Automation in Construction 166 (2024) 105633

Contents lists available at ScienceDirect

Automation in Construction

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

3D reconstruction and measurement of concrete spalling using near-field Photometric stereo and YOLOv8☆

Hamish Dow *, Marcus Perry, Sanjeetha Pennada, Rebecca Lunn, Stella Pytharouli

Department of Civil and Environmental Engineering, University of Strathclyde, Glasgow, G1 1XJ, UK

ARTICLE INFO

Keywords:
Automated inspections
Angled illumination
Directional lighting
ALICS
Defect detection

ABSTRACT

Current concrete spalling detection and measurement methods are sparse; despite recent research and commercial offerings using laser scanners, manual measurement is still the industry standard. This paper presents a spalling 3D reconstruction and measurement method. The method uses images illuminated with angled and directional lighting and three neural networks for photometric stereo 3D mesh generation and spalling volume measurement. The proposed method was benchmarked on a laboratory dataset of spalled concrete slabs against a high-resolution laser scanner, yielding an average height error of 0.0 mm and a standard deviation of 1.3 mm. Volume comparisons showed that with manual input, the method achieved a mean absolute percentage error of 22%. Finally, the proposed technique was compared to manual measurements and benchmarked on a spalled concrete structure against a Trimble X12 laser scanner. This research can provide inspectors with increased data interpretability and reduced imaging time for concrete defect mapping.

1. Introduction

Spalling is the deterioration of the concrete surface layer, characterised by chipping, flaking, or large pieces breaking away. This deterioration results in depressions that expose the underlying aggregates [1]. Poor construction methodologies, excessive loading cycles, or environmental changes can cause spalling. Spalling reduces concrete cover and allows corrosive agents to reach reinforcement faster. If left unnoticed and unrepaired, spalling can become severe and lead to exposed reinforcement, further accelerating corrosion and creating more spalling [2].

Identification of spalling has traditionally used the human eye, with findings recorded using photographs or sketches. Measurement of spalling area and depth usually involves basic tools such as a tape measure and is conducted manually, resulting in many errors [3]. Emerging research and commercial developments have shown the potential of laser scanners for assessing concrete structures and locating spalling [4,5]. However, these devices are costly, slow, and produce data that requires pre-processing by a trained individual [6–8].

Concrete defect detection research using images has primarily focused on crack detection, leaving few novel and innovative methods for spalling imaging severity analysis. A recent publication by McAlorum et al. [9] showed that images illuminated with angled and directional lighting can increase the accuracy of crack detection. The directional lighting images utilised by McAlorum et al. are similar to those used in photometric stereo - a method for producing normal and depth maps using a series of images illuminated from different lighting sources [10]. This paper proposes an angled and directional lighting method for spalling 3D reconstruction and volume measurement that uses a hybrid neural network photometric stereo and YOLOv8 object detection network approach. The technique combines the transparency and ease of understanding of images with the comprehensive data provided by 3D meshes. To the authors' knowledge, this is the first system for spalling identification and volume quantification using only standard Red, Green and Blue channel (RGB) images captured from a fixed position.

A review of the relevant literature is first conducted in Section 2. Following this, Sections 3 and 4 describe the hardware and software components of the proposed method, respectively. The testing methodology is then detailed in Section 5. Section 6 evaluates the performance of the proposed method by comparing it to a highly accurate laser scanner in a controlled laboratory environment. Finally, to assess real-world applicability, the method is tested on a challenging spalling sample found on a concrete structure.

2. Background

2.1. Concrete spalling detection

Image-based concrete defect detection falls into two categories: black-box and white-box techniques [11]. White-box techniques are

☆ This work was funded by the Scottish Funding Council, UK and the University of Strathclyde's Advanced Nuclear Research Centre, UK.
* Corresponding author.
E-mail address: hamish.dow@strath.ac.uk (H. Dow).

https://doi.org/10.1016/j.autcon.2024.105633
Received 31 January 2024; Received in revised form 11 July 2024; Accepted 11 July 2024
Available online 24 July 2024

image-processing mathematical operators applied to an image. They are transparent and modifiable. Black-box techniques are neural-network deep learning approaches. They use machine learning models trained on large amounts of data to conduct tasks such as detection, classification and segmentation.

2.1.1. White-box image analysis methods

German et al. [12] proposed an algorithm for pixel-level segmentation of spalling using local entropy-based thresholding. Their algorithm provides details on spalling and exposed reinforcement size and was benchmarked against manual methods. Depth can also be provided, but this is only relative to reinforcement cover and will not yield results if reinforcement is not exposed. Dawood et al. [13] used a series of image processing techniques (thresholding, gaussian 3D filter, and edge detectors) to segment spalling, with a regression analysis on the pixel intensities to estimate depth. Their method provided promising results but required human input to select a thresholding value for segmentation. The analysis of their method did not consider the accuracy of depths. Yao et al. [14] developed a spalling detection method that uses pixel variances and a second-moment operation to conduct pixel-level segmentation of spalling. While effective on smooth concrete surfaces, the algorithm is not robust to noise, and the testing dataset was of limited size.

2.1.2. Black-box image analysis methods

The semantic segmentation network U-net has shown spalling segmentation results with more robust performance in terms of reliance to noise and data variation in comparison to white-box methods [15]. Black-box approaches can also be multi-class, allowing the same network to detect various defects simultaneously (e.g. spalling, cracking, corrosion and efflorescence) [16]. While effective, pixel-level segmentation approaches are limited to a 2D space and can only provide information on the spalling area, width, and size.

Variations of the "You Only Look Once" (YOLO) object detection model have been extensively used in concrete defect detection [17]. The YOLO version 8 (YOLOv8) model proposed by Jocher et al. [18] is very appropriate for fast and accurate concrete defect detection [19]. Rouf et al. [20] demonstrated that YOLOv8 could provide superior crack and spalling detection results in comparison to other competing object detection models (Faster RCNN, MobileNet-SSD et al.). However, object detection models provide little information on spalling severity other than the approximate surface area.

2.1.3. 3D reconstruction methods

Yang et al. [21] developed a method that used Red, Green, Blue and Depth channel (RGB-D) images and Simultaneous Localization and Mapping (SLAM) to position defects in a 3D space. However, their method did not provide metrics on spalling depth and volume. Other studies have utilised point cloud data (PCD) to calculate spalling volume; Zhang et al. [22] did so using PCD obtained from photogrammetry, but it was not an automated solution as manual intervention was required to pre-process the data for noise. There was no ground truth comparison to find the accuracy of their method. Zhang and Xia [23] proposed a method for identifying corner spalls in PCD by fitting a plane to surfaces. This method required human input and accuracy was not analysed. Zhou et al. [24] proposed an automated spalling detection system using laser scanners; ground truth comparisons only considered spalling area. Kong et al. [25] used a smartphone Light Detection and Ranging (LiDAR) scanner to obtain point cloud data of concrete defects. They used rectangular-shaped synthetic spalling formed with concrete blocks as volume ground truths and found their method to have an average error of 5%; however, it was not an automated solution, and the ground truths did not capture the complex shapes and sizes of spalling. Marchisotti and Zappa [26] trialled time of flight (ToF) sensors mounted to a drone to manually record spalling. Despite long data acquisition and processing times, ground truth comparisons with a laser scanner for height error yielded a standard deviation of 2.5 mm.

Beckman et al. [27] proposed an RGB-D camera method that used a Faster RCNN network to localise spalling and fit a plane to calculate its volume using respective heights of the depth channel. While the methodology is robust, the algorithm produced over- and underestimations of volume on various samples; this was not taken into account in their analysis and meant that overestimations in some samples compensated for underestimations in others, skewing performance results. Furthermore, their ground truths were manually found using water and a ruler for volume and depth, respectively, and the analysis did not consider the produced mesh shape. Mondal et al. [28] proposed a similar RGB-D method; spalling volume ground truth comparisons were made, but the ground truths were from the same RGB-D sensor to analyse the method's segmented spalling area.

As noted by [29,30], RGB-D cameras also pose several issues for automated inspections of concrete such as:

- inability to work outdoors as sunlight can overwhelm the infrared (IR) sensor;
- the reflectance of IR light can change with different materials;
- resolution of depth measurement is considerably smaller than RGB image resolution; and,
- less control of hardware, e.g. lenses.

2.2. Photometric stereo

Photometric stereo, first proposed by Woodham et al. in 1980, is an image processing technique used to estimate the surface normals of objects in a scene by observing those objects under lighting from different angles and directions [10]. Further processing of the surface normals allows the estimation of a 3D surface mesh. While the traditional photometric stereo method is a white-box technique, advancements in computer power and camera technology have allowed the development of black-box neural network alternatives. Most recently, Lichy et al. proposed a fast lightweight photometric stereo capable of handling nearby lighting sources [31]. Photometric stereo has been used in research and industrial applications for defect detection of many materials [32–37]. Most concrete surfaces are non-reflective (Lambertian), making them suitable for use with photometric stereo [38]; however, to the author's knowledge, only one study exists for applied photometric stereo concrete defect detection. Tao et al. [39] utilised photometric stereo for air void detection using surface normals. They proposed two methodologies, one using surface normals and another using 3D depth maps. Their method was successful but only considered the area of the voids, not the volume.

2.3. Gaps in literature and contributions of this research

This review has discussed three related topics: (i) 2D spalling detection methods, (ii) 3D spalling detection methods and (iii) lack of utilisation of photometric stereo for concrete defect detection. 3D spalling detection methods can provide more information on defects; however, current approaches have long data acquisition and processing times, require manual noise removal, and must be operated by a trained individual. Furthermore, previous literature proposing novel spalling 3D reconstruction and measurement methods has overlooked comparisons to ground truths. With fast data acquisition/processing using user-friendly and easily interpretable images, recent advancements in photometric stereo can provide a solution to these research gaps. The contributions of this paper are as follows:

- demonstration of photometric stereo for concrete spalling 3D reconstruction;
- use of a YOLOv8 object detection neural network for spalling bounding box location for mesh volume measurement;

Table 1
Machine vision camera variables and respective values.

Camera variable	Value
Exposure time	200,000 μs
Aperture	f4
Gain	9
Working distance	150 mm
Image resolution	2,448 × 2,048 pixels

- quantitative comparison of proposed methodology's heights and volumes against a laboratory dataset; and,
- quantitative analysis of the proposed methodology against manual measurements and a state-of-the-art commercially available laser scanner on a spalled concrete structure.

3. Hardware

The hardware for this study uses a smaller and more lightweight version of the Adaptive Lighting for the Inspection of Concrete Structures (ALICS) device introduced in [9]. The device, dubbed mini-ALICS and shown in Fig. 1, consists of a "low-poly" dome, which houses light emitting diode (LED) strips and blocks ambient light. The LEDs are manually placed at angles of 50 and 30 degrees incident to the surface at a proximity of 150 mm.

At the centre of this dome is a FLIR Chameleon 1/2 inch sensor machine vision camera paired with a FUJI DF6HA-1S 6 mm lens. Table 1 shows the consistent camera settings used for image acquisition throughout this research.

Mini-ALICS is a contact inspection device. To capture directional lighting images with no ambient lighting from the surrounding environment, the device is placed on a concrete surface. In this work, the system is deployed handheld. However, the lightweight dome design allows for potential remote contact deployment. This could be achieved with robotically deployable contact inspection systems, such as unmanned aerial vehicles (UAVs) (e.g. [40,41]) or wall climbers/crawlers (e.g. [42–44]).

The camera and LEDs are actuated by an onboard Raspberry Pi 4B, which is remotely controlled by a separate master computer. A single 5 V, 2.5 A power supply powers the entire device.

4. Software

The flowchart in Fig. 2 shows the inputs (directional lighting images of a spalled area) and outputs (3D mesh with spalling located) of the proposed method. A detailed description is given in Sections 4.1–4.4.

4.1. NFPS

This research employs the fast near-field photometric stereo (NFPS) method proposed by Lichy et al. [31]. The model involves two neural networks — one for predicting surface normals and a second for predicting depth when given these surface normals.

A complete, detailed description of the NFPS method can be found in the original authors' publication and supplementary material. However, to summarise:

- The surface normals estimation network:
  - Uses reduced-resolution input images illuminated with angled and directional lighting.
  - Utilises "per-pixel lighting" maps to generate a feature vector at a quarter of the input resolution.
  - Performs max pooling on the features of all input images, resulting in a combined feature.
  - Employs a decoder neural network to produce a normal map from the combined feature.

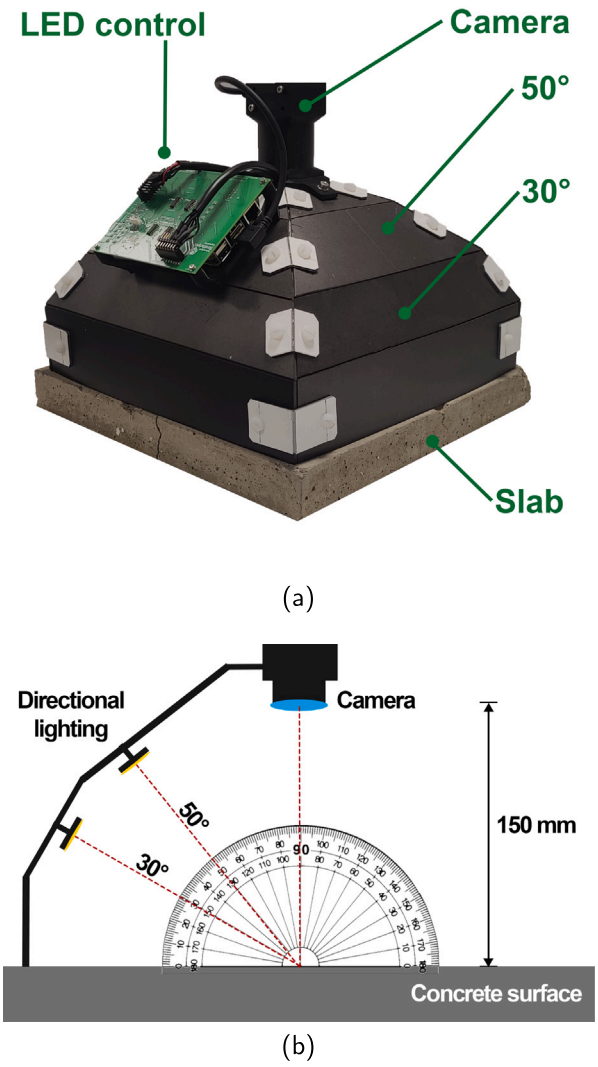

Fig. 1. (a) Mini-ALICS angled and directional lighting image acquisition hardware. Annotations show LED face angles, camera, measured slab and LED acquisition hardware. (b) cross-section of one side of Mini-ALICS, showcasing the projected lighting angles and camera working distance. The actual device has LEDs on all four sides.

- The depth estimation network:
  - Adopts an encoder–decoder network with ResNet architecture.
  - Uses normals generated from the surface normals estimation network to produce a depth map.

The process of the surface normal network through to the depth estimation network is iterated multiple times; with each iteration, the image resolution (width and height) is doubled. After the first iteration, the normal and depth maps from the previous iterations are also used as inputs for their respective networks.

The spalling 3D reconstruction method adopts the same NFPS approach as Lichy et al. [31] using their pre-trained weights; however, to accommodate for the higher-resolution images in this research, the model's initial image resolution for the recursion is increased to 256 × 256. Despite having known lighting positions, auto-calibration is employed to calculate the light location and account for any potential errors introduced from manual dome construction and LED placement (translations in X, Y, Z and rotations in pitch, roll, yaw). While photometric stereo typically assumes point light sources, the

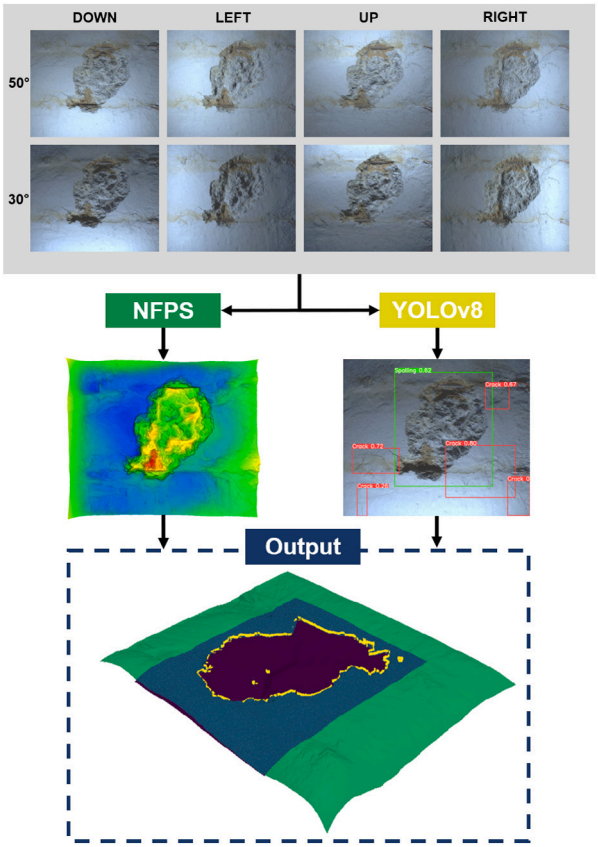

Fig. 2. Flowchart of proposed near-field photometric stereo (NFPS) and YOLOv8 spalling measurement method.

light produced by the LED strips was found to be satisfactory for this specific application.

4.2. YOLOv8

The proposed method uses a standard, unmodified YOLOv8 object detection neural network for region-based identification of concrete spalling. A diagram of the model architecture and open source code is provided by Jocher et al. [18]. The model was trained on an open-source dataset of over 1,900 manually labelled diffused lighting images with six different concrete defect classes (including spalling) [45]. Pre-trained weights from a COCO dataset model were used, and the training was completed with 50 epochs.

Each individually illuminated image is passed to the YOLOv8 network to locate and classify any present defects in the form of bounding boxes. Fig. 15 in the results section of this paper shows four of the eight captured images. Fig. 2 shows all eight captured images and the YOLOv8 output of one.

4.3. Defect localisation

The bounding box coordinates (produced by YOLOv8) of all spalling regions in every angled and directional lighting image are used to localise the defect. An algorithm is used to identify clusters of overlapping bounding boxes across the eight directionally-lit images and remove likely false positives.

If bounding boxes overlap, they are put into a sub-category of bounding boxes. The intersection over union (IoU) of each bounding box in the sub-category is checked using Eq. (1); the box is disregarded if the IoU with other boxes is less than 20%. Following this, the extreme coordinates (top-left, bottom-left, top-right, and bottom-right) of all remaining boxes in the sub-group are calculated and used as the final bounding box. If there is only one bounding box from all input images, this is used as the extreme coordinates. If all image inputs contain multiple bounding boxes but none overlap, they are all disregarded.

4.4. Bounding box defect volume measurement

Similar to the technique used by Beckman et al. [27], the 2D bounding box produced by YOLOv8 surrounding the spalling is used to identify the X, Y, and Z coordinates that outline the perimeter of the spalling area within the mesh. Nearest interpolation is used to estimate the height values within the region of interest that show what the slab would look like if no spalling was present. The non-defective mesh is used to "cap" the spalling and create a mesh with a hollow volume. The volume measurement of the resulting mesh is found using vtk, a mesh measurement package in Python.

5. Methodology

5.1. Laboratory dataset formation

Twenty-five concrete slabs were cast, each measuring 300 mm × 300 mm × 40 mm. To benchmark the proposed method, spalling was intentionally created on twenty-one of these slabs using the following methods:

- manual creation using hammer and chisel (Fig. 3a);
- accelerated corrosion of reinforcing bars (Fig. 3b);
- artificially created efflorescence (Fig. 3c); and,
- removal of 3D-printed moulds inserted during casting (Fig. 3d).

Various spalling depths and shapes were created to ensure variation in the dataset. While all spalling occurred within the central region of the slab for consistency, many were offset and/or displayed asymmetrical distribution across the surface. Fig. 4 illustrates the range and frequency of occurrence of spalling depths throughout the dataset. Spalling depths ranged as deep as 30.3 mm, resulting in a dataset with minimum and maximum volumes of 2,500 mm^3 and 47,500 mm^3, respectively. Several concrete surface finishes and noise-inducing elements (e.g. corrosion staining and efflorescence) were also utilised to ensure the laboratory dataset closely resembled the condition of actual concrete structures found in the built environment.

5.2. 3D mesh acquisition

Using the mini-ALICS hardware, eight images of each slab were captured. Each image was illuminated with lighting individually projected from the up (U), down (D), right (R) and left (L) directions of the image scene at angles of 50 and 30 degrees incident to the surface. The image acquisition time takes 4 s per sample, and the device was placed on the surface to remove ambient lighting from the indoor environment.

For each slab, the captured images were processed by the NFPS model to generate a mesh of the image scene as per the description in Section 4.1.

5.3. Ground truth mesh acquisition and comparison

A stepper motor-driven MICRO-EPSILON scanCONTROL 2700 laser scanner was used to scan the concrete slabs and create a ground truth. The scanner was operated at a rate of 25 profiles per second, capturing 1,600 profiles per scan. The stepper motor speed was set to 0.4 mm/s over a 250 mm distance, and actuation was synchronised with the laser scanner. Due to the scanner measurement range, each slab was scanned in two halves, and the produced point clouds were stitched together, resulting in one point cloud per slab. Finally, a Poisson Surface Reconstruction with an Octree depth of 9 generated a ground truth mesh with units in mm. The accuracy of the scanner was validated by measuring

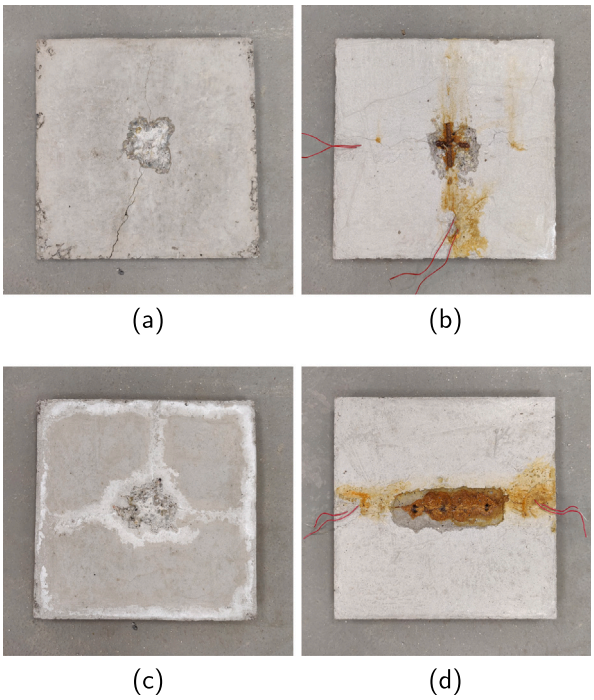

Fig. 3. Spalling dataset variations. (a) chisel carving, (b) reinforcement corrosion, (c) efflorescence and (d) 3D-printed insert removal.

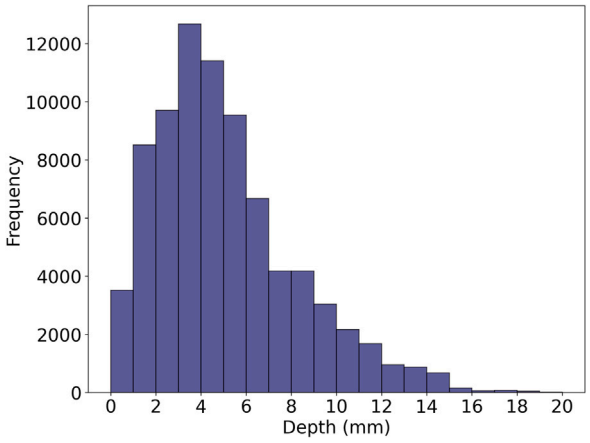

Fig. 4. Histogram of spalling depths throughout the entire dataset. Frequency refers to the number of times a particular depth value (shown on the *x*-axis) appears in the dataset.

the height difference between two machined steel blocks measuring 12.7000 mm (1/2 inch) and 15.8750 mm (5/8 inch) tall, respectively; the manufacturer quotes the accuracy of these blocks as 0.00762 mm (0.0003 inch). Over the 3.175 mm step, the scanner yielded an average distance of 3.331 mm and a standard deviation of 0.122 mm.

Using Cloud Compare, a mesh analysis tool, each mesh produced by the proposed method was aligned to its respective ground truth. The software's "compute mesh distance" tool produced a scalar field array showing the distance from the mesh points to the ground truth.

5.4. Spalling volume ground truth and comparison

The ground truth for the spalling volume of each ground truth 3D mesh slab dataset sample (as detailed in Section 5.3) was determined using Cloud Compare by fitting a plane to the ground truth mesh slab surface and converting it to a point cloud. This point cloud was then cropped to a rectangular shape surrounding the spalling. After this, Cloud Compare's "2.5 Volume" tool was used to compare the "volume removed" when viewing the point cloud change between the plane and the mesh. This value was recorded and repeated for the algorithm output and the respective ground truth of every dataset slab.

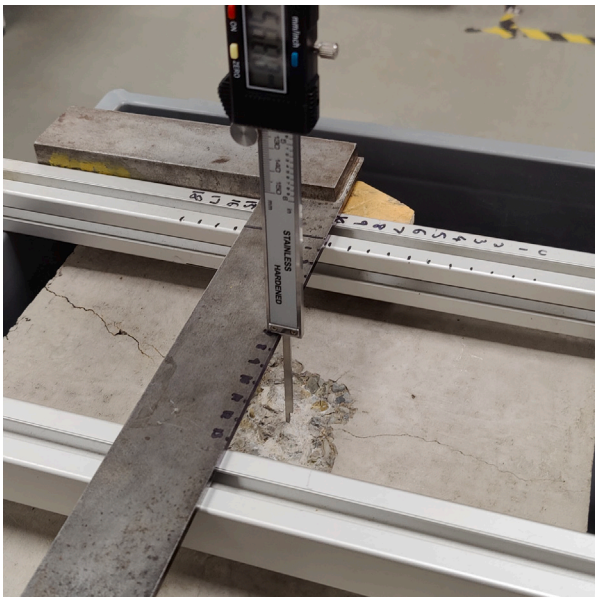

Fig. 5. Probing apparatus used to obtain a manual point cloud.

5.5. Manual spalling depths

Manual measurements of spalling depths on one dataset sample were obtained by probing the spalling on a 90 × 110 mm grid at 10 mm intervals using a height probe. The resulting 99 data points were converted to a point cloud and compared to the respective mesh using the method outlined in Section 5.3. Fig. 5 illustrates the apparatus used to probe the spalling area.

5.6. Spalling bounding box ground truth definition and comparison

Ground truths of spalling bounding boxes were obtained by manually outlining the defect area in data annotation software. Each ground truth, BB_{GT}, was compared to the respective predicted value, BB_P, to calculate intersection over union (IoU) using Eq. (1).

$$IoU = \frac{Area\ of\ Overlap}{Area\ of\ Union} = \frac{Area(BB_{GT} \cap BB_P)}{Area(BB_{GT} \cup BB_P)} \quad (1)$$

5.7. Field trial mesh and comparison mesh acquisition

A field trial of the proposed method was conducted at a concrete structure near Glasgow, UK. A spalled region of the concrete surface was manually identified by eye. The mini-ALICS device was then placed over the spalled region, and images were captured with each individual lighting condition. As shown in Fig. 6, a Trimble X12 laser scanner was used to record a point cloud of the spalled area. All other data processing, analysis and recording were conducted as per laboratory conditions. The data acquisition time of the Trimble X12 was approximately 7 min per sample.

Fig. 7 shows a mesh error map output of the Trimble X12 laser scanner when compared to the MICRO-EPSILON scanCONTROL 2700 laser scanner on a laboratory dataset sample. The meshes of the two scanners are very similar, with most measurements differing by ±0.5 mm.

Fig. 6. Trimble X12 laser scanner data acquisition on field trial sample. The red circle highlights the spalling region.

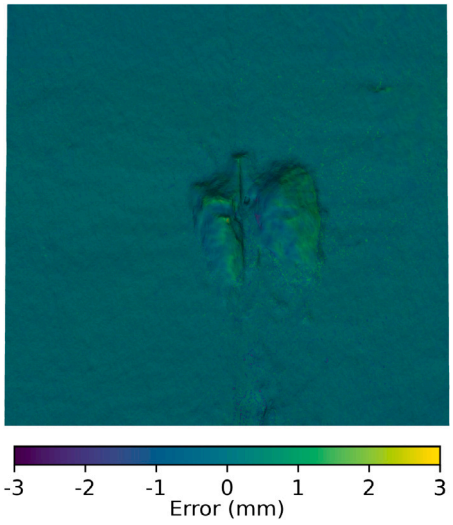

Fig. 7. Error map of Trimble X12 laser scanner mesh compared to MICRO-EPSILON scanCONTROL 2700 laser scanner on a laboratory dataset sample.

6. Results

6.1. Qualitative analysis

Fig. 8 shows one input image and the corresponding mesh output (generated with eight images) for the NFPS model on one sample concrete slab. Qualitatively, the mesh closely resembles the actual depths of the slab, with a flat surface followed by a void in the centre of the scene. However, notable errors occur towards the edge of the mesh where the calculated depth is considerably lower and curved than the actual flat surface. The thicker cracks on the slab's surface are visible on the mesh, but their depths are inaccurate.

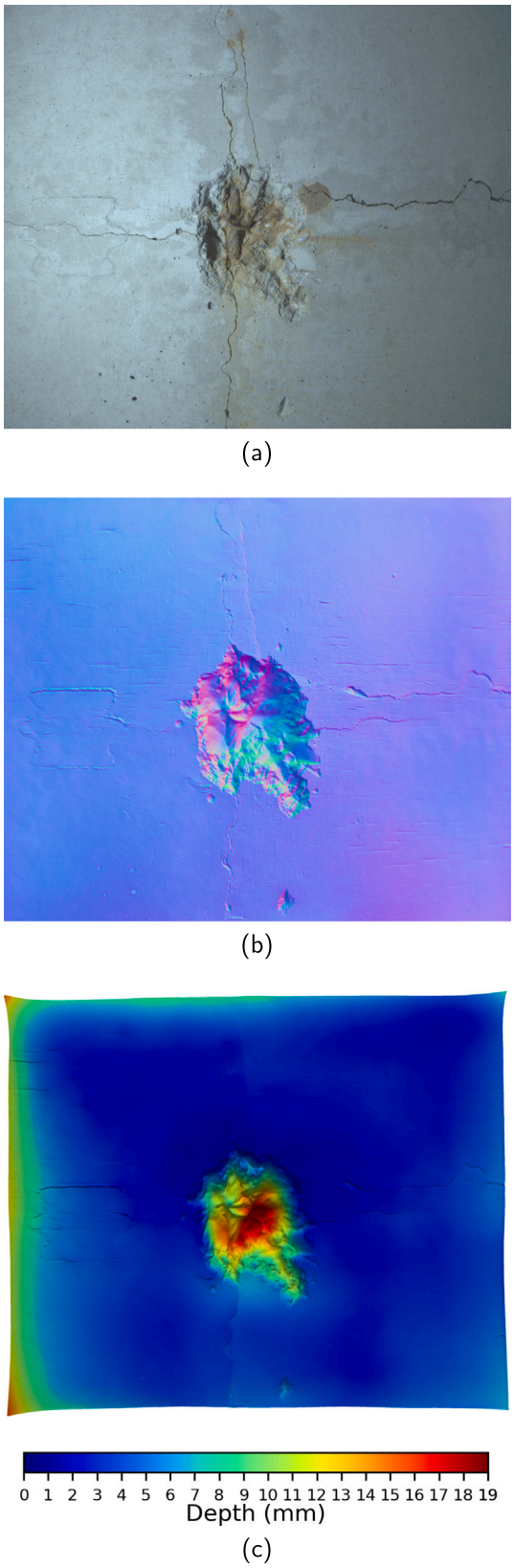

Fig. 8. Spalling dataset sample. (a) captured image with 50-degree angle incident lighting projected from the left direction. (b) NFPS surface normals output. (c) NFPS mesh output; colour map shows elevation change.

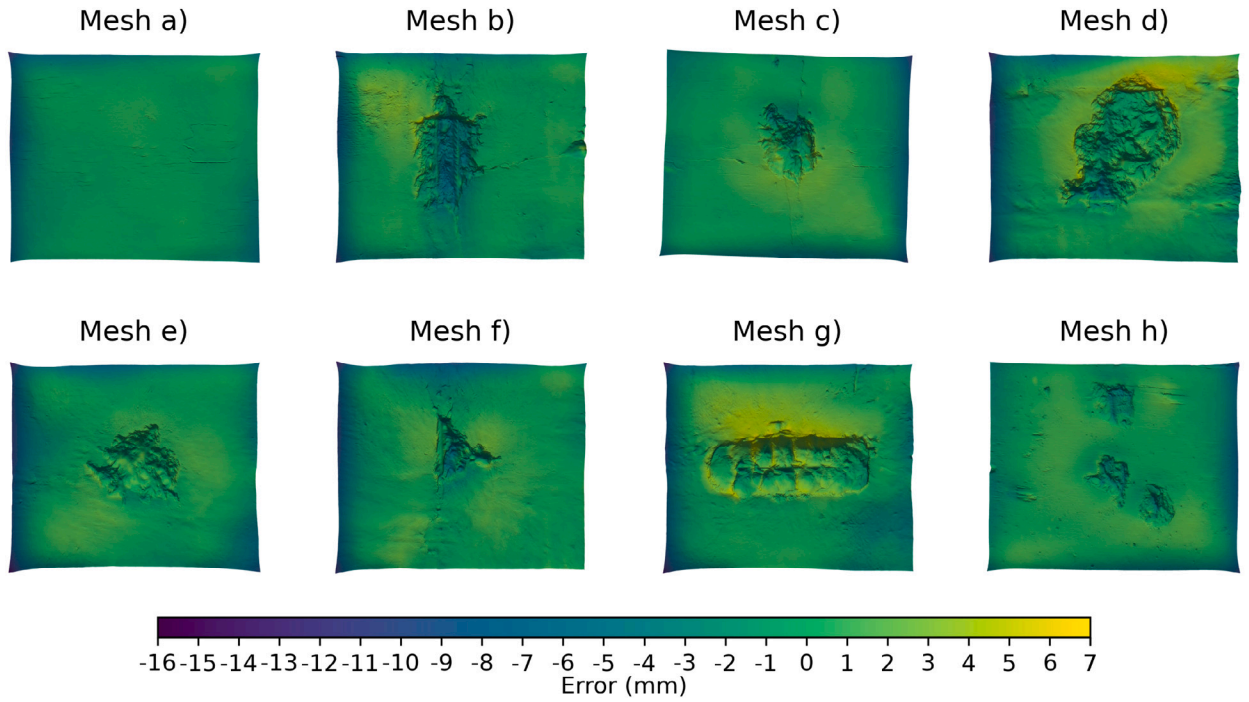

Fig. 9. Height error map of eight dataset samples when compared to laser scanner ground truth. The colour scale is consistent across all samples and units are mm.

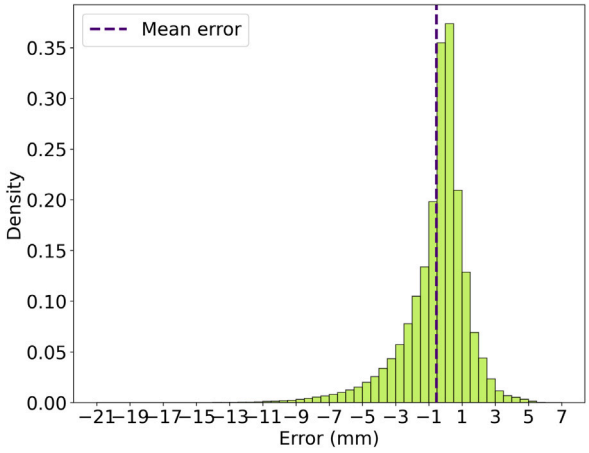

Fig. 10. Histogram of height errors of all dataset samples when compared to their ground truths. Density on the *y*-axis represents a normalised measure of how often an error value occurs.

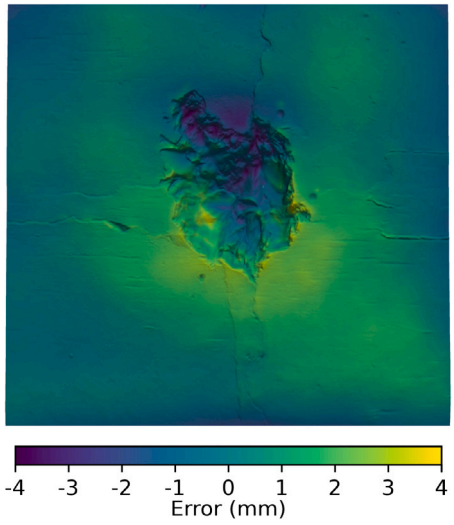

Fig. 11. Error map of one square cropped dataset sample when compared to laser scanner ground truth. Error units are mm.

Fig. 9 presents the height errors of the NFPS method on eight of the twenty-five different dataset samples. Negative values indicate overestimation of depth, and positive values indicate underestimation of depth. In line with earlier analysis, the greatest errors occur at the slab's edges. Notably, the un-spalled slab in Fig. 9 (mesh a) shows that this error arises even in the absence of spalling. On average, it took 40 s per slab dataset sample to generate the surface normals and 3D mesh.

6.2. Quantitative analysis of dataset

The height errors of the NFPS method's 3D meshes, compared to the ground truths for the entire spalling dataset, are displayed as a histogram in Fig. 10. Negative error indicates the NFPS mesh height was below the ground truth (i.e. over estimation of depth); positive error indicates the NFPS mesh height was above the ground truth (i.e. underestimation of depth). The histogram is right-skewed, showing a greater number of height errors in the negative range, indicating a majority overestimation of depths.

Fig. 11 shows the errors of one mesh (Mesh c in Fig. 9) when cropped to a 170 mm × 170 mm square. Cropping has reduced the height error range as the areas towards the edge of the mesh typically have the greatest errors.

Fig. 12 shows a histogram of the height errors of the NFPS method across the entire dataset when cropped to this square resolution. In line with previous analysis, this cropping has reduced the error range and resulted in a more balanced histogram. However, the most significant height errors are still found in the negative range. The histogram can be fitted to a double Weibull distribution, indicating a sharp decline in error with a broad error range. This may be attributed to the randomness of the concrete surface generating noise. The average error of 0.0 mm

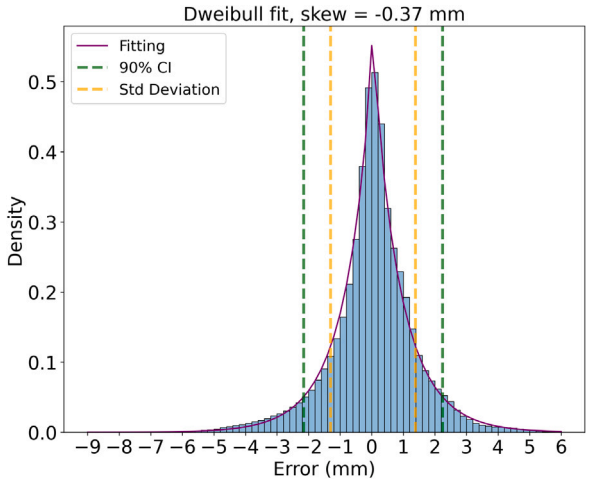

Fig. 12. Histogram of height errors of square cropped dataset. The fitting is a double Weibull distribution. Density on the *y*-axis represents a normalised measure of how often an error value occurs.

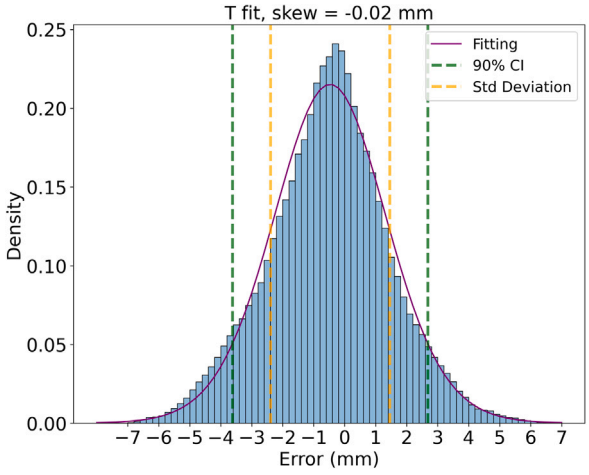

Fig. 13. Histogram of height errors when only considering the spalled regions. The fitting is a T distribution. Density on the *y*-axis represents a normalised measure of how often an error value occurs.

and the standard deviation of 1.3 mm show that the reconstructed mesh heights are consistently accurate on average, with relatively low variability around the mean. The method's standard deviation of 1.3 mm is considerably smaller than the standard deviation of 2.5 mm of the competing ToF approach by Marchisotti and Zappa [26] obtained on a dataset with maximum spalling depth of 31.8 mm. The upper and lower 90% confidence intervals show that height errors commonly fall between approximately −2.1 mm and 2.1 mm.

Fig. 13 illustrates the height errors of the NFPS method throughout the entire dataset when focusing on the data points from the spalled regions of the meshes and excluding data points from the unspalled regions. The NFPS method demonstrated an average error of −0.47 mm with a standard deviation of 1.92 mm. Again, this is notably less than the 2.5 mm standard deviation observed in the ToF approach by Marchisotti and Zappa [26].

The height differences of the NFPS method on a single dataset sample, when compared to manual measurements, are shown in Fig. 14. The manual measurements fall close to those of the NFPS method, with a −3.53 mm and +1.70 mm difference range; however, the introduction of new errors from manual measuring (i.e. human error), could be contributing to this variation.

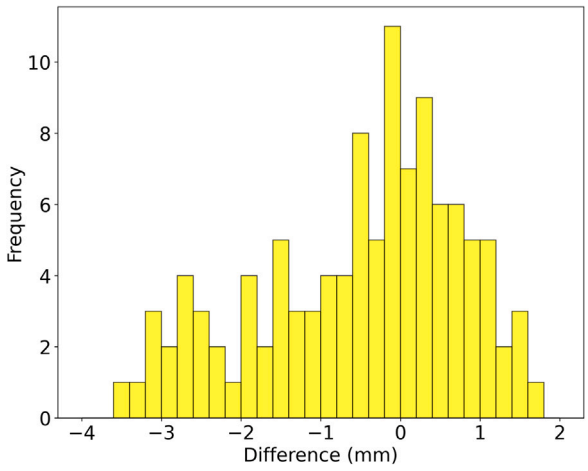

Fig. 14. Histogram of height difference for one data sample when compared to manual measurements. Positive difference indicates the NFPS spalling depth was smaller than the manual measurement depth. Negative difference indicates the NFPS spalling depth was greater than the manual measurement depth. Frequency is the number of times a value (shown on the x-axis) appears in the comparison results.

6.3. Analysis of bounding box measured defect volume

Fig. 15 shows the bounding box output of the YOLOv8 object detection model under four of the eight different illumination methods for one slab. While the model confidence and bounding box locations can change depending on the lighting direction and angle, all spalling areas have correctly been located in all eight images of this slab. The analysed mesh with non-defective plane fittings for this sample is shown in Fig. 16.

Fig. 17 compares the measured volume of the photometric stereo method (when using YOLOv8 defect bounding boxes to outline the defect) to the ground truth volume as calculated from the laser scanner point cloud. The spalling measurements for most of the slabs fall close to their ground truth value; however, for some slabs, the spalling volume has been vastly underestimated, resulting in a mean average percentage error (MAPE) of 41%. The data points highlighted as diamonds in Fig. 17, are from the 3D-printed cast removal spalling. Poor volume measurement on these highlighted samples can be attributed to the YOLOv8 object detection network failing to locate the spalling region, likely due to its few exposed aggregates (which are usually a common indicator of spalling) and irregular shape (e.g. Fig. 3d). The IoU results in Fig. 17 reinforce this analysis and show that as IoU increases, the error between the measured value and ground truth values reduces. Fig. 18 shows a plot of measured volume compared to ground truth volume using manually annotated bounding boxes of spalling location. In this figure, the calculated volumes closely align with their respective ground truth values, reducing MAPE to 22%. This shows that the YOLOv8 model has limitations and should be improved with further training data in order to gain more accurate volume measurements.

The proposed approach offers the benefit of being robust to height variations on the concrete surface that are not from spalling. Alternative methods, for example, thresholding the average height of the mesh, would fail when elevation variations from other defects (e.g. biofouling) are present. Fig. 19 displays one dataset sample featuring multiple pieces of moss and soil placed in the image scene. Despite the sample's noise, the YOLOv8 object detection network has accurately located the spalling region. This accuracy allows effective segmentation of spalling using the proposed method as shown in Fig. 20a. Fig. 20b shows the result of a technique that creates a flat plane at the average height of the mesh. Compared to the YOLOv8 approach, the average height algorithm has resulted in poor segmentation and inaccurate volume measurement, as the moss heights have skewed the average elevation.

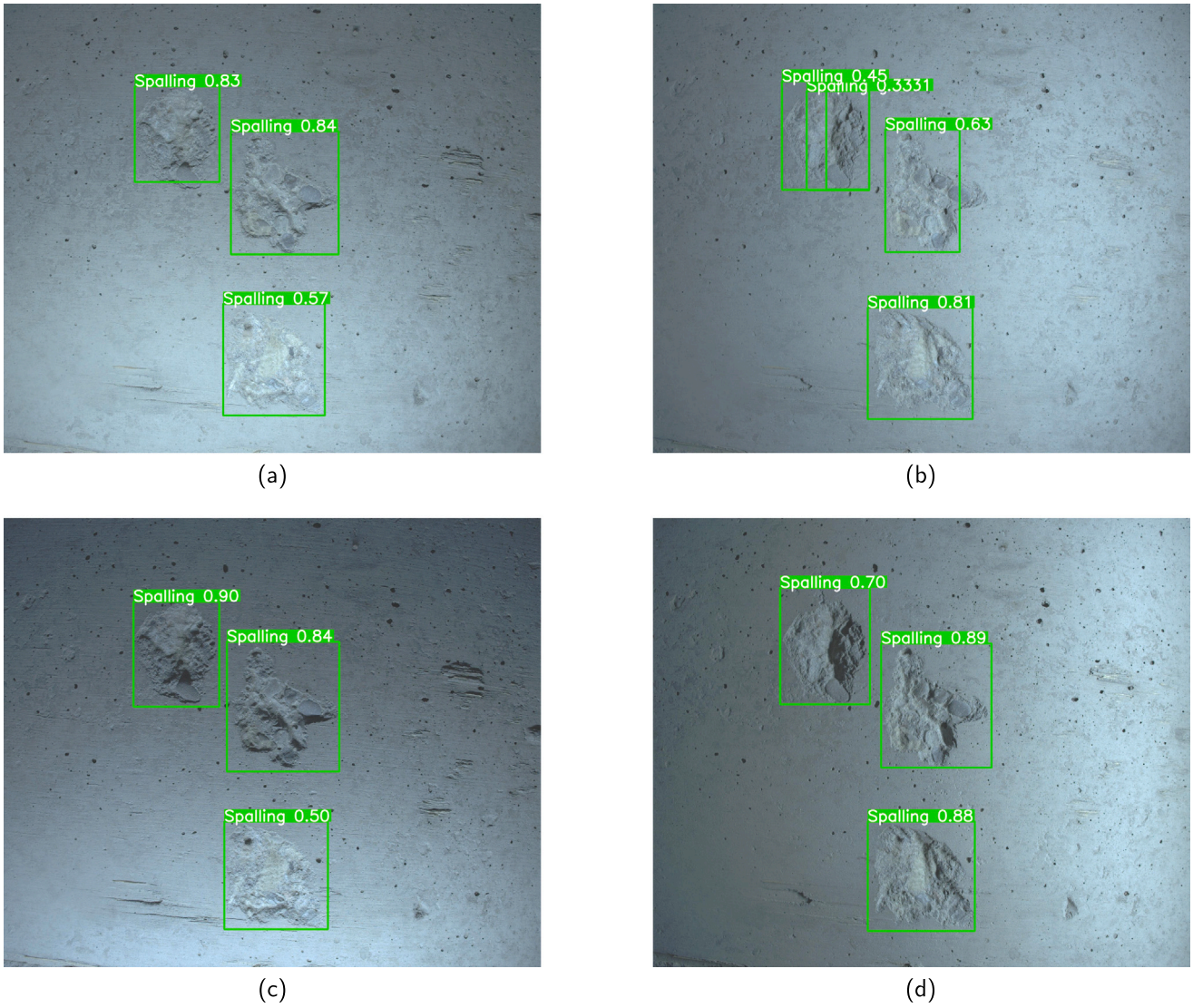

Fig. 15. YOLOv8 output of one spalling dataset sample under 30 and 50 degrees incident lighting. The text shows the respective class and detection confidence of each bounding box.

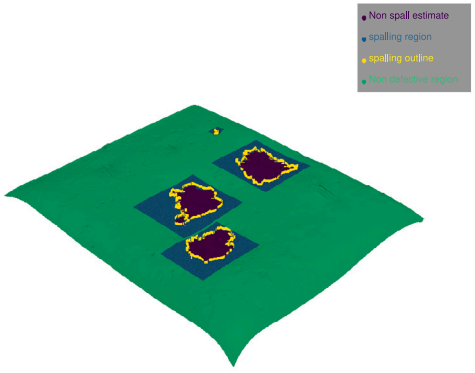

Fig. 16. Detected spalling on a dataset sample.

6.4. Field trial results

Fig. 21 showcases the NFPS mesh and its errors compared to the Trimble X12 on a spalling sample found on an actual concrete structure. The results are close to those of the laboratory tests. The depths of the upper section of the spalling are underestimated by up to 5.33 mm.

The YOLOv8 model struggled on the field trial spalling sample, with no bounding boxes labelled as spalling. Fig. 22 shows the output of one of the angled and directional lighting images, where the spalled area was instead identified as exposed rebar with 0.28 confidence. Manual intervention was used here, and the bounding box from this image was used for the spalling volume analysis.

The measured volume of the field trial spalling sample was 30,373 mm^3 compared to the Trimble X12 volume of 37,137 mm^3. The spalling volume underestimation is likely due to depth underestimations in the upper portion of the spalled region mesh. The proposed method collected all data in 4 s and required an additional 40 s for processing. The Trimble X12 laser scanner took 7 min to acquire the data, and additional time was required to segment and process the mesh.

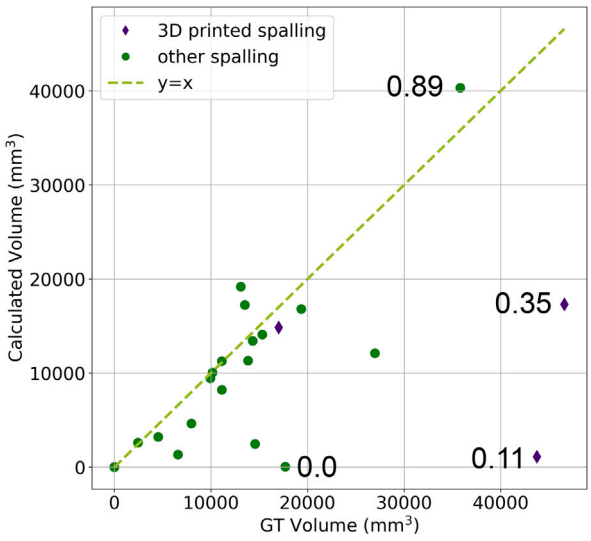

Fig. 17. Comparison of measured volume to ground truth volume when using YOLOv8 bounding boxes to outline the defect location. The datapoint label numbers show the intersection over union (IoU) for the respective dataset samples.

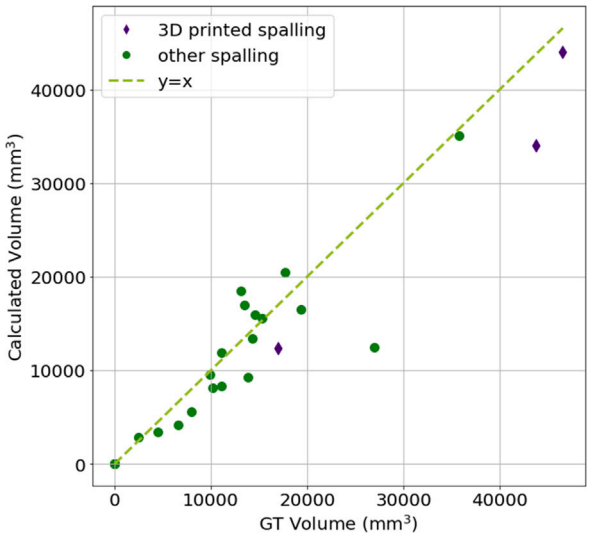

Fig. 18. Comparison of measured volume to ground truth volume when using manually annotated bounding boxes to outline the defect.

7. Conclusion

This paper presented a method for 3D reconstruction and measurement of concrete spalling. The method uses a small form-factor adaptive lighting device (mini-ALICS) to capture multiple images of a concrete surface individually illuminated from various angles and directions. Images are input to a black-box fast lightweight near-field photometric stereo and a multi-class YOLOv8 object detection neural network. The resulting 3D mesh (from photometric stereo) and spalling location coordinates (from YOLOv8) are used to estimate what the concrete surface would look like with no spalling; this acts as a datum for volume measurement.

The method was benchmarked with a more thorough comparison than any previous literature, comparing outputs to a high-resolution laboratory laser scanner on twenty-five defective concrete slabs. Results showed that the proposed method consistently delivered accurate meshes, with an average height error of 0.0 mm and a standard deviation of 1.3 mm. When considering only the spalled region of the mesh,

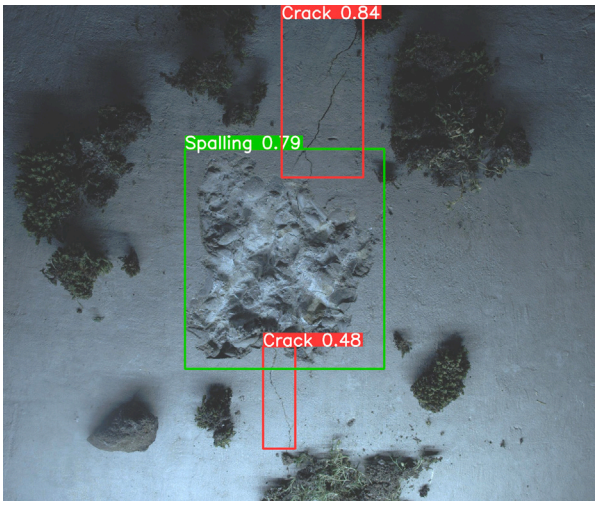

Fig. 19. 30 degree down (D) lighting YOLOv8 bounding box result. The associated text of each bounding box shows the detected class and associated detection confidence.

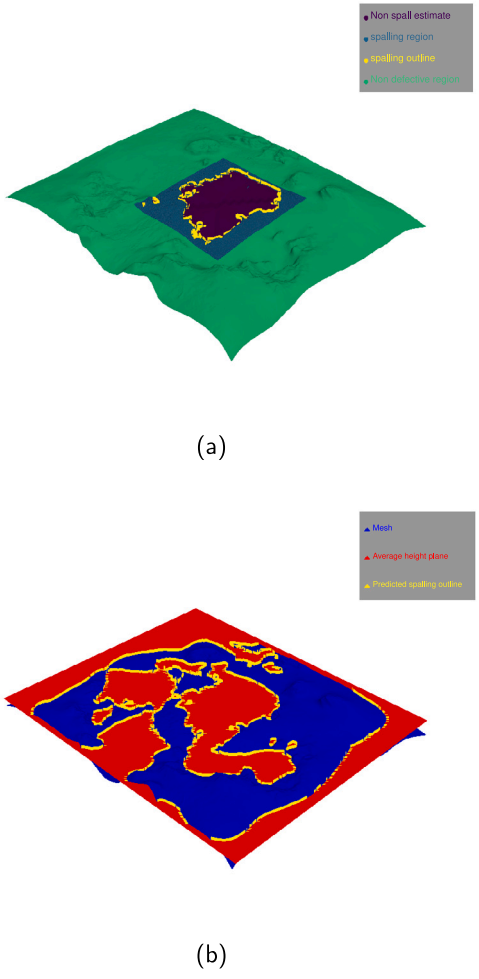

Fig. 20. Detected spalling using (a) the proposed YOLOv8 bounding box approach and (b) alternative average height method.

the average height error and standard deviation changed to −0.47 mm and 1.92 mm, respectively. A comparison with manual measurements yielded similar low-error results. Spalling volume measurements were

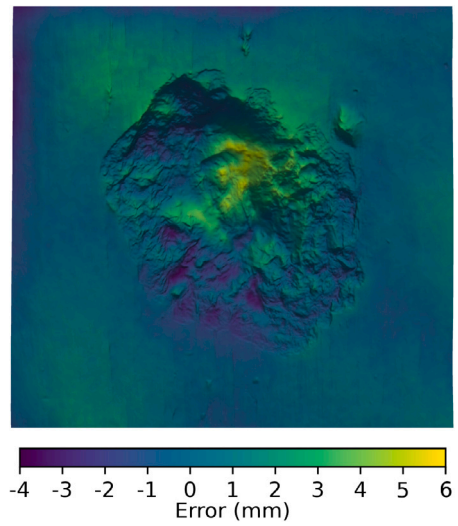

Fig. 21. Error map of NFPS mesh output (square cropped) of a field trial sample when compared to Trimble X12 laser scanner mesh. Error units are mm.

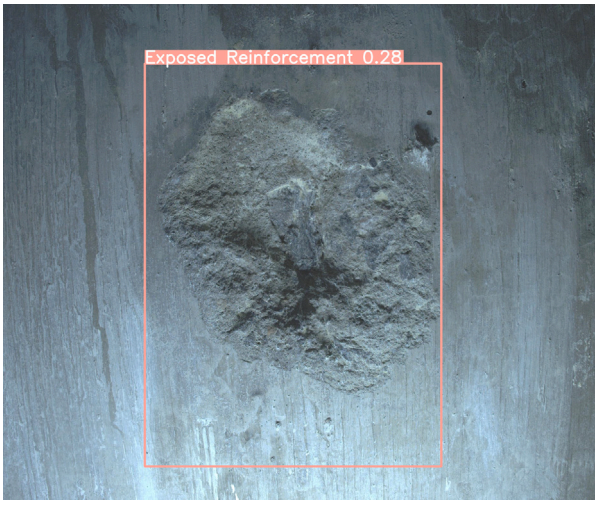

Fig. 22. YOLOv8 detection output of field trial spalling sample with low confidence exposed rebar detection.

also very precise, with flaws resulting from failed bounding box detection by the YOLOv8 network, but this can be attributed to the synthetic nature of some of the slabs. With manual bounding box inputs, the calculated volumes had a mean average percentage error of 22% across the entire dataset. The proposed method was further validated using a comparison with manual measurements and an additional benchmarking on a spalled concrete structure against a Trimble X12 laser scanner. Results from this test closely matched the laboratory findings.

Compared to laser scanners, the proposed approach to spalling 3D reconstruction and measurement allows fast data acquisition (4 s) and transparent, user-friendly inputs that can be easily reviewed by humans (standard image files). It is also fully automatic and requires no human intervention for mesh repair. The method is also cost-effective as it only requires one camera and several low-cost LEDs. Implementing this method in industrial applications can aid the monitoring of concrete structures, ensuring they are safe for use.

Future studies should directly compare the proposed spalling 3D reconstruction method to other approaches, such as depth cameras and point cloud photogrammetry. For a more accurate spalling location using YOLOv8, a larger training set could be used. The normal and depth maps produced from photometric stereo could be utilised as an input layer to the YOLOv8 model. This study focused on comparing spalling within the central region of the images. However, future work should investigate the method's performance with spalling at the edge of images and concrete column corner spalling. Future investigations should also examine how the method performs on reflective (non-Lambertian) polished concrete surfaces and consider errors in the X and Y directions. Comparisons of the method's volume analysis results when using other object detection neural networks (e.g. Faster RCNN, Mobilenet-SSD et al.) should also be made. This investigation also showcased that model detection confidence would vary depending on lighting angle and direction; this is something that should be explored further, to find the optimum lighting angle and direction for detecting different spalls. A limitation of the current system is the requirement for contact with the concrete surface; further investigations should identify the modifications to the apparatus and imaging environment required for non-contact photometric stereo.

CRediT authorship contribution statement

Hamish Dow: Writing – review & editing, Writing – original draft, Software, Methodology, Formal analysis, Conceptualization. **Marcus Perry:** Supervision, Formal analysis, Conceptualization. **Sanjeetha Pennada:** Resources. **Rebecca Lunn:** Supervision, Formal analysis. **Stella Pytharouli:** Supervision, Formal analysis.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Marcus Perry reports financial support was provided by Scottish Funding Council. Hamish Dow has patent pending to University of Strathclyde. If there are other authors, they 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

The data that has been used is confidential.

Acknowledgements

The authors express their gratitude to: **Dr Jack McAlorum:** Assisted with conceptualisation and conducted early photometric stereo experiments. **Prof Gordon Dobie:** Contributed to manuscript proofreading. **Mr Andrew Darby and Mr Feng Zhang:** Formed the defective concrete slabs used in the laboratory dataset of this manuscript.

References

[1] K.D. Hertz, Limits of spalling of fire-exposed concrete, Fire Saf. J. 38 (2) (2003) 103–116, http://dx.doi.org/10.1016/S0379-7112(02)00051-6.
[2] Y. Dong, Performance assessment and design of ultra-high performance concrete (UHPC) structures incorporating life-cycle cost and environmental impacts, Constr. Build. Mater. 167 (2018) 414–425, http://dx.doi.org/10.1016/j.conbuildmat.2018.02.037.
[3] Z. Zhu, S. German, I. Brilakis, Detection of large-scale concrete columns for automated bridge inspection, Autom. Constr. 19 (8) (2010) 1047–1055, http://dx.doi.org/10.1016/j.autcon.2010.07.016.
[4] T. Mizoguchi, Y. Koda, I. Iwaki, H. Wakabayashi, Y. Kobayashi, K. Shirai, Y. Hara, H.-S. Lee, Quantitative scaling evaluation of concrete structures based on terrestrial laser scanning, Autom. Constr. 35 (2013) 263–274, http://dx.doi.org/10.1016/j.autcon.2013.05.022.
[5] M.-K. Kim, H. Sohn, C.-C. Chang, Localization and quantification of concrete spalling defects using terrestrial laser scanning, J. Comput. Civ. Eng. 29 (6) (2015) 04014086, http://dx.doi.org/10.1061/(ASCE)CP.1943-5487.0000415.
[6] C. Popescu, B. Täljsten, T. Blanksvärd, L. Elfgren, 3D reconstruction of existing concrete bridges using optical methods, Struct. Infrastruct. Eng. 15 (7) (2019) 912–924, http://dx.doi.org/10.1080/15732479.2019.1594315.

[7] H. El-Din Fawzy, R. Kandeel, M. Farhan, Detection of deformations in reinforced concrete structures using modern surveying techniques, Alex. Eng. J. 70 (2023) 191–218, http://dx.doi.org/10.1016/j.aej.2023.02.030.

[8] K. Krajníková, J. Smetanková, P. Mésároš, M. Behún, Different approaches in building digitization through the use of 3D laser scanning, in: L. Knapcikova, M. Balog, D. Peraković, M. Periša (Eds.), New Approaches in Management of Smart Manufacturing Systems: Knowledge and Practice, Springer International Publishing, Cham, 2020, pp. 143–158, http://dx.doi.org/10.1007/978-3-030-40176-4_9.

[9] J. McAlorum, H. Dow, S. Pennada, M. Perry, G. Dobie, Automated concrete crack inspection with directional lighting platform, IEEE Sens. Lett. 7 (11) (2023) 1–4, http://dx.doi.org/10.1109/LSENS.2023.3327611, Conference Name: IEEE Sensors Letters.

[10] R.J. Woodham, Photometric method for determining surface orientation from multiple images, Opt. Eng., Bellingham 19 (1) (1980) 139–144, http://dx.doi.org/10.1117/12.7972479.

[11] H. Dow, M. Perry, J. McAlorum, S. Pennada, G. Dobie, Skeleton-based noise removal algorithm for binary concrete crack image segmentation, Autom. Constr. 151 (2023) 104867, http://dx.doi.org/10.1016/j.autcon.2023.104867.

[12] S. German, I. Brilakis, R. DesRoches, Rapid entropy-based detection and properties measurement of concrete spalling with machine vision for post-earthquake safety assessments, Adv. Eng. Inform. 26 (4) (2012) 846–858, http://dx.doi.org/10.1016/j.aei.2012.06.005.

[13] T. Dawood, Z. Zhu, T. Zayed, Detection and quantification of spalling distress in subway networks, in: K.W. Chau, I.Y. Chan, W. Lu, C. Webster (Eds.), Proceedings of the 21st International Symposium on Advancement of Construction Management and Real Estate, Springer, Singapore, 2018, pp. 607–615, http://dx.doi.org/10.1007/978-981-10-6190-5_55.

[14] X. Yao, M. Yao, B. Xu, Automated detection and identification of area-based distress in concrete pavements, in: Proceedings of the 7th International Conference on Managing Pavement Assets, Calgary, Alberta, Canada, 2008.

[15] Y. Xu, W. Qiao, Y. Bao, H. Li, Y. Zhang, Pixel-level damage detection for concrete spalling and rebar corrosion based on U-net semantic segmentation, in: Bridge Maintenance, Safety, Management, Life-Cycle Sustainability and Innovations, CRC Press, 2021, pp. 3319–3326.

[16] C. Zhang, C.-c. Chang, M. Jamshidi, Simultaneous pixel-level concrete defect detection and grouping using a fully convolutional model, Struct. Health Monit. 20 (4) (2021) 2199–2215, http://dx.doi.org/10.1177/1475921720985437.

[17] D. Amirkhani, M.S. Allili, L. Hebbache, N. Hammouche, J.-F. Lapointe, Visual concrete bridge defect classification and detection using deep learning: A systematic review, IEEE Trans. Intell. Transp. Syst. (2024) 1–23, http://dx.doi.org/10.1109/TITS.2024.3365296, Conference Name: IEEE Transactions on Intelligent Transportation Systems.

[18] G. Jocher, A. Chaurasia, J. Qiu, YOLO by ultralytics, Github, 2023, URL https://github.com/ultralytics/ultralytics.

[19] C. Xiong, T. Zayed, E.M. Abdelkader, A novel YOLOv8-GAM-wise-IoU model for automated detection of bridge surface cracks, Constr. Build. Mater. 414 (2024) 135025, http://dx.doi.org/10.1016/j.conbuildmat.2024.135025.

[20] M.A. Rouf, Y. Iwahori, H. Chen, A. Wang, A novel approach for concrete crack and spall detection based on improved YOLOv8, in: Proceedings of the 2023 5th International Conference on Video, Signal and Image Processing, VSIP '23, Association for Computing Machinery, New York, NY, USA, 2024, pp. 149–155, http://dx.doi.org/10.1145/3638682.3638705.

[21] L. Yang, B. Li, W. Li, B. Jiang, J. Xiao, Semantic metric 3D reconstruction for concrete inspection, in: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, CVPRW, IEEE, Salt Lake City, UT, USA, 2018, pp. 1624–16248, http://dx.doi.org/10.1109/CVPRW.2018.00204.

[22] H. Zhang, Y. Zou, E. del Rey Castillo, X. Yang, Detection of RC spalling damage and quantification of its key properties from 3D point cloud, KSCE J. Civ. Eng. 26 (5) (2022) 2023–2035, http://dx.doi.org/10.1007/s12205-022-0890-y.

[23] Y.Z. Xia, 3-d segmentation of concrete spalling in point cloud using unsupervised clustering and plane fitting, in: Life-Cycle of Structures and Infrastructure Systems, CRC Press, 2023, p. 8.

[24] M. Zhou, W. Cheng, H. Huang, J. Chen, A novel approach to automated 3D spalling defects inspection in railway tunnel linings using laser intensity and depth information, Sensors 21 (17) (2021) 5725, http://dx.doi.org/10.3390/s21175725.

[25] Q. Kong, J. Gu, B. Xiong, C. Yuan, Vision-aided three-dimensional damage quantification and finite element model geometric updating for reinforced concrete structures, Comput.-Aided Civ. Infrastruct. Eng. 38 (17) (2023) 2378–2390, http://dx.doi.org/10.1111/mice.12967.

[26] D. Marchisotti, E. Zappa, Feasibility study of drone-based 3-D measurement of defects in concrete structures, IEEE Trans. Instrum. Meas. 71 (2022) 1–11, http://dx.doi.org/10.1109/TIM.2022.3170969.

[27] G.H. Beckman, D. Polyzois, Y.-J. Cha, Deep learning-based automatic volumetric damage quantification using depth camera, Autom. Constr. 99 (2019) 114–124, http://dx.doi.org/10.1016/j.autcon.2018.12.006.

[28] T. Ghosh Mondal, M.R. Jahanshahi, Z.Y. Wu, Deep learning-based RGB-D fusion for multimodal condition assessment of civil infrastructure, J. Comput. Civ. Eng. 37 (4) (2023) 04023017, http://dx.doi.org/10.1061/JCCEE5.CPENG-5197.

[29] Y. Ding, X. Luo, Monocular 2D camera-based proximity monitoring for human-machine collision warning on construction sites, 2023, http://dx.doi.org/10.48550/arXiv.2305.17931.

[30] H. Haggag, M. Hossny, D. Filippidis, D. Creighton, S. Nahavandi, V. Puri, Measuring depth accuracy in RGBD cameras, in: 2013, 7th International Conference on Signal Processing and Communication Systems, ICSPCS, 2013, pp. 1–7, http://dx.doi.org/10.1109/ICSPCS.2013.6723971.

[31] D. Lichy, S. Sengupta, D.W. Jacobs, Fast light-weight near-field photometric stereo, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, Louisiana, USA, 2022, pp. 12612–12621.

[32] M. Ren, X. Wang, G. Xiao, M. Chen, L. Fu, Fast defect inspection based on data-driven photometric stereo, IEEE Trans. Instrum. Meas. 68 (4) (2019) 1148–1156, http://dx.doi.org/10.1109/TIM.2018.2858062.

[33] G. Podrekar, D. Tomaževič, B. Likar, P. Usenik, Model based visual inspection of pharmaceutical tablets with photometric stereo, in: 2017 Fifteenth IAPR International Conference on Machine Vision Applications, MVA, 2017, pp. 133–136, http://dx.doi.org/10.23919/MVA.2017.7986819.

[34] L. Wang, K. Xu, P. Zhou, Online detection technique of 3D defects for steel strips based on photometric stereo, in: 2016 Eighth International Conference on Measuring Technology and Mechatronics Automation, ICMTMA, 2016, pp. 428–432, http://dx.doi.org/10.1109/ICMTMA.2016.109.

[35] E. Weigl, S. Zambal, M. Stöger, C. Eitzinger, Photometric stereo sensor for robot-assisted industrial quality inspection of coated composite material surfaces, in: Twelfth International Conference on Quality Control By Artificial Vision 2015, vol. 9534, SPIE, 2015, pp. 367–374, http://dx.doi.org/10.1117/12.2182750.

[36] Y. Cao, B. Ding, J. Chen, W. Liu, P. Guo, L. Huang, J. Yang, Photometric-stereo-based defect detection system for metal parts, Sensors 22 (21) (2022) 8374, http://dx.doi.org/10.3390/s22218374.

[37] A. Landstrom, M.J. Thurley, H. Jonsson, Sub-millimeter crack detection in casted steel using color photometric stereo, in: 2013 International Conference on Digital Image Computing: Techniques and Applications, DICTA, 2013, pp. 1–7, http://dx.doi.org/10.1109/DICTA.2013.6691532.

[38] J. Blair, B. Stephen, B. Brown, S. McArthur, D. Gorman, A. Forbes, C. Pottier, J. McAlorum, H. Dow, M. Perry, Photometric stereo data for the validation of a structural health monitoring test rig, Data Brief 53 (2024) 110164, http://dx.doi.org/10.1016/j.dib.2024.110164.

[39] 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. Pavem. Eng. 23 (14) (2022) 5168–5185, http://dx.doi.org/10.1080/10298436.2021.2001811.

[40] P.J. Sanchez-Cuevas, P. Ramon-Soria, B. Arrue, A. Ollero, G. Heredia, Robotic system for inspection by contact of bridge beams using UAVs, Sensors 19 (2) (2019) 305, http://dx.doi.org/10.3390/s19020305.

[41] R. Watson, M. Kamel, D. Zhang, G. Dobie, C. MacLeod, S.G. Pierce, J. Nieto, Dry coupled ultrasonic non-destructive evaluation using an over-actuated unmanned aerial vehicle, IEEE Trans. Autom. Sci. Eng. 19 (4) (2022) 2874–2889, http://dx.doi.org/10.1109/TASE.2021.3094966.

[42] L. Yang, B. Li, J. Feng, G. Yang, Y. Chang, B. Jiang, J. Xiao, Automated wall-climbing robot for concrete construction inspection, J. Field Robotics 40 (1) (2023) 110–129, http://dx.doi.org/10.1002/rob.22119.

[43] G.G. Garrido, T.P. Sattar, An autonomous wall-climbing robot for inspection of reinforced concrete structures: SIRCAUR, J. Artif. Intell. Technol. 1 (3) (2021) 188–196, http://dx.doi.org/10.37965/jait.2021.0016.

[44] S.K. Mahmood, S.H. Bakhy, M.A. Tawfik, Propeller-type wall-climbing robots: A review, IOP Conf. Ser.: Mater. Sci. Eng. 1094 (1) (2021) 012106, http://dx.doi.org/10.1088/1757-899X/1094/1/012106.

[45] SHM (Roboflow Universe User), Concrete Defect Detection Dataset, Roboflow, Roboflow Universe, 2023, URL https://universe.roboflow.com/shm/concrete-defect-detection.