Journal of Infrastructure Intelligence and Resilience 4 (2025) 100113

Contents lists available at ScienceDirect

Journal of Infrastructure Intelligence and Resilience

journal homepage: www.journals.elsevier.com/journal-of-infrastructure-intelligence-and-resilience

Automatic PAUT crack detection and depth identification framework based on inspection robot and deep learning method

Fei Hu [a], Hong-ye Gou [a,b,c,*], Hao-zhe Yang [a,e], Huan Yan [a], Yi-qing Ni [d], You-wu Wang [d]

[a] Department of Bridge Engineering, School of Civil Engineering, Southwest Jiaotong University, Chengdu, Sichuan, 610031, China
[b] Key Laboratory of High-Speed Railway Engineering of the Ministry of Education, Chengdu, Sichuan, 610031, China
[c] National Key Laboratory of Bridge Intelligent and Green Construction, Southwest Jiaotong University, Chengdu, Sichuan, 611756, China
[d] Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, China
[e] China Railway Major Bridge Reconnaissance and Design Institute Co., Ltd, Wuhan, 430050, China

ARTICLE INFO

Handling Editor: Yanwei Li

Keywords:
Orthotropic steel bridge deck
Fatigue crack detection
DCGAN
YOLO algorithm
Deep learning
Attention mechanism

ABSTRACT

Orthotropic steel bridge decks (OSD) are widely acclaimed for their lightweight, high load-carrying capacity, and adaptability, making them a popular choice in steel structure bridges. However, the complex nature of their structure makes them susceptible to fatigue cracking, posing significant safety concerns. To address the issues above, this study employs a robot equipped with an ultrasonic phased array probe to automate the detection of internal cracks within Orthotropic Steel Decks (OSD). A Deep Convolutional Generative Adversarial Network (DCGAN) is utilized to augment the training dataset of Phased Array Ultrasonic Testing (PAUT) images. The YOLO series algorithms are applied and compared for crack localization, with YOLO v7-tiny exhibiting the highest accuracy and speed. Integrating attention mechanisms into the YOLO v7-tiny algorithm to facilliate rapid and high-precision crack detection. Analyzing the echo region with an echo intensity bar enabled the identification of crack depth, with an identification error within 5%.

1. Introduction

The orthotropic steel deck (OSD) is a favored structural type in long-span bridge design, known for its lightweight nature and high load-carrying capacity (Karimi and Mirza, 2023) (Liu et al., 2021). However, the presence of fatigue cracks at connection sites, attributed to cyclic vehicle loading and manufacturing defects, poses a significant global concern (Shao et al., 2013) (Fisher and Barsom, 2016) (Cao et al., 2019) (Zhu et al., 2020). These cracks can compromise bridge safety and structural integrity (Zhang et al., 2015). To ensure the safety and longevity of bridge structures, advanced technologies are essential for efficient and accurate identification and evaluation of fatigue cracks.

Traditional inspection techniques like visual, penetration, and magnetic particle methods face challenges in assessing complex orthotropic steel deck (OSD) structures (Thomas, 1995). Ultrasonic testing, a prevalent method for fatigue crack detection (Regazzi et al., 2014) (Bernasconi et al., 2016), is often time-consuming and may overlook minor cracks (Cheng et al., 2017) (Cui et al., 2018). To address these issues, Phased Array Ultrasonic Testing (PAUT) has been developed as an advanced technique with high accuracy in detecting fatigue cracks (Kainuma et al., 2017) (Bernasconi et al., 2022). PAUT utilizes multiple individually controllable elements to generate adjustable beam angles and focal distances, enhancing its scanning and detection capabilities for NDT non-destructive testing (NDT) (McNab and Campbell, 1987). Initially used in medical imaging (Havlice and Taenzer, 1979), PAUT has since been applied to concrete structures (Azar and Wooh, 1999) and has expanded into the engineering field. Recent research has integrated PAUT with neural network technology (Yang et al., 2013) and object detection techniques (Tunukovic et al., 2024), further advancing its applications.

PAUT is a versatile method that uses multiple individually controllable elements to create custom beam angles and focal distances, enhancing imaging capabilities for complex structures (Taheri and Hassen, 2019). It's particularly effective for crack depth identification (Wang et al., 2024) and offers benefits over other depth identification methods like liquid penetration testing (LPT) (Manikandan et al., 2021), magnetic particle testing (MPT) (Li et al., 2020), and eddy current testing (ECT) (Meng et al., 2021), such as portability, speed, accuracy, and automation (Chen et al., 2024). The advent of deep learning (LeCun et al., 2015), especially Convolutional Neural Networks (CNNs), has

* Corresponding author. Department of Bridge Engineering, School of Civil Engineering, Southwest Jiaotong University, Chengdu, Sichuan, 610031, China.
E-mail address: gouhongye@swjtu.edu.cn (H.-y. Gou).

https://doi.org/10.1016/j.iintel.2024.100113
Received 9 May 2024; Received in revised form 6 August 2024; Accepted 21 August 2024
Available online 30 August 2024

significantly improved performance in various engineering areas (Pak and Kim, 2017), including crack detection (Cha et al., 2017), fatigue crack detection (Zhang et al., 2021), and crack depth identification (Laxman et al., 2023). Techniques like wavelet packet transform (Chenglong et al., 2016) and wavelet coefficients (Meng et al., 2017) have also been used to improve detection efficiency.

While deep learning combined with PAUT has been widely applied in crack detection and depth identification (Munir et al., 2019) (Cruz et al., 2017) (Medak et al., 2021), most research still uses a two-stage process (Soviany and Ionescu, 2018). This study introduces a one-stage method that integrates region proposal and object detection within a single neural network, streamlining the process, reducing computational complexity, and improving real-time performance and accuracy, particularly for small and overlapping cracks. This approach is straightforward to implement, potentially enhancing crack detection in steel bridge decks, advancing automated inspection technology, and fulfilling field detection requirements.

Despite the growth in datasets for fatigue cracks (Bianchi and Hebdon, 2022), PAUT datasets are still not sufficiently diverse, necessitating data augmentation (Uhlig et al., 2023). PAUT, a prevalent C-scan method, can apply 2D image augmentation to its scans (Shorten and Khoshgoftaar, 2019). While the finite element method can create synthetic ultrasonic data (Liu et al., 2002), its use in C-scans is limited. GANs, when used wisely, can produce larger datasets (Sekar and Perumal, 2022). GANs have been successful in image style transfer (Zhu et al., 2017), restoration (Liu et al., 2018), and generation (Tian et al., 2021). They are widely used for ultrasonic image generation and processing (Zhang et al., 2018; Nair et al., 2019), including B-scan (Posilović et al., 2022), TOFD (Jiang et al., 2024), and PAUT images (Molinier et al., 2023). DCGAN, a GAN variant, is also popular for dataset creation (Song et al., 2024). It improves efficiency over Pix2Pix GANs (Isola et al., 2017) by replacing pooling layers with strided convolutions, removing fully connected layers, and adding batch normalization. DCGAN is frequently used in engineering to generate crack images (Pei et al., 2021) (Padhi et al., 2022), but its use in PAUT images is less common. This paper uses DCGAN to address the shortage of PAUT crack images, optimizing the hyperparameter num_g_updates for balanced training and stable, quick image generation.

This paper introduces a deep learning-based method for the automated detection of fatigue cracks in orthotropic steel bridge decks, crucial for preventing structural damage and ensuring bridge safety. The method combines region proposal and object detection within a unified neural network, using the YOLO series algorithms for efficient, single-stage detection suitable for real-time field use. By incorporating the echo characteristics of crack regions, it intelligently identifies crack depth, improving the detection of small and overlapping cracks. This approach automates PAUT data analysis, reducing reliance on inspector expertise for more efficient and accurate evaluations. Experiments show that the method's accuracy is within a 5% relative error compared to Time Of Flight Diffraction (TOFT) tests. The research contributes a novel one-stage detection method, enhancing bridge structure safety and longevity.

The remainder of the paper is organized as follows. Section 2 describes the background bridge and detection equipment composed of a Phased Array Ultrasonic Probe and inspection robot. Section 3 proposes a DCGAN model to generate new images in PAUT form. Section 4 compares different YOLO-series model performances with attention mechanisms and explains in detail the implementation process of fatigue crack depth identification and subsequent result correction. Finally, Section 5 includes the conclusion and references cited throughout the paper.

2. Overview

With the increase in service time, under cyclic loading, cracks at the welds, U-ribs, and other positions of the orthotropic steel bridge deck are inevitable. The propagation of cracks will inevitably affect the service performance and lifespan of the structure, and in severe cases, can lead to loss of life and property.

To achieve rapid and high-precision detection of internal cracks, an automated inspection robot has been developed, which, in conjunction with ultrasonic phased array technology, enables unmanned scanning of internal cracks in OSD. The Deep Convolutional Generative Adversarial Network (DCGAN) is utilized to augment the PAUT crack images, while the YOLO series of algorithms is employed for crack localization and depth detection. The aforementioned equipment and algorithms facilitate high-precision localization and identification of cracks in dark environments and areas that are challenging for inspectors to reach. Fig. 1 shows the flowchart of the proposed method.

2.1. Bridge description

Yichang Yangtze River Highway Bridge has a total span of 1188 m and a main span of 960 m, making it a double-tower steel box girder suspension bridge. The bridge's cross-section is in the shape of a fish fin steel box, with the top plate of the steel box girder designed to incorporate additional short ribs, thereby forming an orthotropic steel deck (OSD). This bridge was initially constructed in 1997 and was put into operation in 2001. To date, it has been in service for over two decades. Fig. 2 shows the arrangement of the bridge.

2.2. Crack detection method

2.2.1. PAUT method

To address the research gap and the scarcity of publicly available datasets specifically designed for OSD, the authors of this study performed data collection during the maintenance period of the Yichang Yangtze River Highway Bridge. This study utilized Phased Array Ultrasonic Testing (PAUT) to conduct fatigue crack detection and establish a comprehensive dataset. Simultaneously, Time-of-Flight Diffraction (TOFD) technology was employed on-site for fatigue crack detection. TOFD is a sensitive and accurate method for nondestructive testing of welds, calculating crack depth by measuring pulse travel time and applying trigonometry. Based on the TOFD inspection results, the PAUT data collection primarily focused on the connection between the steel box girder's top deck and longitudinal ribs.

It is worth noting that the orthotropic steel bridge deck commonly exhibits four distinctive types of fatigue cracks: crack_a, crack_b, crack_c, and crack_d, as illustrated in Fig. 3, with the weld joint between the U-rib and the top deck being particularly prone to fatigue-related damage (Wang et al., 2019).

During the data collection process, meticulous attention was given to the selection of appropriate detection parameters, ensuring the efficacy and accuracy of the ultrasonic phased array technique. Table 1 provides a comprehensive overview of the experimental parameters employed for the ultrasonic testing setup.

2.2.2. Detection robot

Although the orthotropic steel deck (OSD) structural system endows the structure with superior mechanical properties, it also inevitably leads to the generation of a large number of welds. These welds are highly susceptible to fatigue cracking during the bridge's service life, posing a significant threat to the safety of bridge operations. However, the interior environment of the steel box girder is dark, making it difficult for maintenance personnel to operate. Traditional image-based crack detection methods have low accuracy. Therefore, a method for rapid crack detection in bridge structures is proposed, this method involves the use of an intelligent inspection robot equipped with an ultrasonic phased array probe. Fig. 4 illustrates the structure of the inspection robot. The U-rib design of the OSD structure is compact, with a central spacing of only 600 mm. Therefore, the robot's locomotion system employs a meticulously designed electromagnetic adhesion

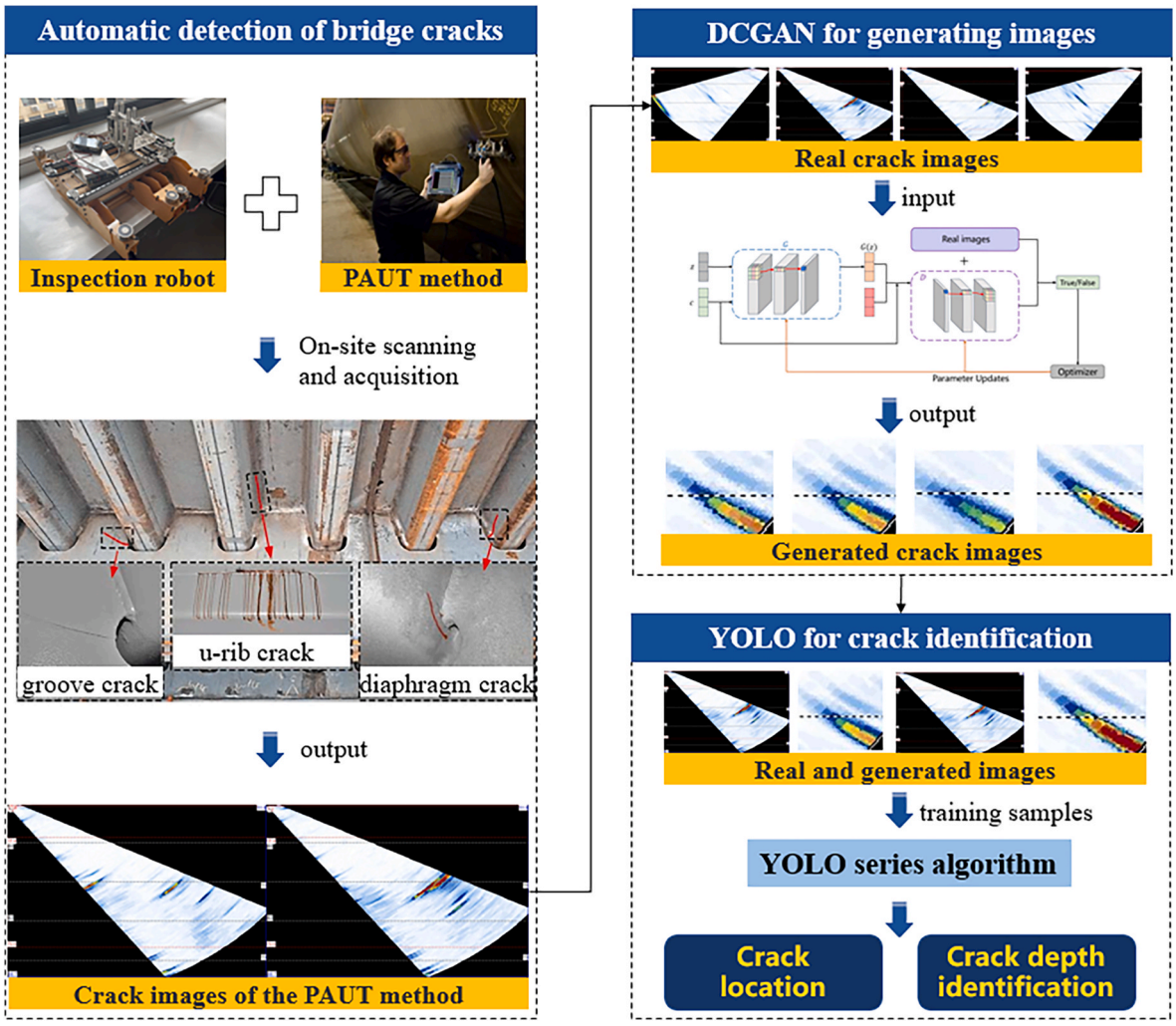


Fig. 1. Flowchart of the proposed framework using automatic detection equipment and deep learning algorithms.

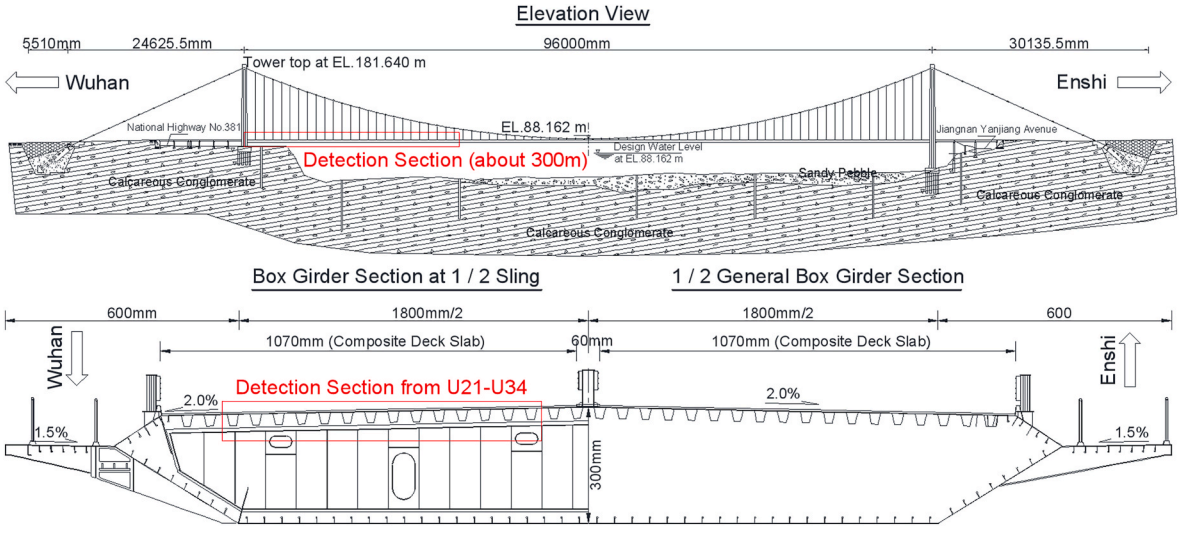


Fig. 2. The layout of spans and the cross-section of the main girder.

mechanism to ensure stable movement of the robot on the U-ribs of the OSD structure. The operation of the electromagnetic coils, which involves on-power adhesion and off-power release, enables the robot to advance in a "creeping" motion similar to that of living organisms. This mode of movement is highly suitable for complex and confined environments within bridge structures.

In Fig. 4(b), the robot features a running system and a probe-holding system. As it advances, the electromagnet group 2 demagnetizes and

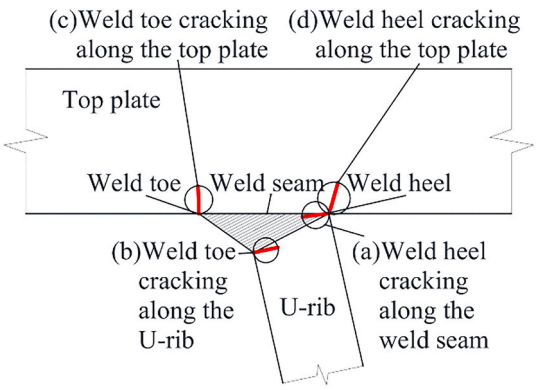


Fig. 3. Common forms of fatigue cracks.

Table 1
Experimental parameters for ultrasonic testing setup.

Parameters	Values	Parameters	Values
Probe Model	5L16-0.6×10-A10-P-2.5-OM	Probe Frequency	5MHz
Probe Type	Linear Array Probe	Fan Scan Angle	42°–72°
Frequency	5 MHz	Pulse Repetition Rate	1000
Chip Number	16	Detection Medium	Carbon Steel
Wafer Spacing	0.6	Material Sound Speed	5900 m/s
Wafer Width	10	Wedge Model	SA10-N55S

shifts forward along the rod. Then, group 2 magnetizes, and groups 1 and 3 are powered off, with the motor propelling groups 1 and 3 forward along the rod. The probe moves forward and backward along the motion axis. Adjusting the screw height alters the elevation of the probe. In an enclosed environment, the robot establishes a connection with external systems via Bluetooth, facilitating remote control capabilities.

Fig. 4(c) illustrates the detection range of the inspection robot. The inspection area includes the u-rib side wall above the probe, the complete top plate, and the diaphragm. By moving along the guide shaft, the probe constructs a V-shaped scanning area, and then by moving along the scan shaft, the V-shaped area moves transversely, achieving full coverage of the top plate and the diaphragm.

2.2.3. Crack distributions

We analyzed the crack patterns from location (a) to (d) based on multiple on-site inspection images, with the typical crack distributions at different locations illustrated in Fig. 5. Among these, weld toe cracking along the U-rib exhibits the least crack development; weld heel cracking along the weld seam is the next in severity; weld toe cracking along the top plate typically features two relatively deep cracks; whereas weld heel cracking along the top plate presents a single crack with significant depth and width. Ordered by severity of cracking, from the most to the least severe: crack_d > crack_c > crack_a > crack_b.

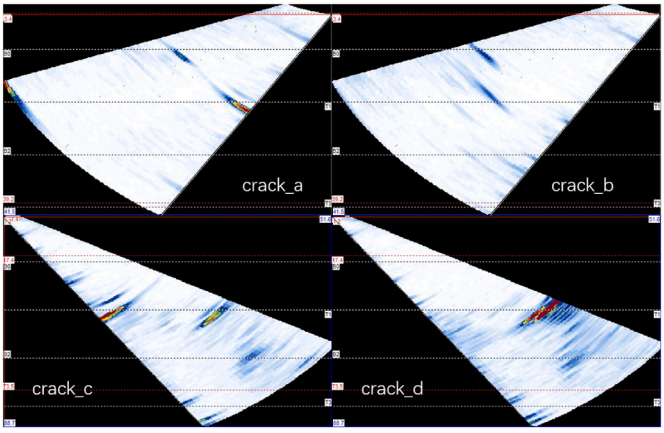


Fig. 5. Typical pattern of the crack distribution of OSDs.

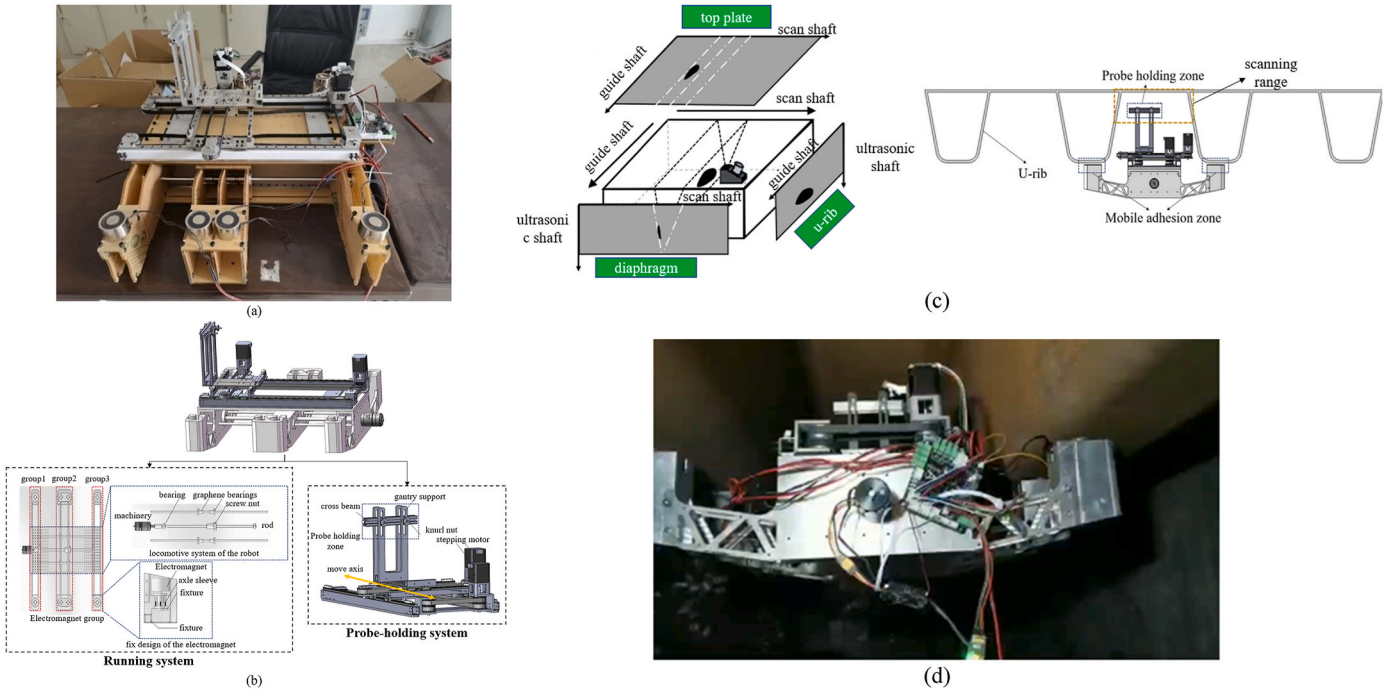


Fig. 4. Crack detection equipment: (a) Inspection robot; (b) Schematic diagram of robotic components; (c) inspection range of the robot; and (d) on-site inspection.

3. Deep Convolutional Generative Adversarial Network (DCGAN)

3.1. Structure of the DCGAN

The DCGAN model includes a generator and a discriminator. Both the generator and discriminator are composed of DCNNs. The architecture of the DCGAN is shown in Fig. 6. Both the generator and the discriminator have abandoned the traditional pooling layers used in CNN. The discriminator retains the overall architecture of the CNN, while the generator replaces the convolutional layers with transposed convolutional layers. Batch Normalization (BN) layers are utilized subsequent to each layer in both the discriminator and the generator of the DCGAN framework. This practice aids in addressing training issues stemming from poor initialization, accelerates the convergence process, and mitigates the risk of overfitting. In the generator, ReLU activation is used in all layers except for the output layer, which uses the Tanh function. The discriminator applies Leaky ReLU across all layers to avoid gradient vanishing. To prevent mode collapse, we have defined the hyperparameter num_g_updates to mitigate the instability in training that may arise from an overly dominant **G** or **D**. Introducing additional stochastic noise aids the generator in evading entrapment within specific patterns, thereby enhancing the model's capacity to generate a diverse array of images.

3.2. Hyper-parameters

The DCGAN is designed to generate images in the size of 256 × 256, with a kernel size of 4 × 4. After multiple training and validation sessions, the input_dim is set to 100, the output_dim is set to 1, and the channels is set to 3. To prevent either the generator or the discriminator from becoming arbitrarily dominant, the parameter num_g_updates was set, such that for every iteration of discriminator training, the generator is trained multiple times, which enables DCGAN to converge quickly with a small number of epochs. This parameter is established at a value of 12 and Adam is used as the optimizer. The main parameters of the DCGAN model are shown in Table 2. The DCGAN employs binary cross-entropy as the loss function.

3.3. Image generation and evaluation

In light of the presence of multiple features such as scanned areas, un-scanned areas, and scanning lines in each PAUT image, we have extracted each crack from the scanned images into individual pictures. This approach ensures that each resulting image contains only the crack and the background. This will enable the model to focus more on the characteristics of the cracks rather than dealing with the complexity of multiple cracks simultaneously. Additionally, this method allows for a more direct assessment of the quality of the generated images. The training set was partitioned into 2776 crack images and 1000 noise images, with each crack image being rotated three times (Guo et al., 2022). After establishing and segmenting the dataset for the DCGAN model, as well as adjusting the parameters, the DCGAN model is trained. Furthermore, four images are generated every 10 epochs. By analyzing the generated images, it is observed that before 50 epochs, the images consisted of randomly scattered color blocks. Subsequently, there is a significant improvement in image quality. After 100 epochs, the image quality approaches that of real images. As iteration progresses, the generated images increasingly resemble reality and demonstrate enhanced diversity. The training process of the DCGAN is depicted in Fig. 7.

To evaluate the model performance and image quality, IS (Inception Score) and FID (Fréchet Inception Distance) are utilized. Inception Score (IS) is a scoring method based on the Inception network (a pre-trained image classifier), which is used to measure the diversity of generated images and the accuracy of classification. The higher the value of IS, the better the quality of the generated image. As shown in Eq. (1).

$$IS = e^{E_{x\sim P_g}[KL(P(y|x)\|P(y))]} \tag{1}$$

Where KL represents Kullback Leibler divergence, an indicator to measure the difference between two probability distributions, $PM(y|x)$ is the category distribution predicted by Inception network for the given image x, $P(y)$ is the uniform distribution of the generated image on all categories in the ideal case, and E represents the expected value, that is, the average value of KL divergence is calculated for all generated images.

Fréchet Inception Distance (FID) is an evaluation method of feature space distance based on the Inception network, which is used to measure the gap between the generated image and the real image. The lower the FID value, the closer the generated image is to the real image. As shown in Eq. (2).

$$FID(r,g) = \left\|\mu_r - \mu_g\right\|^2 + T_r\left(C_r + C_g - 2\left(C_r C_g\right)^{\frac{1}{2}}\right) \tag{2}$$

Where r means real and g means generated, μ denotes the mean value, C represents the covariance, and T represents represents the sum of the elements on the diagonal of the matrix.

Fig. 8 illustrates the model performance based on IS and FID scores. As the training iterations increase, the Inception Score (IS) gradually rises, reaching around 40 at 200 epochs. Subsequently, the IS value tends to stabilize around 40, with a maximum value of 41.27, indicating that the images generated by the DCGAN exhibit diversity. The Fréchet Inception Distance (FID) plummets before 100 epochs and then decreases slowly, eventually stabilizing at 200 epochs, with a minimum value of 7.03. This value suggests that the generated images closely resemble real images.

4. Crack identification by YOLO

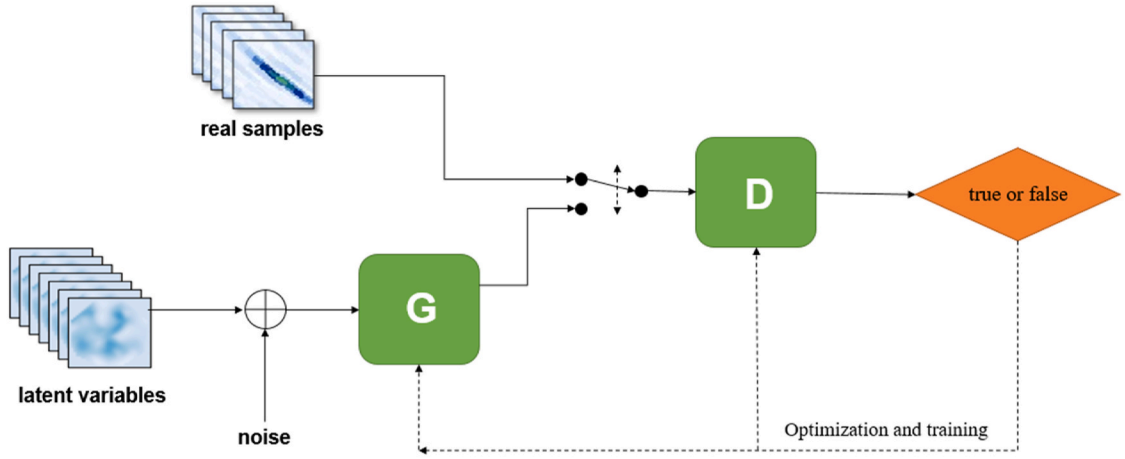
This paper addresses the need for rapid on-site detection within the constraints of mobile device performance by employing a single-stage detection method. The SSD algorithm, while popular, has shown limitations with small objects (Jiao et al., 2019). Thus, we opt for the YOLO algorithm to ensure efficient and accurate detection. Incorporating negative samples during training enables the model to learn to discriminate between relevant fatigue crack regions and irrelevant features.

4.1. YOLO series algorithm

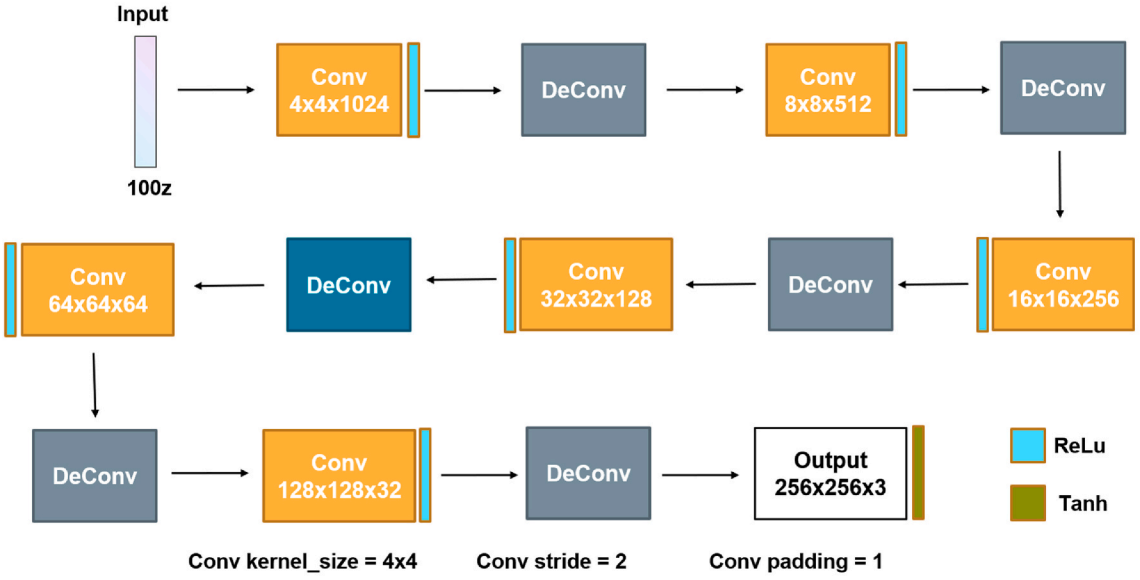
4.1.1. YOLOv7-tiny

YOLO is an acronym for the You Only Look Once target detection algorithm (Wang et al., 2023). To surmount the computational limitations of mobile devices for fatigue crack detection, the model must be streamlined by minimizing parameter count and computational load. The YOLOv5s model, a compact iteration of the YOLOv5 series, incorporates the CSP-DarkNet framework and integrates a Path Aggregation Network (PANet) to optimize feature fusion. YOLOv7 employs several specialized techniques for real-time object detection, such as cross-layer connections and multi-scale prediction, to ensure swift and accurate object detection (Gallo et al., 2023). Additionally, the YOLOv7-tiny model, designed for common GPUs and noted for its rapid average-case execution, is a lightweight adaptation within the YOLOv7 suite. This model adopts the leaky Rectified Linear Unit (ReLU) activation function, as expressed in Eq. (3) (Bochkovskiy et al., 2020), while the YOLOv7 model employs the SiLU (Sigmoid Linear Unit) activation function, as delineated in Eq. (4) (Elfwing et al., 2018), for analogous purposes.

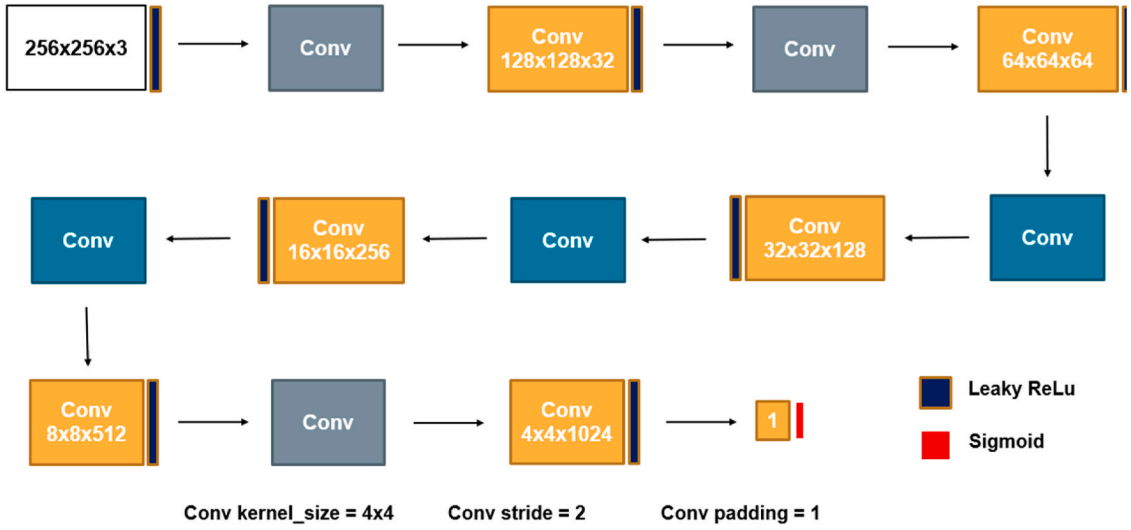
$$\mathrm{f}(x) = \begin{cases} x, x \geq 0 \\ negative\_slope \bullet x, \text{otherwise} \end{cases} \tag{3}$$



(a)



(b)



(c)

Fig. 6. DCGAN model: (a) overall structure of the DCGAN model; (b) structure of the generator; and (c) structure of the discriminator.

Table 2
Parameters of the DCGAN model.

Parameter	Value	Meaning
input_dim	100	generator input dimension
Output_dim	1	discriminator output dimension
epoch	600	Training cycling number
channels	3	color channels (R, G, B)
g_lr	0.0002	learning rate of generator
d_lr	0.0005	learning rate of discriminator
batach_size	32	the size of a batch
num_g_updates	12	number of generator updates
Adam_Beta1	0.5	Parameter for the Adam optimizer

where *negative_slope* is a small constant (In this paper, the value is 0.01).

$$f(x) = x \cdot \frac{1}{1 + e^{-x}} \tag{4}$$

Although not as accurate as two-stage methods like Fast R-CNN (Ren et al., 2016), the YOLOv7-tiny model boasts fewer parameters and higher detection efficiency, which is desirable in the application (Yuhas and Easwaran, 2022). Fig. 9 illustrates the neural network architecture of the YOLOv7-tiny model.

4.2. Training and evaluation

The training process of the YOLOv7-tiny model is composed of the following key components.

4.2.1. Dataset partition

The data set is randomly partitioned into three distinct subsets, namely the training, validation, and test sets, adhering to the ratio of 6:2:2 (Rácz et al., 2021). The training set functions to update the model parameters, while the validation set is employed to supervise the training progress and optimize the hyperparameters. On the other hand, the test collection is used to evaluate the model performance on novel data instances that the model has not encountered before. Fig. 10 shows the annotation schematic and data set partition.

Randomly partitioning the dataset is essential for the model's ability to generalize effectively to novel data and to attain peak performance. This method prevents biases in the training, validation, and test subsets that could adversely affect the model's accuracy. Consequently, the random division of the dataset is a key element in the training process and significantly influences the YOLOv7 model's effectiveness in identifying fatigue cracks within orthotropic bridge decks.

4.2.2. Loss function

The loss function is adjusted to minimize the difference between the predicted bounding boxes and the ground truth labels. Given the high variance in the size and shape of fatigue cracks, the YOLOv7 model uses the loss function consisting of three components: coordinate loss, object confidence loss, and class loss (Zheng et al., 2022).

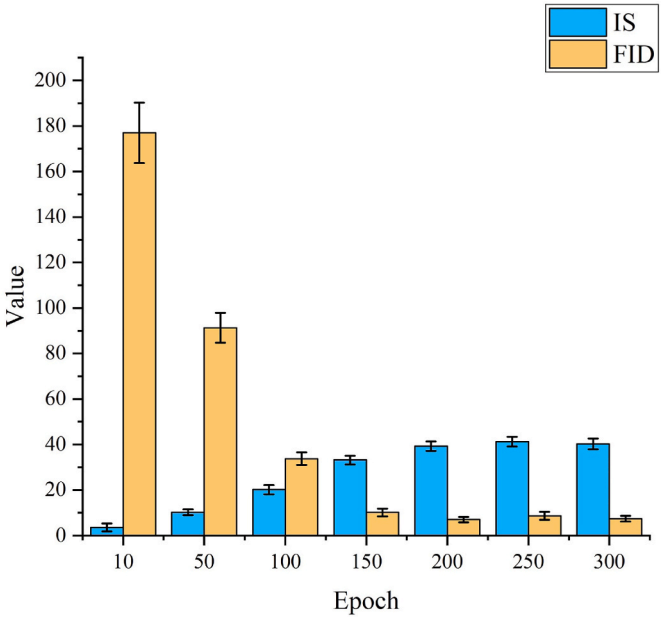


Fig. 8. The DCGAN model performance based on IS and FID scores.

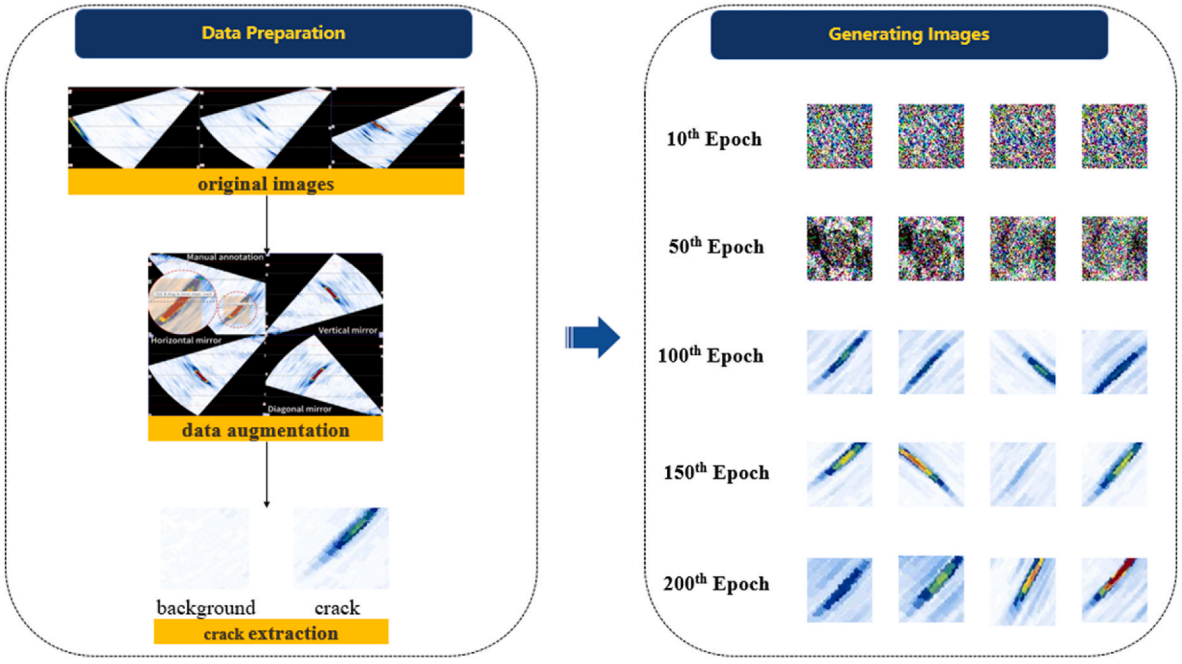


Fig. 7. The training process of the DCGAN model.

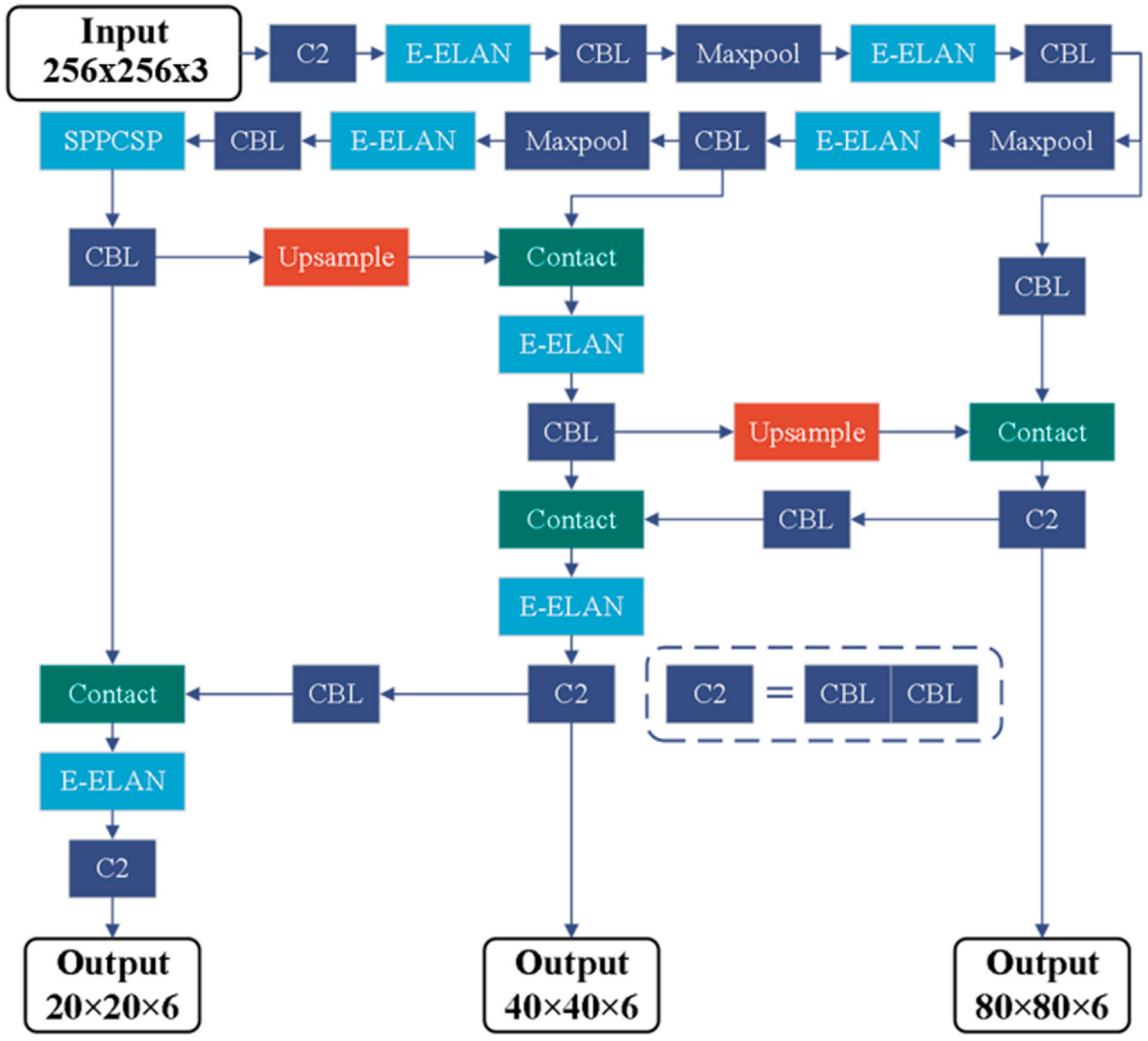


Fig. 9. Neural network structure diagram of YOLOv7-tiny model.

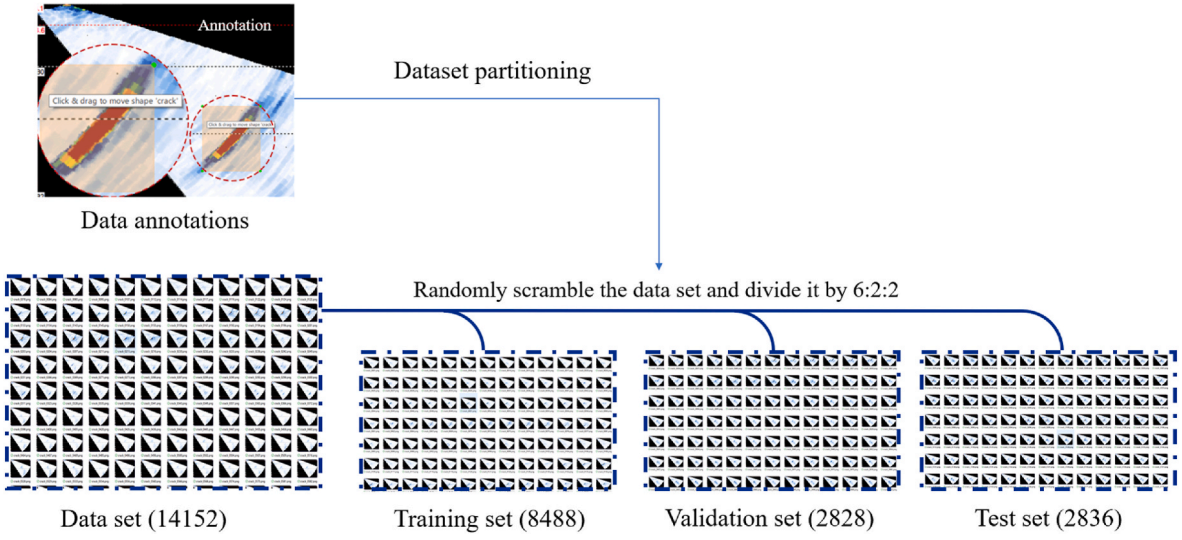


Fig. 10. Annotations and partition of the data set.

1. Coordinate Loss (CIoU Loss)

$$\text{Loss} = 1 - CIoU \tag{5}$$

Where:

$$CIoU = \frac{inter\_area}{union\_area - inter\_area} \tag{6}$$

With:

$$inter\_area = \min\left(\omega^{gt}, \omega^{p}\right) \times \min\left(h^{gt}, h^{p}\right) \tag{7}$$

$$union_{area} = \omega^{gt} \times h^{gt} + \omega^{p} \times h^{p} - inter\_area \tag{8}$$

Where ω^{gt}, h^{gt} are the width and height of the ground truth bounding box, ω^{p} and h^{p} are the width and height of the predicted bounding box.

2. Object Confidence Loss (Binary Cross-Entropy)

$$\text{Loss} = -\left[y^{obj}\log\left(p^{obj}\right) + \left(1 - y^{obj}\right)\log\left(1 - p^{obj}\right)\right] \tag{9}$$

where y^{obj} is the binary indicator for whether an object is present in the grid cell, and p^{obj} is the predicted object confidence score.

3. Class Loss (Binary Cross-Entropy)

$$\text{Loss} = -\sum_{c=1}^{C} y^{c}\log(p^{c}) + (1 - y^{c})\log(1 - p^{c}) \tag{10}$$

where C is the number of classes, y^{c} is the binary indicator for the presence of the true class, and p^{c} is the predicted class probability.

4.2.3. Training parameters

To optimize performance in detecting fatigue cracks in orthotropic steel bridge deck panels with the YOLOv7 model, this study rigorously tuned the hyperparameters detailed in Table 2. These parameters include the learning rate, batch size, weight decay, and momentum. Two optimizers are compared: SGD (Robbins and Monro, 1951) with a learning rate of 0.01, and Adam (Kingma and Ba, 2014) with a learning rate of 0.001. Adam, known for its adaptive learning rates, has typically surpassed SGD in performance across various models, as noted in prior research (Zhang et al., 2019). Through an exhaustive grid search (Liashchynskyi and Liashchynskyi, 2019), a weight decay of 0.0005 and momentum of 0.937 are identified as optimal, enhancing both the model's accuracy and efficiency. Furthermore, advanced data augmentation techniques, including random cropping, flipping, and rotation, are extensively utilized to expand the dataset and bolster the model's generalization capabilities. Table 3 shows the training parameters.

During the training process, the model performance is evaluated using the mean Average Precision (mAP) metric, which factors in the precision, recall, and confidence scores of the model's predicted bounding boxes. The training process entails partitioning the dataset into distinct subsets for training, validation, and testing. It also involves fine-tuning the loss function, optimizing hyperparameters, and assessing the model's accuracy via the mAP. These methodical steps are paramount in enhancing the model's capacity to accurately detect fatigue cracks in orthotropic bridge decks.

The deep learning models were trained on a VM provided by Google Colab. The VM was equipped with an NVIDIA A100-SXM GPU (40G) and 12 Intel(R) Xeon(R) CPUs (2.20 GHz). The PyTorch version used was 1.13.1+cu116. The training environment proved to be well-suited for the conducted experiments, providing reliable computing capabilities for training the deep learning models.

Table 3
Parameters of the YOLO model.

Parameters	Values	Parameters	Values
Optimizer	Adam	Learning rate	0.001
Momentum	0.937	Weight decay	0.0005
Batch size	32	Image size	640*640
Epoch	300	Pretrained	MS COCO

4.2.4. Model evaluation

Performance evaluation is a critical step in developing an intelligent identification model. The metrics used in performance evaluation determine how well the model can detect and classify crack types (Badithela et al., 2022). The following metrics are commonly used in evaluating the performance of object detection models:

Precision and Recall:

$$\text{Precision} = \frac{TP}{TP - FP} \tag{11}$$

$$\text{Recall} = \frac{FP}{TP + FP} \tag{12}$$

F_β score is a variant of F-measure that allows more weight to be given to either precision or recall, depending on the value of $\boldsymbol{\beta}$ (Fawcett, 2006) (Zaidi et al., 2022). It is calculated as Eq. (13):

$$F_\beta = \left(1 + \beta^2\right) \bullet \frac{precision \bullet recall}{\left(\beta^2 \bullet precision\right) + recall} \tag{13}$$

$F1$ score, also known as the Balanced Score, is a specific case of F_β score where β = 1 as Eq. (14), which means that precision and recall are given equal weight:

$$F_1 = 2 \bullet \frac{precision \bullet recall}{precision + recall} \tag{14}$$

Intersection over Union (IoU) and mean Average Precision(mAP): IoU measures the overlap between the predicted bounding boxes and the ground truth bounding boxes (Fawcett, 2006). It is calculated as Eq. (15):

$$\text{IoU} = \frac{Area\ of\ Intersection}{Area\ of\ Union} \tag{15}$$

Mean Average Precision (mAP) is a prevalent metric for assessing performance in object detection tasks. This manuscript employs mAP@0.5:0.95 as the evaluation criterion, which is derived by averaging the mAP values across Intersection over Union (IoU) thresholds ranging from 0.5 to 0.95 in increments of 0.05, as articulated in Eq. (16):

$$\text{mAP@0.5 : 0.95} = \frac{1}{11}\sum_{t=0.5}^{0.95} \text{mAP}_t \tag{16}$$

During the process of training an intelligent identification model for fatigue cracks in orthotropic steel bridge decks, we employed the YOLOv7 and YOLOv7-tiny models with Adam and SGD optimizers, respectively, on a previously established dataset of ultrasonic phased array detection results. To assess the effectiveness of the approach, various metrics, including precision, recall, F1 score, and mAP@0.5:0.95, were computed for both the training and testing phases. Table 4 illustrates the model performance of different YOLO algorithms.

The YOLOv7-tiny model's results are superior in terms of precision, indicating a lower rate of false positives. It also excels in recall, effectively identifying a wide range of true positives. The model's highest F1 score reflects a balanced performance between precision and recall.

Table 4
Model performance of the YOLO algorithms.

Model	Precision	Recall	F1 score	mAP@0.5:0.95
YOLO v7	89.62%	80.14%	0.85	56.18%
YOLO v7-tiny	90.64%	89.36%	0.90	56.57%

When these models are deployed on mobile devices, significant variations in identification speeds are noted. For instance, YOLOv7 necessitates approximately 300 ms to process a single image on an M1 chip. In contrast, YOLOv7-tiny accomplishes the task in merely 50 ms. This discrepancy culminates in a sixfold disparity in identification velocity between YOLOv7-tiny and the other models. Consequently, the more efficient YOLOv7-tiny model is favored for detection tasks.

Fig. 11 presents the experimental outcomes utilizing the YOLOv7-tiny model for identification. Fig. 11 (d) illustrates the echo intensity bar, which increases in intensity from left to right. This study empirically determines a 40% echo intensity threshold for identifying fatigue cracks. The region of the bar associated with fatigue cracks corresponds to the colors beyond the yellow marker on the right. Fig. 11 (a) to Fig. 11 (c) confirm the model's effective detection of multiple fatigue cracks, with labels on the bounding boxes indicating crack followed by a numerical confidence level, signifying the likelihood of the region being a fatigue crack.

Overall, the YOLOv7-tiny model achieves higher recall without compromising the precision and mAP@0.5:0.95 values, as it can be a suitable option for fatigue crack identification in orthotropic steel bridge deck panels.

4.2.5. Attention mechanism

The attention mechanism in deep learning allows models to focus on salient input features while filtering out the irrelevant ones. In the context of the fatigue crack identification model, several attention mechanisms are integrated to boost effectiveness. The Convolutional Block Attention Module (CBAM) (Woo et al., 2018) sequentially refines attention maps across channel and spatial dimensions, enhancing feature representation. Coordinate Attention (CA) (Hou et al., 2021), meanwhile, integrates positional information into channel attention to emphasize crucial spatial areas. Additionally, Omni-Dimensional Dynamic Convolution (ODDC) (Li et al., 2022) utilizes a multi-dimensional attention mechanism across convolutional kernels to achieve superior feature representation. These attention mechanisms collectively improve the model's performance, rendering it highly beneficial for fatigue crack identification tasks. Fig. 12 illustrates the structure of attention mechanisms.

As shown in Fig. 12 (a), the Coordinate Attention (CA) mechanism is composed of two key processes: incorporating coordinate information and creating attention maps. When integrated into the On-Surface Defect (OSD) fatigue crack detection model, the CA mechanism processes the Phased Array Ultrasonic Testing (PAUT) images to produce attention maps for both horizontal and vertical orientations. These maps

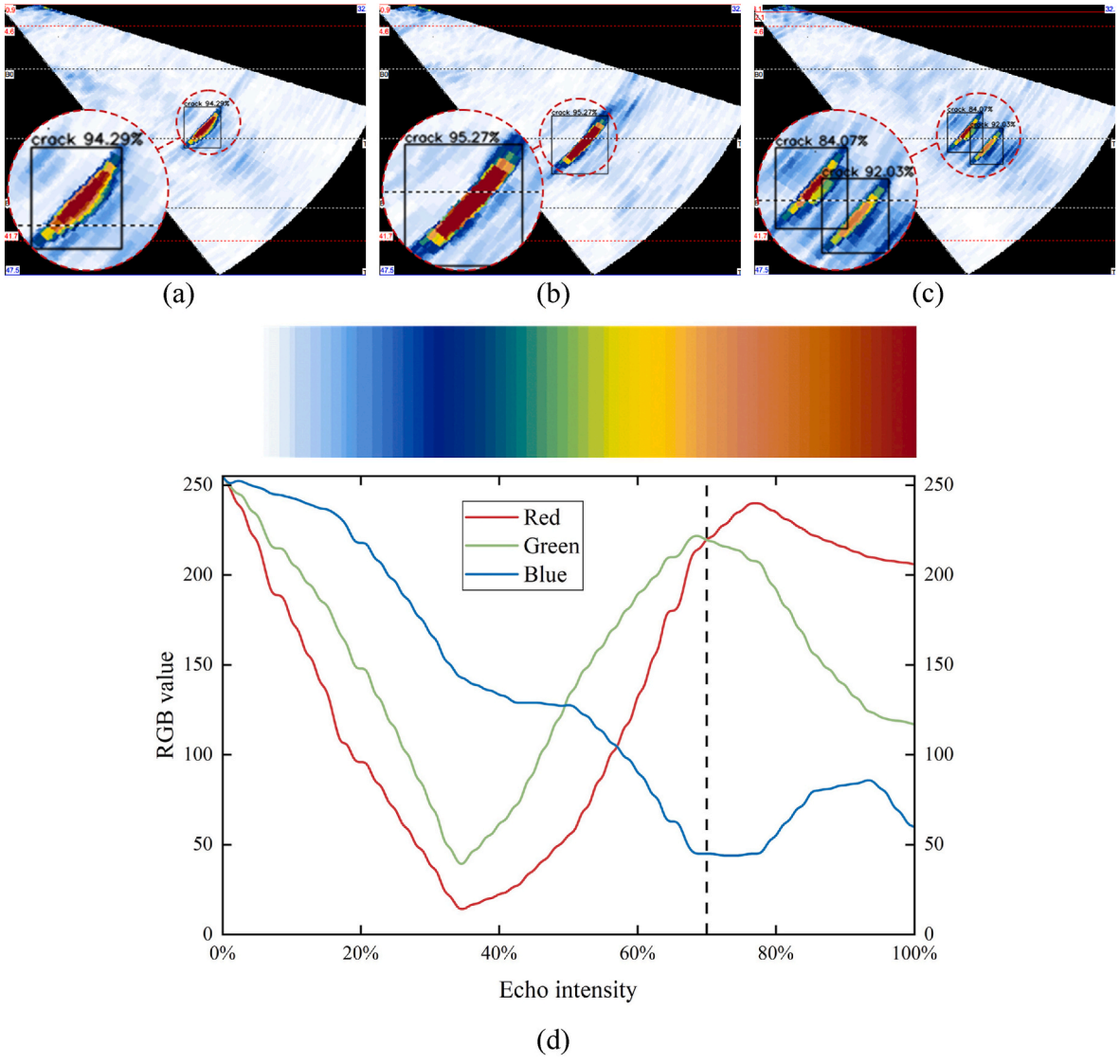


Fig. 11. Fatigue crack identification results using YOLOv7-tiny model: (a) identification results 1; (b) identification results 2; (c) identification results 3; and (d) echo intensity bar.

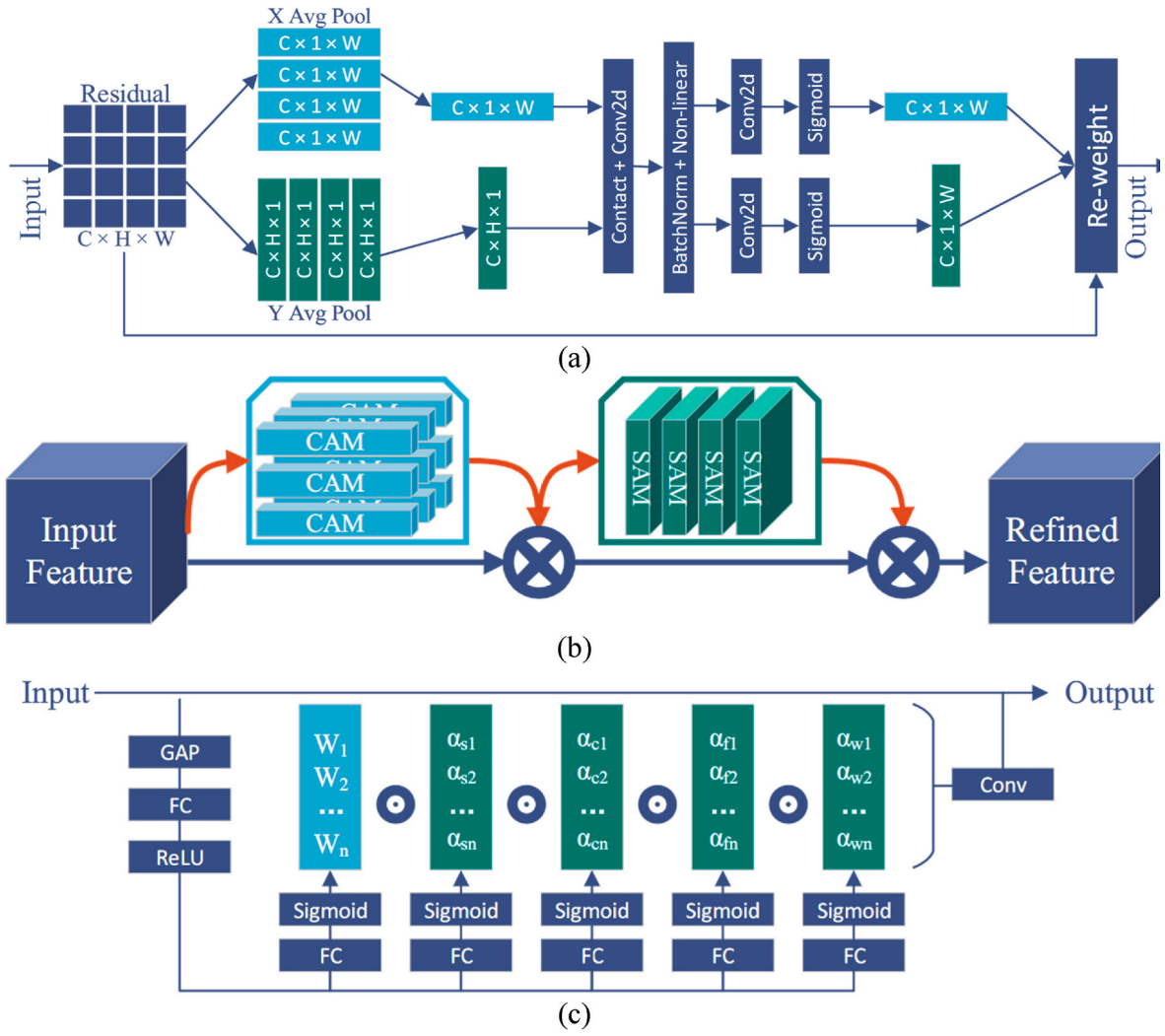


Fig. 12. Structure of attention mechanisms: (a) CA; (b) CBAM; and (c) ODDC.

correspond to arrays in the form of $C \times 1 \times W$ and $C \times H \times 1$. By weighting the original feature map with these attention maps, the model focuses more on features critical to identifying fatigue cracks, thereby improving its ability to detect cracks in PAUT images.

The CBAM attention mechanism, shown in Fig. 12 (b), is composed of two key parts: the Channel Attention Module (CAM) and the Spatial Attention Module (SAM). When incorporated into the OSD fatigue crack detection model, the CAM first evaluates the three color channels of PAUT images to find those most relevant to fatigue cracks, creating a feature map with enhanced importance for these channels. This map is then combined through a multiplication operation and sent to the SAM, which refines the analysis of where fatigue cracks are located within the images. This step-by-step process leads to the intelligent recognition of fatigue cracks in PAUT inspections.

Fig. 12 (c) illustrates that the ODDC attention mechanism is centered around dynamic convolution learning and the integration of multi-directional features. Within the OSD fatigue crack detection model, the ODDC mechanism begins by analyzing the PAUT images through dynamic convolution, which adaptively adjusts feature weights for a weighted summation. It then extracts features along various orientations. The ODDC mechanism subsequently fuses these features with the original feature map via point-wise multiplication, focusing the model's attention on fatigue crack characteristics and thereby improving its recognition capabilities.

To evaluate the effectiveness of attention mechanisms for fatigue crack detection, the YOLOv7-tiny model was augmented with CA, CBAM, and ODDC. As shown in Table 5, the P-R curves of different models are shown in Fig. 13. The integration of these mechanisms significantly improved the model's detection performance.

Table 5
Model performance of the improved YOLO algorithms.

Model	Precision	Recall	F1 score	mAP@0.5:0.95
YOLO v7	89.62%	80.14%	0.85	56.18%
YOLO v7-tiny	90.64%	89.36%	0.90	56.57%
YOLO v7-tiny-CA	90.00%	89.35%	0.90	58.85%
YOLO v7-tiny-CBAM	94.07%	90.00%	0.92	61.76%
YOLO v7-tiny-ODDC	95.56%	92.18%	0.94	62.92%

The results demonstrate that the integration of attention mechanisms into the YOLOv7-tiny model can enhance its performance in detecting fatigue cracks. The CBAM attention mechanism effectively captures local information, while the ODDC attention mechanism exhibits significant recall improvements by considering the complementarity of attention between dimensions. However, the CA attention mechanism did not effectively enhance the model's ability to detect fatigue cracks, mainly due to its focus on global dependent information rather than local information. In comparison with CA and CBAM, ODDC has demonstrated superior performance enhancement, with a precision of 95.56%, which is higher than the 93% accuracy of the fatigue crack

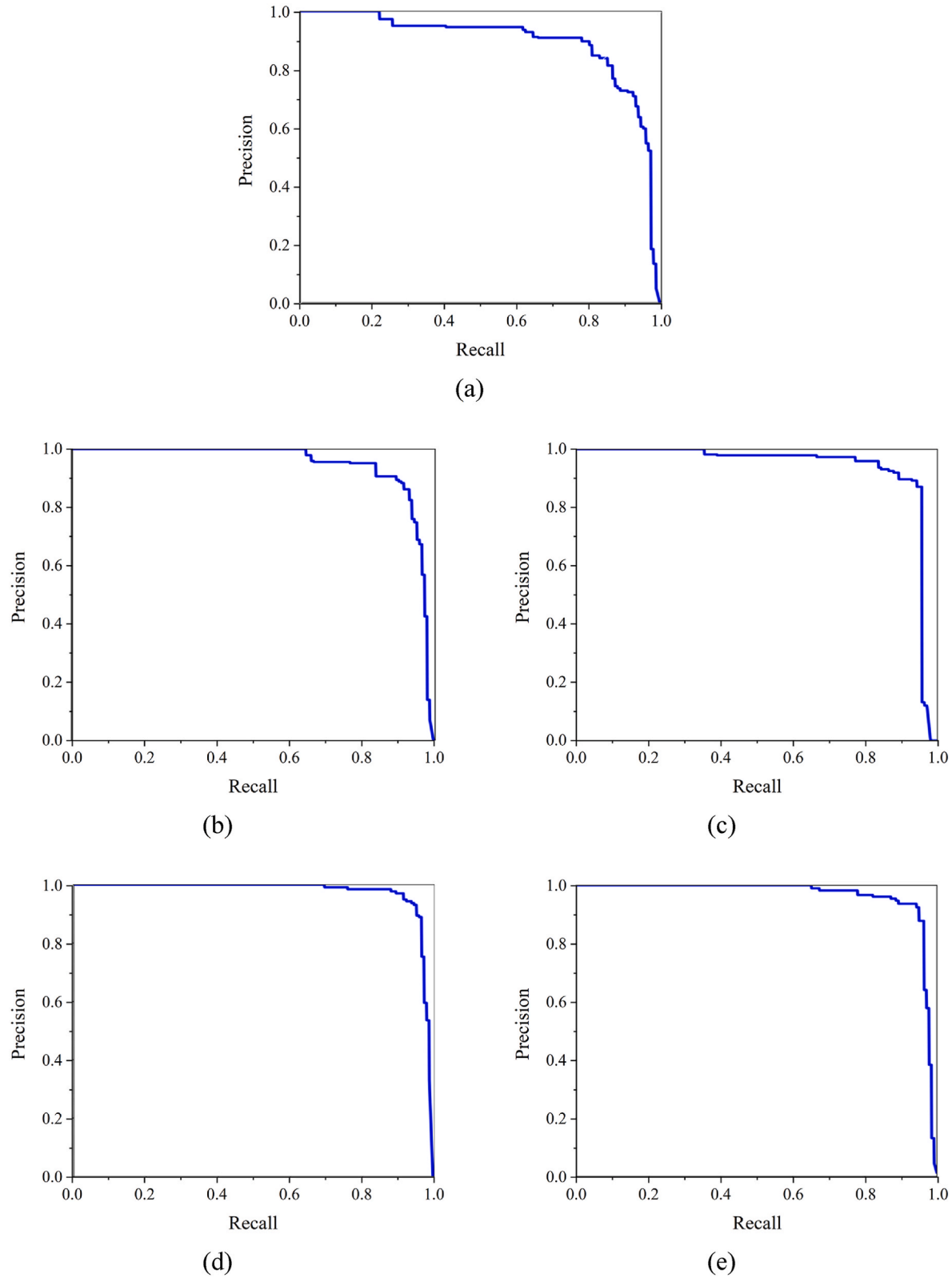


Fig. 13. P-R curves of the different models: (a) YOLO v7; (b) YOLO v7-tiny; (c) YOLO v7-tiny-CA; (d) YOLO v7-tiny-CBAM; and (e) YOLO v7-tiny-ODDC.

monitoring system based on shallow B-CNN configuration (Zhu et al., 2023). The model's detection capability at different confidence thresholds is on par with recently proposed methods (Zhai et al., 2022), with an average mAP0.5:0.95 slightly above 62%. This demonstrates the proposed method's superior reliability and applicability over current practices for precise fatigue crack depth identification in orthotropic steel bridge decks.

4.3. Crack depth identification

The fatigue crack depth identification method proposed in this study builds upon the initial identification of fatigue regions using the YOLOv7-tiny algorithm. Once the fatigue regions are detected, the algorithm proceeds to estimate the depth of the cracks based on the characteristics of the echo area.

4.3.1. Echo area characteristic

For estimating the depth of fatigue cracks, the study analyzes the echo area characteristics within the detected regions. Fig. 11 (d) presents an echo intensity bar chart that illustrates variations in echo intensity across the crack area. Given that ultrasonic phased array imaging of the target area, particularly the fatigue crack region, results in a diffuse pattern, precise identification of the crack region by echo intensity is essential. To this end, the Hue value for each color in the RGB mode is calculated using the formula detailed in Eq. (17), which aids in accurately delineating the fatigue crack region.

$$Hue = \begin{cases} \dfrac{G-B}{\max\{R,G,B\}-\min\{R,G,B\}} \text{if } R = max \\ \dfrac{B-R}{\max\{R,G,B\}-\min\{R,G,B\}} + 120 \quad \text{if } G = max \\ \dfrac{R-G}{\max\{R,G,B\}-\min\{R,G,B\}} + 240 \text{if } B = max \end{cases} \tag{17}$$

The Hue values for each color on the echo intensity bar, as represented in RGB coordinates were calculated and designated within their corresponding color segments. Analysis indicates that the yellow color with RGB values of (198, 210, 63) correlates with an echo intensity of 40%. Furthermore, a critical threshold of 70%, which is primarily derived from the extensive field inspection experience of practitioners, has been established to distinguish between non-crack and crack regions. The critical threshold HUE values are delineated as the 'Critical Line' in Fig. 14. The area to the right of the Critical Line in the detection result image signifies the presence of fatigue cracks.

4.3.2. Identification results

The proposed model has successfully achieved precise identification of fatigue crack regions and utilized an intelligent approach for depth determination. Building upon prior research, an extensive dataset of C-scan images derived from ultrasonic phased array inspections was assembled. As depicted in Fig. 15, the C-scan image encompasses a fatigue crack region, with the non-black area (RGB: 255, 255, 255) indicating the scanned angular range from 37° to 70°.

Utilizing the aforementioned information, an algorithm has been developed by the authors to intelligently ascertain the depth of fatigue cracks within the delineated regions. This algorithm has been refined to augment the precision of identification result labeling. It transitions from a reliance on identification confidence values to a methodology that incorporates the depth of the detected fatigue cracks, thereby yielding more accurate labeling results. The output is presented in Fig. 16. The algorithm-predicted fatigue crack depth of 9.9 mm notably

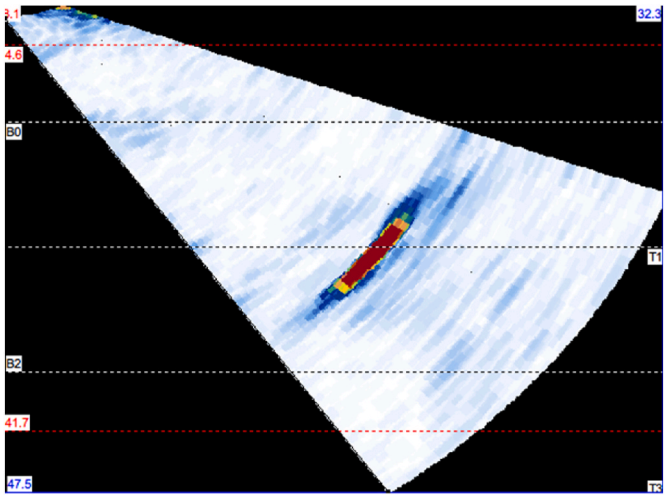


Fig. 15. C-scan image.

differs from the Time-of-Flight Diffraction (TOFD) measurement of 5.8 mm taken in the field, underscoring the need for further refinement of the identification process to enhance accuracy.

4.3.3. Correction

The C-scan image displays three white horizontal dashed lines labeled B0, T1, and B2, which represent the primary, secondary, and tertiary ultrasonic reflection interfaces, respectively. Fatigue cracks are typically located in the regions between B0 and T1, and between T1 and B2, known as the secondary and tertiary echo areas within the context of ultrasonic phased array detection. By integrating this information with the positioning of the ultrasonic phased array probe, a schematic of the scanning process on the top plate has been constructed, as depicted in Fig. 17.

In Fig. 15, fatigue cracks are observed propagating along the top plate, resulting in two successive echo images due to the inherent characteristics of ultrasonic detection. Consequently, the identified region of fatigue cracks intersects with the ultrasonic reflection interface within the C-scan image. Accurately identifying the reflection interfaces B0, T1, and B2 is essential, as fatigue cracks are predominantly situated between B0 and B2. To address this, an algorithm has been designed to pinpoint the location of these reflection interfaces, facilitating the measurement of fatigue crack depths on both sides. While depth detection in the secondary echo area is generally more precise, a safety margin is applied by selecting the greater of the two measured crack depths as the corrected output value for the fatigue crack depth. As shown in Fig. 18.

The crack image within the second echo region corresponds to a depth of 6.0 mm, while the crack image within the third echo region corresponds to a depth of 3.9 mm. Selected the larger depth of 6.0 mm as the identification result for output.

Fig. 19 (b) presents the rectified identification results of 6.0 mm,

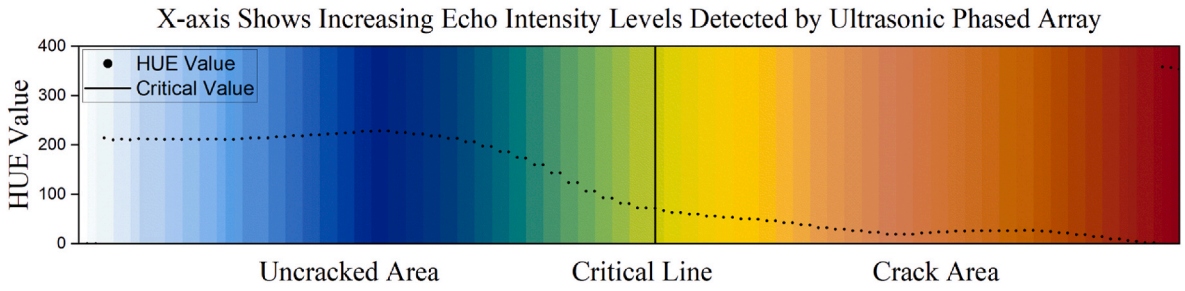


Fig. 14. Echo intensity grade map of ultrasonic phased array identification results.

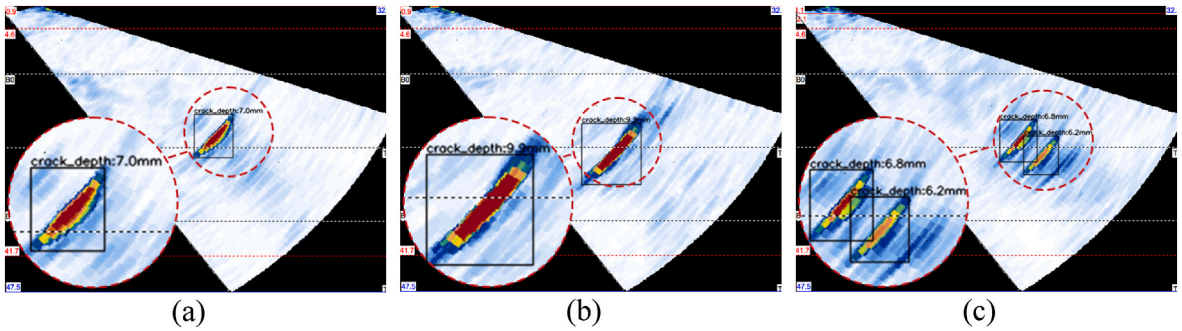


Fig. 16. Fatigue crack depth identification results: (a) identification results 1; (b) identification results 2; and (c) identification results 3.

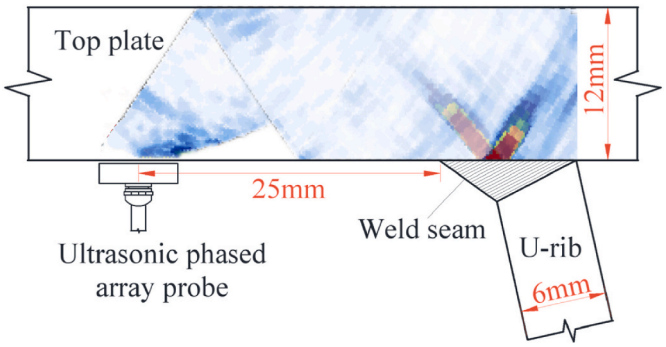


Fig. 17. Schematic diagram of top plate scanning.

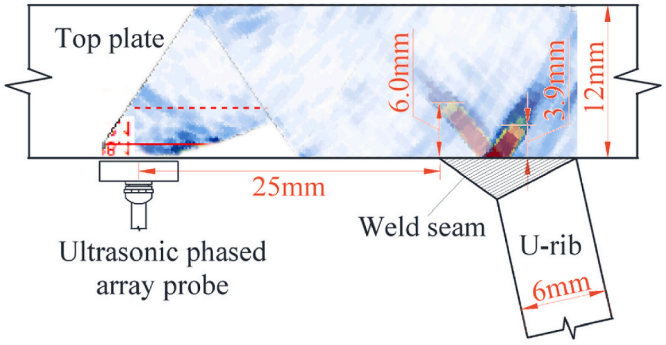


Fig. 18. Corrected top plate scanning.

which closely correspond to the TOFD detection results obtained in the field (i.e., 5.8 mm). The intelligent identification results of the depth of fatigue cracks in Fig. 19 (a) and 19 (c) were corrected from 7.0 mm to 5.5 mm, 6.8 mm–5.5 mm, 6.8 mm and 6.2 mm–4.5 mm and 4.4 mm, respectively. These correction results were within a 5% difference from the field TOFD test results, providing compelling evidence for the accuracy and reliability of the corrected results. Subsequently, the rectified algorithm was employed to intelligently identify the depth of fatigue cracks. This rigorous analysis and correction process significantly improved the accuracy of fatigue crack identification, rendering it more reliable and valuable for practical applications.

To further assess the accuracy of the revised depth detection algorithm, the TOFD scanning results of a specific section at the site were compared with the results of the depth recognition algorithm, as shown in Fig. 20. In the figure, U6 - U15 represent the identification numbers of the U-ribs, and L denotes the length of that section. The results indicate that compared with 80 sets of on-site TOFD scanning results, the maximum error of the revised depth recognition results is 6.4%, and the average error is less than 5%, which can meet the requirements of on-site detection.

The framework integrating PAUT and the YOLOv7-tiny algorithm for crack scanning and identification can acquire the morphology and distribution of OSD cracks in real-time. It can be used to analyze the morphology and propagation trends of the cracks, construct an assessment model for the development of cracks, and predict the future progression of the cracks, thereby providing decision support for the maintenance and reinforcement of bridges. The proposed framework also addresses the issues inherent in traditional OSD fatigue crack detection, such as the difficulty in detecting internal environments, insufficient training samples, low detection efficiency, low level of automation, slow detection speed, and the inability to recover depth information.

5. Conclusions

This paper presents an innovative automated approach for the detection of cracks in orthotropic steel bridge decks. The methodology employs a robotic scanner equipped with an ultrasonic phased array probe to capture images of the cracks. Data augmentation is achieved using a Deep Convolutional Generative Adversarial Network (DCGAN), and crack localization and depth identification are performed utilizing algorithms from the YOLO series. Additionally, attention mechanisms

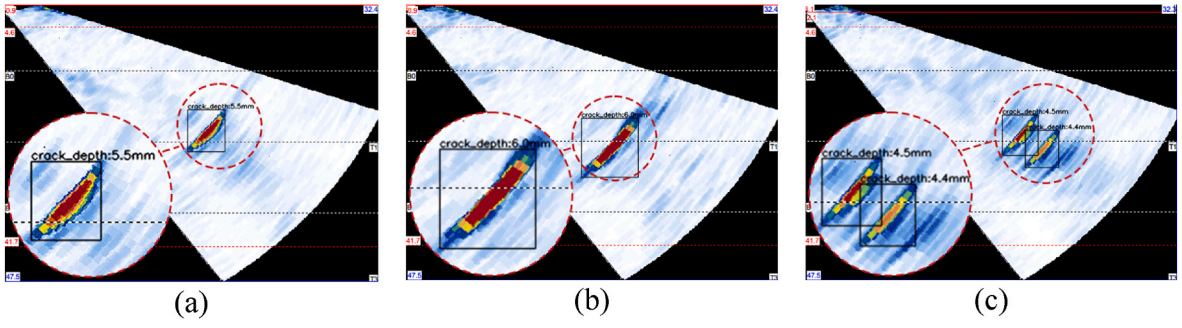


Fig. 19. Corrected fatigue crack depth identification results: (a) identification results 1; (b) identification results 2; and (c) identification results 3.

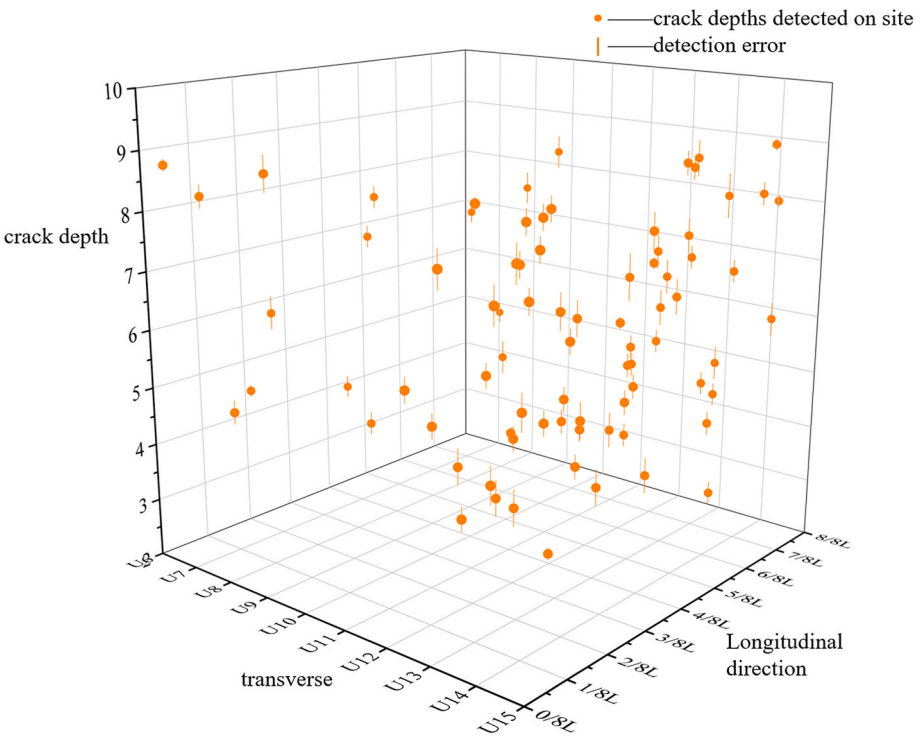


Fig. 20. Error distribution map of crack depth identification compared with TOFD.

are integrated to augment the performance of the YOLO v7-tiny algorithm. The results demonstrate the feasibility and effectiveness of the proposed approach, which holds the potential to enhance the safety and reliability of long-span bridges, reduce the risks of structural failure, and improve transportation efficiency. The main conclusions are drawn.

(1) By employing the algorithm presented in this article, a robotic inspection system equipped with an ultrasonic phased array probe can rapidly scan and identify OSD cracks, thereby minimizing the need for manual operations and achieving high-precision detection in dark and hermetic environments.
(2) The proposed Deep Convolutional Generative Adversarial Network (DCGAN) algorithm achieves a maximum Inception Score (IS) of 41.27 and a minimum Fréchet Inception Distance (FID) of 7.03, indicating that the crack images generated by the DCGAN exhibit diversity and high fidelity.
(3) The YOLOv7-tiny algorithm, as utilized within the article, demonstrates the superior performance among the YOLO algorithms discussed. For crack identification, It achieves a precision rate of 90.64%, a recall rate of 89.36%, an F1 score of 0.90, and an mAP@0.5:0.95 value of 56.57%.
(4) Compared to CA and CBAM, the ODDC attention mechanism yields the greatest performance improvement for the YOLOv7-tiny algorithm. This enhancement results in a crack identification precision of 95.56%, a recall rate of 92.18%, an F1 score of 0.91, and an mAP@0.5:0.95 value of 62.92%, which are better than previous studies
(5) By refining the algorithm to address the intersection of ultrasonic reflection interfaces, the maximum error in the identification of structural crack depth has been reduced to within 5%, indicating that the YOLOv7-tiny algorithm can identify crack depth effectively.

CRediT authorship contribution statement

Fei Hu: Writing – original draft, Software, Data curation. **Hong-ye Gou:** Writing – review & editing, Resources, Project administration, Funding acquisition. **Hao-zhe Yang:** Writing – review & editing, Writing – original draft. **Huan Yan:** Validation, Software, Methodology. **Yi-qing Ni:** Project administration, Investigation, Funding acquisition. **You-wu Wang:** Writing – review & editing, Visualization, Resources.

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.

Acknowledgment

The research was funded by the Chengdu Municipal Bureau of Science and Technology project (grant No. 2023-GH02-00051-HZ), the Sichuan Outstanding Youth Science and Technology Talent Project (grant No. 2022JDJQ0016), the Fund of Science and Technology Project of Transportation in Sichuan Province, China (Grant No. 2022-ZL-02), and the Project of Beijing-Shanghai High Speed Railway Company Limited (Grant No. 2024-11).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.iintel.2024.100113.

References

Azar, L., Wooh, S.C., 1999. Experimental characterization of ultrasonic phased arrays for the nondestructive evaluation of concrete structures. Mater. Eval. 57 (2). https://www.osti.gov/biblio/355660.

Badithela, A., Wongpiromsarn, T., Murray, R.M., 2022. Evaluation metrics for object detection for autonomous systems. arXiv:2210.10298. https://doi.org/10.48550/ARXIV.2210.10298.

Bernasconi, A., Carboni, M., Comolli, L., Galeazzi, R., Gianneo, A., Kharshiduzzaman, M., 2016. Fatigue crack growth monitoring in composite bonded lap joints by a distributed fibre optic sensing system and comparison with ultrasonic testing. J. Adhes. 92 (7–9), 739–757. https://doi.org/10.1080/00218464.2015.1123153.

Bernasconi, A., Martulli, L.M., Carboni, M., 2022. Fatigue crack growth analysis in composite bonded joints by back face distributed strain sensing and comparison with X-ray microtomography. Int. J. Fatig. 154, 106526. https://doi.org/10.1016/j.ijfatigue.2021.106526.

Bianchi, E., Hebdon, M., 2022. Visual structural inspection datasets. Autom. ConStr. 139, 104299. https://doi.org/10.1016/j.autcon.2022.104299.

Bochkovskiy, A., Wang, C.-Y., Liao, H.-Y.M., 2020. YOLOv4: optimal speed and accuracy of object detection. arXiv:2004.10934. https://doi.org/10.48550/ARXIV.2004.10934.

Cao, B., Ding, Y., Song, Y., Zhong, W., 2019. Fatigue life evaluation for deck-rib welding details of orthotropic steel deck integrating mean stress effects. J. Bridge Eng. 24 (2), 04018114. https://doi.org/10.1061/(ASCE)BE.1943-5592.0001344.

Cha, Y., Choi, W., Büyüköztürk, O., 2017. Deep learning-based crack damage detection using convolutional neural networks. Comput. -Aided Civil Infrastruct. Eng. 32 (5), 361–378. https://doi.org/10.1111/mice.12263.

Chen, H., Zhang, Z., Yin, W., Zhou, G., Wang, L., Li, Y., Zhao, C., Wang, C., 2024. Shape characterization and depth recognition of metal cracks based on laser infrared thermography and machine learning. Expert Syst. Appl. 238, 122083. https://doi.org/10.1016/j.eswa.2023.122083.

Cheng, B., Cao, X., Ye, X., Cao, Y., 2017. Fatigue tests of welded connections between longitudinal stringer and deck plate in railway bridge orthotropic steel decks. Eng. Struct. 153, 32–42. https://doi.org/10.1016/j.engstruct.2017.10.016.

Chenglong, S., Fangfang, S., Bixing, Z.H.A.N.G., 2016. Analysis on defect classification by deep neural networks and wavelet packet transform. Acta Acust. 41 (4), 499–506. https://doi.org/10.15949/j.cnki.0371-0025.2016.04.007.

Cruz, F.C., Simas Filho, E.F., Albuquerque, M.C.S., Silva, I.C., Farias, C.T.T., Gouvêa, L.L., 2017. Efficient feature selection for neural network based detection of flaws in steel welded joints using ultrasound testing. Ultrasonics 73, 1–8. https://doi.org/10.1016/j.ultras.2016.08.017.

Cui, C., Bu, Y., Bao, Y., Zhang, Q., Ye, Z., 2018. Strain energy-based fatigue life evaluation of deck-to-rib welded joints in OSD considering combined effects of stochastic traffic load and welded residual stress. J. Bridge Eng. 23 (2), 04017127. https://doi.org/10.1061/(ASCE)BE.1943-5592.0001181.

Elfwing, S., Uchibe, E., Doya, K., 2018. Sigmoid-weighted linear units for neural network function approximation in reinforcement learning. Neural Netw 107, 3–11. https://doi.org/10.1016/j.neunet.2017.12.012.

Fawcett, T., 2006. An introduction to ROC analysis. Pattern Recogn. Lett. 27 (8), 861–874. https://doi.org/10.1016/j.patrec.2005.10.010.

Fisher, J.W., Barsom, J.M., 2016. Evaluation of cracking in the rib-to-deck welds of the bronx–whitestone bridge. J. Bridge Eng. 21 (3), 04015065. https://doi.org/10.1061/(ASCE)BE.1943-5592.0000823.

Gallo, I., Rehman, A.U., Dehkordi, R.H., Landro, N., La Grassa, R., Boschetti, M., 2023. Deep object detection of crop weeds: performance of YOLOv7 on a real case dataset from UAV images. Remote Sens 15 (2), 539. https://doi.org/10.3390/rs15020539.

Guo, P., Meng, X., Meng, W., Bao, Y., 2022. Monitoring and automatic characterization of cracks in strain-hardening cementitious composite (SHCC) through intelligent interpretation of photos. Compos. B Eng. 242, 110096. https://doi.org/10.1016/j.compositesb.2022.110096.

Havlice, J.F., Taenzer, J.C., 1979. Medical ultrasonic imaging: an overview of principles and instrumentation. Proceedings of the IEEE 67 (4), 620–641. In: https://doi.org/10.1109/PROC.1979.11287.

Hou, Q., Zhou, D., Feng, J.. Coordinate attention for efficient mobile network design. https://doi.org/10.1109/CVPR46437.2021.01350.

Isola, P., Zhu, J.Y., Zhou, T., Efros, A.A.. Image-to-image translation with conditional adversarial networks. https://doi.org/10.1109/WIECON-ECE60392.2023.10456447.

Jiang, H., Yang, D., Zhi, Z., Jing, Q., Gao, J., Tao, C., Cheng, Z., 2024. A normal weld recognition method for time-of-flight diffraction detection based on generative adversarial network. J. Intell. Manuf. 35, 217–233. https://doi.org/10.1007/s10845-022-02041-9.

Jiao, L., Zhang, F., Liu, F., Yang, S., Li, L., Feng, Z., Qu, R., 2019. A survey of deep learning-based object detection. IEEE Access 7, 128837–128868. https://doi.org/10.1109/ACCESS.2019.2939201.

Kainuma, S., Yang, M., Jeong, Y.-S., Inokuchi, S., Kawabata, A., Uchida, D., 2017. Experimental investigation for structural parameter effects on fatigue behavior of rib-to-deck welded joints in orthotropic steel decks. Eng. Fail. Anal. 79, 520–537. https://doi.org/10.1016/j.engfailanal.2017.04.028.

Karimi, S., Mirza, O., 2023. Damage identification in bridge structures: review of available methods and case studies. Aust. J. Struct. Eng. 24 (2), 89–119. https://doi.org/10.1080/13287982.2022.2120239.

Kingma, D.P., Ba, J., 2014. Adam: a method for stochastic optimization. arXiv: 1412.6980. https://doi.org/10.48550/ARXIV.1412.6980.

Laxman, K.C., Tabassum, N., Ai, L., Cole, C., Ziehl, P., 2023. Automated crack detection and crack depth prediction for reinforced concrete structures using deep learning. Construct. Build. Mater. 370, 130709. https://doi.org/10.1016/j.conbuildmat.2023.130709.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521 (7553), 436–444. https://doi.org/10.1038/nature14539.

Li, Y., Xu, Y., Shao, M., Gao, H., Tang, C., Lu, Q., 2020. Fast detection using PID pre-interpolation algorithm for magnetic particle testing. IEEE Trans. Magn. 56 (12), 1–7. https://doi.org/10.1109/TMAG.2020.3025050.

Li, C., Zhou, A., Yao, A., 2022. Omni-dimensional dynamic convolution. arXiv: 2209.07947. https://doi.org/10.48550/ARXIV.2209.07947.

Liashchynskyi, P., Liashchynskyi, P., 2019. Grid search, random search, genetic algorithm: a big comparison for nas. arXiv:1912.06059. https://doi.org/10.48550/ARXIV.1912.06059.

Liu, S.W., Huang, J.H., Sung, J.C., Lee, C.C., 2002. Detection of cracks using neural networks and computational mechanics. Comput. Methods Appl. Mech. Eng. 191 (25–26), 2831–2845. https://doi.org/10.1016/S0045-7825(02)00221-9.

Liu, G., Reda, F.A., Shih, K.J., Wang, T.C., Tao, A., Catanzaro, B.. Image inpainting for irregular holes using partial convolutions. https://doi.org/10.1007/978-3-030-01252-6_6.

Liu, Y., Zhang, X., Liu, R., Chen, G., 2021. Design and mechanical properties of steel-UHPC lightweight composite decks. Adv. Civ. Eng. 2021, 1–13. https://doi.org/10.1155/2021/8508795.

Manikandan, K.R., Sivagurunathan, P.A., Ananthan, S.S., Moshi, A.A.M., Bharathi, S.S., 2021. Study on the influence of temperature and vibration on indications of liquid penetrant testing of A516 low carbon steel. Mater. Today: Proc. 39 (4), 1559–1564. https://doi.org/10.1016/j.matpr.2020.05.572.

McNab, A., Campbell, M.J., 1987. Ultrasonic phased arrays for nondestructive testing. NDT. E. Int. 20 (6), 333–337. https://doi.org/10.1016/0308-9126(87)90290-2.

Medak, D., Posilovic, L., Subasic, M., Budimir, M., Loncaric, S., 2021. Automated defect detection from ultrasonic images using deep learning. IEEE Trans. Sonics Ultrason. Eng. 68 (10), 3126–3134. https://doi.org/10.1109/TUFFC.2021.3081750.

Meng, M., Chua, Y.J., Wouterson, E., Ong, C.P.K., 2017. Ultrasonic signal classification and imaging system for composite materials via deep convolutional neural networks. Neurocomputing 257, 128–135. https://doi.org/10.1016/j.neucom.2016.11.066.

Meng, T., Tao, Y., Chen, Z., Avila, J.R.S., Ran, Q., Shao, Y., Huang, R., Xie, Y., Zhao, Q., Zhang, Z., Yin, H., Peyton, A.J., Yin, W., 2021. Depth evaluation for metal surface defects by eddy current testing using deep residual convolutional neural networks. IEEE Trans. Instrum. Meas. 70, 1–13. https://doi.org/10.1109/TIM.2021.3117367.

Molinier, N., Painchaud-April, G., Le Duff, A., Toews, M., Bélanger, P., 2023. Ultrasonic imaging using conditional generative adversarial networks. Ultrasonics 133, 107015. https://doi.org/10.1016/j.ultras.2023.107015.

Munir, N., Kim, H.-J., Park, J., Song, S.-J., Kang, S.-S., 2019. Convolutional neural network for ultrasonic weldment flaw classification in noisy conditions. Ultrasonics 94, 74–81. https://doi.org/10.1016/j.ultras.2018.12.001.

Padhi, B., Reza, M., Shams, M.S., Sai, A.N., 2022. Concrete crack detection using deep convolutional generative adversarial network. In: International Advanced Computing Conference. Springer Nature, Switzerland, Cham, pp. 147–161. https://doi.org/10.1007/978-3-031-35641-4_11.

Pak, M., Kim, S., 2017. A review of deep learning in image recognition. In: 2017 4th International Conference on Computer Applications and Information Processing Technology (CAIPT), pp. 1–3. https://doi.org/10.1109/CAIPT.2017.8320684.

Pei, L., Sun, Z., Xiao, L., Sun, J., Zhang, H., 2021. Virtual generation of pavement crack images based on improved deep convolutional generative adversarial network. Eng. Appl. Artif. Intell. 104, 104376. https://doi.org/10.1016/j.engappai.2021.104376.

Posilović, L., Medak, D., Subašić, M., Budimir, M., Lončarić, S., 2022. Generating ultrasonic images indistinguishable from real images using Generative Adversarial Networks. Ultrasonics 119, 106610. https://doi.org/10.1016/j.ultras.2021.106610.

Rácz, A., Bajusz, D., Héberger, K., 2021. Effect of dataset size and train/test split ratios in QSAR/QSPR multiclass classification. Molecules 26 (4), 1111. https://doi.org/10.3390/molecules26041111.

Regazzi, D., Beretta, S., Carboni, M., 2014. An investigation about the influence of deep rolling on fatigue crack growth in railway axles made of a medium strength steel. Eng. Fract. Mech. 131, 587–601. https://doi.org/10.1016/j.engfracmech.2014.09.016.

Ren, S., He, K., Girshick, R., Sun, J., 2016. Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Trans. Pattern Anal. Mach. Intell. 39 (6), 1137–1149. https://doi.org/10.1109/TPAMI.2016.2577031.

Robbins, H., Monro, S., 1951. A stochastic approximation method. Ann. Math. Stat. 22 (3), 400–407. http://www.jstor.org/stable/2236626.

Sekar, A., Perumal, V., 2022. Cfc-gan: forecasting road surface crack using forecasted crack generative adversarial network. IEEE Trans. Intell. Transport. Syst. 23 (11), 21378–21391. https://doi.org/10.1109/TITS.2022.3171433.

Shao, X., Yi, D., Huang, Z., Zhao, H., Chen, B., Liu, M., 2013. Basic performance of the composite deck system composed of orthotropic steel deck and ultrathin RPC layer. J. Bridge Eng. 18 (5), 417–428. https://doi.org/10.1061/(ASCE)BE.1943-5592.0000348.

Shorten, C., Khoshgoftaar, T.M., 2019. A survey on image data augmentation for deep learning. J. Big Data 6 (1), 60. https://doi.org/10.1186/s40537-019-0197-0.

Song, Q., Liu, L., Lu, N., Zhang, Y., Muniyandi, R.C., An, Y., 2024. A three-stage pavement image crack detection framework with positive sample augmentation. Eng. Appl. Artif. Intell. 129, 107624. https://doi.org/10.1016/j.engappai.2023.107624.

Soviany, P., Ionescu, R.T., 2018. Optimizing the trade-off between single-stage and two-stage deep object detectors using image difficulty prediction. *arXiv:1803.08707*. https://doi.org/10.48550/arXiv.1803.08707.

Taheri, H., Hassen, A.A., 2019. Nondestructive ultrasonic inspection of composite materials: a comparative advantage of phased array ultrasonic. Appl. Sci. 9 (8), 1628. https://doi.org/10.3390/app9081628.

Thomas, G.H., 1995. In: Mignogna, R.B. (Ed.), Overview of Nondestructive Evaluation Technologies, 9th ed., pp. 5–9 https://doi.org/10.1117/12.212554.

Tian, L., Wang, Z., Liu, W., Alsaadi, F.E., Liu, X., 2021. A new GAN-based approach to data augmentation and image segmentation for crack detection in thermal imaging tests. Cogn. Comput. 13 (5), 1263–1273. https://doi.org/10.1007/s12559-021-09922-w.

Tunukovic, V., Mcknight, S., Mohseni, E., Pierce, S.G., Pyle, R., Duernberger, E., Loukas, C., Vithanage, R.K.W., Lines, D., Dobie, G., MacLeod, C.N., Cochran, S., O'Hare, T., 2024. A study of machine learning object detection performance for phased array ultrasonic testing of carbon fibre reinforced plastics. NDT E. Int. 144, 103094. https://doi.org/10.1016/j.ndteint.2024.103094.

Uhlig, S., Alkhasli, I., Schubert, F., Tschöpe, C., Wolff, M., 2023. A review of synthetic and augmented training data for machine learning in ultrasonic non-destructive evaluation. Ultrasonics 134, 107041. https://doi.org/10.1016/j.ultras.2023.107041.

Nair, A.A., Tran, T.D., Reiter, A., Bell, M.A.L., 2019. A generative adversarial neural network for beamforming ultrasound images: invited presentation. The 53rd Annual Conference on Information Sciences and Systems (CISS), Baltimore, USA, pp. 1–6, https://doi.org/10.1109/CISS.2019.8692835.

Wang, Y., Wang, Z., Zheng, Y., 2019. Analysis of fatigue crack propagation of an orthotropic bridge deck based on the extended finite element method. Adv. Civ. Eng. 2019, 1–14. https://doi.org/10.1155/2019/6319821.

Wang, C.-Y., Bochkovskiy, A., Liao, H.-Y.M., 2023. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 7464–7475. Vancouver, Canada. https://doi.org/10.1109/CVPR52729.2023.00721.

Wang, C., Zhu, T., Yang, B., Xiao, S., Yang, G., 2024. A study of fatigue surface crack propagation paths of aluminum alloy butt welds using a Phased-Array Total-Focus imaging technique. Theor. Appl. Fract. Mech. 133, 104572. https://doi.org/10.1016/j.tafmec.2024.104572.

Woo, S., Park, J., Lee, J.-Y., Kweon, I.S., 2018. CBAM: convolutional block attention Module. arXiv:1807.06521. https://doi.org/10.48550/arXiv.1807.06521.

Yang, X., Chen, S., Jin, S., Chang, W., 2013. Crack orientation and depth estimation in a low-pressure turbine disc using a phased array ultrasonic transducer and an artificial neural network. Sensors 13 (9), 12375–12391. https://doi.org/10.3390/s130912375.

Yuhas, M., Easwaran, A., 2022. Demo abstract: real-time out-of-distribution detection on a mobile robot. ArXiv, abs/2211.11520. https://doi.org/10.48550/ARXIV.2211.11520.

Zaidi, S.S.A., Ansari, M.S., Aslam, A., Kanwal, N., Asghar, M., Lee, B., 2022. A survey of modern deep learning based object detection models. Digit. Signal Process. 126, 103514. https://doi.org/10.1016/j.dsp.2022.103514.

Zhai, G., Narazaki, Y., Wang, S., Shajihan, S.A.V., Spencer Jr., B.F., 2022. Synthetic data augmentation for pixel-wise steel fatigue crack identification using fully convolutional networks. Smart Struct. Syst. 29 (1), 237–250. https://doi.org/10.12989/sss.2022.29.1.237.

Zhang, Q.-H., Cui, C., Bu, Y.-Z., Liu, Y.-M., Ye, H.-W., 2015. Fatigue tests and fatigue assessment approaches for rib-to-diaphragm in steel orthotropic decks. J. Constr. Steel Res. 114, 110–118. https://doi.org/10.1016/j.jcsr.2015.07.014.

Zhang, X., Li, J., He, Q., Zhang, H., Luo, J., 2018. In: High-quality reconstruction of plane-wave imaging using generative adversarial network. IEEE, pp. 1–4. https://doi.org/10.1109/ULTSYM.2018.8579877.

Zhang, J., Karimireddy, S.P., Veit, A., Kim, S., Reddi, S.J., Kumar, S., Sra, S., 2019. Why adam beats sgd for attention models. arXiv:1912.03194v1. https://doi.org/10.48550/arXiv.1912.03194.

Zhang, L., Wang, Z., Wang, L., Zhang, Z., Chen, X., Meng, L., 2021. Machine learning-based real-time visible fatigue crack growth detection. Digit. Commun. Netw. 7 (4), 551–558. https://doi.org/10.1016/j.dcan.2021.03.003.

Zheng, J., Wu, H., Zhang, H., Wang, Z., Xu, W., 2022. Insulator-Defect detection algorithm based on improved YOLOv7. Sensors 22 (22), 8801. https://doi.org/10.3390/s22228801.

Zhu, J.Y., Park, T., Isola, P., Efros, A.A.. Unpaired image-to-image translation using cycle-consistent adversarial networks. https://doi.org/10.1109/ICCV.2017.244.

Zhu, Z., Xiang, Z., Li, J., Carpinteri, A., 2020. Fatigue damage investigation on diaphragm cutout detail on orthotropic bridge deck based on field measurement and FEM. Thin Wall. Struct. 157, 107106. https://doi.org/10.1016/j.tws.2020.107106.

Zhu, Y., Sekiya, H., Okatani, T., Tai, M., Morichika, S., 2023. B-CNN: a deep learning method for accelerometer-based fatigue cracks monitoring system. J. Civil Struct. Health Monit. 13, 947–959. https://doi.org/10.1007/s13349-023-00690-9.