Automated surface defect detection of stay cables via UAV route planning and deep learning model

Baoquan Wang [a,b] , Dongming Feng [a,b,c],*, Zhichao Fan [a,b], Hanyu Shi [a,b]

[a] Key Laboratory of Concrete and Prestressed Concrete Structures of the Ministry of Education, Southeast University, Nanjing, 210096, Jiangsu, China
[b] School of Civil Engineering, Southeast University, Nanjing, 210096, Jiangsu, China
[c] National Key Laboratory of Safety, Durability and Healthy Operation of Long Span Bridges, Nanjing, 211189, Jiangsu, China

ARTICLE INFO

ABSTRACT

Stay cables are critical load-bearing components of cable-stayed bridges. To support automated inspection, this paper proposes a framework that integrates UAV flight path planning with deep learning models to detect surface defects in stay cables. A rough three-dimensional (3D) model of the bridge is first constructed via UAV-based oblique photogrammetry, followed by the design of an optimized automatic inspection path. Image data of the stay cables are captured, and image augmentation and enhancement techniques are applied to address the limited sample size. Deep learning models are then used to detect defects, and assess their physical dimensions. A practical engineering case study validates the effectiveness of the proposed inspection path and defect detection methodology. Results show that UAV-based automatic inspection path planning, grounded in 3D model, is feasible and that the Yolov8-seg model effectively detects defects, with the image enhancement and enhancement strategies improving recognition accuracy.

1. Introduction

As a critical load-bearing component in cable-stayed bridges, the safety status of stay cables significantly influences the overall structural performance and safety of the bridge during its service life [1]. Under constantly changing operating conditions, the polyethylene (PE) sheaths surrounding the stay cables are prone to aging and damage [2], leading to surface defects. These defects typically manifest as cracks, scratches, rust, and holes, which expose the internal steel wires to the external environment and create favorable conditions for the infiltration of moisture and corrosive agents. Such an environment, once formed, significantly accelerates the corrosion of the steel wires [3] and directly impacts the service life of the cable-stayed bridge [4]. Corrosion not only markedly accelerates the structural damage and degradation of the mechanical performance of cables, but also exacerbates the potential risks to the overall safety and stability of bridges. This detrimental cycle can result in a significant reduction in the tensile strength of the cables, thereby compromising the bridge's normal load-bearing capacity and ultimately jeopardizing the long-term safety and structural reliability of the bridge [5]. Therefore, the rapid and accurate detection of surface defects in cable-stayed bridges is crucial for evaluating their operational status and informing subsequent maintenance strategies.

Currently, the detection methods for surface defects of stay cables primarily rely on manual inspections [6–9], cable robot inspections [10–15], and sensor-based detection techniques [16–20]. Manual inspections involve inspectors using telescopes, magnifying glasses, or climbing the bridge structure to directly assess cable damage. While this method is easy to operate and offers an intuitive view of cable conditions, it suffers from significant limitations, including the subjectivity of inspectors, low efficiency, and elevated safety risks. Alternatively, manual inspection can also be conducted using inspection trolleys [6], which are mounted on the stay cables to carry inspectors for a more thorough examination. This approach not only enables comprehensive inspection but also facilitates cleaning and maintenance, thus reducing safety risks associated with climbing. However, the high cost of purchasing and maintaining these trolleys, along with potential disruption to bridge traffic, are notable drawbacks. Furthermore, inspection trolleys may be ineffective in areas with complex structures, narrow spaces, or impassable regions. To enhance the automation of cable inspections, researchers have developed robot-based inspection systems. These systems involve robots climbing along the cables, equipped with sensors and cameras for inspection. For example, Ho et al. [11] utilized a climbing robot to travel along the stay cables and capture surface images, which were subsequently analyzed using image processing techniques for damage detection. Cho et al. [12] proposed a

---

multifunctional robot tracking system tailored for suspension bridges, which accurately detects cable surface damage. Li et al. [13] developed a cable surface damage detection system integrating a climbing robot with four fixed cameras, while additional surface defect detection technologies included acceleration sensors [14] and laser scanning [15]. Cable robots can replace manual inspections, autonomously traverse the cables, and efficiently perform comprehensive inspections, thereby improving inspection speed and effectiveness. However, the high cost of these robots, coupled with their reliance on specific cable structures, surface materials, and environmental conditions, limits their applicability. In particular, cables with complex structures, high curvature, or inaccessibility may not be effectively inspected. Sensor-based detection techniques are commonly integrated with signal processing algorithms, employing technologies such as fiber optic sensing [16, 17], ultrasound [18], acoustic emission [19], and magnetic flux leakage [20] to detect defects in stay cables. However, these techniques are generally more suitable for interior defect detection, such as interior corrosion and cracks of cables. These techniques offer notable advantages in terms of real-time performance and detection accuracy, but they also present some application limitations. For instance, fiber optic sensing is hindered by complex wiring requirements and susceptibility to temperature fluctuations, and acoustic emission technology is prone to interference from environmental noise.

Recently, advancements in artificial intelligence (AI) have facilitated the application of deep learning models in bridge health monitoring [21–26], particularly in the widespread use of image processing technology related deep learning models for structural damage detection, encompassing surface defects such as cracks [27–29], holes, and multiple forms of damage [30–32]. For instance, Li et al. [31] applied fully convolutional networks to achieve pixel-level crack detection, identifying four types of damage on concrete surfaces: cracks, peeling, weathering, and holes. Wang et al. [32] employed Faster R-CNN and ResNet101 convolutional neural networks to detect efflorescence and spalling on historic masonry building surfaces, achieving high detection accuracy. Bae et al. [33] combined UAVs with climbing robots to detect cracks in concrete bridges using a deep super-resolution crack network. Feng et al. [34] proposed a dam surface crack detection model based on convolutional neural networks, effectively segmenting crack images. Kim et al. [35] optimized CNN hyperparameters to improve the accuracy of concrete crack detection, integrating UAV technology for real-time inspections. Bukhsh et al. [36] introduced a hybrid strategy of intra-domain and cross-domain transfer learning, validating its advantages on small bridge damage datasets. Jiang et al. [37] developed a wall-climbing UAV system using CNNs to extract crack information. Ding et al. [38] proposed a quantifiable detection method for cracks below 0.2 mm, while Liu et al. [39] used UAVs to inspect bridge pier cracks and suggested correction methods for perspective and geometric distortions. Additionally, the study conducted by Dorafshan et al. [40, 41] revealed that a cost analysis of a 2400-m-long bridge inspection in the United States, carried out in 2017, indicated that UAV-based inspections (with an estimated cost of $20,000) were 66% more cost-effective and 37% faster compared to traditional manual inspection methods (which incurred a cost of $59,000). A further example, the inspection of ITD Bridge Key 21105 [40–42], demonstrates that the cost of traditional manual inspection is $391 per hour, whereas UAV inspections are priced at $200 per hour, making the UAV-based inspections nearly 50% less expensive on an hourly basis. With the ongoing advancements in technology and the increasing commercialization of UAVs, it is anticipated that the cost of UAV inspection equipment will continue to decrease in the future. The deep learning methods mentioned above show great promise in the application of structural health monitoring, offering precise and efficient detection and segmentation of surface defects. However, these methods often rely on large datasets, which poses a significant challenge when acquiring defect images of structures. UAVs equipped with high-definition cameras facilitate the collection of defect images from structures. Although several of the

aforementioned studies have employed UAVs for capturing structural defect imagery, their focus has predominantly been on the acquisition of defect images from flat surfaces, such as bridge piers and buildings, where the spatial distribution of defects is relatively straightforward. These applications require minimal obstacle avoidance during UAV flights and do not necessitate complex flight path planning. In contrast, the challenges associated with maintaining line-of-sight and avoiding collisions, arising from the complex spatial configuration of stay cables, significantly hinder the achievement of the required inspection accuracy and efficiency. Despite the potential of deep learning models, these methods typically require large datasets, presenting a challenge in obtaining structured images. UAVs, equipped with high-definition cameras, offer flexibility and efficiency for inspecting large structures like cable-stayed bridges. However, the complex spatial distribution of stay cables, coupled with challenges in maintaining line-of-sight and avoiding collisions, poses difficulties in meeting the required inspection accuracy and efficiency.

Therefore, flight path planning is of great significance for bridge defect detection using UAV. Firstly, as previously mentioned, stay cables are densely distributed in space, making efficient path planning crucial for ensuring sufficient visibility and avoiding collisions. Secondly, to guarantee accurate defect image acquisition, a certain degree of overlap between consecutive images must be maintained. Non-optimized flight paths may result in excessive redundant waypoints and frequent attitude adjustments, leading to a substantial increase in energy consumption. For instance, optimizing the flight path can minimize unnecessary detours and altitude variations, thus reducing energy consumption. Previous studies have demonstrated that optimized paths can reduce energy consumption by 10% to 37% compared to traditional routes [43–45]. Moreover, for large and complex structures, such as cable-stayed bridges, optimizing the flight path can significantly reduce inspection time, alleviate UAV workload, extend operational lifespan, enhance inspection efficiency, and decrease overall inspection costs. Particularly, in long-term operation and maintenance of bridges, the benefits of energy conservation and emission reduction are more pronounced. In summary, by leveraging UAV-based automated inspection technology, flight path planning facilitates automation, standardization, and obstacle avoidance, thereby improving the safety, reliability, and efficiency of cable inspection processes. Specifically, oblique photography can be employed to generate 3D models of bridges and plan inspection routes accordingly. This technology captures images from various angles, processes them through multi-view synthesis, and performs 3D reconstruction. UAV path planning designs flight routes based on environmental conditions and mission requirements, considering factors such as task type and geographic context, to ensure the safe and efficient completion of tasks. For complex terrains, scholars have proposed path planning strategies based on 3D models, which obtains the rough 3D model of the target scene, and generates an optimized viewpoint set based on the model information. Path planning algorithms then determine the shortest flight route. For instance, Cheng et al. [46] used a simplified building model to optimize the path planning of urban buildings; Dogru et al. [47] developed a flight path planning method that accounts for obstacles and terrain factors, optimizing energy consumption through genetic algorithms to achieve cost-efficient flight trajectories. Mansouri et al. [48] proposed a spatial slicing model-based path planning method for complex environments. Zhang et al. [49] introduced a method that integrates Building Information Modeling (BIM) and Geographic Information System (GIS) data to automate UAV flight and route execution. Han et al. [50] proposed a framework combining oblique photography with point cloud data to extract bridge alignments, demonstrating its effectiveness in multi-view image collection on a multi-span bridge under construction. In summary, while 3D model-based UAV path planning has been extensively applied in fields like surveying and cultural heritage preservation, its use in bridge health monitoring remains underexplored. Although some research has focused on UAV path planning for beam bridge

inspections [50–52], studies on UAV path planning for cable-stayed bridge inspections are scarce [53].

To address the aforementioned research gap, this paper proposes a method for detecting surface defects in stay cables of cable-stayed bridges by integrating UAV flight path planning with deep learning models. First, a preliminary 3D model of the cable-stayed bridge is constructed using UAV oblique photography, and a refined inspection path for the cables is subsequently planned. This is followed by the collection of cable image data. To address the challenge of small sample data, image augmentation and enhancement strategies are employed to expand the dataset. A deep learning-based intelligent detection framework for surface defects of stay cables is then developed using the Yolov8-seg model, allowing for quantitative assessment of the detected defects' physical dimensions. Finally, the proposed UAV inspection path and defect detection method are validated through practical engineering applications. The main contributions of this paper are as follows: (1) the development of an automated UAV inspection path planning method for stay cables based on reconstructed 3D models; and (2) the establishment of an intelligent detection framework for surface defects in stay cables, tailored for small sample datasets. The proposed approach not only mitigates safety risks and reduces manual detection costs but also significantly enhances the efficiency and robustness of defect detection.

The remainder of this paper is organized as follows: Section 2 provides an overview of the proposed methodology's framework. Section 3 presents the proposed UAV flight path planning methods for stay cables inspection, including the 3D reconstruction of the bridge model and two distinct path planning approaches, namely, feature-line based and feature-area based methods. Section 4 introduces the Yolov8-seg deep learning model employed for the detailed segmentation of surface defects on stay cables. Section 5 details the construction of the surface defect dataset using real-world cable images captured by the UAV, presents a comparative experiment of the different Yolov8-seg deep learning model, and outlines the process of quantifying the physical dimensions of the surface defects. In Section 6, the proposed methodology is systematically validated through field engineering practice. Finally, conclusions are drawn in Section 7.

2. Proposed framework

The proposed framework for the automatic detection of surface defects in stay cables, based on UAV flight path planning, is illustrated in Fig. 1. This framework encompasses three main steps: (1) UAV flight path planning for stay cables based on 3D reconstructed bridge model; (2) surface defect detection of stay cables through a deep learning model; and (3) practical validation of the methodology. The first step involves reconstructing a 3D model of the target bridge using oblique photography, followed by the optimization of the UAV inspection path and subsequent image acquisition of the stay cables. The second step includes expanding the small sample stay cables dataset via image augmentation and enhancement strategies, comparing and evaluating the dataset with various Yolov8-seg models, and performing surface defect segmentation and quantitative evaluation of physical dimensions. In the final step, surface defect inspection of the stay cables is conducted on an actual cable-stayed bridge to validate the feasibility and reliability of the proposed methodology in engineering practice. Further details are provided in the subsequent sections.

3. UAV optimal flight route planning for stay cables inspection

This section primarily proposes two optimal flight path planning approaches for UAV-based inspection of surface defects in stay cables. Firstly, a 3D model reconstruction method for cable-stayed bridges utilizing oblique photography technology is introduced. Next, based on the geometric relationship between the bridge cable and the imaging of the UAV camera, the calculation principle for UAV flight waypoint spacing is clarified. Finally, building upon the reconstructed 3D preliminary bridge model and waypoint spacing, two refined flight path planning methods are presented: the feature line-based route planning method and the feature area-based route planning method.

3.1. 3D bridge model reconstruction

To construct a preliminary 3D model of the bridge using oblique photogrammetry, it is essential to initially formulate a comprehensive photogrammetric plan. This includes determination of the critical parameters such as the geographical location of the bridge, the flight altitude and photography angle of the UAV, and the photo overlap ratio, etc. Furthermore, to ensure the reconstruction precision of the model, the oblique photography route must be carefully planned according to the geometric characteristics of the target bridge. For bridges with minimal undulations, such as tied-arch bridges, a single horizontal circular flight path is sufficient. In this scenario, the camera gimbal should be set to a 45° tilt angle, and the flight altitude should be approximately twice the top elevation of the arch bridge. In contrast, for bridges with significant undulations, such as cable-stayed or suspension bridges with prominent cable towers, a dual-level horizontal circular flight path is typically implemented. The first-level flight path is positioned on a horizontal plane 20–30 m above the top of the cable tower, with the gimbal tilt set to 45° to capture images that encompass the entire bridge. To guarantee the integrity of the cable tower's upper structure modeling, a second-level horizontal circular route is implemented on a plane approximately twice the height of the cable tower for supplementary imaging. In addition, oblique photogrammetry typically maintains a certain image overlap ratio to ensure that the captured images exhibit a degree of data redundancy, thereby facilitating comprehensive coverage of the target object [50]. As an illustrative example, the schematic of the oblique photography route for the cable-stayed bridge is presented in Fig. 2, where the blue-cube region represents the reconstruction area for the target bridge model.

The reconstruction of 3D point clouds from UAV images can be conceptualized as the inverse process of imaging, wherein the geometric information of 3D scenes is derived from two-dimensional (2D) image data. Based on the UAV oblique imagery, aerial photogrammetry software tools such as Inpho, Pix4Dmapper, ContextCapture, and DJI Terra can be employed to generate the 3D model of the bridge with realistic textures. This process mainly involves image preprocessing, multi-view image alignment, multi-view image dense matching, the generation of digital surface model (DSM), and texture mapping [54–57].

3.2. Calculation principle of waypoint spacing

Building upon the preliminary 3D reconstruction model of the bridge, refined flight routes for the UAV's oblique photography targeting the bridge cables will be meticulously planned. Initially, the principle for determining the waypoint spacing will be established.

Upon the selection of the UAV-mounted camera, the distance D between the UAV and the cable can be mathematically derived based on the principles of photogrammetric imaging, through the following relationship:

$$D = d \times \frac{GSD}{\mu} \tag{1}$$

where, d represents the focal length of the camera; GSD denotes the ground sampling distance, defined as the actual distance represented by a single pixel in the aerial imagery. The GSD is typically user-defined and tailored to the specific requirements of the aerial photography task; μ corresponds to the pixel size of the camera sensor.

The geometric relationship between the bridge cable and the imaging of the UAV camera is shown in Fig. 3, where the camera sensor size is b mm $\times$ h mm, and the corresponding image shooting range is B mm $\times$ H mm. To ensure a certain level of data redundancy, assuming

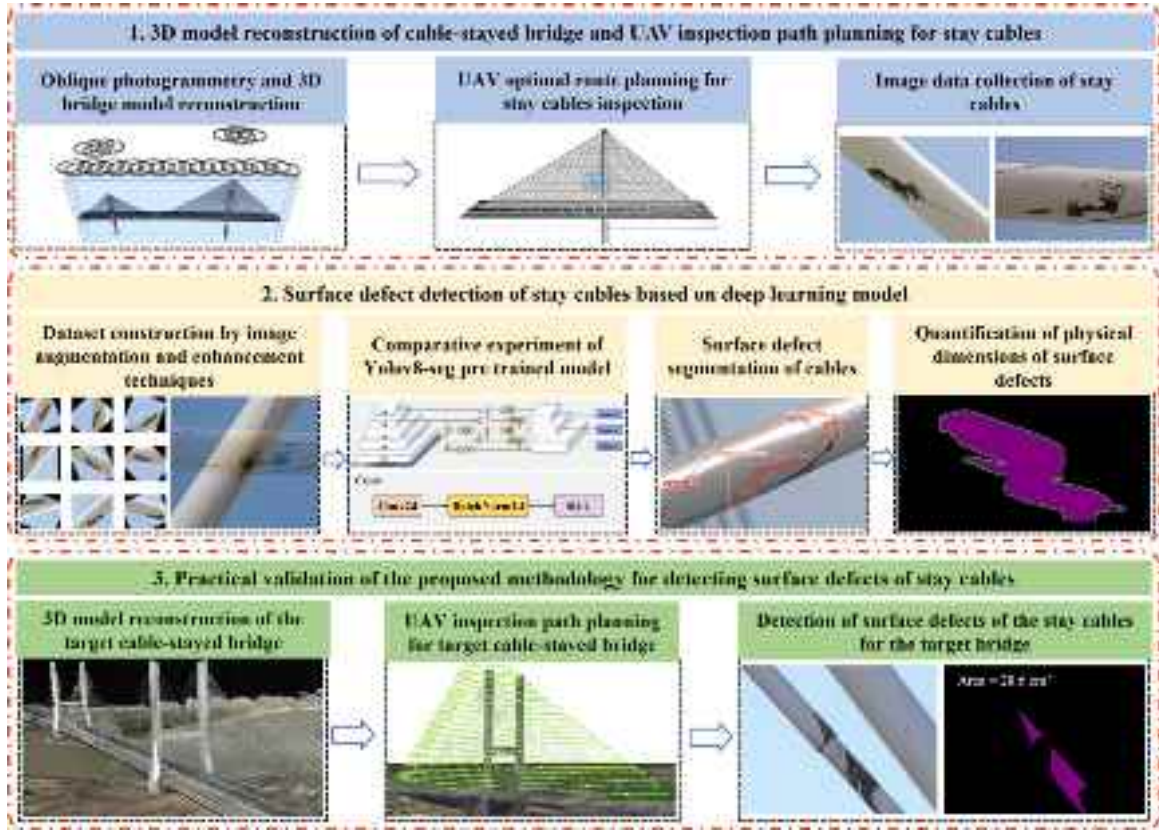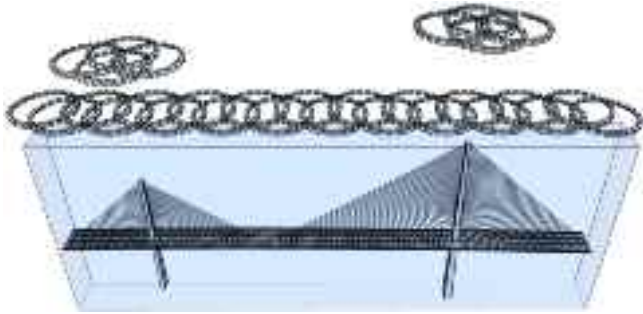**Fig. 1.** Overall framework of proposed methodology.

Fig. 2. Schematic of oblique photography route for cable-stayed bridge.

Fig. 3. Schematic diagram of camera imaging geometry.

the horizontal and vertical overlap rates are α and β, respectively, the horizontal waypoint distance D_{xy} and the vertical waypoint distance D_z can be derived as follows:

$$D_{xy} = (1 - \alpha)H = (1 - \alpha)\frac{Dh}{d} \tag{2a}$$

$$D_z = (1 - \beta)B = (1 - \beta)\frac{Db}{d} \tag{2b}$$

$$\frac{D_z}{D_{xy}} = \frac{\left|z_j - z_{j-1}\right|}{\sqrt{\left(x_j - x_{j-1}\right)^2 + \left(y_j - y_{j-1}\right)^2}} \tag{2c}$$

where, (x_j, y_j, z_j) and $(x_{j-1}, y_{j-1}, z_{j-1})$ are respectively the coordinates of adjacent waypoints j and $j-1$.

In the context of oblique photography, feature point matching between adjacent images constitutes a critical step in the modeling process. To ensure the accuracy of modeli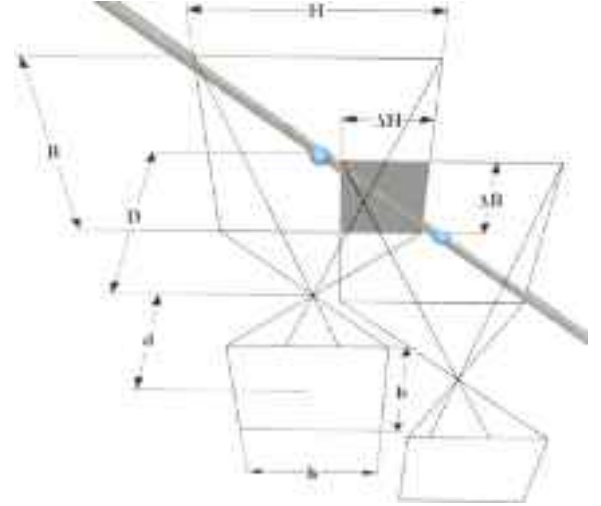ng, it is necessary to maintain a relatively large overlap rate between consecutive images to facilitate sufficient feature point correspondence. However, excessively high overlap rates generally lead to an increased density of waypoints, thereby compromising UAV operational efficiency. Therefore, in practical applications, the appropriate overlap rate should be carefully determined based on the specific requirements of the photography task.

3.3. UAV optimal flight route planning for stay cables inspection

To achieve rapid and efficient acquisition of surface defect images for stay cables, the implementation details of two refined UAV flight
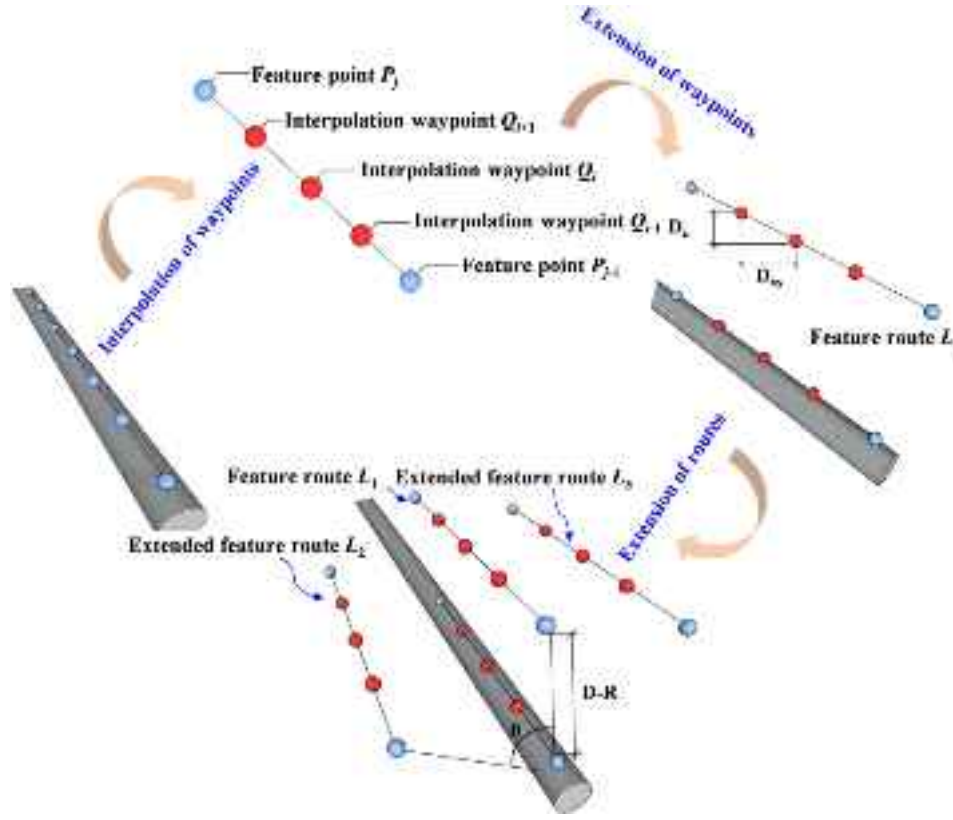
Fig. 4. Feature line-based optimized flight planning.

path planning approaches are proposed in this section, including feature line-based path planning and feature area-based path planning methods.

3.3.1. Feature line-based UAV flight route planning

The feature line-based optimized flight planning strategy aims to design the inspection route for a single cable. In this scenario, the pitch angle of the UAV camera is typically maintained at $0°$, and the object distance D between the UAV and the target cable is determined by Eq. (1), based on the focal length of the UAV camera, the required GSD, and the pixel size of the lens. Taking into account the radius R of the cable, the actual distance between the UAV and the surface of the cable is D-R. Furthermore, to guarantee comprehensive coverage of the cable, both the horizontal and vertical overlap rates between consecutive images must not fall below the predetermined minimum overlap threshold. If the minimum horizontal overlap rate is denoted as v, the corresponding horizontal and vertical waypoint distances can be calculated as:

$$D_{xy} = (1-v)H = (1-v)\frac{(D-R)h}{d} \tag{3a}$$

$$D_z = \frac{\left|z_j - z_{j-1}\right| D_{xy}}{\sqrt{\left(x_j - x_{j-1}\right)^2 + \left(y_j - y_{j-1}\right)^2}} \tag{3b}$$

If the vertical overlap rate $\beta = 1 - d D_z/[(D-R)b] \geq v$, there is no need to recalculate the waypoint distance; Otherwise, recalculate the waypoint distance as:

$$D_z = (1-v)B = (1-v)\frac{(D-R)b}{d} \tag{4a}$$

$$D_{xy} = \frac{\sqrt{\left(x_j - x_{j-1}\right)^2 + \left(y_j - y_{j-1}\right)^2}}{\left|z_j - z_{j-1}\right|} D_z \tag{4b}$$

The feature points on the cable and the radius of the cable are the feature elements for optimal path planning, and the feature points are marked as P_1, P_2,…, P_m. Following the determination of horizontal and vertical waypoint distances, several waypoints Q_1, Q_2,…, Q_n will be interpolated between feature points P_{j-1} and P_j, as shown in Fig. 4. For given interpolated waypoint Q_i (x_i, y_i, z_i) located between feature points P_{j-1} and P_j, the coordinates of interpolated waypoint Q_{i+1} can be derived as:

$$\begin{cases} Q_{i+1} = Q_i + \frac{\sqrt{D_{xy}^2 + D_z^2}}{\left\|\overrightarrow{P_{j-1}P_j}\right\|} \overrightarrow{P_{j-1}P_j}, if \left\|\overrightarrow{Q_iP_j}\right\| \geq \sqrt{D_{xy}^2 + D_z^2} \\ Q_{i+1} = P_j + \frac{\sqrt{D_{xy}^2 + D_z^2} - \overrightarrow{Q_iP_j}}{\left\|\overrightarrow{P_jP_{j+1}}\right\|} \overrightarrow{P_jP_{j+1}}, otherwise \end{cases} \tag{5}$$

Following the interpolation of waypoints, they will be further extended to establish the primary inspection route, designated as L_1. To ensure comprehensive coverage and maintain the integrity of cable imaging from multiple perspectives, as illustrated in Fig. 4, the baseline route L_1 will undergo extension to generate supplementary inspection paths, denoted as L_2 and L_3, etc.

3.3.2. Feature area-based UAV flight route planning

The principle of feature line-based route planning for bridge cable inspection can be effectively extended to feature area-based methodology. The fundamental distinction lies in the spatial coverage, where the feature area-based approach encompasses all cables through designated feature areas, rather than focusing on a single cable in the line-based method. This extension enables efficient route planning for simultaneous inspection of multiple cables. To illustrate this methodology, a single-pylon, single-plane cable-stayed bridge is considered as a representative case, in which the inspection route consists of two distinct components: straight trajectories parallel to the cable plane and the side circular trajectories surrounding the cable plane. As shown in Fig. 5, the implementation of straight route planning involves the following four key steps: (1) Step 1: Determination of feature area. Based on
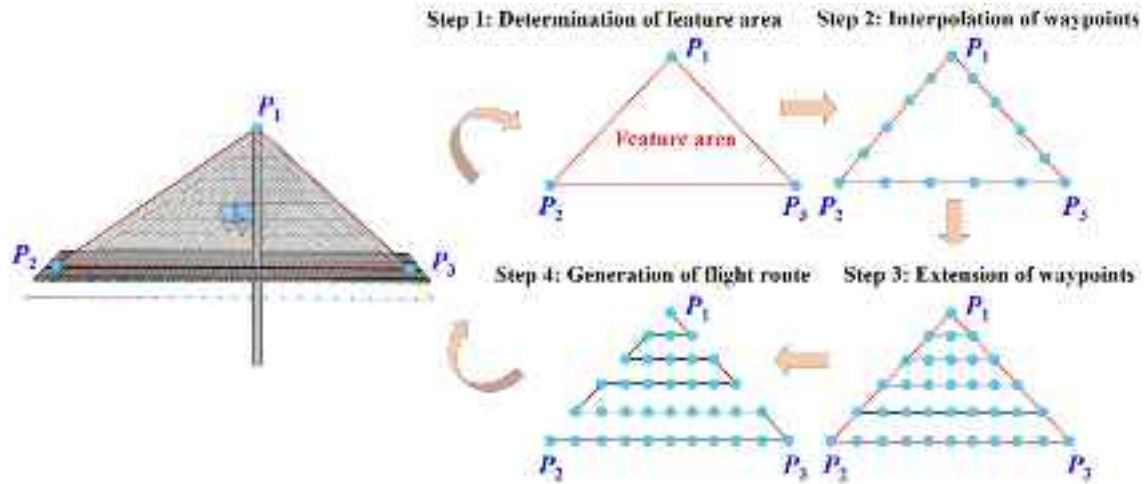
Fig. 5. Feature area-based planning process for straight route inspection of stay cables.

the oblique photography model, the positions of three vertices P_1, P_2, and P_3 on the cable surface are recorded and designated as feature points. The triangular area formed by these three points defines the feature area. (2) Step 2: Interpolation of waypoints. The coordinates of interpolated waypoints along the three edges of the feature area are determined by combining the line equations of P_1P_2, P_2P_3, and P_1P_3 with the required heading overlap rate. (3) Step 3: Extension of waypoints. The interpolated waypoints from each layer are connected horizontally. Then, the waypoints on the connecting lines of each layer are extended, and the position coordinates of these newly extended waypoints are simultaneously determined. (4) Step 4: Generation of flight route. The final flight route is established by defining the start and end positions, followed by sequential connection of all waypoints to form an optimized straight inspection route.

Given that a single-sided straight route is typically insufficient to capture a comprehensive image of cables, dual-sided straight routes are generated, complemented by circular routes connecting the two straight paths to ensure complete coverage of the cable imaging. As illustrated in Fig. 6, the circular waypoints are denoted in red, with the radius R of the circular route matching the flight distance D, where a represents the distance between adjacent circular waypoints.

To ensure the safety of UAV operations, it is crucial to verify the planned path of the oblique photography model upon completion of waypoint planning. In consideration of the traffic conditions on the bridge deck and the height of ancillary facilities such as streetlights, the UAV camera's pitch angle may be appropriately adjusted. Additionally, some waypoints may be omitted to further ensure flight safety. For cable-supported bridges with diverse structural configurations, such as twin-tower cable-stayed bridges, double-cable-plane cable-stayed bridges, the inspection path can be fully optimized by adapting the feature surface geometry and the radius of the circular route to suit specific structural requirements.

4. Deep learning model

To effectively capture the characteristics of surface defects and perform quantitative evaluations for stay cables, this paper adopts the image segmentation-based approach. This methodology allows for the precise measurement of defect size and detailed attributes. In this paper, the Yolov8-seg model will be utilized to achieve accurate and efficient segmentation of surface defects in stay cables.
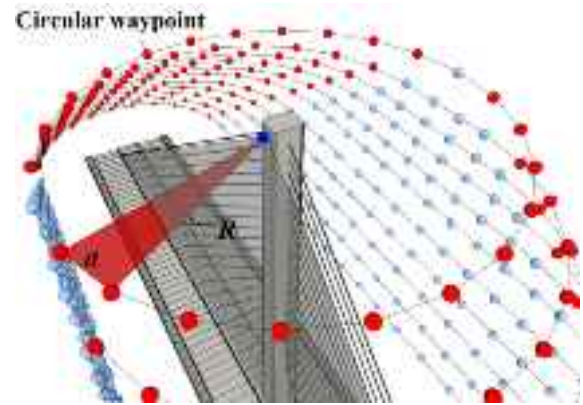

Fig. 6. Schematic diagram of circular waypoints.

4.1. Yolov8-seg model architecture

The Yolov8 algorithm, developed by Ultralytics, is part of the Yolo (you only look once) algorithm family. These are one-stage models known for their high speed and accuracy. Yolov8-seg, the instance segmentation variant of Yolov8, is subdivided into five models based on the number of parameters included: Yolov8n-seg, Yolov8s-seg, Yolov8m-seg, Yolov8l-seg, and Yolov8x-seg. As illustrated in Fig. 7(a), the overall architecture of the Yolov8-seg model comprises three components, i.e., backbone, neck, and head, which collectively enable robust and precise instance segmentation [58].

For the backbone and neck, Yolov8 incorporates several advanced features, such as cross-stage partial connections (CSP), the path aggregation network for feature pyramid networks (PAN-FPN) feature fusion, and spatial pyramid pooling-fast (SPPF) module. CSP is designed to reduce computational costs while simultaneously enhancing or preserving the model's learning capability. It achieves this by splitting the feature map of a layer into two separate parts, processing them independently via convolution, and then merging them afterward, which improves both efficiency and gradient flow during training. Additionally, the PAN-FPN feature fusion technique optimizes the integration of multi-scale features, enhancing the model's ability to detect objects at various scales with increased precision. The SPPF module enhances the capability of the model to handle input images of varying sizes by pooling features at different scales and concatenating them, resulting in a more robust feature representation [59].
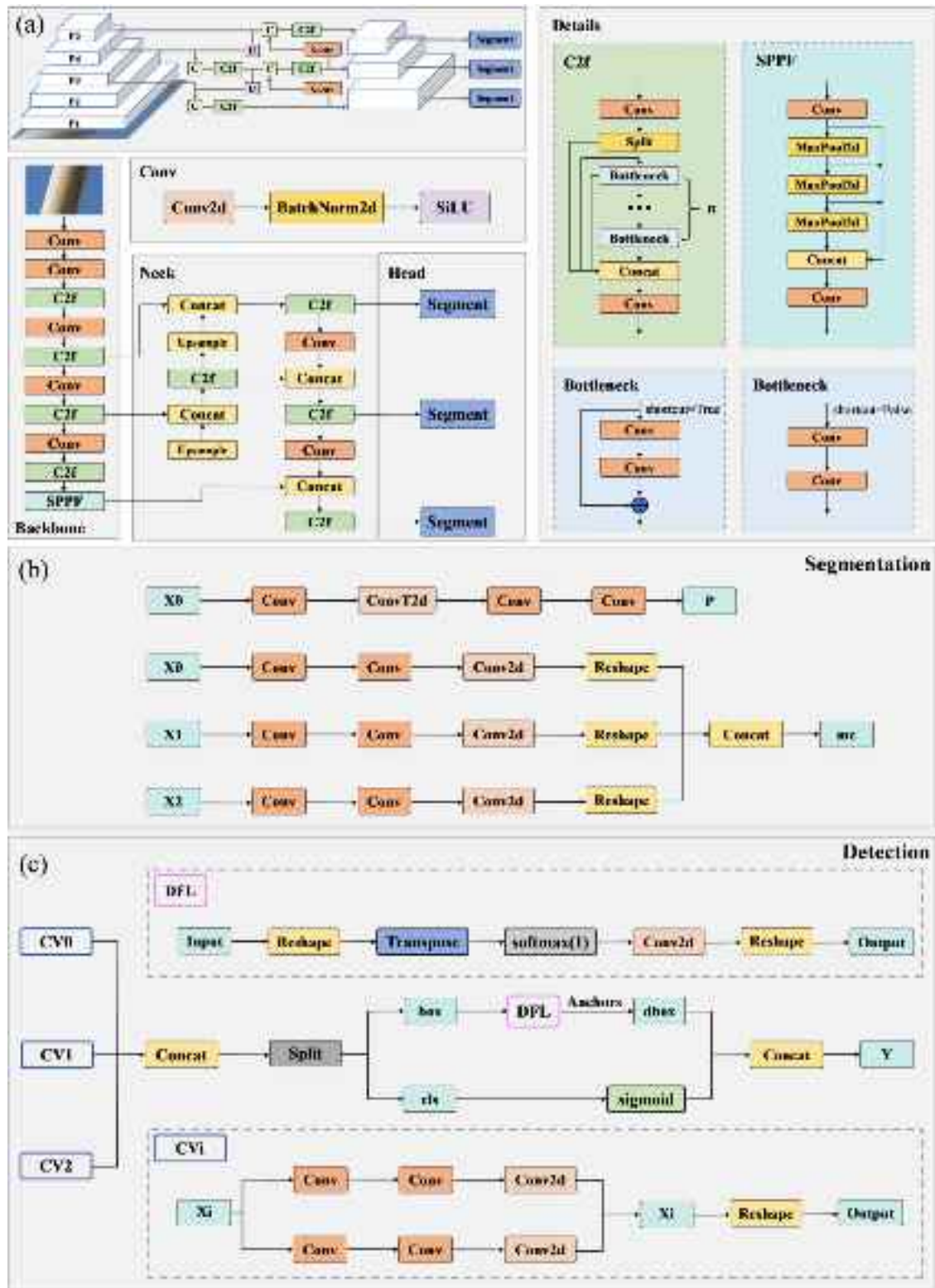
Fig. 7. Architecture of Yolov8-seg model: (a) Overall architecture; (b) Detection module in head structure; (c) Segmentation module in head structure.

Regarding the head structure, the Yolov8-seg model leverages concepts from YOLACT [60]. Unlike standard object detection models, the instance segmentation model includes two parallel modules in the head architecture: a detection module and a segmentation module, as respectively depicted in Fig. 7(b) and (c). In the detection module, Yolov8 adopts an anchor-free model [61] with a decoupled head, which handles object detection, classification, and regression tasks independently, thereby improving overall accuracy [62]. This approach eliminates the need for predefined anchor boxes by directly predicting object centers. The output layer uses the sigmoid function as the activation function for detection scores, which represent the probability of an object being present within the bounding box. In addition to the detection module, the segmentation module contains a prototype mask branch and mask coefficients, which are used to generate instance masks. These masks

offer pixel-level delineation of object boundaries, contributing to precise instance segmentation. More specifically, Yolov8-seg decomposes instance segmentation into two parallel processes: target detection and semantic segmentation. The target detection branch employs a series of convolutional layers to detect individual objects or instances within an image, enabling the model to achieve robust bounding box regression and classification. Meanwhile, the semantic segmentation branch works in parallel, understanding the semantic meaning of each pixel. By integrating these two branches within a unified architecture, Yolov8-seg achieves real-time, high-precision instance segmentation, offering valuable insights into individual objects present in images.

4.2. Loss function

The output of the segmentation head in Yolov8-seg model comprises the predicted category for each anchor box, category confidence, target box positioning, and prototype mask. Consequently, the loss function computation involves the following components: the classification loss L_{cls}, box regression loss L_{box}, distribution local loss (DFL) L_{dfl}, and segmentation loss L_{seg}. The overall loss function is formulated as a weighted sum of the aforementioned losses. Each loss term serves a distinct purpose, and the losses can generally be categorized into classification and regression losses.

The classification loss L_{cls} utilizes a variant of cross-entropy loss, specifically binary cross-entropy (BCE) loss, as defined as:

$$L_{cls} = -W \left[y_n \log \left(x_n \right) + \left(1 - y_n \right) \log \left(1 - x_n \right) \right] \tag{6}$$

where, W denotes the weight, x_n represents the predicted class value, and y_n is the true label. This loss quantifies the error in object classification, reflecting the discrepancy between the predicted and actual class labels.

The regression losses consist of three components, i.e., L_{box}, L_{dfl}, and L_{seg}. L_{box} is computed using complete intersection over union (CIoU) loss [63], which measures the disparity between the predicted bounding box and the actual ground truth bounding box. CIoU loss accounts for the overlap between the two boxes while also incorporating differences in aspect ratios. The expression of the CIoU loss function is formally defined as:

$$L_{box} = 1 - IOU + \frac{d^2 \left(B, B_{gt} \right)}{c^2} + \alpha v \tag{7}$$

where, B and B_{gt} denote the central points of the predicted and ground truth boxes, respectively. d represents the Euclidean distance, and c denotes the diagonal length of the smallest enclosing box that can contain both boxes, as shown in Fig. 8. The hyperparameter α adjusts the weight of the loss, while v measures the aspect ratio consistency, expressed as:

$$\alpha = \frac{v}{(1 - IOU) + v} \tag{8a}$$

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w_{gt}}{h_{gt}} - \arctan \frac{w}{h} \right)^2 \tag{8b}$$

where, w_{gt}, h_{gt}, w and h are the width and height of the target box and prediction box, respectively.

L_{dfl} aims to enhance feature learning by accounting for the relative spatial positions of features. This loss incorporates an orientation factor to better capture spatial distribution, thus preserving finer spatial details in the learned features [64]. The formula for L_{dfl} is presented as:

$$L_{dfl} \left(T_n, T_{n+1} \right) = - \left[\left(y_{n+1} - y \right) \log \left(T_n \right) + \left(y - y_n \right) \log \left(T_{n+1} \right) \right] \tag{9}$$

$$T_{n+1} = \frac{y - y_n}{y_{n+1} - y_n}, T_n = \frac{y_{n+1} - y}{y_{n+1} - y_n} \tag{10}$$

L_{seg} is dedicated to segmentation tasks and evaluates the discrepancy between the instance segmentation produced by the model and
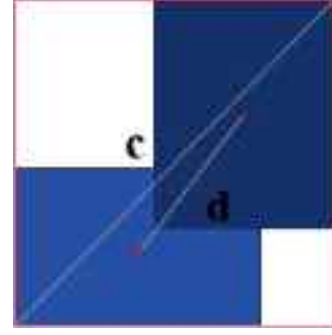

Fig. 8. Bounding box regression corresponding to CIOU loss.

the ground truth segmentation. Finally, the overall loss is expressed as a normalized total loss for the instance segmentation task, formulated as:

$$L = \alpha L_{cls} + \beta L_{box} + \gamma L_{dfl} + \delta L_{seg} \tag{11}$$

where, the parameters α, β, γ, and δ regulate the contribution of each individual loss term.

5. Surface defect detection of bridge cables

This section provides a comprehensive analysis of the detection results for surface defects in stay cables utilizing deep learning models. Firstly, a surface defect dataset for cables is constructed through image enhancement and augmentation techniques, with the corresponding evaluation metrics for segmentation performance clearly outlined. Building on this, the performance of five pre-trained Yolov8-seg models is compared and analyzed for defect segmentation tasks. Finally, the segmentation results of the optimal model are presented, followed by a quantitative analysis of the surface defects, converting the pixel-based measurements into physical dimensions.

5.1. Dataset construction

Typically, bridge cable images are obtained through routine bridge inspection reports or manual/UAV inspections. These images are then processed by filtering out those related to defects, followed by manual annotation to categorize and specify the locations of cable defects. This procedure facilitates the construction of a comprehensive defect dataset. However, the sample size of defect-related data obtained from standard bridge inspection reports is often limited, and the collection of large-scale cable defect data presents significant challenges. To address this, this paper focuses on using UAV to collect surface defect data from the cable-stayed region of a specific cable-stayed bridge, specifically the Tianhou Bridge located in Nanjing, Jiangsu Province, China. Ultimately, 475 high-resolution images (5184 × 3888 pixels) with distinct defect features were generated. Despite this effort, the limited dataset size may still hinder the generalization capacity of the model, potentially leading to overfitting and compromised recognition performance. Furthermore, given that defect features occupy only a small portion of the images, repeated feature extraction and downsampling during model training could lead to the loss of crucial small-target defect information, thereby degrading model accuracy. To mitigate these challenges, this paper employs image augmentation and enhancement techniques in the preprocessing phase. Initially, random cropping, flipping, and rotation are applied to augment the diversity of the defect images. Subsequently, advanced enhancement methods, including Mosaic and Mixup, are utilized to further enhance the dataset. Finally, each image is manually annotated, resulting in a defect-specific image dataset, with each image assigned a corresponding defect label.
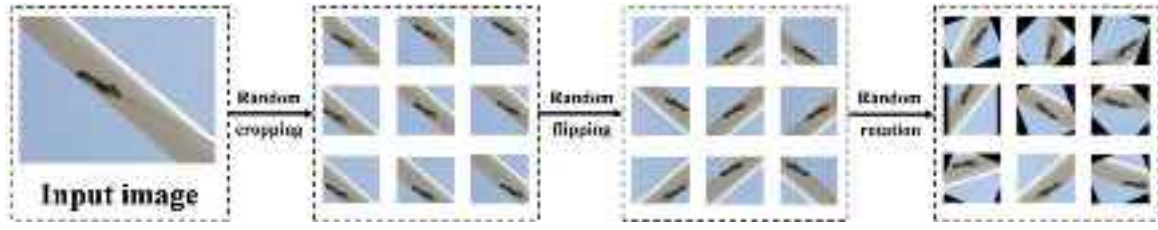
Fig. 9. Image augmentation processing effect.

Fig. 10. Image enhancement effect: (a) Mosaic method; (b) Mixup method.

Fig. 11. Schematic diagram of pixel level annotation for different types of defects: (a) Crazing; (b) Patches.

The image flipping algorithm facilitates dataset augmentation by flipping each pixel across a horizontal or vertical axis of symmetry. The image rotation algorithm, on the other hand, rotates the original image by a predetermined angle around its geometric center, producing images with defect-related features at various orientations, thus increasing data diversity. Given the high resolution of the collected defect images (5184 × 3888), directly utilizing such images for model training would require excessive computational resources. Therefore, an image cropping algorithm is applied to standardize the input size of the defect images. This process involves determining an appropriate cropping size based on a predefined factor and sequentially selecting regions from the original images that meet the size criteria, thereby generating a new set of sub-image blocks. The efficacy of these image augmentation techniques has been illustrated in Fig. 9.

For image enhancement, the Mosaic augmentation technique combines four randomly selected images by cropping and scaling them, and then stitches them together to form a composite image. This method significantly increases the diversity of the training samples and improves the model's robustness in handling complex scenes with overlapping targets. In contrast, the Mixup technique employs image aliasing: it randomly selects images from the dataset and generates new images through linear interpolation. Mixup effectively diversifies the training data while reducing model sensitivity to noise and overfitting, thereby enhancing its generalization ability and robustness. The efficacy of the image enhancement techniques is illustrated in Fig. 10.

Based on the aforementioned data preprocessing techniques applied, the original set of 475 cable defect images was augmented to 3800, thereby creating the final dataset of surface defect images. Prior to model training, manual annotation of the target objects in the images was performed. In this paper, LabelImg was utilized as the data annotation tool, with the surface defects categorized into two classes: Crazing and Patches. As illustrated in Fig. 11, a polygonal rectangular box was employed to annotate these defects at the pixel level, recording both the category and the location of cable defects in each image. Ultimately, a total of 525 samples of crazing defects and 569 samples of patches defects were obtained, resulting in a dataset with a one-to-one correspondence between images and labels. The dataset was then partitioned into a training set (3040 images), a validation set (380 images), and a testing set (380 images) in an 8:1:1 ratio. To ensure a balanced representation of defect images across the training, validation, and testing sets, a stratified sampling approach was implemented during the dataset construction.

5.2. Evaluation metrics

To quantitatively assess the accuracy and effectiveness of the model, four evaluation metrics are adopted, namely precision, recall, mean average precision, and the F_1 score, which are commonly used for segmentation algorithms [65].

Precision (P) evaluates the proportion of true positive predictions relative to all positive predictions made by the model. Recall (R) quantifies the ratio of true positive predictions to the actual number of positive instances in the dataset. AP represents the average segmentation precision of a given category, derived from the Precision–Recall (P-R) curve, where the area under the curve reflects the performance of the model for a specific defect category. The larger the area, the higher AP value, indicating a higher identification accuracy for that category. Mean average precision (mAP) refers to the mean average precision calculated at a threshold of 0.5 ($mAP@0.5$) and across a range from 0.5 to 0.95 ($mAP@0.5$-0.95). This metric evaluates the proficiency of the model in generating precise object segments in its predictions. The F_1 score is another frequently used metric to assess the overall performance of object segmentation models. It harmonizes precision and recall into a single metric, offering an integrated measure of the model's accuracy in identifying true positives while considering both FP and FN errors. The above metrics are mathematically defined as follows:

$$P = \frac{TP}{TP + FP} \tag{12a}$$

$$R = \frac{TP}{TP + FN} \tag{12b}$$

$$AP = \int_0^1 P(R)\mathrm{d}R \tag{12c}$$

$$mAP = \frac{\sum_{i=1}^{n} AP(i)}{n} \tag{12d}$$

$$F_1 = 2 \times \frac{P \cdot R}{P + R} \tag{12e}$$

where, TP denotes true positive samples (correctly identified defects), FP indicates false positive samples (incorrectly predicted as positives), FN refers to false negative samples (positives erroneously predicted as negatives), and n is the total number of categories in the segmentation task.

Table 1
Performance comparison of different Yolov8-seg models.

No.	Model name	P	R	mAP	Memory usage/MB	Test time/ms
1.	Yolov8n-seg	0.724	0.612	0.635	6.4	3.4
2.	Yolov8s-seg	0.764	0.654	0.656	22.7	5.1
3.	Yolov8m-seg	0.790	0.670	0.690	52.2	8.2
4.	Yolov8l-seg	0.701	0.604	0.580	88.0	13.4
5.	Yolov8x-seg	0.570	0.550	0.510	137.0	21.5

5.3. Model comparison experiments

As previously stated, Yolov8-seg offers five pre-trained models with varying parameter sizes: Yolov8n-seg, Yolov8s-seg, Yolov8m-seg, Yolov8l-seg, and Yolov8x-seg. This section presents a comparative analysis of these five models, evaluated using the constructed cable surface defect dataset. Each model was trained and subsequently tested to identify the most suitable model for the dataset at hand. The experiments were conducted on an Ubuntu 20.04 operating system, utilizing Python 3.8, CUDA 11.8, and PyTorch 2.0. The hardware configuration included an RTX 3090 graphics card and an 8-core Intel® Xeon® Gold 6133 CPU @ 2.50 GHz. The experimental parameters were set as follows: an initial learning rate of 0.02, a batch size of 8, an IoU threshold of 0.7, 100 epochs, and an image resolution of 640 × 640.

The results of the comparative experiments are presented in Table 1. It is evident that the Yolov8m-seg model outperforms the other four models in terms of segmentation accuracy, with superior precision, recall, and mAP values. This suggests that, under the current dataset and consistent training parameters, the Yolov8m-seg model, with its balanced parameter size, achieves the highest identification accuracy while maintaining faster detection speeds. Consequently, the Yolov8m-seg instance segmentation model will be adopted for detecting surface defects in stay cables in the subsequent sections.

5.4. Surface defect segmentation results

In this section, sufficient training was conducted on the Yolov8m-seg model over 500 epochs. Training was terminated early if no substantial improvement in performance metrics was observed within 100 consecutive iterations. The evaluation metrics precision P, recall R, and mAP during the training process are shown in Fig. 12, the F_1 score curve is depicted in Fig. 13(a), and the P-R curve for different defects is presented in Fig. 13(b). From Figs. 12 and 13(a), continuous improvements in precision, recall, and mAP on the validation set throughout training can be observed. After 328 epochs, the mAP reaches 0.713 and gradually stabilizes, corresponding to a precision of 0.825 and a recall of 0.674. The model achieves the highest F_1 score of 0.70 for different defects at a confidence threshold of 0.530. As depicted in Fig. 13(b), the AP value of the model for crazing-type defects is 0.684, whereas for patches-type defects, the AP value is 0.743, indicating a higher identification accuracy for patches-type defects. This discrepancy is primarily due to the smaller pixel area occupied by crazing-type defects in the original high-resolution images, making small target recognition more challenging. Additionally, crazing-type defects often suffer from incomplete labeling, further complicating detection. In contrast, the more distinct boundary features of patches-type defects facilitate higher identification accuracy.

Finally, the trained Yolov8m-seg model was evaluated on the test set, and the defect segmentation results are presented in Fig. 14. The normalized confusion matrix for the test results is shown in Fig. 15(a), where each column represents the true category of defects, and each row represents the predicted category. As presented in Fig. 15(b), taking the second column of the confusion matrix as an example, where the true category is patches-type defects, the predicted results show that 5% of instances were misclassified as crazing-type defects (false positives, FP), 75% were correctly predicted as patches-type defects (true positives, TP), and 20% were missed and classified as background (false negatives, FN). Overall, the above results demonstrate that the Yolov8m-seg model performs well in instance segmentation of visible defects.

5.5. Quantification of physical dimensions for surface defects

Quantifying the physical dimensions of surface defects detected in cables is crucial for maintenance decision-making of bridges. To this end, this section first extracts the pixel sizes of detected defect images, and then establishes a mapping relationship from the pixel size to the corresponding physical geometric dimensions, thereby enabling the quantification of defect sizes.

5.5.1. Extraction of pixel information of surface defects

For crazing-type defects, this paper quantifies them by extracting their pixel length. The pixel length is typically derived from the centerline of the image skeleton. Initially, the defect mask image is binarized, followed by the application of the refinement algorithm proposed by Zhang et al. [66] to extract the skeleton's centerline. The pixel length of the crazing is determined by counting the total number of skeleton pixels. In this process, the refinement algorithm assigns a pixel value of 255 to the crazing skeleton and 0 to the background. By normalizing and summing the pixel values of the image, the total pixel length of the crazing is obtained. The calculation formula is as:

$$L_{pixel} = \sum_{x=0}^{H}\sum_{y=0}^{W} \frac{I_{skeleton}(x,y)}{255} \tag{13}$$

where, L_{pixel} represents the pixel length of the crazing defect, while H and W denote the height and width of the image, respectively. $I_{skeleton}(x,y)$ denotes the pixel value of the skeleton centerline at coordinates (x, y).

For patches-type defects, the pixel area is employed to quantify them. The contour of the defect region is obtained using the contour detection method in the OpenCV API from the segmented mask image. As shown in Fig. 16, grayscale processing and pixel counting are employed, traversing the coordinates (x, y) of each pixel to obtain the corresponding image value $f(x, y)$. If the pixel belongs to a defect region, its area is incremented by 1 to calculate the total pixel area of the defect region. The calculation formula is expressed as:

$$A_{pixel} = \sum_{x=0}^{H}\sum_{y=0}^{W} [f(x,y)=1] \tag{14}$$

where, A_{pixel} represents the pixel area of the patches-type defect, with $f(x, y)$ is the pixel value of the image at coordinates (x, y).

5.5.2. Mapping of pixel size to physical dimensions

To accurately determine the physical size of the defect, it is essential to establish a mapping relationship between pixel coordinates and physical coordinates. In practical UAV-based defect detection, the physical size Aphy of the defect area is typically estimated based on imaging principles, as formulated:

$$A_{phy} = A_{pixel} \times \frac{bhD^2}{BHd^2} \tag{15}$$

where, the camera sensor size is $b \times h$, the resolution of the UAV image is $B \times H$, d represents the camera focal length, and D denotes the distance from the camera to the cable surface. However, this assumes that the distance D is known. Herein, an adaptive mapping method for converting pixel size to physical size of defects is introduced, eliminating the need for prior knowledge of D. Initially, edge detection is applied to the image using the Canny algorithm. Then, the Hough
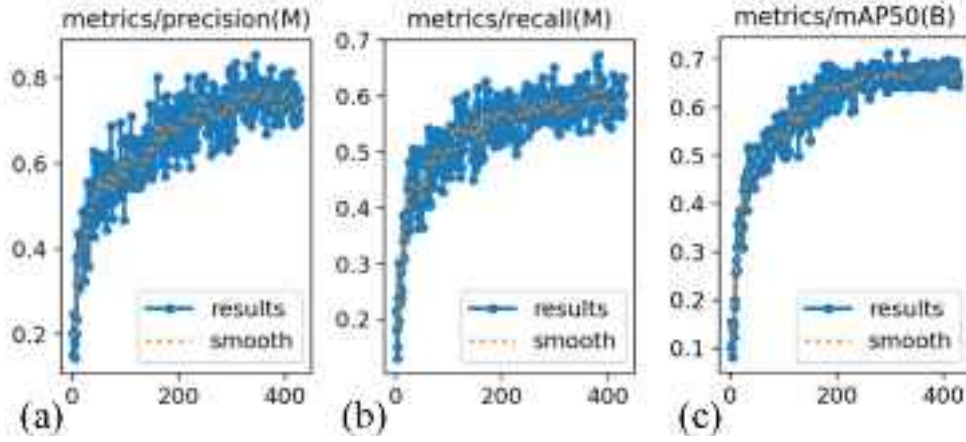
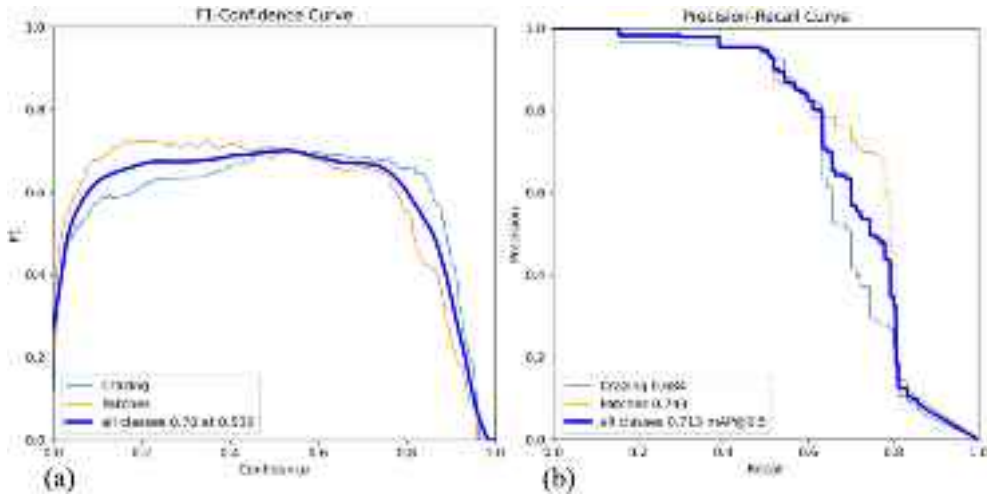**Fig. 12.** (a) Precision P, (b) recall R, and (c) mAP curves for Yolov8m-seg model training.

Fig. 13. (a) F1 score and (b) P-R curve for Yolov8m-seg model training.

Fig. 14. Segmentation results of surface defect of stay cables on the test set.

line detection method is employed to fit the line formed by the edges, allowing for the calculation of the average pixel size l of the cable diameter, as illustrated in Fig. 17. By combining the known design size L with the average pixel size l of the cable diameter, the scaling factor λ for mapping pixel size to physical size can be computed. Finally, the physical dimensions of the defect are derived as:

$$A_{phy} = A_{pixel} \times \lambda = A_{pixel} \times \frac{L}{l} \tag{16}$$

It is crucial to note that the precise determination of the geometric dimensions of surface defects on cables is not the primary focus of
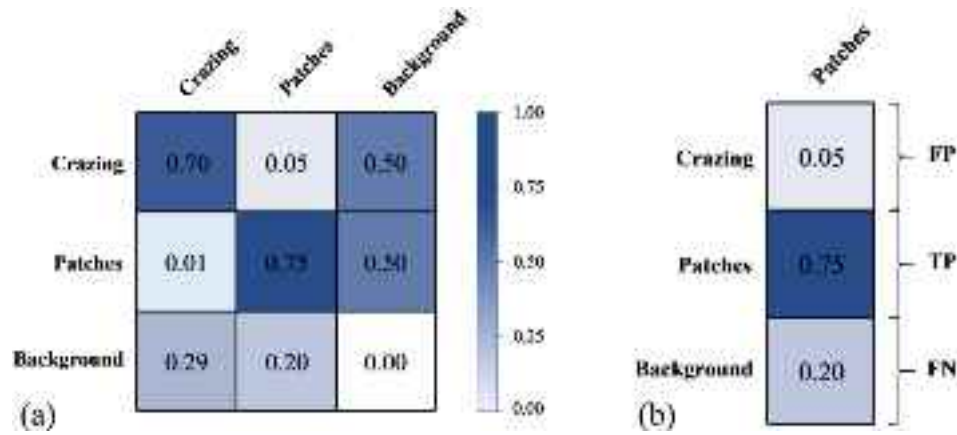
Fig. 15. Confusion matrix for instance segmentation of different defects on the test set: (a) Confusion matrix; (b) Patches type defect detection results.
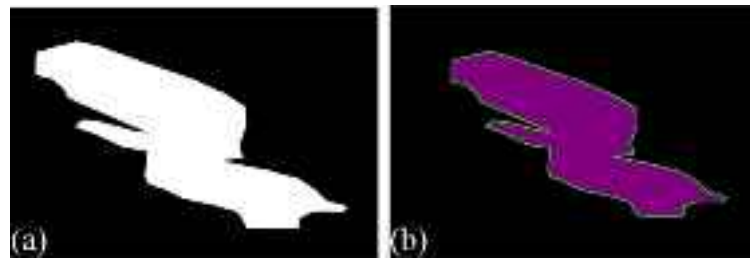
Fig. 16. Extraction of pixel area of patches-type defects: (a) Patches-type defect mask; (b) Contour area extraction.
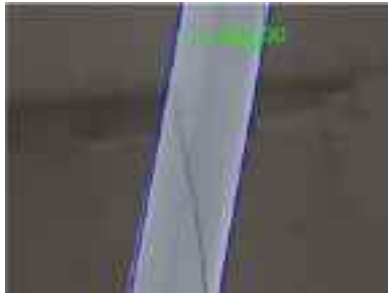
Fig. 17. Extraction of average pixel diameter of cables.

this study. However, given current engineering practices, the proposed method for estimating the physical size of surface defects can offer valuable insights for maintenance decision-making. Consequently, the potential distortion of defect images due to the curvature of cylindrical cables has not been addressed in this work. For a more accurate quantification of the physical dimensions, further reference can be made to [67].

6. Field validation

The effectiveness of the proposed methodology for the practical detection of surface defects in stay cables will be systematically validated in this section. First, relevant information about the target bridge and UAV equipment is provided, followed by the details of the oblique photography used during testing and the resulting 3D reconstruction of the bridge. Subsequently, a comparative analysis is conducted to evaluate the effectiveness of UAV flight route generation based on two path planning methods. Building on this, defect segmentation results, derived from UAV-collected image data using the deep learning model, are presented and physically quantified. Finally, the limitations of

this paper are discussed, and some directions for future research are outlined.

6.1. Test information

In this section, the Huai'an Bridge is selected for the on-site validation of the proposed methodology. Located on the Huai'an-Suqian Expressway in Jiangsu Province, China, the Huai'an Bridge spans the Huaihe River and is a super-large bridge. As illustrated in Fig. 18(a) and (b), the main bridge is a double-tower, double-cable-plane prestressed concrete cable-stayed structure, with a total length of 674 m and a span configuration of (152+370+152) m. The bridge towers feature an H-shaped design, with 31 pairs of cables on each tower. The cables on the left tower are labeled as HL01 (for the Huai'an side) and SL01 (for the Suqian side), while those on the right tower are designated HR01 and SR01, respectively.

For image acquisition, the DJI M300-RTK UAV is employed, equipped with cameras Zenmuse P1 and Zenmuse H20T. The M300-RTK utilizes a real-time kinematic (RTK) technique to achieve positioning accuracy of 1 cm horizontally and 1.5 cm vertically. The Zenmuse P1 camera facilitates continuous aerial imaging across five directions through pan–tilt rotation, significantly enhancing the efficiency of oblique photography. The Zenmuse H20T camera comprises a 20-megapixel zoom camera, a 12-megapixel wide-angle camera, a thermal imaging camera, and a laser rangefinder, enabling the simultaneous capture of visible light, wide-angle, and thermal images. Consequently, the P1 camera is used for efficient modeling during oblique photography, while the H20T camera is employed in refined routes of UAV to ensure the image quality of the bridge cables. The test equipment is shown in Fig. 18(c).

6.2. 3D reconstruction result

For the rough 3D model reconstruction of bridges, UAV oblique photography circular flight routes typically offer a broad modeling
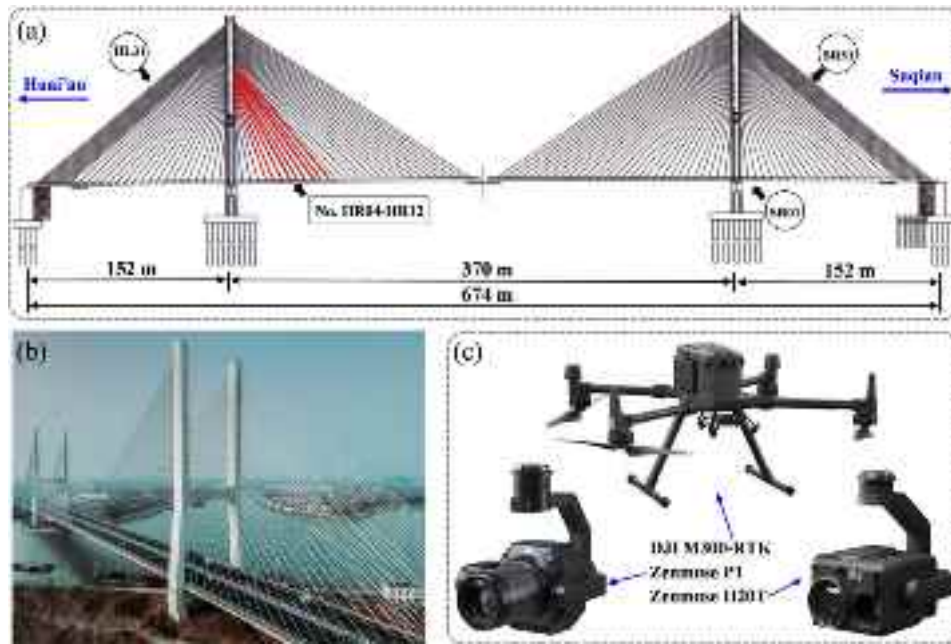
Fig. 18. Field test information: (a) Schematic diagram of facade layout of tested bridge; (b) Real-life picture of tested bridge; (c) UAV and cameras.

Fig. 19. Reconstructed rough 3D model of target bridge by UAV oblique photography.

coverage and rapid data acquisition, enabling efficient construction with a small amount of image data. Therefore, based on the M300-RTK UAV equipped with a Zenmuse P1 camera, efficient circular route is conducted to capture bridge images and establish a rough 3D model of the target bridge. During the test, the flight altitude of the UAV was set to the height relative to the takeoff point, with the circular flight path altitude of 120 m, the tower top supplementary path altitude of 140 m, an overlap rate of 80%, a GSD of 4.2 cm/pixel, and a total flight duration of 24 min and 45 s. The flight control system received waypoint positions and execution instructions in KML file format, guiding the UAV to each waypoint to capture aerial images of the bridge. In total, 359 oblique images were collected in this process. To enhance modeling accuracy, images of poor quality were excluded. Based on these image data, the final 3D reconstruction model of the target bridge is generated, as depicted in Fig. 19. As can be seen, the model can accurately reflect the 3D shape and textural features of the bridge, providing a foundation for the optimal route planning of the UAV inspection for bridge cables.

6.3. Flight route planning results for UAV inspection

Based on the reconstructed rough 3D bridge model, the results of two UAV flight path planning methods are presented. Meanwhile, a comparative analysis is conducted on the two methods, highlighting their characteristics, applicability, and detection efficiency.

6.3.1. Feature line-based flight route planning result

The UAV flight route is comprised of a series of high-precision GNSS-based point coordinates. By using the reconstructed 3D bridge model, the flight path of the UAV can be planned, and further multi-view images can be acquired. Due to the large number of cables, HR12 from the cables marked in red in Fig. 18(a), is selected as a representative for path planning herein. The M300 RTK UAV, equipped with a Zenmuse H20T camera, captures cable images at a target resolution of 2 mm/pixel. The UAV gimbal pitch angle is set to 0°, with a 30° angle between adjacent flight routes. The lens zoom is set to 5x, and the image overlap rate is 20%. The UAV operates at a distance of 15 m from the cable, with 40 waypoints per route, resulting in a total of 120 images per cable. Following the feature-based flight route planning process, the planned UAV flight route is finally presented in Fig. 20.

6.3.2. Feature area-based flight route planning result

This section takes the southeast cable region of the target bridge as an example to plan the cable inspection UAV path based on feature area. Similarly, the Zenmuse H20T camera was used to capture cable images, with a target resolution of 2 mm/pixel, an overlap rate of 50%, a 5x lens zoom, and a UAV distance of 20 m from the cable. Based on these parameters, a total of 27 straight flight paths are calculated, with horizontal and vertical waypoint spacings of 3.89 m and 5.18 m, respectively. Due to some flight restrictions imposed by bridge maintenance agency, the UAV not being allowed to invade the internal space of the bridge, only partial circular waypoints were generated. The final planned UAV cable inspection path based on feature area is shown in Fig. 21.

6.3.3. Comparative analysis for two route planning approaches

In this test, the feature line-based route planning method is employed to capture 1200 images of cables HR04-12, while the feature area-based method captures 1365 images for the southeast cable region.
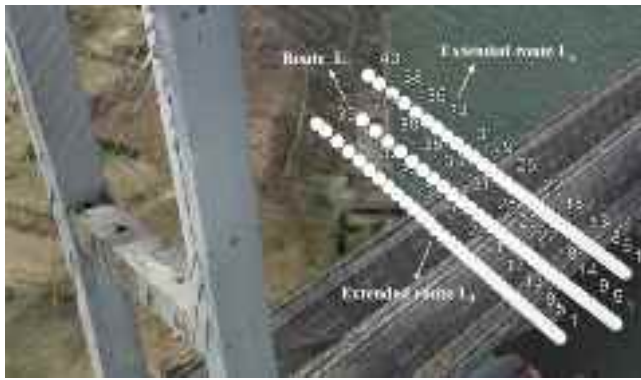
Fig. 20. Planned UAV route for cable inspection based on feature line method.

Fig. 21. Planned UAV route for cable inspection based on feature area method.

A comparative analysis of the two methods is provided in Table 2. It can be seen that the feature line-based method ensures that each cable is centered within the frame, and each image corresponds precisely to a specific cable. In contrast, the feature area-based method often captures multiple cables within a single frame, complicating the identification of individual cable numbers and necessitating further differentiation. Furthermore, it was observed that the feature area-based path planning approach was capable of inspecting 62 cables within 1 h and 30 min, whereas the feature line-based path planning method involved planning three distinct paths for each cable to cover the surface from various angles. As a result, defect images for only nine cables were meticulously captured within approximately 1 h and 20 min. Consequently, the feature line-based method is associated with relatively poor operational coherence and longer image acquisition time, whereas the feature area-based method facilitates continuous operation, reduces image acquisition flight time, and improves the inspection efficiency of multiple cables. In this regard, the feature area-based approach is more suitable for rapid inspections of the entire cable surface, while the feature line-based method is ideal for more detailed inspections of individual cables. Overall, both approaches effectively detect visible cable defects, and the selection of the appropriate inspection method should be based on specific practical requirements, it is advisable for operators to plan multiple routes for each cable to enhance the integrity of data collection from various angles.

6.4. Result analysis of surface defect detection

The collected cable images by UAV were analyzed for defect detection using the trained Yolov8-seg model. Due to the good maintenance of the bridge, only a few cables exhibited localized patches-type defects. Representatively, some defect detection results are displayed in Table 3. These results demonstrate that the trained Yolov8-seg effectively

detects surface defects in cables, thereby validating the effectiveness of the proposed methodology.

6.5. Limitations and future work

This paper provides a preliminary validation of the proposed framework through an on-site inspection of surface defects on cables of cable-stayed bridges. However, there are still two critical limitations that warrant clarification and further exploration in future research: (1) For long-span bridges, particularly those with spans exceeding 1000 m, there are some limitations in the flight path control of a single UAV due to environmental factors and current UAV technology. Specifically, large-span bridges, often located in windy areas such as those crossing rivers or valleys, experience fluctuations in wind speed and airflow, as well as extreme weather conditions like strong winds, all of which can interfere with UAV flight stability. These disruptions may result in deviations from the flight trajectory and reduced stability in image acquisition, which in turn compromises the quality of surface defect imaging. Additionally, during long-range flights, the accumulation of small positional deviations can degrade the accuracy of path control, increasing the risk of missed or redundant images of the target cable. Furthermore, the limited endurance of a single UAV restricts its coverage range, which in turn lowers detection efficiency. (2) The two path planning approaches proposed in this paper do not fully consider the issue that all local areas around the cylindrical cable's surface could fall within the central region of the camera's field of view, thereby ensuring clear imaging with high resolution and minimal distortion. The feature area-based path planning method captures overall images from both sides of the plane formed by the cable, while the feature line-based path planning method extends the flight path from multiple perspectives to enhance the integrity of the cable defect image acquisition as much as possible. However, it must be acknowledged that the current methods do not yet fully resolve the challenge of achieving optimal imaging coverage for cylindrical cables at all angles and without blind spots.

To overcome the aforementioned limitations and challenges, future research will integrate adaptive algorithms to develop a more robust real-time correction mechanism for UAV flight paths to enhance the path control precision. Additionally, the exploration of multi-UAV collaborative operations will be undertaken to enhance path control precision, reduce the burden of single UAV route control, and improve detection efficiency. Furthermore, efforts will be directed towards refining flight path planning to ensure coverage of the cylindrical stay cables from as many angles as possible, thereby enabling comprehensive imaging for surface defect detection.

7. Conclusions

This paper employed UAV oblique photogrammetry to rapidly construct a preliminary 3D model of a cable-stayed bridge. Building on this model, two refined automatic inspection path-planning methods for stay cables were proposed. A surface defect detection method for stay cables, utilizing the Yolov8-seg model, was developed for small-sample datasets. The proposed methodology was applied to inspect surface defects on the cables of a real-world cable-stayed bridge. Based on this methodology, the following conclusions were drawn:

(1) Two path optimization methods for UAV-based automatic inspection of stay cables, utilizing feature lines and feature surfaces, were developed based on the 3D model of the bridge. The effectiveness of the proposed path planning approach was validated through a field inspection test on the stay cables.
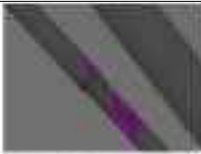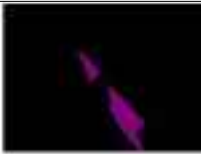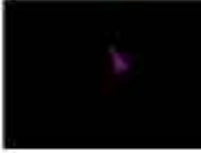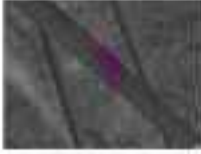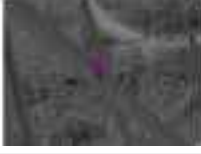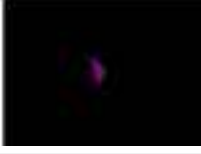
(2) The small sample dataset of stay cables was augmented using image augmentation and enhancement techniques. Comparative experiments were conducted on the augmented dataset using five variations of the Yolov8-seg model with different parameter sizes. Results indicated that the Yolov8m-seg model achieved the highest defect

Table 2
Comparison of route planning methods for cable inspection of target bridges.

Item	Feature line-based route planning	Feature area-based route planning
Inspection range	HR04-12 (9 cables)	Southeast cable region (62 cables)
Image example		
Overlap rate	20%	50%
Number of waypoints	1200	1365
Time for flight	1 h 20 min	1 h 30 min
Time for planning	30 min	10 min

Table 3
Visualization of partial results of cable defect detection for target bridge.

No.	Original image	Detection result	Defect segmentation	Defect area
1.				28.6 cm^2
2.				5.4 cm^2
3.				15.3 cm^2
4.				6.7 cm^2

segmentation accuracy, with a mAP value of 0.713 and segmentation precision of 0.825.

(3) A pixel-to-physical size mapping method for cable stay defects, based on edge detection techniques, was proposed. The physical measurements of patch-type and crazing-type defects segmented by the Yolov8 model were performed, enabling quantitative evaluation of surface defects.

(4) Two route planning methods and a pre-trained Yolov8m-seg model were applied to conduct engineering inspections of surface defects on a cable-stayed bridge's cables. The results demonstrated that the proposed method efficiently and accurately detects surface defects on the cables.

CRediT authorship contribution statement

Baoquan Wang: Writing – original draft, Validation, Software, Methodology, Investigation, Data curation. **Dongming Feng:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition, Conceptualization. **Zhichao Fan:** Validation, Software, Methodology, Investigation, Formal analysis, Data curation. **Hanyu Shi:** Visualization, Validation, Investigation.

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.

Acknowledgments

Data availability

Data will be made available on request.

References

[1] N. Gimsing, C. Georgakis, Cable Supported Bridges: Concept and Design, third ed., John Wiley & Sons., ISBN: 9781119978237, 2012, http://dx.doi.org/10.1002/9781119978237.

[2] W. Wickramasinghe, D. Thambiratnam, T. Chan, Damage detection in a suspension bridge using modal flexibility method, Eng. Fail. Anal. 107 (2020) 104194, http://dx.doi.org/10.1016/j.engfailanal.2019.104194.

[3] M. Zhang, S. Huang, P. Li, K. Shah, X. Zhang, Application of dehumidification as anti-corrosion technology on suspension bridges: A review, Appl. Therm. Eng. 199 (2021) 117549, http://dx.doi.org/10.1016/j.applthermaleng.2021.117549.

[4] Y. Ding, X. Ye, Y. Su, X. Zheng, A framework of cable wire failure mode deduction based on Bayesian network, Struct. 57 (2023) 104996, http://dx.doi.org/10.1016/j.istruc.2023.104996.

[5] S.L. Chen, D. Chen, R.D.A. Sannasiraj, L.H. Zhang, Engineering reliability-based condition assessment for stay cables using non-destructive interferometric radar, Int. J. Struct. Stab. Dyn. 24 (14) (2024) 2450154, http://dx.doi.org/10.1142/S0219455424501542.

[6] M. Elliott, E. Heymsfield, Inspection of Luling bridge cable stays: Case study, J. Constr. Eng. Manag. 129 (2) (2003) 226–230, http://dx.doi.org/10.1061/(ASCE)0733-9364(2003)129:2(226).

[7] S. Hou, B. Dong, H. Wang, G. Wu, Inspection of surface defects on stay cables using a robot and transfer learning, Autom. Constr. 119 (2020) 103382, http://dx.doi.org/10.1016/j.autcon.2020.103382.

[8] X. Long, X. Gui, W. Zhan, Surface defect detection method of high speed railway bridge cables based on optical fiber sensing technology, in: 2020 IEEE International Conference on Industrial Application of Artificial Intelligence, IAAI, Harbin, China, 2020, pp. 549–554, http://dx.doi.org/10.1109/IAAI51705.2020.9332902.

[9] Ministry of Transport of the People's Republic of China, Technical Specifications for Maintenance of Highway Cable-Supported Bridge: JTG/T 5122-2021, People's Communications Publishing House Co. Ltd., Beijing, China, 2021, https://xxgk.mot.gov.cn/2020/jigou/glj/202112/t20211209_3630469.html (Accessed 09 December 2021).

[10] H. Kajiwara, N. Hanajima, K. Kurashige, Y. Fujihira, Development of hanger-rope inspection robot for suspension bridges, J. Robot. Mechatron. 31 (6) (2019) 855–862, http://dx.doi.org/10.20965/jrm.2019.p0855.

[11] H. Ho, K. Kim, Y. Park, J. Lee, An efficient image-based damage detection for cable surface in cable-stayed bridges, NDT E Int. 58 (2013) 18–23, http://dx.doi.org/10.1016/j.ndteint.2013.04.006.

[12] K. Cho, Y. Jin, H. Kim, H. Moon, J. Koo, H. Choi, Multifunctional robotic crawler for inspection of suspension bridge hanger cables: Mechanism design and performance validation, IEEE/ASME Trans. Mechatron. 22 (1) (2017) 236–246, http://dx.doi.org/10.1109/TMECH.2016.2614578.

[13] X. Li, C. Gao, Y. Guo, F. He, Y. Shao, Cable surface damage detection in cable-stayed bridges using optical techniques and image mosaicking, Opt. Laser Technol. 110 (2019) 36–43, http://dx.doi.org/10.1016/j.optlastec.2018.07.012.

[14] K. Nguyen, N. Van Bach Pham, T. Dao, Damage detection of cables in cable-stayed bridges using vibration data measured from climbing robot, Adv. Struct. Eng. 25 (13) (2022) 2705–2721, http://dx.doi.org/10.1177/13694332221105700.

[15] S. Genna, C. Leone, V. Tagliaferri, Characterization of laser beam transmission through a High Density Polyethylene (HDPE) plate, Opt. Laser Technol. 88 (2017) 61–67, http://dx.doi.org/10.1016/j.optlastec.2016.08.010.

[16] H. Li, J. Ou, Z. Zhou, Applications of optical fibre Bragg gratings sensing technology-based smart stay cables, Opt. Lasers Eng. 47 (10) (2009) 1077–1084, http://dx.doi.org/10.1016/j.optlaseng.2009.04.016.

[17] D. Li, Z. Zhou, J. Ou, Development and sensing properties study of FRP-FBG smart stay cable for bridge health monitoring applications, Meas. 44 (4) (2011) 722–729, http://dx.doi.org/10.1016/j.measurement.2011.01.005.

[18] C. Schaal, S. Bischoff, L. Gaul, Damage detection in multi-wire cables using guided ultrasonic waves, Struct. Heal. Monit. 15 (3) (2016) 279–288, http://dx.doi.org/10.1177/1475921716642747.

[19] H. Zejli, L. Gaillet, A. Laksimi, S. Benmedakhene, Detection of the presence of broken wires in cables by acoustic emission inspection, J. Bridg. Eng. 17 (6) (2012) 921–927, http://dx.doi.org/10.1061/(ASCE)BE.1943-5592.0000404.

[20] F. Xu, X. Wang, H. Wu, Inspection method of cable-stayed bridge using magnetic flux leakage detection: principle, sensor design, and signal processing, J. Mech. Sci. Technol. 26 (3) (2012) 661–669, http://dx.doi.org/10.1007/s12206-011-1234-x.

[21] B. Wang, Y. Zeng, D. Feng, Deep learning-based damage assessment of hinge joints for multi-girder bridges utilizing vehicle-induced bridge responses, Eng. Struct. 333 (2025) 120148, http://dx.doi.org/10.1016/j.engstruct.2025.120148.

[22] B. Wang, Y. Zeng, D. Feng, J. Li, Track vertical irregularity estimation for railway bridges using a novel and lightweight deep-learning architecture, Veh. Syst. Dyn. 63 (8) (2024) 1597–1624, http://dx.doi.org/10.1080/00423114.2024.2375031.

[23] Z. Sun, X. Wang, T. Han, H. Huang, X. Huang, L. Wang, Z. Wu, Pipeline deformation monitoring based on long-gauge FBG sensing system: missing data recovery and deformation calculation, J. Civ. Struct. Heal. Monit. 15 (2025) 2433–2453, http://dx.doi.org/10.1007/s13349-025-00943-9.

[24] Z. Sun, X. Wang, T. Han, L. Wang, Z. Zhu, H. Huang, J. Ding, Z. Wu, Pipeline deformation prediction based on multi-source monitoring information and novel data-driven model, Eng. Struct. 337 (2025) 120461, http://dx.doi.org/10.1016/j.engstruct.2025.120461.

[25] Z. Sun, Y. Li, L. Su, D. Niu, D. Luo, W. He, S. Xie, Investigation of electrical resistivity for fiber-reinforced coral aggregate concrete, Constr. Build. Mater. 414 (2024) 135011, http://dx.doi.org/10.1016/j.conbuildmat.2024.135011.

[26] Y. Li, Z. Sun, Y. Li, H. Yang, X. Liu, W. He, A vision transformer-based method for predicting seismic damage states of RC piers: Database development and efficient assessment, Reliab. Eng. Syst. Saf. 263 (2025) 111287, http://dx.doi.org/10.1016/j.ress.2025.111287.

[27] S. Bhowmick, S. Nagarajaiah, A. Veeraraghavan, Vision and deep learning-based algorithms to detect and quantify cracks on concrete surfaces from UAV videos, Sens. 20 (21) (2020) 6299, http://dx.doi.org/10.3390/s20216299.

[28] S. Tang, Z. Chen, Scale-space data augmentation for deep transfer learning of crack damage from small sized datasets, J. Nondestruct. Eval. 39 (3) (2020) 70, http://dx.doi.org/10.1007/s10921-020-00715-z.

[29] Q. Yang, X. Ji, Automatic pixel-level crack detection for civil infrastructure using U-Net++ and deep transfer learning, IEEE Sens. J. 21 (17) (2021) 19165–19175, http://dx.doi.org/10.1109/JSEN.2021.3089718.

[30] V. Hoskere, Y. Narazaki, T. Hoang, B. Spencer, MaDnet: Multi-task semantic segmentation of multiple types of structural materials and damage in images of civil infrastructure, J. Civ. Struct. Heal. Monit. 10 (5) (2020) 757–773, http://dx.doi.org/10.1007/s13349-020-00409-0.

[31] S. Li, X. Zhao, G. Zhou, Automatic pixel-level multiple damage detection of concrete structure using fully convolutional network, Comput.-Aided Civ. Infrastruct. Eng. 34 (7) (2019) 616–634, http://dx.doi.org/10.1111/mice.12433.

[32] N. Wang, X. Zhao, P. Zhao, Y. Zhang, Z. Zou, J. Ou, Automatic damage detection of historic masonry buildings based on mobile deep learning, Autom. Constr. 103 (2019) 53–66, http://dx.doi.org/10.1016/j.autcon.2019.03.003.

[33] H. Bae, K. Jang, Y. An, Deep super resolution crack network (SrcNet) for improving computer vision-based automated crack detectability in situ bridges, Struct. Heal. Monit. 20 (4) (2020) 1428–1442, http://dx.doi.org/10.1177/1475921720917227.

[34] C. Feng, H. Zhang, H. Wang, S. Wang, Y. Li, Automatic pixel-level crack detection on dam surface using deep convolutional network, Sens. 20 (7) (2020) 2069, http://dx.doi.org/10.3390/s20072069.

[35] B. Kim, N. Yuvaraj, K. Sri Preethaa, R. Arun Pandian, Surface crack detection using deep learning with shallow CNN architecture for enhanced computation, Neural Comput. Appl. 33 (15) (2021) 9289–9305, http://dx.doi.org/10.1007/s00521-021-05690-8.

[36] Z. Bukhsh, N. Jansen, A. Saeed, Damage detection using in-domain and cross-domain transfer learning, Neural Comput. Appl. 33 (24) (2021) 16921–16936, http://dx.doi.org/10.1007/s00521-021-06279-x.

[37] S. Jiang, J. Zhang, Real-time crack assessment using deep neural networks with wall-climbing unmanned aerial system, Comput.-Aided Civ. Infrastruct. Eng. 35 (6) (2020) 549–564, http://dx.doi.org/10.1111/mice.12519.

[38] W. Ding, H. Yang, K. Yu, J. Shu, Crack detection and quantification for concrete structures using UAV and transform, Autom. Constr. 152 (2023) 104929, http://dx.doi.org/10.1016/j.autcon.2023.104929.

[39] Y. Liu, X. Nie, J. Fan, X. Liu, Image-based crack assessment of bridge piers using unmanned aerial vehicles and three-dimensional scene reconstruction, Comput.-Aided Civ. Infrastruct. Eng. 35 (5) (2020) 511–529, http://dx.doi.org/10.1111/mice.12501.

[40] S. Dorafshan, M. Maguire, Bridge inspection: human performance, unmanned aerial systems and automation, J. Civ. Struct. Heal. Monit. 8 (2018) 443–476, http://dx.doi.org/10.1007/s13349-018-0285-4.

[41] S. Dorafshan, R.J. Thomas, C. Coopmans, M. Maguire, A practitioner's guide to small unmanned aerial systems for bridge inspection, Infrastructures 4 (72) (2019) 72, http://dx.doi.org/10.3390/infrastructures4040072.

[42] H. Li, Y. Chen, J. Liu, Z. Zhang, H. Zhu, Unmanned aircraft system applications in damage detection and service life prediction for bridges: A review, Remote. Sens. 14 (17) (2022) 4210, http://dx.doi.org/10.3390/rs14174210.

[43] X. Wang, J. Dong, Y. Yao, K. Hu, J. Cui, Y. Luo, J. Zhang, S. Niu, Energy consumption optimization in trajectory planning for fuel cell hybrid uavs based on HMPC, Adv. Robot. Res. (2025) e2500045, http://dx.doi.org/10.1002/adrr.202500045.

[44] G. Ahmed, T. Sheltami, A. Mahmoud, Energy-efficient multi-UAV multi-region coverage path planning approach, Arab. J. Sci. Eng. 49 (2024) 13185–13202, http://dx.doi.org/10.1007/s13369-024-09295-w.

[45] X. Zhang, Z. Ji, L. Chen, Y. Lyu, UAV path planning via semantic segmentation of 3D reality mesh models, Drones 9 (8) (2025) 578, http://dx.doi.org/10.3390/drones9080578.

[46] P. Cheng, J. Keller, V. Kumar, Time-optimal UAV trajectory planning for 3D urban structure coverage, in: 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems, Nice, France, 2008, pp. 2750–2757, http://dx.doi.org/10.1109/IROS.2008.4650988.

[47] S. Dogru, L. Marques, Energy efficient coverage path planning for autonomous mobile robots on 3D terrain, in: 2015 IEEE International Conference on Autonomous Robot Systems and Competitions, Vila Real, Portugal, 2015, pp. 118–123, http://dx.doi.org/10.1109/ICARSC.2015.23.

[48] S. Mansouri, C. Kanellakis, E. Fresk, D. Kominiak, G. Nikolakopoulos, Cooperative coverage path planning for visual inspection, Control. Eng. Pr. 74 (2018) 118–131, http://dx.doi.org/10.1016/j.conengprac.2018.03.002.

[49] C. Zhang, F. Wang, Y. Zou, J. Dimyadi, B. Guo, L. Hou, Automated UAV image-to-BIM registration for building facade inspection using improved generalised Hough transform, Autom. Constr. 153 (2023) 104957, http://dx.doi.org/10.1016/j.autcon.2023.104957.

[50] Y. Han, D. Feng, W. Wu, X. Yu, G. Wu, J. Liu, Geometric shape measurement and its application in bridge construction based on UAV and terrestrial laser scanner, Autom. Constr. 151 (2023) 104880, http://dx.doi.org/10.1016/j.autcon.2023.104880.

[51] F. Wang, Y. Zou, E. Castillo, Y. Ding, Z. Xu, H. Zhao, J. Lim, Automated UAV path-planning for high-quality photogrammetric 3D bridge reconstruction, Struct. Infrastruct. Eng. 20 (10) (2024) 1595–1614, http://dx.doi.org/10.1080/15732479.2022.2152840.

[52] X. Li, Y. Chen, Z. Chen, Z. Huang, Coverage path planning of bridge inspection with Unmanned aerial vehicle, Eng. Appl. Artif. Intell. 156 (Part B) (2025) 111253, http://dx.doi.org/10.1016/j.engappai.2025.111253.

[53] Z. Xia, J. Shu, W. Ding, Y. Gao, Y. Duan, C. Debono, V. Prakash, D. Seychell, R. Borg, Complete-coverage path planning for surface inspection of cable-stayed bridge tower based on building information models and climbing robots, Comput.-Aided Civ. Infrastruct. Eng. (2025) 1–23, http://dx.doi.org/10.1111/mice.13469.

[54] S. Jiang, C. Jiang, W. Jiang, Efficient structure from motion for large-scale UAV images: a review and a comparison of SfM tools, ISPRS J. Photogramm. Remote. Sens. 167 (2020) 230–251, http://dx.doi.org/10.1016/j.isprsjprs.2020.04.016.

[55] M. Pollefeys, Visual 3D Modeling from Images, University of North Carolina, Chapel Hill, USA, 2004, https://www.cvg.ethz.ch/teaching/compvis/2012/tutorial.pdf.

[56] H. Fathi, F. Dai, M. Lourakis, Automated as-built 3D reconstruction of civil infrastructure using computer vision: achievements, opportunities, and challenges, Adv. Eng. Inform. 29 (2015) 149–161, http://dx.doi.org/10.1016/j.aei.2015.01.012.

[57] Y. Xu, J. Zhang, UAV-based bridge geometric shape measurement using automatic bridge component detection and distributed multi-view reconstruction, Autom. Constr. 140 (2022) 104376, http://dx.doi.org/10.1016/j.autcon.2022.104376.

[58] A. Twinanda, S. Shehata, D. Mutter, J. Marescaux, M. De Mathelin, N. Padoy, Endonet: A deep architecture for recognition tasks on laparoscopic videos, IEEE Trans. Med. Imaging 36 (1) (2017) 86–97, http://dx.doi.org/10.1109/TMI.2016.2593957.

[59] H. Lou, X. Duan, J. Guo, H. Liu, J. Gu, L. Bi, H. Chen, DC-YOLOv8: small-size object detection algorithm based on camera sensor, Electron. 12 (10) (2023) 2323, http://dx.doi.org/10.3390/electronics12102323.

[60] D. Bolya, C. Zhou, F. Xiao, Y. Lee, YOLACT: Real-time instance segmentation, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, ICCV, 2019, pp. 9157–9166, http://dx.doi.org/10.1109/ICCV.2019.00925.

[61] O. Khare, S. Gandhi, A. Rahalkar, S. Mane, YOLOv8-based visual detection of road hazards: potholes, sewer covers, and manholes, in: 2023 IEEE Pune Section International Conference, Pune, India, 2023, pp. 1–6, http://dx.doi.org/10.1109/PuneCon58714.2023.10449999.

[62] J. Su, Y. Qin, Z. Jia, B. Liang, MPE-YOLO: enhanced small target detection in aerial imaging, Sci. Rep. 14 (2024) 17799, http://dx.doi.org/10.1038/s41598-024-68934-2.

[63] Z. Zheng, P. Wang, Distance-iou loss: Faster and better learning for bounding box regression, Proc. AAAI Conf. Artif. Intell. 34 (07) (2020) 12993–13000, http://dx.doi.org/10.1609/aaai.v34i07.6999.

[64] X. Li, C. Lv, Generalized focal loss: Towards efficient representation learning for dense object detection, IEEE Trans. Pattern Anal. Mach. Intell. 45 (3) (2022) 3139–3153, http://dx.doi.org/10.1109/TPAMI.2022.3180392.

[65] X. Yin, Y. Chen, A. Bouferguene, H. Zaman, M. Al-Hussein, L. Kurach, A deep learning-based framework for an automated defect detection system for sewer pipes, Autom. Constr. 109 (2020) 102967, http://dx.doi.org/10.1016/j.autcon.2019.102967.

[66] T. Zhang, C. Suen, A fast parallel algorithm for thinning digital patterns, Commun. ACM 27 (3) (1984) 236–239, http://dx.doi.org/10.1145/357994.358023.

[67] J. Li, C. Wang, P. Li, S. Li, S. Li, Surface damage detection of cable stays based on PointRend model using unmanned aerial vehicles, Struct. Infrastruct. Eng. (2024) 1–17, http://dx.doi.org/10.1080/15732479.2024.2363832.