Construction and Building Materials 324 (2022) 126717

Contents lists available at ScienceDirect

Construction and Building Materials

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

Deep learning based automated segmentation of air-void system in hardened concrete surface using three dimensional reconstructed images

Jueqiang Tao [a], Haitao Gong [a], Feng Wang [a,*], Xiaohua Luo [a], Xin Qiu [b], Jinli Liu [c]

[a] Ingram School of Engineering, Texas State University, San Marcos, USA
[b] Department of Road and Traffic Engineering, Zhejiang Normal University, Jinhua, China
[c] Department of Geography and Environmental Studies, Texas State University, San Marcos, USA

ARTICLE INFO

Keywords:
3D reconstruction
DCNN
Semantic segmentation
Air voids
Hardened concrete

ABSTRACT

The automated air-void detection methods specified in the ASTM C457 require the aid of contrast enhancement which is time consuming and labor intensive. This study investigated the utilization of three-dimensional (3D) reconstruction and Deep Convolution Neural Network (DCNN) methods to detect the air voids in hardened concrete surfaces without the use of contrast enhancement. The experimental results showed that the DCNN could accurately distinguish air voids from hardened concrete images with the detection accuracy of over 0.9 in only less than a minute. The accuracy rates for air content, specific surface, and spacing factor were 0.92, 0.91, and 0.89, respectively.

1. Introduction

Air voids are small air bubbles that are embedded in the hardened concrete paste. The existence of a well-distributed air-void system in a hardened concrete pavement is critical for maintaining the freeze-thawing performance of the concrete infrastructure. Air-void petrographic analysis provides evidence on whether a concrete infrastructure can resist freeze–thaw damages in cold regions. Among the air-void analysis procedures, the detection of air voids in the hardened concrete surface is one of the most important requirements. Well-trained petrographic raters are required for conducting the petrographic analysis, while the analysis results are subjective and could lead to biased concrete performance estimations [1]. In addition, more than 2,000 air-void observations are required for each concrete sample, which makes the air-void analysis a time-consuming and labor-intensive process [1].

Numerous automating efforts have been conducted to reduce human labor by adopting contrast-enhancement methods and computer image analysis techniques. The conventional contrast-enhancement method applies black ink and white powders to make the concrete surface black and air voids white. The RapidAir 457 procedure proposed by Pade et al. was the most widely used contrast-enhancement based automated system [2]. The prepared concrete surface needs to be carefully examined by a trained petrographer using a stereomicroscope to evaluate: 1) if the surface enhancement is acceptable and 2) more importantly to blacken and cover any pores and cracks present in coarse and fine aggregates with a very thin tip black sharpie. The inappropriate pigments or powder sizes adopted for contrast enhancement can lead to a biased result [3]. Some dark and thick pigments can fill up tiny air voids and make the air voids inaccessible to the white powders [4]. The grain size of powders is decisive for the minimum air-void size that can be observed. Fine powder particles may get stuck in tiny defects on the concrete surface and are hard to be wiped away, which lead to an overestimation of the air-void content. Whereas, coarse powders can create a 'wall effect' around the perimeter of small air voids and cause an underestimation of the air-void content [3]. X-ray Computed Tomography (CT) is a typical non-destructive method that has been widely adopted for air-void analysis [5]. However, the high expense of using X-ray CT equipment makes it unlikely for practical use.

The air voids in a hardened concrete surface have similar greyscale values as the paste in the concrete surface. However, unlike the solid phase which provides only color information, the air voids are hollows in the hardened concrete surface. The three-dimensional (3D) information (e.g. depth and gradient) of air voids is valuable for air voids detection. Consequently, a new approach to automated air-void segmentation that is free of contrast enhancement is using 3D reconstruction techniques. Wolter et al. [4] proposed an air-void segmentation method without contrast enhancement by introducing the surface profile information. The photogrammetry method was adopted for

* Corresponding author at: Ingram School of Engineering, 601 University Dr., San Marcos, TX 78666, USA.
E-mail address: f_w34@txstate.edu (F. Wang).

https://doi.org/10.1016/j.conbuildmat.2022.126717
Received 14 November 2021; Received in revised form 26 January 2022; Accepted 31 January 2022
Available online 4 February 2022

estimating the depth of each pixel in the concrete images. However, the proposed method was not accurate enough. Tao et al. [6] captured the 3D concrete surface information using the photometric stereo method. The 3D surface captured valuable information for air-void segmentation. The 3D reconstruction techniques provided an inexpensive solution to automated air-void segmentation without using contrast enhancement. However, reflective concrete surface and transparent aggregates still cause inaccuracies to 3D reconstruction, and impact the segmentation results [6].

As a subarea in artificial intelligence (AI), deep learning (DL) has achieved great success in semantic segmentation. During the semantic segmentation process, a classification label is predicted on each pixel, which may greatly fulfill the objectives of air-void segmentation. The deep convolutional neural networks (DCNN), which is an important branch in DL, shows good potential in detecting target objects in noisy images at pixel resolution. Another advantage of DCNN is the end-to-end segmentation manner. Significant fewer human interventions are required to be involved in the segmentation work. Song et al. [7] made an innovative attempt at using DCNN method to segment the phases in hardened concrete surfaces. The proposed air-void segmentation method outperformed the contrast-enhancement method and achieved much higher accuracy in the boundary area between each phase. However, the air voids still need to be highlighted using orange chalk powders, which made the method not fully automatic.

The objective of this paper is to propose an end-to-end automated segmentation method that could detect air voids in concrete surfaces without contrast enhancement. In the study, an air-void image dataset including a set of surface normal images and air-void annotations was first developed. The surface normal images were generated from the surface normal vectors of concrete surfaces which were estimated using a 3D reconstruction technique. The annotations were first obtained using the contrast-enhancement method and then refined by human raters. Consequently, as a DCNN model, U-Net was trained using the air-void dataset for detecting air voids in hardened concrete images. Finally, the segmentation results were manually evaluated using a set of accuracy measurement indexes.

2. Literature review

2.1. 3D reconstruction methods

Generally, the state-of-the-art 3D reconstruction approaches can be classified as passive and active methods. Passive 3D reconstruction techniques reconstruct the 3D surface of an object without introducing new energy into the environment [8]. Numerous technologies and methods employed this approach, including multi-view stereo [9], structure from motion [10], light-field cameras [11], and space-carving techniques [12]. Binocular stereo is the most common multiview stereo approach [13]. Two cameras are utilized to capture pictures from slightly different two viewpoints. By analyzing the disparity between the objects in the two pictures, the relative depth can be calculated. However, calculating the disparity is not always straightforward for a computer vision system. The well-known correspondence problem induces the difficulty in locating matching points in the two images. In the case when an object surface with low texture needs to be captured, the structure light technique, which is one of the active 3D imaging approaches, can be introduced to aid the 3D reconstruction by creating projected light patterns. The projected patterns help the computer vision system to solve the correspondence problem. Wolter et al. [4] made the first study on 3D air-void segmentation by using a photogrammetry method. The working principle of photogrammetry is similar to binocular stereo. The research investigated the potential of utilizing 3D reconstruction for air-void segmentation, but large differences were observed between the experiment results and the ground truth. A possible reason could be that the concrete surfaces were texture-less, which increased the difficulty of solving the correspondence problem. Space-carving and light-field systems can overcome the correspondence problem. However, space-carving systems require many different views and may fail to reconstruct the crowded areas [11]. Light-field systems rely on expensive camera technology to capture high-resolution data and thus make the air-void analysis system not cost-effective [14].

Active 3D imaging approaches introduce outside energy sources to help 3D reconstruction and overcome many problems of the passive approaches. The time-of-flight 3D laser scanner is an active scanner that uses laser light to probe the subject [15]. The core of this type of scanner is a time-of-flight laser rangefinder. The laser rangefinder finds the distance of a surface by timing the round-trip time of a pulse of light. The laser rangefinder only detects the distance of one point in its direction of view. Thus, the scanner scans its entire field of view one point at a time by changing the range finder's direction of view to scan different points. The advantage of time-of-flight range finders is that the method is capable of operating over very long distances [16]. The disadvantage of time-of-flight range finders is relatively low accuracy and low resolution. Due to the high speed of light, timing the round-trip time is difficult and the accuracy of the distance measurement is relatively low. Triangulation is another active 3D imaging approach [17]. Triangulation laser emits a laser on the subject and exploits a camera to look for the location of the laser dots. Depending on how far away the laser strikes a surface, the laser dot appears at different places in the camera's field of view. The laser dot, the camera, and the laser emitter form a triangle, which makes the depth measurement possible. Triangulation laser scanners are susceptible to occlusions and air-void regions may be obscured by air-void edges. Photometric stereo is an active imaging technique that is low-cost and can achieve high image resolutions and fast capture speeds [18]. The photometric stereo method estimates the 3D surface of objects based on the relationship between image intensity and the surface normal under various lighting directions. The photometric stereo method has the key advantage of achieving automation while reducing test time, which is a cost-effective and real access to high-resolution 3D images, easy to implement, and robust to reconstruct on textured or texture-less surfaces.

2.2. Deep semantic segmentation methods

As a subset of machine learning, the deep learning based semantic segmentation, which aims to classify an image at pixel resolution, has achieved significant success in many image segmentation related fields, including autonomous driving [19], pavement condition survey [20], face recognition [21], and image search engines [22]. The key advantage of deep learning based semantic segmentation techniques is the ability to learn appropriate feature representation of pixels in each category in an end-to-end manner. The deep learning techniques substantially improved the accuracy and efficiency of a semantic segmentation task.

R. Girshick et al. [23] proposed a region-based convolutional neural networks (RCNN). The method first utilized selective search [24] to extract numerous object regions, and then a set of features were extracted from each of the extracted regions. Finally, a classifier was utilized to classify regions into each category. Compared with conventional hand-crafted methods, the RCNN was able to address more complicated tasks and achieved a higher accuracy. A 30% improvement was found compared with the previous best model. However, the RCNN also suffers from many drawbacks for image segmentation tasks. Hariharan et al. [25] argued that the network of RCNN was actually fine-tuned to classify bounding boxes, making it suboptimal to extract foreground features. Guo et al. [26] stated that the features extracted by RCNN did not contain sufficient spatial information, which leads to fuzzy boundaries in segmented images. Many improvements have been made to address these issues [25,27,28]. Long et al. [29] proposed the Fully Convolutional Network (FCN) which was the basis of many state-of-the-art deep learning based semantic segmentation methods. They replaced the fully connected layers of various CNNs like AlexNet [30],

Table 1
Description of experimental specimens.

Specimen No.	Total scan	Description of material constituents	
		Coarse aggregate	Fine aggregate
1	3 1(test)	Limestone	Quartz, limestone, chert, granite and feldspar
2	2 1(test)	Limestone	Manufactured sand, quartz, feldspar and chert
3	1 1(test)	Limestone, quartz and chert	Quartz, limestone, sandstone, igneous and siliceous
4	3 1(test)	Limestone, siliceous, igneous chert and quartzite	Limestone, siliceous, igneous chert and quartzite
5	3 1(test)	Sandstone, limestone and igneous	Quartz, limestone, sandstone, igneous and siliceous

VGG [31], GoogLeNet [32], and ResNet [33] with fully convolutional layers. The structure first realized end-to-end image semantic segmentation at pixel level. While the conventional FCN model did not consider the global context information, the model inherently limited the spatial precision for semantic segmentation. Mostajabi et al. [34] and Szegedy et al. [35] illustrated the importance of adopting global context information for accurate image segmentation. As for semantic segmentation, per-pixel classification was often ambiguous in the presence of only local information. However, the task became much simpler if contextual information, from the whole image, was available. Chen et al. [36] introduced conditional random field (CRF) into FCN and proposed DeepLab. The CRF significantly refined object boundaries in the segmented image with improved efficiency. Dilated convolutions [37] expanded the receptive field of CNN by enlarging convolution filters without increasing parameters. The key advantage of dilated convolutions was improving the ability of global information integration without additional computation cost. The multi-scale context aggregation module [38], improved DeepLab [39], and the ENet [40] all adopted dilated convolutions as a method to integrate global information. Feature fusion is another way to enlarge the receptive field of CNN. Liu et al. [41] proposed the ParseNet which concatenated global features with local features to form combined features. The combined features were then convoluted for classification. Chen et al [39] utilized Atrous Spatial Pyramid Pooling to combine the output of dilated convolutions with various dilation rates together to enlarge the field-of-view without increasing the number of parameters.

3. Data acquisition

3.1. Hardened concrete samples

The hardened concrete specimens were provided by the concrete laboratory of the Texas Department of Transportation (TxDOT). Considering the fact that the appearance of the concrete specimens could affect the semantic segmentation results, the concrete specimens with various aggregate types and cement types were selected as experimental specimens. All the selected concrete specimens were drilled and sampled from an in-service concrete pavement structure. The details of the experimental specimens are described in Table 1. The specimen surfaces were polished according to the specifications in ASTM C457. To fit the field-of-view of the photometric stereo system, the original samples were sub-sliced into 4 cm × 5 cm small pieces and a region of 3 cm × 4.5 cm was captured as the field-of-view. Consequently, A total of 12 pieces of sliced concrete samples were utilized for the training purpose. For each category of the hardened concrete samples, one slice was utilized for testing purposes. There was a total of 5 pieces of sliced concrete samples utilized for the testing purpose.

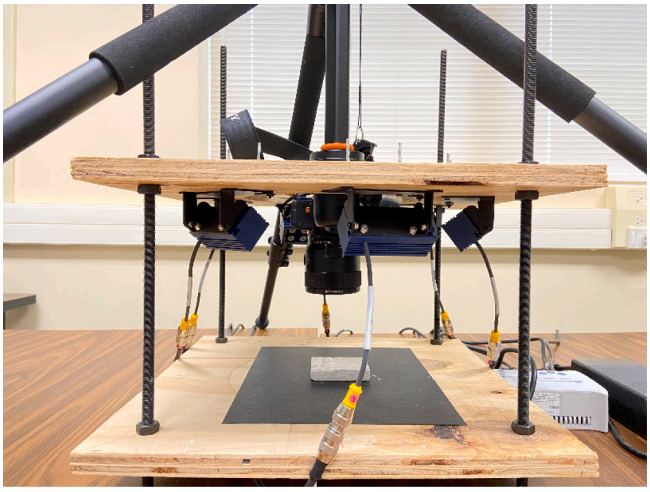

Fig. 2. Photometric stereo system with six LED lights.

3.2. Data annotation and registration

The labels of training data were first annotated using a contrast-enhancement method, and then manually refined. Acrylic ink and a rubber brayer were adopted to blacken the polished concrete surfaces [3]. The applied acrylic ink could generate a thin dark layer without filling out air voids. In case some aggregates cannot be ideally painted, the missed regions were carefully re-painted by a marker pen. The specimens were then left to air dry at room temperature for 30 min. After the ink was dried thoroughly, a barium sulfate powder with an average particle size of 3 μm was used to highlight the air voids into white color. The barium sulfate powders were scattered on the hardened concrete surface and then pressed into air voids using hand fingers. The excess

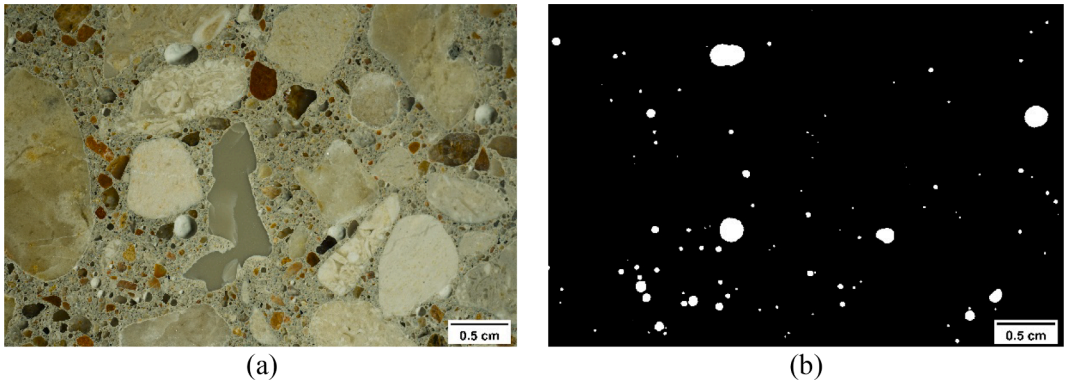

Fig. 1. An example of manually annotated air-void image: (a) Hardened concrete surface, (b) Annotated air voids.

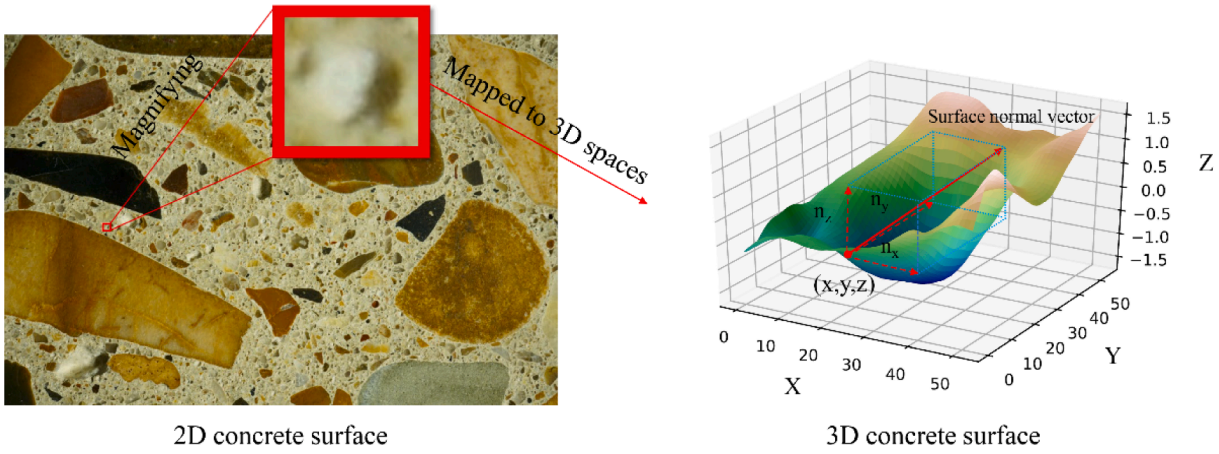

Fig. 3. Illustration of surface normal vector on concrete surface. n_x, n_y, and n_z are the components of surface normal vector in × y and z directions at the point (x, y,z.)

powders were removed with the edge of a silicone spatula. The images of the contrast enhanced concrete surface were captured using the photometric stereo system that is shown in Fig. 2. The system consisted of six LED lights and was designed for illuminating the concrete samples in different directions. The details of the setup and utilization of the photometric stereo system are described in '3D Reconstruction – Photometric Stereo' section. All six LED lights were lighted to generate a uniform illumination on the concrete surfaces. Finally, an image-processing software (ImageJ) was used to segment the air voids from the enhanced concrete images by setting a gray value threshold. The Otsu method was utilized to provide an optimal threshold. In the case when the Otsu method did not generate an ideal threshold, the generated threshold may be manually adjusted. The non-air-void regions in the concrete images such as cracks, voids in aggregates and the region with residual barium sulfate powder were double-checked and removed by the rater using Adobe Photoshop. It is worth mentioning that the contrast enhancement process is only used for data preparation. During the 3D reconstruction and image segmentation process, no contrast enhancement is required for concrete samples.

The raw concrete images and the enhanced concrete images were

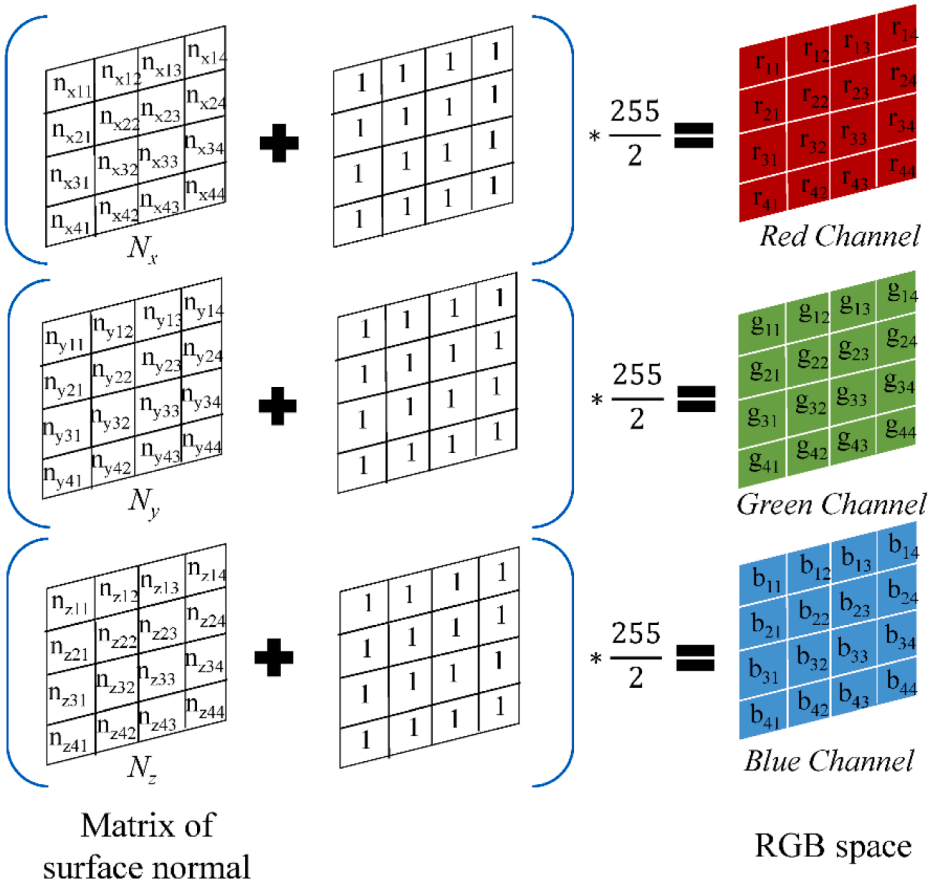

Fig. 4. Mapping surface normal to RGB space. N_x, N_y, and N_z are surface normal matrices that contain the components that are involved in ×, y, and z directions. The *r*, *g*, and *b* are pixel intensities in red, blue, and green channels. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

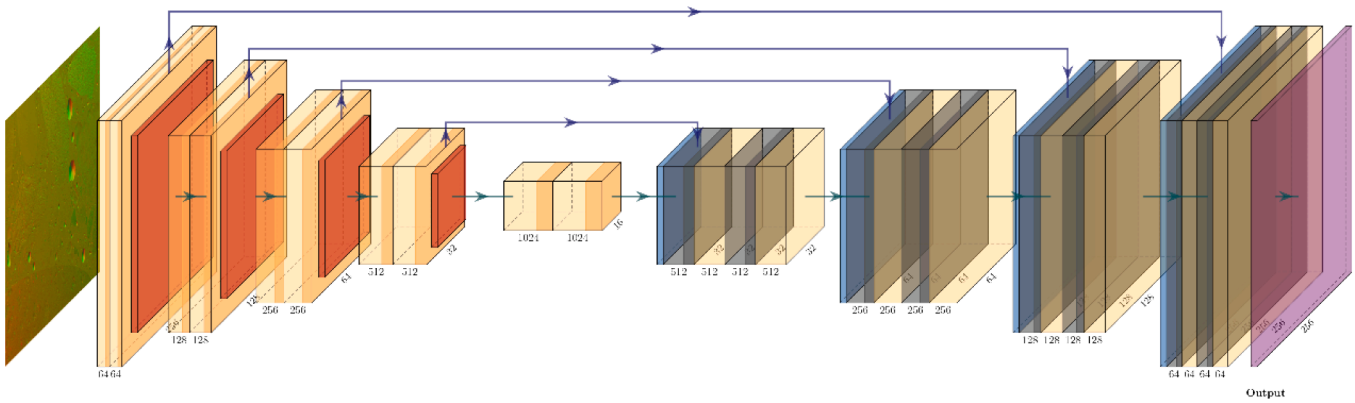

Fig. 5. U-Net for image with resolution of 256 pixels × 256 pixels.

captured in two different scans in sequence. The hardened concrete surface was first scanned using the photometric stereo system to obtain the 3D surface normal image of the hardened concrete surface. The concrete specimen was then taken away from the testbed of the photometric stereo system for contrast-enhancement procedure. After the enhancement procedure, the concrete sample was relocated to the testbed and scanned to capture the contrast-enhanced image. Even though careful locating was exercised to ensure the hardened concrete was aligned to the same position as the first scan, slight displacements were still observed between the two scans. The DCNNs require accurate annotations, and the labels and image features are expected to correspond at pixel level. Therefore, the images of the two scans were manually adjusted using Adobe Photoshop to match up at each pixel in the two scans. An example of a hardened concrete image and its annotated air-void image is shown in Fig. 1.

4. Methodology

4.1. 3D reconstruction – Photometric stereo

As discussed in the research by Tao et al., various photometric stereo methods were compared for the extraction of 3D air-void information [6]. The conventional photometric stereo method that was proposed by Woodham outperformed the other photometric stereo methods and could extract the gradient of air voids [6]. The Woodham's photometric stereo method, which is shown in Equation (1), utilizes the relationship between incoming lighting direction $L \in \mathscr{R}^{k\times 3}$, surface normal $N \in \mathscr{R}^{3\times 1}$, and observed intensity $I \in \mathscr{R}^{k\times 1}$ to compute the surface normal of each pixel [18].

$$\begin{bmatrix} i_1 \\ i_2 \\ \vdots \\ i_k \end{bmatrix} = \begin{bmatrix} L_1 \\ L_2 \\ \vdots \\ L_k \end{bmatrix} \cdot \begin{bmatrix} n_x \\ n_y \\ n_z \end{bmatrix} \tag{1}$$

$k \in \mathscr{R}$ is the number of lighting directions. In this study, a

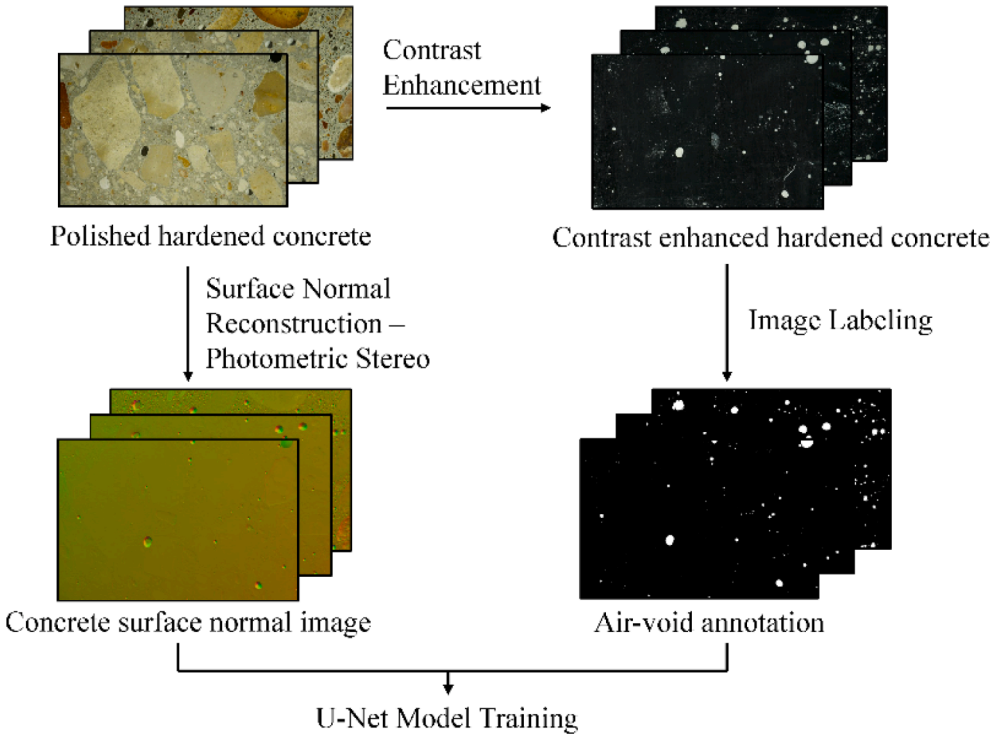

Fig. 6. Steps of U-Net model training. The concrete surface normal images and air-void annotations were utilized for the training process.

Table 2
Augmentation strategies.

Augmentation Strategy	Flipping	Rotating	Scaling	FillMode	Images for Training
Augmentation 1	√	√	√	wrap	27,000
Augmentation 2	√	√	√	constant	27,000
Augmentation 3	√	×	×	×	18,000

photometric stereo system with 6 LED lights, as shown in Fig. 2, was used for 3D reconstruction. A Sony Charge-Coupled Device (CCD) camera with a resolution of 42 megapixels was adopted for concrete surface image capturing. The system achieved a resolution of 5.6 μm/pixel for concrete images. The details of the photometric stereo system were presented in a previous paper by the authors [6].

A diagram of an estimated surface normal vector on hardened concrete surface is shown in Fig. 3. The components n_x, n_y, and n_z in the computed surface normal vector were then normalized to (-1, 1) scale. Whereas the DCNNs are designed for RGB images and the intensity of pixels in each channel is between 0 and 255. To ensure the surface normal compatible with the DCNNs, the n_x, n_y, and n_z of each pixel were mapped from (-1, 1) to (0, 255). An example of the mapping process with a 4 pixels × 4 pixels image is shown in Fig. 4. For example, n_{x11}, n_{y11}, and n_{z11} are the components of the surface normal at pixel (1,1) on ×, y, and z directions. The r_{11}, g_{11}, and b_{11}, which are mapped by n_{x11}, n_{y11}, and n_{z11}, are the pixel intensity of the pixel (1,1) in red, blue, and green channels, respectively.

4.2. Semantic segmentation – U-Net

U-Net, which is a variant of FCN and improved with skipped connections, was adopted for air-void segmentation in this study [42]. As shown in Fig. 5, the U-Net consists of an encoder structure and a decoder structure. Skip connections between the encoder and decoder combine lower-level features with higher-level features. The combined features can improve pixel-level localization. The U-Net architecture has been validated to be powerful for image segmentation. It is currently one of the most used algorithms in biomedical image segmentation [43] and has been successfully extended to the other semantic segmentation tasks in many other fields [44]. In addition, the U-Net model can generate a comparable result using a small dataset. In the study, the algorithm was coded and implemented with TensorFlow, an open-source deep learning library in Python. The training processes were conducted on the Google Colab Pro which provides Graphics Processing Units (GPUs) for deep learning purposes and one NVIDIA® Tesla® V100 GPU with 16 GB of RAM. A flow chart that includes the major works of training the U-Net model is shown in Fig. 6.

There are 1,941,105 trainable parameters incorporated in the U-Net model. The images were randomly cropped into 256 pixels × 256 pixels small pieces and 10,200 cropped images were generated. 80% of the crops were adopted as training data and 20% of the crops were adopted as validation data. During the training process, cross-entropy was selected as the loss factor to evaluate the discrepancy between the training results and labels after each epoch. The Adam optimizer was adopted for updating the weights in U-Net.

4.3. Image Augmentation

To make the training process more efficient, a set of data augmentation procedures were conducted on the training data. Considering the shapes of the air voids can be a critical differentiation factor for distinguishing the air voids, the ineffective modifications can lead to a decreased segmentation accuracy. For example, the compressing and stretching modifications, which change the height and width ratio of the air-void image, change the shape of the air voids and thus make the air voids confused with the air-void like noises. In this research, the effectiveness of using various image augmentation procedures was evaluated. Random combinations of flipping, rotating, and scaling modification were applied. Scaling and rotating operations significantly increased the number of images for the training. However, both scaling and rotating operations generated a margin between the augmented images and the edges of the pictures. Consequently, two fill modes were introduced to fill the margins. Three image augmentation strategies, as shown in Table 2, were developed to investigate the best augmentation strategy for the segmentation of air-void images. An example of the augmented image using different augmentation strategies is shown in Fig. 7.

4.4. Accuracy measurement

An accuracy measurement procedure that is similar to the Modified Point Count method [45] was utilized for evaluating the accuracy of the proposed method. Song et al. also adopted a similar accuracy measurement procedure in one of their previous research studies [7]. A 100 × 100 dot matrix was generated and appended to both the segmented images and the raw concrete images. The pixels in the raw concrete images that are corresponding to the appended dot-matrix were manually observed by an experienced petrographer. According to the observation, the dots in the dot matrix were labeled as air voids and non-air voids. The dots in the dot matrixes that were appended to the segmented images were also labeled by identifying the color of the corresponding pixel in the segmented images. The dot that was appended to a white pixel was labeled as air voids. The dot that was appended to a black pixel was labeled as non-air voids. The labeling process for the segmented images was done automatically using a program coded in Python. Consequently, accuracy measurements including MIoU (Mean of Intersection over Union), P (precision), R (Recall), and F_1, which can be calculated by Equations 2–7, were utilized to evaluate the accuracy of the segmentation results.

$$IoU_{airvoids} = \frac{TP}{TP + FP + FN} \tag{2}$$

$$IoU_{non-air-voids} = \frac{TN}{TN + FN + FP} \tag{3}$$

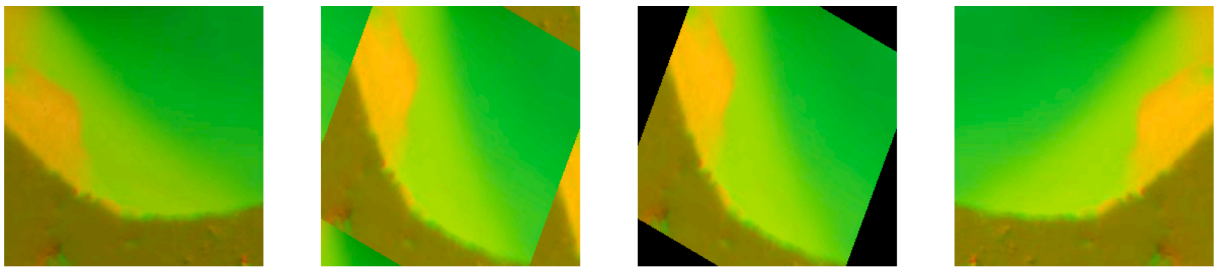

Fig. 7. Augmented images using different augmentation strategies: (a) Original image, (b) Augmentation 1, (c) Augmentation 2, (d) Augmentation 3.

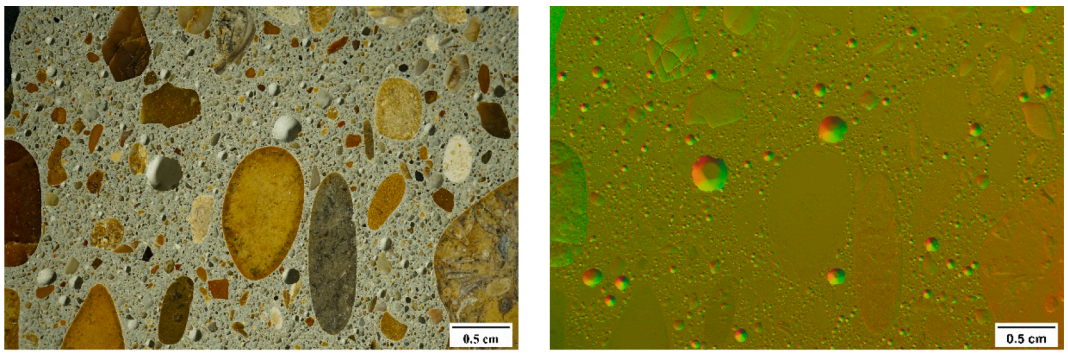

Fig. 8. The air voids in original concrete surface image and surface normal image: (a) Original concrete surface image, (b) Mapped surface normal image.

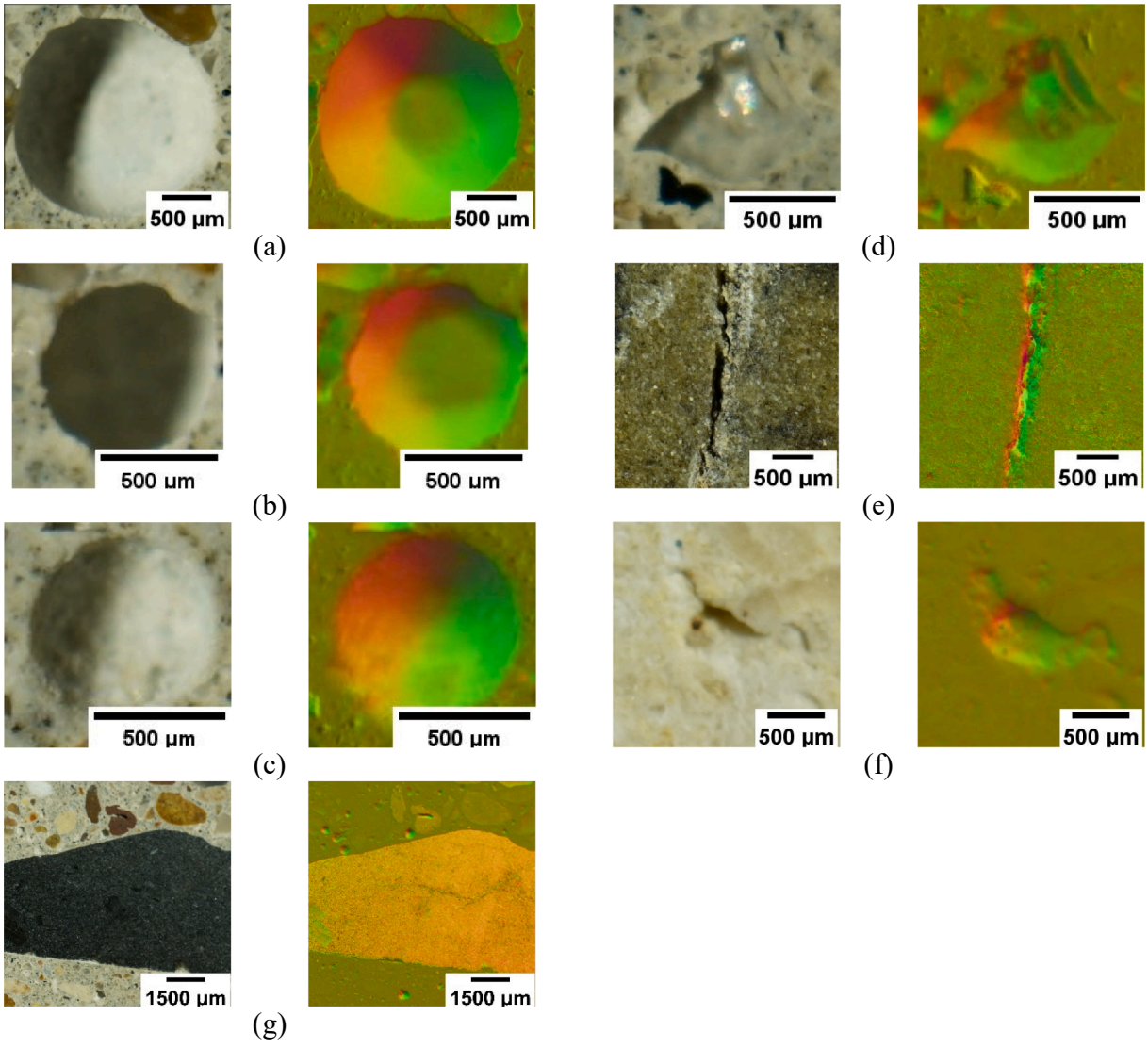

Fig. 9. Air-void appearances and air-void like noises generated by components on concrete surfaces. (a) Air void 1, (b) Air void 2, (c) Air void 3, (d) Transparent aggregate, (e) Cracks in aggregate, (f) Void in aggregate, (g) Drak aggregate.

$$MIoU = \frac{IoU_{airvoids} + IoU_{non-airvoids}}{2} \quad (4)$$

$$P = \frac{TP}{TP + FP} \quad (5)$$

$$R = \frac{TP}{TP + FN} \quad (6)$$

$$F_1 = \frac{2TP}{2TP + FP + FN} \quad (7)$$

Table 3
Training results of U-Net model using different augmentation strategies.

Augmentation Strategy	Epochs	Learning Rate	Best Train MIoU	Best Val MIoU
Augmentation 1	780	0.001	0.681	0.610
Augmentation 2	780	0.001	0.691	0.647
Augmentation 3	780	0.001	0.814	0.706

TP is the percentage of dots that are correctly segmented as air voids; *TN* is the percentage of dots that are correctly segmented as non-air-void; *FP* is the percentage of dots that are incorrectly segmented as air voids; *FN* is the percentage of dots that are incorrectly segmented as non-air-void.

5. Results and analysis

5.1. 3D reconstruction results of photometric stereo method

An example of the raw concrete surface and the mapped surface normal image is shown in Fig. 8. Compared with the original concrete surface image, the surface normal image increased the contrast in uneven areas. The areas with a slant surface normal can be distinguished by identifying the color changes on the surface normal map. The areas with uniform pale green are the solid phase (aggregates and paste). The round areas with large color variations in a circle are the air voids.

As shown in Fig. 8, the air voids in the mapped surface normal image present a clear pattern and can be easily identified by naked eyes. Fig. 9 presents various appearances of air voids and air-void like noises on concrete surface normal images. As shown in Fig. 9, the regions of some dark or transparent aggregates also present a variation of color. The photometric stereo method estimates the surface normal of a target object by the intensity of reflected light. Under various lighting directions, a slant surface presents a great intensity variation, while a flat surface generates an identical surface intensity. The dark aggregates were apt to produce specularities under a specific lighting angle and thus lead to a biased slant surface normal estimation. For the transparent aggregates, the lights are transmitted down to the bottom of the aggregates and reflected by the paste. Biased slant surface normal estimations were produced by the transparent aggregates. Consequently, the photometric stereo method inaccurately estimates the normal information in the region within some transparent aggregates and dark aggregates. The biased estimation generated air-void like appearances in the surface normal map. The similarity made the automated identification of air voids in hardened concrete a challenge. In addition, the air voids are not the only 'hollows' in concrete surfaces. The voids and cracks in aggregates are another kind of 'hollows' in concrete surfaces that can be mistaken as air voids.

The air voids with different sizes and depths presented different appearances. Inside some deep air voids, a 'flat region' can be observed. The 'flat regions' were caused by occlusions. The lights were blocked by the edge of air voids and did not reach the bottom of the air voids. Thus, the photometric stereo system failed to capture the information at the bottom of the air voids. The system assumed those regions as a flat plane because there was no intensity variation captured in the blocked regions. On the other hand, in shallow air voids, the color variation was too little and sometimes can be mistaken as non-air-void regions. Therefore, even though the air voids were highlighted in the surface normal image, the diversity of air-void appearances and various air-void like noises made the air-void detection a tough work.

5.2. Evaluation of different Augmentation strategies

Table 3 presents the evaluation results of the U-Net model trained with different image augmentation strategies. All 3 models were trained for 780 epochs to ensure convergence. As shown in Table 3, the training result of Augmentation 1 is almost the same as the evaluation result of Augmentation 2. Whereas a 3.7% difference is observed between the evaluation results of validation data. The comparison between Augmentation 1 and Augmentation 2 indicates that the 'constant' fill mode helps the U-Net model improve robustness. By comparing Augmentations 1 and 2 to Augmentation 3, the U-Net model which was trained with the images augmented with the strategy of Augmentation 3, is significantly better than the other two U-Net models. It is worth noting that the number of images generated by Augmentation 3 is 2/3 of the total images generated by Augmentation 1 or Augmentation 2. A reasonable explanation for the observation is that the shape information, which is important for distinguishing of air voids, is well retained in the training data using the strategy of Augmentation 3. Consequently, the strategy of Augmentation 3 was used in the research.

5.3. Air-void segmentation results of U-Net model

The U-Net was trained using surface normal images and annotated air-void masks. The model was trained for 780 epochs and the training work took about 15 h to accomplish. The variations of MIoU and loss during the training process were recorded in each epoch and presented in Fig. 10.

The loss estimates the discrepancy of predicted results and labels, and a lower loss indicates a better segmentation performance. The loss

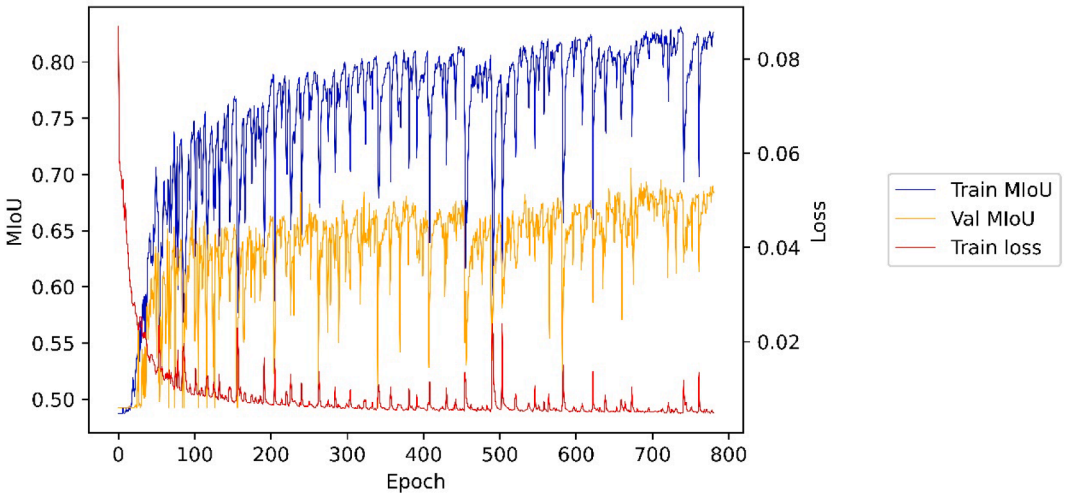

Fig. 10. The loss and MIoU curves of the U-Net model training.

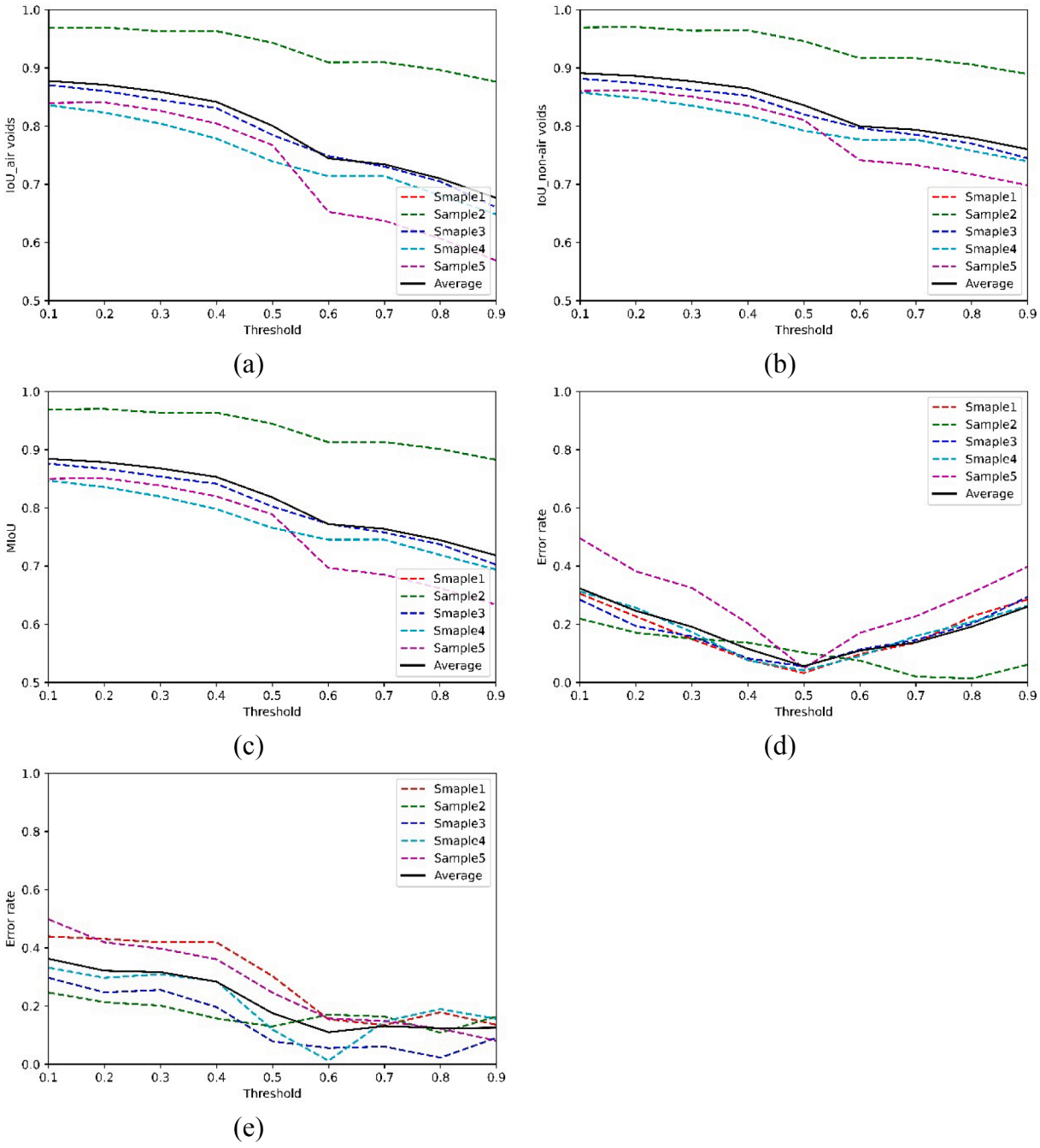

Fig. 11. Accuracy measurement of test samples with different thresholds: (a) $IoU_{air\ voids}$, (b) $IoU_{non\text{-}air\ voids}$, (c) MIoU, (d) Air content, and (e) Spacing factor.

curve presents some fluctuations and shows a decreasing trend. The variation trend of the loss curves slows down with the increase of the number of epochs and becomes stabilized when the loss value approaches 0.005. The train_MIoU and val_MIoU measure the similarity of predicted results and labels for the training set and validation set, respectively, and a higher MIoU indicates a better segmentation performance. The MIoUs of the training dataset and validation dataset are combining, and an increasing trend is found for both curves. The loss curve reversely correlates with the MIoU curves. As shown in Fig. 10, the decreasing trend of the loss curve stabilizes after 650 epochs, which indicates that the model is converged.

The size of the input of the U-Net model was 256 pixels × 256 pixels which was much smaller than the size of the raw concrete surface images. Consequently, the raw concrete surface images were cropped into 256 pixels × 256 pixels small pieces and fed into the trained U-Net, one cropped image at a time to generate an output. After all the cropped images were processed, the processed images were then stuck together as the segmentation result of the raw concrete surface image. The U-Net model computed on each pixel and output a 256 × 256 matrix. Each value in the matrix was from 0 to 1, where 0 means the pixel was most likely from the background (non-air-void) and 1 means the pixel was most likely from the target object (air voids). Some raw output values that are close to 0 are found inside of some air-void regions. To make a binary air-void segmentation, the selection of a threshold between 0 and 1 is needed to classify a pixel as air void or non-air void. The accuracy measurements of the test samples with different thresholds are presented in Fig. 11. As the threshold increases, the IoUs show slight downward trends. Consequently, 0.1 was chosen as the threshold for achieving the best MIoU. To evaluate the influence of the threshold to air-void parameters, the estimated air-void parameters of each threshold

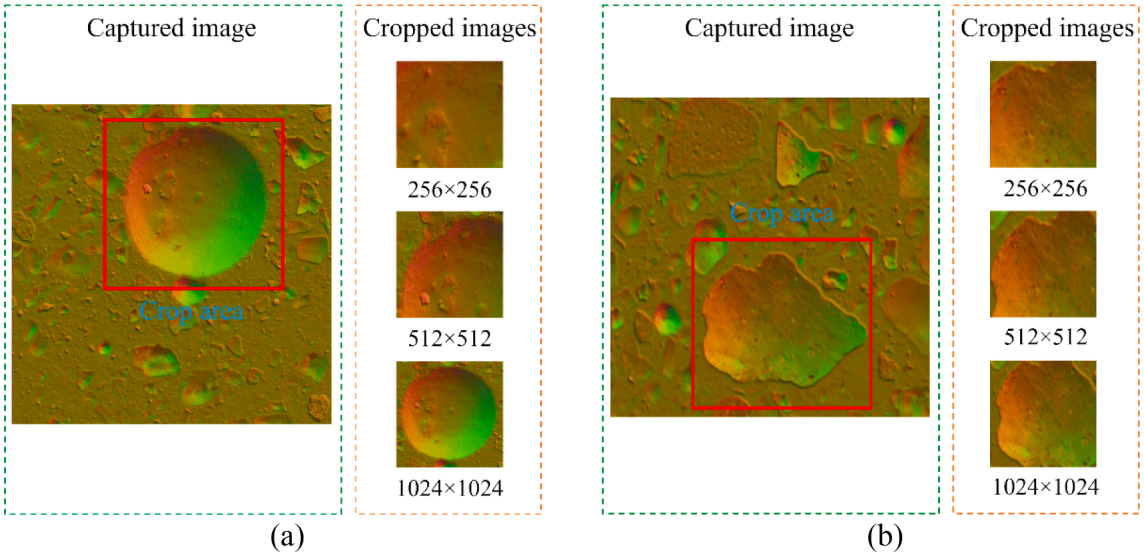

Fig. 12. Random cropping using different cut sizes: (a) Air voids, (b) Aggregates.

were also calculated. As the concrete samples did not come with the measured paste content, 27%, which is recommended as a good estimation in FHWA petrographic manual, was used as the estimated paste content [46]. The air-void parameters were measured using the Procedure B Modified Point-Count Method. As shown in Fig. 11(d) and 11 (e), it can be observed that the best-fit threshold for air content is 0.5 and the best-fit threshold for spacing factor is 0.6. The results indicate that the importance of the segmentation error in different locations is not the same. By observing the segmentation results, it can be found that a lower threshold helps to reduce the FN segmentation error inside of large air voids, while a lower threshold increases the FP segmentation errors which are mainly FP fine air voids. Most of the FP fine air voids are caused by the misidentified fine aggregates. The discretely distributed fine FP air voids have a significant impact on the measurement of air content and spacing factor. It is worth to mention that even though a lower threshold has a significant influence on FP fine air voids, but it barely impacts the segmentation result of TP fine air voids. Consequently, the correct segmentation of fine air-voids and aggregates is highly important for the accurate measurement of air-void system.

As discussed previously, many air-void like noises can be generated in surface normal images using photometric stereo methods. The trained model correctly identified most of the biased regions and only a small percent of the biased regions was incorrectly identified as air voids. Some well-rounded fine aggregates can be false positively segmented as air voids, while the percentage of those FP segmentations is not significant. In addition, most of the voids and cracks in aggregates were well identified as non-air voids. However, some well-rounded voids in aggregates were still incorrectly segmented as air voids. Also, some missing air voids were observed in the segmentation result. Most of the missing air voids were deep air voids. Those air voids generally could be easily identified using naked eyes, but these kinds of air voids presented a significantly different appearance in the surface normal map, which caused interference to the correct air-void image segmentation. One of the reasons that lead to the missing segmentation could be the quantity of training data was not sufficient for the U-Net to learn the pattern of tiny or shallow air voids. Another explanation could be the 256 pixels × 256 pixels sliced training images were not big enough to entirely contain some large-size air voids. Consequently, the U-Net model failed to capture valid features to represent the large size air voids. As shown in Figs. 12, 3 image sizes were selected to make a random cropping from an image with a large air void and an image with a large aggregate. It can be clearly observed that the air void and the aggregate can barely be differentiated from the 256 pixels × 256 pixels cropped images. As the image size becomes larger, the patterns of air voids and aggregates can be identified more easily.

To furtherly quantify the observations of the segmentation results. The accuracy measurements for the testing samples are presented in Table 4. The testing samples were first scanned using the photometric stereo system and the surface normal images were then mapped from the estimated surface normal vectors. All the air-void segmentation results were output by the trained U-Net512 model based on the surface normal images. Equations 2–7 were adopted for calculating the accuracy indices. As shown in Table 4, the average of the MIoU of five testing samples is 0.884, which indicates that the proposed method could detect the air voids in hardened concrete surface with a relative good accuracy. The average FP is 0.009 and the average P is 0.990, which indicate that the proposed method could differentiate the air voids from most of the air-void like noises and only a small portion of air-void like noises was incorrectly identified as air voids. The average FN is 0.115 and is almost 10 times as great as FP. In addition, R is 0.885 which is nearly 0.1 less than P. Both FN and R indicate that the misidentification was the major source of segmentation errors.

Consequently, the segmentation results of the concrete surface images using the U-Net model trained with different cropped image sizes (256 × 256, 512 × 512, and 1024 × 1024) are presented in Table 5. After various experiments, it can be observed that the image size has an

Table 4
Accuracy measurement for testing samples.

Specimen No.	FP	FN	TP	TN	P	R	F_1	$IoU_{air\ void}$	$IoU_{non\ air\text{-}void}$	MIoU
1	0.007	0.121	0.879	0.993	0.992	0.879	0.932	0.873	0.886	0.880
2	0.004	0.027	0.973	0.996	0.996	0.973	0.984	0.969	0.970	0.969
3	0.019	0.113	0.887	0.981	0.979	0.887	0.931	0.871	0.882	0.876
4	0.006	0.159	0.841	0.994	0.993	0.841	0.911	0.836	0.858	0.847
5	0.007	0.154	0.846	0.993	0.991	0.846	0.913	0.840	0.860	0.850
Average	0.009	0.115	0.885	0.991	0.990	0.885	0.934	0.878	0.891	0.884

Table 5
U-Net models trained with different image sizes.

Model	Image Size	Augmentation Strategy	Images for Training	Batch Size	Best Train MIoU	Best Val MIoU
U-Net256	256 × 256	Augmentation 3	18,000	32	0.814	0.706
U-Net512	512 × 512	Augmentation 3	10,187	16	0.794	0.724
U-Net1024	1024 × 1024	Augmentation 3	1,906	4	0.825	0.707

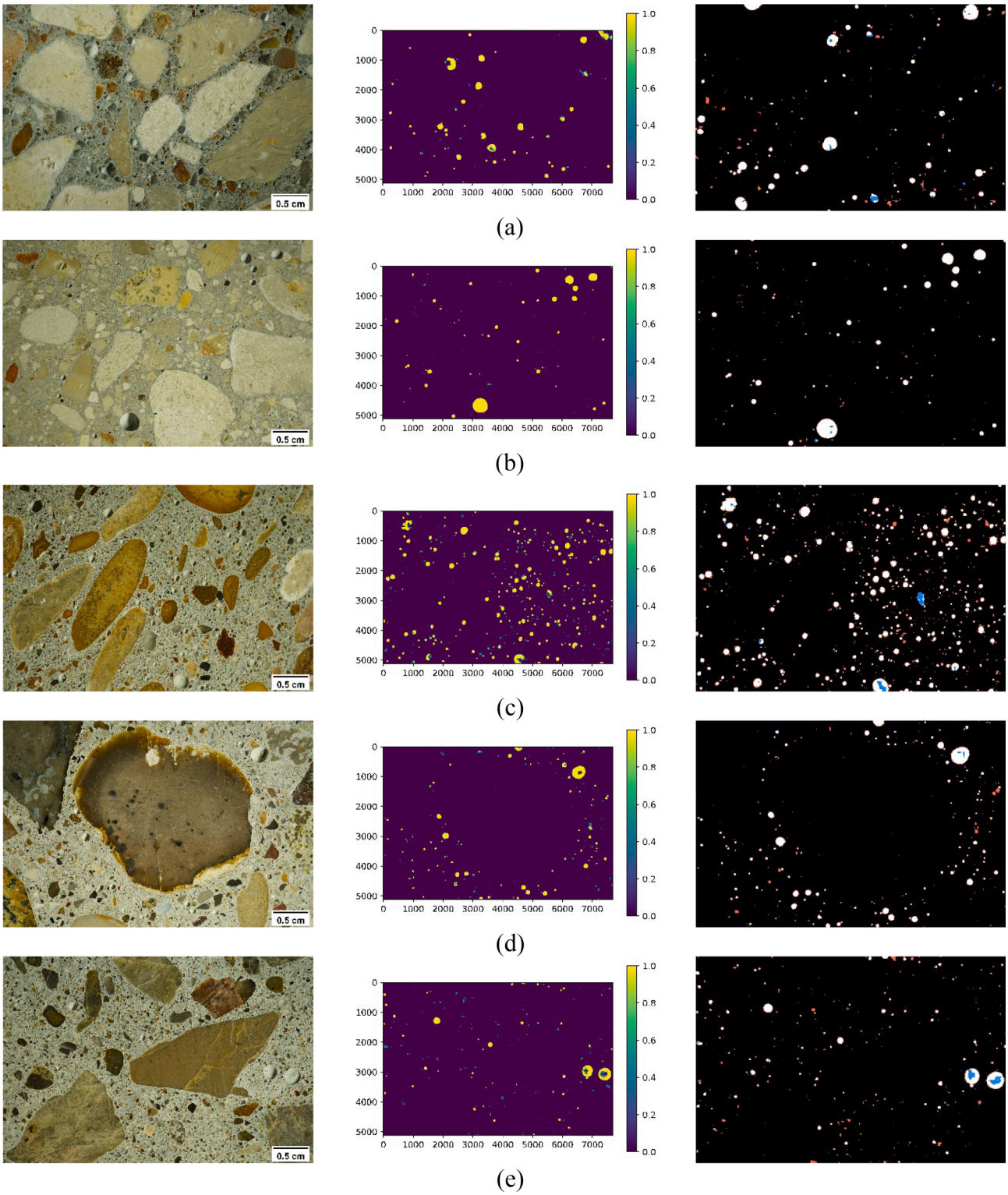

Fig. 13. Air-void segmentation results of test concrete samples using U-Net512: (a) Sample 1, (b) Sample 2, (c) Sample 3, (d) Sample 4, (e) Sample 5. The sequence of each image in the subplot is raw hardened concrete image, raw segmentation result (raw output by the U-Net model), and binary segmentation result.

Table 6
Accuracy measurement for testing samples using U-Net512.

Specimen No.	FP	FN	TP	TN	P	R	F_1	$IoU_{air\ void}$	$IoU_{non\ air\text{-}void}$	MIoU
1	0.008	0.115	0.885	0.992	0.991	0.885	0.935	0.878	0.890	0.884
2	0.003	0.014	0.986	0.997	0.997	0.986	0.991	0.983	0.983	0.983
3	0.019	0.092	0.908	0.981	0.980	0.908	0.943	0.892	0.899	0.895
4	0.006	0.073	0.927	0.994	0.994	0.927	0.959	0.922	0.927	0.924
5	0.006	0.115	0.885	0.994	0.993	0.885	0.936	0.879	0.891	0.885
Average	0.008	0.082	0.918	0.992	0.991	0.918	0.953	0.911	0.918	0.914

Table 7
Air-void Parameters Measured on U-Net512 results and Ground Truth using Point Count Method.

		Air content (%)		Specific surface (mm^{-1})		Spacing factor (mm)	
		Measurement value	Error (%)	Measurement value	Error (%)	Measurement value	Error (%)
1	U-Net512	1.64	6.49	13.152	25.75	0.288	25.32
	Ground truth	1.54		10.459		0.386	
2	U-Net512	1.63	11.64	10.397	1.29	0.367	11.57
	Ground truth	1.46		10.264		0.415	
3	U-Net512	4.08	2.77	17.644	10.16	0.086	11.67
	Ground truth	3.97		16.016		0.098	
4	U-Net512	1.6	11.11	16.194	2.15	0.24	11.89
	Ground truth	1.44		15.854		0.272	
5	U-Net512	1.32	7.32	15.491	7.45	0.304	0.68
	Ground truth	1.23		16.738		0.302	
Average error (%)			7.87		9.36		12.23

impact on the air-void segmentation performance. The U-Net512, which was trained using 512 pixels × 512 pixels images, showed a slightly better performance in generalization than the other two models. Even though a higher train accuracy is observed for U-Net256, the U-Net512 performs better in validation data.

Fig. 13 shows the air-void segmentation results using U-Net512. In the raw segmentation results, the output value of a purple pixel is close to 0 and the output value of a yellow pixel is close to 1. The segmented concrete images with the threshold of 0.1 are presented in the binary segmentation results of Fig. 11. TP pixels are marked in white color, TN pixels are marked in black color, FN pixels are marked in blue color, and FP pixels are marked in red color. Compared with the segmentation results using U-Net256, the large air voids can be identified more completely. Also, some air-void like noises can be better differentiated using U-Net512. The detailed accuracy measurements of U-Net512 are estimated and presented in Table 6. As shown in Table 6, the average MIoU is 0.914 and indicates that U-Net512 model was quite capable of precisely distinguishing the air voids from the non-air-void regions. Compared to the segmentation measurements of U-Net256, the average MIoU of U-Net512 is 0.03 higher than the average MIoU of U-Net256. Additionally, the U-Net512 significantly reduces the number of misidentified air voids. The FN is reduced by 0.03 and the R is increased by 0.03. While the U-Net512 does not show a significant impact on the FP and P. Both FP and P are slightly increased by 0.001. The size of the minimum air void that can be segmented by U-Net512 is around 22 μm (4-pixel length). Table 7 presents the measurement of air-void parameters using U-Net512 results and ground truth. The air-void parameters are measured using the Procedure B Modified Point-Count Method. 27% is assumed as the measured paste content. As shown in Table 7, the average measurement errors of air content, specific surface, and spacing factor for the three hardened concrete specimens are 7.87%, 9.36% and 12.23%, respectively.

Compared with contrast-enhancement based method, the proposed method also achieved a significant improvement in efficiency and comparable accuracy. For a concrete surface image with a resolution of 7953*5304 pixels, the construction of the surface normal image of concrete surface took 10 to 15 s to accomplish, and the segmentation of a concrete surface normal image took 15 to 20 s to accomplish. The segmentation time for air voids in hardened concrete is significantly reduced to around half a minute.

6. Conclusions

This study investigated the potential of combining deep learning and 3D reconstruction techniques to detect air voids in hardened concrete surfaces, extracted and processed from concrete pavement structures. The proposed method can automatically detect the air voids in hardened concrete surfaces without the use of contrast enhancement, which reduces the labor intensiveness and improves the time efficiency. The research has achieved the following: 1) An air-void dataset, which consists of a group of surface normal images and air-void annotations, was developed for DCNN training. 2) The surface normal images of concrete surfaces were captured using a 3D reconstruction method. In the surface normal images, the air voids were automatically highlighted by color variations without the use of contrast-enhancement method. 3) A DCNN model was trained using the developed air-void dataset for air-void detection. 4) A rigorous manual verification was conducted to obtain a reliable reference for evaluating the segmentation accuracy pixel-to-pixel. The research results have shown that:

(1) The 3D reconstruction method can capture depth variations and automatically highlight the air voids in the hardened concrete surface, while some transparent aggregates, dark aggregates, and voids in aggregates are highlighted as well. Most of the air-void like noises can be effectively identified using the proposed method.
(2) The appearance of air voids in surface normal images can vary. Some deep air voids can generate a significantly different appearance in the surface normal images, which could greatly harm the segmentation results. Increase the number of training data that contains deep air voids or refine the illumination angle of the photometric stereo system can help improve the segmentation results of deep air voids.
(3) Shapes and gradient variations of air voids are two important characteristics for air-void segmentation. Inappropriate augmentation methods can change the appearances of air voids in training data set and then cause a poor generalization. The selection of image augmentation methods has a significant impact on the air-void segmentation results.
(4) The over 0.9 average MIoU of the DCNN segmentation indicates that the DCNN has a good potential in extracting the features

from the surface normal images and making predictions at pixel level. The U-Net model can correctly segment most of the air-void regions. The mis-identified air voids are the key error resources. The FP segmentation of some fine aggregates has the most significant impact on the measurement of air-void system.

CRediT authorship contribution statement

Jueqiang Tao: Conceptualization, Methodology, Software, Investigation, Validation, Writing – original draft, Writing – review & editing. **Haitao Gong:** Funding acquisition, Software. **Feng Wang:** Conceptualization, Supervision, Writing – review & editing, Funding acquisition. **Xiaohua Luo:** Writing – review & editing. **Xin Qiu:** Conceptualization, Writing – review & editing. **Jinli Liu:** Writing – review & editing, Investigation.

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.

Acknowledgement

The authors would like to thank the funding supports received from Texas State University, the National Science Foundation [Grant Number #2051192], and the Dwight David Eisenhower Transportation Fellowship Program of the Federal Highway Administration.

References

[1] F. Saucier, R. Pleau, D. Vézina, Precision of the air void characteristics measurement by ASTM C 457: results of an interlaboratory test program, Can J Civil Eng 23 (5) (1996) 1118–1128.
[2] C. Pade, U. Jakobsen, J. Elsen, A new automatic analysis system for analyzing the air void system in hardened concrete, Illinois, International cement microscopy associationMetropolis, 2002, pp. 204–213.
[3] P. Fonseca, G. Scherer, An image analysis procedure to quantify the air void system of mortar and concrete, Mater Struct 48 (2015) 3087–3098.
[4] S. Wolter, F.A.H. Uhre, M.T. Hasholt, V.A. Dahl, F. Anton, Air void analysis of hardened concrete by means of photogrammetry, Constr Build Mater 226 (2019) 953–964.
[5] S. Stock, Recent advances in X-ray microtomography applied to materials, Int Mater Rev 53 (2008) 129–181.
[6] J. Tao, H. Gong, F. Wang, X. Luo, X. Qiu, Y. Huang, Automated Image Segmentation of Air Voids in Hardened Concrete Surface Using Photometric Stereo Method, INT J PAVEMENT ENG, (in press).
[7] Y.u. Song, Z. Huang, C. Shen, H. Shi, D.A. Lange, Deep learning-based automated image segmentation for concrete petrographic analysis, Cement Concrete Res 135 (2020) 106118, https://doi.org/10.1016/j.cemconres.2020.106118.
[8] G. Bianco, A. Gallo, F. Bruno, M. Muzzupappa, A comparative analysis between active and passive techniques for underwater 3D reconstruction of close-range objects, Sensors-Basel 13 (2013) 11007–11031.
[9] M.P. Pound, A.P. French, E.H. Murchie, T.P. Pridmore, Automated recovery of three-dimensional models of plant shoots from multiple color images, Plant Physiol 166 (2014) 1688–1698.
[10] S. Jay, G. Rabatel, X. Hadoux, D. Moura, N. Gorretta, In-field crop row phenotyping from 3D modeling performed using Structure from Motion, Comput Electron AGR 110 (2015) 70–77.
[11] G. Bernotas, L.C. Scorza, M.F. Hansen, I.J. Hales, K.J. Halliday, L.N. Smith, M. L. Smith, A.J. McCormick, A photometric stereo-based 3D imaging system using computer vision and deep learning for tracking plant growth, GIGASCIENCE 8 (2019) giz056.
[12] J.A. Gibbs, M. Pound, A.P. French, D.M. Wells, E. Murchie, T. Pridmore, Plant phenotyping: an active vision cell for three-dimensional plant shoot reconstruction, Plant Physiol 178 (2018) 524–534.
[13] B. Biskup, H. Scharr, U. Schurr, U. Rascher, A stereo imaging system for measuring structural parameters of plant canopies, Plant Cell Environ 30 (2007) 1299–1308.
[14] F.-C. Huang, D.P. Luebke, G. Wetzstein, The light field stereoscope, SIGGRAPH emerging technologies (2015) pp. 24:21.
[15] M. Herrero-Huerta, R. Lindenbergh, W. Gard, Leaf movements of indoor plants monitored by terrestrial LiDAR, Front Plant Sci 9 (2018) 189.
[16] P.J. Besl, N.D. McKay, Method for registration of 3-D shapes, Sensor fusion IV: control paradigms and data structures, International Society for Optics and Photonics (1992) 586–606.
[17] S. Paulus, H. Schumann, H. Kuhlmann, J. Léon, High-precision laser scanning system for capturing 3D plant architecture and analysing growth of cereal plants, Biosyst Eng 121 (2014) 1–11.
[18] R.J. Woodham, Photometric method for determining surface orientation from multiple images, Opt Eng 19 (1980), 191139.
[19] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, B. Schiele, The cityscapes dataset for semantic urban scene understanding, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 3213–3223.
[20] A. Zhang, K.C.P. Wang, Y. Fei, Y. Liu, S. Tao, C. Chen, J.Q. Li, B. Li, Deep Learning–Based Fully Automated Pavement Crack Detection on 3D Asphalt Surfaces with an Improved CrackNet, J COMPUT CIVIL ENG, 32 (2018) 04018041-04018041-04018041-04018014.
[21] W. Liu, Y. Wen, Z. Yu, M. Li, B. Raj, L. Song, Sphereface, Deep hypersphere embedding for face recognition, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 212–220.
[22] J. Wan, D. Wang, S.C.H. Hoi, P. Wu, J. Zhu, Y. Zhang, J. Li, Deep learning for content-based image retrieval: A comprehensive study, in: Proceedings of the 22nd ACM international conference on Multimedia, 2014, pp. 157–166.
[23] R. Girshick, J. Donahue, T. Darrell, J. Malik, Rich feature hierarchies for accurate object detection and semantic segmentation, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2014, pp. 580–587.
[24] J.R. Uijlings, K.E. Van De Sande, T. Gevers, A.W. Smeulders, Selective search for object recognition, Int J Comput Vision 104 (2013) 154–171.
[25] B. Hariharan, P. Arbeláez, R. Girshick, J. Malik, Simultaneous detection and segmentation, European Conference on Computer Vision, Springer (2014) 297–312.
[26] Y. Guo, Y.u. Liu, T. Georgiou, M.S. Lew, A review of semantic segmentation using deep neural networks, Int J Multimed INF R 7 (2) (2018) 87–93.
[27] J. Dai, K. He, J. Sun, Convolutional feature masking for joint object and stuff segmentation, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 3992–4000.
[28] K. He, X. Zhang, S. Ren, J. Sun, Spatial pyramid pooling in deep convolutional networks for visual recognition, IEEE T Pattern Anal 37 (9) (2015) 1904–1916.
[29] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 3431–3440.
[30] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, Adv. Neural Inf. Process. Syst. 25 (2012) 1097–1105.
[31] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv preprint arXiv, 2014.
[32] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 1–9.
[33] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.
[34] M. Mostajabi, P. Yadollahpour, G. Shakhnarovich, Feedforward semantic segmentation with zoom-out features, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 3376–3385.
[35] C. Szegedy, S. Reed, D. Erhan, D. Anguelov, S. Ioffe, Scalable, high-quality object detection, arXiv preprint arXiv:. 2014.
[36] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, A.L. Yuille, Semantic image segmentation with deep convolutional nets and fully connected crfs, arXiv preprint arXiv:. 2014.
[37] S. Zhou, J.-N. Wu, Y. Wu, X. Zhou, Exploiting local structures with the kronecker layer in convolutional networks, Computer Vision and Pattern Recognition (2015).
[38] F. Yu, V. Koltun, Multi-scale context aggregation by dilated convolutions, Computer Vision and Pattern Recognition (2015).
[39] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, A.L. Yuille, Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs, IEEE T PATTERN ANAL 40 (2017) 834–848.
[40] A. Paszke, A. Chaurasia, S. Kim, E. Culurciello, Enet: A deep neural network architecture for real-time semantic segmentation, Computer Vision and Pattern Recognition, 2016.
[41] W. Liu, A. Rabinovich, A.C. Berg, Parsenet: Looking wider to see better, Computer Vision and Pattern Recognition (2015).
[42] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, International Conference on Medical image computing and computer-assisted intervention, Springer, 2015, pp. 234-241.
[43] S.W. Chang, S.W. Liao, KUnet: microscopy image segmentation with deep unet based convolutional networks, 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC), IEEE, 2019, pp. 3561-3566.
[44] S.L.H. Lau, E.K.P. Chong, X.u. Yang, X. Wang, Automated pavement crack segmentation using u-net-based convolutional neural network, IEEE ACCESS 8 (2020) 114892–114899.
[45] ASTM, Standard Test Method for Microscopical Determination of Parameters of the Air-Void System in Hardened Concrete: C 457/C 457M-12, ASTM, 2012.
[46] H.N. Walker, D.S. Lane, P.E. Stutzman, Petrographic methods of examining hardened concrete: A petrographic manual, United States. Federal Highway Administration, Office of Infrastructure Research and Development, 2006.