Experimental study on intelligent detection for surface defects on earth-rockfill dams based on UAV images and TLS point clouds

Hongyu Ren [a] , Rui Pang [a,b,*] , Weijie Huang [a], Bin Xu [a,b]

[a] School of Infrastructure Engineering, Dalian University of Technology, Dalian 116024, China
[b] State Key Laboratory of Coastal and Offshore Engineering, Dalian University of Technology, Dalian 116024, China

A B S T R A C T

This study proposes an intelligent detection solution for surface defects on earth-rockfill dams, including unmanned aerial vehicle (UAV), terrestrial laser scanning (TLS) and Deep Convolutional Generative Adversarial Network (DCGAN)-based defect intelligent detection method. To this end, an earth-rockfill dam simulation platform is constructed to explore the defect detection performance of UAV and TLS on earth-rockfill dam, and the impacts of various factors are evaluated by changing the ground control point arrangement or weather scene. Additionally, a DCGAN-based method for generating defect images is proposed for data augmentation of limited on-site images, which provides a new approach for the pain point of over-reliance on datasets for object detection in practical engineering. By learning from small batches of defect images, plenty of similar images are generated. The results indicate that integrating UAV and TLS enables high-precision and rapid detection of surface defects on earth-rockfill dams. The defect images generation method effectively improves the detection accuracy of object detection model under limited samples.

1. Introduction

Earth-rockfill dams, due to the long-term coupling effects of various factors such as earthquake, rainfall, hydrostatic pressure, temperature and material aging, commonly exhibit defects such as crack, spalling, collapse, slippage, and precipitate [1,2], as illustrated in Fig. 1. This will inevitably result in a compromised structural integrity and pose a threat to the safe operation and maintenance of earth-rockfill dams [3–5]. Currently, periodic manual inspections are commonly used to detect and assess surface defects on earth-rockfill dams. However, manual inspections are characterized by low accuracy, low efficiency, and high costs, and are inconvenient for complex structures such as spillways [6]. To enhance the accuracy and automation of defect detection, researchers have employed contact methods such as ground-penetrating radar [7,8] and distributed optical fibers [9] to assess the safety and structural integrity of dam structures. However, given the large scope of earth-rockfill dams and the high randomness and complexity of panel and soil damage, contact detection methods face challenges in detecting earth-rockfill dams rapidly and comprehensively. Consequently, researchers move their focus to non-contact monitoring methods.

Non-contact monitoring based on multi-source sensing involves the effective integration of unmanned aerial vehicle (UAV) photogrammetry, three-dimensional (3D) laser scanning, ground-based synthetic aperture radar (SAR), InSAR, etc. By utilizing fusion data analysis, efficient and comprehensive health monitoring of large-scale structures is enabled, effectively addressing the limitations of individual monitoring techniques. Among them, UAV photogrammetry and 3D laser scanning based on computer vision are widely considered to be one of the most effective technologies for structural health monitoring. With continuous advancements in sensor technology and compitility, these methods have significant potential for both research and practical applications [11–14].

UAV photogrammetry is highly efficient, flexible and convenient, enabling rapid acquisition of high-resolution images from multiple angles, providing an intuitive representation of the actual features [15–18]. 3D laser scanning, including terrestrial laser scanning (TLS), airborne LiDAR, SLAM, is characterized by large coverage, rapid data acquisition, and robust anti-interference capabilities [19,20]. It can directly capture the 3D coordinates of targets, providing a visual representation of their spatial geometric features [21,22]. Based on these two techniques researchers have carried out related research and practical applications. Xu et al. [23] proposed an integrated space-air-
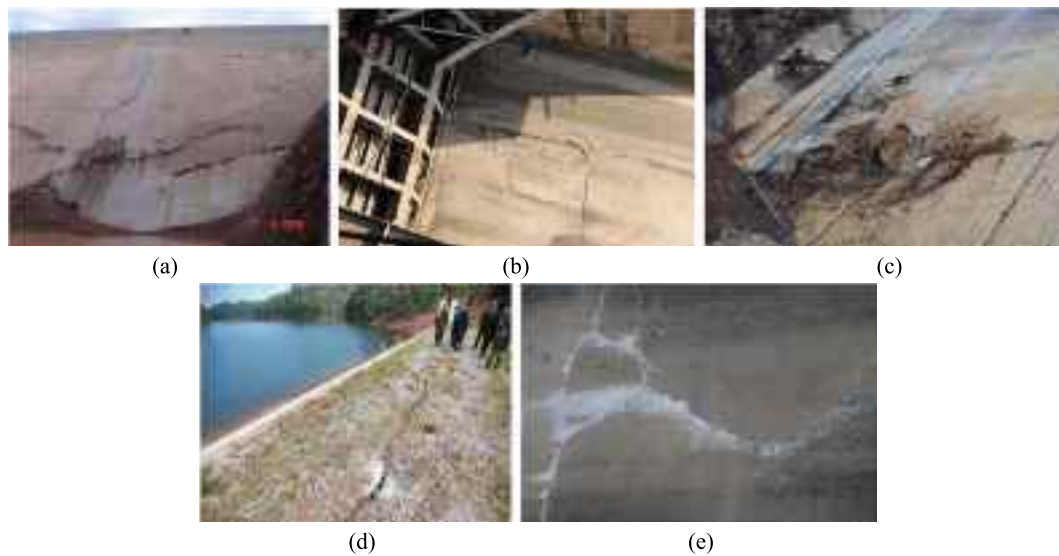

Fig. 1. Surface defects on dams: (a) crack; (b) spalling [10]; (c) collapse; (d) slippage; (e) precipitate [10].

ground investigation system, using UAV and airborne LiDAR to conduct detailed detection of high-risk sections of geological hazards and major geological hazard sites. Li et al. [24] quantitatively evaluated the rockfall hazard of high rocky slopes after earthquake based on TLS. Bolourian et al. [25] proposed a path planning method for bridge detection by using optical detection and airborne LiDAR, which can provide potential locations of surface defects to efficiently obtain sampling points in vertical and overlapping views. In the field of dam defects detection, Zhao et al. [10] proposed an emergency detection model for dams based on UAV images, aimed at obtaining 3D dam models with realistic texture and dimensional information. These results indicate that UAV photogrammetry and 3D laser scanning facilitate efficient and cost-effective health monitoring of large-scale structures. The fusion analysis of the two helps to analyze and assess defects more accurately. However, earth-rockfill dams exhibit unique defect characteristics, such as crack, spalling, collapse, slippage, and precipitate. Differences in defect characteristics make defect detection and analysis methods for other large structures may not be applicable to earth-rockfill dams. Moreover, research on UAV-based 3D reconstruction and TLS point cloud detection of earth-rockfill dams is not comprehensive enough. It is not clear how effective it is in detecting various types of defects and how subjective and objective factors, such as ground control points (GCPs) and weather, affect the detection performance. TLS point clouds and UAV images have some complementarities in technology. Through data fusion, the advantages of UAV images and TLS point clouds can be fully combined to make up for the limitations of a single data source. However, there are few studies on multi-source fusion modeling of earth-rockfill dams. Therefore, it is necessary to conduct model tests to investigate the effect of TLS detection and multi-source fusion modeling on the defect detection of earth-rockfill dams, as well as the influence of each factor on the detection performance.

Furthermore, in the above studies on UAV and 3D laser scanning, the detection of structural defects is mainly based on manual analysis. Manual analysis has the disadvantage of being inefficient and prone to fatigue when the workload is large. Object detection algorithms based on deep learning can address this issue, and the detected defects can be mapped onto a 3D model [26]. The integration of these two technologies facilitates the visualization of structural defects and provide technical support for research related to digital twins.

Object detection is an advanced research direction in the field of deep learning vision. Among them, the YOLO model, as a relatively mature object detection model, is widely used in industry and engineering for defect detection [27,28]. The basic process of the deep learning-based dam defect detection involves training an object detection model with dam defect images and then deploying the model on a vision machine such as UAVs to achieve intelligent autonomous detection. This method has enhanced the intelligence level of defect detection [26,29]. But due to the weak generalization ability and robustness of the object detection models, models trained on datasets from other dams often fail to meet the desired detection performance when applied to a specific dam, and the corresponding dataset must be built according to the actual project to reinforcement training of the model. Nonetheless, when building dam-specific datasets, there are often insufficient training samples due to the limited on-site images. Thus, a data augmentation method for dam defects is necessary to reduce the dependance on on-site data.

This study aims to explore an intelligent detection solution based on UAV and TLS for surface defect detection of earth-rockfill dams. The framework of the proposed intelligent detection solution is shown in Fig. 2. First, the experimental platform, devices, and the experimental procedures are introduced. Next, to evaluate the effectiveness of UAV and TLS in detecting surface defects on earth-rockfill dams, the proposed intelligent detection solution is applied on the earth-rockfill dam model.
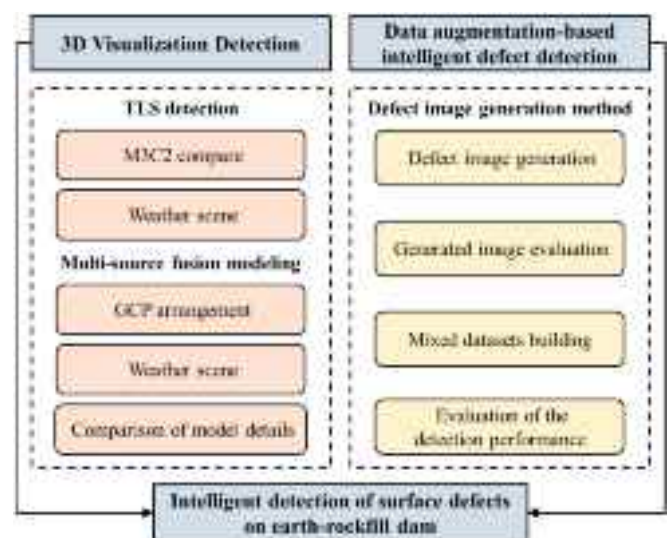

Fig. 2. Framework of the proposed intelligent detection solution for surface defects on earth-rockfill dam.
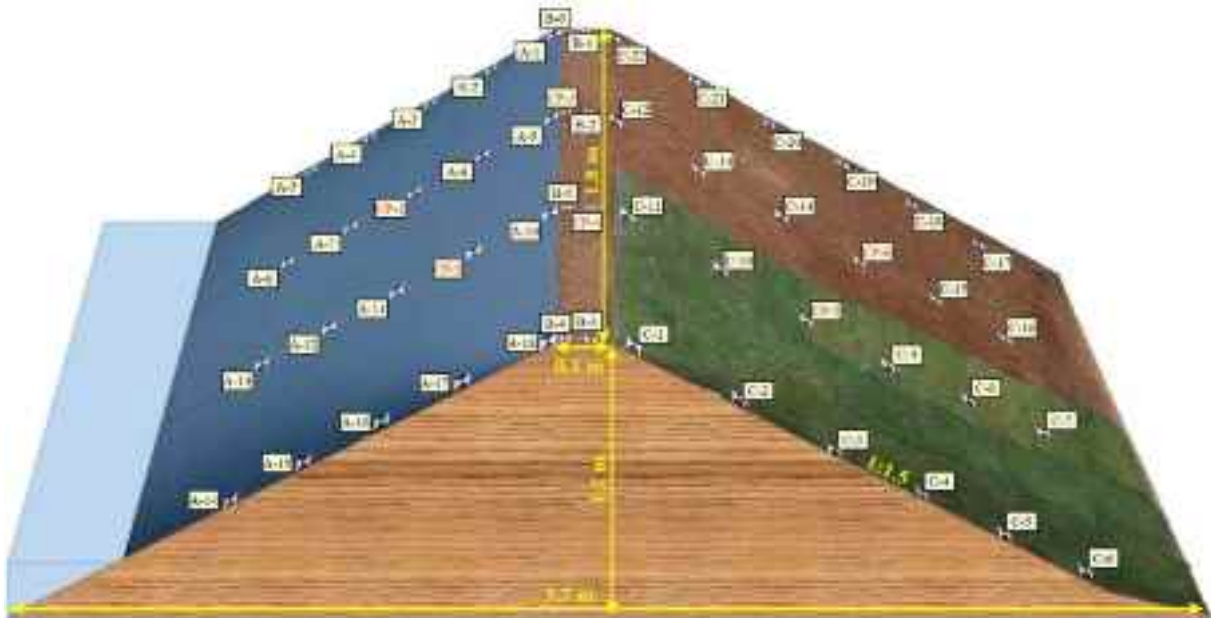
Fig. 3. View of the earth-rockfill dam model.

Table 1
Basic physical parameters of clay.

Dry density (g/ cm^3)	Liquid limit (W_L)	Plastic limit (W_P)	Liquidity index (I_L)	Natural moisture content (ω)
1.8	36.9	21.2	0.27	25.4 %

The detection performance of TLS point clouds and multi-source fusion model are investigated, and the impacts of various factors are evaluated by changing the GCP arrangements or weather scenes. In addition, to overcome the problem of limited training samples for object detection in practical engineering, a Deep Convolutional Generative Adversarial Network (DCGAN)-based method for generating defect images of earth-rockfill dam is proposed for data augmentation of limited on-site images, thereby enhancing the defects detection performance of the object detection model.

2. Experimental design

2.1. Surface defects simulation platform for earth-rockfill dams

In a steel model box with dimensions of 5.1 m × 1.8 m × 1.2 m (length × width × height), an earth-rockfill dam model was constructed with a dam axis length of 1.8 m, a height of 1.2 m, a dam crest width of 0.1 m, and a slope ratio of 1:1.5, as shown in Fig. 3. The main part of the model was composed of clay from Benxi, Liaoning, China, with its physical parameters detailed in Table 1. 6 PVC panels (30 mm thick) were placed on the upstream surface to simulate concrete faces. Half of the downstream surface was bare, while the other half was covered with grass, representing two cases: a vegetated dam slope and a bare dam slope.

52 black-and-white tags (3 cm × 3 cm) were marked on the model surface as GCPs and check points (CPs), 20 of which were marked on the upstream, 8 on the crest, and 24 on the downstream.

To simulate defects on the surface of earth-rockfill dams, after completing data acquisition for the intact dam, various types of defects were artificially created on both the upstream and downstream. The upstream defects included cracks, collapse, spalling, and precipitate, with specific arrangements and sizes shown in Fig. 4a. Besides, different

depths were set for the three types of defects: cracks (0.5 cm, 1.0 cm, 1.5 cm, 2.0 cm), collapse (0.5 cm, 1.0 cm, 2.0 cm), and spalling (0.5 cm, 1.0 cm, 1.5 cm). The downstream defects included slippage, cracks, and collapse, which were placed on both vegetated dam slop and bare dam slope, with specific arrangements shown in Fig. 4b.

To simulate rainfall, an artificial rainfall device was set on the model box, which consisted of 4 spray nozzles, 2 rows of brackets, water pumps, spray nozzles spaced 0.6 m apart along the direction of the dam axis, and brackets spaced 1.6 m apart. Among them, the spray nozzles had a hole diameter of 0.1 mm, a spray diameter of 1 m, and a flow rate of 25 ml/ min.

2.2. Experimental devices

DJI Phantom 4 RTK UAV: Used to acquire images of the earth-rockfill dam, the UAV has a maximum endurance of 30 min, a 1-inch CMOS, an image resolution of 5472 × 3648, a lens with a focal length of 8.8 mm, and is equipped with a Real-time kinematic (RTK) module with 1 cm accuracy (horizontal), 1.5 cm accuracy (horizontal), 1.5 cm accuracy (vertical) (Fig. 5a).

EasyScan T10 TLS: Used to acquire true color point clouds with a ranging accuracy of 1 cm@10 m, a laser wavelength of 905 nm, a spot frequency of 320,000 pts/s, and a camera resolution of 6080 × 3040 (Fig. 5b).

Qianxun SRmini RTK: Used with TLS and Total Station to acquire global coordinates of point clouds, GCPs and CPs with 8 mm accuracy (horizontal), 15 mm accuracy (vertical) (Fig. 5c).

Trimble S5 Total Station: Used to measure 3D coordinates of GCPs and CPs with an accuracy of 2 mm and a measuring range of 1300 m (Fig. 5d).

Sanliang PP720 lux meter: Used for monitoring lux values during data acquisition by UAV and TLS, with a range of 0 to 200,000 Lux (Fig. 5e).

2.3. Experimental procedures

It is well known that the performance of optical devices is heavily dependent on ambient light and rainfall can affect the reflectivity of the dam surface. Therefore, in this study, light and rainfall are considered as variables to investigate the influence of rainfall on TLS point cloud
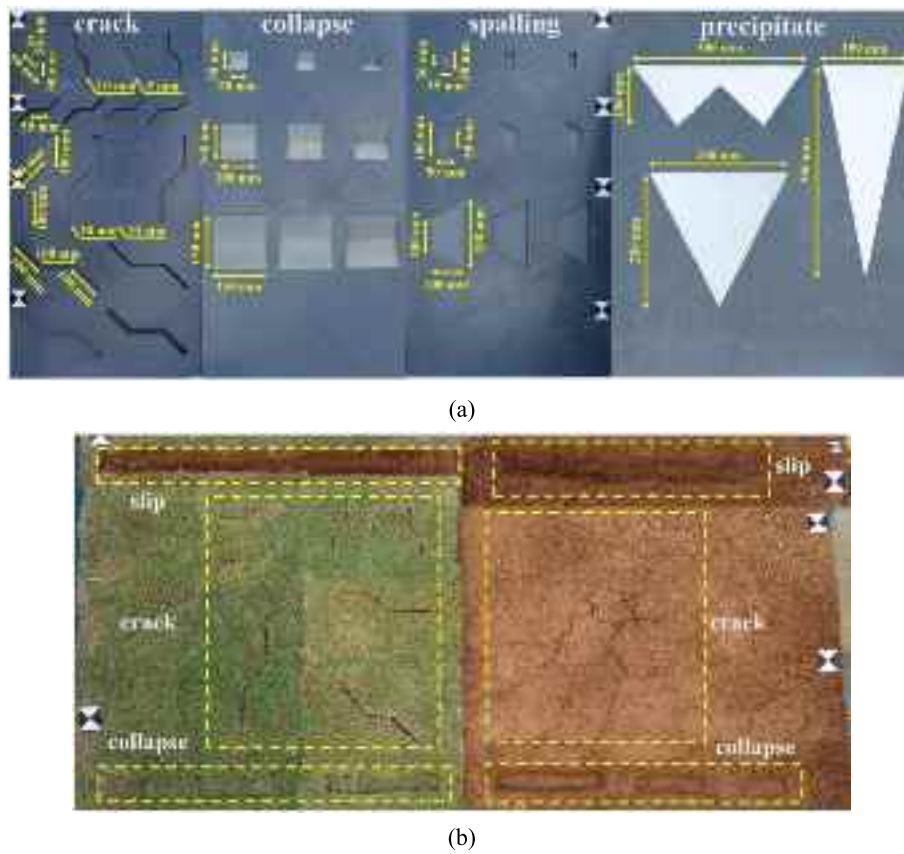
(a)

(b)

Fig. 4. Distribution of dam surface defects: (a) defects on upstream dam; (b) defects on downstream dam.

detection, as well as the effects of light and rainfall on the accuracy of 3D modeling of earth-rockfill dams. Three weather scenes are set: daytime & cloudy, sunset & cloudy, and daytime & rainy. Data acquisition is conducted on both the intact dam model and the defective dam model under these scenes. The data acquisition scheme is presented in Table 2. The illumination is monitored by an illuminometer and divided into regular illumination (7500–15000 Lux) and weak illumination (400–800 Lux) based on the Lux values, which ensures that the illumination is within these two intervals when acquiring UAV images. Rainfall is controlled by the rainfall device, which is turned on for 15 min before the rain test, and data acquisition starts immediately after the rain.

In this study, a combination of orthophoto and oblique photography is used to acquire the surface images of the earth-rockfill dam model. The UAV flight path is shown in Fig. 6, with the flight height operated at 1.5 m perpendicular to the dam, ensuring an overlap rate of over 85 %. The overall point cloud is generated by merging the two-site TLS point clouds of the model.

The data acquisition process is as follows: Total station and RTK measurements—15 min of artificial rainfall (rainfall scene only)—TLS scanning—UAV photogrammetry.

3. Multi-source fusion detection of surface defects on earth-rockfill dams

3.1. Detection of surface defects on earth-rockfill dams based on TLS point cloud

TLS can efficiently and accurately acquire 3D point cloud data of earth-rockfill dams in almost all weather scenes [30–32]. In order to verify the feasibility and accuracy of TLS point cloud for surface defect detection of earth-rockfill dams, the TLS point clouds under different weather scenes are compared using the M3C2 algorithm, and the effectiveness of defect detection based on the TLS point cloud and the influence of rainfall on detection performance are explored. The details are presented in the following sections.

3.1.1. Comparison of point clouds

The multiscale model-to-model cloud comparison (M3C2) algorithm provided by CloudCompare v2.12.4 software was used to analyze the TLS point cloud data.

As a robust algorithm for directly calculating the distance between two point clouds, the M3C2 algorithm can directly detect the changes of complex terrain on the point cloud without mesh delineation, and is less affected by spatial point density, surface roughness, and different sampling locations when calculating the changes [33,34]. The specific implementation process of the algorithm is as follows:

(1) Calculation on core points: Use a set of calculation "core" points for which one distance and confidence interval is calculated.

(2) Calculation of surface normals in 3D: For any given core point i, a normal vector is defined for each cloud by fitting a plane to the neighbors NNi of that cloud that are within a radius $D/2$ of i. The standard deviation of the distance of the neighbors NNi to the best fit plane is recorded and used as a measure of the cloud roughness $\sigma_i(D)$.

(3) Distance calculation between the two clouds: Once the normal is defined for the core point i, it is used to project i onto each cloud at scale d by defining a cylinder of radius $d/2$ whose axis goes through i and which is oriented along the normal vector. Projecting each of the subsets on the axis of the cylinder gives two distributions of distances. The mean of the distribution gives the average position, i_1 and i_2, and the two standard deviations give a

Fig. 5. Experimental devices: (a) DJI Phantom 4 RTK UAV; (b) EasyScan T10 TLS; (c) Qianxun SRmini RTK; (d) Trimble S5 Total Station; (e) Sanliang PP720 illuminometer.

Table 2
Data acquisition scheme.

Dam model	Weather scene	Acquisition equipment	
		UAV	TLS
Intact dam	Daytime & cloudy	✓	✓
	Sunset & cloudy	✓	—
	Daytime & rainy	✓	✓
Defective dam	Daytime & cloudy	✓	✓
	Sunset & cloudy	✓	—
	Daytime & rainy	✓	✓

Note: ✓indicates the equipment used for data acquisition under the corresponding scene.

local estimate of the point cloud roughness $\sigma_1(d)$ and $\sigma_2(d)$. The $L_{M3C2}(i)$ distance is given by the distance between i_1 and i_2.

3.1.2. Performance of defect detection

TLS point clouds of intact and defected dams in cloudy weather calculated by the M3C2 algorithm for comparison to detect the location and size of the defects. The maximum M3C2 distance was limited to 5 cm.

The point cloud detection results of the upstream are shown in Fig. 7a. Point cloud detection is not effective for precipitates due to the small thickness of the precipitate. For linear defects such as cracks, point cloud comparison can detect the location, size, and depth of large cracks (30 cm in length), but it is not effective for small cracks (15 cm in length). For block-like defects such as spalling and collapse, point cloud comparison can effectively detect the location, size, and depth of the defect. Therefore, to investigate the effective detection range and

accuracy of the point cloud, the box-selected defects in Fig. 7a are taken as an example. Their detection depth and size values were obtained by manually counting the M3C2 comparison results, as shown in Table 3. In the results of the three types of defects, the error values for depth are around 0.1 cm, indicating that the point cloud performs well in detecting the depth. In terms of crack size detection, the point cloud detection is not effective, and only the length value of cracks with a depth greater than 0.5 cm can be effectively detected. For the size detection of collapses and spalling, the detection results are similar to the actual value for spalling and collapse with a depth greater than 0.5 cm, with a maximum error of −1.578 cm. However, for spalling and collapses with a depth of 0.5 cm, the maximum error is −4.295 cm, the detected sizes are significantly smaller than the actual values. It can be seen that TLS point clouds are highly sensitive to depth; the greater the depth, the higher the detection accuracy.

The point cloud detection results of the downstream are shown in Fig. 7b. For large-scale defects such as slippages and soil collapses, the TLS point cloud can effectively detect their location, size, and depth informations. For cracks and grass collapse, the TLS point cloud is less effective due to the lower surface uniformity of the vegetated dam slope and bare dam slope compared to the upstream panel.

It can be seen that the TLS point cloud can quickly detect the location and approximate size of the defects, but there are some limitations, and the detection effect is more sensitive to the depth, size and surface uniformity of the measured dam, and it is not effective in detecting the precise size of the defects.

3.1.3. Influence of rainfall on TLS point cloud detection

The flood season is often accompanied with rainfall. To ensure the safe operation of earth-rockfill dams, it is unavoidable to conduct TLS point cloud detection on the dam after or even during rainfall in
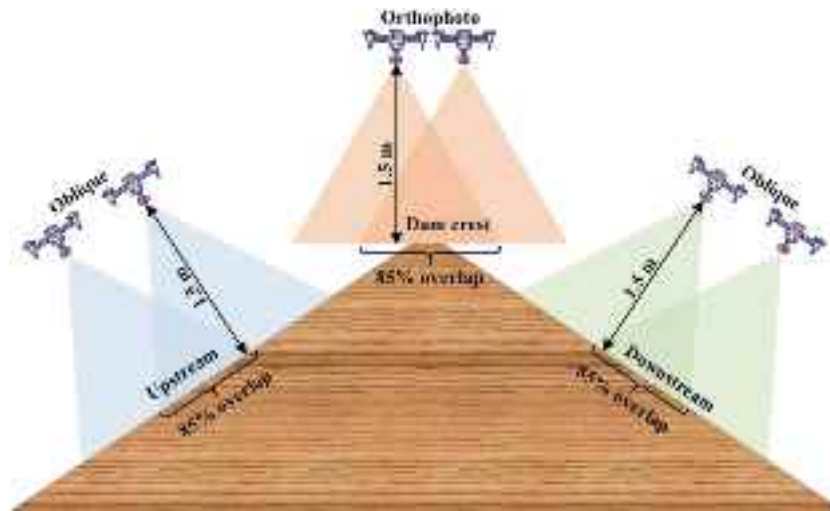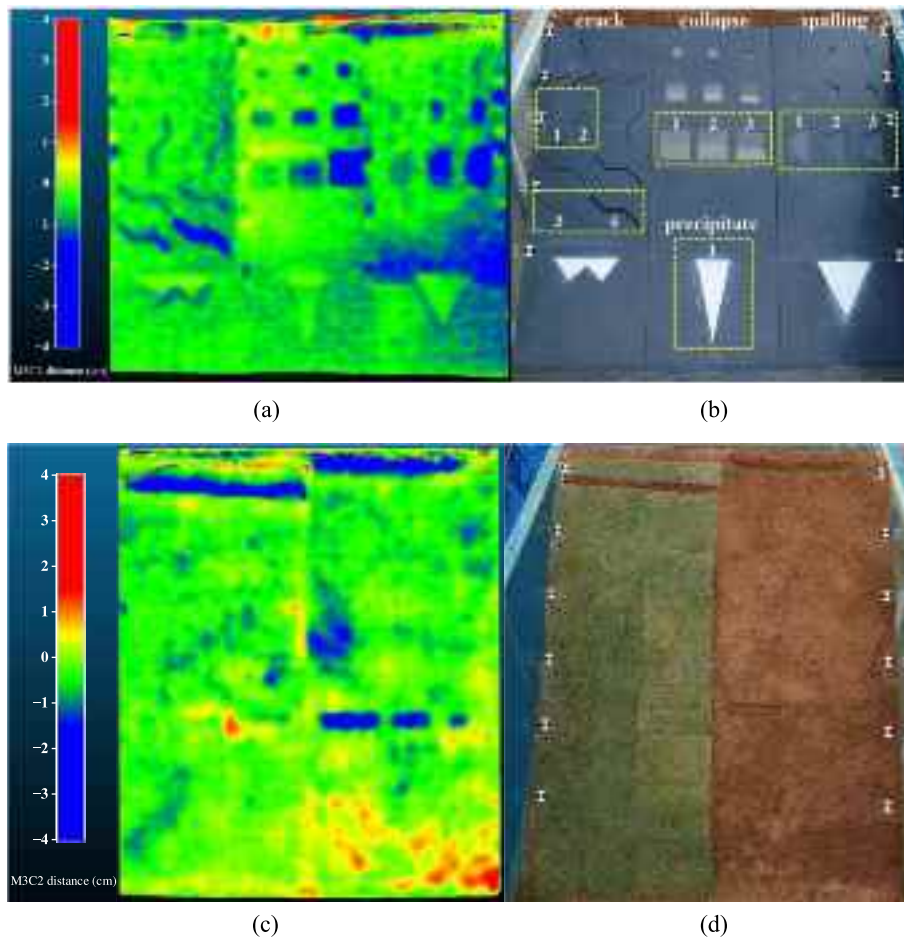
Fig. 6. UAV flight path.

Fig. 7. Point cloud detection effect: (a) The M3C2 distance distribution of the upstream; (b) distribution of the upstream defects; (c) the M3C2 distance distribution of the downstream; (d) distribution of downstream defects.

practical engineering applications. Therefore, the point clouds of the intact dam and the defective dam after rainfall were compared with those of the intact dam on a cloudy day to analyze the influence of rainfall on the detection performance.

The detection result of the intact dam after rainfall are shown in Fig. 8a. It can be seen that large raindrop-like and ripple-like anomalies appear on the upstream and downstream of the dam, respectively, which is related to the interference of TLS by the water accumulated on the dam after the rain. The accumulated water causes the laser beam to diverge, resulting in a decrease of rang accuracy [35]. The point cloud detection results of the defective dam after rainfall are shown in Fig. 8b, rainfall interferes with the detection of all types of defects to various

Table 3
Measured values and actual values from point cloud for the boxed defects.

Type of the defect	Measured value (cm)			Actual value (cm)			Error (cm)		
	Depth	Length	Width	Depth	Length	Width	Depth	Length	Width
Crack 1	0.498	26.793	0.911	0.5	30	2	−0.002	−3.207	−1.089
Crack 2	1.015	30.347	1.706	1.0	30	2	0.015	0.347	−0.294
Crack 3	1.421	29.724	2.719	1.5	30	2	−0.079	−0.276	0.719
Crack 4	2.075	30.680	4.473	2.0	30	2	0.075	0.680	2.473
Collapse 1	0.539	10.705	11.216	0.5	15	15	0.039	−4.295	−3.784
Collapse 2	1.062	13.707	13.422	1.0	15	15	0.062	−1.293	−1.578
Collapse 3	2.170	15.925	16.375	2.0	15	15	0.170	0.925	1.375
Spalling 1	0.485	17.514	7.608	0.5	20	10	−0.015	−2.486	−2.392
Spalling 2	1.109	19.005	8.829	1.0	20	10	0.109	−0.995	−1.171
Spalling 3	1.576	20.863	10.976	1.5	20	10	0.076	0.863	0.976

(a)

(b)
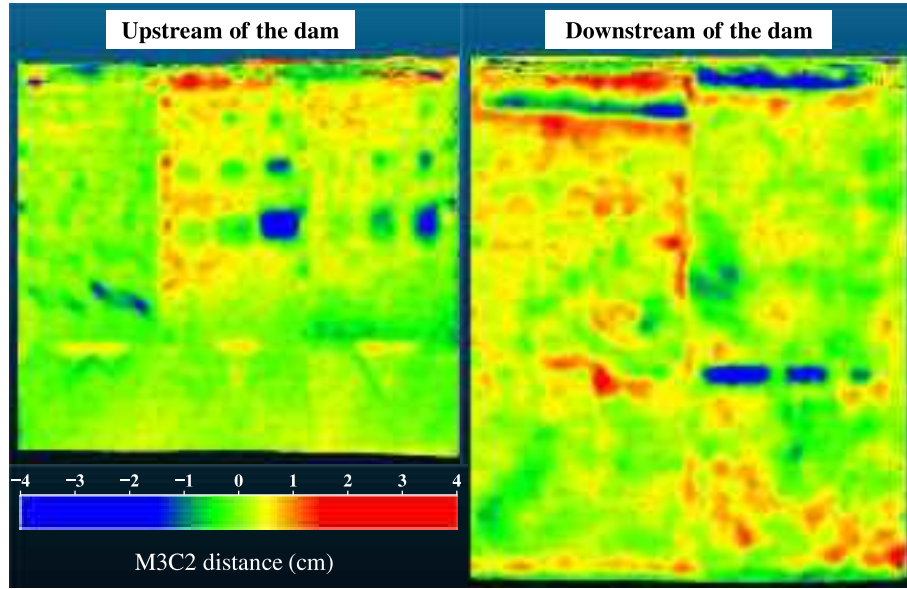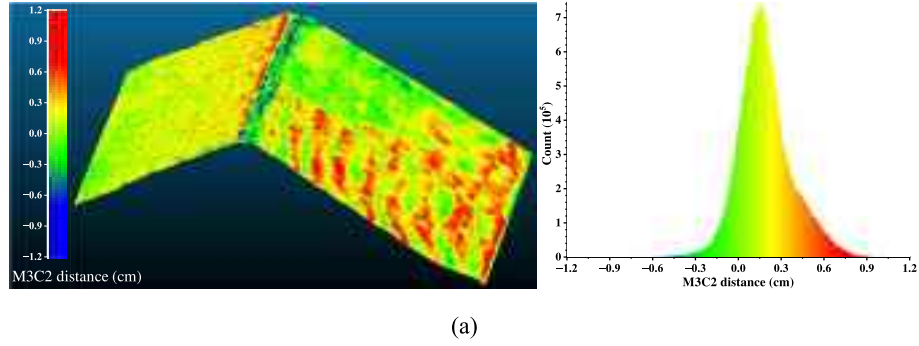
Fig. 8. Point cloud detection effect after rainfall: (a) the M3C2 distance distribution and value histogram of the intact dam; (b) the M3C2 distance distribution of the defective dam.

degrees, further reducing the effective detection range of point cloud detection to only large-scale defects. The point cloud detection results before and after rainfall are evaluated using well-known error metrics such as Standard Deviation (STD), Mean Absolute Error (MAE) and RMSE according to the following Equations:

$$STD = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(D(i)-\overline{D})^2} \tag{1}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|D(i)| \tag{2}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}D(i)^2} \tag{3}$$

where N denotes the number of observed data points of the sample, D denotes the M3C2 distance value of each point to the corresponding reference point cloud, $\overline{D}$ is the average value of M3C2 distance.
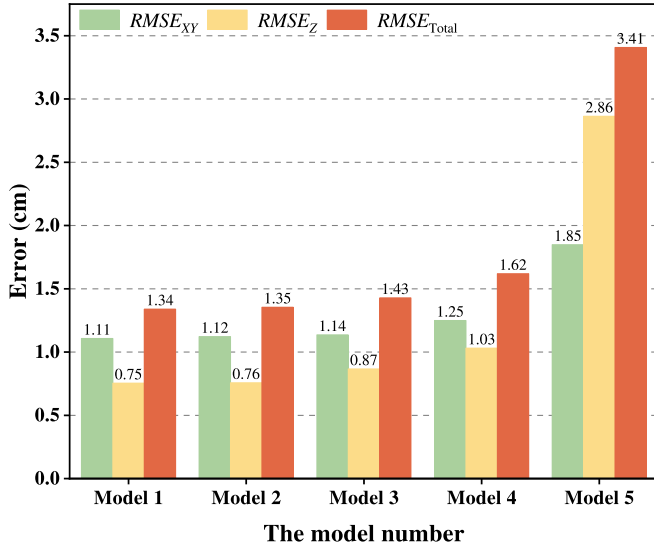
Table 4
Evaluation of point cloud detection results before and after rainfall.

Weather scene	Metrics (cm)		
	STD	MAE	RMSE
Daytime & cloudy	0.2476	0.1863	0.3372
Daytime & rainy	0.4175	0.3111	0.4367

Table 5
The GCP and CP arrangement for each model.

Model number	GCPs			CPs
	Upstream	Crest	Downstream	
Model-1	A-1 ~ A-18	B-1 ~ B-6	C-1 ~ C-22	CP-1 ~ CP-6
Model-2	A-1 ~ A-5, A-14 ~ A-18	B-1, B-3, B-4, B-6	C-1 ~ C-6, C-17 ~ C-22	CP-1 ~ CP-6
Model-3	A-1, A-18	B-1, B-3, B-4, B-6	C-1, C-22	CP-1 ~ CP-6
Model-4	—	B-1, B-3, B-4, B-6	—	CP-1 ~ CP-6
Model-5	—	—	—	CP-1 ~ CP-6

Fig. 9. The image alignment accuracy of the model under different GCP arrangements.

As shown in Table 4, from the evaluation metrics, the point cloud detection results before and after the rainfall deviate from each other and the detection data are discretized due to the interference of surface water accumulation on the TLS point cloud.

In practical engineering applications, TLS point cloud detection should be avoided as much as possible after or during rainfall. When the situation is urgent, supplementary detection should be conducted once the weather is cloudy to prevent missed detections.

3.2. 3D reconstruction of earth-rockfill dam based on multi-source data fusion

A 3D model with realistic texture features can comprehensively display the distribution of surface defects on the earth-rockfill dam, as well as the size of the defects. UAV images are used for the 3D reconstruction of the earth-rockfill dam model, and the influence of GCP arrangement and weather scene on model accuracy is discussed. Furthermore, TLS point clouds are combined for a multi-source fusion

model to further enhance the model details. Finally, a typical application of the multi-source fusion model to the engineering scene is illustrated.

3.2.1. Basic theory of 3D reconstruction

3D reconstruction refers to the use of machine vision theories such as feature point matching and multi-view geometry algorithms to process multiple image data with overlapping of a certain area to construct a 3D model with realistic textures [18,20]. Based on the texture, color differences, and positional relationships in UAV images, combined with RTK position information, the optical projection inversion is achieved to restore the surface information and geometric topology of the specified object. 3D reconstruction techniques include: feature extraction and matching, Structure from Motion (SfM) [36], Multi-View Stereo (MVS) [37], Image-Based modeling, and Image-Based rendering. Among these, SfM and MVS are crucial steps in the 3D reconstruction. SfM is primarily responsible for recovering camera poses and sparse 3D point clouds from overlapping images through methods such as feature detection, camera calibration, and bundle adjustment (BA) [38], which optimizes both camera parameters and 3D point positions to minimize reprojection errors. MVS, on the other hand, refines and densifies the 3D reconstruction by generating dense point clouds or meshes using techniques like depth map estimation, stereo matching, and surface reconstruction. Together, these processes enable the transition from UAV and TLS data to detailed and accurate 3D models.

In this study, Agisoft PhotoScan Professional, version 1.4.4 (PS) is used for the aerial triangulation of UAV images and TLS point clouds, and ContextCapture v10.20 (CC) is used for 3D modeling. The combined use of PS and CC effectively integrate the advantages of PS in aerial triangulation and CC in high-quality modeling, thereby enhancing the efficiency and quality of 3D reconstruction [39,40].

3.2.2. Influence of GCP arrangement on model accuracy

It is well known that the GCPs arrangement has a significant impact on the UAV image alignment accuracy and the efficiency of 3D reconstruction [10,40–42]. The GCPs arrangement required for the 3D reconstruction task of earth-rockfill dams is different from that of concrete dams due to the differences in dam structure types. Existing researches are mainly focused on concrete dams, making it difficult to find suitable GCP distributions as references. To analyze the influence of GCP arrangement on earth-rockfill dam model accuracy and optimize modeling efficiency, five detection models based on different GCP arrangements are proposed, as shown in Table 5. Model-1 has a high density of GCPs and extends to the bottom of the dam, so the UAV-based 3D model shows excellent performance [42]. Therefore, Model-1 is defined as the standard model. The middle of the dam is vulnerable and the GCPs may be damaged due to the emergence defects. Thus, Model-2, Model-3, and Model-4 are proposed, which place GCPs only on the sides of the dam, to optimize modeling efficiency with fewer GCPs. Model-5, serving as a control group, has no control point. The CPs for each model are set identically, with two CPs placed in the middle sections of the upstream, dam crest, and downstream to analyze the overall accuracy of the models.

The UAV image alignment accuracy is evaluated by CPs, and the root mean square error (RMSE) formula is used to calculate the difference between model coordinates and measured coordinates. The RMSE in the vertical direction is denoted as $RMSE_Z$, in the horizontal direction as $RMSE_{XY}$, and the total RMSE as $RMSE_{Total}$. The expressions are as follows:

$$RMSE_Z = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(Z_{UAV}(i) - Z_u(i)\right)^2} \tag{4}$$

$$RMSE_{XY} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[\left(X_{UAV}(i) - X_u(i)\right)^2 + \left(Y_{UAV}(i) - Y_u(i)\right)^2\right]} \tag{5}$$
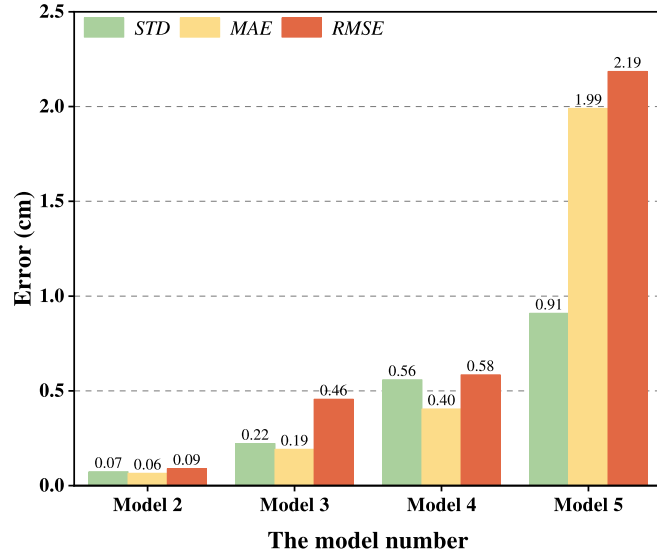
(a)
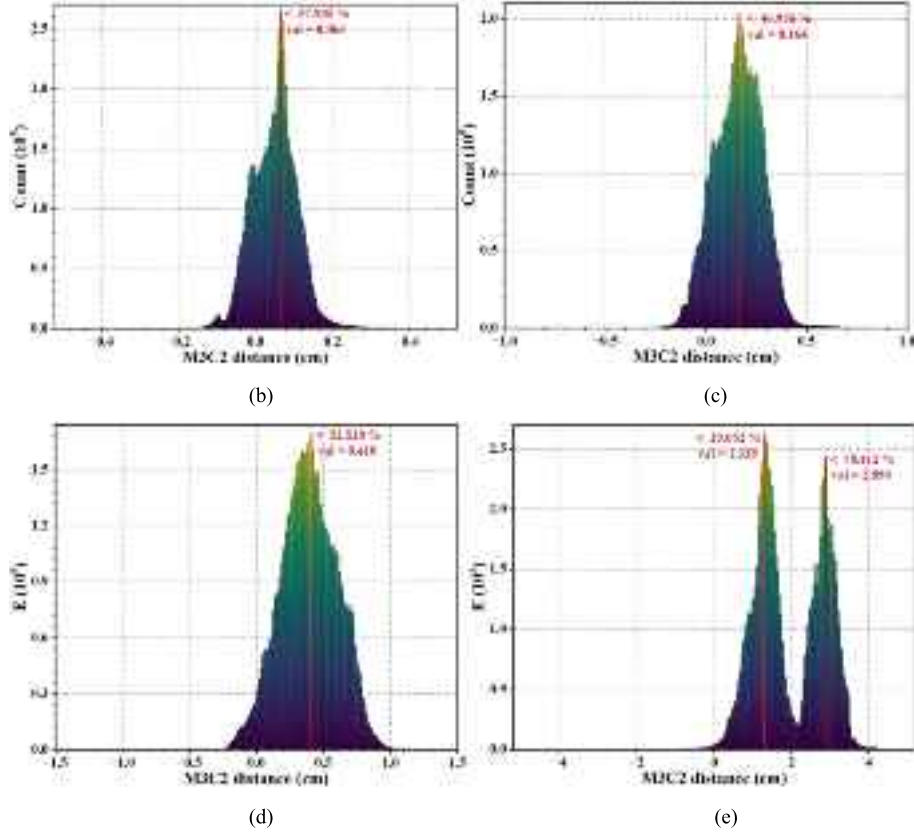
(b)

(c)

(d)

(e)

Fig. 10. The M3C2 distance between each model and the standard model: (a) the M3C2 evaluation results of each model; (b) the value histogram of Model-2; (c) the value histogram of Model-3; (d) the value histogram of Model-4; (e) the value histogram of Model-5.

$$RMSE_{\text{Total}} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[(X_{\text{UAV}}(i)-X_{\text{u}}(i))^2 + (Y_{\text{UAV}}(i)-Y_{\text{u}}(i))^2 + (Z_{\text{UAV}}(i)-Z_{\text{u}}(i))^2\right]}$$

(6)

where X_{UAV}, Y_{UAV}, Z_{UAV} denote model coordinates; X_{u}, Y_{u}, Z_{u} denote total station measured coordinates; and N denotes the number of CPs.

The analysis results of each model are shown in Fig. 9. For Model-1,

as the standard model with the highest accuracy, $RMSE_{XY}$ and $RMSE_Z$ are 1.11 cm and 0.75 cm, respectively, and $RMSE_{\text{Total}}$ is 1.34 cm. For Model-2, which has no GCP in the central area, $RMSE_{XY}$ and $RMSE_Z$ are 1.12 cm and 0.76 cm, respectively, with $RMSE_{\text{Total}}$ being 1.35 cm, an increase of only 0.01 cm compared to Model-1. For Model-3, the accuracy decreases in all directions, and the $RMSE_{\text{Total}}$ is 1.43 cm, which is an increase of 0.09 cm compared to Model-1. It is because the number of GCPs has been drastically reduced, and this arrangement of GCPs is insufficient to reconstruct a high-accuracy model. For Model-4, which has only four GCPs at the dam crest, the accuracy in all directions significantly

Table 6

The defect dimension measurement error of the model under different GCP arrangements.

Model number	MAE of crack (cm)		MAE of collapse (cm)		MAE of spalling (cm)		MAE of precipitate (cm)	
	Length	Width	Length	Width	Length	Width	Length	Width
Model-1	0.231	0.067	0.082	0.138	0.074	0.055	0.211	0.185
Model-2	0.253	0.073	0.089	0.147	0.085	0.063	0.237	0.201
Model-3	0.286	0.937	0.115	0.197	0.098	0.071	0.277	0.251
Model-4	0.354	0.111	0.126	0.223	0.132	0.087	0.324	0.298
Model-5	0.511	0.168	0.194	0.377	0.199	0.161	0.597	0.556
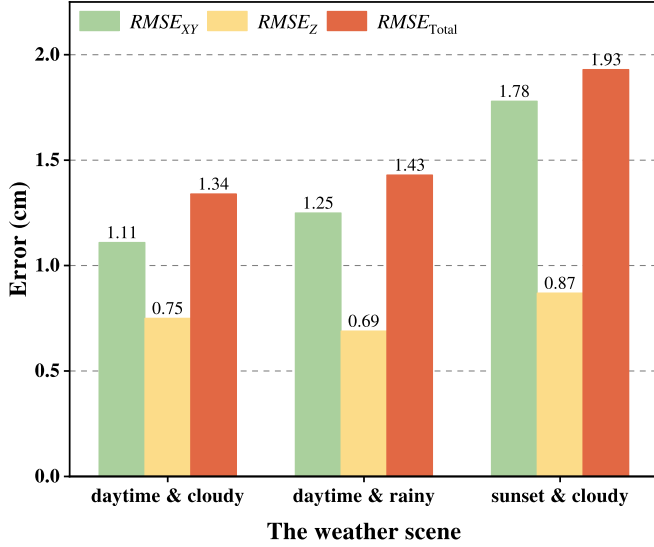
Fig. 11. The image alignment accuracy of the model under different weather scene.

decreased. Compared to Model-1, $RMSE_{XY}$, $RMSE_Z$ and $RMSE_{Total}$ increased by 0.14 cm, 0.28 cm, and 0.28 cm, respectively. The vertical error increase is twice that of the horizontal direction, indicating that accuracy of upstream and downstream cannot be maintained by the only GCPs at the dam crest. For Model-5, which has no GCP, $RMSE_{Total}$ reached 3.41 cm. Compared to Model-4, its accuracy decreased by more than half.

Considering that the RMSE analysis based on discrete CPs is not comprehensive enough to evaluate the overall accuracy of the model, the model accuracy is further evaluated using the generated point clouds from the 3D models. The M3C2 method in Section 3.1.1 is used to calculate the point cloud distance between each model and the standard model. The point cloud distance reflects the differences between the model and the standard model, the larger the point cloud distance from the standard model, the lower the model accuracy. Similar to the accuracy analysis in [43–45],the model accuracy is evaluated using well-known error metrics such as STD, MAE and RMSE.

According to Fig. 10, the *STD*, *MAE* and *RMSE* evaluation metrics shows that the overall change rule of the accuracy of each model is consistent with the results of accuracy analysis based on CPs. The average value of M3C2 distance of Model-2, Model-3 and Model-4 gradually increases with the decrease of the number of GCPs, and the

overall accuracy decreases steadily. For Model-5, which has no GCP, the average M3C2 distance increases greatly, indicating that GCP arrangement has a significant effect on improving model accuracy.

Furthermore, the dimensions of the defects boxed in Fig. 7b are measured on each model, and the defect dimension measurement errors were calculated.

As shown in Table 6, the dimension measurement errors of all types of defects gradually increase with the decrease in the number of GCPs. Moreover, for Model-5 with no GCPs, the measurement error is significantly larger than that of other models with GCPs.

From the analysis of model accuracy, the model accuracy varying patterns under various GCP arrangements are identified, and the accuracy and feasibility of the proposed GCP arrangements are validated. The engineering application scenes of each detection model are as follows:

(1) Model-1 can be used to establish a high-precision standard model as a reference for other models;
(2) Model-2 can be used for routine inspections to improve modeling efficiency with little loss of model accuracy;
(3) In the event of major disasters such as dam overtopping, dam failure, or earthquakes, it is necessary to quickly assess dam safety. Depending on the site damage, Model-3 or Model-4 can be selected, significantly improving modeling efficiency.

3.2.3. Influence of weather scene on model accuracy

In practical engineering, UAV data acquisition will inevitably face various weather scenes. Therefore, light and rainfall are considered as variables, and their effects on model accuracy are investigated.

The GCP arrangement of each scene is consistent with Model-1 in Section 3.2.2. The UAV image alignment accuracy under different scenes are shown in Fig. 11. For the daytime & cloudy scene, as the standard scene with the highest accuracy, $RMSE_{XY}$ and $RMSE_Z$ are 1.11 cm and 0.75 cm, respectively, and $RMSE_{Total}$ is 1.34 cm. For the daytime & rainy scene, the accuracy decreases in all directions, with the $RMSE_{Total}$ is 1.43 cm, which is 0.09 cm higher than daytime & cloudy scene. The accumulation of water on the dam surface after rainfall increased the difficulty of matching feature points, thus impacting the UAV image alignment accuracy. For the sunset & cloudy scene, the illumination value is about 1/12 of the daily & cloudy scene. Compared to the daily & cloudy scene, $RMSE_{XY}$, $RMSE_Z$, and $RMSE_{Total}$ increased by 0.67 cm, 0.12 cm, and 0.59 cm, respectively, with a total error increase of 44 %. Under weak illuminance, feature point matching becomes difficult, leading to a significant decrease in accuracy.

Similar to Section 3.2.2, the dimensions of the defects boxed in Fig. 7b are measured on each model, and the defect dimension

Table 7

The defect dimension measurement error of the model under different weather scenes.

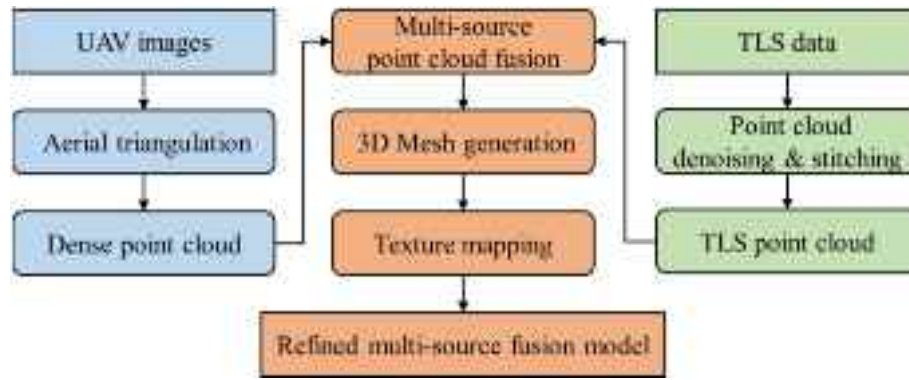| Weather scene	MAE of crack (cm)		MAE of collapse (cm)		MAE of spalling (cm)		MAE of precipitate (cm)	
	Length	Width	Length	Width	Length	Width	Length	Width
Daytime & cloudy	0.231	0.067	0.082	0.138	0.074	0.055	0.211	0.185
Daytime & rainy	0.307	0.098	0.120	0.205	0.115	0.077	0.296	0.213
Sunset & cloudy	0.415	0.129	0.151	0.279	0.152	0.116	0.376	0.355

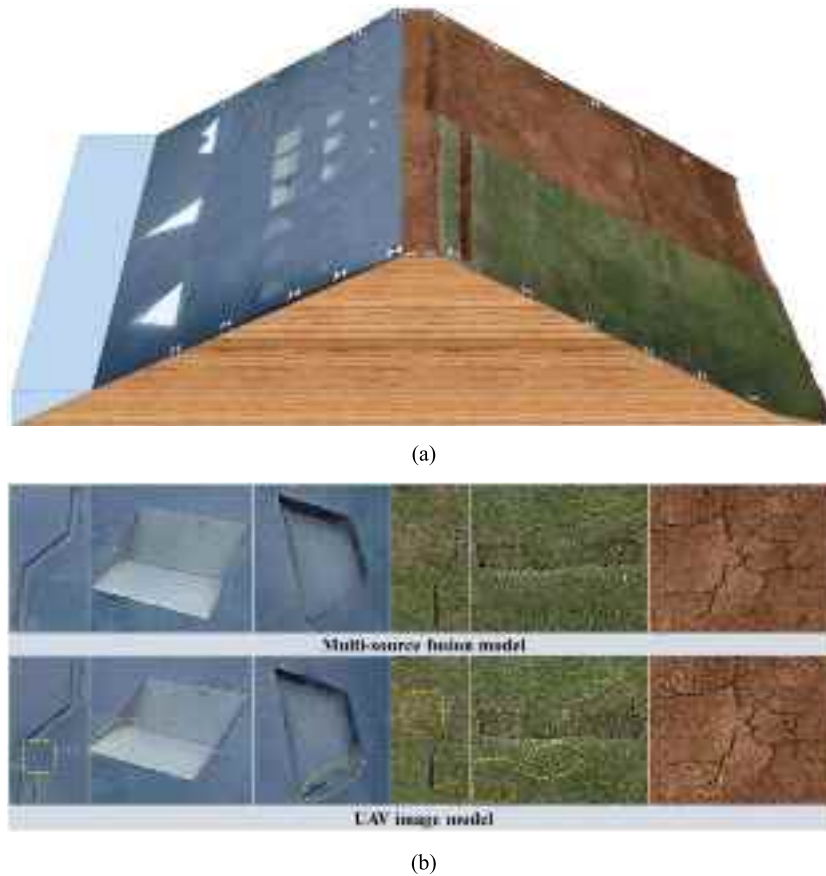**Fig. 12.** The process of multi-source fusion modeling.

(a)

(b)

Fig. 13. Multi-source fusion model of detective dam and its detailed effects: (a) multi-source fusion model of the detective dam; (b) comparison of the detailed effects between multi-source fusion model and UAV image model.

measurement errors were calculated. As shown in Table 7, the variation trend of the defect dimension measurement errors of the models is basically consistent with the above model accuracy analysis results based on CPs.

In engineering applications, weather scene should be closely monitored during UAV data acquisition, as model accuracy is optimal during daily & cloudy scene. In emergency situations, detections can be conducted after rainfall with minimal effect on model accuracy, whereas detection in poor illuminated scene should be avoided whenever possible.

In addition, the comparison shows that TLS point cloud detection and UAV-based 3D reconstruction are complementary in terms of application scenes and accuracy. The dense point cloud generated

during the 3D reconstruction process can also be used for point cloud detection. Therefore, TLS point cloud detection and UAV-based 3D reconstruction technology can be combined for the detection and analysis of surface defects of earth-rockfill dams in poor illuminated or post-rainfall scene.

3.2.4. Multi-source data fusion modeling

The UAV-based 3D reconstruction facilitates the efficient acquisition of 3D information of earth-rockfill dams and enables the creation of 3D models with realistic texture features. However, due to the complex structure and terrain distribution characteristics of surface defects on earth-rockfill dams, the perspective occlusion of UAV images easily causes local distortion and a decrease in model accuracy [46,47]. TLS

Table 8
The defect dimension measurement error of the multi-source fusion model and UAV image model.

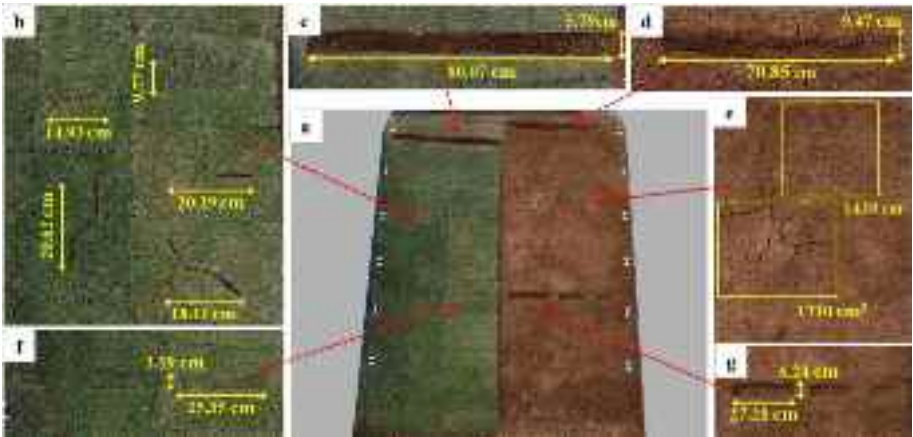| 3D model	MAE of crack (cm)		MAE of collapse (cm)		MAE of spalling (cm)		MAE of precipitate (cm)	
	Length	Width	Length	Width	Length	Width	Length	Width
Multi-source fusion model	0.219	0.059	0.063	0.120	0.059	0.034	0.184	0.169
UAV image model	0.253	0.073	0.089	0.147	0.085	0.063	0.237	0.201

Fig. 14. Downstream of the dam fusion model and its defects defection results: (a) downstream of the dam 3D model; (b) cracks on vegetation-covered slope; (c) dam slippage on vegetation-covered slope; (d) dam slippage on soil slope; (e) cracks on soil slope dam slippage slip on soil slope; (f) collapses on vegetation-covered slope; (g) collapses on soil slope.
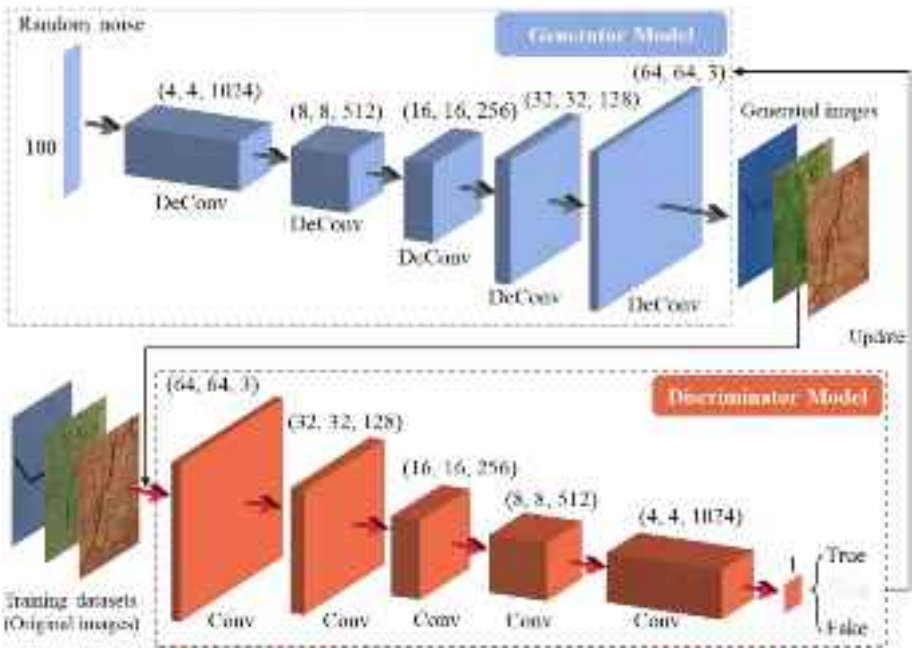

Fig. 15. Schematic diagram of DCGAN network structure.

Table 9
Relevant variables in *Precision*, *Recall* metrics.

Real results	Predicted results	
	Positive sample	Negative sample
Positive sample	TP (True Positive)	FN (False Negative)
Negative sample	FP (False Positive)	TN (True Negative)

point clouds can directly detect the surface positions of objects, providing a natural advantage in obtaining the geometric features and depth information of earth-rockfill dam [48]. To enhance the geometric details of the model, the fusion of UAV images and TLS point clouds is used to reconstruction a refined 3D model of defective dam.

As illustrated in Fig. 12, the multi-source fusion modeling begins with aerial triangulation of UAV images, followed by the generation of a dense point cloud. Meanwhile, TLS data is preprocessed, including point cloud denoising and stitching, to produce a complete and high-quality TLS point cloud. Next, point cloud data is coarsely registered based on
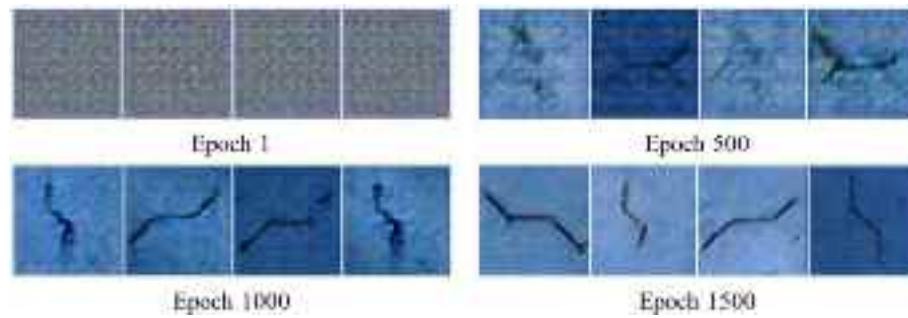
Fig. 16. Crack-panel images generated by DCGAN in multiple training epochs.

the geographic reference coordinate system and GCPs, followed by fine registration using the Iterative Closest Point (ICP) algorithm. Finally, triangulated mesh generation and texture mapping are performed on the fused point cloud to construct a refined multi-source fusion model with high geometric accuracy and detailed texture information. The core of the multi-source fusion modeling is the fusion of TLS point clouds with dense point clouds using the ICP algorithm. Through iterative optimization, the ICP algorithm calculates the optimal rigid transformation, including rotation and translation matrices, to minimize Euclidean distance errors between the two point clouds, thereby enabling precise point cloud alignment and seamless fusion [49].

Based on the research results from Sections 3.2.2 and 3.2.3, the GCP arrangement of the defective dam is consistent with Model-2 in Section 3.2.2. The acquisition of UAV images and TLS point clouds is conducted during daytime & cloudy scene. The multi-source fusion model is shown in Fig. 13a. Compared with 3D model generated solely from UAV images, it can be seen that the multi-source fusion model has a significant improvement in the control of local distortions and blurring of the cracks, collapses, and spallings on the upstream and downstream, as shown in Fig. 13b.

To analyze the measurement accuracy of the multi-source fusion model, the dimensions of the defects boxed in Section 3.1.2 were measured based on the multi-source fusion model. The defect dimension measurement errors are shown in Table 8, and the results indicate that the multi-source fusion model can detect the dimension of the defects. In terms of measurement accuracy. Compared with the UAV image model, the defect dimension measurement accuracy of the multi-source fusion model is further improved, except for the slightly larger measurement error for cracks, the measurement error for other defects remains within 0.2 cm, which meets the measurement requirements for general dam defects. Notably, the size of precipitates was successfully detected from the multi-source fusion model, which was not achieved by TLS point cloud detection.

After verifying the accuracy of the multi-source fusion model, a demonstration defects detection of the downstream slope is conducted based on the dam fusion model. As depicted in Fig. 14, defects on the downstream slope can be intuitively identified and measured. The multi-source fusion model can not only realize the accurate positioning of defects, but also get the accurate dimensions of defects and continuously grasp its development status. It makes the detection of dam defects simple, efficient and cost-effective.

Based on the study above, the following detection solutions can be used in practical engineering applications:

(1) Daily rapid TLS scanning of dam to detect localized defects and deformations by comparison with previous TLS data.
(2) Periodically, or when the TLS point cloud detects anomalies, the UAV is used to conduct rapid inspections and construct a multi-source fusion model to determine the location and size of the defects.

(3) If necessary, a high-precision dam model can be generated with the arrangement of GCPs for detailed measurement of each defect.

4. DCGAN-based defect image generation method for earth-rockfill dams

Due to the differences in materials and appearances between different dams, the characteristics of surface defects vary. When the model trained on the defect datasets of other dam is applied to detect the defects of a specific dam, it often faces problems such as poor model generalization ability and insufficient accuracy. If only the defect images of the dam to be tested are used as dataset samples, the limited samples often lead to poor model characterization. To address this issue, a DCGAN-based method for generating defect images of earth-rockfill dam is proposed for data augmentation of limited on-site images. A generative adversarial network is introduced for data augmentation, i.e. the DCGAN model is used for feature learning from small batches of on-site defect images, thereby generating similar but not identical images of dam defects, effectively augmenting the dataset samples. And the YOLO object detection model is used to validate the object detection performance on the augmented dataset.

DCGAN is a generative adversarial deep learning model proposed by Alec Radford et al. [50] in 2015, which is mainly applied in image generation and restoration, and video generation, etc. After years of application and research, it has achieved remarkable success in various image generation tasks [51–54]. Particularly for the image generation tasks with small-sample datasets in the field of engineering detection, DCGAN can effectively generate high-quality defect images even with relatively limited samples, providing a stable performance in data augmentation. Given the successful engineering applications of the model, this study applies it to the task of augmenting small-sample defect datasets of earth-rockfill dam, fully leveraging the powerful similar image generation capabilities of the DCGAN model.

As shown in Fig. 15, the core idea of the DCGAN model is to generate new data with both real features and some differences through adversarial learning between the Generator and the Discriminator. The process starts with random noise input into the Generator, which extends dimensions to generate high-dimensional images corresponding to the noise. These images are then input into the Discriminator, which identifies them as real or fake. If the Discriminator identifies a fake sample and feeds back to the Generator, the Generator will optimize the weights to create more realistic outputs in next iteration. If the Discriminator misidentifies a generated image as real, it will optimize its weights to enhance the judgment accuracy. The two models converge in the adversarial process, i.e. the authenticity of the Generator samples and the discriminative ability of the Discriminator are dynamically balanced.

You Only Look Once (YOLO) is a CNN-based object detection model. Its core idea is to transform the object detection task into a regression problem, directly predicting the bounding boxes and classes of objects
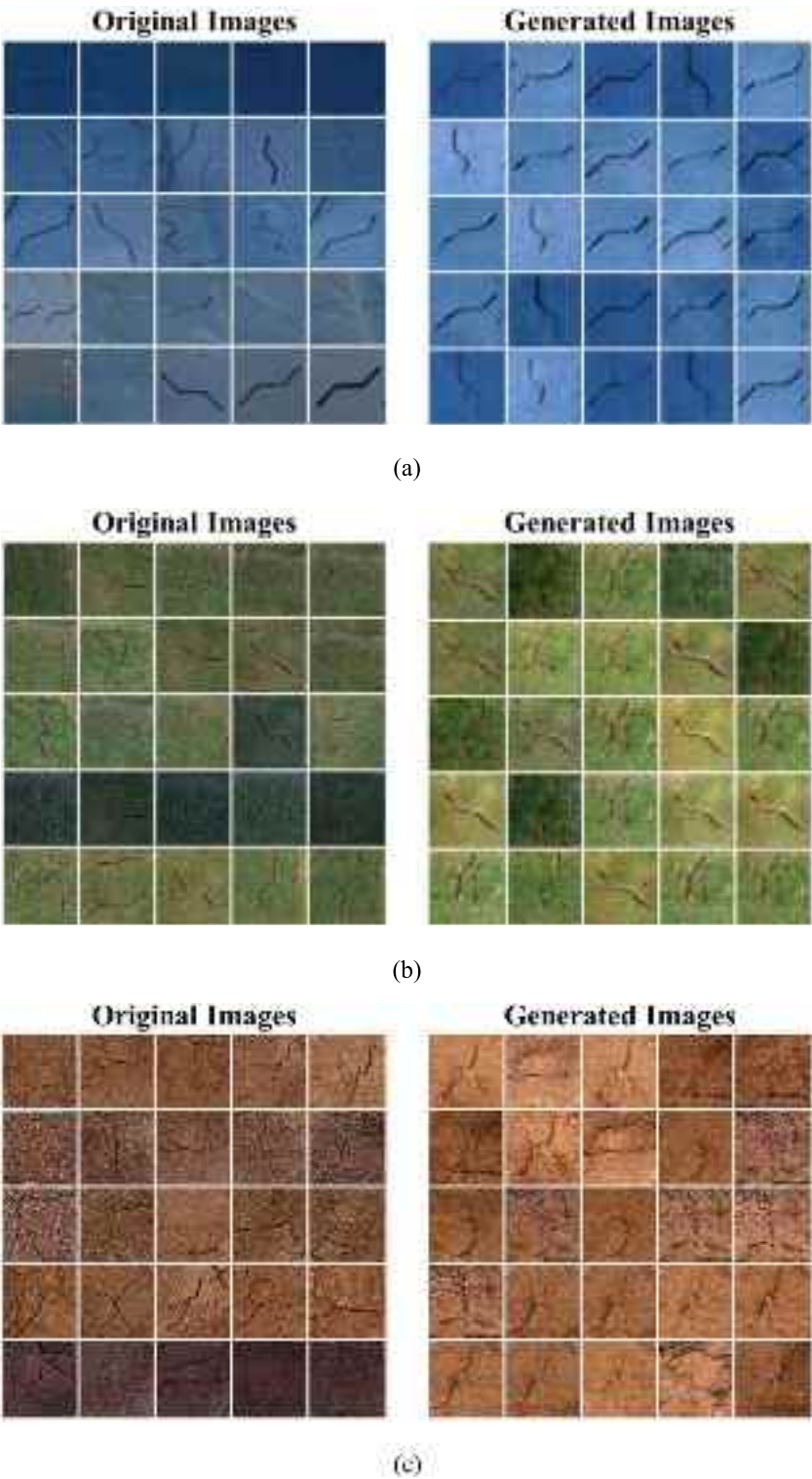
(a)

(b)

(c)

Fig. 17. Comparison between the original images and the generated images: (a) crack-panel; (b) crack-grass; (c) crack-soil.

Table 10
Evaluation results of the generated images.

Metrics	Type of the defect			
	All	Crack-panel	Crack-grass	Crack-soil
FID	179.34	180.64	177.52	179.87
SSIM	0.77	0.79	0.76	0.75
PSNR	17.31	18.56	17.19	16.17

through a single neural network [55]. YOLOv5 is a mature model in YOLO series and widely used in industrial applications. In addition to inheriting the design advantages of previous YOLO models, it also puts forward more outstanding improvements and innovations in image pre-processing and network structure. In this study, the YOLOv5 model is used to validate the object detection performance based on DCGAN generated datasets.

Table 11
Detection results of different models under the same test dataset.

Model	Training dataset			AP			mAP	Relative mAP (%)
	Total	Original	DCGAN	Crack-panel	Crack-grass	Crack-soil		
Model I	1650	1650	0	0.970	0.991	0.500	0.820	—
Model II	150	150	0	0.861	0.944	0.327	0.711	86.71
Model III	1650	150	1500	0.896	0.987	0.362	0.748	91.22

4.1. Dataset building and training strategy

4.1.1. DCGAN dataset and training strategy

The cracks, one of the main defects of the dam, are used as an example to build the dataset. The samples consist of crack images acquired from three different regions, including panel, vegetated slope, and bare slope, with 50 images for each region. To reduce computility consumption during DCGAN model training and improve the efficiency of data augmentation, each defect image is resized to 64×64.

Based on the dataset of surface defects on the earth-rockfill dam model, DCGAN is used to train the corresponding model. Subsequently, 500 similar images are generated from 50 image samples of each type of crack, and a total of 1500 similar images are generated. These images can be added as an extended dataset to the dataset used to train the crack detection model, thereby enhancing the training effectiveness.

The training parameters of the DCGAN model are set as follows, the number of channels is 3, the learning rate is 0.0005, batch size is 64, Adam optimizer is chosen, and the model generates the corresponding number of images by adjusting the offset noise of the fixed noise input. The DCGAN training is conducted on a NVIDIA GeForce RTX 4070 Super graphics card with 12 GB of memory, taking 5 min 20 s for 1500 epochs.

4.1.2. YOLO dataset and training strategy

Datasets used to train the YOLO model are built based on the original crack images and images generated by DCGAN. The object detection datasets is in VOC format, and labeled by the Labelme annotation software. Three types of datasets are built, i.e., a) Dataset-I: a dataset consisting of all original images; b) Dataset-II: a dataset consisting of the original images used to train the DCGAN model; and c) Dataset-III: a dataset consisting of the training images of DCGAN fused with the generated images at a ratio of 1:10 [56]. Both the Dataset-I and the Dataset-III have the same number of samples, totaling 1650 images, where the Dataset-II contains 150 original images and 1500 generated images.

To verify the data augmentation effect of integrating DCGAN into the YOLO object detection model, this study is based on the YOLOv5 model, which is trained on the Dataset-I, Dataset-II and Dataset-III respectively. And the obtained models are compared to explore the object detection performance of the models trained by the three datasets, aiming to further validate the effectiveness of DCGAN.

The training parameters for the YOLO model are set as follows, the input images are resized to a resolution of 640×640, the number of channels is 3, the learning rate is 0.0005, the batch size is 32, Adam optimizer is chosen, and the lightweight YOLOv5s model is selected as the pre-trained model.

4.2. Evaluation metrics

Based on the datasets and training strategy in Section 4.1, the DCGAN models and YOLO models are trained, and some appropriate evaluation metrics are selected to evaluate the model performance.

In this study, the Fréchet inception distance (FID), structural similarity index (SSIM) and peak signal-to-noise ratio (PSNR) metrics are used to evaluate the images generated by the DCGAN model.

The FID is an image quality evaluation metric commonly used to evaluate images generated by GAN networks [57]. By comparing the distribution differences between generated images and real images, it evaluates their realism and diversity. A lower FID score indicates a smaller distance between the real and generated images. The expression is as follows:

$$FID = \left\| \mu_r - \mu_g \right\|^2 + Tr(\Sigma_r + \Sigma_g - 2(\Sigma_r \Sigma_g)^{1/2}) \tag{7}$$

where μ_r and μ_g denote the average of the original and generated images, respectively, and Σ_r and Σ_g denote their corresponding covariance matrices; $Tr(\cdot)$ is the trace of the matrix.

The SSIM is a metric for measuring the structural similarity between two images [58]. It not only considers the difference between pixels but also comprehensively measures the similarity in terms of image brightness, contrast, and structural information. The expression is as follows:

$$\text{SSIM}(x, y) = \frac{(2\mu_x \mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{8}$$

where μ_x and μ_y denote the mean values of images x and y, respectively; σ_x and σ_y denote the standard deviations of images x and y, respectively; σ_{xy} denotes the covariance between images x and y; and C_1 and C_2 are constants used to keep the results stable. The SSIM value ranges from $[-1, 1]$, and the closer the SSIM value is to 1, the higher the structural similarity between the original image and the generated image.

The PSNR, which is commonly used in image evaluation, serves as an indicator of the similarity and quality between the original image and generated image. It provides a quantitative evaluation of the DCGAN model. The expressions are as follows:

$$MSE = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} \left[I_1(i,j) - I_2(i,j) \right]^2 \tag{9}$$

$$PNSR = 10 \cdot log_{10} \left(\frac{MAX^2}{MSE} \right) \tag{10}$$

where MSE is mean square error of the original image $I_1(i, j)$ and generated image $I_2(i, j)$; m and n are the height and width of the image, respectively; and MAX is the maximum value of the pixels in the image; a higher PSNR value indicates a smaller difference between two images.

The evaluation metrics used for the object detection model in this study include *Precision*, *Recall*, and class-wide mean Average Precision (*mAP*). The expressions are as follows:

$$Precision = \frac{TP}{TP + FP} \tag{11}$$

$$Recall = \frac{TP}{TP + FN} \tag{12}$$

where the meanings of *TP*, *FN*, and *FP* are shown in Table 9. Among them, positive samples indicate that the target is correctly identified and negative samples indicate that the target is misidentified and missed.

The overall performance of the model is mainly evaluated by *mAP*, and the expression is as follows:

$$mAP = \frac{1}{k} \sum_{i=1}^{k} AP_i \tag{13}$$

where k denotes the number of defect types, the value of *AP* is the area of

the Precision-Recall (PR) curve.

4.3. Comparative analysis of original and generated images

The visual inspection of the DCGAN training process is shown in Fig. 16. Starting with images of random noise, the DCGAN model extracted features from original crack images, continuously adjusting network parameters to generate similar images. At 500 epochs, some prominent features appeared in the generated images, but there were too much noise and a lack of detail. When it reached 1000 epochs, the generated images began to take on the crack features, but the features remained blurred. After 1500 epochs of training, the noise was significantly reduced, and the generated images showed more details, becoming similar to the original crack images.

These features can enhance the generalization ability of the object detection model, thereby improving detection accuracy.

Fig. 17 shows the effects of crack images generated by DCGAN. The SSIM, PNSR, FID metrics are used to evaluate the image quality of the generated images, and the results are shown in Table 10. It can be seen that the generated images are very similar to the original images in appearance, with subtle differences in feature details and color relationships. Specifically, the SSIM values (ranging from 0.75 to 0.79) indicate high structural similarity between the generated and original images, with crack-panel images achieving the highest SSIM (0.79). The PSNR values (ranging from 16.17 to 18.56) suggest that crack-panel images also exhibit the best pixel-level accuracy (18.56), while crack-soil images show relatively lower precision (16.17). Additionally, the FID values (ranging from 177.52 to 180.64) demonstrate that crack-grass images have the closest feature distribution to original images (FID = 177.52). The model demonstrates stable performance in generating crack images across three different regions, with the best results observed for crack-panel images. This will enhance the characterization capabilities and the detection accuracy of the trained object detection model.

To evaluate the detection performance of the models, defect detection tests were conducted on the three trained object detection models based on a unified test dataset. The detection performance of the models trained on different datasets is compared to that of the model trained on Dataset-I by calculating the 'Relative mAP', which denotes the comparative ratios of mAP metrics of a model to that of the model trained on Dataset-I. A higher Relative mAP indicates that the object detection performance of a model closely matches that of the model trained on Dataset-I.

The results are shown in Table 11. Among them, Model-I represents the model trained on Dataset-I, achieving the best detection performance for various types of cracks, with a mAP of 0.82. Model-II is trained on Dataset-II, due to the limited number of training samples, its detection performance significantly differs from that of Model-I, with the lowest mAP of 0.711 and a Relative mAP of only 86.71 %. Model-III is trained on Dataset-III, showing significant improvement in detection performance for all types of cracks compared to Model-II, with a mAP of 0.748 and a Relative mAP of 91.22 %. This indicates that the detection performance of Model-III is close to that of Model-I.

The results indicate that incorporating an additional 1500 generated images into the training dataset significantly improved the defect detection performance of the trained model compared to training with only 150 original images. This significantly reduces the demand for a large amount of original image samples, further demonstrating the effectiveness of DCGAN-generated images.

5. Conclusion

This study aims to advance an intelligent detection solution for surface defects on earth-rockfill dams, including TLS point cloud detection, multi-source fusion modeling, and a DCGAN-based defect intelligent detection method. An experimental study is conducted to explore the surface defect detection performance of TLS point cloud and multi-source fusion modeling on earth-rockfill dams, and the impacts of various factors are evaluated. Meanwhile, a DCGAN-based method for generating defect images of earth-rockfill dam is proposed for data augmentation of limited on-site images, thereby enhancing the defects detection performance of the object detection model.

The proposed intelligent detection solution optimizes the visual detection performance of surface defects on earth-rockfill dams, and provides a new perspective for the intelligent detection of earth-rockfill dam surface defects under limited data samples. The primary conclusions are as follows:

(1) TLS point clouds can quickly detect the location and approximate size of defects, but the performance in defect dimension measurement is not good. The detection performance is highly sensitive to the depth, size, and surface uniformity of the defects, which should be combined with UAV photogrammetry to achieve precise detection of multi-scale and various types of defects. Rainfall has a significant effect on TLS detection performance. In practical engineering, TLS detection should be avoided after or during rainfall.

(2) The GCP arrangement and weather scene are major influences on model accuracy. According to accuracy analysis results under different GCP arrangements, four GCP arrangements are proposed for the corresponding engineering application scenes, the GCPs greatly enhances model accuracy. The accuracy of the models under different weather scenes is compared. Model accuracy is best in daily & cloudy scene. Rainfall has less effect on model accuracy, while the illuminance has a significant effect. For weak illumination or post-rainfall scenes, a combination of TLS detection and 3D reconstruction is recommended to enhance measurement accuracy.

(3) The multi-source fusion model is conductive to fully combining the technical advantages of UAV image and TLS point cloud to optimize the geometric details of the 3D model. Compared with the UAV image model, the multi-source fusion model shows obvious improvement in the control of local distortion and blurring of various types of defects, and has higher dimensional measurement accuracy. It helps to effectively control the location, dimension, and development status of defects, making the operation and maintenance of earth-rockfill dams more efficient and cost-effective.

(4) The DCGAN model is introduced for data augmentation of limited on-site images, similar defect images are generated to expand the datasets. The YOLOv5 model is used for validation, demonstrating that the DCGAN model can extract the features from small-sample defect images on earth-rockfill dams, and the generated similar defect images significantly enhance the performance of defect detection model.

There are still some limitations in this study. Since the experiments are conducted on the earth-rockfill dam model, the dam materials cannot be consistent with the real dam. Meanwhile, the setting of defect geometries and dimensions as well as the simulation of weather scenes, are somewhat idealized.

In future research, we will consider to implement the proposed intelligent detection solution in practical earth-rockfill dam projects. The UAV flight paths and TLS acquisition methods should be optimized to provide standardized approaches for the construction of multi-source fusion models and defect detection. Moreover, improvements are necessary for the DCGAN and object detection algorithms to optimize their performance in engineering applications. With the advancement of BIM and digital twin technologies, a safety evaluation and early warning system for earth-rockfill dams will be established based on the intelligent detection solution in future.

CRediT authorship contribution statement

Hongyu Ren: Writing – review & editing, Writing – original draft, Methodology, Data curation. **Rui Pang:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization. **Weijie Huang:** Writing – review & editing, Methodology, Conceptualization. **Bin Xu:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization.

Declaration of competing interest

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

Acknowledgements

Data availability

Data will be made available on request.

References

[1] F. Kang, J. Li, S. Zhao, Y. Wang, Structural health monitoring of concrete dams using long-term air temperature for thermal effect simulation, Eng. Struct. 180 (2019) 642–653, https://doi.org/10.1016/j.engstruct.2018.11.065.

[2] Y. Li, R. Pang, B. Xu, Y. Zhou, Jaya-ICSM: a rapid inverse method driven by monitoring data for concrete-faced rockfill dams static displacement simulation, Adv. Eng. Inform. 57 (2023) 102080, https://doi.org/10.1016/j.aei.2023.102080.

[3] B. Xu, Z. Rong, R. Pang, W. Tan, B. Wei, A novel method for settlement imputation and monitoring of earth-rockfill dams subjected to large-scale missing data, Adv. Eng. Inform. 62 (2024) 102642, https://doi.org/10.1016/j.aei.2024.102642.

[4] M. Xu, R. Pang, Y. Zhou, B. Xu, Seepage safety evaluation of high earth-rockfill dams considering spatial variability of hydraulic parameters via subset simulation, J. Hydrol. 626 (2023) 130261, https://doi.org/10.1016/j.jhydrol.2023.130261.

[5] Z. Rong, R. Pang, B. Xu, Y. Zhou, Dam safety monitoring data anomaly recognition using multiple-point model with local outlier factor, Autom. Constr. 159 (2024) 105290, https://doi.org/10.1016/j.autcon.2024.105290.

[6] M.A. Akbar, U. Qidwai, M.R. Jahanshahi, An evaluation of image-based structural health monitoring using integrated unmanned aerial vehicle platform, Struct. Control Health Monit. 26 (1) (2019) e2276.

[7] D.P. Bigman, D.J. Day, Ground penetrating radar inspection of a large concrete spillway: a case-study using SFCW GPR at a hydroelectric dam, Case Stud. Constr. Mater. 16 (2022) e00975, https://doi.org/10.1016/j.cscm.2022.e00975.

[8] X. Xu, J. Wu, J. Shen, Z. He, Case study: application of GPR to detection of hidden dangers to underwater hydraulic structures, J. Hydraul. Eng. 132 (1) (2006) 12–20, https://doi.org/10.1061/(ASCE)0733-9429(2006)132:1(12).

[9] J. Chen, F. Cheng, F. Xiong, Q. Ge, S. Zhang, An experimental study: fiber Bragg grating–hydrothermal cycling integration system for seepage monitoring of rockfill dams, Struct. Health Monit. 16 (1) (2017) 50–61, https://doi.org/10.1177/1475921716661874.

[10] S. Zhao, F. Kang, J. Li, C. Ma, Structural health monitoring and inspection of dams based on UAV photogrammetry with image 3D reconstruction, Autom. Constr. 130 (2021) 103832, https://doi.org/10.1016/j.autcon.2021.103832.

[11] C.V. Dung, L.D. Anh, Autonomous concrete crack detection using deep fully convolutional neural network, Autom. Constr. 99 (2019) 52–58, https://doi.org/10.1016/j.autcon.2018.11.028.

[12] D. Feng, M.Q. Feng, Computer vision for SHM of civil infrastructure: from dynamic response measurement to damage detection – a review, Eng. Struct. 156 (2018) 105–117, https://doi.org/10.1016/j.engstruct.2017.11.018.

[13] R.S. Adhikari, O. Moselhi, A. Bagchi, Image-based retrieval of concrete crack properties for bridge inspection, Autom. Constr. 39 (2014) 180–194, https://doi.org/10.1016/j.autcon.2013.06.011.

[14] Z. Wang, J. Zhu, T. Ma, Review on monitoring of pavement subgrade settlement: influencing factor, measurement and advancement, Measurement 237 (2024) 115225, https://doi.org/10.1016/j.measurement.2024.115225.

[15] B.F. Spencer, V. Hoskere, Y. Narazaki, Advances in computer vision-based civil infrastructure inspection and monitoring, Engineering 5 (2) (2019) 199–222, https://doi.org/10.1016/j.eng.2018.11.030.

[16] K.L. Cook, An evaluation of the effectiveness of low-cost UAVs and structure from motion for geomorphic change detection, Geomorphology 278 (2017) 195–208, https://doi.org/10.1016/j.geomorph.2016.11.009.

[17] W. Yi, M. Sutrisna, H. Wang, Unmanned aerial vehicle based low carbon monitoring planning, Adv. Eng. Inform. 48 (2021) 101277, https://doi.org/10.1016/j.aei.2021.101277.

[18] W.W. Immerzeel, P.D.A. Kraaijenbrink, J.M. Shea, A.B. Shrestha, F. Pellicciotti, M.F.P. Bierkens, S.M. de Jong, High-resolution monitoring of Himalayan glacier dynamics using unmanned aerial vehicles, Remote Sens. Environ. 150 (2014) 93–103, https://doi.org/10.1016/j.rse.2014.04.025.

[19] L. Yang, H. Cai, Enhanced visual SLAM for construction robots by efficient integration of dynamic object segmentation and scene semantics, Adv. Eng. Inform. 59 (2024) 102313, https://doi.org/10.1016/j.aei.2023.102313.

[20] S. Chen, G. Fan, J. Li, Improving completeness and accuracy of 3D point clouds by using deep learning for applications of digital twins to civil structures, Adv. Eng. Inform. 58 (2023) 102196, https://doi.org/10.1016/j.aei.2023.102196.

[21] L. Ramos-Alcázar, M. Marchamalo-Sacristán, R. Martínez-Marín, Estimating and plotting TLS midrange precisions in field conditions: application to dam monitoring, Int. J. Civ. Eng. 15 (2) (2017) 299–307, https://doi.org/10.1007/s40999-016-0093-3.

[22] J. Zhu, T. Bu, T. Ma, X. Huang, F. Chen, Raster-based point cloud mapping of defective road marking: toward automated road inspection via airborne LiDAR, J. Transp. Eng. Part B: Pavements 150 (2) (2024) 4024015, https://doi.org/10.1061/JPEODX.PVENG-1410.

[23] Q. Xu, X. Dong, W. Li, Integrated space-air-ground early detection, monitoring and warning system for potential catastrophic geohazards, Geomat. Inf. Sci. Wuhan Univ. 44 (07) (2019) 957–966, https://doi.org/10.13203/j.whugis20190088.

[24] H. Li, X. Li, W. Li, S. Zhang, J. Zhou, Quantitative assessment for the rockfall hazard in a post-earthquake high rock slope using terrestrial laser scanning, Eng. Geol. 248 (2019) 1–13, https://doi.org/10.1016/j.enggeo.2018.11.003.

[25] N. Bolourian, A. Hammad, LiDAR-equipped UAV path planning considering potential locations of defects for bridge inspection, Autom. Constr. 117 (2020) 103250, https://doi.org/10.1016/j.autcon.2020.103250.

[26] S. Zhao, F. Kang, J. Li, Concrete dam damage detection and localisation based on YOLOv5s-HSC and photogrammetric 3D reconstruction, Autom. Constr. 143 (2022) 104555, https://doi.org/10.1016/j.autcon.2022.104555.

[27] J. Zhu, J. Zhong, T. Ma, X. Huang, W. Zhang, Y. Zhou, Pavement distress detection using convolutional neural networks with images captured via UAV, Autom. Constr. 133 (2022) 103991, https://doi.org/10.1016/j.autcon.2021.103991.

[28] Z. Yu, Y. Shen, C. Shen, A real-time detection approach for bridge cracks based on YOLOv4-FPM, Autom. Constr. 122 (2021) 103514, https://doi.org/10.1016/j.autcon.2020.103514.

[29] J. Li, X. Lu, P. Zhang, Q. Li, Weak feature crack detection in high-resolution concrete dam surface image with LarkMNet, Measurement 238 (2024) 115327, https://doi.org/10.1016/j.measurement.2024.115327.

[30] K. Mirzaei, M. Arashpour, E. Asadi, H. Masoumi, Y. Bai, A. Behnood, 3D point cloud data processing with machine learning for construction and infrastructure applications: a comprehensive review, Adv. Eng. Inform. 51 (2022) 101501, https://doi.org/10.1016/j.aei.2021.101501.

[31] Y. Lu, S. Wang, S. Fan, J. Lu, P. Li, P. Tang, Image-based 3D reconstruction for Multi-Scale civil and infrastructure Projects: a review from 2012 to 2022 with new perspective from deep learning methods, Adv. Eng. Inform. 59 (2024) 102268, https://doi.org/10.1016/j.aei.2023.102268.

[32] K.N. Poku-Agyemang, A. Reiterer, Weighted multiple point cloud fusion, PFG – J. Photogramm. Remote Sens. Geoinf. Sci. 93 (1) (2025) 65–78, https://doi.org/10.1007/s41064-024-00310-1.

[33] D. Lague, N. Brodu, J. Leroux, Accurate 3D comparison of complex topography with terrestrial laser scanner: application to the Rangitikei canyon (N-Z), ISPRS-J. Photogramm. Remote Sens. 82 (2013) 10–26, https://doi.org/10.1016/j.isprsjprs.2013.04.009.

[34] D. Moon, S. Chung, S. Kwon, J. Seo, J. Shin, Comparison and utilization of point cloud generated from photogrammetry and laser scanning: 3D world model for smart heavy equipment planning, Autom. Constr. 98 (2019) 322–331, https://doi.org/10.1016/j.autcon.2018.07.020.

[35] A. Filgueira, H. González-Jorge, S. Lagüela, L. Díaz-Vilariño, P. Arias, Quantifying the influence of rain in LiDAR performance, Measurement 95 (2017) 143–148, https://doi.org/10.1016/j.measurement.2016.10.009.

[36] M.J. Westoby, J. Brasington, N.F. Glasser, M.J. Hambrey, J.M. Reynolds, 'Structure-from-Motion' photogrammetry: a low-cost, effective tool for geoscience applications, Geomorphology 179 (2012) 300–314, https://doi.org/10.1016/j.geomorph.2012.08.021.

[37] Y. Furukawa, J. Ponce, Accurate, dense, and robust multiview stereopsis, IEEE Trans. Pattern Anal. Mach. Intell. 32 (8) (2010) 1362–1376, https://doi.org/10.1109/TPAMI.2009.161.

[38] B. Triggs, P.F. McLauchlan, R.I. Hartley, A.W. Fitzgibbon, in: Bundle Adjustment: A Modern Synthesis, Springer, Berlin, 2000, pp. 298–372.

[39] C.M. Yeum, S.J. Dyke, Automated region-of-interest localization and classification for vision-based visual assessment of civil infrastructure, Struct. Health Monit. 18 (3) (2019) 675–689, https://doi.org/10.1177/1475921718765419.

[40] M.R. James, S. Robson, S. D'Oleire-Oltmanns, U. Niethammer, Optimising UAV topographic surveys processed with structure-from-motion: ground control quality, quantity and bundle adjustment, Geomorphology 280 (2017) 51–66, https://doi.org/10.1016/j.geomorph.2016.11.021.

[41] E. Ferrer-González, F. Agüera-Vega, F. Carvajal-Ramírez, P. Martínez-Carricondo, UAV photogrammetry accuracy assessment for corridor mapping based on the

number and distribution of ground control points, Remote Sens. 12 (15) (2020) 2447, https://doi.org/10.3390/rs12152447.

[42] E. Ridolfi, G. Buffi, S. Venturi, P. Manciola, Accuracy analysis of a dam model from drone surveys, Sensors 17 (8) (2017) 1777, https://doi.org/10.3390/s17081777.

[43] M. Mohammadi, M. Rashidi, V. Mousavi, A. Karami, Y. Yu, B. Samali, Quality evaluation of digital twins generated based on UAV photogrammetry and TLS: bridge case study, Remote Sens. 13 (17) (2021) 3499, https://doi.org/10.3390/rs13173499.

[44] V. Mousavi, M. Khosravi, M. Ahmadi, N. Noori, S. Haghshenas, A. Hosseininaveh, M. Varshosaz, The performance evaluation of multi-image 3D reconstruction software with different sensors, Measurement 120 (2018) 1–10, https://doi.org/10.1016/j.measurement.2018.01.058.

[45] V. Mousavi, M. Varshosaz, M. Rashidi, W. Li, A new multi-criteria tie point filtering approach to increase the accuracy of UAV photogrammetry models, Drones 6 (12) (2022) 413, https://doi.org/10.3390/drones6120413.

[46] J. Kang, D. Kim, C. Lee, J. Kang, D. Kim, Efficiency study of combined UAS photogrammetry and terrestrial LiDAR in 3D modeling for maintenance and management of fill dams, Remote Sens. 15 (8) (2023) 2026, https://doi.org/10.3390/rs15082026.

[47] J. Li, Y. Peng, Z. Tang, Z. Li, Three-dimensional reconstruction of railway bridges based on unmanned aerial vehicle–terrestrial laser scanner point cloud fusion, Build.-Basel 13 (11) (2023) 2841, https://doi.org/10.3390/buildings13112841.

[48] M. Nap, S. Chiorean, C. Cira, M. Manso-Callejo, V. Păunescu, E. Șuba, T. Sălăgean, Non-destructive measurements for 3D modeling and monitoring of large buildings using terrestrial laser scanning and unmanned aerial systems, Sensors 23 (12) (2023) 5678, https://doi.org/10.3390/s23125678.

[49] P.J. Besl, H.D. Mckay, A method for registration of 3-D shapes, IEEE Trans. Pattern Anal. Mach. Intell. 14 (2) (1992) 239–256.

[50] A. Radford, L. Metz, S. Chintala, Unsupervised representation learning with deep convolutional generative adversarial networks, Comput. Sci. (2015).

[51] H. Qin, D. Zhang, Y. Tang, Y. Wang, Automatic recognition of tunnel lining elements from GPR images using deep convolutional networks with data augmentation, Autom. Constr. 130 (2021) 103830, https://doi.org/10.1016/j.autcon.2021.103830.

[52] Y. Lin, F. Fan, J. Zhang, J. Zhou, P. Liao, H. Chen, Z. Deng, Y. Zhang, DHI-GAN: improving dental-based human identification using generative adversarial networks, IEEE Trans. Neural Netw. Learn. Syst. 34 (12) (2023) 9700–9712, https://doi.org/10.1109/TNNLS.2022.3159781.

[53] M. Zhang, R. Hu, J. Mo, Z. Xiang, Z. Zhou, A cross-domain state monitoring method for high-speed train brake pads based on data generation under small sample conditions, Measurement 226 (2024) 114074, https://doi.org/10.1016/j.measurement.2023.114074.

[54] H. Zhong, S. Yu, H. Trinh, Y. Lv, R. Yuan, Y. Wang, Fine-tuning transfer learning based on DCGAN integrated with self-attention and spectral normalization for bearing fault diagnosis, Measurement 210 (2023) 112421, https://doi.org/10.1016/j.measurement.2022.112421.

[55] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, You only look once: unified, real-time object detection, Comput. Vis. Pattern Recogn. (2016).

[56] H. Tan, Z. Guo, Z. Lin, Y. Chen, D. Huang, W. Yuan, H. Zhang, J. Yan, General generative AI-based image augmentation method for robust rooftop PV segmentation, Appl. Energy 368 (2024) 123554, https://doi.org/10.1016/j.apenergy.2024.123554.

[57] M. Hensel, H. Ramsauer, T. Unterthiner, B. Nessler, S. Hochreiter, GANs trained by a two time-scale update rule converge to a local Nash equilibrium, Adv. Neural Inf. Process. Syst. 30 (2017) 5.

[58] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Trans. Image Process. 13 (4) (2004) 600–612, https://doi.org/10.1109/TIP.2003.819861.