Data-Centric Engineering (2026), 7: e10
doi:10.1017/dce.2026.10042

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

Denoising image point clouds using segmentation and synthetic data for enhanced structural health analysis of tunnels

Ran Zhang[1], Wei Lin[2], Chao Wang[1], Brian Sheil[3], Zhongqiang Liu[4] and Zili Li[1]

[1]University College Cork, Ireland
[2]Tongji University, China
[3]University of Cambridge, UK
[4]Norwegian Geotechnical Institute, Norway
Corresponding author: Zili Li; Email: zili.li@ucc.ie

Received: 02 March 2025; **Revised:** 20 September 2025; **Accepted:** 02 November 2025

Keywords: denoising; point cloud; segmentation; structural health monitoring; synthetic dataset

Abstract

Point clouds derived from UAV photogrammetry are a cost-effective alternative to LiDAR for infrastructure inspections, but they often include both structural and non-structural elements that complicate analysis. Traditional denoising filters remove outliers indiscriminately and frequently erode edges, making it difficult to preserve the curved tunnel lining while distinguishing bolts, access gates, or pipelines. In contrast, segmentation-based approaches leverage geometric context to explicitly separate lining surfaces from ancillary components, thereby enabling more accurate deformation analysis and structural assessment. To that end, this paper presents a novel approach for denoising image point clouds using a synthetic training dataset to address the scarcity of labeled public data for enhancing point cloud quality. Unlike other denoising approaches that rely on projections or assume points lie on a predefined surface shape, this segmentation-based denoising method retains only meaningful points in their original locations, allowing for more accurate analysis of deformation. Enhanced by synthetic training datasets, the application of the proposed denoising method to a road tunnel image point cloud and a subway tunnel terrestrial laser scanning point cloud demonstrates its potential to enhance point cloud quality in tunnels with diverse geometries and point cloud data resources, even when data are limited. The method achieves an 80% mean intersection over union for both the road tunnel and the subway tunnel from manual annotation. This enables an improvement in structural deformation analysis at the mm level.

Impact statement

This study presents a novel 3D deep learning approach for improving unmanned aerial vehicle-based tunnel inspections by addressing the limitations of point cloud accuracy. Using a synthetic dataset for training, the method effectively denoises point clouds by removing erroneous points and preserving crucial structural details. This approach enhances the accuracy and efficiency of tunnel monitoring, offering a scalable solution for capturing tunnel geometry and surface deformations in complex underground environments. The proposed technique significantly improves the feasibility of automatic tunnel inspections, supporting better infrastructure maintenance and safety assessment.

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/dce.2026.10042 Published online by Cambridge University Press

1. Introduction

Image-based inspection has become an essential method for monitoring and maintaining tunnel infrastructure, particularly as the deterioration of such structures continues to escalate (Attard et al., 2018; Attard et al., 2018; Li et al., 2021). The monitoring of tunnel spatial structural deformation forms a crucial element of its structural health assessment (Xu and Yang, 2020; Yang and Xu, 2021). Whilst traditional point-wise measurement techniques, such as deformation monitoring nodes, have been widely used, they do not capture the complete deformation profiles of tunnels, providing limited insight into their structural health (Wang et al., 2023). Unmanned aerial vehicles (UAVs) have become a popular and scalable alternative for tunnel structural inspection, enabling continuous monitoring of large tunnel surfaces and delivering comprehensive data for structural health assessments, even in challenging, geometrically complex tunnels (Li et al., 2015; Zhang et al., 2023).

UAV photogrammetry produces point clouds as a direct output, enabling the extraction of spatial information and detailed analysis of a tunnel's observable (inner) surfaces. By fitting geometric primitives (e.g., circles or ellipses) to the curved tunnel lining, UAV-derived point clouds enable geometric analysis of cross-sections and the detection of deviations from the expected geometry. These deviations can indicate potential tunnel deformations and support reliable structural health assessment. The concept of extracting deformations from UAV-generated digital surface models (DSMs) originated from Pietersen's work on airport pavement inspections, where an accuracy of 6 mm was achieved (Pietersen, 2022). Similarly, UAV inspections of bridges (Chen et al., 2019) and other infrastructure (Lattanzi and Miller, 2017; Greenwood et al., 2019; Keizer et al., 2022) have achieved cm-level accuracy. For tunnels, UAV inspections have been conducted in a cross-passage twin tunnel section, along with geometric analysis, to assess its ageing deformational performance (Zhang et al., 2024). A key benefit of UAV-generated point clouds is their ability to provide orthomosaics, which include color features, and DSMs that precisely capture the positions of surface points.

While photogrammetry-based point clouds offer a flexible, scalable and cost-effective solution for infrastructure monitoring, one significant limitation is their lower accuracy compared to other methods (Gaspari et al., 2022). Tunnel inspections, in particular, present additional challenges due to photogrammetry inside the tunnels, which increases the likelihood of artifacts, rough surfaces and outliers being included in the generated model (Wolff et al., 2016). Such noise occurs near both hyper-spectral areas, such as structural edges, and low-spectral areas, such as artifacts. The latter is typically located within 15 cm of the true surface point (Peterson et al., 2019). However, research has shown that surface fitting techniques can reduce the mean error to as low as 1 mm, approaching the accuracy of terrestrial laser scanning (TLS) data (Peterson et al., 2019). Nonetheless, cameras often struggle to capture features in areas with uneven lighting, resulting in cavities or densely populated regions in the point cloud, depending on what the computer vision algorithm detects (Gevaert et al., 2017). These variations in point cloud density complicate the denoising process, particularly around edges where surfaces of different colors meet. It is acknowledged improved data collection campaigns, such as enhanced illumination, higher magnification, or more controlled UAV flight paths, can reduce some of these artifacts. However, even under well-planned conditions, photogrammetry is inherently subject to noise from imaging principles (e.g., texture variation, occlusions, and reflections). In practice, inspection campaigns inside tunnels are often constrained by safety requirements, limited lighting, and access restrictions, meaning that ideal acquisition conditions cannot always be achieved. This makes robust post-processing and denoising approaches a necessary complement to improved data collection practices. A further challenge, often overlooked, is that tunnel point clouds typically capture not only the curved lining surfaces but also numerous ancillary elements such as bolts, pipelines, access gates, and temporary equipment. For reliable structural health assessment, it is essential to separate the lining geometry from these non-structural components. Traditional outlier removal filters cannot perform this discrimination, as they treat all deviations as noise. Segmentation-based deep learning approaches, however, can learn to distinguish between lining and non-lining elements, enabling both denoising and more precise structural analysis. It is important to note that the aim of segmentation in this study is to

https://doi.org/10.1017/dce.2026.10042 Published online by Cambridge University Press

isolate the tunnel lining for geometric analysis, rather than to directly detect surface defects such as cracks or water leakages.

Research on point cloud denoising has recently gained significant momentum (Sun et al., 2023), with a focus on improving both accuracy and efficiency. Manual denoising is often the first step in many point cloud processing workflows (Zeybek and Şanlıoğlu, 2019), but it is subjective and time-consuming (Zhou et al., 2022). Traditional filters, such as statistical outlier removal (SOR) filters (Balta et al., 2018; Ortiz-Coder and Sánchez-Ríos, 2020), are widely used but can inadvertently remove important points (Zeybek, 2021), particularly at corners and edges in tunnel structures where clustered noise is sometimes mistaken for part of the actual structure (Zhang et al., 2023). Recent research suggests that noisy points can be effectively detected and segmented by leveraging their spatial information (Vetrivel et al., 2015). Segmentation-based methods have also been shown to outperform Gaussian filters and traditional feature denoising techniques, as they are more capable of learning the geometry of tunnels and identifying obstacles (Rhee and Kim, 2016). Despite this progress, several critical challenges remain. Segmentation-based approaches specifically tailored for tunnel structures are not well-established, creating uncertainty about how such models should be developed, trained, and applied in real-world tunnel inspection scenarios. The lack of segmentation models designed for tunnels raises concerns regarding their performance, especially in complex environments with curved surfaces and inconsistent lighting. Additionally, the existing denoising techniques struggle to differentiate between actual structural features and noise in data, particularly in areas where point cloud density fluctuates or where visual features are occluded or distorted. This challenge is compounded by the absence of well-annotated, diverse datasets representing realistic tunnel conditions, limiting the development and validation of more accurate denoising solutions.

To address this issue, this paper introduces a new 3D deep learning (DL) approach that leverages synthetic data for tunnel point cloud cleaning, addressing the issue of low accuracy in image-based point clouds and improving the feasibility of UAV-based automatic tunnel inspection. The synthetic dataset is generated without requiring manual annotation. This dataset accounts for data deficiencies, such as cavities and noise clusters near tunnel surfaces, which are caused by the low quality of visual features (Sapronova et al., 2023). It also addresses the lack of well-annotated datasets collected under similar inspection conditions. The DL segmentation model is then trained using synthetic data and evaluated by real-world datasets. Once trained, the segmentation model can process new tunnel point clouds to produce denoised outputs in real time. The effectiveness of the proposed denoising method is validated through application to two real-world cases: (a) the Dublin port tunnel (DPT) cross-passage twin-tunnel, and (b) the open-source Wuxi dataset comprising Metro Line 4 tunnels from Wuxi, China. Unlike traditional methods that project all points onto an assumed tunnel surface, this approach accurately measures the tunnel's cross-sectional geometry by removing erroneous points while preserving the integrity of the tunnel lining in the final results. In summary, the proposed segmentation-based deep learning framework not only suppresses noise but, more importantly, separates curved tunnel linings from surrounding components, thereby preserving geometric fidelity for reliable geometric analysis and primitive fitting to support deformation assessment.

2. Synthetic data generation and DL-based denoising

2.1 Denoising workflow

In this paper, a 3D DL denoising framework is proposed to reduce the time and cost associated with traditional manual denoising processes (see Figure 1). This workflow presents an adaptive denoising methodology for diverse tunnel point cloud quality, while the trained model is tailored to the specific inspection environment by the input parameters. The following section demonstrates this process with an example from the VCP 16 section in the DPT dataset. The flowchart of this framework consists of four phases: (1) point cloud visual assessment, (2) synthetic data generation, (3) training with the RandLA-Net DL architecture, and (4) application to the infrastructure point clouds.

https://doi.org/10.1017/dce.2026.10042 Published online by Cambridge University Press

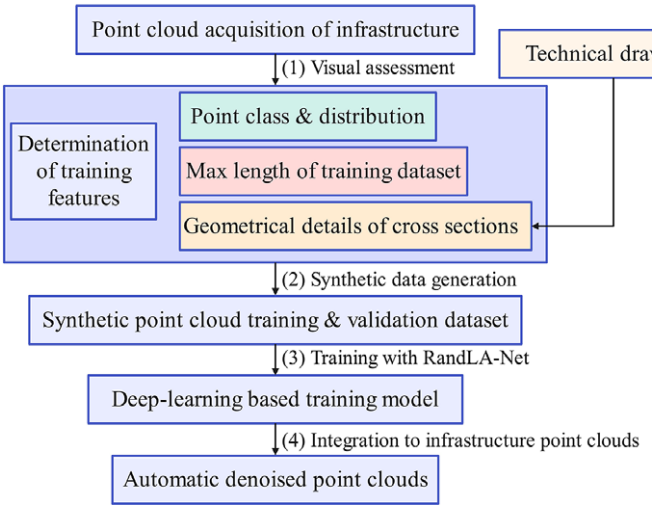


Figure 1. *Flowchart of proposed denoising methodology (algorithm in (Hu et al., 2022; Lin et al., 2024)).*

Table 1. *Input parameters adopted for the synthetic data generation algorithm*

(Step 1)	Dimensions of tunnel cross section (lining, pavement and crown),
Tunnel model creation	Model length
(Step 2)	Lining cavity-to-tunnel point number ratio (*ratio_crop_lin*),
Cavity point removal	Pavement cavity-to-tunnel point number ratio (*ratio_crop_pav*),
	Cavity size modeled as circles with radius ~ U_1 (0, r), generated via a spatial Poisson process
(Step 3)	Lining noise-to-tunnel point number ratio (*ratio_noise_lin*),
Noise addition	Pavement noise-to-tunnel point number ratio (*ratio_noise_pav*),
	Noise cluster size modeled as rectangles with side lengths l_1, l_2 ~ U_2 (0, a), generated via a spatial Poisson process
	Noise cluster resolution (*res_noise*)

In the first phase, key parameters for synthetic data generation are quantified as per Table 1. The tunnel cross-section adopts either circular or elliptical profiles, vertically truncated at a 2.5 m vertical clearance to emulate road pavement obstruction. Cavity defects are defined by volume ratio with diameters 0-r cm, whereas noise clusters follow spatial Poisson distribution having rectangular shapes. In the second phase, the synthetic data are generated in three steps: building the uniform tunnel section, removing cavity points and adding noise clusters, as described in the following content:

1) Tunnel model creation: the geometry of the tunnel cross-section can be defined according to the dimensions specified in the tunnel technical drawings, including elements such as lining, pavement, and crown. The tunnel length is determined to maintain a balance between the computational load for training and the amount of training information required.
2) Cavity removal: The cavity in a tunnel point cloud model represents the absence of points that should have been captured in the tunnel model. It results from insufficient texture near smooth tunnel surfaces. The extent of cavities is quantified by proportionally removing points in a circular shape from the tunnel model initially created. Two groups of key parameters in Table 1 determine

https://doi.org/10.1017/dce.2026.10042 Published online by Cambridge University Press

the condition of cavity defects: the ratio of cavity points to total tunnel points and the physical dimensions (diameter of the circle) of the non-textured areas on the tunnel surface.

3) Noise addition: Similarly, noisy points are spurious artifacts that are not part of the true tunnel structure. Such noise is introduced by proportionally adding points in rectangular-shaped areas to mimic noise clusters. A random generation of the cavities and the noise is achieved by a spatial Poisson distribution with similar parameters in Table 1.

For instance, in the DPT dataset, an appropriate model length is determined to be 40 m according to the tunnel dimension. This model length provides enough segmentation data for training and saves computational costs. After selecting the input parameters, the synthetic data generation algorithm produces labeled point clouds. This phase simulates realistic noise distributions caused by image stitching artifacts, thereby removing the need for manual annotation.

In the third phase, the DL model is trained using the RanDLA-Net algorithm, enabling segmentation of the point cloud based on spatial information.

2.2 Synthetic dataset preparation

Figure 2 illustrates an example synthetic dataset generated for DPT comprising three main classes of points: pavements (red), linings (green), and noise (blue). In this way, the proposed method produces fundamental characteristics of real-world point cloud data. In this study, the generation of the training and validation datasets is achieved using the synthetic data generation algorithm. There are 16 synthetic training datasets and 4 validation datasets. Advantages of synthetic point clouds include a substantial increase in efficiency and accuracy of annotation compared to a manual approach, particularly for noise which is commonly the minority class in a highly imbalanced dataset, whose annotation inevitably subjects to human errors. Additionally, the original real DPT dataset alone is insufficient for effective training, whereas the use of synthetic data ensures sufficient data volume.

2.3 3D DL method

In the fourth phase of the flowchart in Figure 1, the 3D DL algorithm is trained on the synthetic training dataset to generate a prediction model of the inspected area. The 3D DL RandLA-Net model (Hu et al., 2022;

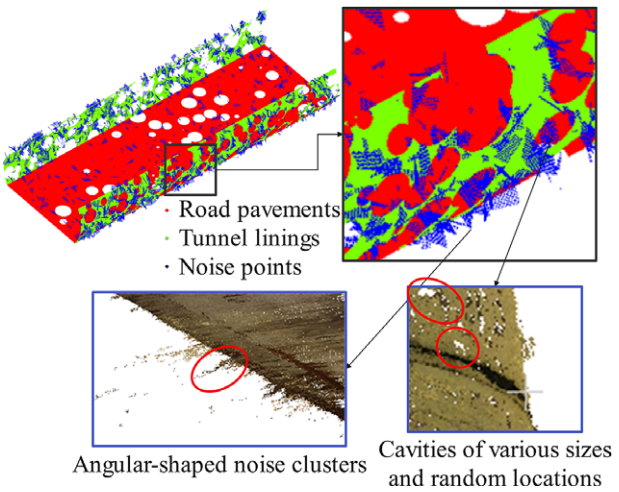


Figure 2. *Illustration of one realization of the synthetic dataset containing noise clusters and cavities (green: tunnel linings, blue: noise, red: road pavement), and zoomed image point clouds from VCP 16 test dataset.*

https://doi.org/10.1017/dce.2026.10042 Published online by Cambridge University Press

Lin et al., 2024) is adopted to train and validate with synthetic datasets, due to its advantage of processing large-scale point clouds while preserving complex geometric structures. Figure 3(a) illustrates the underlying principles of the RandLA-Net architecture. First, random down-sampling is used to reduce the computational cost whilst retaining key features. From left to right, the layers (1) to (3) in Figure 3 (a) describe two local feature aggregation processes. Local spatial encoding (L) is used to extract valuable spatial information and attentive pooling (A) is used for information aggregation. As a result, one point in the colored circles is used to represent several neighbor points, and this process progressively increases the receptive field size with minimal information loss.

Figure 3(b) provides a down-sampled example from the synthetic dataset where the DL method reduces the number of points by three orders of magnitude whilst retaining important features. In contrast, traditional filters use a fixed number of points to find those close to the centers of point clusters, regardless of the real-world geometry of the point clusters. For instance, Figure 4 presents a Sigma SOR filter, which statistically analyses the K-nearest neighboring points (p_i) to each point $p_j$to calculate the threshold distance for point denoising (K = 3 in this particular example). The analysis is conducted to establish a threshold formula:

$$D_{max} = i + n_\sigma \cdot \sigma \tag{1}$$

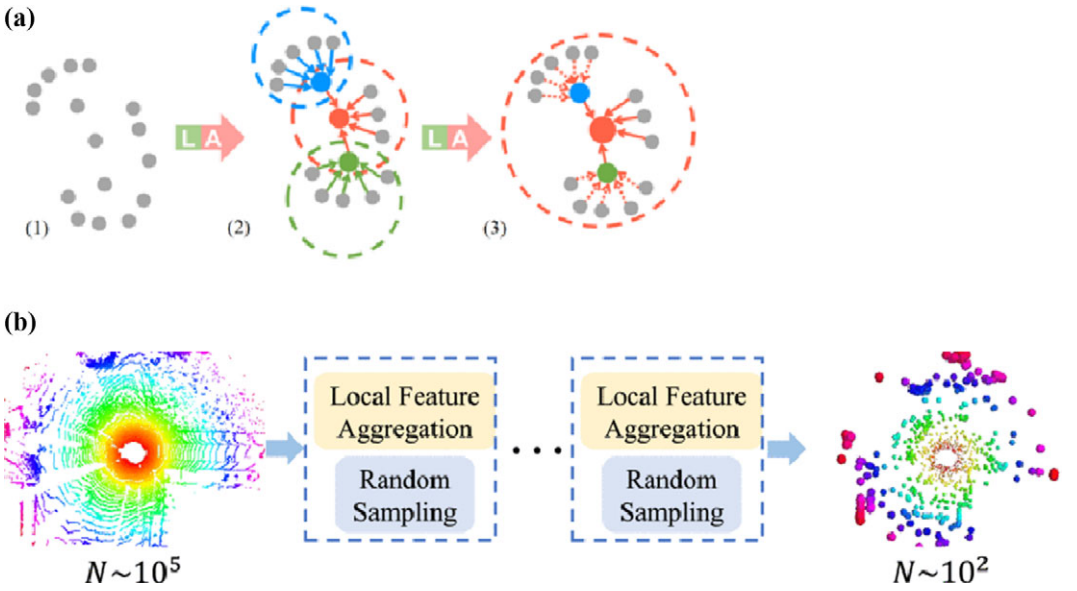


Figure 3. *Principle of RandLA-Net (a) Schematic illustration and (b) Down-sampling performance (Hu et al., 2022).*

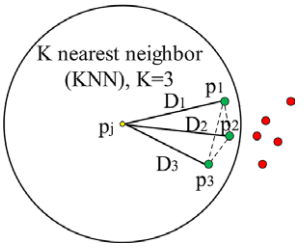


Figure 4. *Schematic illustration of the principle of traditional filters.*

https://doi.org/10.1017/dce.2026.10042 Published online by Cambridge University Press

where i is the global average distances from each point to its K nearest neighbor points, and n_σ is a user-defined multiplier of the deviations (σ). For p_i, if the distance between p_i and $p_j(D_i)$ exceeds threshold D_{max}, it will be considered noise (red points in Figure 4). This method presupposes that noise points are distributed in a uniform manner around the surface, which is not the case around corners in tunnels. Consequently, the Sigma SOR filter struggles to preserve geometric features in intricate underground structures, even though it may be well-suited for denoising smooth surfaces.

The RandLA-Net is trained using a batch size of 4 and a learning rate of 0.01. The network has 5 layers with subsampling ratios of (Attard et al., 2018; Xu and Yang, 2020), and the output dimensions of the layers are [16, 64, 128, 256, 512]. The above hyperparameters apply to all case studies. Once trained, the model is applied to identify noise in the unseen raw real-world infrastructure point cloud.

2.4 Evaluation metrics

Model training was performed on a consumer-grade computer with NVIDIA GeForce RTX 3080 Ti GPU and Intel core i9 12900K CPU. Training of a 40 m model over 100 epochs takes around 16 hours. The cross-entropy loss (Schult et al., 2023; Kolodiazhnyi et al., 2024) is used for the semantic segmentation of the point clouds, and the overall accuracy (OA) and mean intersection over union (mIoU) are used as evaluation metrics for these imbalanced datasets. These metrics are calculated to evaluate the training performance in each iteration (Rainio et al., 2024):

$$OA = \frac{\sum_{i=1}^{n} TP_i}{n} \tag{1}$$

$$IoU = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP} + \mathrm{FN}} \tag{2}$$

$$mIoU = \frac{1}{n}\sum_{i=1}^{n} IoU_i \tag{3}$$

Where OA is overall accuracy, n is the total number of classes, IoU (and IoU_i) is Intersection over Union (and for the ith class), TP (and TP_i) is the number of true positives (and for the i-th class), FP is the number of false positives, FN is the number of false negatives, and $mIoU$ represents the mean intersection over union.

3. Real-world image-based point clouds

The proposed point cloud methodology is based on point segmentation and verified with one image dataset (DPT) and one TLS dataset (Wuxi). The former is acquired with a UAV in a road tunnel and the latter is from a LiDAR scanner in a subway tunnel. Unlike the TLS dataset, image-based point clouds are less accurate in point cloud position information but contain color information. Section 3 introduces the generation of this kind of point cloud from images to show the error forms and characteristics.

3.1 Data acquisition

DPT is one of the longest urban road tunnels in Europe. The DPT twin tunnel structure exhibits tunnel lining deformation and surface defects that have developed over time, necessitating regular structural health condition assessments. The inspection area covers the Southbound Bore (SB) and the Northbound Bore (NB), connected by vehicle cross passages (VCPs), as shown in Figure 1. Inspections focused on the layby tunnel within the "VCP 16" area measuring 40 m in length along the traffic direction, which displayed the most significant surface defects compared to the other sections of the tunnel. In the zone of interest (the inspection area), there are additional adjoining structures, such as lay-by tunnels and VCPs at

https://doi.org/10.1017/dce.2026.10042 Published online by Cambridge University Press

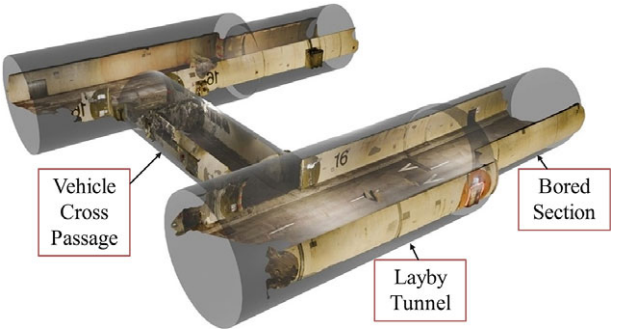


Figure 5. *3D visualization of DPT VCP 16 section derived from UAV images (left SB connected to right NB via VCP).*



Figure 6. *UAV inspection area as part of VCP 16 section in DPT.*

intervals of 1 km, components, such as fire hoses every 65 m, and telephone niches every 250 m. The DPT lining encompassing the abovementioned structures and components, was inspected, resulting in a 3D digital model of the inspected section (see Figure 5).

The visual data acquisition platform used is a C2 class DJI Mavic 2 Enterprise Dual. Images were captured using a pre-programmed reactive UAV specifically adapted for automatic and stable inspections in road tunnels such as DPT (Zhang et al., 2024). Figure 6 illustrates the UAV inspection environment under dim lighting conditions. The tunnel linings and road pavement are visible, but the tunnel crown lacks sufficient visual details and is excluded from the UAV inspection. The images were taken from a distance of approximately 3 m from the tunnel lining surface with a resolution of 4056 × 3040 pixels. A total of 2494 images were used for tunnel model reconstruction, totaling 13.3 GB. The inspection required multiple flights for safety reasons, and to account for possible changes in the inspection environment, with 30% of the images were recaptured, potentially introducing artifacts and necessitating data cleaning. The presence of noise in image point clouds introduces noticeable surface roughness, adversely affecting the extraction of spatial information from the tunnel linings. In the context of tunnel health condition assessment, it is imperative to utilize accurate data, free from noise, to ensure reliable outcomes.

3.2 Point cloud generation

The image point cloud is generated to capitalize on the paired color and positional information, thereby facilitating the concurrent identification of defects and deformation. Figure 7 outlines the primary image processing workflow containing steps (1) to (5) illustrated by Figure 8(1) to (5) for generating a UAV

https://doi.org/10.1017/dce.2026.10042 Published online by Cambridge University Press

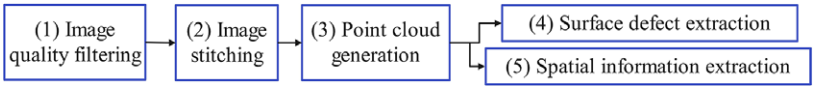


Figure 7. *Overview of point cloud processing workflow.*

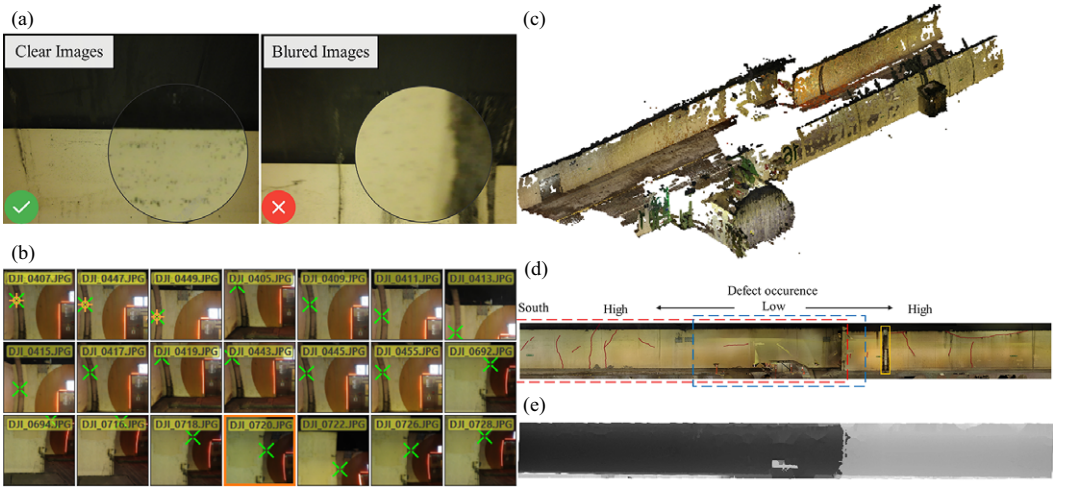


Figure 8. *Illustration of point cloud processing steps: (a) Image quality filtering, (b) Image stitching from multiple views (Orange cross: tie points; Green cross: control points), (c) Point cloud generation, (d) Defects from orthomosaic, and (e) Depth image from DSM.*

image point cloud from individual images. The image quality filtering step (1) removes blurry images that could cause misalignment of hyperspectral features like edges or corners (Figure 8(1)). The image stitching step (2) aligns images in 3D space using feature matching to create digital models of tunnel linings using the software Pix4Dmapper (Zhang et al., 2024), with Figure 8(2) showing the software interface during this process. The numeric label in the top-left corner (e.g., "DJI_0407") corresponds to the original image sequence number containing a specific example feature occurrence. The subsequent step (3) in Figure 8(3), point cloud generation, produces point clusters containing color or brightness information. The step (4), defect and spatial information extraction, yields paired datasets crucial for structural health analysis. Surface defects are identified and highlighted with distinct colors in Figure 8(4) (cracks in red, and water leakage in yellow box). In this workflow, image-based analysis is used to detect defects such as cracks and water leakages. The point cloud then provides positional information to accurately locate and size these defects. The DL-segmented point clouds, however, are not used for direct defect detection, but rather to ensure geometric fidelity of the tunnel lining for deformation analysis. Tunnel lining is divided into various zones according to the density of defects. The structural spatial information is incorporated into the DSM in Figure 8(5), which is displayed as depth images and can be extracted using GIS software.

The tunnel model reconstruction process is highly feature-sensitive, particularly to elements such as tunnel fissures, structural components, and edges, leading to non-uniform point cloud density. Under low-light conditions, UAV drifting can result in the generation of cavities, artifacts, and rough surfaces. These issues often stem from conflicts or the lack of hyperspectral information. Potential approaches to enhance the quality of tunnel model reconstruction include addressing geofencing-related point absences, improving sensor quality, and increasing the number of inspection views. While higher-quality sensors and additional inspection views could alleviate some of these limitations, the constrained tunnel environment and safety-related restrictions often prevent optimal acquisition strategies. As such, developing denoising

https://doi.org/10.1017/dce.2026.10042 Published online by Cambridge University Press

techniques that can operate effectively on imperfect datasets is essential for practical tunnel monitoring workflows.

3.3 Image point cloud deficiencies

Image point cloud deficiencies, including cavities, noise, and spurious rough surfaces, can be identified from the spatial distribution of the points. Figure 9 gives the extracted cross-sections of the tunnel lining from the manually denoised output, with data sourced from both the bored sections and the layby sections. Figure 9(a) shows the overall extraction map in the plan view, where two groups are marked with A–F and (1)–(4), respectively:

1) Labels A to F represents the extraction positions used to investigate tunnel behavior near the VCP structure, consistent with relevant prior research (Zhang et al., 2024). These extraction points are approximately evenly spaced to ensure data consistency while avoiding obstacles;

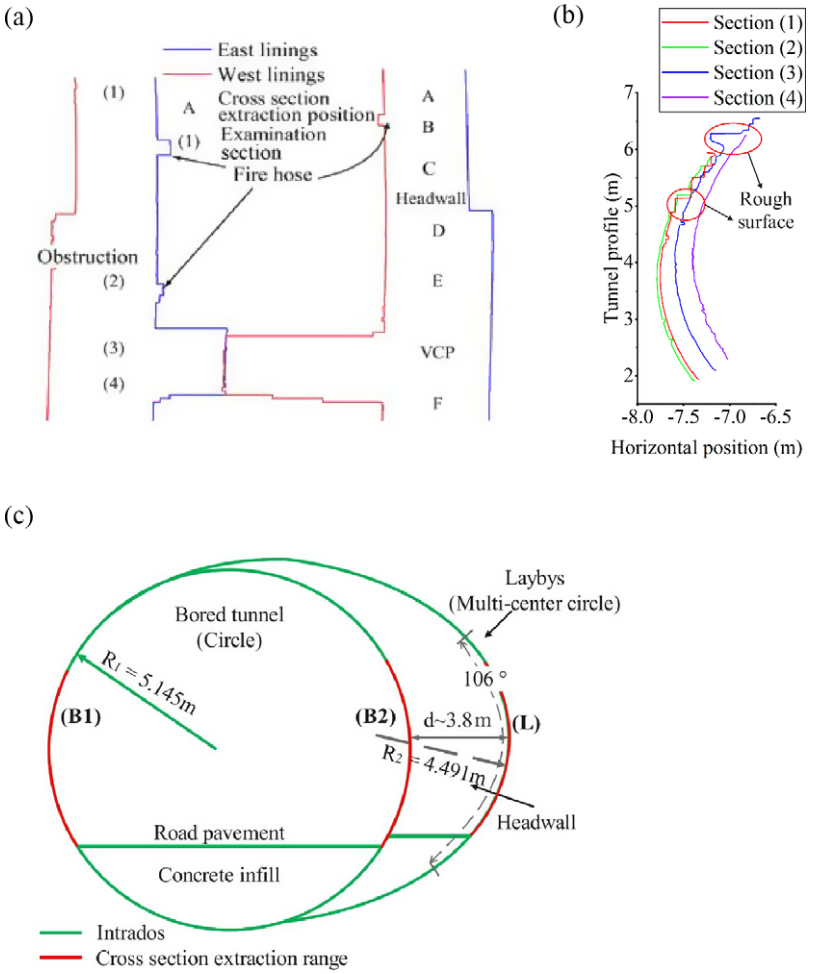


Figure 9. *Example of poor-quality cross-sectional extractions: (a) overall extraction map in plan view (two groups marked A to F and (1)–(4)); (b) tunnel lining profiles at examination sections (1)–(3) showing noise clusters; (c) details of tunnel circle fitting (on bored sections (B1 and B2) and layby sections (L)).*

https://doi.org/10.1017/dce.2026.10042 Published online by Cambridge University Press

2) Labels (1)–(4) represent examination sections, which are control sample locations used to evaluate the effectiveness of noise removal through circle fitting analysis. Specifically, Figure 9(b) shows the tunnel profile at examination Sections (1) to (3), with noise clusters highlighted.

A high-quality point cloud enables quantification of surface defects and the assessment of tunnel geometry. However, when comparing the tunnel's design specification with the acquired inspection data at key locations (see Figure 9(a)), the original DSM derived from the image point cloud inevitably contains greater errors than the tunnel deformation measurements. Therefore, these measurements must be denoised. Figure 9(c) illustrates the tunnel linings, with the bored sections labeled B1 and B2 and the layby sections denoted L. Also shown are the dimensions of the fitted tunnel diameter in the bored sections (modeled as circles) and layby areas (modeled as multi-center circles). Cross-sectional extractions (see Figure 9(b)) were used to represent the inspected part of the tunnel's configuration. Based on the as-built design (see Figure 9(c)), these sections should exhibit smooth circular or multi-center circular surfaces under undeteriorated tunnel condition.

The scale of noise is evaluated from the characteristics of the circles fitted to the DSM extractions. While the noise assessment is influenced by tunnel deformation, the noise clusters, which typically present as surface roughness, occur on a scale larger than the actual deformation of the tunnel lining. Hence, the noise level is assessed using the confidence interval of circle radii and centers, as listed in Table 1. A closer fit to the designed geometry of the bored and layby sections (R_1 and R_2 respectively) indicates fewer instances of clustered noise and better reconstruction accuracy, while a smaller confidence interval indicates less influences from the noise clusters.

In this study, manual point denoising was performed for comparison with later automated methods. It is worth noting that such manual processes remain subjective, prohibitively labor-intensive and introduce file encoding issues. Outlier removal is focused on areas near the edges of the tunnel linings and in areas with color variation. Comparisons of noise-free cross-sections in Table 1 reveals important structural health insights. For example, a slightly decreasing radius from sections D to F is observed, which suggests greater local compression near layby sections. This may be attributed to the diminished soil arching effect, as the soil is replaced by the cross-passage tunnel (Li et al., 2015). Between the twin tunnels, higher stiffness is observed on the east sides of both bores, especially over the right bracelet-shaped tunnel bend, due to increased compression resistance. Additionally, localized strengthening is observed around the headwall structures. Findings from previous research (Chao, 2023) for the same inspection area verify this deformation mechanism data. This information is critical for structural analysis and integrity assessment, making denoising an essential step to improve accuracy (Table 2).

Table 2. *Circle fit details of SB cross sections A–F (center coordinates (x_c, y_c) and radius r) together with mm-level fitting error*

		x_c /m	y_c /m	r /m
Bored tunnel	Section A	4.009 ± 5 E−4[a]	1.711	5.119
(Circle fit derived from the combined	Section B	3.838 ± 1 E−3	1.865	5.103
data of B1 and B2, $R_1 = 5.145m$)	Section C	3.749 ± 8 E−4	1.950	5.097
	Headwall	3.708	2.017	5.115
Laybys	Section D	3.663 ± 1 E−3	5.597 ± 7 E−3	4.459 ± 6 E−3
(Circle fit derived from L, part of	Section E	3.628 ± 1 E−3	5.493 ± 6 E−3	4.437 ± 6 E−3
the circular segment, $R_2 = 4.491m$)	VCP	3.494 ± 2 E−3	5.405 ± 9 E−3	4.381 ± 9 E−3
	Section F	3.380 ± 1 E−3	5.347 ± 6 E−3	4.273 ± 6 E−3

[a]Standard errors <5E-4 are omitted for brevity; larger errors are listed for circle fitting uncertainty.

https://doi.org/10.1017/dce.2026.10042 Published online by Cambridge University Press

4. Testing on DPT dataset

4.1 Experimental setting and model performance

The proposed denoising method is tested on two real-world cases: DPT and Wuxi metro. Application to DPT is first considered here. Each synthetic training and validation dataset contains 102,400 points. The parameters of the noise-free data in the synthetic model were chosen to mirror those of the original point cloud dataset, ensuring the denoised data was representative. A uniform cross-section from the layby area, representing the majority of the tunnel's length (see Figures 5 and 8), was used in the synthetic dataset to build the tunnel geometry. The cross section's multi-center circle shape was simplified to an ellipse with major and minor axes of 7.42 m and 5.18 m, respectively. The ellipse was truncated at heights of ±2.8 m to represent the ceiling and pavement. The tunnel model has a total length of 40 m. Its inner width varies at the headwall and VCP. The trained model was tested on both the bored and layby sections, and the results were combined to produce a denoised point cloud covering the entire tunnel length. The original image point cloud is down-sampled to maintain a resolution of 10 mm to preserve details and for training efficiency; the synthetic model shares the downsampling voxel size of 0.01 m.

The noise and cavity distribution in the DPT model are configured based on preliminary assessments from the image point cloud data. The noise is added with *ratio_noise_lin* set to 5% and *ratio_noise_pav* set to 1%. It is introduced at a spatial resolution of 0.05 m, and the dimensions of noise rectangular shapes are randomly ranging from 0 × 0 m to 0.5 × 0.5 m. Additionally, cavities are introduced to emulate imperfections in the image data acquisition, with *ratio_crop_lin* set to 50% and *ratio_crop_pav* to 10%. The radii of the circular cavities are randomly distributed between 0 m and 1 m.

Figure 10 illustrates the training process on the synthetic datasets. It shows that the DL model converges quickly and achieves good performance on both the training and validation data over 100 epochs, achieving a high mIoU score. Importantly, both training and validation loss decrease sharply and plateau to small and comparable labels indicating effective training without overfitting and thus is able to generalize between the datasets. After approximately 20 epochs, these evaluation metrics remain steady, indicating that it has reached an optimal state of learning. The consequent dataset prediction takes several minutes per model, showcasing advantage of efficiency for long-term or large dataset prediction over manual labeling.

Figure 11 shows the original raw point cloud (Figure 11(a)), the DL prediction of the primary segmented data in the DPT southbound bore (Figure 11(b)), and the DL-segmented data with noise removed (Figure 11(c)). The DL segmentation results reveal the present of noise around the tunnel lining. The road pavement and tunnel linings are generally well-segmented, but the narrow pedestrian area on the road pavement is not recognized, as it was not modeled in the synthetic data simulator. After removing the

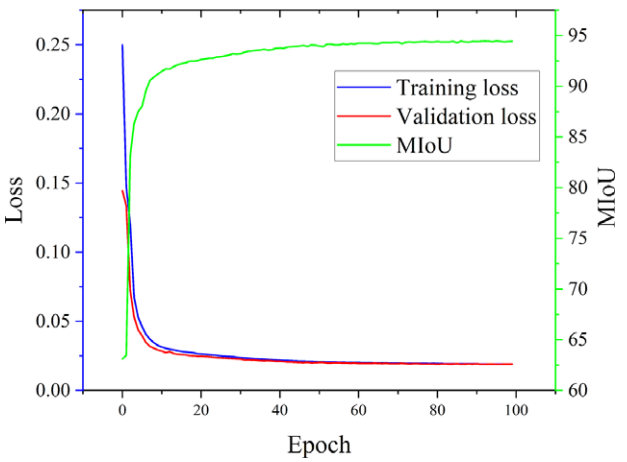


Figure 10. *Training process of RandLA-Net on the DPT synthetic dataset over 100 epochs.*

https://doi.org/10.1017/dce.2026.10042 Published online by Cambridge University Press

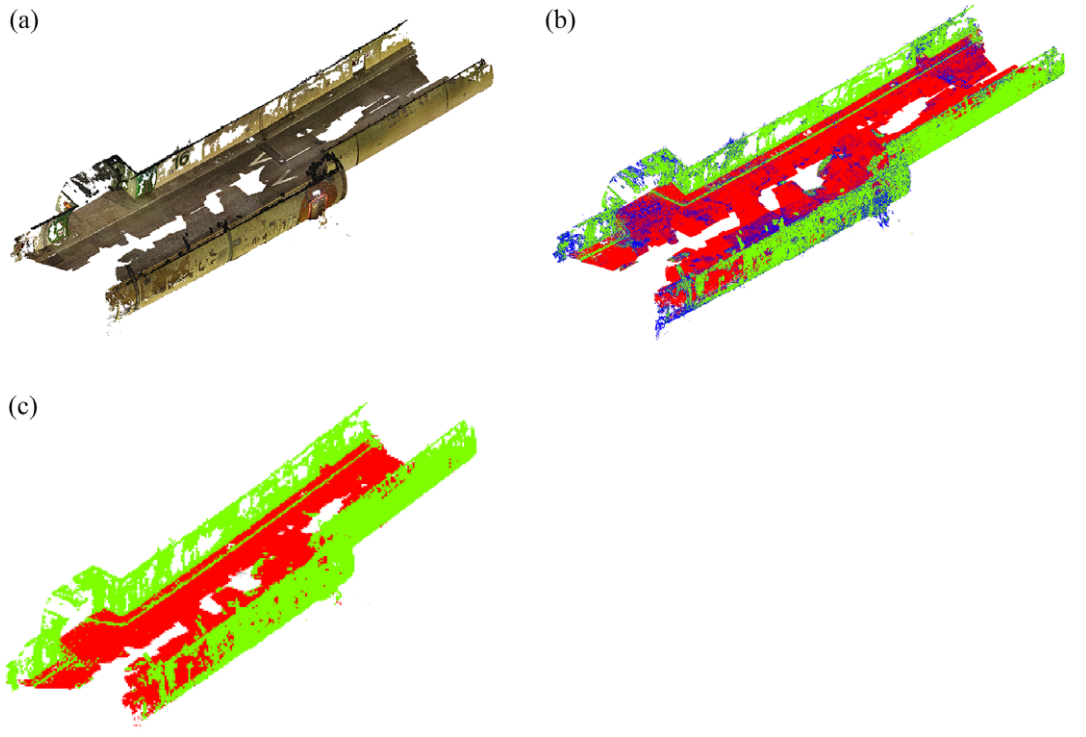


Figure 11. *Data visualization of the point cloud applied to SB in VCP 16: (a) Original raw real point cloud data, (b) DL-segmented data into primary classes, (c) DL-segmented following removal of noise.*

noise class, Figure 11(c) reduces noise points near low fidelity areas in the original model, including the southern part with many cavities, structural complex areas like fire door areas, and VCP areas.

4.2 Qualitative observations

DPT visualization model contains bulk noise and artifacts (similar to comet tail artifacts in medical imaging). Bulk Noise is observed as randomly distributed noise cover the majority of the model surface. It causes a lower resolution and grainy texture from the noises of the sensors. The artifacts, on the contrary, occurs more frequently on structural edges as fading lines due to various reasons, taking the incapacity in processing as an example.

Figure 12 illustrates the segmentation of the NB point cloud. Figure 12(a) presents the entire structure, focusing on two areas: the road pavement, highlighted by the green box and the tunnel lining edge, highlighted by the red box. A closer examination of them is shown in Figure 12(b) and (c), respectively. The data is presented in two classes: the predicted noise-free structure (black points) and the predicted noise (blue points). And that the resulting noise-free points in pavement section span an absolute range of 0.2 m and the noise spans an absolute range of 2 m. In contrast, for tunnel lining, the noise-free point and the noise span an absolute range of 0.2 m and 0.5-meter range, respectively. Overall, this technique effectively reduces artifact noise in targeted areas, enhancing the clarity and precision of the structural model.

Figure 13 illustrates the segmented photogrammetric point cloud of the SB layby tunnel generated from multi-temporal UAV image acquisitions across three inspection dates (September 2022 to March 2023). Dense noise clusters observed behind the fire door (highlighted in blue) predominantly originate from two synergistic factors: (1) temporal redundancy caused by repeated UAV flights capturing transient objects, including parked maintenance vehicles, temporary equipment installations, and dynamic door

https://doi.org/10.1017/dce.2026.10042 Published online by Cambridge University Press

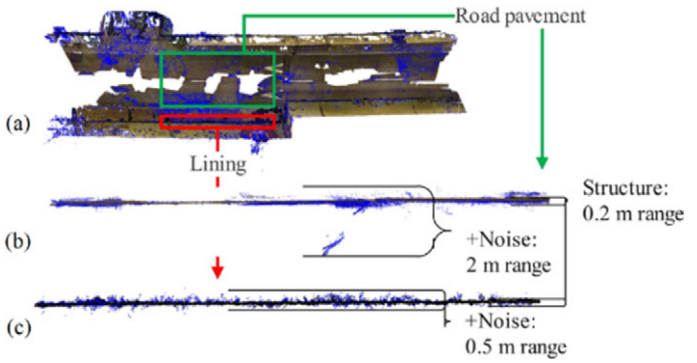


Figure 12. *Deep learning-based point cloud denoising shown in (a) the whole southbound, (b) a 14 m long road pavement, and (c) crown-lining border.*

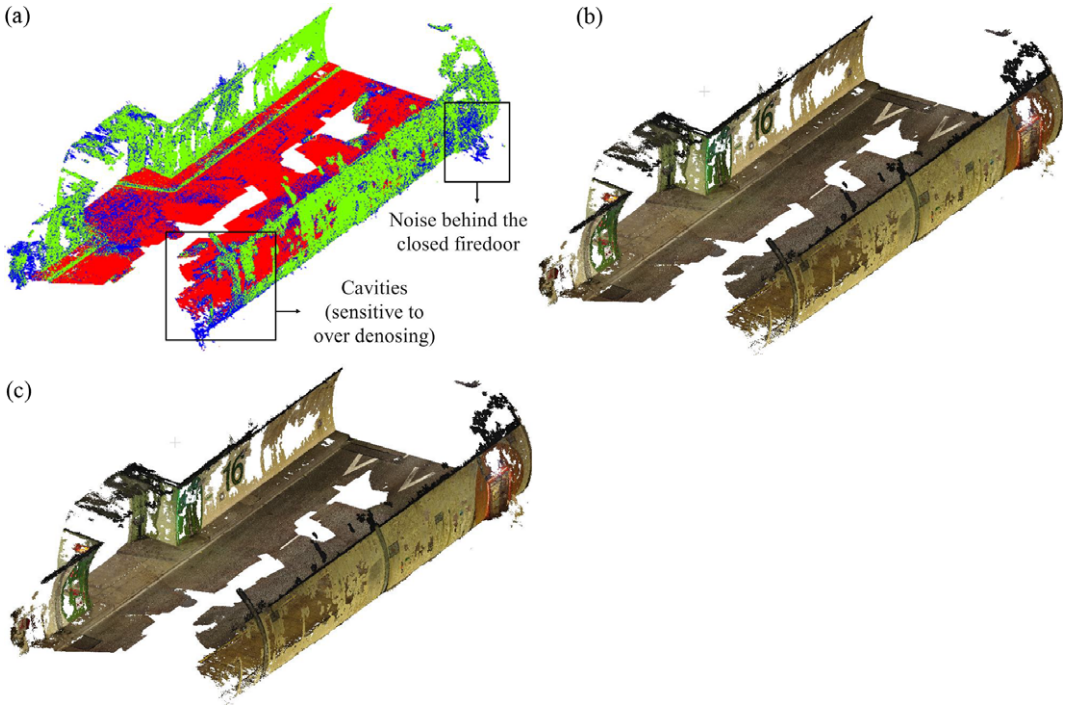


Figure 13. *Segmentation results of the (a) SPT SB layby tunnel, with two zoomed-in examples of noise: (b) near cavities, and (c) behind the fire door.*

positions (open/closed states during interventions), which were misinterpreted as persistent structural features during Structure-from-Motion (SfM) processing and (2) environmental variability induced by LED lighting upgrades midway through the monitoring period, resulting in illumination fluctuations between scans. These photometric inconsistencies degraded feature matching accuracy during image registration, propagating positional errors as false 3D points.

Figure 13 provides a three-dimensional visualization of the point cloud, where the different classes are again identified by the different colors. Noise points are particularly noticeable near the edges of the tunnel

https://doi.org/10.1017/dce.2026.10042 Published online by Cambridge University Press

lining and on top of the pavement. These areas correspond to artifacts near structural edges or erroneous data captured behind the fire door during scanning. The proposed effectively eliminated 7232 spurious points (84% of the initial 8560 noise points) in the fire door's closed-state region, improving the reconstructed model's geometric accuracy (RMSE reduction: from 2.6 m to 0.2 m). Moreover, it can be seen that the present DL method effectively retains two structural and geometric details, such as the junction between tunnel walls and the road surface.

In contrast, Figure 14 shows the point cloud using a traditional SOR approach for three different combinations of K and σ, representing the number of points used for mean distance estimation and the

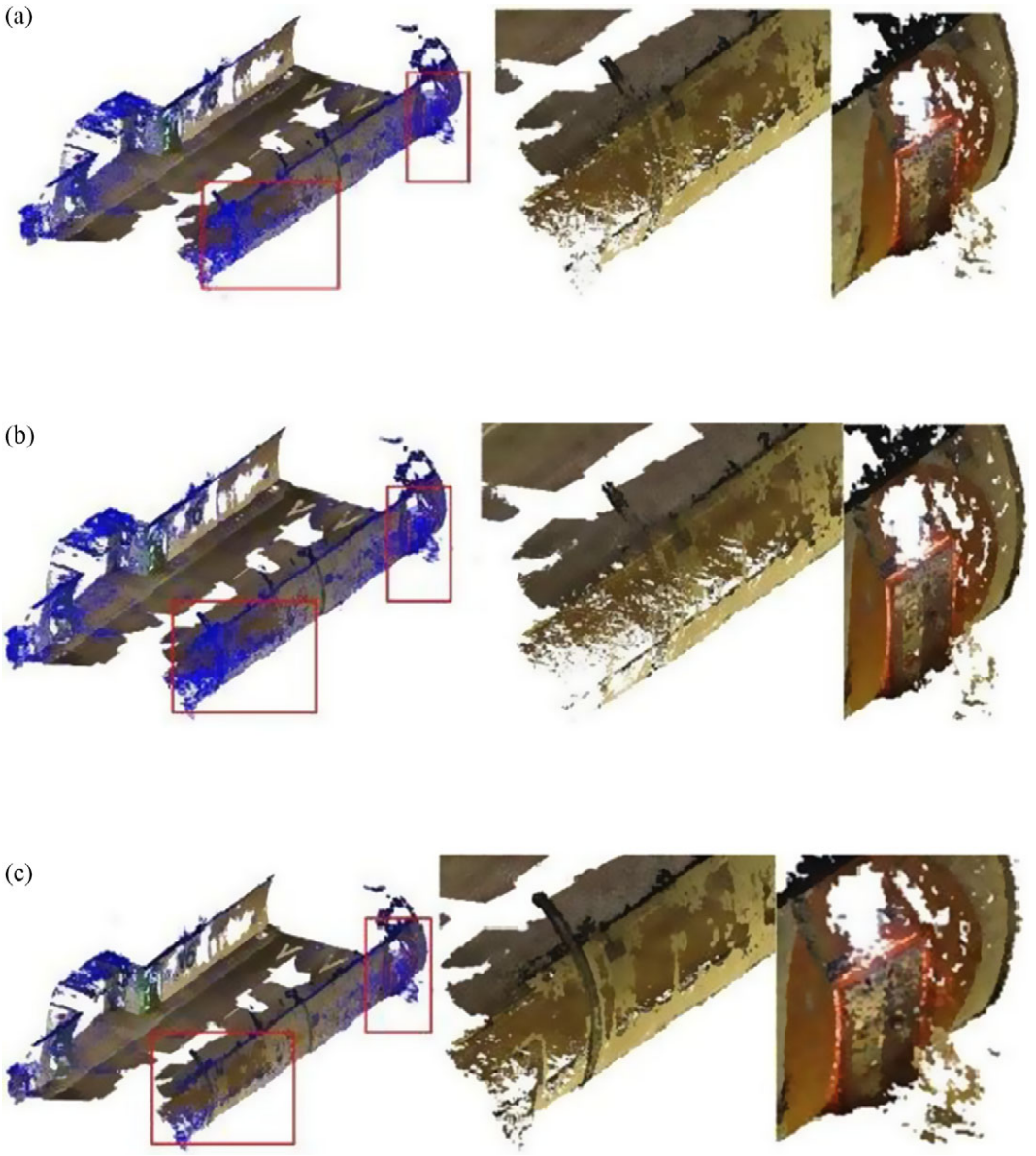


Figure 14. *Sensitivity analysis of SOR parameters defined in Eq. 1: (a)* $\mathrm{K} = 10$, $\sigma = 1$*: over-smoothing regime with critical detail loss on tunnel lining (b)* $\mathrm{K} = 5$, $\sigma = 1$*, over-smoothing regime (c)* $\mathrm{K} = 5$, $\sigma = 2$*: under-filtering regime with residual noise clusters near fire door joints.*

https://doi.org/10.1017/dce.2026.10042 Published online by Cambridge University Press

factor of deviation for outlier detection, respectively. The SOR method's reliance on neighborhood-based statistical thresholds (Eq.1) introduces edge-blurring artifacts, particularly in geometrically complex regions. As shown in Figure 14, under $K = 10$, $\sigma = 1$ and $K = 5$, $\sigma = 1$, tunnel lining textures are eroded despite achieving better noise removal effect behind fire door. This demonstrates a critical trade-off: larger K improves global denoising but sacrifices local geometric fidelity. This is because traditional methods treat all points in the data equally, without considering whether a point is part of a flat surface, edge, or corner. These traditional denoising methods often lose detail in sharp features like corners, while the proposed method preserves geometric details, making it ideal for precision-critical applications.

4.3 Quantitative tunnel geometry examinations

Figure 15 compares tunnel cross-sectional geometries before and after denoising at four inspection locations: (1) shown in Figure 15(a) and (2)–(4) in Figure 15(b). Table 3 quantifies the circle-fitting performance, deviations in center coordinates (maximum shift: 201 mm) and radii (mean difference: 169 mm) between denoised and raw point clouds. It reflects inherent uncertainties in tunnel scanning conditions. These magnitudes are attributable to dynamic acquisition artifacts and denoising algorithmic tolerance. Notably, these sub-200 mm discrepancies represent operational-level precision for infrastructure inspection, where critical alert thresholds are typically set at >500 mm (ASCE 95-2023). The

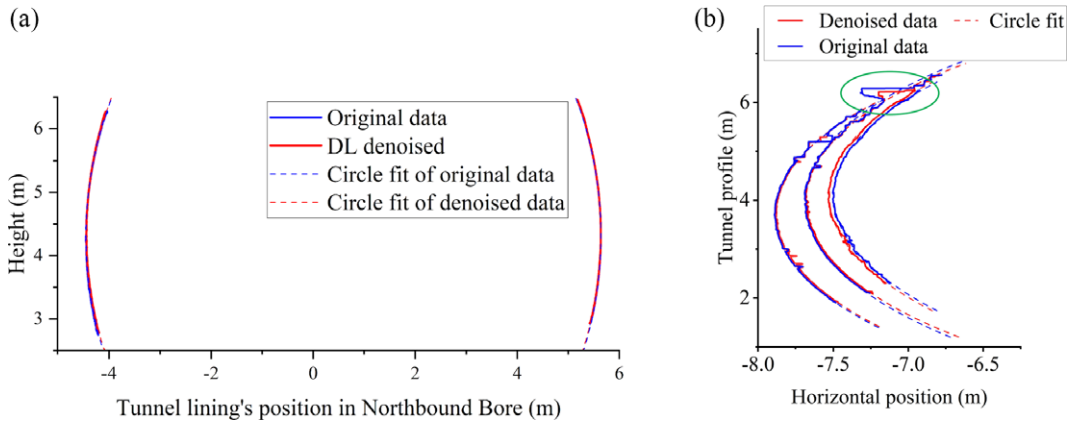


Figure 15. *Comparison of two fitted circles' shapes before and after DL denoising in Northbound Bore (NB) (a) Examination section (1) (b) Examination sections (2), (3), (4) from right to left.*

Table 3. *Circle fit details of two example cross sections before and after DL denoising in NB (/m)*

		x_c	y_c	r	Reduced Chi-Sqr	R^2
Bored tunnel	Original data	4.313	0.600	5.041	4 E−5	1
Section (1)	Denoised data	4.349	0.602	5.036	1 E−5	1
Laybys	Original data	4.103 ± 1 E−3	−3.068 ± 9 E−3	4.325 ± 9 E−3	7 E−5	0.99995
Section (2)	Denoised data	4.079 ± 1 E−3	−3.081 ± 8 E−3	4.315 ± 8 E−3	5 E−5	0.99996
Laybys	Original data	3.970 ± 5 E−3	−3.122 ± 3 E−2	4.445 ± 3 E−2	0.002	0.999
Section (3)	Denoised data	3.974 ± 2 E−3	−3.323 ± 1 E−2	4.276 ± 1 E−2	4 E−4	0.99977
Laybys	Original data	3.701 ± 2 E−3	−3.569 ± 2 E−2	4.217 ± 1 E−2	3 E−4	0.9998
Section (4)	Denoised data	3.724 ± 1 E−3	−3.538 ± 9 E−3	4.245 ± 9 E−3	1 E−4	0.99993

https://doi.org/10.1017/dce.2026.10042 Published online by Cambridge University Press

preserved correlation between denoised/raw datasets (Pearson's r = 1) confirms methodological consistency despite absolute value differences.

Denoising significantly improves fitting accuracy in high-noise regions (Sections 3) and 4). Particularly, it reducing local horizontal errors from 0.4 m to 0.2 m at most, while the radius narrows at most 17 cm on section (3) circled area in Figure14(b). Statistically, DL denoising reduces the reduced Chi-Square values by 74% in bored sections and at least 28% in layby sections with noise, with both models achieving coefficient of determination (R^2) values exceeding 0.99. These millimeter-scale changes in the 3.5 m × 0.5 m local inspection zone, demonstrate critical sensitivity improvements for deformation assessment. Future extensions could incorporate elliptical fittings to analyze localized deformation patterns, particularly in regions showing asymmetric profile shifts.

5. Testing on Wuxi dataset

5.1 Experimental setting

To evaluate the generalizability of the proposed workflow, an additional experiment was conducted utilizing TLS data from the Wuxi dataset (Lin et al., 2024). The tunnel comprises prefabricated concrete lining rings with a 2.75 m radius, horizontally truncated at the invert to interface with the road surface. Each segmented ring extends 1.5 m axially and incorporates engineered joints. The dataset also includes adjacent structures, including bolts, pipelines, and road surfaces. This high-fidelity data enables robust segmentation of individual lining segments, particularly for detecting service-stage deformations such as joint misalignments >2 mm or radial deviations exceeding 1.5% of the nominal radius.

The synthetic training dataset was generated with systematic noise simulation configured to approximate real-world TLS measurement artifacts. Lining noises was incorporated at a ratio of 5% (*ratio_noise_lin* = 0.05) along structural edges, while road surface noise was introduced at 1% density (*ratio_noise_pav* = 0.01). These noise elements were spatially distributed at a resolution of 0.05 m, with rectangular noise clusters randomly sampled within dimensional bounds of 0 × 0 m to 0.5 × 0.5 m to simulate irregular scanning noises. Cavity artifacts were explicitly excluded from the model on the smooth tunnel surfaces. For the RandLA-Net implementation, a 5-layer architecture was employed with progressive downsampling ratios of (Attard et al., 2018; Xu and Yang, 2020), corresponding to feature dimension expansions from 16 to 512 channels. Training utilized a batch size of 4 and an initial learning rate of 0.01, optimized through gradient descent with momentum stabilization to mitigate high-noise sensitivity.

Figure 16 illustrates a synthetic point cloud where the structural component of interest (the lining) is identified in red and all other points (noise plus pavement) are again identified in blue. A limitation of this synthetic dataset is the difficulty in incorporating details such as tracks, bolts and pipelines. Each training dataset comprises 204,800 points, following downsampling with a voxel size of 0.04 m. The model

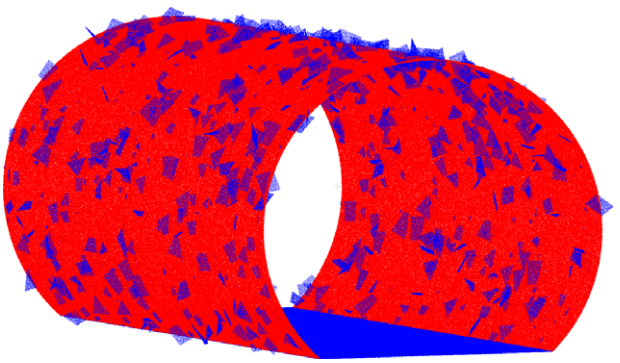


Figure 16. *An exemplar synthetic point cloud in the Wuxi dataset.*

https://doi.org/10.1017/dce.2026.10042 Published online by Cambridge University Press

validation results on the synthetic data achieve an OA of 98.88% and an mIoU of 97.18%, highlighting the method's exceptional accuracy and efficacy in segmenting the synthetic data.

5.2 DL model assessment

Figure 17 compares the DL denoising results (Figure 17(1)) to the manually segmented ground truth (Figure 17(2)). It can be seen the method demonstrates a high level of accuracy in removing non-lining structures, particularly bolts, pipelines, and tracks, which are not prioritized in the training dataset as primary structural details. However, certain discrepancies between the DL predictions and the ground truth are evident, particularly in the form of missegmented fragments within the road pavement. Furthermore, continuous pipelines near the surface at the tunnel crown and 2 m height were misclassified as tunnel linings. When applied to the real-world test set, the method achieves satisfactory results with an OA of 89.68% and an mIoU of 80.06%. While there is a slight decrease in performance when transitioning from synthetic to real-world data, these results still indicate robustness and an ability to generalize to different datasets well. Reasons for these differences in performance can be attributed to the expediency of the modeling process, which did not provide other data artifacts such as miscellaneous obstructions.

5.3 Model comparison

To validate the effectiveness of the RandLA-Net deep learning network adopted in this paper, representative approaches were used for comparisons, i.e., ensemble learning and Point Transformer.

Since there is no intensity or color information in the synthetic data, the input features of the base model consist of three-dimensional coordinates (X, Y, Z) and normal vectors (Nx, Ny, Nz). The normal vectors were calculated using 10 nearest points. The K-nearest neighbors (KNN), gradient boosting (GB), support vector machine (SVM), extra trees (ET), random forest (RF), and bootstrap aggregating (bagging) algorithms were chosen as the base models. The basemodels were trained independently using the training set. The hyperparameters in the basemodels were optimized by grid search. Then the outputs of basemodels were input into the metamodel for another round of training. The metamodel adopts the stacking strategy and uses logistic regression. The model training workflow was the same as that in the previous work (Peng et al., 2024). As a result, the OA and mIoU of the meta model are 84.43% and 74.96%, respectively. The results highlight the adaptability of the adopted deep learning algorithms for the specific task. The potential reason is that the synthetic data includes only basic geometric information, hindering the ensemble learning method from effectively learning the advanced features in the real-world point clouds.

Additional comparison was implemented by replacing the local feature aggregation module in RandLANet with Point Transformer (Zhao et al., 2021). All other experiment setups and hyperparameters were kept unchanged to ensure rigorous comparison. The OA and mIoU of the trained model using Point

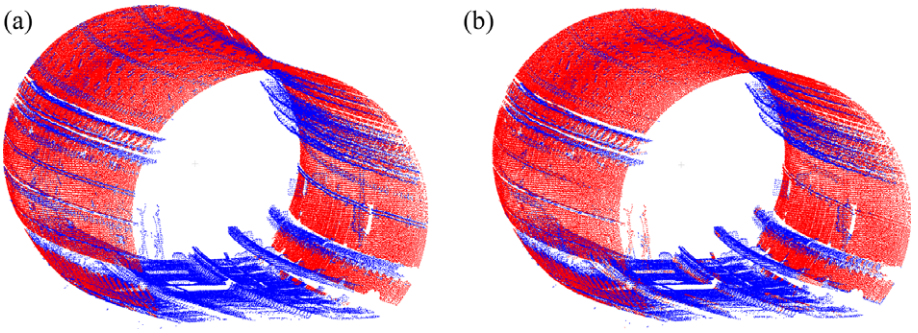


Figure 17. *Segmentation results and ground truth in the Wuxi Dataset (a) Segmented result using machine learning (b) Ground truth from manual annotation.*

https://doi.org/10.1017/dce.2026.10042 Published online by Cambridge University Press

Transformer are 88.51% and 77.65%, respectively, showing compromised performance compared to RandLA-Net. The results indicate that the feature aggregation module in RandLA-Net, designed with a self-attention mechanism, is more suitable for tunnel point cloud segmentation.

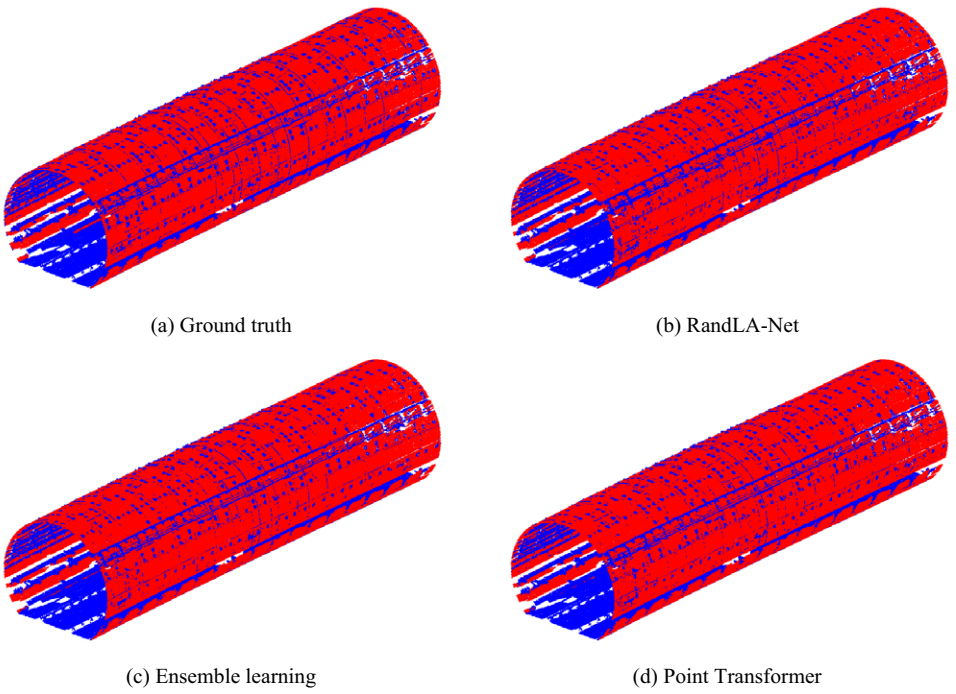
The segmentation performance of the compared approaches is listed in Table 4. It is proven that the synthetic data are effective for different approaches, with mIoU more than 74%, although there are slight differences in segmentation performance. Further, the segmentation results of the compared approaches are presented in Figure 18. Despite the numerical differences in the evaluation metrics, all three compared approaches yield satisfactory performance in recognizing structural points in the tunnel point clouds, demonstrating the feasibility of synthetic data.

6. Computational efficiency and accuracy analysis

This section evaluates the trade-off between computational cost and denoising accuracy for both the DPT and Wuxi datasets. For experienced annotators, manual denoising can take hours per 10 m per bore in DPT dataset, depending on the local noise distribution. In contrast, the automated denoising using the trained 3D DL models takes approximately 2 minutes when running on an NVIDIA 3060 Ti 16G GPU, showcasing significant improvements in efficiency and automation.

Table 4. *Segmentation performance of compared approaches*

Approach	OA (%)	mIoU (%)
RandLA-Net	89.68	80.06
Ensemble learning	86.43	74.96
Point transformer	88.51	77.65



(a) Ground truth (b) RandLA-Net

(c) Ensemble learning (d) Point Transformer

Figure 18. *Segmentation results of compared approaches.*

https://doi.org/10.1017/dce.2026.10042 Published online by Cambridge University Press

In terms of training efficiency, the DPT dataset required 16 hours to complete 100 epochs, while the Wuxi dataset needed 21 hours under similar hardware configuration. This difference can be attributed to the varying complexity and size of the two datasets. Despite these variations, the proposed method maintained a consistent balance between computational efficiency and denoising performance across both datasets.

It is important to note that the proposed method has been applied to subsampled point clouds to optimize computational efficiency. Specifically, only 1% of the original DPT point cloud, equivalent to 0.1 million points, was used for training, validation, and testing. While subsampled point clouds proved satisfactory for the present analyses, training on larger datasets may require significantly greater memory and computational resources. This suggests that while the proposed method is highly efficient for moderate-scale datasets, further optimization may be necessary to handle extremely large-scale point clouds, such as those in the Wuxi dataset, without incurring prohibitive training costs.

Regarding denoising quality, manual denoising often struggles to achieve consistent and satisfactory results due to its subjective nature and lack of reproducibility. For instance, two manually denoised outputs may exhibit no commonality in the eliminated noise clusters. On the other hand, once a DL denoising model is trained and its output is saved, it can be reused repeatedly for extended inspection distances, making the DL method effective for large-scale applications. For the DPT dataset, the proposed model achieved a 95.2% mIoU. Similarly, for the Wuxi dataset, the model demonstrated a comparable accuracy of 80.06% mIoU.

7. Conclusion

UAV image point cloud data provide valuable information for assessing infrastructure health. Although it offers lower accuracy compared to LiDAR point clouds, UAV data can be highly advantageous for specific inspection conditions, such as inaccessible or hazardous areas, surface texture and color analysis, and cost-effective infrastructure monitoring. However, its utility is often compromised by noise and spurious data artifacts. To efficiently and effectively denoise infrastructure point clouds, this paper has described a DL segmentation approach involving the use of synthetic training datasets to identify unwanted noise. The proposed method was tested on two real-world tunnel datasets. It should be emphasized that improved photogrammetric data acquisition (e.g., optimized lighting, higher-resolution imaging) remains an important avenue for reducing artifacts at the source. Our work is therefore complementary: by addressing residual noise that persists under realistic, often sub-optimal field conditions, the proposed denoising method enhances the reliability of UAV-based tunnel inspections when ideal data quality cannot be guaranteed. The main findings of this study are as follows:

(1) The use of synthetic data has proven highly beneficial for preparing training datasets, particularly in challenging inspection environments. This approach enhances model robustness and enables more flexible training conditions. The proposed workflow eliminates the need for data annotation, reducing implementation time and manual effort, thereby increasing its practicality for real-world applications.
(2) Noise in UAV image point clouds was primarily categorized into edge artifacts and spurious data. The deep learning-based segmentation was shown to effectively identify and remove the majority of such noise, achieving a high mIoU score of approximately 95. On road pavements, the spatial range of noise was reduced from 5 m to 0.2 m, while on tunnel linings, it decreased from 0.5 m to 0.2 m. Large ‘bulk’ noise was reduced by 84%, with 7232 out of 8560 noise points successfully removed.
(3) Compared to manually denoised DSMs, the proposed denoising method delivered superior results in deformation extraction. In highly noisy areas, the method significantly improved tunnel surface fitting. The radius of fitting circles was shown to vary on a scale of 5 mm to 170 mm, with model precision improving by between 11% and 55%, and the reduced chi-square value improving by

https://doi.org/10.1017/dce.2026.10042 Published online by Cambridge University Press

between 24% and 73%. Additionally, the proposed method outperforms SOR filters in preserving corner points of tunnel structures and demonstrates less dependence on data quality.

(4) The proposed deep learning denoising method exhibited robust performance across various point cloud datasets, including DPT and the Wuxi dataset, achieving an 80% mIoU score and effectively identifying key features in both synthetic and real-world data. The method adapts well to datasets with different noise characteristics, with minor performance variations attributed to missing elements, such as pedestrians, in the training data. Future enhancements will prioritize expanding the diversity of training datasets to further improve accuracy.

(5) The main limitation of this study lies in its reliance on a single noise type and distribution. Noise in this study was generated randomly along surfaces, whereas in real-world scenarios, noise is often concentrated around cavities, structural edges, and fire doors. Incorporating these specific noise types into synthetic datasets could significantly improve the model's performance on noise clutter. Future work will focus on diversifying noise types and distribution rules in synthetic data generation, ensuring greater applicability to tunnel inspection conditions.

This study highlights the potential of deep learning techniques to enhance the quality and utility of UAV image point cloud data. By effectively reducing noise and improving accuracy, these methods can contribute significantly to reliable and efficient infrastructure health assessments. Beyond noise suppression, a key advantage of the proposed segmentation approach is its ability to isolate tunnel linings from non-structural elements such as bolts, access gates, and pipelines. This separation is essential for precise structural analysis, as it ensures that deformation measurements are derived solely from the true lining geometry rather than being distorted by ancillary features. This capability represents a substantial improvement over traditional denoising filters and underscores the broader value of deep learning segmentation for tunnel inspection. Future research will refine these techniques, focusing on understanding the distribution likelihood of data quality issues, such as cavities and noise, and improving synthetic data generation processes.

Data availability statement. Data can be found at https://zenodo.org/records/18471574

Acknowledgments. The authors would like to thank Dr Enok Cheon and Dr Dieter Issler for their useful assistance with coding suggestions and computational platforms.

Author contribution. Conceptualization: B.S., R.Z., and W.L. Methodology: W.L. and R.Z. Data curation: R.Z. Data visualization: R.Z. Writing original draft: R.Z. Writing review & editing: B.S., C.W., and Z.L. Supervision: Z.L. Funding acquisition: Z.L. All authors approved the final submitted draft.

Funding statement. This research was supported by grants from the China Scholarship Council; The authors acknowledge the support from Science Foundation Ireland for a research grant under the Frontiers for the Future Programme [grant number: 21/FFP-P/10090]. Funding was partly provided by the basic funding to NGI from The Research Council of Norway.

Competing interests. None.

Ethical standard. The research meets all ethical guidelines, including adherence to the legal requirements of the study country.

References

Attard L, **Debono CJ**, **Valentino G and Di Castro M** (2018) Tunnel inspection using photogrammetric techniques and image processing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing 144*, 180–188. https://doi.org/10.1016/j.isprsjprs.2018.07.010.

Attard L, **Debono CJ**, **Valentino G and Di Castro M** (2018) Vision-based change detection for inspection of tunnel liners. *Automation in Construction 91*, 142–154. https://doi.org/10.1016/j.autcon.2018.03.020.

Balta H, **Velagic J**, **Bosschaerts W**, **De Cubber G and Siciliano B** (2018) Fast statistical outlier removal based method for large 3D point clouds of outdoor environments. *IFAC-Papers 51*(22), 348–353. https://doi.org/10.1016/j.ifacol.2018.11.566.

Chao W (2023) Monitoring and Modelling the Long-term Performance of Dublin Port Tunnel. University College Cork, Cork. Available at https://cora.ucc.ie/items/0e823787-d639-42b6-a4b7-3e3612d4d7b8 (accessed 16 July 2024)

Chen S, **Laefer DF**, **Mangina E**, **Zolanvari SMI and Byrne J** (2019) UAV bridge inspection through evaluated 3D reconstructions. *Journal of Bridge Engineering 24*(4), 05019001. https://doi.org/10.1061/(ASCE)BE.1943-5592.0001343.

Gaspari F, **Ioli F**, **Barbieri F**, **Belcore E and Pinto L** (2022) Integration of UAV-Lidar and UAV-photogrammetry for infrastructure monitoring and bridge assessment. https://doi.org/10.5194/isprs-archives-XLIII-B2-2022-995-2022. Available at https://re.public.polimi.it/handle/11311/1217562 (accessed 16 July2024)

Gevaert CM, **Persello C**, **Sliuzas R and Vosselman G** (2017) Informal settlement classification using point-cloud and image-based features from UAV data. *ISPRS Journal of Photogrammetry and Remote Sensing 125*, 225–236. https://doi.org/10.1016/j.isprsjprs.2017.01.017.

Greenwood WW, **Lynch JP and Zekkos D** (2019) Applications of UAVs in civil infrastructure. *Journal of Infrastructure Systems 25*(2), 04019002. https://doi.org/10.1061/(ASCE)IS.1943-555X.0000464.

Hu Q, **Yang B**, **Xie L**, **Rosa S**, **Guo Y**, **Wang Z**, **Trigoni N and Markham A** (2022) Learning semantic segmentation of large-scale point clouds with random sampling. *IEEE Transactions on Pattern Analysis and Machine Intelligence 44*(11), 8338–8354. https://doi.org/10.1109/TPAMI.2021.3083288.

Keizer R, **Dubay R**, **Waugh L and Bradley C** (2022) Robotic camera positioning system for inspection of tunnels: a conceptual design. In *2022 IEEE International Systems Conference (SysCon)*, pp. 1–8. https://doi.org/10.1109/SysCon53536.2022.9773797. Montreal, Canada: IEEE. Available at https://ieeexplore.ieee.org/document/9773797 (accessed 15 January 2024)

Kolodiazhnyi M, **Vorontsova A**, **Konushin A and Rukhovich D** (2024) *Top-down beats bottom-up in 3D instance segmentation. In Presented at the Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pp. 3566–3574. Available at https://openaccess.thecvf.com/content/WACV2024/html/Kolodiazhnyi_Top-Down_Beats_Bottom-Up_in_3D_Instance_Segmentation_WACV_2024_paper.html (accessed 12 August 2024)

Lattanzi D and Miller G (2017) Review of robotic infrastructure inspection systems. *Journal of Infrastructure Systems 23*(3), 04017004. https://doi.org/10.1061/(ASCE)IS.1943-555X.0000353.

Li D, **Xie, Q**, **Gong, X**, **Yu, Z**, **Xu, J**, **Sun, Y**, **Wang, J** (2021) Automatic defect detection of metro tunnel surfaces using a vision-based inspection system. *Advanced Engineering Informatics 47*, 101206. https://doi.org/10.1016/j.aei.2020.101206.

Li Z, **Soga K and Wright P** (2015) Long-term performance of cast-iron tunnel cross passage in London clay. *Tunnelling and Underground Space Technology 50*, 152–170. https://doi.org/10.1016/j.tust.2015.07.005.

Lin W, **Sheil B**, **Zhang P**, **Zhou B**, **Wang C and Xie X** (2024) Seg2Tunnel: A hierarchical point cloud dataset and benchmarks for segmentation of segmental tunnel linings. *Tunnelling and Underground Space Technology 147*, 105735. https://doi.org/10.1016/j.tust.2024.105735.

Ortiz-Coder P and Sánchez-Ríos A (2020) An integrated solution for 3D heritage modeling based on videogrammetry and V-SLAM technology. *Remote Sensing 12*, 1529. https://doi.org/10.3390/rs12091529.

Peng X, **Lin P**, **Xia Q**, **Yu L and Wang M** (2024) A new method for recognizing discontinuities from 3D point clouds in tunnel construction environments. *Tunnelling and Underground Space Technology 152*, 105955. https://doi.org/10.1016/j.tust.2024.105955.

Peterson S, **Lopez J and Munjy R** (2019) Comparison of UAV imagenary-derived point cloud to terrestrial laser scanner point cloud. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences IV-2-W5*, 149–155. https://doi.org/10.5194/isprs-annals-IV-2-W5-149-2019.

Pietersen RA (2022) *Automated method for airfield pavement condition index determination*. Thesis, Massachusetts Institute of Technology. Available at https://dspace.mit.edu/handle/1721.1/144963 (accessed 2 September 2024)

Rainio O, **Teuho J and Klén R** (2024) Evaluation metrics and statistical tests for machine learning. *Scientific Reports 14*(1), 6086. https://doi.org/10.1038/s41598-024-56706-x.

Rhee S and Kim T (2016) Dense 3D point cloud generation from UAV images from image matching and global optimization. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XLI-B1*, 1005–1009. https://doi.org/10.5194/isprs-archives-XLI-B1-1005-2016.

Sapronova A, **Unterlass PJ**, **Dickmann T**, **Hecht-Méndez J and Marcher T** (2023) Prediction of geological conditions ahead of the tunnel face: Comparing the accuracy of machine learning models trained on real and synthetic data. In Correia AG, Azenha M, Cruz PJS, Novais P and Pereira P (eds), *Trends on Construction in the Digital Era*. Cham: Springer International Publishing, pp. 76–86. https://doi.org/10.1007/978-3-031-20241-4_6.

Schult J, **Engelmann F**, **Hermans A**, **Litany O**, **Tang S and Leibe B** (2023) *Mask3D: mask transformer for 3D semantic instance segmentation. In 2023 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 8216–8223. https://doi.org/10.1109/ICRA48891.2023.10160590. Available at https://ieeexplore.ieee.org/abstract/document/10160590. [Accessed Aug. 12, 2024]

Sun S, **Chen C**, **Wang Z**, **Zhou J**, **Li L**, **Yang B**, **Cong Y**, **Wang H** (2023) *Real-time UAV 3D image point clouds mapping*. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences X-1-W1-2023*, 1097–1104, https://doi.org/10.5194/isprs-annals-X-1-W1-2023-1097-2023

Vetrivel A, **Gerke M**, **Kerle N and Vosselman G** (2015) Segmentation of UAV-based images incorporating 3D point cloud information. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XL-3/W2*, 261–268. https://doi.org/10.5194/isprsarchives-XL-3-W2-261-2015.

Wang C, **Friedman M and Li Z** (2023) 'Monitoring and assessment of a cross-passage twin tunnel long-term performance using wireless sensor network. *Canadian Geotechnical Journal*. https://doi.org/10.1139/cgj-2022-0224. Available at https://cdnsciencepub.com/doi/abs/10.1139/cgj-2022-0224 (accessed 29 May 2023)

https://doi.org/10.1017/dce.2026.10042 Published online by Cambridge University Press

Wolff K *et al.* (2016) Point cloud noise and outlier removal for image-based 3D reconstruction. In *2016 Fourth International Conference on 3D Vision (3DV)*. Stanford, CA: IEEE, pp. 118–127. https://doi.org/10.1109/3DV.2016.20. Available at http://ieeexplore.ieee.org/document/7785084/ (accessed 31 October 2024)

Xu X and Yang H (2020) Vision measurement of tunnel structures with robust modelling and deep learning algorithms. *Sensors 20*(17), 4945. https://doi.org/10.3390/s20174945.

Yang H and Xu X (2021) Structure monitoring and deformation analysis of tunnel structure. *Composite Structures 276*, 114565. https://doi.org/10.1016/j.compstruct.2021.114565.

Zeybek M (2021) Inlier point preservation in outlier points removed from the ALS point cloud. *Journal of the Indian Society of Remote Sensing 49*(10), 2347–2363. https://doi.org/10.1007/s12524-021-01397-4.

Zeybek M and Şanlıoğlu İ (2019) Point cloud filtering on UAV based point cloud. *Measurement 133*, 99–111. https://doi.org/10.1016/j.measurement.2018.10.013.

Zhang J, **Wu J and Wang H** (2023) Quality analysis and improvement of 3D panoramic stitching sequence image. In *2023 3rd International Conference on Consumer Electronics and Computer Engineering (ICCECE)*, pp. 445–449. Guangzhou, China: IEEE. https://doi.org/10.1109/ICCECE58074.2023.10135274. Available at https://ieeexplore.ieee.org/abstract/document/10135274 (accessed 30 October 2024)

Zhang R, **Hao G**, **Zhang K and Li Z** (2023) Unmanned aerial vehicle navigation in underground structure inspection: A review. *Geological Journal 58*(6), 2454–2472. https://doi.org/10.1002/gj.4763.

Zhang R, **Hao G**, **Zhang K and Li Z** (2024) Reactive UAV-based automatic tunnel surface defect inspection with a field test. *Automation in Construction 163*, 105424. https://doi.org/10.1016/j.autcon.2024.105424.

Zhang R, **Wang C and Li Z** (2024) Automatic Inspection and Assessment of a Cross-Passage Twin Tunnel Using UAV. *Automation in Construction 175*, 106200. https://doi.org/10.1016/j.autcon.2025.106200

Zhao H, **Jiang L**, **Jia J**, **Torr P and Koltun V** (2021) Point transformer. In *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*. Montreal, Canada: IEEE.

Zhou L, **Sun G**, **Li Y**, **Li W and Su Z** (2022) Point cloud denoising review: From classical to deep learning-based approaches. *Graphical Models 121*, 101140. https://doi.org/10.1016/j.gmod.2022.101140.

Cite this article: Zhang R, Lin W, Wang C, Sheil B, Liu Z and Li Z (2026). Denoising image point clouds using segmentation and synthetic data for enhanced structural health analysis of tunnels. *Data-Centric Engineering*, 7, e10. doi:10.1017/dce.2026.10042

https://doi.org/10.1017/dce.2026.10042 Published online by Cambridge University Press