

Research Article

Automatic Detection of Surface Defects on Underwater Pile-Pier of Bridges Based on Image Fusion and Deep Learning

Shaofei Jiang ^{1,2}, Wei Wang ¹, Zhenheng Su ¹ and Shengxian Wang ¹

¹College of Civil Engineering, Fuzhou University, Fuzhou 350108, Fujian, China

²Fujian Provincial Key Laboratory on Multi-Disasters Prevention and Mitigation in Civil Engineering, Fuzhou 350108, China

Correspondence should be addressed to Shaofei Jiang; cejsf@fzu.edu.cn

Received 9 December 2022; Revised 1 April 2023; Accepted 8 May 2023; Published 1 June 2023

Academic Editor: Yi-Qing Ni

Copyright © 2023 Shaofei Jiang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

As an important part of the bridge structure system, the underwater pile-pier structure usually occurs various defects on its surfaces due to its complex hydrological environment. The existing conventional defect detection approaches exist two aspects of problems: (1) insufficient definition and color distortion of the underwater images, and (2) low efficiency and error-prone. To solve these problems, this paper proposed the target defect detection model by integrating the image-fusion enhancement algorithm and the deep learning algorithm. Firstly, by analyzing the reasons for the degradation of the underwater images, the ACE (automatic color equalization) and CLAHE (contrast limited adaptive histogram equalization) algorithms are selected to enhance the image, respectively. Secondly, the two enhanced images are fused based on the point sharpness weight, and then the fusion results are further sharpened by the USM (unsharp mask) algorithm, thus obtaining the final fused images. Thirdly, 3,200 fused images are taken as the training set, by adopting the YOLOv3 algorithm to train the detection model, and then the training model is validated and tested by the other each 400 fused images, thus building up the target automatic detection model of underwater pile-pier surface defects. Finally, a series of comparison and discussion were conducted to validate the effectiveness of image-fusion and the robustness and effectiveness of the target detection model. The results found that the target detection model has excellent robustness against noise and effectiveness in the surface defect detection. This indicates that the image-fusion approach proposed in this paper can effectively enhance the image features, and the target detection model is feasible, robust, and effective in the automatic detection of surface defects on underwater pile-pier structures.

1. Introduction

The number of existing bridges in service in China has exceeded 900,000, and the proportion of bridges over 30 years will soar from less than 20% in 2014 to 62.7% in 2044 [1]. The explosive growth of old bridges shows that China's bridges have generally entered into a rapid aging stage. During the service process of bridge underwater pile-pier structure, it is constantly affected by several factors, such as current scouring, ship collision, and wave force, which often lead to the appearance of various surface defects, such as cracks, exposed reinforcements, holes, and swellings [2]. With the accumulation of defects, bridge collapse accidents will frequently occur in case surface defects of the bridge underwater pile-pier structure are not detected and found in

time. Therefore, it is particularly urgent and important to carry out the defect detection of underwater pile-pier structures, thus providing accurate and effective data for the damage analysis and evaluation of underwater pile-pier structures [3, 4].

At present, the underwater structure detection of bridges is mainly completed by professional divers or underwater robots carrying equipment to take photos or videos of pile-pier structures [5]. Due to the influence of light and water quality, the images obtained by equipment are blurry, distorted in color, and full of various noises. In addition, the number of defect images obtained is large. These facts bring about low efficiency, high subjectivity, and poor recognition precision when adopting conventional approaches [6]. To efficiently obtain precise recognition results of surface

defects, numerous scholars have investigated on them from two aspects in recent years. One aspect is to enhance the image quality of surface defects, and the other is to develop more effective and intelligent target detection models.

A great number of approaches have been presented for image enhancement. Generally, wavelet transform, bilateral filtering, and Retinex-based approaches are commonly used for underwater image enhancement. For example, Guraksin et al. [7] proposed an underwater image enhancement approach based on the wavelet transform and differential evolution algorithm. This approach effectively improves the visibility and image quality; however, it causes the image to be darker overall. Hassan et al. [8] presented a Retinex-based enhancing approach to enhance underwater images. This algorithm improves the overall underexposure of the images while preserving edge detail, however, the problem of image color distortion is still inevitable. To improve the definition of images, Dan et al. [9] proposed the bilateral filter function with a controllable kernel function to estimate the illumination intensity of images. The experimental results show that this approach can remarkably improve the definition of underwater images. From the abovementioned theory, we can find that the abovementioned approaches can only enhance a certain characteristic of the image and cannot improve the overall effect of the image. As a result, several scholars tried to solve this problem by image fusion with different enhanced images. For example, Zhou et al. [10] proposed a fusion enhancement approach for underwater images based on white balance, guided filtering, and multiexposure sequence techniques to improve image dark details and solve the overenhancement problem of a single algorithm, and yet, it ignored the relationship between degradation and scene depth. Gao et al. [11] proposed an underwater image enhancement approach based on multiscale fusion, which fuses local contrast-corrected images with sharpened images to solve the problem of low contrast and color distortion in underwater images. However, this approach has corresponding restrictions on the field of application, and only local details can be enhanced. In conclusion, none of these approaches can be directly applied to enhance the images of underwater pile-pier structures surface defects. It is necessary to develop corresponding image fusion enhancement approaches for the purpose of solving the problems, such as enhancing contrast and optimizing detail features of defect images.

The deep learning algorithm has been widely applied in target detection in recent years, due to its characteristics of high efficiency, objectivity, and precise recognition accuracy. Zhang et al. [12] were the first to apply deep learning techniques in bridge surface defect detection and proposed the application of convolutional neural network (CNN) algorithm in bridge crack image recognition. It is preliminarily proved that the approach based on deep learning can solve the problem of bridge defect detection. In addition to the classification of the bridge surface defect, it is more important to locate the defect. Yang et al. [13] proposed a vision-based automated method for surface condition identification of concrete structures, consisting of pretrained convolutional neural networks (CNNs), transfer learning,

and decision-level image fusion, to improve the accuracy of crack detection. Afterwards, Yang et al. [14] presented a data-driven model based on 2D convolutional neural networks and the improved bird swarm algorithm to evaluate the torsional capacity of reinforced concrete beams, and the results found that the proposed model outperformed other machine learning models, building codes, and empirical formulas. Cha et al. [15] applied the two-stage target detection algorithm, namely, faster R-CNN, to accomplish the classification and localization of the defects. Due to the insufficient detection efficiency of the two-stage target detection algorithm, it is unable to meet the needs of real-time detection in engineering. To solve this problem, Deng et al. [16] applied the YOLOv2 one-stage target detection algorithm in the detection of cracks in concrete. The experimental results show that using the YOLOv2 algorithm to detect cracks can indeed significantly improve the detection efficiency, but this approach has poor detection performance for targets with large scale. After Joseph and Ali [17] proposed the YOLOv3 target detection algorithm with both detection speed and accuracy, Zhang et al. [18] applied this algorithm to concrete bridge surface defect detection, thus realizing the efficient and accurate detection of common surface defects. Afterwards, Pan and Yang [19] combined the YOLOv3 and CNN algorithm to establish the real-time detection model. The developed model was used to monitor the bolt rotation angle and the results showed that the detection accuracy could reach more than 90%. Liu et al. [20] proposed the modified YOLOv3 model to automatically detect pavement crack and found that the detection effect of the model is higher than other state-of-the-art methods. Through the research of YOLOv3 algorithm, the Darknet53 network is used as the backbone due to its excellent feature extraction capabilities and inference speed in YOLOv3 algorithm, and the multiscale feature map output by the neck module is conducive to detecting objects of different scales. Compared with R-CNN series algorithms, it is found that the YOLOv3 algorithm can maintain a high detection speed while ensuring accurate detection [21, 22]. It is clear that this algorithm is the ideal one for underwater pile-pier defect detection of bridges. However, for the current detection approaches based on deep learning, the existing models and algorithms cannot be directly transplanted and applied for bridge underwater pile-pier structures. For the harsh environment, complicated noise, and blurred defect image details, it is indispensable to train and build up the target detection models for specific environments and defect categories.

To solve the abovementioned problems, this paper presents an automatic detection model by integrating images enhancement and deep learning, which is applicable to detecting and locating the surface defects of bridge underwater pile-pier structures. First of all, this paper proposed the image enhancement approach based on pixel-level fusion, which simultaneously reduces the blurriness of underwater images and strengthens the clarity of defect contours through increasing contrast and correcting color deviation, thus improving the image overall quality and enhancing the defect detail features. Next, the target

automatic detection model was built up by integrating the YOLOv3 algorithm and image enhancement approach, which can mine and learn the defect features in the images than other methods. Finally, a series of comparison and discussion were conducted to validate the effectiveness of image-fusion and the robustness and effectiveness of the target detection model. The paper is organized as follows. Section 2 introduces the image enhancement approach; Section 3 presents the target automatic model; Section 4 involves experimental verification of the model; and the conclusions are made in Section 5.

2. Image Enhancement Approach

2.1. Conventional Approaches of Underwater Image Enhancement. When light propagates underwater, the phenomenon of light absorption and scattering will occur due to the propagation characteristics of light [23]. This further leads to several problems such as insufficient contrast, [24, 25] color distortion, [26, 27], and uneven brightness distribution [28] in underwater images. The abovementioned problems restrict the practical application of underwater images in the defect detection of bridge underwater pile-pier structures. To solve these problems, a number of image enhancement approaches have been developed. Herein, two main common-used conventional approaches are reviewed briefly as follows.

2.1.1. Contrast Limited Adaptive Histogram Equalization. Contrast limited adaptive histogram equalization (CLAHE) [29] is employed to realize contrast enhancement by expanding the gray range. Generally, the algorithm divides the image into blocks and realizes histogram transformation by calculating the transformation function of each pixel neighborhood, which can reduce the loss of image details. In addition, the CLAHE algorithm also restricts the height of the gray histogram by clipping and redistribution, which can effectively solve the problems of excessive detail enhancement and noise amplification. The process of the CLAHE algorithm is as follows:

- (1) Divide the original image into several subregion images according to the image size
- (2) Establish the histogram $H(x)$ of each subregion
- (3) Calculate the clipping amplitude T :

$$T = c \times \frac{H \times W}{M}, \quad (1)$$

where c is the acquisition coefficient; H and W are the numbers of pixels in the height and width direction of the subregion image, respectively; M is the gray level

- (4) Fill the part above the threshold T into the bottom of the histogram and then obtain a new histogram $H'(x)$
- (5) Reconstruct the gray value by the bilinear interpolation calculation for different subregion images

All in all, the CLAHE algorithm has the capability to balance the brightness distribution and significantly improve the contrast, while being less effective in color correction.

2.1.2. Automatic Color Equalization Algorithm. To address the problem that the CLAHE algorithm is not satisfactory in brightness enhancement and color restoration, automatic color equalization (ACE) algorithm [30] emerged as the times require. The algorithm considers the spatial location relation between color and brightness in the image; the pixel values of the enhanced images are obtained by differential calculating the relative light-dark relationship between the target pixels and the surrounding pixels, and finally, the final pixel values are corrected so that the enhanced image has excellent color restoration.

The ACE algorithm is mainly divided into two steps. The first step is to adjust the image domain: substitute the pixel brightness value of the original underwater image I_z into formula (2), and the intermediate image R_z is

$$R_z(k) = \sum_{q=S, q \neq k} r[I_z(k) - I_z(q)]/d(k, q), \quad (2)$$

where $R_z(k)$ is the brightness value of pixel point; $I_z(k) - I_z(q)$ is the brightness value difference of two different pixel points; $d(k, q)$ is the distance function; and $r(*)$ represents the brightness performance function.

The second step is dynamic expansion: adjust the dynamic range of the intermediate image R_z and obtain the final target image O_z

$$O_z(k) = \text{round}[127.5 + s_z \times R_z(k)], \quad (3)$$

where $O_z(k)$ is the brightness value of pixel point; $\text{round} (*)$ is the rounding function; s_z is the slope of $[(m_z, 0), (M_z, 255)]$ of the line segment, where m_z, M_z is calculated as follows:

$$\begin{aligned} m_z &= \min [R_z(k)], \\ M_z &= \max [R_z(k)]. \end{aligned} \quad (4)$$

Through the abovementioned two steps, it can be achieved to correct the image color deviation and improve the overall brightness.

2.2. Image Enhancement Approach Based on Pixel-Level Fusion

2.2.1. Drawbacks of Conventional Enhancement Approaches. Due to the diversity of underwater image degradation reasons, images from different underwater environments need to be enhanced by different algorithms, so a single image enhancement approach can only solve a certain aspect of the problem. Through the comparison of the above enhancement algorithms, it is found that the ACE algorithm can effectively achieve color restoration, correct color deviation, and enhance brightness significantly, but the effect on contrast enhancement is not ideal. On the contrary, the CLAHE algorithm can significantly improve the contrast of underwater images and balance the brightness distribution, but it does not perform well in color restoration and the

overall brightness enhancement. Obviously, both the ACE algorithm and the CLAHE algorithm have ideal complementarity. Therefore, this paper presents an image pixel-level fusion enhancement approach by integrating the ACE algorithm and the CLAHE algorithm.

2.2.2. Image Fusion Enhancement Algorithm Based on Point Sharpness Weight. To obtain images with the better definition, the point sharpness value of different enhanced images is calculated and selected as the fusion weight. The calculation formula of point sharpness value is as follows:

$$E(G) = \frac{\sum_{i=1}^{m \times n} \sum_a^8 (dG/dx)^2}{m \times n}, \quad (5)$$

where m and n are the length and width of images, respectively; dG/dx is the rate of gray level; and $E(G)$ is the calculated point sharpness value.

The steps of the image fusion approach proposed in this paper are as follows:

- (1) Enhance the original images of underwater pile-pier structures by the ACE and CLAHE algorithms, respectively and obtain two enhanced images
- (2) Adopt the improved point sharpness formula to calculate the point sharpness value of the two enhanced images and normalize it as their weight value, respectively. The calculation formula is as follows:

$$W_A = \frac{E(G)_A}{E(G)_A \times E(G)_C}, \quad (6)$$

$$W_C = \frac{E(G)_C}{E(G)_A \times E(G)_C}, \quad (7)$$

where $E(G)_A$ and $E(G)_C$ are the point sharpness values of the image enhanced by the ACE and CLAHE algorithm, respectively; W_A and W_C are the corresponding image weight coefficients, respectively

- (3) Decompose the two enhanced images into three single RGB channel images and fuse the corresponding channel values by the weight coefficients from Formulas (6) and (7)
- (4) Recombine the three fused single-channel images to obtain the final fused image

The image fusion process based on the point sharpness value weight is shown in Figure 1.

To further reduce the noise interference, the USM algorithm is adopted to further sharpen the fused image. More specifically, after the Gaussian blur processing is performed on the input image, the extracted high-frequency components are multiplied by the sharpening coefficients and then resuperimposed on the input image; finally, the resuperimposed image is filtered and denoised, respectively. The calculation formula is as follows:

$$g(a, b) = [f(a, b) + \omega \cdot h(a, b)] \otimes g_\sigma, \quad (8)$$

where $f(a, b)$ is the input image; $h(a, b)$ is the high-frequency component; ω is the sharpening coefficient, usually the value is 0.6; and g_σ and \otimes represent filter denoising and convolution operations, respectively.

2.3. Verification of the Enhancement Approaches

2.3.1. Image Results from Enhancement Approaches. To verify the effectiveness of the fusion algorithm proposed in this paper, two common surface defects of underwater structures are given as examples: crack (a-1) and exposed reinforcement (a-2). After being scanned and photographed by the underwater visible camera, the original images of surface defects are obtained. The ACE, CLAHE, and the fusion algorithm proposed in this paper are applied to enhance the acquired defect images of the underwater pile-pier structures, respectively. The results are shown in Figure 2.

It can be seen from Figure 2 that the image enhanced by the fusion algorithm proposed in this paper is the best, ACE is the second best, and CLAHE is the worst. As shown in Figures 2(d-1) and 2(d-2), the fused images can not only highlight the crack and reinforcement details but also recover the concrete surface pores and hollows distinctly. This indicates that the fusion algorithm proposed in this paper can solve the problems of blurring, indistinguishable contours, and color distortion of the original image; furthermore, it is beneficial to feature extraction of the image content and defect discrimination. Correspondingly, the overall color is well recovered from the images enhanced by the ACE algorithm in Figures 2(b-1) and 2(b-2); however, there still exist problems of local darkness and low contrast at the periphery of the images. Note also that the images enhanced by the CLAHE algorithm in Figures 2(c-1) and 2(c-2) have a better effect of defogging and enhanced contrast, while the color correction on the concrete surface has no significant effect.

In summary, the image-fusion enhancement algorithm proposed in this paper combines the advantages of the ACE algorithm and the CLAHE algorithm, which is suitable for image enhancement of bridge underwater pile-pier structure surface defect.

2.3.2. Comparison and Discussion. To observe and quantitatively assess the efficiency of the enhancement approach proposed in this paper, SIFT [31] (scale invariant feature transform) approach was employed to quantitate images from different enhancement approaches. The essence of the SIFT approach is firstly to find feature points on different scale-spaces, then to calculate the gradient directions of the feature points, and to adopt the calculated gradient directions to build up match relationships between images on different scale spaces finally. The image quality is evaluated according to the number of feature points and matching relationships. Specifically, the more feature points and matching relationships found, the higher the image quality is.

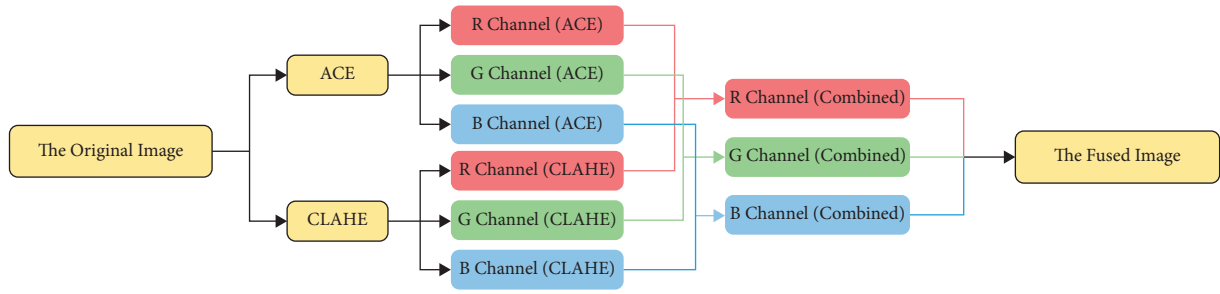


FIGURE 1: Image fusion algorithm process.

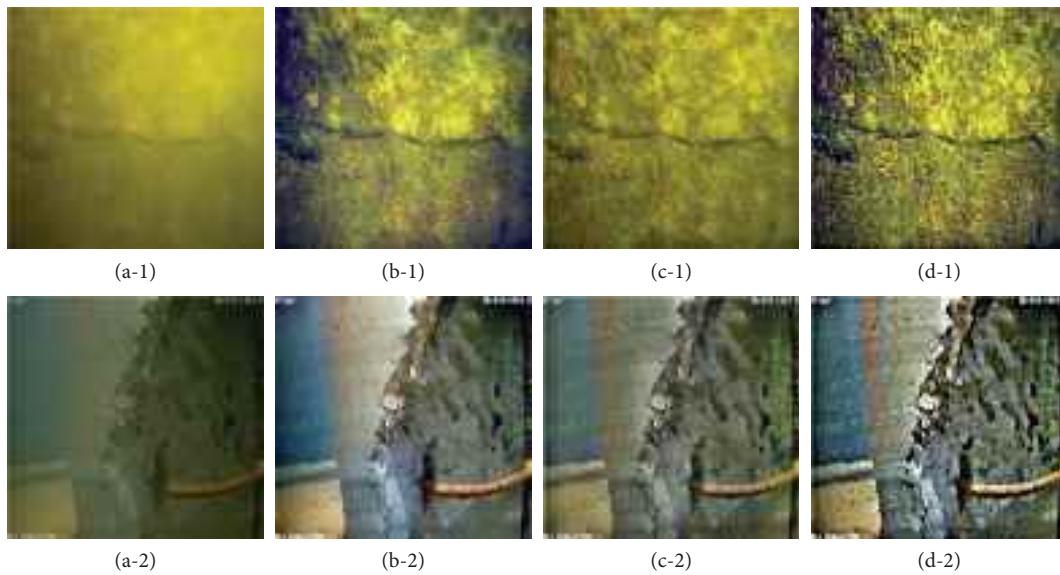


FIGURE 2: The image enhancement results: (a) the original image, (b) image enhanced by ACE algorithm, (c) image enhanced by CLAHE algorithm, and (d) the fused image: (1) the crack image and (2) the exposed reinforcement image.

This SIFT approach generally includes three steps. (1) Extract feature points; (2) Locate feature points and determine feature gradient directions; (3) Find several pairs of feature points that match each other, and establish the corresponding relationship.

On the basis of the abovementioned steps, enhanced image feature points and matching relationships are calculated and depicted in Figures 3 and 4 with the ACE algorithm, the CLAHE algorithm, and fusion algorithm proposed in this paper, respectively. Here, the yellow numbers represent the feature points and the green lines represent the matching relationships.

It is obviously seen in terms of the number of feature points and matching relationships that the images enhanced by the fusion algorithm are the best, the ones by the ACE algorithm are the second place, the ones by the CLAHE algorithm rank third, and the original ones are the worst. Especially for the crack image, as shown in Figures 3(a-1)–3(d-1), the number of feature matching points is increased from zero to thousands after the enhancement by the fusion algorithm. From Figures 3(a-2)–3(d-2), the image feature points of exposed reinforcements are increased to hundreds after enhancement by the fusion algorithm, while the

original image has only 10 feature points. The specific numerical values of the feature points and matching relationships are demonstrated in Figure 4.

In conclusion, the images enhanced by the fusion algorithm proposed in this paper have more feature points and better matching performance than those of other enhancement algorithms. It can be proved that the fusion algorithm proposed in this paper is effective and feasible, which can significantly improve the detail feature information of the bridge underwater pile-pier surface defect images. It is conducive to the target detection model to extract the defect features and thus improving the detection effect of the target automatic detection model.

3. Target Automatic Detection Model

According to the analysis and summary of the abovementioned, the target automatic detection model is presented in this paper. Firstly, the underwater pile-pier surface defect images are obtained by the underwater visible camera. The acquired images are then amplified by rotating, flipping, and scaling transformations. Afterward, the actual damage locations on the images are marked with regions. Finally, the

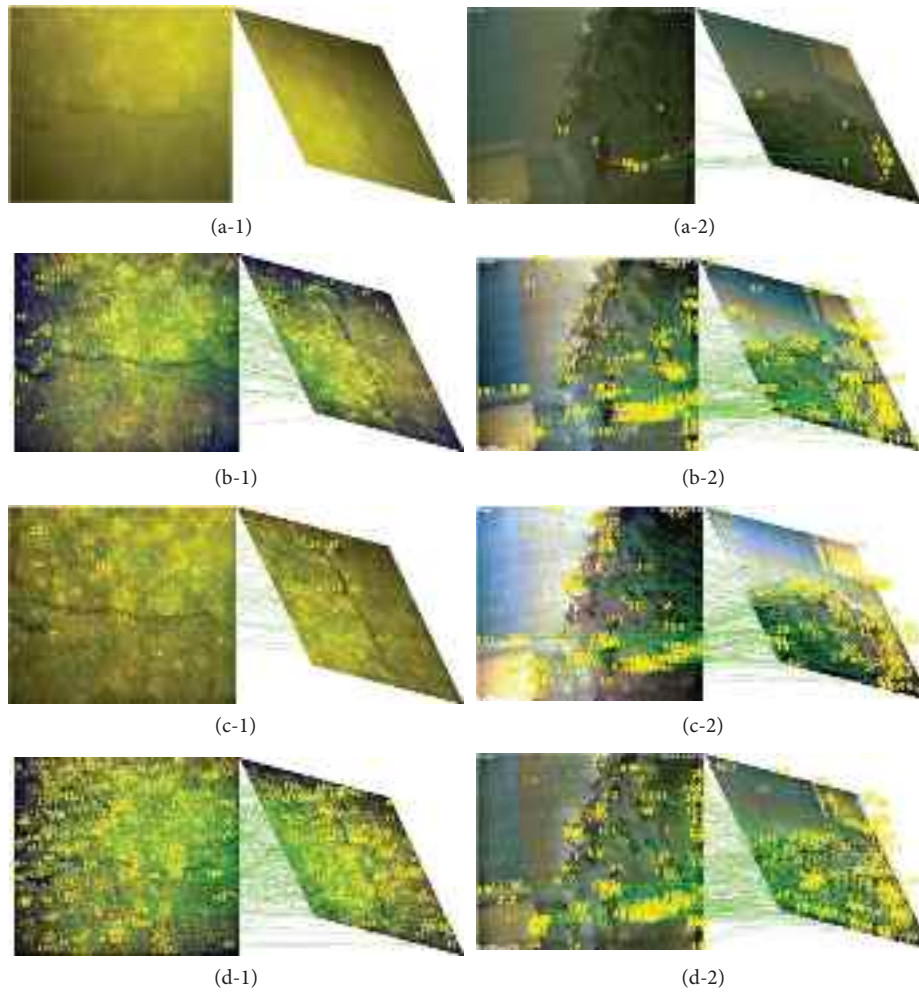


FIGURE 3: The matching results of Figure 2 based on the SIFT feature matching approach: (a) the original image feature matching, (b) the ACE algorithm enhanced image feature matching, (c) the CLAHE algorithm enhanced image feature matching, and (d) the fused image feature matching: (1) the crack image and (2) the exposed reinforcement image.

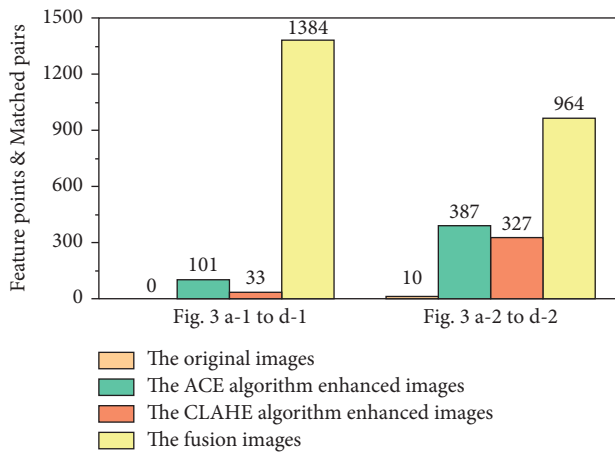


FIGURE 4: The number of feature points and matched pairs.

target automatic detection model is trained by the YOLOv3 algorithm and the enhanced images, and the multicategory underwater defect regions are regarded as the target to be

detected, thus building up the target automatic detection model for the underwater pile-pier surface defects. The model-built up can achieve the precise classification and localization of the target defect.

3.1. Data Processing

3.1.1. Underwater Defect Image Data Acquisition. In this paper, four common surface defects of bridge underwater pile-pier structures are selected from the underwater pile-pier images collected in the laboratory as the database, namely, cracks, exposed reinforcements, holes, and swellings. Meanwhile, the database is randomly divided into the training set, the validation set, and the testing set. The training set is used for feature learning and training the parameter weights of the model; the validation set is used for adjusting the hyperparameters of the model and for preliminary evaluation of the trained model; and the function of the testing set is to test the model on data that has not been trained and validated and to evaluate the overall performance of the model for target recognition and localization.

3.1.2. Data Augmentation. The target detection model has a deep structure and a large number of parameters, which requires numerous of data to participate in training so as to update the weights to improve the generalization ability of the model. It is very difficult to obtain enough data; however, the data augmentation approach can effectively solve this problem. Among them, affine transformation refers to the approach of cutting, flipping, scaling, and rotating images. It is one of the most commonly used data augmentation approaches. In this paper, the operation of rotation, flip, and scaling in affine transformation is used to increase the number of data samples. The image obtained by the data augmentation approach is shown in Figure 5.

3.1.3. Data Region Labeling. To achieve the defect automatic detection by the target detection model built up in this paper, the defect regions and defect categories in the images need to be marked manually. The acquired images are labeled by the software labellmg, thus completing the data region labeling. The operation steps are as follows.

Firstly, open the image annotation tool labellmg, click “open” to load the image and then select “Create RectBox” to select the objects in the image, afterwards enter the corresponding defect category; finally, click “save” to save the data as a corresponding “xml” file. Wherein “crack” corresponds to cracks, “exre” corresponds to exposed reinforcements, “hole” corresponds to holes, and “swelling” corresponds to swellings.

The operation interface of labeling the sample image with Labellmg is shown in Figure 6.

3.2. The Model Network Structure. The target automatic detection model proposed in this paper is built up by integrating the YOLOv3 algorithm and image-fusion enhancement approach. The steps are firstly to extract features from the input enhanced images through the feature extraction network Darknet53 to generate the corresponding feature map, then to score the targetability of the content contained in the feature map through the anchor box, and finally to predict the category, location, and confidence of the detected target. Because the model is capable of fusing features on each scale to achieve prediction on feature maps of three different scale sizes, it can significantly enrich the information of the feature maps, making the network model learn more features and improving the detection performance of the model. Simultaneously, the residual structure is added to the network to prevent from gradient disappearance and gradient explosion caused by too deep network structure and too many parameters. The network structure of the model in this paper is shown in Figure 7.

Among them, the feature extraction module is the core part of the model in this paper, which determines the performance of the whole network model. The model adopts the Darknet53 network with deeper network layers and more convolutional layers and adds the residual network to solve the problem of nonconvergence of network training.

3.2.1. Resblock. Resblock consists of CBL (Conv2D_BN_Leaky) and Res_unit, which are the basic components of this network structure. Among them, the function of the CBL component is feature extraction as well as downsampling, and Res_unit ensures that the training will not be nonconvergence due to the deep network structure. The Darknet53 network of the feature extraction module contains five different Resblock units, thus enabling effective extraction of features from the target defect even in the case of the deep network layer.

3.2.2. CBL Component. The CBL component contains the convolutional layer, the BN (batch normalization) layer, and the activation layer. Its main function is to extract features from the images and to recognize the category and location of the defect.

The convolutional, BN, and activation layers of this model are specified as follows:

- (1) The convolutional layer is the most important structure in the target detection algorithm, which contains several different convolutional kernels. Each element of the constituent convolutional kernels corresponds to a weight coefficient and a deviation coefficient, and its main function is to perform dot product operations with image data by the convolutional kernels thus achieving feature extraction. The calculation formula is as follows:

$$x_i^{(l)} = \sum x_i^{(l-1)} \times k_{ij}^{(l)} + b_i^{(l)}, \quad (9)$$

where $x_i^{(l)}$ is the i^{th} output of the l^{th} layer, $x_i^{(l-1)}$ is the i^{th} output of the upper layer, $k_{ij}^{(l)}$ is the convolutional kernel of the l^{th} layer, and $b_i^{(l)}$ is the i^{th} deviation coefficient of the l^{th} layer;

- (2) The function of the BN layer is mainly to normalize the image data before inputting it to the next layer, which enables to reduce the variability between data. The calculation formula is

$$\hat{x}^{(k)} = \frac{x^{(k)} - \mu[x^{(k)}]}{\sigma[x^{(k)}]}, \quad (10)$$

where $\hat{x}^{(k)}$ is the results after normalization, μ is the data mean value, and σ is the data standard deviation;

- (3) The activation layer provides the network with nonlinear modeling capability. Only when the network model contains the activation function, the deep network has the ability to learn nonlinear mapping in layers. Otherwise, it is difficult to effectively model the data with the nonlinear distribution. The activation layer in this paper adopts the Leaky ReLU function. The calculation formula is as follows:

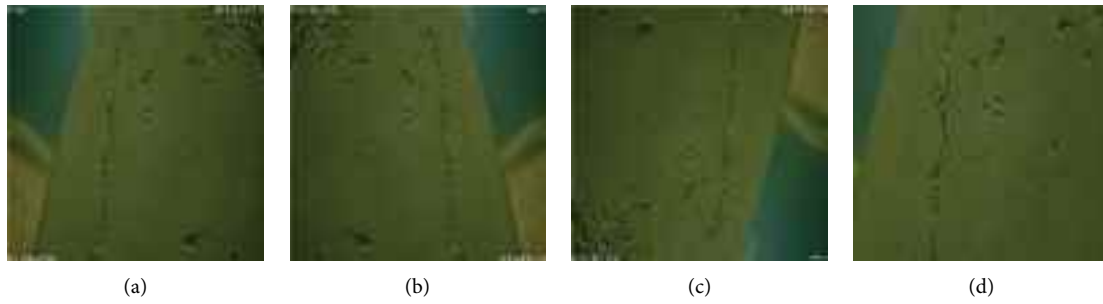


FIGURE 5: Image data augmentations: (a) The original image; (b) image processed by horizontal flip; (c) image processed by rotating 180°; (d) image processed by enlargement.



FIGURE 6: Labeling defect images with LabelImg.

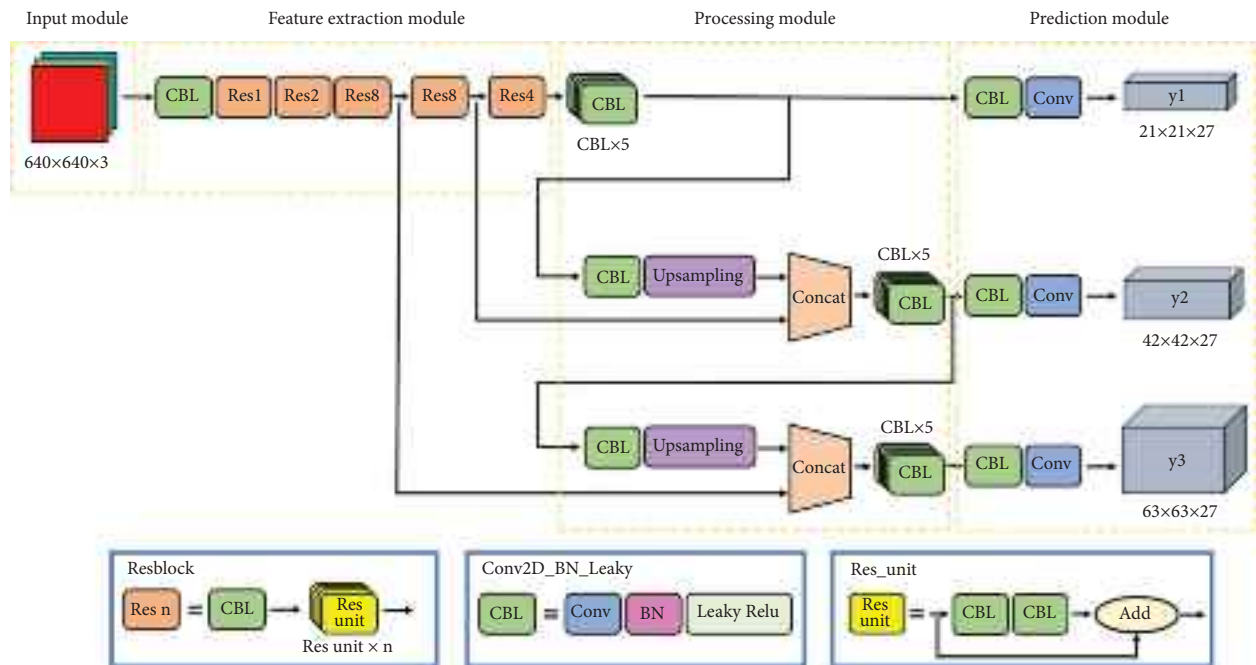


FIGURE 7: The model network structure.

$$L(x) = \begin{cases} \max(0, x), & x \geq 0, \\ \alpha x, & x < 0, \end{cases} \quad (11)$$

where $\max()$ is the function of the maximum value and α takes the value of 0.01.

3.2.3. Res_unit. The network model introduces the skip connection between every two layers of CBL, which can be activated from a certain layer of the network and fed back to a deeper layer of the network immediately. This skip connection can prevent the phenomenon of gradient disappearance and gradient explosion due to the network depth, which leads to nonconvergence of training. Therefore, the skip connection can ensure that the deep network model is able to be trained effectively and gain precise results.

3.3. The Loss Function. The loss function is applied to determine whether the model training has converged or not. The model training is stopped when the loss value reaches a certain threshold, and the model is considered to have achieved the target effect. The loss function also acts as an important tool to evaluate the difference between the actual value and the prediction result of the model and can provide the direction for the optimization of the model. The loss function in this paper mainly consists of confidence loss, classification loss, and bounding box loss. The calculation formula is as follows:

$$L = \lambda_1 L_{\text{conf}} + \lambda_2 L_{\text{class}} + \lambda_3 L_{\text{loc}}, \quad (12)$$

where L_{conf} is the confidence loss, L_{class} is the classification loss, L_{loc} is the bounding box loss, and $\lambda_1, \lambda_2, \lambda_3$ are the balance coefficients.

3.3.1. Target Confidence Loss Function. The target confidence refers to the probability that the target to be predicted is in the rectangular recognition box. This paper adopts the binary cross entropy loss function, which is

$$\begin{aligned} L_{\text{conf}}(\hat{x}_i, y_i) &= -\sum (y_i \ln(\hat{x}_i) + (1 - y_i) \ln(1 - \hat{x}_i)), \\ \hat{x}_i &= \text{sigmoid}(x_i) \\ &= \frac{1}{1 + e^{-x_i}}, \end{aligned} \quad (13)$$

where x_i is the predicted value of the target to be detected, \hat{x}_i is the sigmoid probability of the predicted value, and y_i is the presence or absence of the target to be predicted in the prediction box, taking the value of 0 or 1, herein 0 and 1 represents with absence and presence, respectively.

3.3.2. Target Classification Loss Function. Although the targets to be detected in this paper are four types of defects (cracks, exposed reinforcements, holes, and swellings), it is worth noting that the classification loss function is still adopted the binary cross entropy loss function. The reason

for this is that only positive samples have target classification loss. That is to say, when one type of target defect is detected in the recognition box, the other three types of defects are considered as the same category target defect that is absent in the recognition box. The calculation formula is as follows:

$$\begin{aligned} L_{\text{class}}(\hat{x}_{ij}, y_{ij}) &= -\sum_{i \in \text{loc}} \sum_{i \in \text{class}} (y_{ij} \ln(\hat{x}_{ij}) \\ &\quad + (1 - y_{ij}) \ln(1 - \hat{x}_{ij})), \\ \hat{x}_{ij} &= \text{sigmoid}(x_{ij}) \\ &= \frac{1}{1 + e^{-x_{ij}}}, \end{aligned} \quad (14)$$

where x_{ij} is the predicted value of the target to be detected, \hat{x}_{ij} is the sigmoid probability of the predicted value, and y_{ij} is the presence or absence of the j^{th} defect in the i^{th} target detection box, taking the value of 0 or 1, herein 0 and 1 represents with absence and presence, respectively.

3.3.3. Target Localization Loss Function. The target localization loss function of the algorithm in this paper adopts the sum of squared error loss function, which is the sum of squares of the difference value between the true value and the predicted value. The calculation formula is as follows:

$$\begin{aligned} L_{\text{loc}}(x, y, w, h, I) &= \sum_{i \in \text{loc}} \sum_{i \in \text{class}} \left[(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 \right. \\ &\quad \left. + (w_i - \hat{w}_i)^2 + (h_i - \hat{h}_i)^2 \right], \end{aligned} \quad (15)$$

where x_i, \hat{x}_i are the actual and predicted values of the horizontal coordinates of the center point of the target detection box, y_i, \hat{y}_i are the actual and predicted values of the vertical coordinates of the center point of the target detection box, w_i, \hat{w}_i are the actual and predicted values of the width of the target detection box, and h_i, \hat{h}_i are the actual and predicted values of the height of the target detection box.

4. Experiment Verification

4.1. Data Acquisition. At present, there is no open-source database for the images of surface defect on bridge underwater pile-pier structures; therefore, it is necessary to collect images from experiments and practical engineering. Images in this paper are mainly obtained from two approaches. The first approach is to cast pile-pier components with common surface defects in the laboratory and to place them in the pool, and then the defect images are obtained by the underwater visible camera; the second approach is mainly through the detection of underwater pile-pier on site (on-site detection of Wulongjiang Bridge in Fuzhou, China) and collection on line (detection reports on the underwater structures of bridges in Fujian province). In total, this paper collected 800 original images, of which 669 were from experiments and 131 were from practice engineering or the networks.

Through the investigation of numerous bridges across the hydrological environment in Fujian province (Jinshan Bridge, Minqing bridge, Jimei bridge, etc.), it is found that there are four types of the most common and influential surface defects for bridge underwater pile-pier structures, namely, cracks, exposed reinforcements, holes, and swellings. As a consequence, the abovementioned four types of defects are mainly simulated on the cast pile-pier components. Pool and partial components with defects are shown in Figure 8. Herein, the underwater visible camera was used to collect the surface defect images of the underwater pile-pier structures.

4.2. Software and Hardware Configuration. Since the training phase of the target detection model needs consuming a high degree of computer resources and taking a long training time, the cloud server was employed to train the target detection model. The operating environment configuration in this paper is as follows: the operating system is Linux Ubuntu-4ubuntu0.3, the programming language is Python, the framework is Pytorch, and the graphics card is GeForce RTX 3090 with 23G memory.

4.3. Training and Validation Phases of the Target Automatic Detection Model. There are 800 images obtained by the experiment or practical engineering collected, which contains 243 hole images, 272 crack images, 138 exposed reinforcement images, and 147 swelling images. After data augmentation method, 3200 images were generated through various operations. Among them, 800 images augmented by horizontal flip operation, 800 images augmented by image enlargement operation, 800 images augmented by image scaling operation, and 800 images augmented by rotation operation. 3,200 images, 80 percent of the 4,000 fused image samples, were randomly selected as the training set, and each 400 images (10% of the 4,000 fused images) were selected as the test set and validation set, respectively. The input image size is set to 640×640 pixels, the initial learning rate is 0.01, the momentum is 0.937, the weight decay is 0.0005, the batch size is set to 128, the backbone network is Darknet53, and the number of the training epochs is 500. The loss curve of each module is shown in Figure 9.

From the observation of Figure 9, the following characteristics can be found:

- (1) The more times the model is trained, the smaller the training and validation loss values are. Especially in the first 100 epochs of training, the loss curve decreases rapidly; afterwards, the loss values hardly change in both the training and validation phases. This indicates that the model learns numerous feature information, and the weight parameters change significantly after the training phase.
- (2) The convergence effect reaches the ideal situation when the epoch is 500. At this moment, the loss curve is already close to the horizontal level. This indicates the difference between the predicted value and the actual value is extremely small, and this phenomenon also appears in all other three diagrams in Figure 9.
- (3) For both the bounding box loss curve and the classification loss curve, the iterations on both the training set and validation sets converge well. This indicates that as the epoch increases, the model has a better ability to locate and classify the target defect. The binary cross-entropy loss function is adopted to train the classification ability of the model in this paper. This means that the loss value will decrease remarkably when the model correctly classified the defect type. That is to say, the model can rapidly acquire the ability as to how to correctly classify defects, which is why the classification loss value can steadily converge around 0. Compared with the classification loss value, the bounding box loss value converges steadily around 0.02. The reason for this is that certain errors will occur between the localization box from the predicted model and the rectangular box by manual labeling; furthermore, the localization box will have a larger range than the actual target defect. However, the loss value of bounding box converges to 0.02. This also indicates that the target detection model has remarkable localization ability.
- (4) Although there appears certain deviation between the training set loss and the validation set loss in the confidence loss curve, the difference is not obvious and whose values converge below 0.02. This indicates that there exists the slight overfitting of the model in the aspect of confidence, but the phenomenon of slight overfitting does not affect the overall recognition performance of the model. This can be proven in the following test set results.

To sum up, the target automatic detection model proposed in this paper has excellent convergence performance in the training and validation phases. It is capable of adequately extracting the effective feature information in the fusion images and can intelligently and efficiently recognize the target defect on both the training and validation sets.

4.4. Testing Phase of the Target Automatic Detection Model. To verify the generalization ability of the trained target detection model, 400 images from the test set were input into the trained and validated model. After the detection results were obtained from the automatic detection model, three indices were employed to evaluate the performance of the model.

4.4.1. Fusion Images Detection Results. The images from the test set were recognized, classified, and localized by the trained model, and the bounding boxes with defect categories and confidence were outputted. The partial detection results are shown in Figure 10.

The detection results demonstrate that the target detection model presented in this paper is capable to achieve automatic detection for the underwater pile-pier structure surface defect images after a comparison of Figures 10(a)

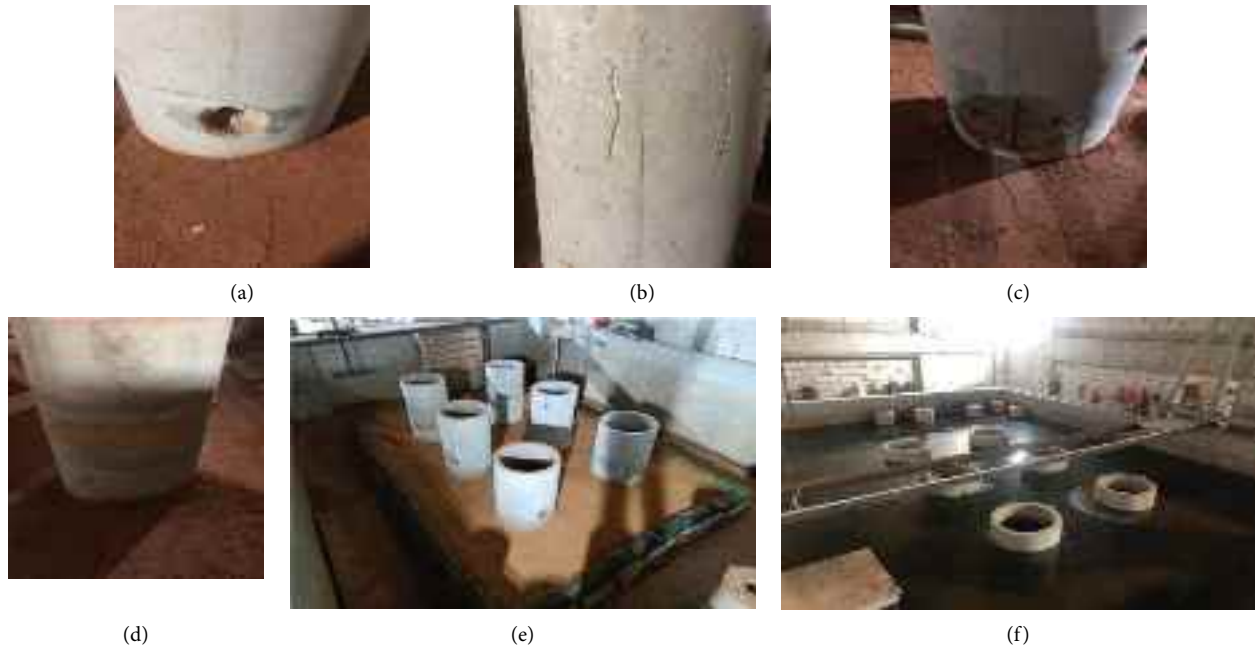


FIGURE 8: Pile and pier components: (a) hole, (b) crack, (c) exposed reinforcement, (d) swelling, (e) component layout, and (f) underwater maintenance of components.

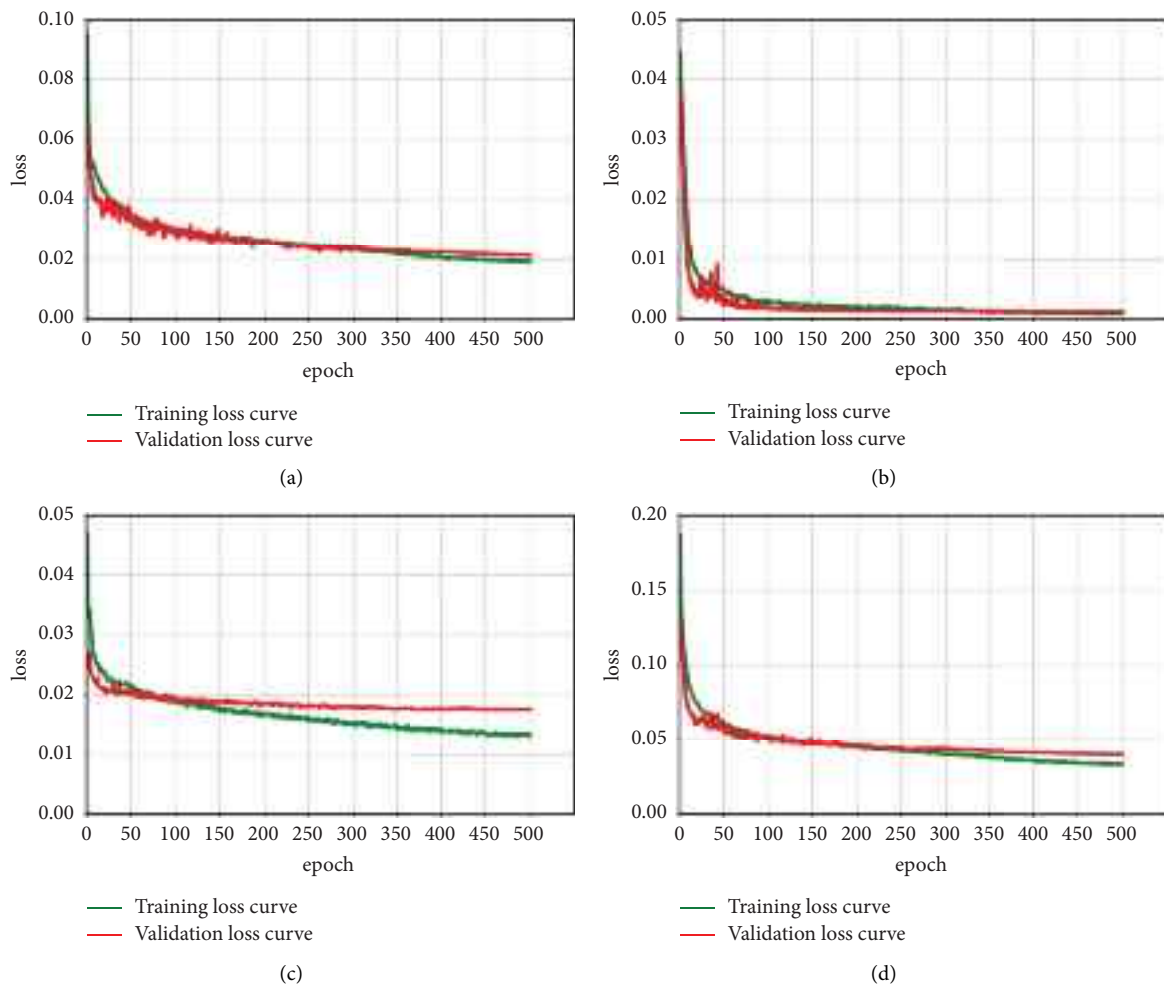


FIGURE 9: Loss curves of different modules in the network model: (a) bounding box loss curve, (b) classification loss curve, (c) confidence loss curve, and (d) total loss curve.

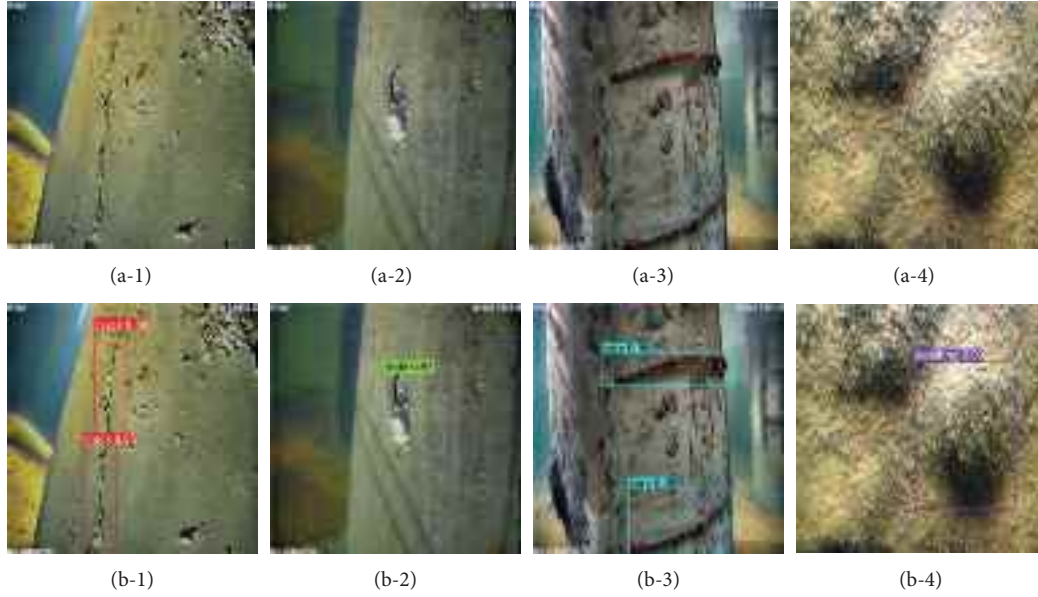


FIGURE 10: Detection results of testing set: (a) input enhanced images and (b) detection results: (1) crack, (2) hole, (3) exposed reinforcement, and (4) swelling.

and 10(b). Wherein the rectangular box in the images indicates the location of the detected defect in Figure 10(b), and the label of the box displays the type and confidence of the defect. Obviously, the confidence of four types of defects is around 0.95, without classification errors and localization bias. This indicates that the model has the outstanding ability to process defect image characteristics information and to generalize it for the purpose of predicting underwater pile-pier structure surface defect.

4.4.2. Model Performance Evaluation Index. To quantitatively evaluate the performance of target automatic detection model, four evaluation indices were used in this paper, namely, the recall (R), precision (P), average precision (AP), and mean average precision (mAP). They are

$$\begin{aligned}
 R &= \frac{TP}{TP + FN}, \\
 P &= \frac{TP}{TP + FP}, \\
 AP &= \int_0^1 P(R) dR, \\
 mAP &= \frac{\sum_{i=1}^m AP_i}{m},
 \end{aligned} \tag{16}$$

where R is the ratio of the number of detected targets to the total number of targets; P is the ratio of the number of detected targets to the number of all detected targets; TP is the number of correct detections of the target defect; FN is the number of target defects that are incorrectly detected as other defects; FP is the number of other defects detected as target defects; AP is the area of the curve surrounded by the horizontal and vertical

coordinates with recall (R) and precision (P); mAP is the mean value of the average precision (AP) of all defect categories.

The four evaluation indices were employed to assess the performance of the model, and the evaluation results are also given.

(1) Evaluation Indices of P and R . The P and R of target detection model trained in this paper are as follows.

From Table 1, it is seen that the mean value of the P reaches 95.19%. Herein, the maximum value of P is 98.63% for the swelling defect, while the minimum value of P is 90.48% for exposed reinforcement defect. This implies that the false detection rate of the four types of defects is extremely low. However, the mean value of R reaches 88.04%. This indicates that the model rarely misses defect detection, and all types of defects presented in the image can be generally detected.

All in all, it is evident that the model built up in this paper has not only a high correct identification rate, but also a low probability of missing detection.

(2) Evaluation Indices of AP and mAP . The AP of target detection model trained in this paper is as follows.

It is obvious in Figure 11 that the AP of cracks, holes, exposed reinforcements, and swellings reach 94.29%, 97.94%, 90.90%, and 99.91%, respectively. And, the mAP value reaches 95.76%. These values are all above 90% and even the AP of swelling defect reaches 99.91%. This reveals that the target detection model has the excellent ability of recognition and classification for all types of defects.

In conclusion, the target automatic detection model proposed in this paper is feasible and effective, which enables to meet the actual engineering requirements.

TABLE 1: The results of precision and recall.

Defect types	P (%)	R (%)
Crack	93.14	84.02
Hole	98.50	97.04
Exre	90.48	80.00
Swelling	98.63	98.63
Mean	95.19	88.04

The significance of bold values is the average values of precision and recall of the above four defects, respectively.

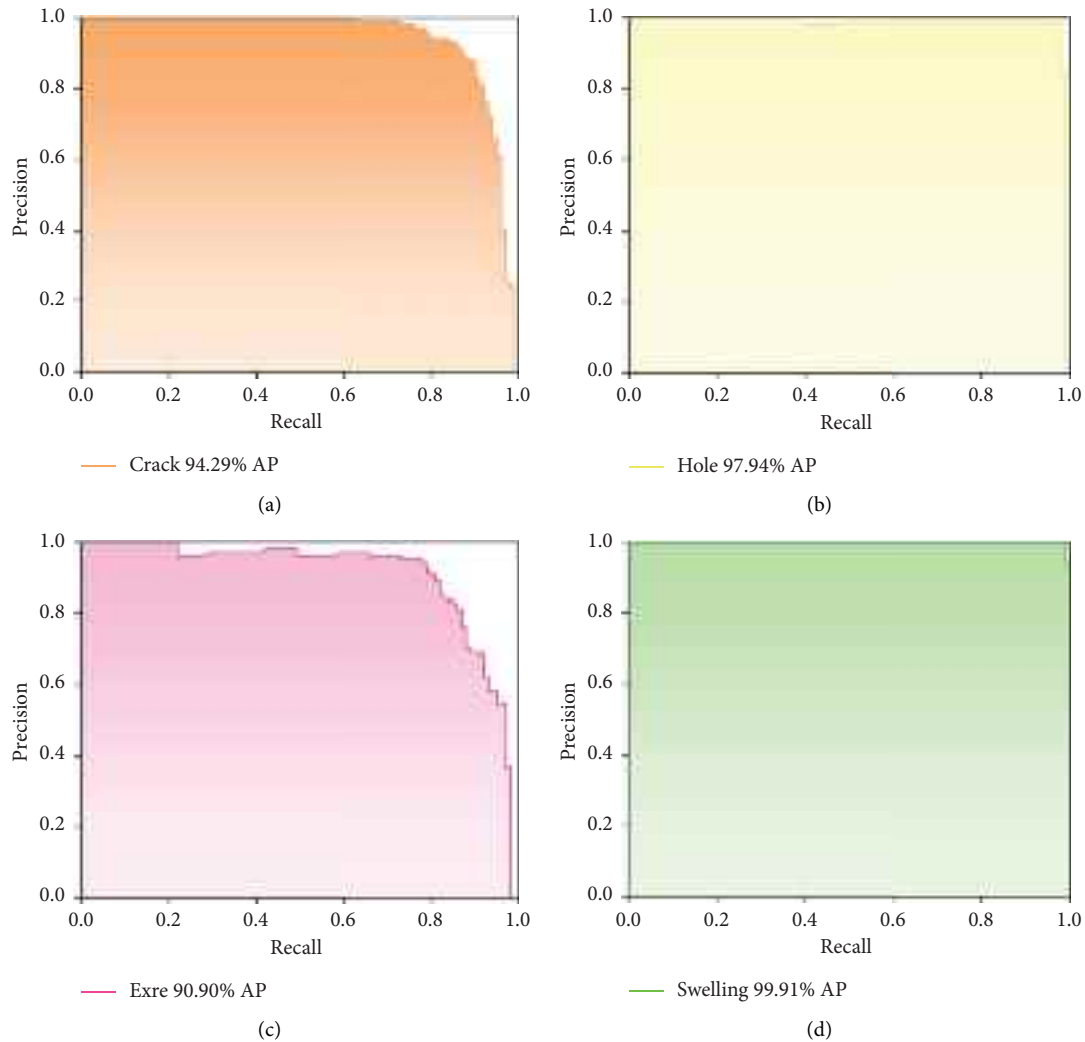


FIGURE 11: The AP of detection results: (a) the AP of cracks, (b) the AP of holes, (c) the AP of exposed reinforcements, and (d) the AP of swellings.

4.5. Comparison and Discussion. A series of comparison and discussion were conducted to validate the effectiveness of image-fusion, the robustness and effectiveness of the target detection model. Herein, it is divided into three parts. In the first part, recognition effect is compared between images without and with fusion. The second part examines the robustness of model under different noise. In final, the third part discusses the effectiveness of the target detection model.

4.5.1. Images without and with Fusion. A comparison and discussion were made between the model presented in this paper using the images without and with the fusion enhancement algorithm. Herein, the same number of images and the same algorithm were used to build up the target detection model for the original images. The partial detection results between the original images and the fused images are shown in Figure 12.

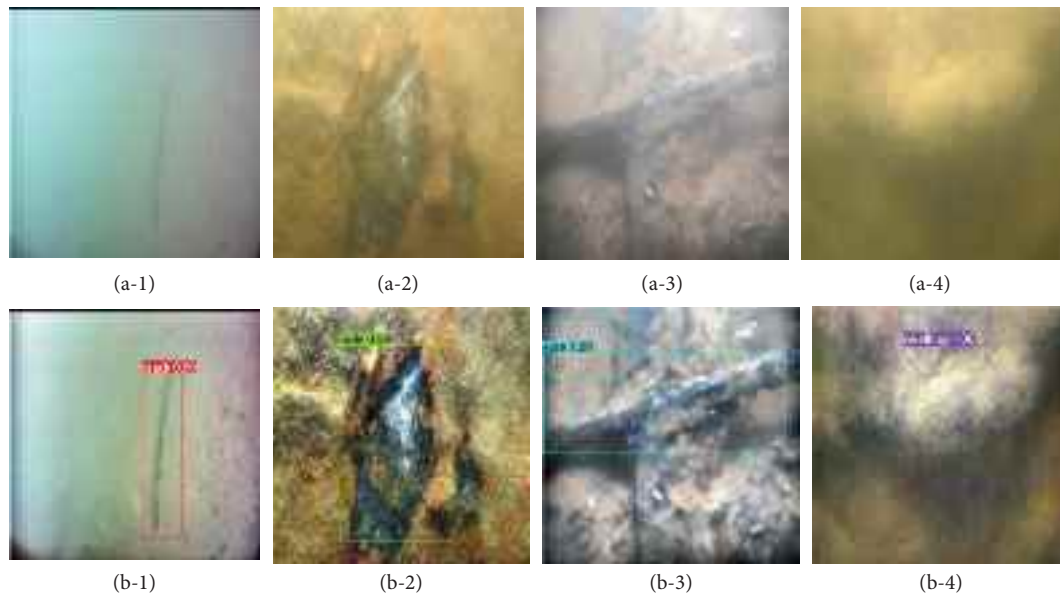


FIGURE 12: The comparison of detection results between the original image and the fused image: (a) the detection result of the original image, (b) the detection result of the fused image: (1) crack, (2) hole, (3) exposed reinforcement, and (4) swelling.

It is worth noting that in Figure 12, for the same defect image, the detection results of original images failed to recognize the defect. However, the detection results of the fused images can successfully detect all the defects with accurate classification and localization, and all confidence is between 0.99 and 1. Therefore, it is intuitively seen that the image-fusion enhancement algorithm proposed in this paper can enhance the overall quality of original images and strengthen the defect detail features. It is conducive to feature extraction and detection performance of the target automatic detection model.

The overall detection performance of the models trained separately by the original and fused images is still evaluated by adopting the AP and mAP. The comparison of the model detection performance is shown in Figure 13.

As can be seen from Figure 13, the indices of the target detection model trained by the fused images are both higher than those of the original images in terms of AP. Among them, the largest increment in AP value is 20.39% for exposed reinforcements, while the smallest one is 3.93% for holes. This indicates that the image-fusion enhancement algorithm has the most significant enhancement effect on exposed reinforcement. From the overall perspective, the mAP of the model trained by fused images is 95.76%, which is 11.79% higher than those of the original images. It demonstrates that the images enhanced by the image fusion algorithm proposed in this paper can effectively improve the detection performance of the target automatic detection model.

4.5.2. Noise Effect. In order to test the robustness of the target model against noise effect, Gaussian noise was added to the fusion images. The variance values of Gaussian noise were taken as 0.1, 0.2, 0.3, 0.4, and 0.5, respectively. Five groups of fusion images with Gaussian noise of different

variances were input into the model proposed in this paper. The final recognition accuracy results are shown in the Figure 14.

It is obvious from Figure 14 that as the noise variance increases, the mAP indices gradually decrease. When the noise variance is less than 0.4, the model proposed in this paper has excellent noise-tolerance capacity and robustness, and the mAP is over 80.48%. Moreover, the recognition accuracy of the target detection model is as high as 75.94% even if the noise variance reaches 0.5. It proved that when the defect image becomes more blurred with the influence of noise, the model proposed can still reliably identify, classify, and locate defects. This indicates that the model has excellent robustness and recognition accuracy.

4.5.3. Different Detection Algorithms Effect. The same 3,200 fusion images were employed to train other three models using SSD (single shot MultiBox detector), fast R-CNN (fast region-based convolutional neural network), YOLOv2 algorithms, respectively. A comparison was made among the model proposed in this paper and other three models. Figure 15 depicts the results of testing set with other 400 fusion images.

As can be seen from Figure 15, all four target detection algorithms have excellent recognition capacity, and their mAP values are more than 88%. More specifically, the YOLOv3 algorithm proposed in this paper ranks the first with 95.19%, the fast R-CNN algorithm ranks the second with 90.47%, the YOLOv2 algorithm ranks the third with 89.42%, and the SSD algorithm ranks the fourth with 88.70%. Based on the results of this study, it can be concluded that the model built up using the YOLOv3 algorithm in this paper has demonstrated exceptional accuracy and effectiveness in detecting underwater structures of bridges.

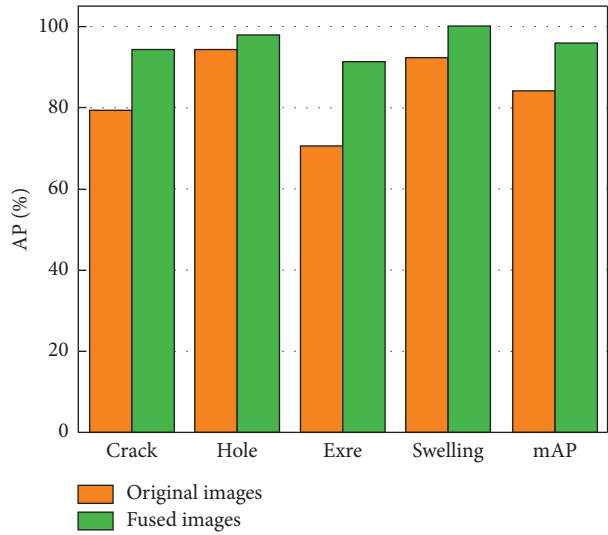


FIGURE 13: Comparison of detection performance.

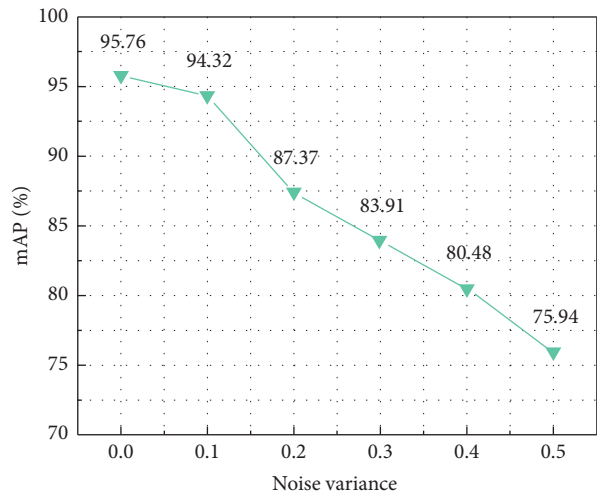


FIGURE 14: The indices of model under variance noise.

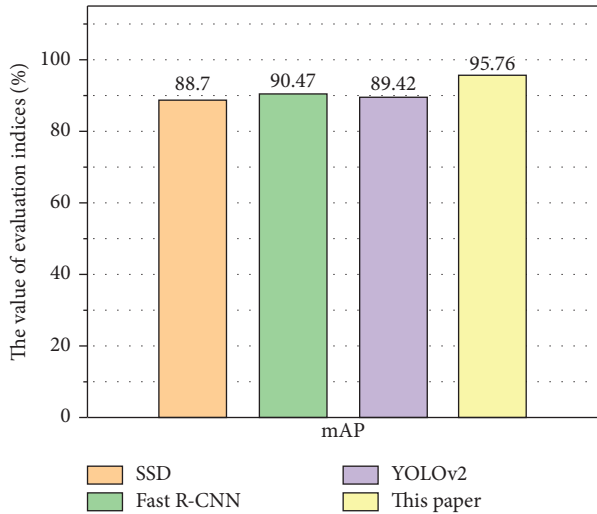


FIGURE 15: The indices of different detection algorithms.

5. Conclusions

This paper firstly proposed an image pixel-level fusion algorithm based on point sharpness weights by analyzing the problems of underwater imaging. This algorithm fuses and enhances the collected images of underwater pile-pier structure surface defects. The fused images were then adopted to build up the target automatic detection model, to realize automatic detection of underwater pile-pier structure surface defects. The main conclusions are as follows:

- (1) This paper proposes the point sharpness weight-based image fusion algorithm that is combined the advantages of the ACE algorithm and the CLAHE algorithm. The results based on the SIFT feature matching approach show that the fusion algorithm can significantly improve the contrast and definition of underwater images and strengthen the images feature information. It is conducive to feature extraction for target detection model.
- (2) This paper proposes the target detection model by integrating the image-fusion enhancement algorithm and the YOLOv3 algorithm. Experimental results show that the model can achieve automatic detection of the underwater pile-pier structure surface defect. This indicates that the model proposed in this paper provides a new intelligent detection technology and is applicable to the identification of bridge underwater pile-pier surface defect.
- (3) The target detection model built-up in this paper can effectively recognize and locate underwater pile-pier surface defects. Four indices, namely, the precision (P), recall (R), average precision (AP), and mean average precision (mAP) are employed to validate the effectiveness of image-fusion and the robustness and effectiveness of the target detection model.

It is evident that the image quality of underwater pile-pier surface defects can be improved and the automatic detection can be effectively performed by the model proposed in this paper. This provides approach and technical support for the automatic detection of underwater pile-pier surface defect. However, the target detection model proposed in this paper can only be used to recognize and locate four common surface defects of bridges underwater pile-pier structures. Furthermore, this model does not support a quantitative evaluation of the defects inside the structure, such as crack size and depth and area of swelling and hole area, which are crucial for assessing the remaining load capacity and service life of bridge structures. These works will be further investigated for us in the future.

Data Availability

The data used to support the findings of this study cannot be shared freely because of third-party rights patient privacy and commercial confidentiality.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This work was helped by some friends, and the authors would also like to thank Associate Research Sheng Shen, Ms Ya-mian Zeng, Dr. Jian-bin Luo for the experiments and the anonymous referees for their constructive comments and suggestions. This study was funded by the Key Project of Fujian Natural Science Foundation, under Grant no. 2022J02016, Shaofei Jiang, and Guiding Key Project for the Social Development of Fujian Province, China, under Grant no. 2020Y0015, Sheng Shen.

References

- [1] Y. Zhang and D. Yang, "Why are detection and monitoring so important?" *China Highway*, vol. 16, pp. 30–33, 2020.
- [2] B. Yang, B. Han, H. Zhao, T. Gu, and T. Cai, "Multi dimensional disease intelligent detection device for underwater pier column structure through machine learning," in *Proceedings of the 2022 IEEE International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA)*, pp. 544–549, Changchun, China, February 2022.
- [3] L. Bao, C. Zhao, X. Xue, and L. Yu, "Improved dark channel defogging algorithm for defect detection in underwater structures," *Advances in Materials Science and Engineering*, vol. 13, 2020.
- [4] Y. Zhu, C. Jiang, Y. Zhang, F. Xiong, F. He, and X. Fang, "Temperature tracer method for crack detection in underwater concrete structures," *Structural Control and Health Monitoring*, vol. 27, no. 9, Article ID e2595, 2020.
- [5] X. Li, H. Sun, T. Song, T. Zhang, and Q. Meng, "A method of underwater bridge structure damage detection method based on a lightweight deep convolutional network," *IET Image Processing*, vol. 16, pp. 3893–3909, 2022.
- [6] W. Zhao, Y. Liu, J. Zhang, Y. Shao, and J. Shu, "Automatic pixel-level crack detection and evaluation of concrete structures using deep learning," *Structural Control and Health Monitoring*, Article ID e2981, 2022.
- [7] G. E. Guraksin, O. Deperlioglu, and U. Kose, "A novel underwater image enhancement approach with wavelet transform supported by differential evolution algorithm," *Nature Inspired Optimization Techniques for Image Processing Applications*, vol. 150, pp. 255–278, 2019.
- [8] N. Hassan, S. Ullah, N. Bhatti, H. Mahmood, and M. Zia, "The Retinex based improved underwater image enhancement," *Multimedia Tools and Applications*, vol. 80, no. 2, pp. 1839–1857, 2021.
- [9] L. Dan, M. Xi, and J. Sun, "Underwater image enhancement algorithm based on controllable nuclear bilateral filtering Retinex," *Chinese Journal of Quantum Electronics*, vol. 35, no. 1, p. 7, 2018.
- [10] J. Zhou, D. Zhang, and W. Zhang, "Underwater image enhancement method via multi-feature prior fusion," *Applied Intelligence*, vol. 52, Article ID 16435, 2022.
- [11] F. Gao, K. Wan, Z. Yang, Y. Wang, and Q. Zhang, "Underwater image enhancement based on local contrast correction and multi-scale fusion," *Journal of Marine Science and Engineering*, vol. 9, no. 2, p. 225, 2021.

- [12] L. Zhang, Y. Fan, Y. D. Zhang, and Y. J. Zhu, "Road crack detection using deep convolutional neural network," in *Proceedings of the 2016 IEEE international conference on image processing (ICIP)*, pp. 3708–3712, Phoenix, AZ, USA, September 2016.
- [13] Y. Yang, B. Samali, M. Rashidi, M. Mohammadi, T. N. Nguyen, and G. Zhang, "Vision-based concrete crack detection using a hybrid framework considering noise effect," *Journal of Building Engineering*, vol. 61, Article ID 105246, 2022.
- [14] Y. Yang, S. Liang, B. Samali et al., "Torsional capacity evaluation of RC beams using an improved bird swarm algorithm optimised 2D convolutional neural network," *Engineering Structures*, vol. 273, Article ID 115066, 2022.
- [15] Y.-J. Cha, W. Choi, G. Suh, S. Mahmoudkhani, and O. Büyüköztürk, "Autonomous structural visual inspection using region-based deep learning for detecting multiple damage types," *Computer-Aided Civil and Infrastructure Engineering*, vol. 33, no. 9, pp. 731–747, 2018.
- [16] J. Deng, Y. Lu, and C.-S. L. Vincent, "Imaging-based crack detection on concrete surfaces using You Only Look once network," *Structural Health Monitoring*, vol. 20, no. 2, pp. 484–499, 2021.
- [17] R. Joseph and F. Ali, "Yolov3: an incremental improvement," 2018, <https://arxiv.org/abs/1804.02767>.
- [18] C. Zhang, C.-C. Chang, and M. Jamshidi, "Concrete bridge surface damage detection using a single-stage detector," *Computer-Aided Civil and Infrastructure Engineering*, vol. 35, no. 4, pp. 389–409, 2020.
- [19] X. Pan and T. Y. Yang, "Image-based monitoring of bolt loosening through deep-learning-based integrated detection and tracking," *Computer-Aided Civil and Infrastructure Engineering*, vol. 37, no. 10, pp. 1207–1222, 2022.
- [20] J. Liu, X. Yang, S. Lau et al., "Automated pavement crack detection and segmentation based on two-step convolutional neural network," *Computer-Aided Civil and Infrastructure Engineering*, vol. 35, no. 11, pp. 1291–1305, 2020.
- [21] B. Benjdira, K. Taha, A. Koubaa, A. Ammar, and K. Ouni, "Car detection using unmanned aerial vehicles: comparison between faster r-cnn and yolov3," in *Proceedings of the 2019 1st International Conference on Unmanned Vehicle Systems-Oman (UVS)*, pp. 1–6, IEEE, Muscat, Oman, February 2019.
- [22] S. Shiv Kumar, M. Wang, D. M. Abraham et al., "Deep learning-based automated detection of sewer defects in CCTV videos," *Journal of Computing in Civil Engineering*, vol. 34, no. 1, Article ID 04019047, 2020.
- [23] H. Lu, Y. Li, Y. Zhang, M. Chen, S. Serikawa, and H. Kim, "Underwater optical image processing: a comprehensive review," *Mobile Networks and Applications*, vol. 22, no. 6, pp. 1204–1211, 2017.
- [24] J. Zhou, X. Wei, J. Shi, W. Chu, and W. Zhang, "Underwater image enhancement approach with light scattering characteristics," *Computers & Electrical Engineering*, vol. 100, Article ID 107898, 2022.
- [25] S. Anwar and C. Li, "Diving deeper into underwater image enhancement: a survey," *Signal Processing: Image Communication*, vol. 89, Article ID 115978, 2020.
- [26] J. Zhou, T. Yang, W. Chu, and W. Zhang, "Underwater image restoration via backscatter pixel prior and color compensation," *Engineering Applications of Artificial Intelligence*, vol. 111, Article ID 104785, 2022.
- [27] C. O. Ancuti, C. Ancuti, C. De Vleeschouwer, and P. Bekaert, "Color balance and fusion for underwater image enhancement," *IEEE Transactions on Image Processing*, vol. 27, no. 1, pp. 379–393, 2017.
- [28] T. Li, S. Rong, X. Cao, Y. Liu, L. Chen, and B. He, "Underwater image enhancement framework and its application on an autonomous underwater vehicle platform," *Optical Engineering*, vol. 59, no. 8, Article ID 083102, 2020.
- [29] K. Zuiderveld, "Contrast limited adaptive histogram equalization," *Graphics Gems*, vol. 474–485, 1994.
- [30] A. Rizzi, "A new algorithm for unsupervised global and local color correction," *Pattern Recognition Letters*, vol. 24, no. 11, pp. 1663–1677, 2003.
- [31] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.