scientific reports

OPEN

A novel edge crop method and enhanced YOLOv5 for efficient wind turbine blade damage detection

Boyu Feng[1], Bo Liu[2], Li Song[1,3]✉, Yongyan Chen[1,3], Xiaofeng Jiao[4] & Baiqiang Wang[1]

Accurately and rapidly detecting damage to wind turbine blades is critical for ensuring the safe operation of wind turbines. Current deep learning-based detection methods predominantly employ the gathered blade images directly for damage detection. However, due to the slender geometry of wind turbine blades, non-blade background information accounts for a considerable proportion of the captured images with complex background features, affecting the detection of blade damage. To address this challenge, we propose a novel edge cropping method combined with an enhanced YOLOv5s network for detecting damage in wind turbine blades, termed Edge Crop and Enhanced YOLOv5 (EC–EY). The edge cropping method adaptively modifies the cropping stride by the edge features of both sides of the blade, thereby procuring image content that predominantly encompasses the blade region. This procedure effectively mitigates the interference from complex background features and augments the utilization of image pixels. Furthermore, the enhanced YOLOv5 network incorporates the global attention mechanism into the head section of the network and substitutes the original SPPF module with an attention-based intra-scale feature interaction module. The EC–EY aims to improve the detection accuracy for small and variable-shape damages in wind turbine blades. EC–EY achieved excellent performance on a dataset of wind turbine blade damage collected in western Inner Mongolia. Notably, the edge cropping method significantly improves the accuracy of wind turbine blade damage detection.

Nowadays, the development of wind energy utilization has propelled the extensive installation of wind turbines[1]. As the key to capturing wind energy, the wind turbine blade dictates whether the turbine can function effectively and safely[2]. Therefore, the real-time and rapid monitoring of blade damage is critical. Otherwise, damage will reduce the efficiency of wind power generation and negatively impact economic performance.

Methods for detecting wind turbine blade damage can be categorized into contact and non-contact-based[3]. Contact-based monitoring, such as strain and vibration monitoring, requires sensors to be physically attached to the blades for detection. Manual installation using suspended baskets is necessary for turbines not yet equipped with such systems[4–6]. Non-contact-based monitoring methods, such as visual inspection and noise measurement, allow damage detection without physical contact with blades, thus reducing economic costs and improving practicality[7–9]. This study focuses on a visual monitoring method using non-contact-based techniques to provide real-time, rapid, and intuitive detection of wind turbine blade damage.

Many researchers have contributed significantly to visually inspecting wind turbine blades[10–17]. These researchers have proposed image processing or deep learning damage detection methods by considering the characteristics of wind turbine blade damage, such as the small proportion of damage within the blade image. However, because of the slender geometry of wind turbine blades, background information inevitably occupies a portion of the computational resources, which reduces the efficiency of image processing and pixel utilization and makes damage detection more challenging. To address this issue, scholars have attempted to solve specific problems under certain scenarios[12,13]. However, these methods lack generalizability and cannot handle wind turbine blade images with varying pixel resolutions and morphological variations.

[1]College of Energy and Power Engineering, Inner Mongolia University of Technology, Hohhot 010051, China. [2]College of Data Science and Application, Inner Mongolia University of Technology, Hohhot 010051, China. [3]Key Laboratory of Renewable Energy in Inner Mongolia, Hohhot 010051, China. [4]Inner Mongolia Power Science Research Institute, Hohhot 010051, China. ✉email: songli@imut.edu.cn

In order to eliminate background information from images and improve the utilization of image pixels during deep learning detection, this study proposes an image-cropping method based on edge detection, called edge crop. Traditional edge detection algorithms, such as the first-order Sobel and Robert operators, and second-order edge detection methods, such as the Canny and Laplacian operators, have been widely used in image processing[18]. Inspired by the Arithmetic Optimization Algorithm[19,20], the edge crop method employs the Canny operator from traditional edge detection techniques to detect the edges of wind turbine blades. It selects two edge lines from multiple detected boundaries to crop the image. This simple data processing eliminates background information and enhances the utilization of image pixels. The experimental results demonstrate that the edge crop method significantly improves the detection accuracy of wind turbine blade damage.

Because wind turbine blade damage can exhibit variable shapes even within the same damage category, this study enhanced the YOLOv5 network algorithm based on variable-shape damage to better adapt it to blade damage detection tasks. YOLOv5, introduced by the Ultralytics team in 2020[21], has gained widespread recognition for its robust mechanism and modification adaptability. YOLOv5s, as a lightweight model, maintains high accuracy while controlling the model parameters, making it particularly suitable for drone deployment. In response to the diverse shapes of wind turbine blade damage, this study incorporates attention-based intra-scale feature interaction (AIFI) and the global attention mechanism (GAM). The experimental results demonstrate that the enhanced YOLOv5 outperforms the original version in terms of damage-detection accuracy.

In summary, this study presents a novel edge crop method for image processing and an enhanced YOLOv5s model for detecting blade damage named EC–EY. Experiments demonstrate that the EC–EY eliminates the impact of background interference in blade damage detection, enhancing detection accuracy and pixel utilization. The main contributions of this study are as follows:

1. A novel image-processing method, edge crop, is proposed. Edge crops effectively eliminate most of the background interference, allowing for the extraction of wind turbine blade regions from complex environments and significantly improving damage detection accuracy. Compared to the existing global crop method, edge crop demonstrates a cutting efficiency twice that of the global crop, making it more suitable for processing wind turbine blade images.
2. An enhanced YOLOv5 algorithm is proposed for detecting variable-shaped damage to wind turbine blades. In response to the variable shapes of wind turbine blade damage, the AIFI is integrated into the bottom of the backbone, and the GAM is inserted into the head of YOLOv5 to enhance the model. These improvements have increased the accuracy of blade damage detection.
3. A dataset collected from the western region of Inner Mongolia is used to validate the EC-EY method. The experimental results show that EC-EY achieves state-of-the-art performance on this dataset.

Related work

To process wind turbine blade images, Tang et al.[22] utilized the edge information of wind turbine blades to locate their straight edges using Hough line detection. They implemented adaptive segmentation of the blades based on the Otus thresholding method and morphological operations in the grab-cut algorithm, achieving an accurate segmentation of the blade image from the background. Ali et al.[23] employed the Canny operator for edge detection in wind turbine blade images and explored the feasibility of online detection without turbine shutdown. Zhang et al.[12] used the Grab-cut algorithm to eliminate the background of wind turbine blade images and employed the Canny operator and Hough Linear Inspection to filter two edge lines and align the blade image vertically by rotating it, facilitating data processing. Zhou et al.[13] proposed the regression crop method, which utilizes the LineNet method for auto-selecting blade areas, predicts the regression line, and crops the blade image along that line to a fixed size, enhancing detection accuracy. However, the aforementioned image processing methods have not simultaneously addressed the issues of background and pixels or employed neural networks leading to intricate processing, making it challenging to meet the application requirements for wind turbine blade damage detection tasks in deep learning.

Deep learning-based object detection methods can be classified into two categories[24]. The first includes CNN-based models, such as YOLO[25], R-CNN[26], Faster R-CNN[27], and Mask R-CNN[28]. Several researchers have adapted these networks to detect wind turbine blade damage. Zhu et al.[14] proposed MI-YOLO based on YOLOv5 by modifying the backbone and introducing a C3TR module to improve the detection accuracy of surface cracks. Dimitri et al.[15] compared the performance of different CNN architectures with varying training data sizes and concluded that ResNet50 yielded the best results. Zhang et al.[16] compared YOLOv3, YOLOv4, and Mask R-CNN and concluded that Mask R-CNN was the optimal wind turbine blade defect detection model. Chao et al.[17] proposed a semi-supervised object detection method based on YOLOv4 and GANs, improving the loss function and introducing a scSE attention mechanism to enhance the network's feature extraction capabilities. The second category consists of transformer-based models such as DETR[29], Deformable-DETR[30], and RT-DETR[31]. Although these models were introduced in 2023 and have not yet been widely applied in wind turbine blade damage detection, significant progress has been made in other domains. Ge et al.[32] designed a position-encoding network with multi-scale feature maps and redefined a loss function to detect defects in wood veneers. Cheng et al.[33] developed PCB-DETR, a model for PCB defect detection based on Deformable-DETR, which outperformed Faster R-CNN, YOLOv3, and Deformable-DETR in PCB defect detection. Liu et al.[34] proposed a model for detecting bearing defects, Bearing-DETR, based on RT-DETR, incorporating methods such as Deformable Large Kernel Attention to improve the model's performance in detecting bearing defects. CNN-based models and transformer-based models are used to capture the local and global features, respectively. The integration of these strengths has emerged as a research topic in wind turbine blade damage detection.

Method

Damage to wind turbine blades is small and variable in shape, which poses a significant challenge to its detection. To address this, we propose an edge crop data-processing method and an enhanced YOLOv5 network called EC-EY.

Edge crop

Edge cropping is a method for cropping wind turbine blade images based on edge detection, as shown in Fig. 1. This method can extract and filter straight edges on both sides of a wind turbine blade and crop the image based on the relationship between the two lines, effectively addressing the issues of background interference and low pixel utilization. This method consists of three steps: horizontally aligning the image, filtering out straight edges on both sides of the blade, and cropping the wind turbine blade image. In the first step, the Canny operator is applied to the original image for edge detection, followed by the Hough transform to detect straight lines in the image. By limiting the difference in slope between the two lines, denoted as the first value $\triangle k$, the image is horizontally aligned. In the second step, after horizontally aligning the image, the Canny operator and Hough line transform are applied again to process the image. The coordinates of the two endpoints of all of the detected lines are recorded in an array $p = \{(x_i, y_i)|i = N\}$. Utilizing mathematical relationships, specifically the five values listed below, the straight lines corresponding to the two edges of the blade are filtered out. In the third step, the wind turbine blade image is cropped based on the vertical coordinates of the filtered lines corresponding to the upper edge of the blade.

This method performs relational filtering of edge lines to extract suitable edge lines, which are divided into five values:

a. Parallel relationship between the two lines. According to the analysis of the wind turbine blade image's edge features, most of the blade image's edge lines are approximately parallel. The two straight lines selected using this method should exhibit slope differences within a specific range. This value is used to rotate the image and position the wind turbine blade horizontally in the next processing step.

$$k = \frac{(y_2 - y_1)}{(x_2 - x_1)} \tag{1}$$

$$\Delta k = k_{i+1} - k_i \tag{2}$$

b. Two wire lengths. The selected line in the wind turbine blade image should be of a certain length to ensure that the calculation can be performed. Additionally, lines that might potentially affect the results can be filtered to obtain appropriate lines on both sides.
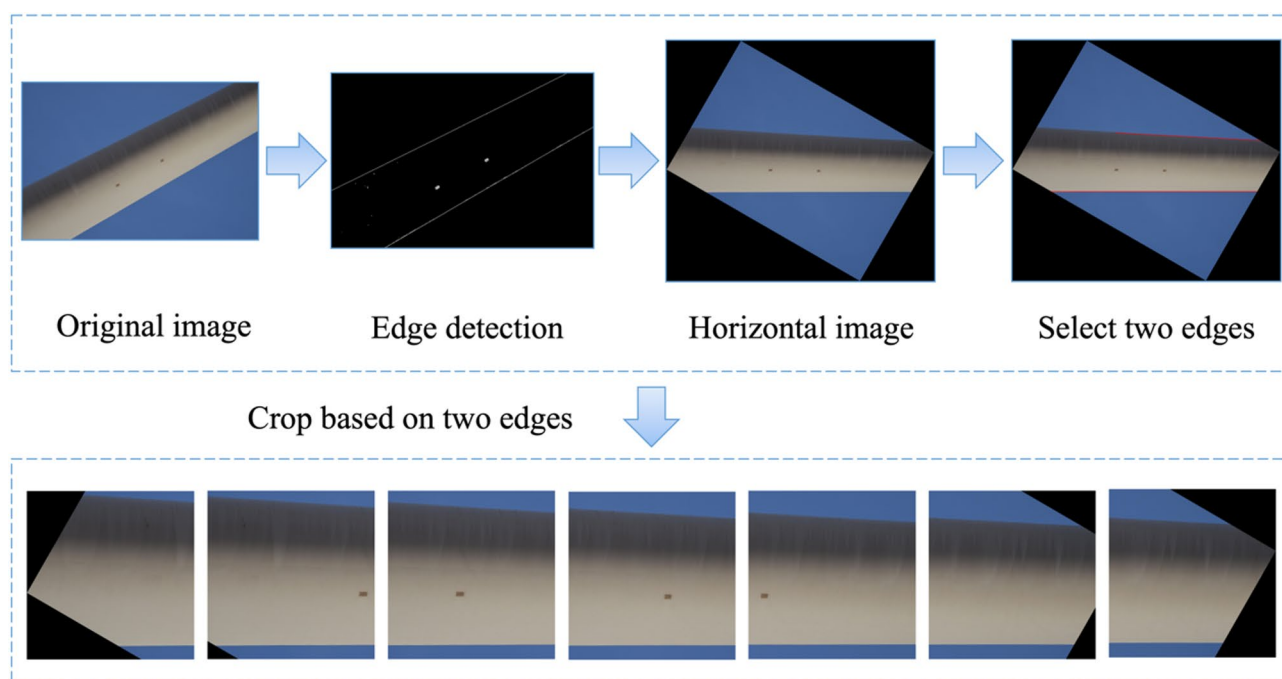

Fig. 1. Edge crop data-processing method.

$$l = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \tag{3}$$

c. Distance between two lines. The two lines at the edges of the wind turbine blade image should be maintained at a certain distance to prevent them from appearing on only one side. The distance between the two lines should be adjusted based on the blade's aspect ratio and proportions.

$$b_i = y_i - k_i x_i \tag{4}$$

$$d = \frac{|b_2 - b_1|}{\sqrt{k^2 + 1}} \tag{5}$$

d. Calculate $\overline{y}$ of all lines. After the image is flipped, the required line segment can be approximated as a horizontal line, implying that its position can be determined by its y-value. The average value of y for all line segments is calculated. One line on the left side of the wind turbine blade had a y-coordinate greater than this average, whereas the line on the right side had a y-coordinate smaller than the average. This method effectively excludes certain environmental interferences.

$$\overline{y} = \frac{\sum_1^n y_i}{n} \tag{6}$$

e. The slope of this line is approximately zero. For an inverted image, the slope of one of the lines is bound to be close to zero, and this line should be selected.

$$k_i \approx 0 \tag{7}$$

After the first step of the horizontal alignment of the image, the Canny operator and Hough line transform are applied. Based on the five aforementioned values, two suitable lines representing the edges of the wind turbine blades are filtered. Because the blade is horizontally aligned, the slope of one edge line is approximately zero, whereas the other edge line can be determined using the slope difference, the distance between the two lines, the length of the lines, and by calculating the average vertical coordinates of all of the lines and ensuring that the vertical coordinates of the two lines are above and below this average, respectively. After selecting the two lines, the edge cropping can be adjusted according to the different requirements of the image cropping size. It can be cropped to a fixed size, such as 640×640 or 1280×1280, with the step size set to 640 or 1280. Alternatively, cropping can be performed based on the distance between the two lines, which corresponds to the height of the wind turbine blade in the pixels. Half of this value could be used as the step size, making it more adaptable to datasets with varying pixel sizes.

Compared to the global sliding window crop (global crop), the edge crop is more suitable for cropping wind turbine blade images, as shown in Fig. 2. The global crop is a technique that crops the entire image along a fixed path. However, it may crop unnecessary background areas when applied to wind turbine blade images. It can only perform cropping with a fixed size, making it unsuitable for wind turbine blade image tasks. By contrast, edge crop is more appropriate for wind turbine blade images. Instead of starting from the image's top-left corner, as in the global crop method, the edge crop begins at the blade's edges. Moreover, edge crop adapts the cropping stride and image size based on the blade edge features, resulting in more suitable datasets for detection models.

Enhanced YOLOv5

Convolution operations are adept at extracting local information from images, which makes the convolution-based YOLOv5 network effective at identifying local damage features for wind turbine blade damage detection tasks. However, because of the small and variable-shaped characteristics of wind turbine blade damage, relying solely on learning local features is insufficient to accomplish this task effectively. To address this issue, we propose two improvement measures. First, AIFI[31] is replaced by SPPF on the bottom of the backbone to improve the detection performance of multi-scale damage. Second, GAM[35] is added to the head part, which interacts with the spatial and channel parts, reducing information loss and improving the final accuracy. The improved network structure is shown in Fig. 3, where the red box highlights the network-modified parts.

AIFI

AIFI is a self-attention-based intra-scale interaction module that utilizes a multi-head self-attention mechanism to capture damaged features in parallel. It can capture more comprehensive features and enhance the ability to extract abstract features through a feedforward network (FFN), thus improving the detection performance for variable-shape damage.

The AIFI structure is shown in Fig. 4. The feature maps generated after convolution are input into the module, which applies the multi-head attention mechanism to perform linear transformations via matrix operations, resulting in feature representations. The Add and Norm module prevented degradation, improved training stability, and accelerated convergence. Next, the FFN module, which relies on the GELU activation function, performs nonlinear transformations to enhance learning capability and obtain better feature representations. Finally, the Add and Norm module is applied again to ensure the detection accuracy did not decrease. This module directly captures features on high-level feature maps, outperforming low-level feature without sufficient
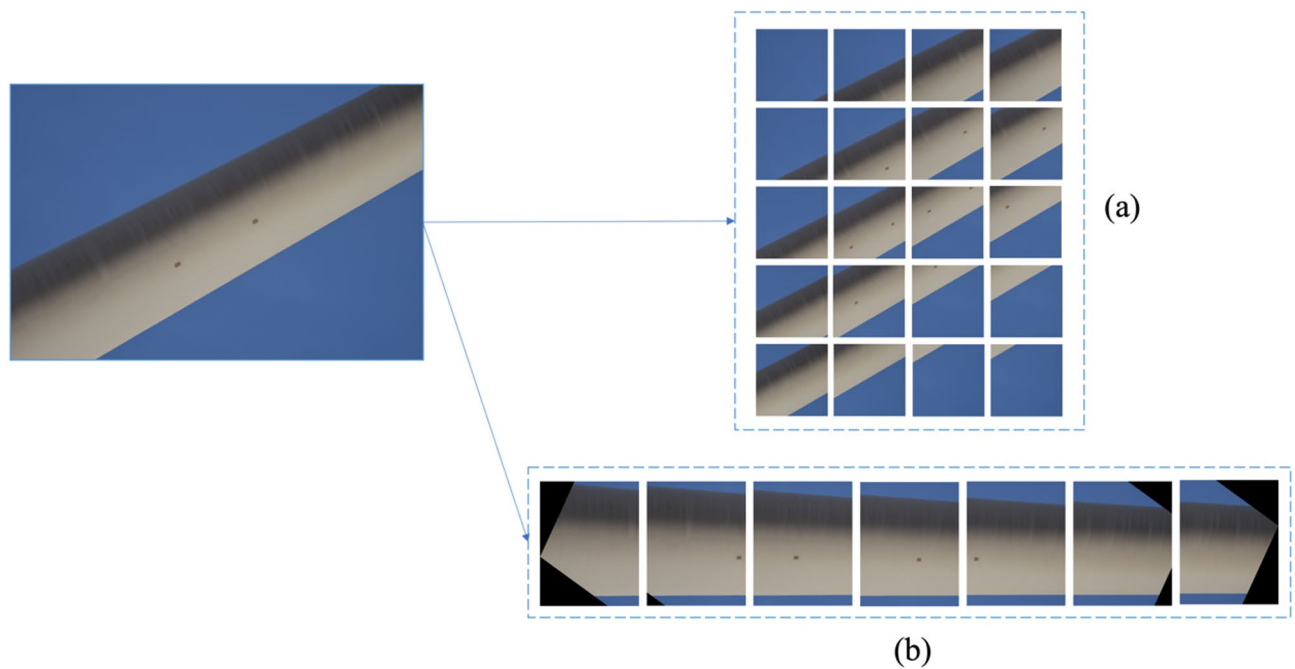
Fig. 2. Comparison of edge crop and global crop: (**a**) global crop; (**b**) edge crop.
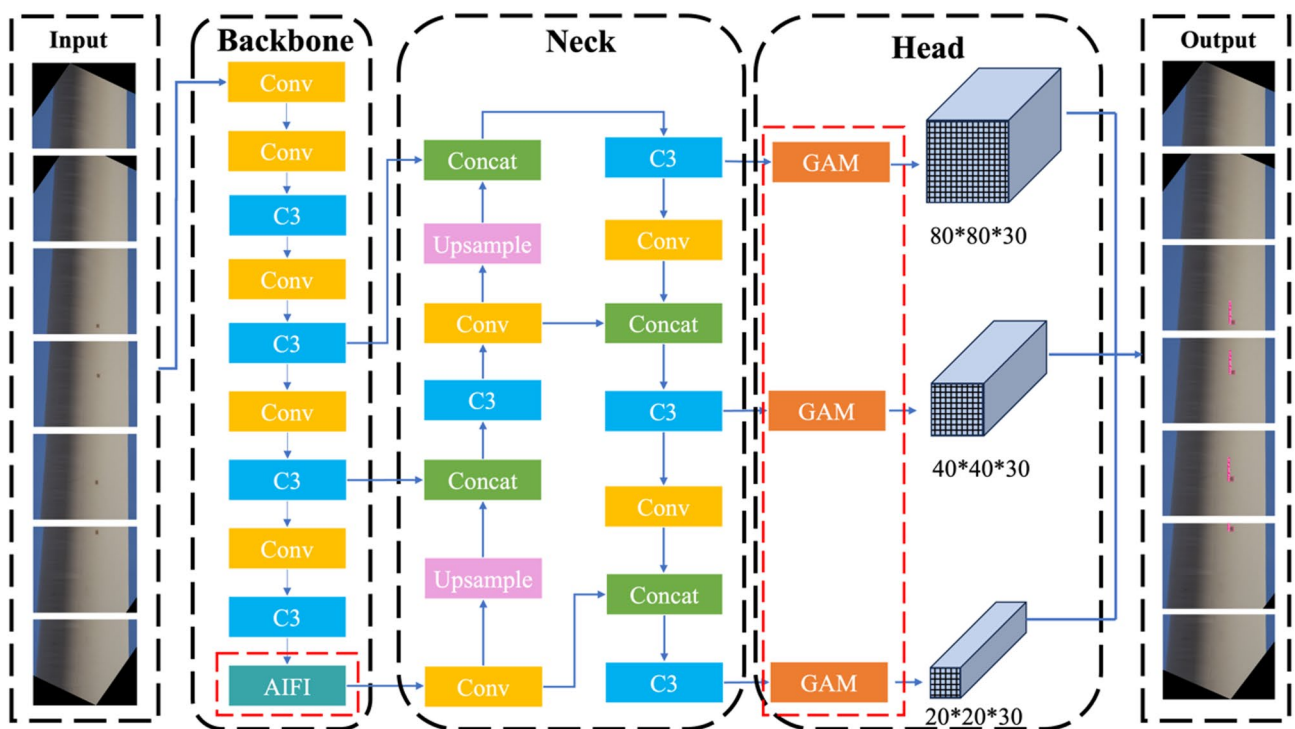

Fig. 3. Overview of improved YOLOv5 network structure.

semantic information[31]. Therefore, this study used the AIFI module to replace the SPPF module at the bottom of the backbone to enhance YOLOv5.

GAM
The structure of the GAM (Global Attention Mechanism) is illustrated in Fig. 5. The GAM can be divided into two parts: channel and spatial. The channel attention mechanism transforms the information into three dimensions. It amplifies the channel-spatial dependency through the MLP, effectively integrating channel
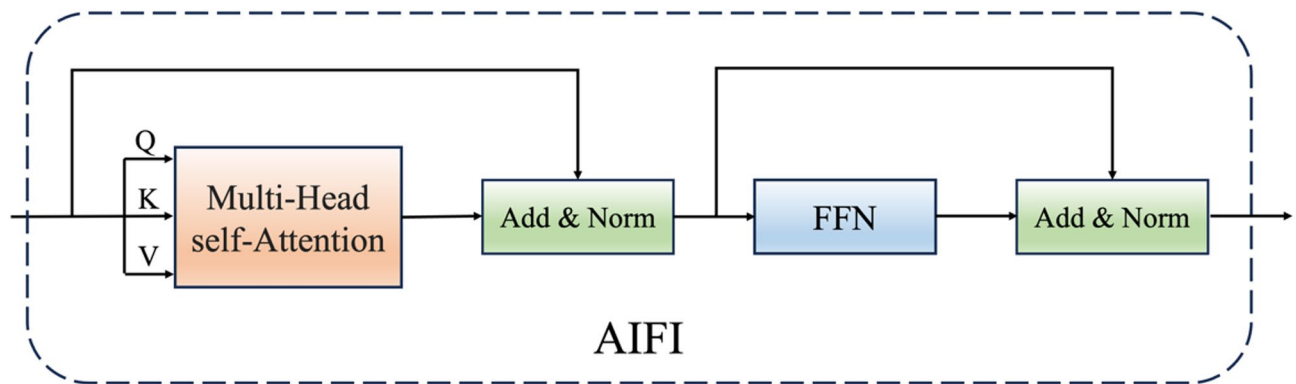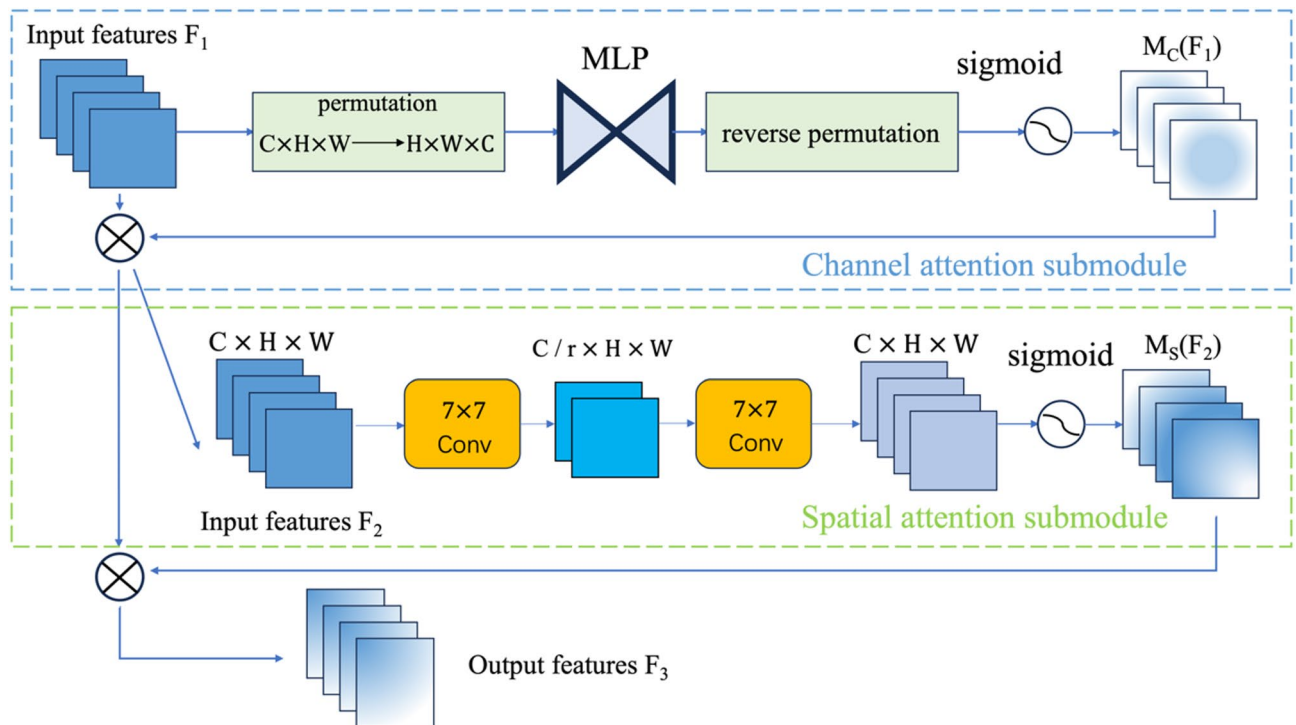
Fig. 4. Structure of AIFI.

Fig. 5. Structure of GAM.

information with spatial information, adding cross-dimensional information, and capturing important features across the three dimensions. Two convolutional layers are used for spatial information fusion in the spatial attention mechanism, employing the same reduction ratio as that in the channel attention module. Unlike other attention mechanisms such as CBAM[36], it does not use pooling operations, thereby preserving the spatial information as much as possible, which is suitable for blade damage characteristics and beneficial for detecting wind turbine blade damage.

The channel attention mechanism of the GAM rearranges the image channels, converting the [c, h, w] to [h, w, c] channels. It transforms the image arrangement of three BGR channels and utilizes MLP to learn nonlinear features, thereby establishing deeper relationships. By comparing the [c, h, w] and [h, w, c] channels, as shown in Fig. 6, it can be observed that [h, w, c] represents a normally captured photograph, whereas [c, h, w] is an image divided into three BGR channels. Although [c, h, w] splits the image into three parts, thus making it easier to discern any channel changes, it also lacks some information. Therefore, when compared to learning only from [c, h, w] channels, introducing [h, w, c] channels can capture features more comprehensively and improve the detection effectiveness of blade damage.

To better incorporate the GAM module, the superiority of the two schemes is explored by introducing the GAM into the head or at the bottom of the backbone. Through the analysis of network principles, it is determined that placing three GAM modules on three feature maps of varying sizes within the head allows for more effective
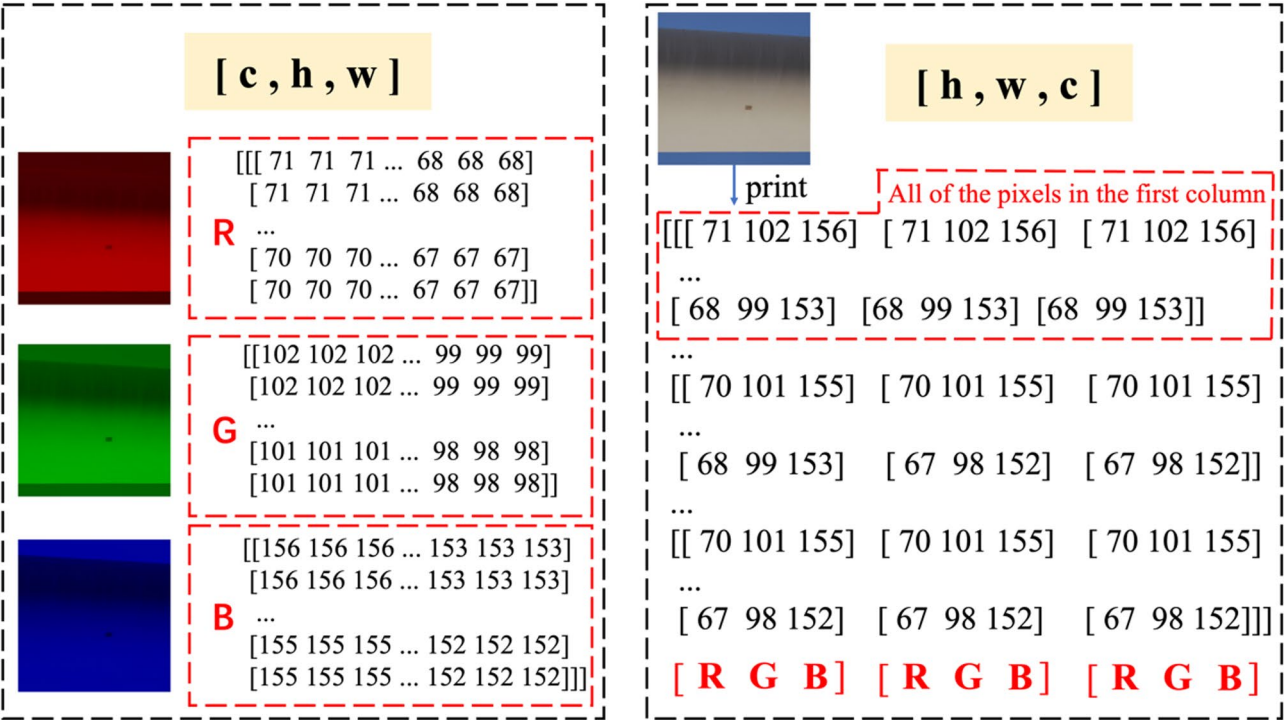
Fig. 6. Comparison of [c, h, w] and [w, h, c] channels on wind turbine blade images.

detection of small damages to wind turbine blades. In contrast, positioning a single GAM module at the bottom of the backbone limits the analysis to the smallest feature map alone, which does not align with the wind turbine blade damage characteristics. Therefore, this paper opted to place three GAM modules on the head.

Experiment design and results

Three experiments are conducted to validate the edge crop method and improve the accuracy of wind turbine blade damage detection. The first experiment involved detecting 1269 images to identify the standard values of the five values. In the second experiment, datasets are created using the original, global crop-processed, and edge crop-processed images, which are then input into the YOLOv5s network for detection. The performances of these datasets are compared to evaluate their detection effectiveness. The third experiment focused on enhancing the YOLOv5 network based on variable-shape damage and optimizing it for wind turbine blade damage detection tasks.

Experiment of the edge crop
Edge crop program verification

The edge crop, programmed by Python, is depicted in Algorithm 1.

Input： Origin image

Output： Multiple cropped images

Extracting lines from blade images by Canny and Hough functions；

Filter out two parallel lines and rotate the image based on the slope of the lines；

Filter two straight lines on sides of the blade edge based on five value and draw；

Crop the blade image based on the two lines；

Value	Standard value range
Δk	$|\Delta k| < 0.17$
	$l > 50$
d	$d > 300$
$\overline{y}$	$y_i < \overline{y} < y_{i+1}$
k_i	$|k_i| < 0.1$

Table 1. The standard values tested by experiments.

Image resolution	Proportion
3936×2624	40%
7952×5304	20%
5008×3336	10%
7719×5149	10%
Other	20%

Table 2. Image pixel statistics.

Fig. 7. Two images that edge crop is having trouble processing.

A total of 1269 wind turbine blade images are tested in this study. The testing environment is Python 3.8 and OpenCV 4.9.0.80. The results yielded approximate standard values, as listed in Table 1.

Because the pixel sizes of the datasets in this study are inconsistent, the five values are not fixed and require adjusting according to the images. The experimental results are as follows: Using the Standard value range in Table 2, 1026 of 1269 images are processed with a processing rate of 80.85%. After adjusting the value range based on the characteristics of different images, 1226 images are processed, yielding a processing rate of 96.61%. The remaining 3.39% of the images are difficult to process even with the adjusted data. Therefore, the edge crop method proposed in this study, which is based on edge detection, can be applied to most wind turbine blade images.

Challenging images for edge crop
Difficult-to-process images are analyzed to verify the robustness of the edge crop. Figure 7 shows typical images.

Figure 7a depicts a portion of the wind turbine blade, and a complete image of the blade is not captured. Therefore, it is not possible to apply two straight-edged lines. However, the edge of the image can be used as another edge to crop the blade image, as illustrated in Fig. 8a, where the red line represents the selected edge. In Fig. 7b, owing to factors such as the shooting angle, blade shape, and deformation, the edges on both sides of the blade are not straight, thus making it difficult to crop, as shown in Fig. 8b.

In summary, it is concluded that the edge crop method proposed in this paper only applies to wind turbine blade images in which both edges of the blade are straight. For more complex images, more parameters must be adjusted. However, owing to the Canny operator, the background of the blade had little influence on the selection of the edge lines. As long as there is a noticeable gradient change, the method can still process images effectively, even when affected by factors such as the image capture angle or low illumination on cloudy days. In the dataset used in this study, 96.61% of wind turbine blade images are processed using the proposed edge-crop method.

Compare edge crop with global crop
Data set composition
This study's wind turbine blade dataset is from western Inner Mongolia, China. A total of 39,251 images were captured using drones, and 1269 high-quality images were retained after filtering, which depicted operational
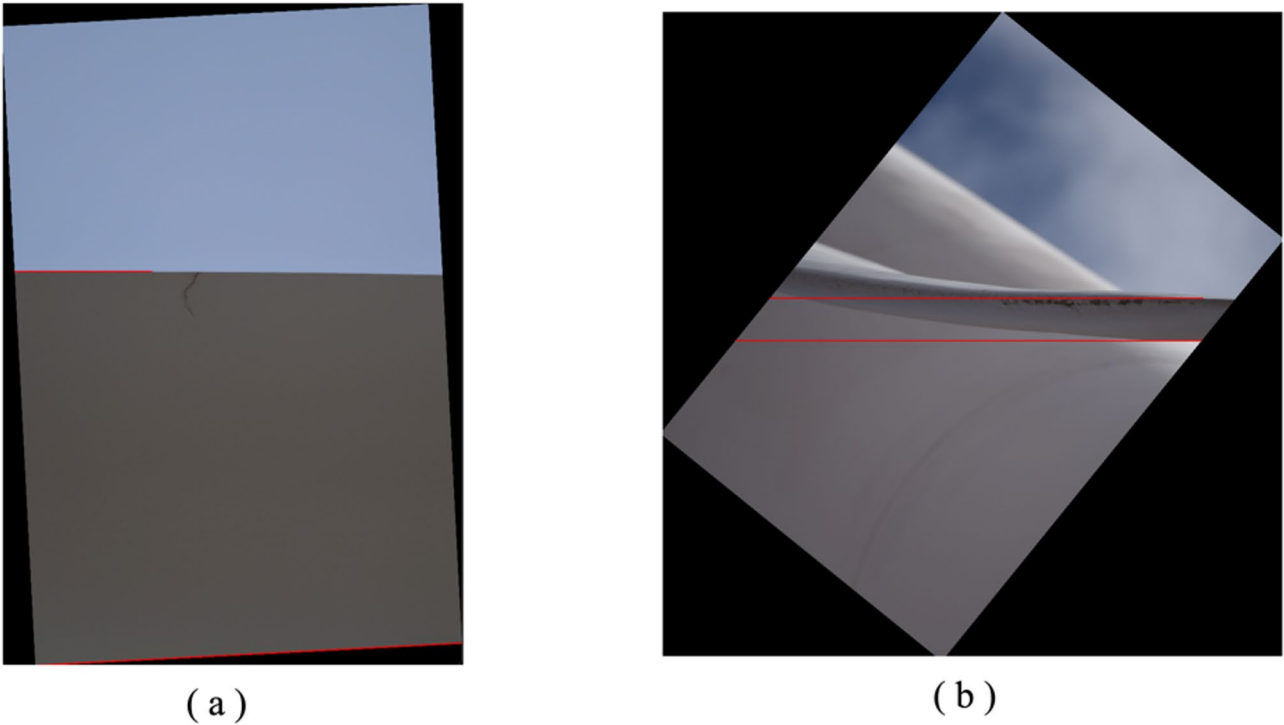
Fig. 8. Figure 7a and b processing.

	Lacquer peels off	Scratch	Stains	Crackle	Lightning strike
Train set	1244	102	315	104	35
Test set	94	12	56	15	5

Table 3. The data instances of five types of damage in the original dataset.

	Lacquer peels off	Scratch	Stains	Crackle	Lightning strike
Train set	1244	306	315	312	245
Test set	94	12	56	15	5

Table 4. The data instances of five types of damage in the expanded dataset.

blades that had sustained damage. In the dataset of 1269 images, all image resolutions are greater than 10 megapixels, with high-resolution images exceeding 40 megapixels. The images are categorized and organized according to their resolutions, as listed in Table 2. Consequently, the challenge lies in processing these high-resolution images to ensure their efficient utilization becomes a pertinent issue.

The original dataset is divided into training and test sets at 14:3:3, as shown in Table 3. Data on lightning strikes, scratches, and cracks are relatively scarce among the five types of damage. Therefore, this study enhanced the data to increase diversity, prevent overfitting during deep learning training, and reduce the model's generalization ability. Two augmentation methods are used for scratches and cracks: sharpening and contrast enhancement. Owing to the extreme scarcity of lightning strike datasets, seven augmentation methods are employed: sharpening, contrast enhancement, Gaussian noise, median blur, motion blur, brightness increase, and saturation adjustment. These augmentation methods are applied only to the training set. To prevent augmentation from causing any accuracy distortion, only real images are detected, and no augmentation is performed on the test and validation sets. Consequently, an augmented dataset is obtained, as shown in Table 4.

Following a discussion with the engineer, the damage is classified into five categories based on its characteristics, as shown in Fig. 9. Lacquer peels off refers to the peeling of the surface coating on the blade, which is mainly concentrated at the blade's leading edge. A scratch refers to superficial scratches on the blade caused by sand, stones, or other materials that have minimal impact on the blade. Stains refer to the contamination of the blade surface, such as oil stains, which have little effect on the blade. Crackles are cracks that can damage the internal
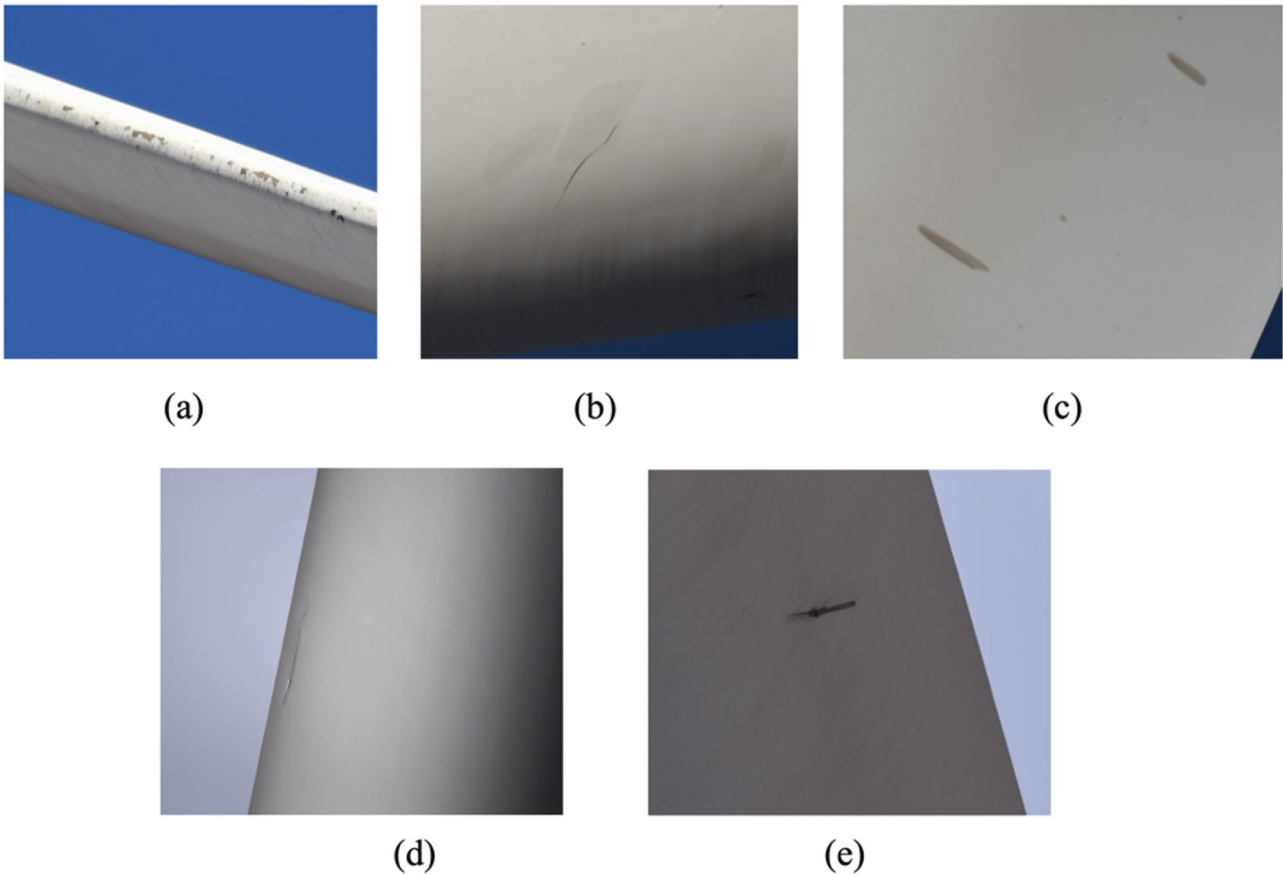
Fig. 9. Five types of damage to wind turbine blades: (**a**) lacquer peels off; (**b**) scratch; (**c**) stains; (**d**) crackle; (**e**) lightning strike.

	Lacquer peels off	Scratch	Stains	Crackle	Lightning strike
Train set	3649	699	940	726	521
Test set	245	25	104	32	11

Table 5. The data instances of five types of damage after global crop.

	Lacquer peels off	Scratch	Stains	Crackle	Lightning strike
Train set	2280	210	419	248	90
Test set	140	14	64	18	6

Table 6. The data instances of five types of damage after edge crop.

structure of a wind turbine blade. A lightning strike refers to a part of a wind turbine blade being struck by lightning.

As shown in Tables 5 and 6, the edge crop resulted in fewer data points than the global crop after applying the same dataset, with a ratio of approximately 1:2. The edge crop offers three advantages over the global crop. First, the edge crop can specifically crop the wind turbine blade area without cutting the background, which makes data processing more convenient. Second, the edge crop allows cropping based on the distance between the two edge lines of the wind turbine blade in the image. However, the global crop cannot adapt to pixel changes in wind turbine blade images. The crop image size in the edge crop can be selected, such as 640×640, 1280×1280, or the method used in this study, where the distance between the two lines is taken as the height and width of the image. Half of the distance between the two lines is taken as the step size of the width for clipping, which is adjusted according to the processing requirements. Third, the edge crop method has significantly higher cropping efficiency than the global crop method, achieving approximately double the efficiency.

Experimental environment
The experimental environment is based on Linux, and the GPU is an NVIDIA GeForce RTX 4060 Ti (16G). Model training and data testing are performed using Pytorch version 1.8.2 and CUDA version 12.4. The imported image size is 640×640, the batch size is 16, the epoch is 300, and the weight is yolov5s.pt, as provided by the official. Standard values are used for the other hyperparameters.

Evaluation index
To compare the accuracy of the algorithm network, the average accuracies of mAP0.5 and mAP0.25 are used as precision indices. The formulas are as follows:

$$P = \frac{TP}{TP + FP} \tag{8}$$

$$R = \frac{TP}{TP + FN} \tag{9}$$

$$AP = \int_0^1 P(R)dR \tag{10}$$

$$mAP = \frac{1}{|n|}\sum_{i=1}^n AP_i \tag{11}$$

In the wind turbine blade damage-detection task, the positive sample is the damage, and the negative sample is the background. In the preceding formula, True Positive (TP) indicates the number of positive samples that are correctly predicted, False Positive (FP) indicates the number of positive samples that are incorrectly predicted, and False Negative (FN) indicates the number of negative samples that are incorrectly predicted. Precision (P) represents the correct number of samples to be predicted in all prediction results, Recall (R) represents the correct number of samples to be predicted in all positive samples, and Average Precision (AP) represents the area of the curve enclosed by the accuracy P and recall rate R of a single injury category. Mean Average Precision (mAP) indicates the average AP for all damage categories. mAP0.5 represents the mean AP values of all damage categories when the intersection ratio between the prediction and real frames is 0.5, and mAP0.25 represents the mean AP values of all damage categories when the intersection ratio between the prediction and real frames is 0.25.

Experimental results and analysis
The images are classified into three categories, as listed in Tables 4, 5, and 6. A represents the original image; B represents the image processed using global crop; and C represents the image processed using edge crop. YOLOv5s network training is used to test the accuracy of the three image types, as listed in Table 7.

The following experimental table is divided into five items: 1) Case, 2) training set, 3) test set, 4) mAP0.5, and 5) mAP0.25. This experiment is divided into four cases, corresponding to separate tests of three datasets: the image clipped by the edge crop is added to the original image as extended data to check whether it is improved in Case 4 to judge the training mode that is most suitable for the clipping method proposed in this study. Because the wind turbine blade damage detection task is an engineering task, the frame selection of the wind turbine blade damage does not need to be too accurate, and the key is to detect the damage and not miss it. Therefore, the mAP0.5 and mAP0.25 are selected for comparison in this study.

The comparison in Table 6 shows that the edge crop proposed in this study shows a significantly larger improvement than the original image. This aligns with the analysis that wind turbine blade damage is often small, concluding that cropped images are beneficial for detecting small damages. Compared with the global crop, the edge crop also performs slightly better in accuracy. As shown in Case 4, when the edge-cropped and original images are used together as the training set and then tested on the original test set A, it is found that, compared to Case 1, using the edge-cropped images for augmentation slightly improved the mAP.

Enhanced YOLOv5
To solve the problem of variable-shape damage of wind turbine blades, the YOLOv5 network is improved by replacing the AIFI at the bottom of the backbone and adding a GAM to the head. Experiments and ablation studies are conducted to compare and validate the enhanced YOLOv5 algorithm with other algorithms in object detection, thereby demonstrating the effectiveness and advancement of the enhanced YOLOv5 method in the wind turbine blade damage detection domain.

Case	Train set	Test set	mAP0.5	mAP0.25
1	A	A	73.4	79.9
2	B	B	84.8	89.0
3	C	C	87.5	91.3
4	A+C	A	74.8	81.0

Table 7. Comparison experiment between global crop and edge crop.

Method	mAP0.5[%]					mAP0.5[%]	Model size (M)	FPS
	Crackle	Lacquer peels off	Lightning strike	Scratch	Stains			
Faster R-CNN	30.6	31.6	65	67.8	68.8	52.7	108.3	12
SSD	30.4	41.2	74.8	78	70.1	58.9	92.6	31
YOLOX	49.7	32.2	99.5	90.9	67.4	68.0	68.52	21
YOLOv3-tiny	49.7	36.6	99.5	87.5	76.9	70.0	17.4	150
YOLOv5s	53.1	48.8	99.5	90.9	74.8	73.4	13.8	128
YOLOv8s	39.1	35.4	99.5	99.5	78.2	70.3	21.4	125
[37]	74.1	47	99.5	99.5	73.9	78.8	–	–
RT-DETR-L	85.7	46.5	99.5	95.5	80.6	81.6	66.2	53
EC-EY	88.0	79.8	99.5	99.5	84.8	90.3	21.5	82

Table 8. Comparison of the EC-EY and other algorithms.

Case	Edge crop	GAM	AIFI	mAP0.5	mAP0.25	Parameters (M)
1				73.4	81.0	7.03
3	✓			87.5	91.3	7.03
5		✓		79.4	83.4	9.31
6			✓	78.9	81.6	8.47
7		✓	✓	81.0	84.3	10.76
8	✓	✓	✓	90.3	93.9	10.76

Table 9. Ablation experiment.

Experimental environment and evaluation index
The experimental environment is the same as the previous experiment, with two additional evaluation metrics: frames per second (FPS) and Model size. The number of image FPS that the model could detect is used as the speed index to determine the speed of the algorithm network. The memory size occupied by the model is used as an evaluation index to determine its complexity.4.3.2 Comparative experiment.

Among the various models listed in Table 8, EC-EY is compared with various classic models in object detection and with previously published results from our research group that utilized the same dataset.

The experimental results show that mAP0.5 of EC-EY is the highest. Compared with the original YOLOv5s model, EC-EY effectively improved the damage accuracy of lacquer peels off by 31%, reaching the lead of the faulting model among all models. Moreover, the dataset contained a large amount of highly robust data. For crackles, mAP0.5 improved by 34.9%, achieving the highest performance among the models used; for stains, lightning strikes, and scratches, mAP0.5 showed negligible difference and is within the normal range. Compared to other models, the EC-EY achieved the highest accuracy and maintained a relatively small model size, the FPS is generally sufficient to meet the requirements of wind turbine blade damage detection tasks.

Ablation experiment
To verify the improved method's effectiveness, ablation experiments are conducted by training the models separately. As shown in Table 9, eight cases are included in the experiment. Because the effectiveness of the edge crop image processing method has already been verified, this experiment only conducted ablation studies on GAM and AIFI, as demonstrated in cases 5, 6, and 7. The final results are obtained by combining the three improvement methods, as shown in case 8.

Table 8 shows that the GAM and AIFI improvement schemes are effective, and the effect is better after joining the network, which verifies the effectiveness of the two schemes in improving the damage accuracy. Compared with the original model, after adding the GAM to the head, mAP0.5 is 6% higher than the original model. After replacing SPPF with AIFI at the bottom of the backbone, mAP0.5 is 5.5% higher, and when both improvements are added, mAP0.5 is 7.6% higher than the original model. After adding the edge crop, mAP0.5 increased by 16.9%, and the number of parameters in the model increased slightly. The experimental results show that the three improvements proposed in this study are effective according to wind turbine blades' small and variable-shape damage characteristics.

As shown in Fig. 10, compared to Case 1, after cropping, the lacquer peels off and crackle damage in Cases 2 And 3, the mAP0.5 increased significantly, and the stains showed some improvement. However, because of the small amount of data, scratches and lightning strikes consistently have high detection rates, making them less meaningful for reference purposes. Case 3, which uses the edge crop method, outperformed Case 2, which uses the global crop method, in terms of mAP0.5. In the section on enhanced YOLOv5, that is, Cases 5, 6, and 7, it can be observed that when compared to the original model (Case 1), the enhanced YOLOv5s (Case 7) show a significant improvement in detecting crack damage. This corresponds to the effective crack damage detection performance of RT-DETR, thereby highlighting the excellence of the AIFI module derived from RT-DETR in the field of crack detection. By integrating the edge crop and enhanced YOLOv5 (Case 8), it can be
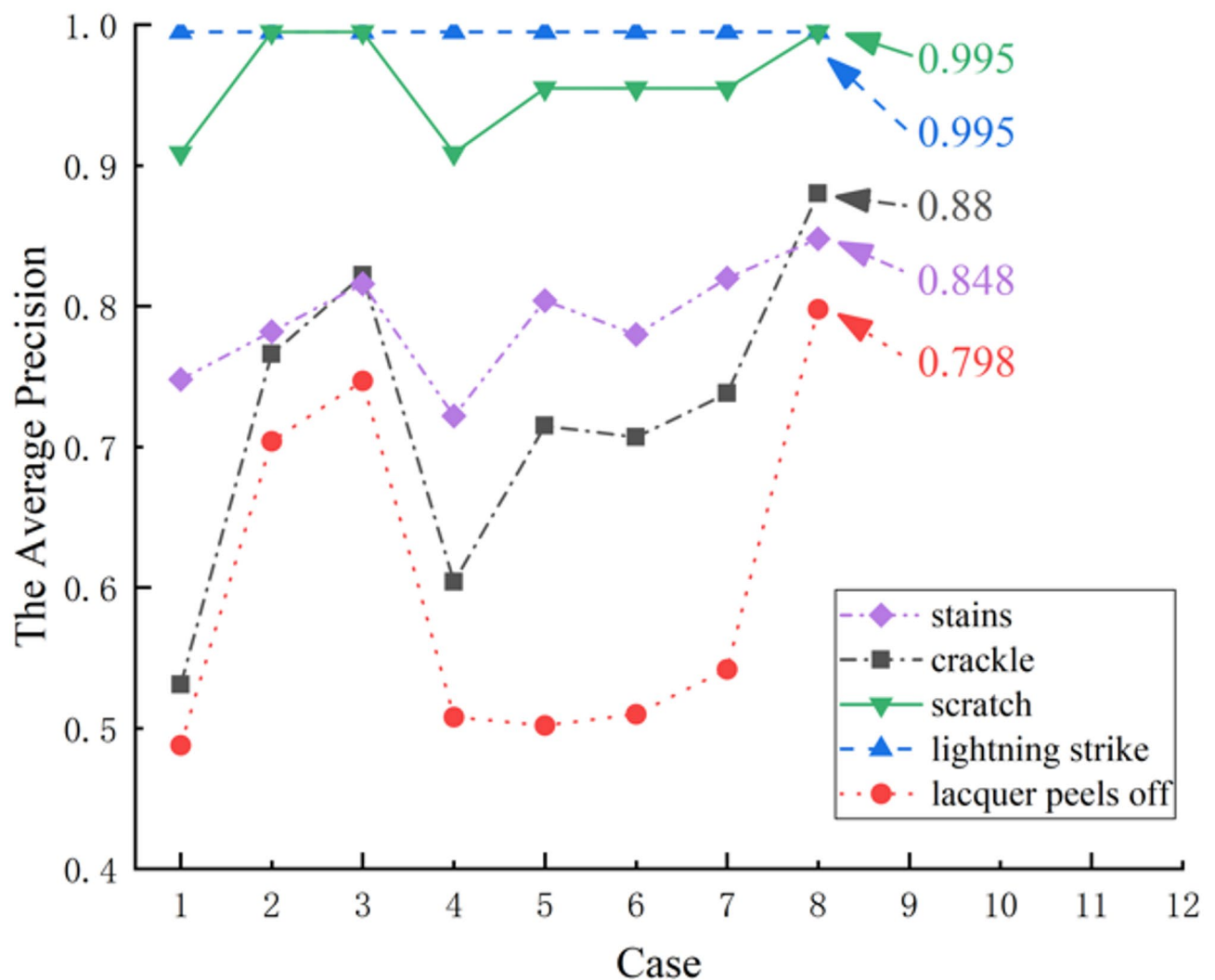
Fig. 10. Average accuracy of eight cases corresponding to five types of damage.

found that compared with Case 3, which only uses the edge crop, Case 8 is improved. Therefore, the edge crop and enhanced YOLOv5 proposed in this study are effective and satisfy the requirements for wind turbine blade damage detection tasks.

Visualization of experimental results
To provide a clearer comparison between the EC-EY and other methods, the detection differences are illustrated in Fig. 11. Specifically, (a) shows a manually annotated original image with multiple damages in a complex environment. (b) displays the detection results using Faster R-CNN, where four damages are identified; however, two are false positives: one detected the background, and the other detected a blank area. (c) the detection results using YOLOv5s identified only two instances of paint peeling damage. (d) shows the results obtained using enhanced YOLOv5s, the results of detecting three instances of paint peeling damages are superior to YOLOv5s. Finally, (e) illustrates the effect of EC-EY, and it can be observed that this approach detects all damages marked manually and identifies some damage that is not initially discovered manually, with clearer results shown in Fig. 12.

Conclusion and discussion

To accurately detect damage to wind turbine blades, this study proposes two methods tailored to address small and variable-shape damage characteristics: edge crop for image processing and enhanced YOLOv5 for network optimization, termed EC-EY. The edge crop method isolates the blade from the background, eliminating most of the background influence, thereby improving the detection accuracy and optimizing image pixel utilization. This approach primarily mitigates the challenge posed by the small extent of the damage to wind turbine blades. The enhanced YOLOv5 method enhances feature learning by introducing a GAM into the head and replacing the SPPF module with the AIFI module. These modifications strengthen the ability of the network to capture nonlinear and long-range features, effectively addressing the issue of variable-shape damage.
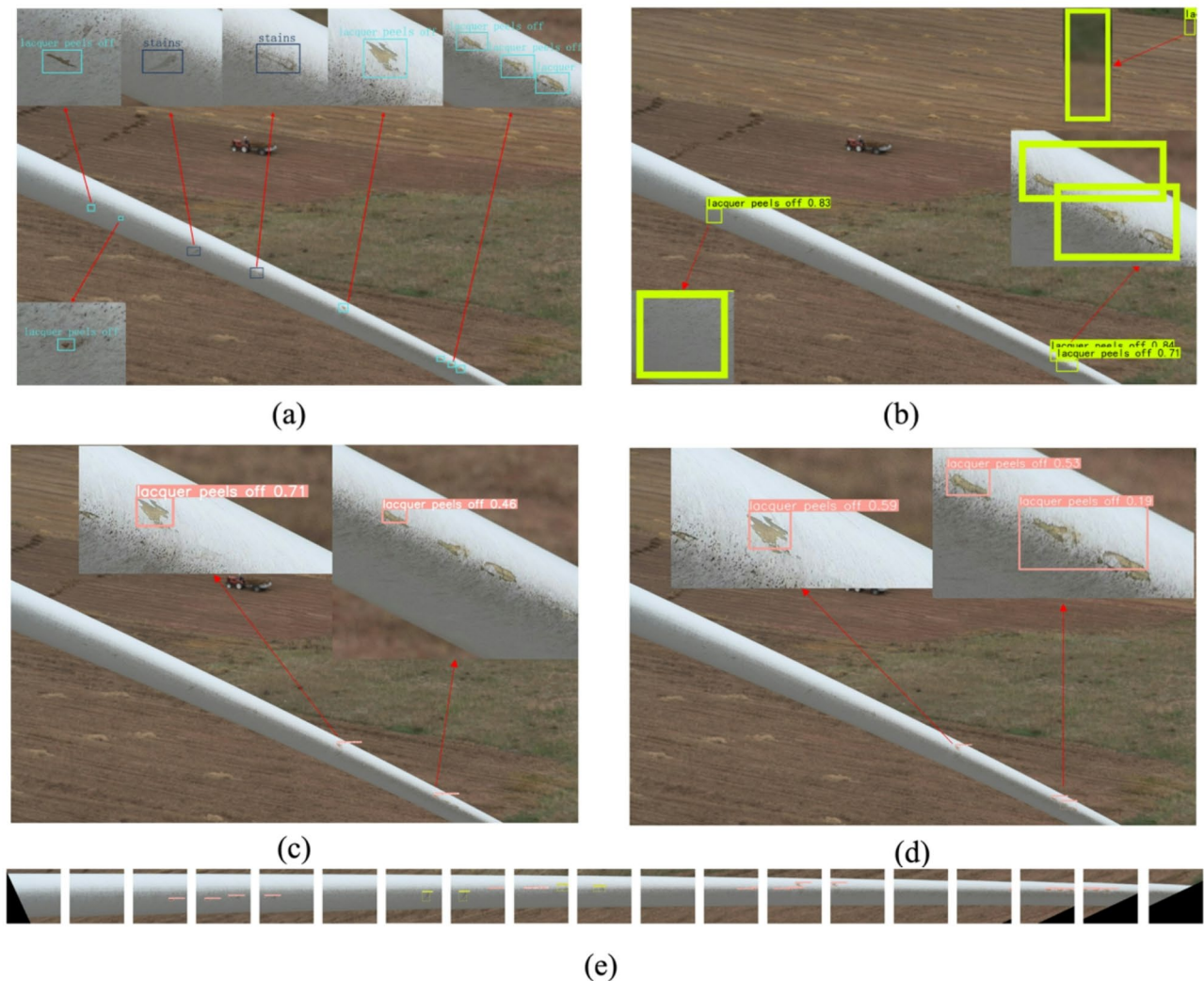
Fig. 11. Comparison of methods: (**a**) Original image; (**b**) Faster R-CNN; (**c**) YOLOv5s; (**d**) enhanced YOLOv5s; (**e**) EC–EY.

The experimental results showed that the edge crop method outperformed the global crop in terms of cropping effectiveness and detection accuracy. As a preprocessing step, edge crop improves detection performance and can be used to expand image datasets. This method effectively resolves the challenge of detecting small-sized damage, such as paint peeling, among the five damage types on wind turbine blades, laying the foundation for comprehensive damage detection. The conceptual contribution of this study lies in the proposition of five mathematical relationships that facilitate the selection of the two edge lines of a blade following Canny edge detection.

Comparative experiments with various algorithms demonstrated that the enhanced YOLOv5s achieves high detection accuracy, maintains a compact model size, and ensures fast detection speeds, making it suitable for variable-shape damage detection of wind turbine blades in western Inner Mongolia. Combining enhanced YOLOv5s with the edge crop method, referred to as the EC-EY method, significantly improves accuracy while maintaining the model size and detection speed, effectively meeting the requirements of wind turbine blade damage detection tasks.

However, the edge crop method faces challenges in rapidly and automatically processing all images. In future research, we will explore adaptive Canny operator approaches to optimize edge detection, aiming to eliminate parameter tuning steps and achieve fully automated processing of all images. Furthermore, experimental validation will be conducted on novel algorithms to resolve this engineering challenge ultimately.
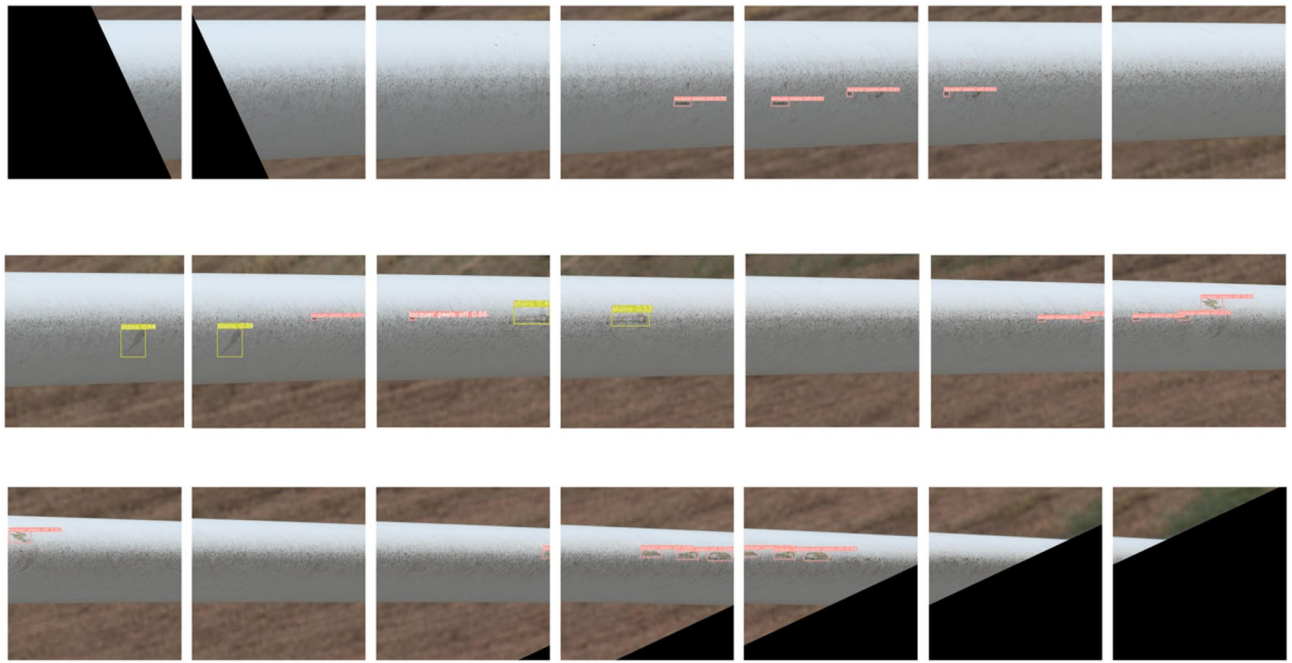
Fig. 12. Clear version of Fig. 11e.

Data availability

Data is provided within the manuscript or supplementary information files.Due to commercial reasons, we cannot release the dataset publicly, if anyone wants to request other data from this study, please contact the email: songli@imut.edu.cn.

References

1. Sun, S., Wang, T. & Chu, F. In-situ condition monitoring of wind turbine blades: A critical and systematic review of techniques, challenges, and futures. *Renew. Sustain. Energy Rev.* **160**, 112326. https://doi.org/10.1016/j.rser.2022.112326 (2022).
2. Wang, W. et al. Review of the typical damage and damage-detection methods of large wind turbine blades. *Energies* **15**(15), 5672. https://doi.org/10.3390/en15155672 (2022).
3. Chu, H.-H & Wang, Z.-Y. A vision-based system for post-welding quality measurement and damage detection. *Int. J. Adv. Manuf. Technol.* **86**(9–12), 3007–3014. https://doi.org/10.1007/s00170-015-8334-1 (2016).
4. Sierra-Pérez, J., Torres-Arredondo, M. A. & Güemes, A. Damage and nonlinearities detection in wind turbine blades based on strain field pattern recognition. FBGs, OBR and strain gauges comparison. *Compos. Struct.* **135**, 156–166. https://doi.org/10.1016/j.compstruct.2015.08.137 (2016).
5. Theresa, L. & Alexander, B. Vibration-based fingerprint algorithm for structural health monitoring of wind turbine blades. *Appl. Sci. Basel* **11**(9), 4294–4294. https://doi.org/10.3390/app11094294 (2021).
6. Ying, D. et al. Damage detection techniques for wind turbine blades: A review. *Mech. Syst. Signal Process.* **141**, 106445. https://doi.org/10.1016/j.ymssp.2019.106445 (2020).
7. Xiaowen, S. et al. Review on the damage and fault diagnosis of wind turbine blades in the germination stage. *Energies* **15**(20), 7492. https://doi.org/10.3390/en15207492 (2022).
8. Marco, C. & Cecilia, S. Non-destructive techniques for the condition and structural health monitoring of wind turbines: A literature review of the last 20 years. *Sensors* **22**(4), 1627–1627. https://doi.org/10.3390/s22041627 (2022).
9. Aminzadeh, A. et al. Non-contact inspection methods for wind turbine blade maintenance: Techno–economic review of techniques for integration with industry 4.0. *J. Nondestruct. Eval.* https://doi.org/10.1007/s10921-023-00967-5 (2023).
10. Li, H., Wang, W. & Wang, M. A review of deep learning methods for pixel-level crack detection. *J. Traffic Transp. Eng.* **9**(6), 945–968. https://doi.org/10.1016/j.jtte.2022.11.003 (2022).
11. Bi, Li. & Quanjie, G. Defect detection for metal shaft surfaces based on an improved YOLOv5 algorithm and transfer learning. *Sensors* **23**(7), 3761–3761. https://doi.org/10.3390/s23073761 (2023).
12. Zhang, R. & Wen, C. SOD-YOLO: A small target damage detection algorithm for wind turbine blades based on improved YOLOv5. *Adv. Theory Simul.* https://doi.org/10.1002/adts.202100631 (2022).
13. Zhou, W. et al. Wind turbine actual defects detection based on visible and infrared image fusion. *IEEE Trans. Instrum. Meas.* **72**, 1–8. https://doi.org/10.1109/TIM.2023.3251413 (2023).
14. Xiaoxun, Z. et al. Research on crack detection method of wind turbine blade based on a deep learning method. *Appl. Energy* **328**, 120241. https://doi.org/10.1016/j.apenergy.2022.120241 (2022).
15. Denhof, D. et al. Automatic optical surface inspection of wind turbine rotor blades using convolutional neural networks. *Procedia CIRP* **81**, 1166–1170 (2019).
16. Jiajun, Z., Georgina, C. & Jason, W. Image enhanced mask R-CNN: A deep learning pipeline with new evaluation measures for wind turbine blade damage detection and classification. *J. Imaging* **7**(3), 46–46. https://doi.org/10.3390/jimaging7030046 (2021).
17. Huang, C., Chen, M. & Wang, L. Semi-supervised surface defect detection of wind turbine blades with YOLOv4. *Glob. Energy Interconnect.* **7**(3), 284–292. https://doi.org/10.1016/j.gloei.2024.06.010 (2024).

18. Narges, K. & Vida, M. A critical review and comparative study on image segmentation-based techniques for pavement crack detection. *Construct. Build. Mater.* **321**, 126162–126162. https://doi.org/10.1016/j.conbuildmat.2021.126162 (2022).
19. Sumika, C. et al. Parallel structure of crayfish optimization with arithmetic optimization for classifying the friction behaviour of Ti-6Al-4V alloy for complex machinery applications. *Knowl.-Based Syst.* **286**, 111389. https://doi.org/10.1016/j.knosys.2024.111389 (2024).
20. Sumika, C. & Govind, V. A synergy of an evolutionary algorithm with slime mould algorithm through series and parallel construction for improving global optimization and conventional design problem. *Eng. Appl. Artif. Intell.* **118**, 105650. https://doi.org/10.1016/j.engappai.2022.105650 (2023).
21. Hussain, M. YOLO-v1 to YOLO-v8, the rise of YOLO and its complementary nature toward digital manufacturing and industrial damage detection. *Machines* **11**(7), 677. https://doi.org/10.3390/machines11070677 (2023).
22. Tang, Z. et al. Adaptive segmentation method for wind turbine blades combining Hough line detection and Grab-cut algorithm. *J. Electron. Measur. Instrum.* **35**(4), 161–168. https://doi.org/10.13382/j.jemi.B2003392 (2021).
23. Ali, A. Application of quad-copter target tracking using mask based edge detection for feasibility of wind turbine blade inspection during uninterrupted operation. (2023).
24. Li, J. et al. Survey of transformer-based object detection algorithms. *Comput. Eng. Appl.* **59**(10), 48–64. https://doi.org/10.3778/j.issn.1002-8331.2211-0133 (2023).
25. Joseph, R.,Santosh, D., Ross, G., et al. You only look once: Unified, real-time object detection. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 779–788 (2016).
26. Ross, G., Jeff, D., Trevor, D., et al. Rich feature hierarchies for accurate object detection and semantic segmentation. In *27th IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 580–587 (2014).
27. Shaoqing, R. et al. 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 (2017).
28. He, K. M. et al. Mask R-CNN. *IEEE Trans. Pattern Anal. Mach. Intell.* **42**(2), 386–397. https://doi.org/10.1109/TPAMI.2018.2844175 (2020).
29. Nicolas, C., Francisco, M., Gabriel, S., et al. End-to-End object detection with transformers. arXiv, (2020).
30. Zhu, X., Su, W., Lu, L., et al. Deformable detr: Deformable trans- formers for end-to-end object detection. 6(7) (2020). arXiv:2010.04159
31. Yian, Z., Wenyu, L., Shangliang, X. DETRs beat YOLOs on real-time object detection. arXiv, https://doi.org/10.48550/arXiv.2304.08069 (2023).
32. Yilin, G., Dapeng, J. & Liping, S. Wood veneer damage detection based on multiscale DETR with position encoder net. *Sensors* **23**(10), 4837–4837. https://doi.org/10.3390/s23104837 (2023).
33. Yang, C. & Daming, L. An image-based deep learning approach with improved DETR for power line insulator damage detection. *J. Sens.* https://doi.org/10.1155/2022/6703864 (2022).
34. Minggao, L. et al. Bearing-DETR: A lightweight deep learning model for bearing damage detection based on RT-DETR. *Sensors* **24**(13), 4262–4262. https://doi.org/10.3390/s24134262 (2024).
35. Yichao, L., Zongru, S. & Nico, H. Global attention mechanism: Retain information to enhance channel-spatial interactions. arXiv (2021).
36. Sanghyun, W., Jongchan, P., Joon-Young, L. et al. CBAM: Convolutional block attention module. arXiv (2018).
37. Xing, Y. et al. Research on wind turbine blade damage detection based on image recognition technology. *J. Eng. Thermophys.* **46**(1), 92–97 (2025).

Acknowledgements

Author contributions

Boyu Feng: Methodology, Software, Writing—original draft. Bo Liu: Writing—reviewing & editing, Conceptualization. Li Song: Data Curation, Funding acquisition. Yongyan Chen: Supervision. Xiaofeng Jiao: Data Curation. Baiqiang Wang: Validation.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-04882-9.

Correspondence and requests for materials should be addressed to L.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.