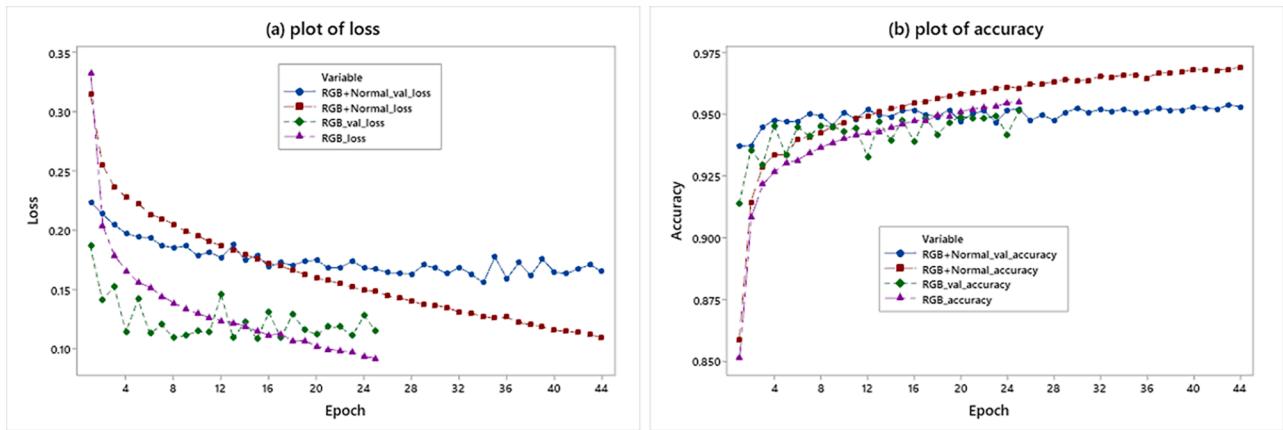
And for Case 4 (North-South), the GPS coordinates of a joint  $J_{ij}$  can be determined with Eq. (1d),

hazards and normal sidewalk joints with different labels while the values of vertical displacement are linked with each point object. In addition, cropped segments of the annotated sidewalk trip hazard and joint images are attached to all points, like in Fig. 4, where a cropped

$$\begin{cases} Longitude_{J_{ij}} = \left\{ \begin{array}{l} Longitude_{start} + (Longitude_{end} - Longitude_{start}) \times (ratio_x - 0.5 + 1 - ratio_y) \\ Latitude_{start} + (Latitude_{end} - Latitude_{start}) \times ratio_x \end{array} \right. \\ \end{cases} \quad (1d)$$

**Table 6**  
U-Net model layers and output shapes.

Layer (type)	Output shape	Parameter number	Connected to
input_1 (InputLayer)	(128, 128, 3)	0	
conv2d_1 (Conv2D)	(128, 128, 64)	1792	input_1
conv2d_2 (Conv2D)	(128, 128, 64)	36,928	conv2d_1
max_pooling2d_1 (MaxPooling2D)	(64, 64, 64)	0	conv2d_2
conv2d_3 (Conv2D)	(64, 64, 128)	73,856	max_pooling2d_1
conv2d_4 (Conv2D)	(64, 64, 128)	147,584	conv2d_3
max_pooling2d_2 (MaxPooling2D)	(32, 32, 128)	0	conv2d_4
conv2d_5 (Conv2D)	(32, 32, 256)	295,168	max_pooling2d_2
conv2d_6 (Conv2D)	(32, 32, 256)	590,080	conv2d_5
max_pooling2d_3 (MaxPooling2D)	(16, 16, 256)	0	conv2d_6
conv2d_7 (Conv2D)	(16, 16, 512)	1,180,160	max_pooling2d_3
conv2d_8 (Conv2D)	(16, 16, 512)	2,359,808	conv2d_7
dropout_1 (Dropout)	(16, 16, 512)	0	conv2d_8
max_pooling2d_4 (MaxPooling2D)	(8, 8, 512)	0	dropout_1
conv2d_9 (Conv2D)	(8, 8, 1024)	4,719,616	max_pooling2d_4
conv2d_10 (Conv2D)	(8, 8, 1024)	9,438,208	conv2d_9
dropout_2 (Dropout)	(8, 8, 1024)	0	conv2d_10
up_sampling2d_1 (UpSampling2D)	(16, 16, 1024)	0	dropout_2
conv2d_11 (Conv2D)	(16, 16, 512)	2,097,664	up_sampling2d_1
concatenate_1 (Concatenate)	(16, 16, 1024)	0	dropout_1, conv2d_11
conv2d_12 (Conv2D)	(16, 16, 512)	4,719,104	concatenate_1
conv2d_13 (Conv2D)	(16, 16, 512)	2,359,808	conv2d_12
up_sampling2d_2 (UpSampling2D)	(32, 32, 512)	0	conv2d_13
conv2d_14 (Conv2D)	(32, 32, 256)	524,544	up_sampling2d_2
concatenate_2 (Concatenate)	(32, 32, 512)	0	conv2d_6, conv2d_14
conv2d_15 (Conv2D)	(32, 32, 256)	1,179,904	concatenate_2
conv2d_16 (Conv2D)	(32, 32, 256)	590,080	conv2d_15
up_sampling2d_3 (UpSampling2D)	(64, 64, 256)	0	conv2d_16
conv2d_17 (Conv2D)	(64, 64, 128)	131,200	up_sampling2d_3
concatenate_3 (Concatenate)	(64, 64, 256)	0	conv2d_4, conv2d_17
conv2d_18 (Conv2D)	(64, 64, 128)	295,040	concatenate_3
conv2d_19 (Conv2D)	(64, 64, 128)	147,584	conv2d_18
up_sampling2d_4 (UpSampling2D)	(128, 128, 128)	0	conv2d_19
conv2d_20 (Conv2D)	(128, 128, 64)	32,832	up_sampling2d_4
concatenate_4 (Concatenate)	(128, 128, 128)	0	conv2d_2, conv2d_20
conv2d_21 (Conv2D)	(128, 128, 64)	73,792	concatenate_4
conv2d_22 (Conv2D)	(128, 128, 64)	36,928	conv2d_21
conv2d_23 (Conv2D)	(128, 128, 2)	1154	conv2d_22
conv2d_24 (Conv2D)	(128, 128, 1)	3	conv2d_23
Total parameters		31,032,837	



**Fig. 6.** Plot of model training and validation with RGB and RGB + Normal, (a) loss, and (b) accuracy.

sidewalk segment is centered at the middle point of joint  $J_j$ . As the sidewalk centerline is along the horizontal direction in the sidewalk trip hazards image, the cropped segment needs to rotate 0° for Case 1, rotate 180° for Case 2, rotate 90° for Case 3, and rotate 270° for Case 4 in counterclockwise. Consequently, the rotated segment presents the sidewalk in the correct direction, a demo can be found in (Jiang, 2021b).

#### 4. Experimental results and discussions

Experiments were performed to validate the feasibility of the proposed approaches, including the concrete slab joints segmentation with deep learning model, the algorithm for joint extraction and vertical displacement measurement, as well as trip hazard detection and mapping.

##### 4.1. Preparation of the sidewalk joint data set

To validate the proposed approach, four sidewalk paths (i.e., P<sub>1</sub>, P<sub>2</sub>, P<sub>3</sub>, P<sub>4</sub>, see Fig. 5) on the UWM campus were scanned using an iPad Pro (12.9-in., 4th generation, with a built-in LiDAR sensor) to prepare the model training and testing dataset.

##### 4.1.1. Sidewalk scanning

The scanning was conducted with a resolution of 10 mm and other setting in Table 3, and a textured point cloud is generated by a processing tool (3D Scanner App). Then, it is necessary to check the start and end points of the scanned sidewalk and the obtained point cloud, especially when the sidewalk path is long, because part of the point cloud may be lost due to technique issues of the point cloud processing tool. Thus, the following strategies are proposed to obtain geo-coordinates for the scanned sidewalk paths:

- (1) Place features of points, lines, or polygons to annotate the scanning start and end points, path, or region in the sidewalk public reporting platform by users (see Fig. 2(c)), and the scanned files also could be attached. The alternative option is taking two photos on the scanned sidewalk at the start and end points with a smart phone. The GPS coordinates of the start and end points are recorded in the two images. As a result, the GPS coordinates can be extracted from image properties, and used to improve the automatic level of trip hazard mapping.
- (2) Scan an entire sidewalk path (ending with corners or turning points) in a single scanning if possible. Turning points can be manually located on the aerial imagery basemap in the Web GIS if GPS coordinates are not recorded or not accurately recorded during the scan.

(3) Plan breaking points before the scanning if breaks are necessary. Breaking points should be on or next to noticeable reference objects, such as sidewalk intersections, isolated trees, building corners and entrances. Those reference objects should be easy to manually locate on the aerial imagery basemap in the Web GIS as well.

##### 4.1.2. Feature image generation

The obtained point clouds were exported as LAS files and imported into a point cloud processing tool, Autodesk ReCap, for visualization and sidewalk plane alignment (in case the scanned sidewalks have a noticeable slope). By setting the display point size as 2, the orthographic view is close to true orthoimages with few gaps in sparse point regions. Then, sidewalk feature images, including orthographic views of RGB and normal features for the aligned point clouds, were created via screenshot. As a point cloud was kept in the same viewpoint and zoom scale, the captured screenshots of RGB and normal views have the same pixel coordinates. Examples of sidewalk feature images are shown in Fig. 2(d), where the RGB view shows points with camera captured colors, and the Normal view displays normal vectors of points in color (Autodesk, 2021; Jiang et al., 2021a).

As previous research showed that the integrated features had better segmentation performance than single feature (Jiang et al., 2021a), this research also generated integrated feature images. Specifically, the 6-channel integrated feature images were generated by assembling RGB and Normal information, which have R, G, and B color information in the first three channels, and the normal information in the following three channels. The elevation image was not used for integrated feature image creation because elevations may change along the joint as shown in Fig. 1. In the experiment, segmentation performance was compared for models trained with RGB and RGB + Normal features respectively to investigate whether adding the normal feature can improve the sidewalk concrete slab joint detection. To prepare the groundtruth annotation for the segmentation task, binary pixelwise joint label images were manually created via the labeling tool and steps presented in (Jiang et al., 2021b). When preparing the groundtruth labels, only the concrete slab joints that are perpendicular to the sidewalk centerlines were created by assigning a label value of 0 for joint pixels and a label value of 255 for non-joint pixels. The prepared model training and testing data sets are downloadable in (Jiang, 2021a).

##### 4.1.3. Training and validation data

In this paper, two west-east direction sidewalk paths, P<sub>1</sub> and P<sub>2</sub>, were used for model training, in which only up-down joints were labeled (while data of P<sub>3</sub> and P<sub>4</sub> are only for testing the trained model). To enrich the training dataset, the following data augmentation (DA) strategies were conducted in preparing the 128 × 128-pixel training

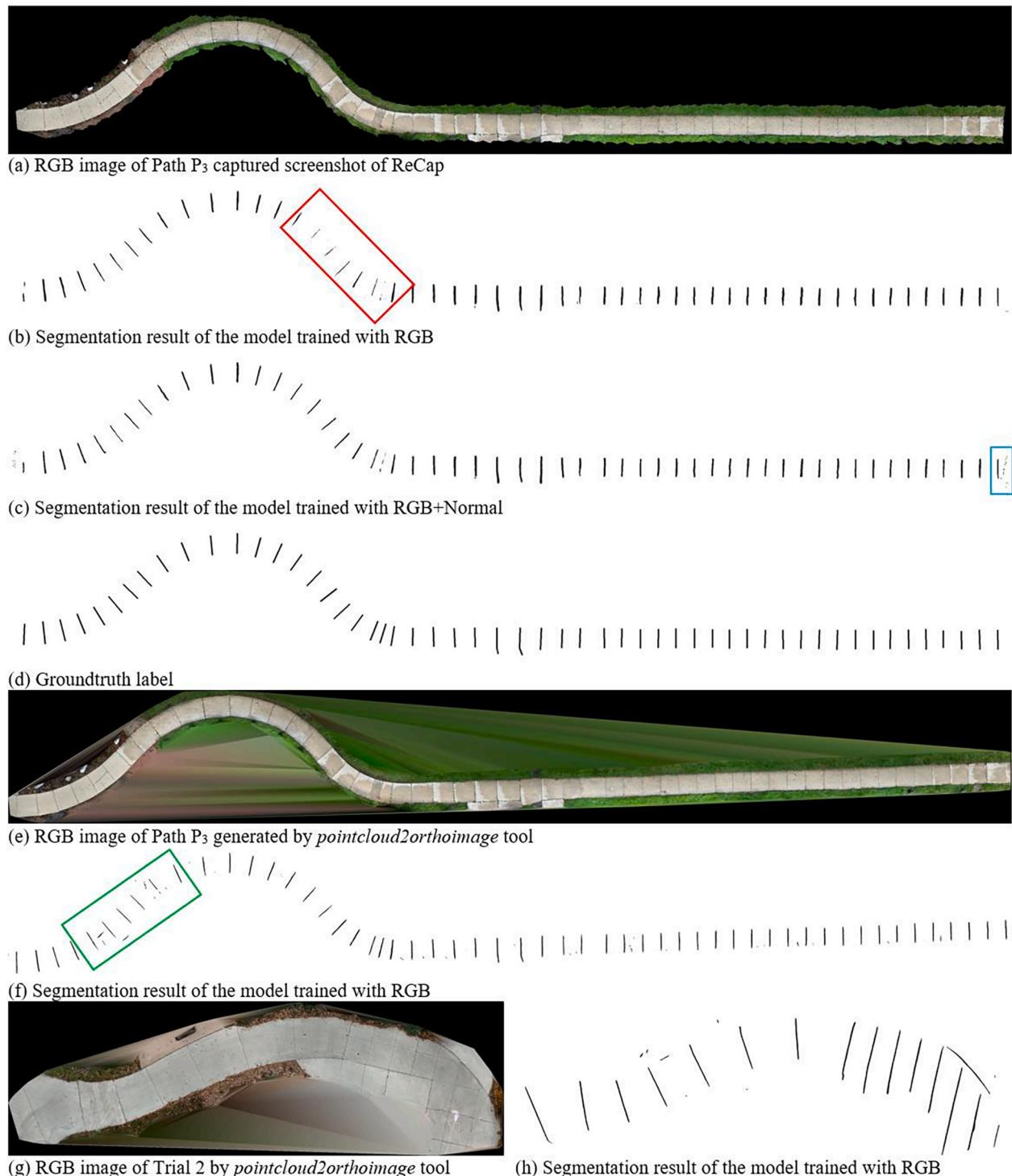
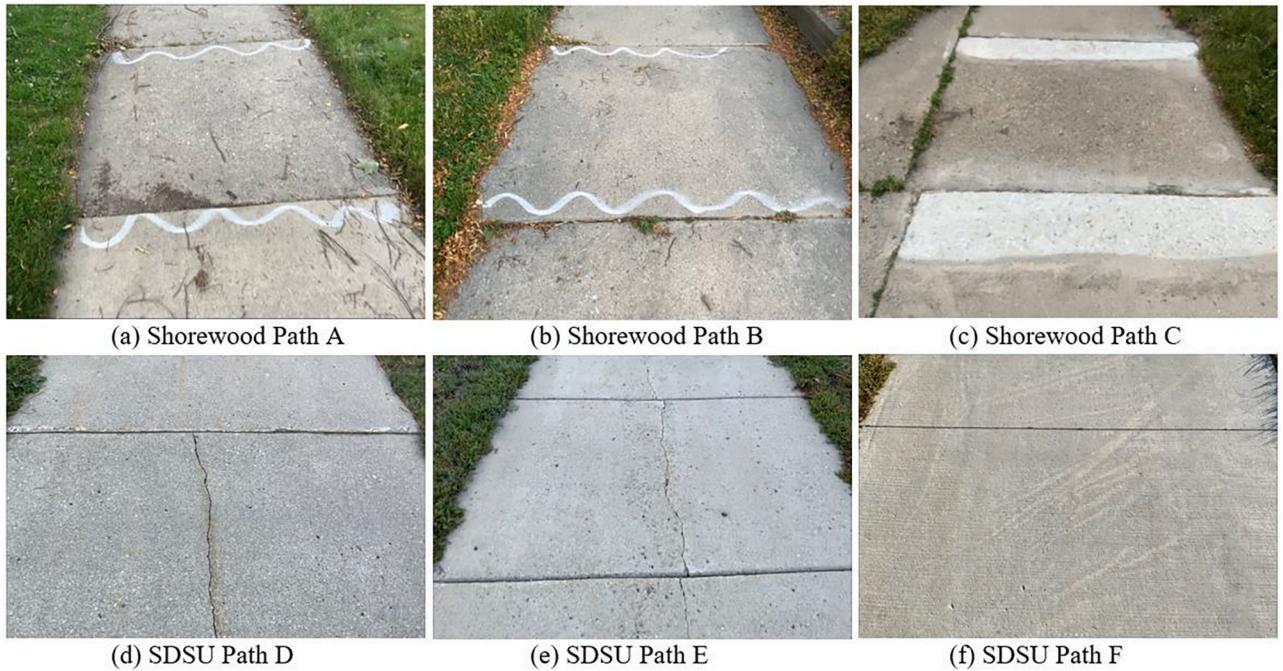


Fig. 7. Comparison of segmentation results.

**Table 7**  
Evaluation results.

Sidewalk Path	Pixel Accuracy		Non-joint IoU		Joint IoU	
	RGB	RGB + Normal	RGB	RGB + Normal	RGB	RGB + Normal
P <sub>1</sub>	0.9933	0.9955	0.9931	0.9954	0.7413	0.8163
P <sub>2</sub>	<b>0.9958</b>	<b>0.9976</b>	<b>0.9957</b>	<b>0.9976</b>	<b>0.8112</b>	<b>0.8880</b>
P <sub>3</sub>	0.9940	0.9941	0.9940	0.9940	0.5469	0.5517
P <sub>4</sub>	0.9896	0.9900	0.9894	0.9898	0.6658	0.6875

images and labels. (a) Randomly flip feature image and label in one of the following options: horizontal, vertical, both horizontal and vertical, non-flipping. (b) Randomly resize the flipped feature image and label in the range of [0.5,2.5]. (c) Randomly rotate the resized feature image and label in the range of [-30,30] degrees. (d) Either randomly conduct the perspective transformation of the feature image and label (keep left, right, top, or bottom edge the same), or not. (e) Cut black margins from the transformed feature image and label, and pad the remaining feature image and label to be multiples of 128 pixels. (f) Randomly adjust the padded feature image's brightness, color, contrast, or sharpness in the



**Fig. 8.** (a) and (b) grinding eligible concrete slabs with marked trip hazards, (c) and (d) repaired concrete slabs without marked trip hazards, (e) a curved sidewalk Path E in good condition with groover cut contraction joints, (f) a newly constructed sidewalk Path F with sawcut contraction joints.

range of [0.5, 1.5], adjustments are not applied to the normal feature and label. (g) Rotate the adjusted feature image and label by 0° and 180° (because sidewalk paths are always rotated in a horizontal direction, and only joints perpendicular to the centerline are considered in this paper). (h) Crop the two sets of rotated feature images and labels into 128 × 128-pixel small-patches (which have 50% overlap among adjacent ones) by moving a 128 × 128-pixel slide window with a stride of 64-pixels in both width and height directions, skipping blank windows.

By repeating the above DA steps with several rounds (skipping the random processing steps in the first round to keep the original feature image and label, and then, runs all steps for remaining times), the created model training data sets would have a high variety of size, shape, color, orientation, and views of concrete slabs and joints. This study ran the DA for 51 rounds and generated 80,806 RGB (3-channel) and ground truth label samples, which were temperately saved in the RAM. Since the RGB + Normal sample is larger than the RGB sample, a fewer number of round DA (i.e., 41 rounds) was applied, which generated 62,284 RGB + Normal (6-channel) and label samples (see Table 5).

#### 4.2. Training and testing results of the segmentation model

After preparing the dataset, the U-Net models were constructed with software packages of Keras 2.3.1, Python 3.6.8, OpenCV 3.4.2 and TensorFlow-GPU 1.14, and ran on a workstation of 96 GB RAM and 4 × 11 GB GPUs (GeForce RTX 2080 Ti) for model training. The U-Net model architecture was well described in (Ronneberger et al., 2015), and the code can be found in (Zhi, 2019). With a 128 × 128-pixel RGB sample, the detailed U-net model layers and output shapes are shown in Table 6, where the hidden layers “conv2d\_1” to “conv2d\_23” (kernel size 3 × 3) use an activation function ReLU for faster model training; the two dropout layers are used to prevent overfitting; the four concatenate layers are used to combine the feature-maps (tensors) from two different layers as a new feature-map (tensors); and the output layer “conv2d\_24” (kernel size 1 × 1) uses the Sigmoid activation function to create label pixels in the range of 0 to 1 (Chollet, 2020b; Jiang & Bai, 2020b; Jiang, Bai, & . , 2020; Ronneberger et al., 2015). Similarly, for any RGB + Normal sample, the output shape of the “input\_1” layer is (128, 128, 6), and the parameter number of the “conv2d\_1” layer is 3520, and the

remaining shapes and parameter numbers are the same as Table 6.

Moreover, for the model training, the optimized configurations in (Jiang et al., 2021a) were adopted in this research. In detail, the Adam optimizer (learning rate 0.0001) and binary cross-entropy loss function were used. Each model was set to be trained up to 100 epochs, batch size was set as 256. In addition, 10% of samples were randomly selected to validate the model in each training epoch. Meanwhile, early stopping criteria was used to avoid model overfitting, which would stop the model training once the validation loss does not decrease for 10 epochs.

##### 4.2.1. Training results

Fig. 6 shows the plots of loss (binary cross-entropy loss) of the U-Net training and validation process with two different types of datasets (RGB and RGB + Normal). The training of both models was stopped before the 100th epoch which avoided the model overfitting with the early stopping criteria of the validation loss has not been decreased for 10 epochs. In detail, the RGB model reached the smallest validation loss of 0.1083 at the 15th epoch, and had an ending validation loss of 0.1151 at the 25th epoch; the the RGB + Normal model achieved the smallest validation loss of 0.1561 at the 34th epoch, and got an ending validation loss of 0.1653 at the 44th epoch.

Moreover, in Fig. 6, the training and validation accuracy was measured by Keras “accuracy” which calculates how often predictions equal labels for the 128 × 128-pixel small-patches (Chollet, 2020a). With the additional 10 epochs training, the RGB model slightly increased the validation accuracy from 0.9478 to 0.9517, and the RGB + Normal model slightly improved the validation accuracy from 0.9521 to 0.9531. The restuls indicate the RGB + Normal model’s validation accuracy is only slightly better than the RGB model. However, training the U-Net model with the RGB + Normal dataset (56,055 samples, 140 s/epoch in average, 44 epoches in total, and about 103 min) costs much more time than the RGB dataset (72,725 samples, 170 s/epoch in average, 25 epoches in total, and approximate 71 min). Thus, the perforamces of the two trained U-Net models will be further evaluated with the testing dataset. Furthermore, since both models achieved satisfactory performance, with an accuracy over 0.95 in the end, the saved models at the ending epochs were used for testing and generating label prediction in this research.

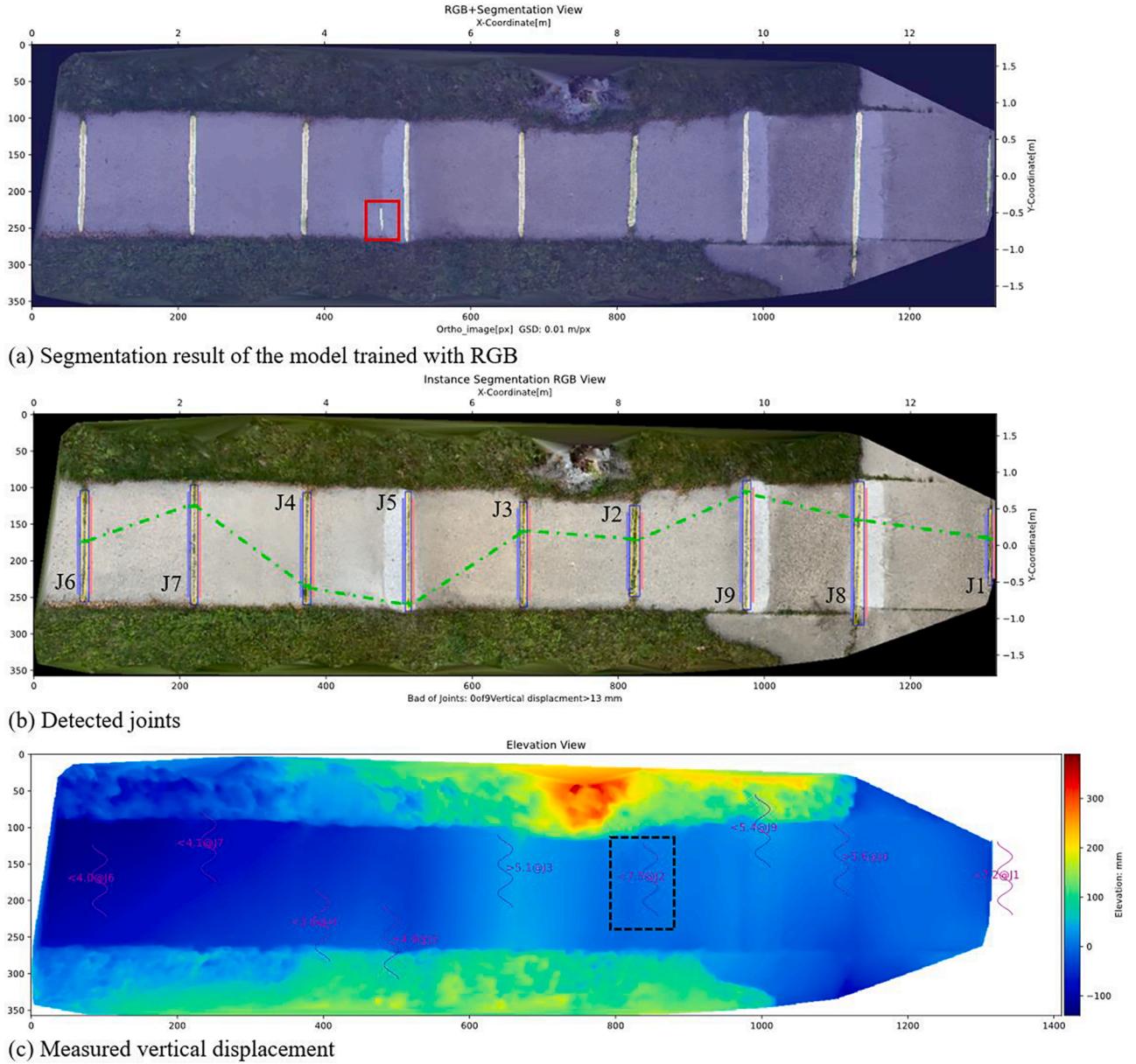


Fig. 9. Experimental results of Path C with repaired concrete slabs.

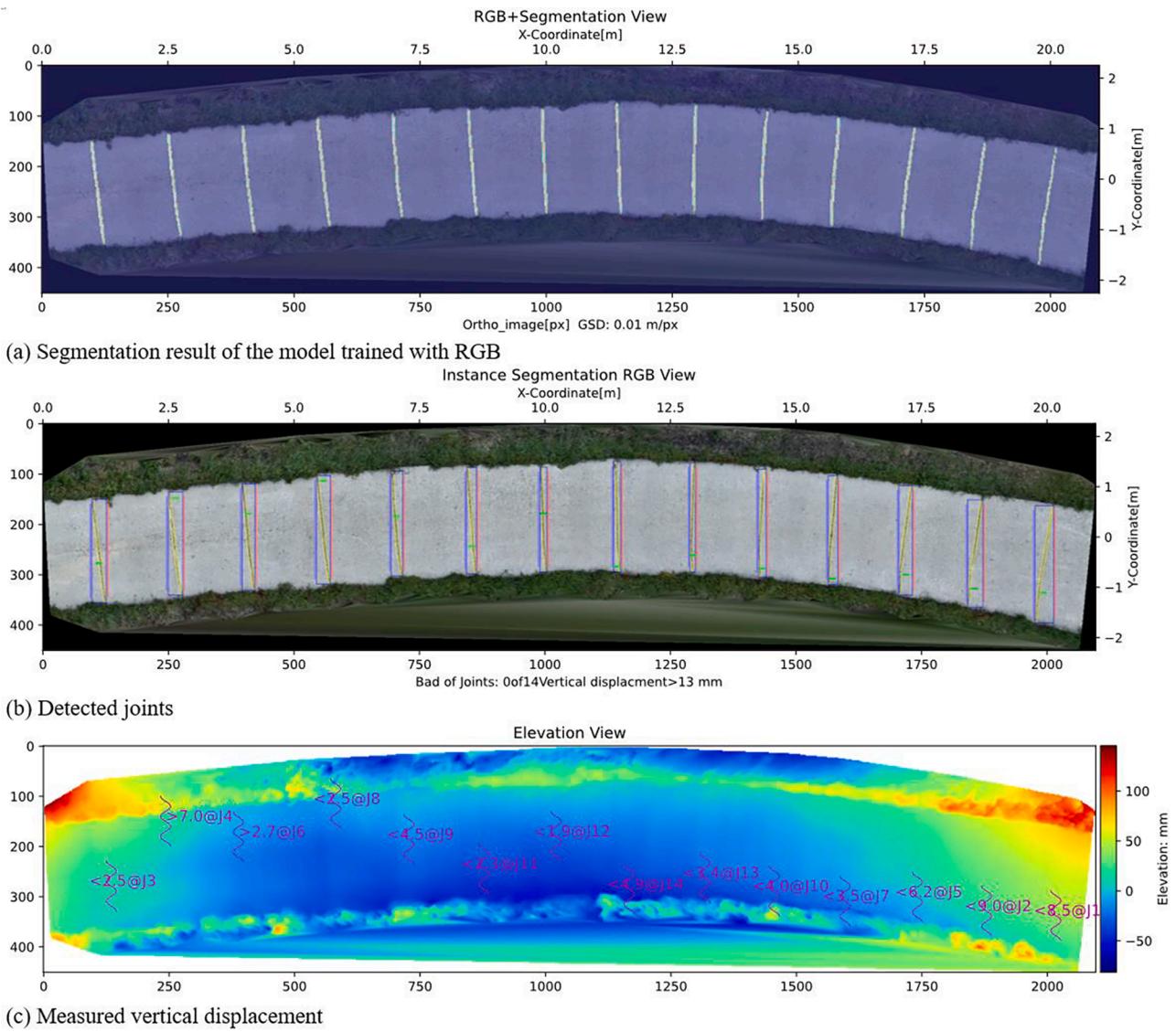
#### 4.2.2. Testing results

After training the models, data of the curved sidewalk path P<sub>3</sub> (see Fig. 7 (a), where joints are presented as oblique joints and different from training data sets of sidewalk path P<sub>1</sub> and P<sub>2</sub>) and the south-north direction sidewalk path P<sub>4</sub> were used to test the well-trained RGB model and RGB + Normal model. Data of P<sub>3</sub> and P<sub>4</sub> were fed into the trained U-Net model, which then generated the 128 × 128-pixel small patch predictions. For example, Fig. 7 (a) has dimensions of 3,456 × 512-pixel, which was disassembled into 371 small patches (because the width direction has  $2 \times 3,456/128 - 1 = 53$  and the height direction contains  $2 \times 512/128 - 1 = 7$  slide windows with a 64-pixel stride), then the U-Net model generated 371 output patches.

As the output patches were 50% overlapped with each other, only central parts of the output patches were assembled (Jiang et al., 2021b) to obtain the large-sized segmented images, which were compared to the ground truth label images to evaluate the segmentation accuracy. Pixel accuracy, non-joint IoU (Intersection over Union) and joint IoU were used as the evaluation metrics. The evaluation results in Table 7 show that the model trained with RGB + Normal dataset performed slightly

better than the model trained with only RGB dataset for all the four sidewalk paths. This conclusion is also supported by the validation accuracy plotted in Fig. 6 (b). The first two rows P<sub>1</sub> and P<sub>2</sub> in Table 7 indicate the U-Net model in sidewalk joint segmentation has better performance than the U-Net model in pavement cracking segmentation (Jiang et al., 2021a), which has an average pixel accuracy of 0.9817, average non-crack IoU of 0.9813, and average cracking IoU of 0.5728 in validation. In addition, in sidewalk joint segmentation, the integrated RGB + Normal feature is slightly better than the singular RGB feature because joints are straight lines with uniform normal features, while in pavement cracking segmentation, the singular RGB feature is better than the integrated RGB + Normal feature because cracks are irregular curves with more complicated normal features.

Additionally, Table 7 shows pixel accuracies obtained by the two models are similar when testing on each sidewalk path data. Similarly, there is only tiny or no difference of the non-joint IoUs for the two models. In contrast, relative larger differences were observed in joint IoUs obtained by the two models. In addition, both models always achieved the highest accuracy (in terms of all the three metrics) for



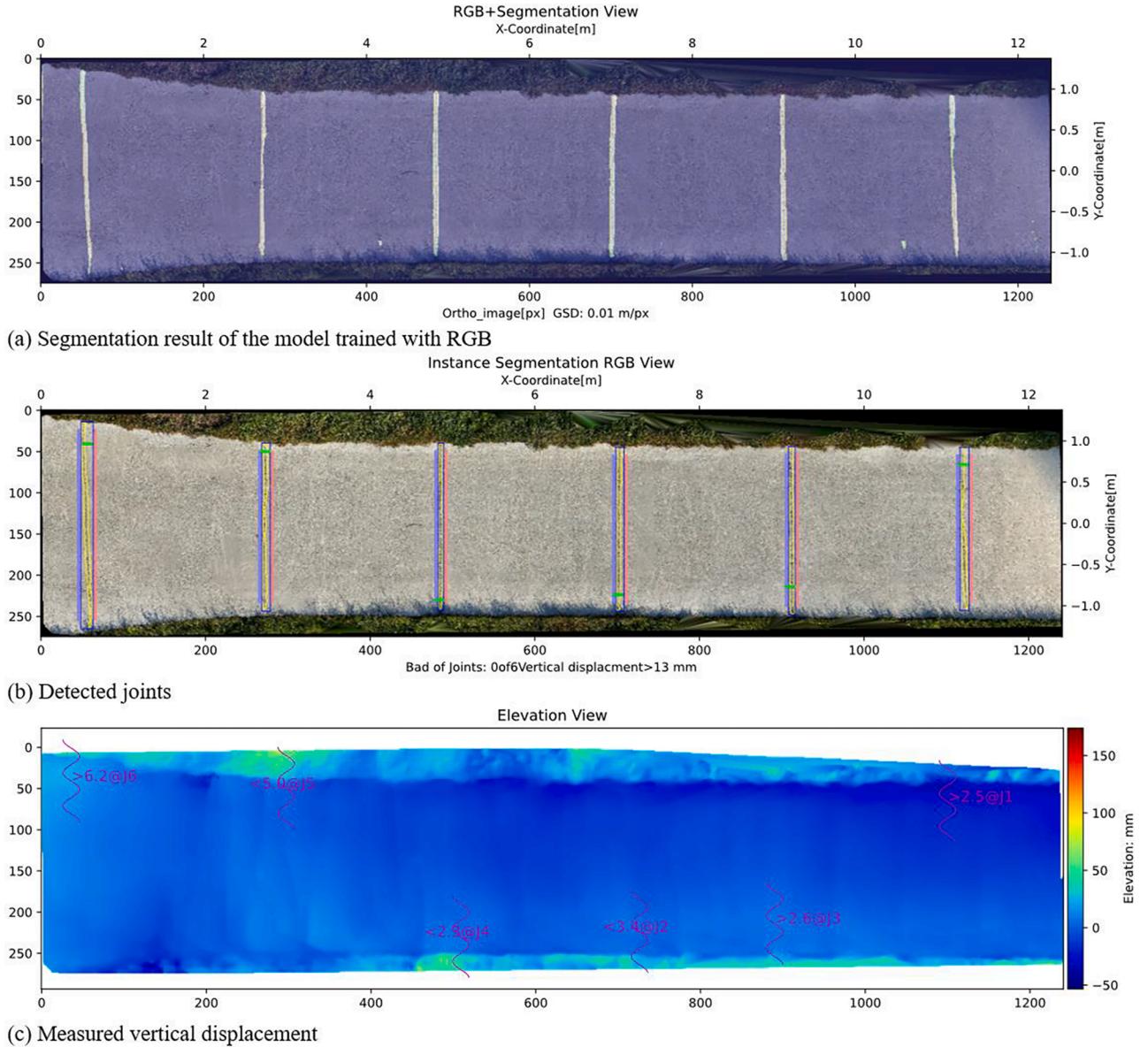
**Fig. 10.** Experimental results of Path E with groover cut contraction joints.

segmenting the data of P<sub>2</sub> while the lowest for P<sub>4</sub> (in terms pixel accuracy and non-joint IoU) and P<sub>3</sub> (joint IoU). One possible reason for such results is that data of P<sub>3</sub> and P<sub>4</sub> are not included in the training dataset and the joints of P<sub>3</sub> are oblique joints which are different from those in P<sub>1</sub> and P<sub>2</sub>.

Furthermore, Fig. 7 (b) and (c) show the applied DA made the oblique joints in the curved section of sidewalk path P<sub>3</sub> detectable in the RGB and RGB + Normal models, respectively. The widths of the detected joints are smaller than the manually annotated label image (in Fig. 7 (d)), which resulted in the smallest joint IoUs in Table 7 when testing the models on sidewalk path P<sub>3</sub>. That is reasonable because the manually created label images used line segments with a constant width to represent all joints without considering the actual joint widths. Moreover, the geospatial resolution values of GSD are different among training and testing data sets (Jiang, 2021a), where the concrete slab widths (joint lengths) are about 72, 143, 66 and 96-pixel in sidewalk paths P<sub>1</sub>, P<sub>2</sub>, P<sub>3</sub> and P<sub>4</sub>, respectively. The applied DA has randomly scaled ratios of (0.5, 2.5), which made the joint detectable in feature images with different GSDs. For future application, increasing the size of feature images (i.e., using a small GSD) to increase the number of joint pixels could be considered for improving performance of pixelwise segmentation, which is discussed later.

#### 4.2.3. Testing on images generated by the pointcloud2orthoimage tool

The RGB images created by the *pointcloud2orthoimage* tool (Fig. 7 (e) and (g)) were processed by the well-trained RGB model with the disassembling and assembling algorithm (Jiang et al., 2021b). The Fig. 7 (e) has dimensions of 8,539 × 1,128-pixel (GSD = 1 cm/pixel), which was disassembled into 2,261 overlapped 128 × 128-pixel small-patches (because the image was padded into 8,576 × 1,152-pixel first, and then 2 × 8,576/128-1 = 133 columns and 2 × 1,152/128-1 = 17 rows, and a total of 2,261 small-patches were generated), and then the same number of 128 × 128-pixel U-Net output patches were used to assemble the Fig. 7 (g). The corresponding output segmented label image in Fig. 7 (f) has dimensions of 8,539 × 1,128-pixel (GSD = 1 cm/pixel), which are much larger than Fig. 7 (b) - (d) with dimensions of 3,456 × 512-pixel. The results show that the proposed deep learning-based image segmentation can process a very large sidewalk RGB image to produce pixelwise label image, in which all joints were well segmented, and several cracks were detected in Fig. 7 (f) as well. Large cracks on concrete slabs would cause the trip hazards (see Table 1), while small cracks can be discarded by setting an area threshold in the joint extraction, as discussed later. Moreover, the label image prediction in Fig. 7 (h) shows the segmentation model trained with RGB dataset successfully segmented sawcut joints that are perpendicular to concrete slab



**Fig. 11.** Experimental results of Path F with sawcut contraction joints.

centerlines and skipped joints along centerline directions. Since the segmentation performances are very similar between the models trained with integrated RGB + Normal and singular RGB datasets as shown in Table 7, this paper used the U-Net model trained with RGB dataset for the remaining joint extraction tasks.

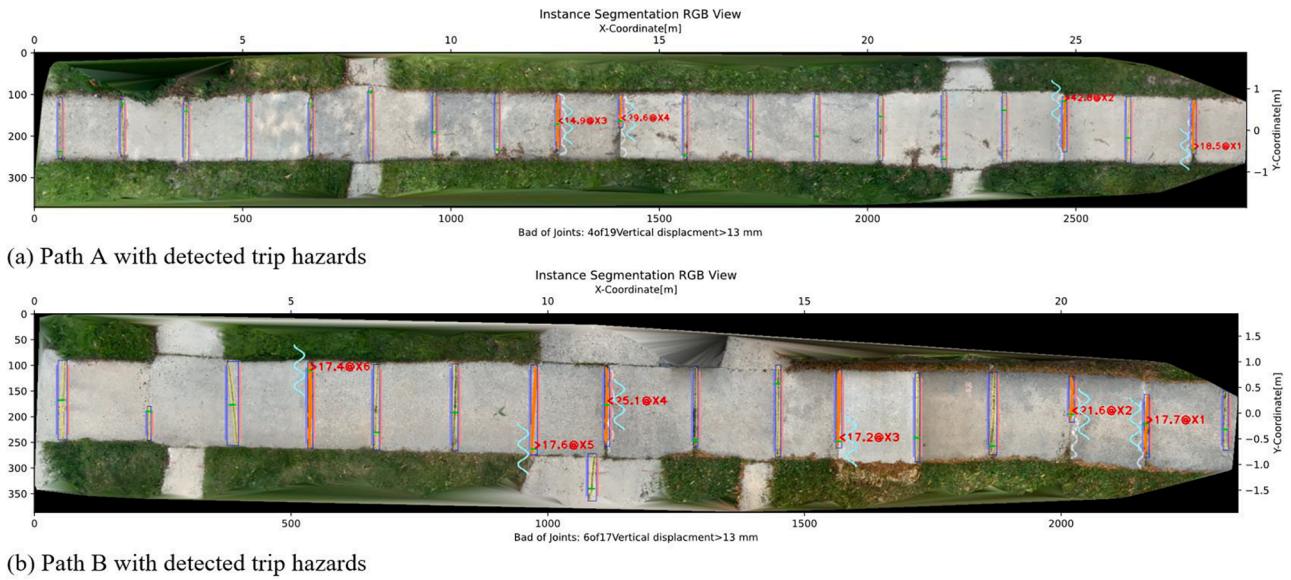
#### 4.3. Results of joint extraction and vertical displacement measurement

Typically, sidewalk concrete slab joints include contraction (control) joint, isolation (expansion) joint, and construction joint, which a width range from 3 mm to 20 mm with different construction methods and tools. The authors scanned additional old and new concrete sidewalks to test the developed joint extraction and vertical displacement measurement algorithm (in Fig. 3). First, two manually surveyed old sidewalk Paths A and B and one repaired (with grinding) sidewalk Path C were scanned (with resolution of 5 mm and other settings in Table 3) at the Shorewood Village in the USA. Shorewood Village conducted the sidewalk replacement program in the year of 2021, and the grinding eligible concrete slabs were marked with white wavy lines on the edges of the slabs if their vertical displacement is between 19.05 and 38.1 mm (3/4

and 1 ½ inches), indicating the trip hazards, as shown in Fig. 8 (a) and (b). In addition, three relatively new sidewalks (without marked trip hazards) with sawcut (Fig. 8 (d) and (f)) and concrete groover cut (Fig. 8 (e)) contraction joints were scanned on the SDSU campus. The scanned point clouds were automatically converted into orthoimages and elevation data via the *pointcloud2orthoimage* tool with GSD = 1 cm/pixel.

##### 4.3.1. Experimental results of non-trip hazard sidewalks

The joint segmentation result for Path C by the well-trained RGB model-based is shown in Fig. 9 (a), where all joints perpendicular to the sidewalk centerline were successfully detected, and any joints codirect to the sidewalk path centerline direction were not detected as designed. Fig. 9 (b) shows extracted nine joints which are imposed on the RGB image, where the linked dash-dotted line shows the maximum vertical displacement points are irregularly distributed, and the noise annotated in Fig. 9 (a) was discarded as its area is smaller than the threshold of 200-pixel. Fig. 9 (c) shows the measured vertical displacement of each joint, where a maximum vertical displacement of 7.5 mm was detected at the extracted joint J2. As all joints have vertical displacements less



**Fig. 12.** Experimental results of Paths A and B with trip hazards.

than the trip hazard criterion of 13 mm (1/2 in.), no wavy line was marked in Fig. 9 (b). This evaluation result shows the grinding work at joints J5, J8 and J9 was well performed, as the vertical displacements are 4.9, 5.6 and 5.4 mm respectively, all of which are less than 6.4 mm (1/4 in.) to satisfy ADA Standards (in Table 1). The annotated region in Fig. 9 (c) contains several instances of elevation noise because it is a sparse point region. The authors applied a median filter (OpenCV, 2021b) with a size of  $5 \times 5$ -pixel and removed some noises, which smoothed the elevation data but kept all edges of elevation changes. To avoid this issue, obtaining a dense point cloud is necessary, which means moving forward slowly over the joint when scanning the sidewalk is required.

Moreover, Fig. 10 shows the results (i.e. segmentation, extracted joints and vertical displacement) of a curved sidewalk Path E which is in good condition with groover cut contraction joints, and Fig. 11 shows the same set of results of a newly constructed sidewalk Path F with sawcut contraction joints. Evaluation results show all concrete slabs in these two paths are in a good condition with vertical displacements less than 13 mm (1/2 in.).

#### 4.3.2. Experimental results of trip hazard sidewalks

The grinding eligible sidewalk Paths A and B were evaluated, and results are shown in Fig. 12. For sidewalk Path A in Fig. 12(a), the developed method correctly identified four potential trip hazards, and the generated wavy lines match with the manually performed evaluation results in the same concrete slab edges. For sidewalk Path B in Fig. 12(b), six trip hazards were identified, i.e., their vertical displacement exceeded the criterion of 13 mm. The largest three vertical displacement values occurred at X4, X2 and X1, and match with the manual evaluation results. Since Shorewood Village's policies (in Table 1) only identify vertical displacements larger than 19.05 mm (3/4 in.), the other three identified joints with vertical displacements (17.2, 17.4, 17.6 mm) less than 19.05 mm were not marked on Path B. In Fig. 12 (a) and (b), several extracted joints missed the parts where joints were covered by vegetation and dead leaves, as shown in Fig. 8 (a) and (b). One potential approach to avoid that is to clean the sidewalk before scanning. Since the repaired sidewalk Path C is much cleaner than A and B, the joint extraction is better as well. Another approach is preparing additional joint label images for the covered conditions as labeling the joint label images is easier than labeling the point cloud.

Furthermore, a trip hazard was successfully detected on the joint edge of Path D and annotated in Fig. 13 (a) with a vertical displacement

of 14.7 mm. The possible reason is that, Fig. 8 (d) shows that only the concrete slab edge was grinded for Path D, and the repairing was not thoroughly performed as with path C in Fig. 8 (c). Moreover, the authors tried a smaller GSD of 5 mm/pixel to improve the pixelwise segmentation for Path D, in which the broken contraction joints were extracted as two separated joints in Fig. 13 (a). The developed *pointcloud2orthoimage* tool created a new RGB orthoimage in Fig. 13 (c), which has double the pixel size of the one in Fig. 13 (a). As a result, the broken contraction joints were extracted as a single joint in Fig. 13 (c). The maximum vertical displacement was slightly decreased from 14.7 mm to 12.9 mm, hence made the trip hazard disappear in Fig. 13 (c). However, this small difference is reasonable because: the edge offset parameter in Table 4 was not changed, the left slab edge ( $X_l = x - 3$ ) and right slab edge ( $X_r = x + w + 3$ ) are closer to the joint with the smaller GSD; then, the elevations of edges were measured on the grinded slope with a slightly smaller maximum vertical displacement.

#### 4.4. Long path trip hazards detection, mapping and geo-visualization

The sidewalk joint segmentation and extraction results in Section 4.3 further confirmed the trained U-Net model with RGB dataset has good performances on the *pointcloud2orthoimage* tool generated RGB images. Those RGB images ( $GSD = 1$  cm/pixel) have the maximum width of 2,910-pixel, which is the longest (29.1-m) sidewalk Path A in Fig. 12 (a). In this section, SfM and LiDAR are compared for scanning long sidewalk paths, and the joint and trip hazard mapping is discussed later.

##### 4.4.1. Trip hazards detection of long paths using SfM and LiDAR

As mentioned in Table 2, SfM photogrammetry is another alternative scanning method to obtain a sidewalk full-width as-is condition. To test the feasibility of the camera and SfM photogrammetry method, three trials of sidewalk scanning were conducted. In Trial 1, there were 1,136 images of a long singular walk path manually captured with a smart phone (Apple iPhone SE). The images were used to generate point cloud using a photogrammetry software (Pix4Dmapper) and SfM tool, VisualSfM (Wu, 2011). However, both methods failed to generate the point cloud file that continually represents the straight and flat sidewalk surfaces that have been scanned. Moreover, Pix4Dcatch (an application for ground 3D scans from mobile devices) was used to assist the sidewalk scanning in Trials 2 and 3.

In Trial 2, the authors walked twice on a short sidewalk (about 20 m) and obtained 100 images (with effective overlaps: 3.14 images per pair,

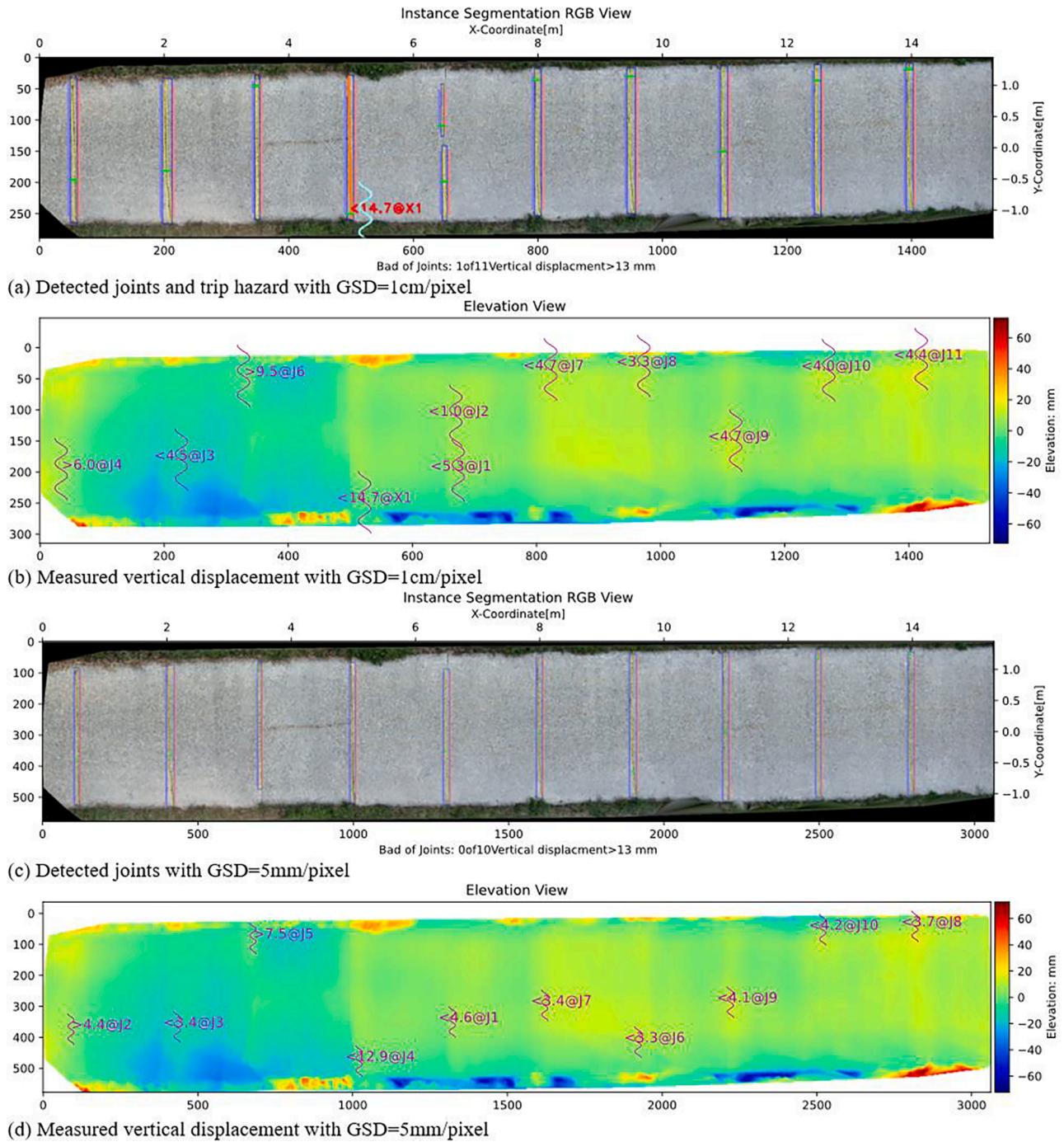


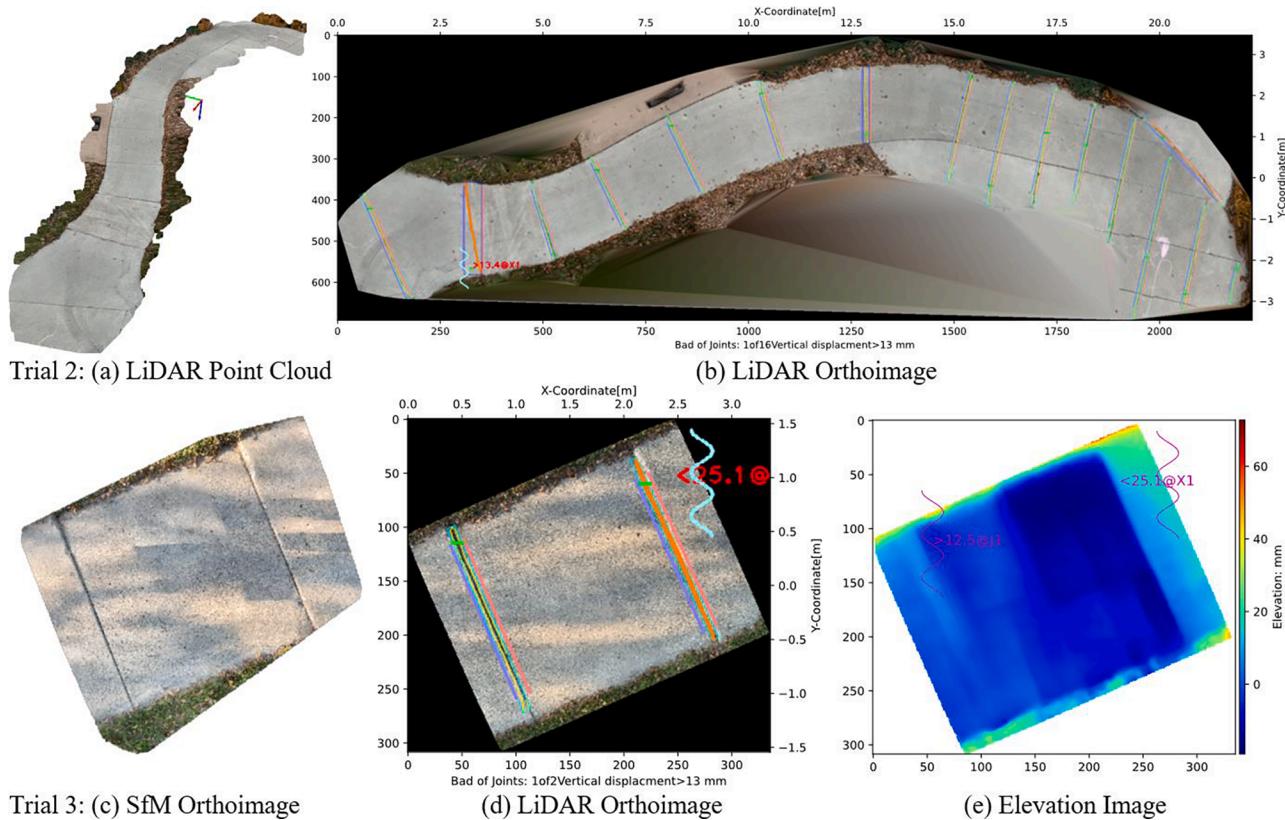
Fig. 13. Experimental results of Path D with different GSDs.

GSD: 1.019 cm/pixel). The ReCap Photo took 65 min but only produced an area of 0.103 m<sup>2</sup> of point cloud. In contrast, the LiDAR point cloud scanned by iPhone Pro has full coverage of the short sidewalk, as shown in Fig. 14 (a). Following that, the created orthoimage (GSD = 1 cm/pixel) from the LiDAR data, segmented image, and extracted joints are shown in Fig. 7 (g), Fig. 7 (h), and Fig. 14 (e), respectively. All joints perpendicular to sidewalk centerline were detected, joints across several concrete slabs were extracted as single long joints, while joints along centerline directions were skipped as per the design.

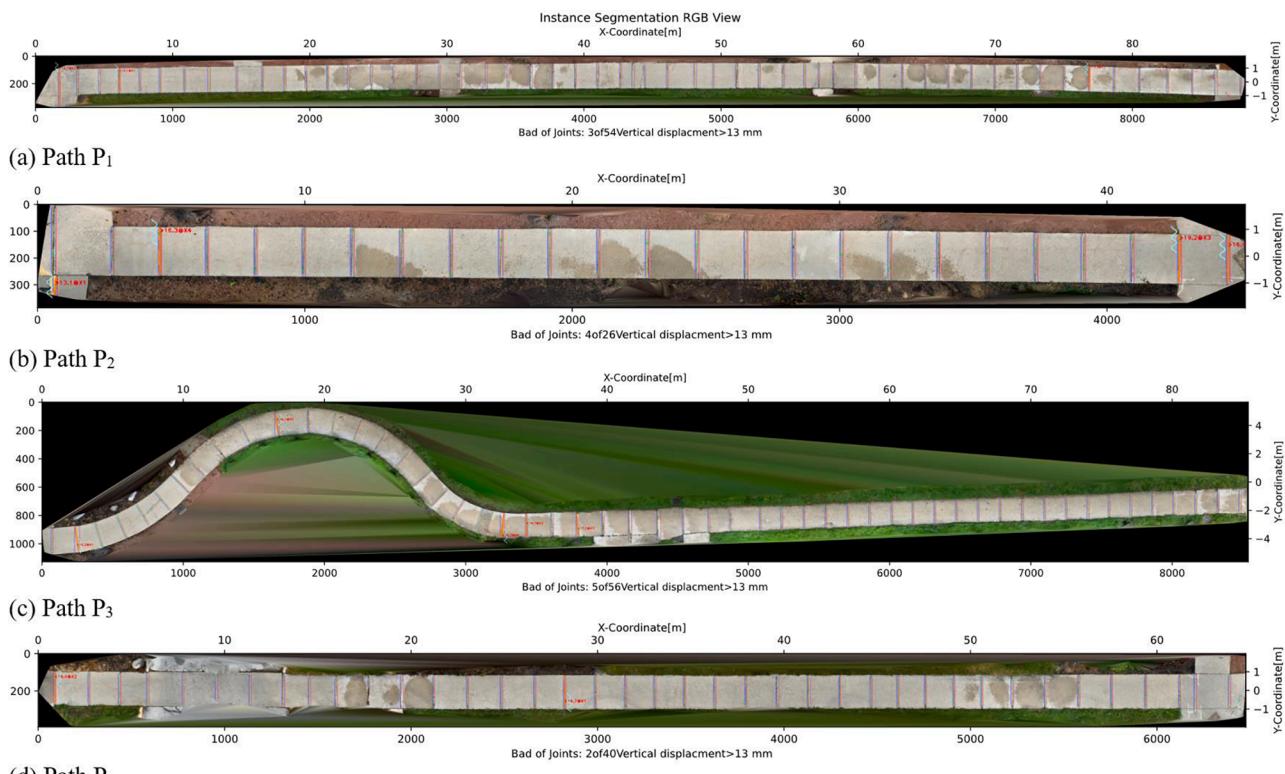
In Trial 3, two concrete slab joints were scanned and 100 images (effective overlaps: 4.37 images per pair, GSD: 1.001 cm/pixel) were obtained. The ReCap Photo took 61 min to generate the SfM photogrammetric point cloud with the orthoimage shown in Fig. 14 (c), which

covers an area of 6.148 m<sup>2</sup> and the two concrete slab joints. The same joints were also scanned with an iPad Pro using built-in LiDAR sensor. Fig. 14 (d) and (e) show the LiDAR point cloud converted orthoimage and the elevation image, which have a GSD = 1 cm/pixel. In Trial 3, the LiDAR and SfM method had equal scanning time, but the latter was much slower than former in converting images to textured point cloud. Thus, camera and SfM photogrammetry could be used for scanning a piece of trip hazard by property owners and concerned citizens, but for reporting a long sidewalk path, LiDAR would be a better option. According to the developed methods presented in this paper, the extracted joints and measurements of vertical displacements have the same length as the actual joints in Fig. 14.

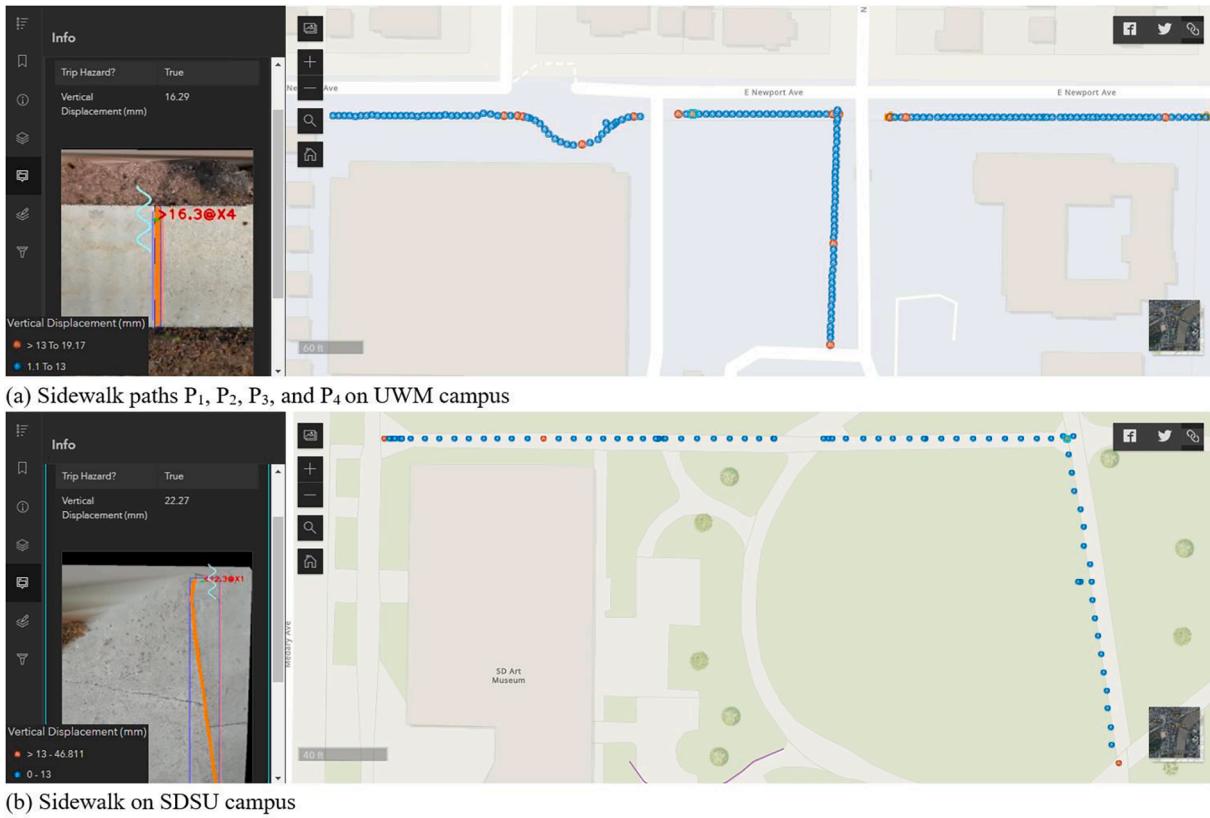
For scanning a long sidewalk path (like in Fig. 7 (e)), which has an



**Fig. 14.** Experimental results on Trials 2 and 3 to compare SfM and the proposed method.



**Fig. 15.** Images with detected and annotated trip hazards for path P<sub>1</sub>, P<sub>2</sub>, P<sub>3</sub>, and P<sub>4</sub>.



**Fig. 16.** Mapped joint and trip hazards with annotated image segments and specific information. (Map Data © 2021 Esri).

approximate length of 90 m along the sidewalk centerline, using a resolution of 10 mm is recommended, because the point cloud processing tool e.g. 3D Scanner App has limited storage for the point cloud file in a single scan. By setting GSD = 1 cm/pixel, the point cloud converted elevation data is smooth without gaps. The point clouds of the four scanned long sidewalk paths P<sub>1</sub>, P<sub>2</sub>, P<sub>3</sub> and P<sub>4</sub> (see Fig. 5, concrete slab width about 1.8 m) were converted to feature images via the developed *pointcloud2orthoimage* tool with GSD = 1 cm/pixel. All concrete slab joints were extracted, and vertical displacements were measured via the developed joint extraction and vertical displacement measurement algorithm. The created images of annotated trip hazards are shown in Fig. 15, in which 54, 26, 56, and 40 joints were extracted, and three, four, five and two trip hazards were identified in paths P<sub>1</sub>, P<sub>2</sub>, P<sub>3</sub>, and P<sub>4</sub>, respectively.

#### 4.4.2. Mapping and geo-visualization of the results in Web GIS

GPS coordinates of the start and end points of the four long paths were manually obtained from the Web GIS platform i.e., ArcGIS Online because they are corners and intersections. Based on the GPS coordinates of start and end points, the sidewalk Paths P<sub>1</sub> and P<sub>2</sub> belong to Case 1 (scanned from west to east), Path P<sub>3</sub> is Case 2 (scanned from east to west), and Path P<sub>4</sub> is Case 3 (scanned from south to north). The GPS coordinates of the middle points of each joint were calculated via Eq. (1) and the middle point was added to the Web GIS platform to represent the joint. Meanwhile, the joint segments were cropped from the annotated trip hazards images, and rotated zero degrees for Paths P<sub>1</sub> and P<sub>2</sub> (see Fig. 16(a)), rotated 180° for Path P<sub>3</sub> (see Fig. 4(a)), and rotated 90° for Path P<sub>4</sub> (see Fig. 4(b)). In the end, both the cropped joint image and specific displacement value were attached to each point of the joint in the Web GIS platform, as shown in Fig. 16 and also available in (Jiang, 2021b). This method can not only visualize and facilitate sidewalk assessment, but also help monitor and analyze the long-term sidewalk condition changes if scanning data are obtained continuously.

## 5. Conclusion

This paper developed and tested a sidewalk trip hazard detection and geo-visualization method that can automatically assess concrete slab deficiencies after obtaining the point clouds via a low-cost LiDAR scanner. Firstly, low-cost mobile LiDAR devices were used to scan sidewalks to obtain the point cloud data, which were then converted to RGB images using the develop tool. Second, a deep learning-based segmentation model U-Net was trained with the sidewalk images to segment concrete joints in the image. Afterwards, joints were extracted from the segmented image and vertical displacements for each joint were evaluated, based on which potential trip hazards were identified and specific information was geo-visualized in Web GIS platform. The experiment results demonstrated the effectiveness of the proposed method. Specifically, the segmentation model performed well for segmenting different types of joints in images (with a highest joint IoU of 0.88) and all the vertical displacement conditions were accurately and comprehensively detected. It was found that integrating the RGB feature with the Normal feature can improve the joint segmentation accuracy of the deep learning model, but the improvement was not significant. For future application, using the point cloud converted orthoimages is sufficient to detect joints. In this study, the segmentation model trained with a few images of straight sidewalks with groover cut contraction (control) joints and the corresponding joint label images already obtained good performance, but adding extra images, such as vegetation covered joints, to enrich the dataset will be considered for future application. Compared to the methods (in Table 2) in existing studies, scanning the as-is condition of the sidewalk with a mobile device is convenient and faster in achieving full-width coverage.

The main contributions of this study are as follows:

- (1) Developed more efficient detection and mapping of sidewalk trip hazards by mobile devices. A *pointcloud2orthoimage* algorithm and

tool was developed (code available in (Jiang, 2022)), which can generate large-size high-resolution sidewalk feature images of orthoimage and elevation image from the mobile devices (e.g., iPad and iPhone Pro) scanned LiDAR point cloud automatically. Then, a deep learning-based image segmentation method was developed for concrete slab joint extraction and vertical displacement measurement, which can evaluate performance or condition of constructed concrete slabs, such as the sidewalk trip hazards detection presented in this paper.

- (2) Developed a sidewalk reporting and management platform based on Web GIS. With the sidewalk public reporting port, property owners and concerned citizens can report community sidewalk deficiencies, such as trip hazards, crack, spalling and chipping, gap at edge, and vertical displacement. With the sidewalk management function, facility management agencies can maintain the updated sidewalk performance information for maintenance planning. Moreover, technicians can use the mapped joints and trip hazards information to fast locate them on jobsite. After the repairing, superintendent can inspect the repairing quality, and update the sidewalk performance information to the Web GIS platform. In addition, wheelchair users can find the best path based on the geo-mapped sidewalk deficiencies.

## 6. Data availability statement

The training and testing data sets are available in (Jiang, 2021a). The *pointcloud2orthoimage* code is available in (Jiang, 2022). The U-Net code are available in (Zhi, 2019). Other Python codes are available from the corresponding author upon reasonable request.

## CRediT authorship contribution statement

**Yuhan Jiang:** Conceptualization, Methodology, Software, Visualization, Investigation, Writing - original draft. **Sisi Han:** Investigation, Writing - original draft. **Dapeng Li:** Software. **Yong Bai:** Conceptualization. **Mingzhu Wang:** Investigation, Writing - review & editing.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

Data will be made available on request.

## Acknowledgements

The authors are thankful to the Shorewood Sidewalk Program. The authors are thankful to Mr. Ashwani Verma (undergraduate student at South Dakota State University) for his help in scanning sidewalks on SDSU campus. The deep learning workstation was financially supported by the McShane Endowment fund at Marquette University.

## References

- Access Board. (2013). (Proposed) Public Rights-of-Way Accessibility Guidelines. Retrieved June 22, 2021, from <https://www.access-board.gov/prowag/>.
- Ali, L., Valappil, N. K., Kareem, D. N. A., John, M. J., & Al Jassmi, H. (2019). Pavement Crack Detection and Localization using Convolutional Neural Networks (CNNs). *2019 International Conference on Digitization (ICD)*, 217–221. IEEE. doi:10.1109/ICD47981.2019.9105786.
- Alipour, M., Harris, D. K., & Miller, G. R. (2019). Robust Pixel-Level Crack Detection Using Deep Fully Convolutional Neural Networks. *Journal of Computing in Civil Engineering*, 33(6), 04019040. [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000854](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000854)
- Augustaukas, R., & Lipnickas, A. (2019). Pixel-wise Road Pavement Defects Detection Using U-Net Deep Neural Network. In *2019 10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*, 1, 468–471. IEEE. doi:10.1109/IDAACS.2019.8924337.
- Autodesk. (2021). 3D View. <http://help.autodesk.com/view/RECAP/2018/ENU/>.
- Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(12), 2481–2495. <https://doi.org/10.1109/TPAMI.2016.2644615>
- CCRPC, (Champaign County Regional Planning Commission). (2016). Sidewalk Network Inventory and Assessment for the Champaign Urbana Urbanized Area. Retrieved June 22, 2021, from <https://ccrcpc.org/wp-content/uploads/2016/02/SidewalkNetworkInventoryAssessment.pdf>.
- Chen, L. C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). Encoder-decoder with atrous separable convolution for semantic image segmentation. In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Springer Verlag. [https://doi.org/10.1007/978-3-030-01234-2\\_49](https://doi.org/10.1007/978-3-030-01234-2_49)
- Cheng, J. C. P., & Wang, M. (2018). Automated detection of sewer pipe defects in closed-circuit television images using deep learning techniques. *Automation in Construction*, 95(June), 155–171. <https://doi.org/10.1016/j.autcon.2018.08.006>
- Chollet, F. (2020a). Accuracy metrics. Retrieved July 1, 2020, from [https://keras.io/api/metrics/accuracy\\_metrics/](https://keras.io/api/metrics/accuracy_metrics/).
- Chollet, F. (2020b). Concatenate layer. Retrieved June 13, 2020, from [https://keras.io/api/layers/merging\\_layers/concatenate/](https://keras.io/api/layers/merging_layers/concatenate/).
- City of Middleton. (2021a). Sidewalk Deficiencies Examples. Retrieved June 22, 2021, from <https://www.cityofmiddleton.us/271/Sidewalk-Deficiencies-Examples>.
- City of Middleton. (2021b). Sidewalk Maintenance Program. Retrieved June 22, 2021, from <https://www.cityofmiddleton.us/207/Sidewalk-Maintenance-Program>.
- City of Sioux Falls. (2017). Sidewalk Network Inventory and Assessment A Self-Evaluation for the City of Sioux Falls, SD. Retrieved June 22, 2021, from <https://gis.2.siouxfalls.org/data/ada/SelfEvaluationReportSidewalkNetworkInventoryandAssessment.pdf>.
- Cole. (2013). ADA Transition Planning. Retrieved June 22, 2021, from [http://starodub.com/pdf/ADA\\_cutsheet\\_NEW\\_102313.pdf](http://starodub.com/pdf/ADA_cutsheet_NEW_102313.pdf).
- Dadrasjavan, F., Zarrinpanjeh, N., Ameri, A., Engineering, G., & Branch, Q. (2019). Automatic Crack Detection of Road Pavement Based on Aerial UAV Imagery. *Preprints*, 1–16. <https://doi.org/10.20944/preprints201907.0009.v1>
- Department of Justice. (2010). 2010 ADA Standards for Accessible Design. Retrieved June 22, 2021, from <https://www.ada.govregs2010/2010ADAStandards/2010ADAstandards.htm>.
- Dung, C. V., & Anh, L. D. (2019). Autonomous concrete crack detection using deep fully convolutional neural network. *Automation in Construction*, 99(December 2018), 52–58. <https://doi.org/10.1016/j.autcon.2018.11.028>
- Edmondson, V., Woodward, J., Lim, M., Kane, M., Martin, J., & Shyha, I. (2019). Improved non-contact 3D field and processing techniques to achieve macrotexture characterisation of pavements. *Construction and Building Materials*, 227, Article 116693. <https://doi.org/10.1016/j.conbuildmat.2019.116693>
- Fan, R., Bocus, M. J., Zhu, Y., Jiao, J., Wang, L., Ma, F., ... Liu, M. (2019). Road Crack Detection Using Deep Convolutional Neural Network and Adaptive Thresholding. In *2019 IEEE Intelligent Vehicles Symposium (IV), 2019-June(Iv)* (pp. 474–479). IEEE. <https://doi.org/10.1109/IVS.2019.8814000>.
- Frackleton, A., Grossman, A., Palinginis, E., Castrillon, F., Elango, V., & Guensler, R. (2013). Measuring Walkability: Development of an Automated Sidewalk Quality Assessment Tool. *Suburban Sustainability*, 1(1). <https://doi.org/10.5038/2164-0866.1.1.4>
- Gézero, L., & Antunes, C. (2019). Road Rutting Measurement Using Mobile LiDAR Systems Point Cloud. *ISPRS International Journal of Geo-Information*, 8(9), 404. <https://doi.org/10.3390/ijgi8090404>
- Ji, A., Xue, X., Wang, Y., Luo, X., & Xue, W. (2020). An integrated approach to automatic pixel-level crack detection and quantification of asphalt pavement. *Automation in Construction*, 114(March), Article 103176. <https://doi.org/10.1016/j.autcon.2020.103176>
- Jiang, Y., & Bai, Y. (2020a). Determination of Construction Site Elevations Using Drone Technology. In *Construction Research Congress* (pp. 296–305). Reston, VA: American Society of Civil Engineers. <https://doi.org/10.1061/978078442865.032>
- Jiang, Y., & Bai, Y. (2020b). Estimation of Construction Site Elevations Using Drone-Based Orthoimagery and Deep Learning. *Journal of Construction Engineering and Management*, 146(8), 04020086. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0001869](https://doi.org/10.1061/(ASCE)CO.1943-7862.0001869)
- Jiang, Y., & Bai, Y. (2021). Low-High Orthoimage Pairs-Based 3D Reconstruction for Elevation Determination Using Drone. *Journal of Construction Engineering and Management*, 147(9), 04021097. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0002067](https://doi.org/10.1061/(ASCE)CO.1943-7862.0002067)
- Jiang, Y., Bai, Y., & Han, S. (2020). Determining Ground Elevations Covered by Vegetation on Construction Sites Using Drone-Based Orthoimage and Convolutional Neural Network. *Journal of Computing in Civil Engineering*, 34(6), 04020049. [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000930](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000930)
- Jiang, Y., Han, S., & Bai, Y. (2021a). Building and Infrastructure Defect Detection and Visualization Using Drone and Deep Learning Technologies. *Journal of Performance of Constructed Facilities*, 35(6), 04021092. [https://doi.org/10.1061/\(ASCE\)CF.1943-5509.0001652](https://doi.org/10.1061/(ASCE)CF.1943-5509.0001652)
- Jiang, Y., Han, S., & Bai, Y. (2021b). Development of a Pavement Evaluation Tool Using Aerial Imagery and Deep Learning. *Journal of Transportation Engineering, Part B: Pavements*, 147(3), 04021027. <https://doi.org/10.1061/JPEODX.0000282>

- Jiang, Y. (2021a). Dataset. Retrieved July 14, 2021, from <https://www.yuhanjiang.com/dataset>.
- Jiang, Y. (2021b). Sidewalk Management. Retrieved October 23, 2021, from <http://www.yuhanjiang.com/research/IM/SW>.
- Jiang, Y. (2022). Pointcloud2Orthoimage. Retrieved February 22, 2022, from <https://www.yuhanjiang.com/research/FM/PC/P2I>.
- Kearney, S. P., Coops, N. C., Sethi, S., & Stenhouse, G. B. (2020). Maintaining accurate, current, rural road network data: An extraction and updating routine using RapidEye, participatory GIS and deep learning. *International Journal of Applied Earth Observation and Geoinformation*, 87(September 2019), 102031. <https://doi.org/10.1016/j.jag.2019.102031>
- Kim, H., & Ahn, C. R. (2016). Feasibility of Using Pedestrians' Physical Stability to Detect Defects in a Sidewalk. In *16th International Conference on Computing in Civil and Building Engineering* (pp. 1389–1396).
- Kim, H., Ahn, C. R., & Yang, K. (2016). A people-centric sensing approach to detecting sidewalk defects. *Advanced Engineering Informatics*, 30(4), 660–671. <https://doi.org/10.1016/j.aei.2016.09.001>
- Laan Labs. (2021). 3D Scanner App - Help. Retrieved June 22, 2021, from <https://www.3dscannerapp.com/help/>.
- Li, Z., Cheng, C., Kwan, M.-P., Tong, X., & Tian, S. (2019). Identifying Asphalt Pavement Distress Using UAV LiDAR Point Cloud Data and Random Forest Classification. *ISPRS International Journal of Geo-Information*, 8(1), 39. <https://doi.org/10.3390/ijgi8010039>
- Liu, Z., Cao, Y., Wang, Y., & Wang, W. (2019). Computer vision-based concrete crack detection using U-net fully convolutional networks. *Automation in Construction*, 104 (April), 129–139. <https://doi.org/10.1016/j.autcon.2019.04.005>
- Majidifard, H., Adu-Gyamfi, Y., & Buttler, W. G. (2020). Deep machine learning approach to develop a new asphalt pavement condition index. *Construction and Building Materials*, 247(February), Article 118513. <https://doi.org/10.1016/j.conbuildmat.2020.118513>
- Maniat, M. (2019). *Deep Learning-Based Visual Crack Detection Using Google Street View Images* (The University of Memphis). Retrieved from The University of Memphis <https://search.proquest.com/docview/2376764289/accountid=147090%0A>.
- Open3D. (2020). Point cloud. Retrieved June 22, 2021, from <http://www.open3d.org/docs/release/tutorial/geometry/pointcloud.html>.
- OpenCV. (2021a). Contour Features. Retrieved June 22, 2021, from [https://docs.opencv.org/3.4/dd/d49/tutorial\\_py\\_contour\\_features.html](https://docs.opencv.org/3.4/dd/d49/tutorial_py_contour_features.html).
- OpenCV. (2021b). Smoothing Images. Retrieved June 22, 2021, from [https://docs.opencv.org/3.4/dc/dd3/tutorial\\_gaussian\\_median.blur.bilateral.filter.html](https://docs.opencv.org/3.4/dc/dd3/tutorial_gaussian_median.blur.bilateral.filter.html).
- Palmer, B. (2020). Types of Concrete Joints and Their Purpose. Retrieved October 24, 2021, from <https://www.concretenetwork.com/concrete-joints/>.
- Protopapadakis, E., Voulodimos, A., Doulamis, A., Doulamis, N., & Stathaki, T. (2019). Automatic crack detection for tunnel inspection using deep learning and heuristic image post-processing. *Applied Intelligence*, 49(7), 2793–2806. <https://doi.org/10.1007/s10489-018-01396-y>
- Roberts, R., Inzerillo, L., & Di Mino, G. (2020). Exploiting low-cost 3D imagery for the purposes of detecting and analyzing pavement distresses. *Infrastructure*, 5(1). <https://doi.org/10.3390/infrastructures5010006>
- Rodriguez, J. (2021). Types of Concrete Joints and Placing Tips. Retrieved October 24, 2021, from <https://www.thespruce.com/types-of-concrete-joints-845022>.
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (pp. 234–241). Springer Verlag. [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28).
- Shelhamer, E., Long, J., & Darrell, T. (2017). Fully Convolutional Networks for Semantic Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(4), 640–651. <https://doi.org/10.1109/TPAMI.2016.2572683>
- Shorewood Village Board. (2019). Policy No. 36 concrete replacement criteria. Retrieved June 22, 2021, from <https://villageofshorewood.org/DocumentCenter/View/7381/Policy-36-rev-06172019?bidId=>.
- Song, W., Jia, G., Zhu, H., Jia, D., & Gao, L. (2020). Automated Pavement Crack Damage Detection Using Deep Multiscale Convolutional Features. *Journal of Advanced Transportation*, 2020(ii), 1–11. <https://doi.org/10.1155/2020/6412562>
- Starodub Inc. (2009a). Starodub's Ultra-Light Inertial Profiler ULIPr (RoLine Laser Line Scan Version). Retrieved June 24, 2021, from <http://www.starodub.com/pdf/ULIPr.pdf>.
- Starodub Inc. (2009b). ULIP Ultra-Light Inertial Profiler. Retrieved June 22, 2021, from <http://www.starodub.com/ulip.html>.
- Tan, Y., Cai, R., Li, J., Chen, P., & Wang, M. (2021). Automatic detection of sewer defects based on improved you only look once algorithm. *Automation in Construction*, 131 (August), Article 103912. <https://doi.org/10.1016/j.autcon.2021.103912>
- Village of Shorewood. (2021). Sidewalk Program. Retrieved June 30, 2021, from <https://villageofshorewood.org/262/Sidewalk-Program>.
- Wang, M., & Cheng, J. C. P. (2020). A unified convolutional neural network integrated with conditional random field for pipe defect segmentation. *Computer-Aided Civil and Infrastructure Engineering*, 35(2), 162–177. <https://doi.org/10.1111/mice.12481>
- Wikipedia. (2022a). Lidar. Retrieved May 28, 2022, from <https://en.wikipedia.org/wiki/Lidar>.
- Wikipedia. (2022b). Structure from motion. Retrieved May 28, 2022, from [https://en.wikipedia.org/wiki/Structure\\_from\\_motion](https://en.wikipedia.org/wiki/Structure_from_motion).
- Wu, C. (2011). VisualSfM : A Visual Structure from Motion System. Retrieved June 22, 2021, from <http://cvcu.me/vsfm/index.html>.
- Yang, Q., Shi, W., Chen, J., & Lin, W. (2020). Deep convolution neural network-based transfer learning method for civil infrastructure crack detection. *Automation in Construction*, 116(May), Article 103199. <https://doi.org/10.1016/j.autcon.2020.103199>
- Yates, B., Fouts, L., Sehgal, S., & Mcren, D. (2017). Sidewalk Repair Cost Estimation Project. Retrieved June 22, 2021, from [http://icap.sustainability.illinois.edu/files/projectupdate/4207/ADA\\_Sidewalk\\_repair\\_costs.pdf](http://icap.sustainability.illinois.edu/files/projectupdate/4207/ADA_Sidewalk_repair_costs.pdf).
- Zhang, K., Zhang, Y., & Cheng, H.-D. (2020). CrackGAN: Pavement Crack Detection Using Partially Accurate Ground Truths Based on Generative Adversarial Learning. *IEEE Transactions on Intelligent Transportation Systems*, 1–14. <https://doi.org/10.1109/tits.2020.2990703>
- Zhao, H., Shi, J., Qi, X., Wang, X., & Jia, J. (2017). Pyramid Scene Parsing Network. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017-Janua* (pp. 6230–6239). IEEE. <https://doi.org/10.1109/CVPR.2017.660>
- Zhi, X. (2019). Implementation of deep learning framework – Unet, using Keras. Retrieved July 1, 2020, from <https://github.com/zhihuao/unet>.
- Zhou, S., & Song, W. (2020b). Robust Image-Based Surface Crack Detection Using Range Data. *Journal of Computing in Civil Engineering*, 34(2), 04019054. [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000873](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000873)
- Zhou, S., & Song, W. (2020a). Deep learning-based roadway crack classification using laser-scanned range images: A comparative study on hyperparameter selection. *Automation in Construction*, 114(October 2019), 103171. <https://doi.org/10.1016/j.autcon.2020.103171>
- Zou, Q., Zhang, Z., Li, Q., Qi, X., Wang, Q., & Wang, S. (2019). DeepCrack: Learning Hierarchical Convolutional Features for Crack Detection. *IEEE Transactions on Image Processing*, 28(3), 1498–1512. <https://doi.org/10.1109/TIP.2018.2878966>