Segmentation and grade evaluation of corrosion on hydraulic steel gates based on image-level labels

Wenheng Zhang[1,2] · Yuqi Zhang[1,2] · Qifeng Gu[1] · Huadong Zhao[1,2]

Abstract

Machine vision offers distinct advantages, such as enhanced efficiency and precision, in the segmentation and assessment of corrosion on hydraulic steel gates. This study addresses the challenge of demanding a substantial amount of pixel-level annotated data in machine vision-based corrosion segmentation and assessment approaches. To tackle this issue, a novel weakly supervised method for corrosion segmentation and assessment in hydraulic steel gates is proposed, leveraging class labeling. The technique employs a class activation map to pinpoint regions containing corrosion seeds and to train a network to capture semantic affinity relations. Subsequently, the concept of region growing is adopted to propagate semantic information across the entire image. The average feature vector of the seed region serves as the corrosion feature, enabling precise segmentation of corroded areas and circumventing the laborious pixel-level annotation process. Additionally, a fine-grained corrosion classification network is established and trained using salt spray corrosion test data to accurately evaluate the corrosion severity. To validate the proposed method's accuracy, a dataset of steel gate corrosion images is curated based on real-world operational scenes. Experimental results demonstrate that, in practical scenarios, the segmentation method presented in this paper achieves a segmentation intersection ratio of 62.37% in corrosion, without pixel-level annotation. This performance closely approaches the performance of mainstream fully supervised methods. Additionally, the corrosion grade evaluation method proposed in this study achieves an accuracy of 95.77%.

1 Introduction

In recent years, the development of water conservancy projects has driven the widespread adoption of hydraulic steel gates in water resource management and flood control. However, environmental factors and extended usage have exacerbated the corrosion issue on the surface of hydraulic steel gates [1], thereby escalating a threat to their operational efficiency and remaining lifespan [2]. The increasing severity of surface corrosion on hydraulic steel gates due to environmental influences and prolonged usage poses a potential menace to their performance and longevity. Traditional methods for detecting corrosion on hydraulic steel gates primarily rely on manual inspections and visual observations, a practice that not only increases maintenance expenditures but also fosters inefficiencies and delays [3]. Additionally, the scope of manual inspection is constrained by the operational environment, particularly in challenging settings like high-altitude, underwater, or high-temperature conditions, where personnel well-being may be compromised, elevating the associated risks [4]. Machine vision-based techniques for corrosion assessment offer the merits of automation, heightened efficiency, and precision, establishing a compelling advantage in the realm of corrosion segmentation and evaluation concerning hydraulic steel gates [5].

Based on the characteristics of machine vision technology, several researchers have endeavored to employ machine vision technology to address challenges in corrosion segmentation and assessment. The study conducted by Bastian et al. involved designing a customized convolutional neural network tailored to classify pipeline images based

✉ Huadong Zhao
  zh_zzu0829@163.com

1 School of Mechanical and Power Engineering, Zhengzhou University, 450001 Zhengzhou, China

2 Intelligent Manufacturing Research Institute of Henan Province, 450001 Zhengzhou, China

on corrosion levels. Despite achieving a remarkable 98.8% accuracy in the binary classification problem of distinguishing the presence of corrosion, it did not encompass the classification of corrosion degree and localization of corroded areas [6]. Chen et al. utilized the Faster R-CNN neural network to pinpoint corrosion regions on a large crane structure captured by a UAV using a target detection algorithm. This approach facilitated real-time detection through communication between the UAV and a ground station [7]. Based on the above research, Yu et al. utilized the lightweight characteristics of the Yolov3-tiny model to create a corrosion target detection method based on mobile devices [8]. Atha and Jahanshahi from Purdue University evaluated the performance of U-Net, DeepLab, PSPNet, and RefineNet in semantic segmentation of corrosion using 600 finely labeled high-resolution corrosion images [9]. Liu et al. constructed a VGG19 and Faster R-CNN neural network to delineate the extent of surface and edge corrosion of large oceanic steel structures [10]. Wang devised a semantic segmentation network employing U-Net as the primary architecture, utilizing corrosion images of the Sutong Bridge steel structure. This culminated in the quantitative analysis of corrosion areas within individual images [11]. These studies accomplished target detection or pixel-level segmentation of corroded areas, yet both necessitated extensive, meticulously labeled image datasets. However, corrosion segmentation tasks linked to hydraulic steel gate scenarios encounter distinct challenges. First and foremost, in-service hydraulic equipment undergoes routine corrosion removal and anticorrosion paint applications [12], creating a dearth of large-scale image datasets illustrating the corrosion progression. Moreover, steel gates endure alternating wet and dry environments, persistent darkness, a variety of corrosion types, and intricate corrosion boundaries during the prolonged service. Capturing and annotating images within this context is a laborious, time-intensive endeavor, demanding specialized expertise and invariably incurring subjectivity and inconsistencies in actual execution. These factors make constructing a comprehensive, high-quality corrosion image dataset profoundly challenging [13]. Although many researchers have already laid the foundation for deep learning corrosion segmentation by manually annotating corrosion images, this approach suffers from challenges such as high cost of annotation, difficulty in data acquisition, and limitations in specific environments. This study aims to bridge this knowledge gap by proposing a method for corrosion segmentation and assessment of hydraulic steel gates without pixel-level annotations to address the challenges of existing methods.

Weakly supervised learning approaches offer succinct labeling information, offering an effective means to address the challenges of limited data volume and intricate data annotation in hydraulic steel gate corrosion segmentation. In contrast to fully supervised methods demanding substantial

labeling, weakly supervised semantic segmentation techniques typically leverage labeling information that is more readily accessible, such as bounding box labels[14], graffiti labels [15], and image-level labels [16, 17]. Among these, image-level labels are extensively adopted due to their simplicity in annotation and cost-effectiveness. However, image-level labels lack crucial object localization details in segmentation tasks, leading to suboptimal model performance. To remedy this, Class Activation Map (CAM) [18] has been employed. Ahn utilizes the rudimentary localization information acquired from CAM as a seed region. This technique generates a probability transfer matrix of pixels grounded in the semantic correlation among image pixels, which is subsequently diffused across the image based on the random walk principle. This process rectifies the imprecisions of CAM [19]. However, the random walk method is susceptible to noise and indistinct regions [20]. Considering that corrosion segmentation tasks involving hydraulic gates frequently encompass noise and blurred features, such as uneven illumination and water mark shadows, the application of this method in the hydraulic gate corrosion segmentation domain poses challenges.

In this study, a corrosion segmentation and assessment methodology for hydraulic steel gates without requiring pixel-level annotations is introduced. Initially, the corrosion images collected from the Heiquan Reservoir site were used as a dataset to train the corrosion binary classification network to generate a CAM for each corrosion image. Subsequently, a semantic affinity network is constructed by extracting highly confident regions from the CAM, which serves as supervisory information. This network is utilized to refine the corrosion seed regions, resulting in an accurate corrosion segmentation outcome. Next, employing the minimum bounding rectangle of the corrosion's connected domain, the corrosion area is isolated and input into a fine-grained corrosion classification network to assess the extent of corrosion. For the training of the fine-grained corrosion classification network, a dataset comprising images depicting the corrosion process is compiled through salt spray corrosion tests. These images facilitate the training of an efficient and precise corrosion classification network, tailored for grading the identified corrosion regions. The approach investigated in this paper achieves pixel-level segmentation of corrosion on hydraulic steel gates, simultaneously evaluating the degree of corrosion within the corroded regions. Our proposed method is not limited to the field of hydraulic steel gate corrosion assessment. By adopting a weakly supervised learning approach, we successfully cope with some of the difficulties in image annotation and provide a more cost-effective and efficient image annotation solution for other domains. This generally makes our approach more widely applicable and able to cope with the challenges of image labeling in different

environments and objects. Thus, our research not only fills the knowledge gap in the field of corrosion assessment of hydraulic steel gates but also provides an innovative approach to the field of image annotation. We believe that the popularization and application of this method will bring new possibilities for future research and practice.

The subsequent sections are organized as follows. In Sect. 2, we commence by outlining the overarching framework of the proposed methodology. Subsequently, we delve into the particulars of training the pixel affinity network through the utilization of class activation maps, alongside the elucidation of how fine segmentation of corroded areas is realized. Following this, we expand upon the bilinear corrosion-level assessment network, which is trained utilizing comprehensive corrosion image data derived from salt spray corrosion experiments. Moving on to Sect. 3, we substantiate the efficacy of the proposed approach using on-site corrosion image data from the Heiquan reservoir gate located in Xining City. Concluding in Sect. 4, we offer an overview encompassing the strengths and limitations of our proposed methodology, while also outlining avenues for future optimization.

2 Methodology

The methodology of this study comprises three primary components. The first component involves CNN1 and the Grad-CAM algorithm. Utilizing annotated images of metal gates with labels corrosion and non-corrosion, a classification network is trained to determine the presence of corrosion in the images. Subsequently, the CAM (Class Activation Map) technique is applied to extract coarse segmentation results. The second component employs CNN2 as a feature extractor to capture high-dimensional feature information. Using the coarse segmentation results obtained from the first component as labels, a corrosion segmentation

model based on pixel affinity is trained. The third component involves simulating real working conditions through a salt spray corrosion experiment to generate a dataset of corrosion grade images. A bilinear convolutional neural network model is trained on this dataset to evaluate the grade of corrosion in the segmentation results. The overarching structure is depicted in Fig. 1.

2.1 Weakly supervised corrosion segmentation method

2.1.1 Computation of class activation maps

Grad-CAM is extensively employed in weakly supervised semantic segmentation. It efficiently retrieves the position information of the target in the image using only image category labels [21]. Grad-CAM represents a refinement of the conventional CAM approach. CAM is a method used to interpret predictions of deep learning models in image classification tasks. The core idea involves taking the feature maps of the last convolutional layer, applying global average pooling to obtain a weight vector for each channel, and then multiplying this vector with the weights of the classification layer. This process results in weighted sums for each spatial position, which are then applied to the original image to generate a class activation map. The formulation governing the class activation map is given in Eq. 1:

$$M_c(x, y) = \sum_{k=1}^{n} w_k^c f_k(x, y). \tag{1}$$

The feature map with n channels is globally average pooled to generate a feature vector comprising n elements, where w_k signifies the weight of the k th element within this vector concerning the classification task. Denoting $f_k(x, y)$ as the activation value at position (x, y) within the k th channel of the feature map, the region of interest
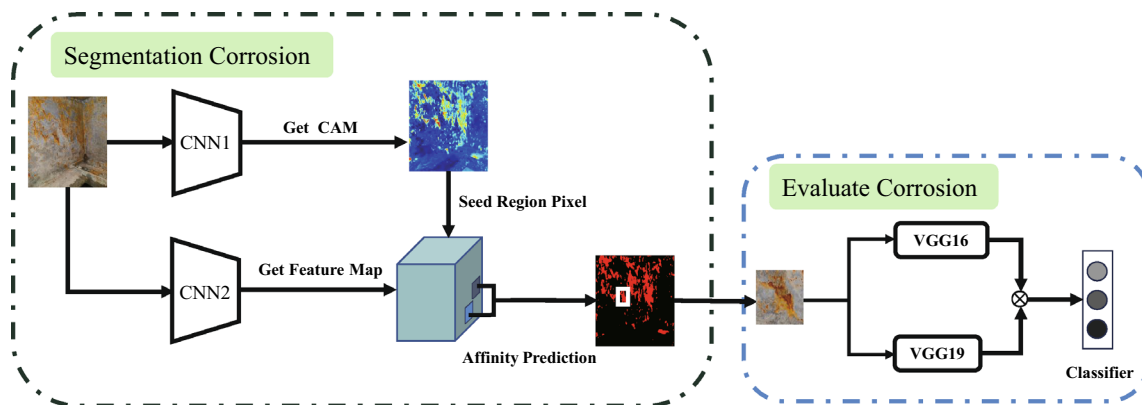

Fig. 1 Overall structure of the corrosion segmentation and assessment network

pertinent to corrosion classification can be derived from the image using Eq. 1. However, this approach presents certain limitations. Specifically, due to the application of global average pooling, the computation of only the CAM corresponding to the last feature map is feasible. Moreover, obtaining the corresponding CAM mandates the retraining of the neural network.

Grad-CAM, renowned for its enhanced generalization capability, has gained widespread adoption. Its schematic representation is depicted in Fig. 2. The foundational concept behind Grad-CAM remains consistent with that of CAM. It involves calculating a weighted feature map by employing weights aligned with each channel of the feature map. However, the distinction lies in the weight computation process. While CAM entails retraining to acquire the weights through the replacement of the fully connected layer with the Global Average Pooling (GAP) layer, Grad-CAM employs gradients of the global average to compute the weights. This differentiation is evident in Eq. 2, where the gradient is derived through the automatic differentiation mechanism inherent in the backpropagation algorithm:

$$\beta_k^c = \frac{1}{i \times j} \sum_i \sum_j \frac{\partial y^c}{\partial A_{ij}^k}. \tag{2}$$

The y^c is the output of the classification model, which represents the probability that the image is of the target category. By applying the backpropagation function, we can compute the gradient of each element A_{ij}^k in the target layer's feature map A with respect to the output y^c. Here, i and j denote the spatial coordinates within the feature map, specifying the position of the element in question. Subsequently, the resulting gradient matrix, bearing the same dimensions as the feature map, undergoes global average pooling to derive the importance weight β_k^c instead

of the weigh w_k^c employed in the CAM method. This substitution not only enables the acquisition of class activation maps for distinct layers within the neural network but also effectively circumvents the necessity of network retraining. Following the successful implementation of the Grad-CAM algorithm, it was subjected to testing using corrosion images of gates under field conditions. We employ Vgg16 as the foundation for our classification network, referred to as CNN1 in Fig. 1. The structure of CNN1 is illustrated in Fig. 3.

The CAM is generated using the convolutional layer before the pooling layer in the model to obtain five CAMs with dimensions 448*448, 224*224, 112*112, 56*56, and 28*28. We employ a stepwise up-sampling and summation strategy to progressively enhance the resolution of the CAM. Beginning with a 28×28 feature map, we up-sample it to 56×56 and add it to the original 56×56 feature map, generating a refined map. This process iterates, sequentially up-sampling and summing, until a final 448×448 CAM is achieved. This method effectively integrates information from multiple resolutions, providing a detailed and comprehensive visualization of class-specific activations. The outcomes of these tests are showcased in Fig. 4b and c.

2.1.2 Pixel affinity label extraction

The class activation map typically captures responses primarily in the most discriminative region of the object intended for segmentation. However, this approach is plagued by issues such as unclear boundaries and suboptimal segmentation accuracy. Moreover, it fails to yield accurate positional information of the target. Following the computation of the class activation map, regions exhibiting higher confidence are extracted from the original CAM map using Eqs. 3 and 4:
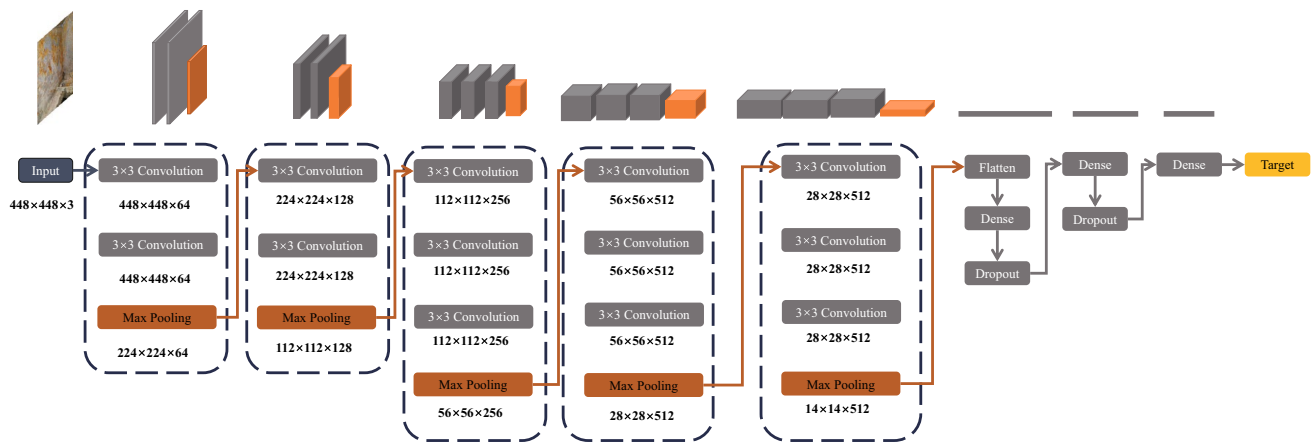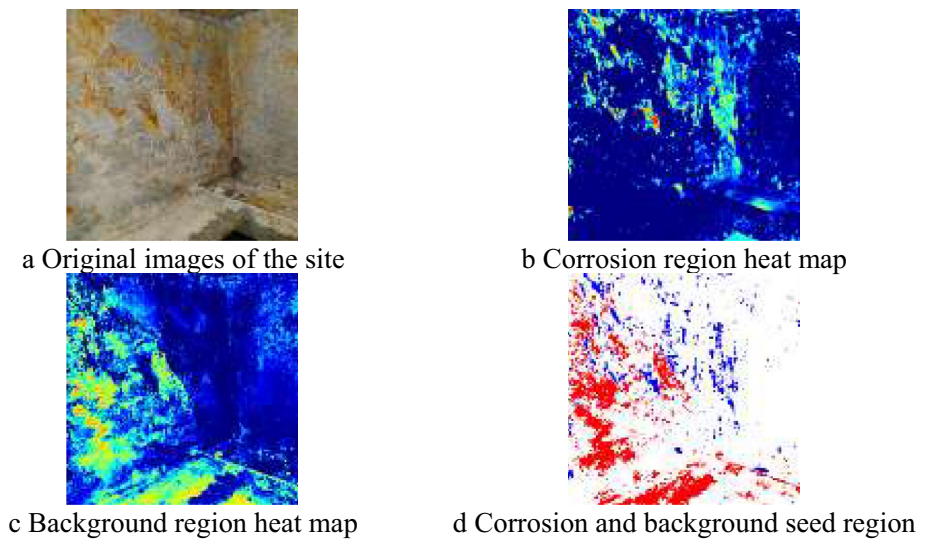
Fig. 2 Schematic diagram of Grad-CAM
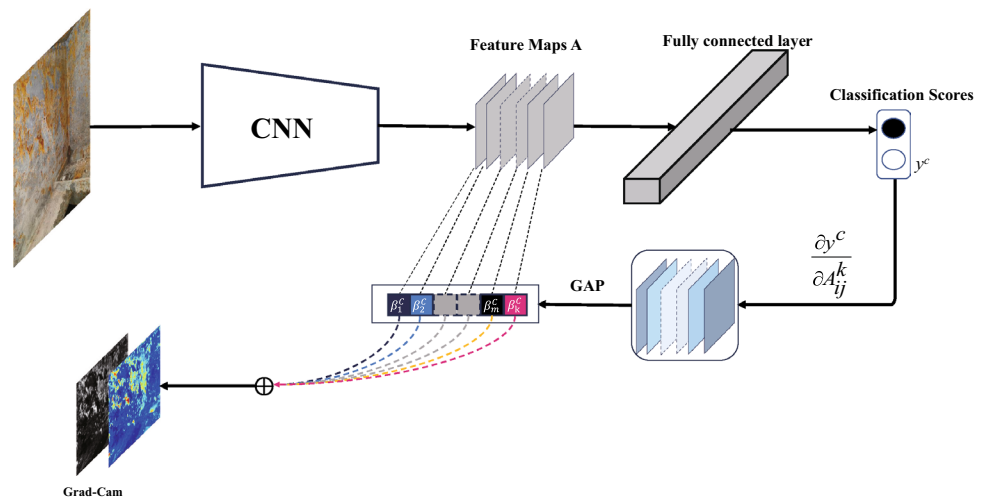
Fig. 3 VGG16 network structure

Fig. 4 Class activation map with confidence region extraction

a Original images of the site

b Corrosion region heat map

c Background region heat map

d Corrosion and background seed region

$$M_{bg}(x,y) = \left\{ 1 - \max_{c \in C} M_c(x,y) \right\}^{\alpha}, \qquad (3)$$

$$M_{rust}(x,y) = \left\{ M_c(x,y) \right\}^{\alpha}, \qquad (4)$$

Where α represents the parameter for refining the confidence score, M_{bg} denotes the activation score of the background class, and M_{rust} denotes the activation score of the corrosion class, adjusted by α. In this study, we set the parameter α to 32. Further discussion on the selection of parameter α is included in Sect. 3.2. If the corrosion confidence score after α-square reduction remains larger than the original background score, the region is considered to be a corrosion confidence region. The same principle is used for the selection of the background confidence region. By employing the abovementioned formula, the corrosion and non-corrosion regions can be confidently isolated from the image. In Fig. 4d the red region represents the

high-confidence corrosion region, the blue region represents the high-confidence background region, and the white area is an area that cannot be distinguished based on CAM as to whether it is corroded or not. After the above method, the corrosion and non-corrosion regions with high confidence can be effectively extracted, thus improving the segmentation accuracy.

After obtaining the high-confidence corrosion regions and high-confidence non-corrosion regions shown in Fig. 4d, We need to create a bar of data like [(x_1, y_1), (x_2, y_2), 1] from the image, where (x_1, y_1), (x_2, y_2) are points randomly taken from high-confidence corrosion regions and high-confidence non-corrosion regions. The label is 0 if (x_1, y_1) is located in high-confidence corrosion regions and (x_2, y_2) is located in non-corrosion regions. The label is 1 if both (x_1, y_1) and (x_2, y_2) are located in high-confidence corrosion regions, and the label is also 1 if both (x_1, y_1) and (x_2, y_2) are located in non-corrosion regions.

2.1.3 Training of pixel affinity network

Following the completion of label extraction, the subsequent step involves constructing and training the pixel affinity network. The structure of the pixel affinity network is illustrated in Fig. 5, serving as a detailed explanation of "CNN2" and the process of "Get Feature Map" in Fig. 1.

The input image is first passed through a feature extractor to obtain a feature map. In the forward propagation process, 1×1 convolutional layers are added at different locations to extract the features. Subsequently, these features are merged to form a feature representation that covers more comprehensive information. The resulting feature map is then up-sampled to match the dimensions of the original input image. The corroded and non-corroded confidence regions obtained in Sect. 2.1.3 are utilized as labels for the feature vectors at the corresponding locations in the convolutional feature map. A model for distinguishing whether the feature vectors belong to the same class is then trained. The concept of the region growing algorithm is harnessed through the utilization of the pixel affinity network to refine the corrosion seed region across the entire image. The image segmentation algorithm rooted in region growing commences from a seed point or seed region, progressively determining whether adjacent pixels pertain to the same class or not [22]. We designed a cosine similarity-based loss function to measure the similarity between the predicted and target feature vectors, as shown in Eq. 5:

$$L = 1 - \frac{\vec{a} \cdot \vec{b}}{\vec{a} \times \vec{b}}, \tag{5}$$

where a and b are two vectors, $\vec{a}$ and $\vec{b}$ are the modulus of the two vectors respectively, and $\vec{a} \cdot \vec{b}$ is the dot product of the two vectors. This loss function has a value between 0 and 2, 0 when the two vectors are identical and 2 when the two vectors are completely different. This loss function measures the similarity between the two input vectors. We get the local labels of the image by means of high-confidence regions,

compute the loss only on the labeled pixels, and then use this loss to update the parameters of the model. By this method, our model can perform effective image segmentation with only category labels.

The segmentation model in this paper draws on the principle of region growing algorithm in the application phase. We use the first model to get the confidence region coordinates and the second model to get the feature map. Then, the average feature vectors corresponding to the confidence region coordinates are computed on the feature map as the seed, and it is judged whether the seed belongs to the same class as each feature vector of the non-confidence region, so as to complete the segmentation task on the whole image. This process establishes pixel affinity across the entire image, propelled by the propagation of classification results, thereby engendering the expansion of the CAM seed region. The progression of this process is illustrated in Fig. 6. In the corrosion area correction process, the red color represents the corrosion area and black color represents the background area.

2.2 Construction of the corrosion grade assessment method

2.2.1 Corrosion classification dataset

After completing the segmentation of corrosion areas, a grade valuation must be conducted on these areas. However, due to the difficulty in obtaining image datasets of corrosion grades in engineering sites, salt spray corrosion experiments were used to simulate the actual working environment of gates, and image datasets of corrosion grade were obtained through experiments. This method provides us with a feasible way to evaluate the degree of corrosion. The salt spray corrosion test serves as a prevalent experimental technique for emulating the corrosion process of steel. In this section, we delve into the categorization of corroded steel plates, leveraging image information derived from the corrosion process as observed in salt spray corrosion tests.

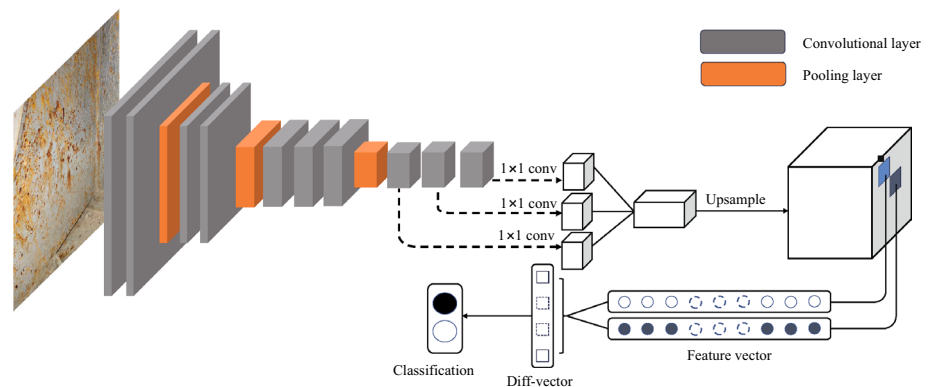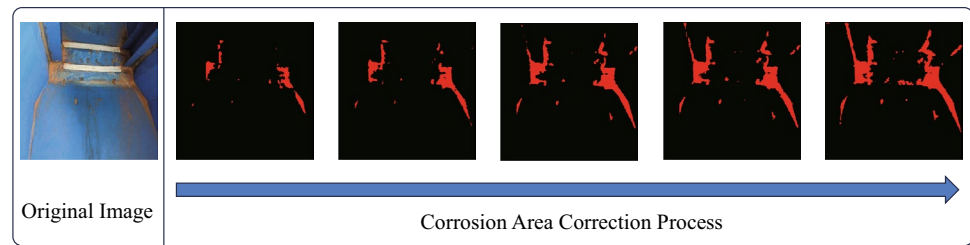**Fig. 5** Structure of the pixel affinity network

Fig.6 Example of seed region correction

Original Image

Corrosion Area Correction Process

Aligned with the gate's design specifications and on-site research of operational conditions, the favored material for the steel gate is Q235. As a result, this study employs 20 samples measuring 200× 200 mm, all composed of Q235. This selection not only aligns with the structural dimensions of the salt spray chamber but also enhances the scope of image capture. The parameters for the accelerated salt spray corrosion experiment adhere to the guidelines outlined by Wang et al.[23]. The test solution is configured as a neutral solution containing a 5% mass fraction of sodium chloride, with environmental parameters for the salt spray test chamber established accordingly. The precise parameter settings are detailed in Table 1.

Initially, the salt spray corrosion platform was commissioned to validate its operational integrity. Subsequently, the designated steel plate specimen for testing was positioned within the salt spray test chamber, as depicted in Fig. 8. A preconfigured 5% mass fraction sodium chloride solution was introduced into the brine tank. Following the experimental protocol, environmental parameters were configured in the control panel of the salt spray test chamber. Subsequently, the lid of the salt spray test chamber was sealed to isolate the external environment from the internal operational environment, thus initiating the acceleration of the corrosion process on the surface of the steel plate specimen. Ultimately, upon reaching the predetermined experimental duration, the steel plate specimens were individually removed and placed in a dry location. Subsequently, image acquisition was conducted after a 12-h interval.

The aforementioned experiments emulate the corrosion progression of steel plate specimens within a controlled environment and yield a series of image data. This method not only ensures the scientific validity and reproducibility

of the experiment but also provides image samples depicting varying levels of corrosion and morphologies, thereby enhancing the diversity of the dataset.

2.2.2 Image classification

Upon completing the acquisition of corrosion images, it becomes imperative to classify the corrosion levels based on the established standards, serving as the guidance for corrosion-level classification. In alignment with the Chinese national standard GB/T 8923.1–2011[24], a total of 785 corrosion images derived from the salt spray test is categorized into three distinct classes, forming the basis for training the corrosion class classification network. The evolution of corrosion is visualized through image slices, as illustrated in Fig. 7. The descriptions of the corrosion classes in GB/T 8923.1–2011 are qualitative, and example pictures are given for categorization. In the first 6 days of corrosion images as Grade A, the original surface of the oxide layer gradually falls off. In the 7–16 days of corrosion images as Grade B, the stage of corrosion color deepening, the original surface has been completely removed, and hence the emergence of local corrosion pits. In the 17–26 days of images as Grade C, the previous stage of corrosion layer began to fall off, and continues to corrode downward, and local pits become larger. After this analysis, we categorized all collected images into corrosion grades, aligning with the time intervals illustrated in Fig. 8.

2.2.3 Corrosion classification networks

Within the realm of corrosion image classification, the disparity in inter-class feature distinctions among varying corrosion classes is relatively slight, whereas intra-class feature distinctions are substantial. This necessitates that the classification network is adept at discerning nuanced feature differences across different corrosion class images. Achieving fine-grained feature learning entails an elevated capacity for perceptual acuity and representation within the network. To address this challenge effectively, the Bilinear Convolutional Neural Network (BCNN) proves instrumental [25]. BCNN conducts feature extraction from images through two branch networks, subsequently employing bilinear pooling to amalgamate features from these branches. This component

Table 1 Salt spray test chamber experimental parameter settings

Parameters	Retrieve a value
Salt spray acidity and alkalinity	Neutral
Test chamber temperature	35 °C
Saturation temperature	47 °C
Test time	12 h
Interval time	1 h

Fig. 7 Schematic diagram of salt spray accelerated corrosion test

a Uncorroded steel plate

b Steel plates corroded in salt spray environment

Grade A

0-6days

Grade B
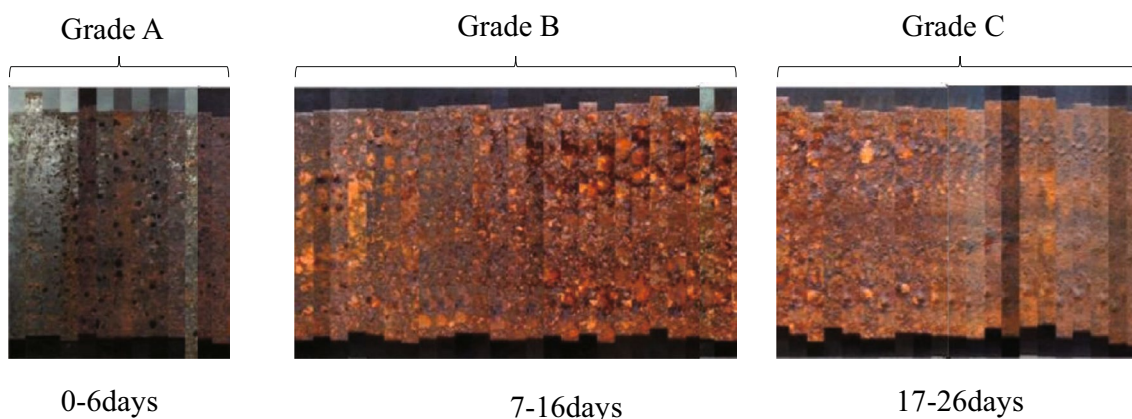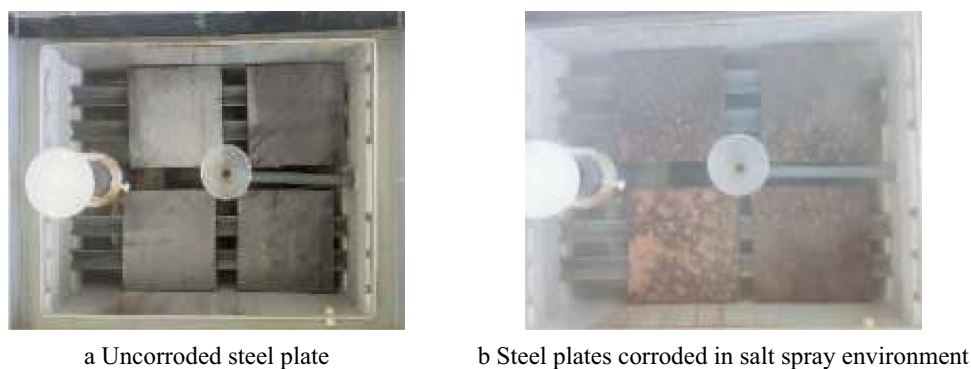

7-16days

Grade C

17-26days

Fig. 8 Corrosion development process of steel in salt spray test

introduces a corrosion image classification model tailored for corrosion-level assessment by amalgamating the attention mechanism and bilinear pooling. The architectural configuration of the network is shown in Fig. 9.

The model encompasses three primary components: bicolor space image input, bilinear feature extraction network, and bilinear feature vector classifier. On the input front, informed by Khayatazad's research [26], the saturation component within the HSV color space is responsive to overall corrosion color variations, while the blue component within the RGB spectrum responds to localized corrosion luminance alterations. Accordingly, a bilinear convolutional neural network is harnessed in this study, aiming to enhance image classification task efficacy by capturing global features and inter-feature interaction information. The network encompasses two separate neural networks for extracting images in the HSV and RGB color spaces. Subsequently, a bilinear pooling operation fosters a dot product computation between the two feature maps, capturing non-linear inter-feature interactions. These bilinear pooled features then channel into a fully connected layer, facilitating further integration and fine-grained feature learning. Ultimately, corrosion classification outcomes are furnished via an output layer, affording precise corrosion area class assessments.

Figure 10 shows in detail the logical relationship of the proposed method in this article. The input of the corrosion segmentation model is the image, and the output is the segmentation result of the corrosion area in the image. The corrosion-level evaluation model takes the output of the corrosion segmentation model as input information to output the corrosion level of the corrosion area. The corrosion grade assessment section is used after the corrosion segmentation is completed. In the output graph, the red area represents Grade C corrosion, the blue area represents Grade B corrosion, and the green area indicates Grade A corrosion.

3 Algorithm validation for field conditions

3.1 Corrosion segmentation dataset

The experimental dataset employed for the corrosion segmentation analysis primarily originates from the Heiquan Reservoir and its scaled-down hydraulic experimental platform model situated in Xining City, Qinghai, China, as shown in Fig. 11. The focal components within the experimental setting encompass the arc gate of the water discharge outlet and the planar maintenance gate. Notably, the arc

gate bears a weight of approximately 16.7 tons, featuring dimensions of 3.0 m in width, 4.0 m in height, and a gate panel length of 6.0 m. Functioning under a design head of 65 m, the cumulative water pressure can ascend to 6500 KN. Likewise, the planar gate, weighing around 15.5 tons, boasts dimensions of 3.0 m by 4.0 m, while accommodating

a total water pressure of 8560 KN under a design head of 57.5 m. In a comprehensive endeavor to grasp the intricacies of hydraulic steel gate corrosion, this study was conducted within the genuine operational environment of the Heiquan Reservoir and the corresponding experimental platform, utilizing image acquisition protocols. The experimental
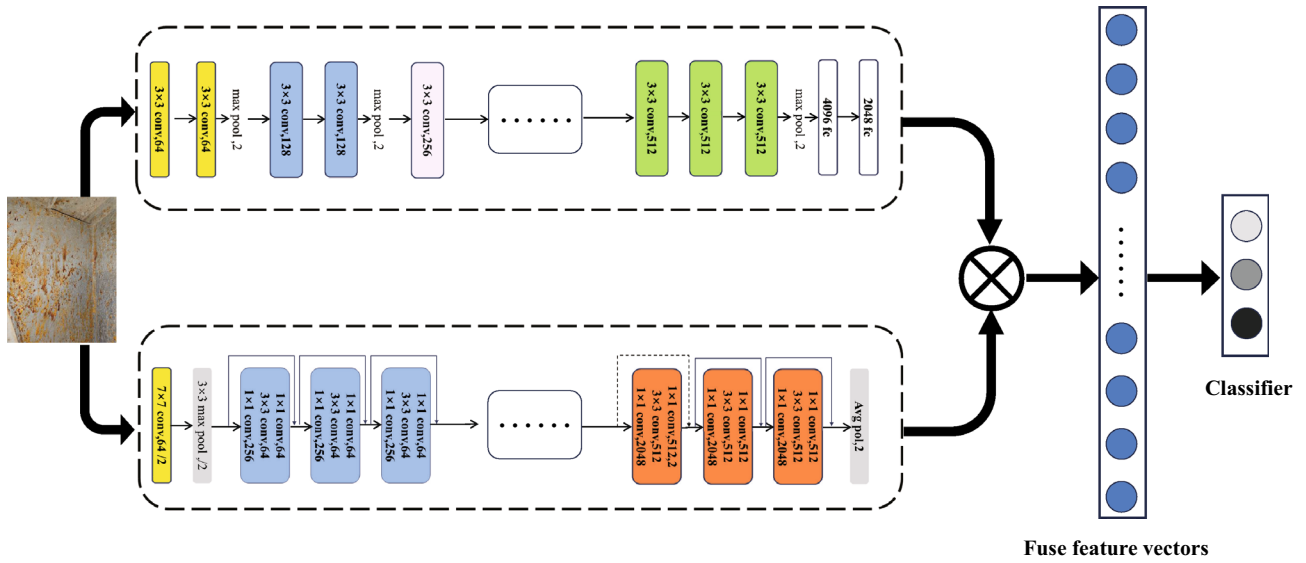

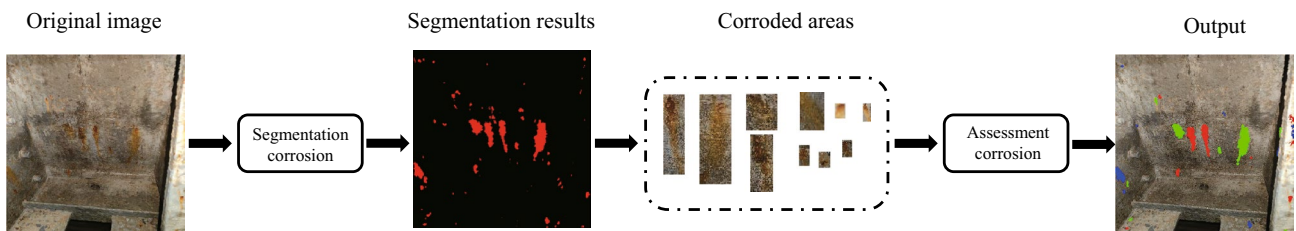**Fig. 9** Fine-grained corrosion image classification model architecture

Fig. 10 Corrosion area segmentation and grade assessment

Fig. 11 Field gate and experimental model of Heiquan Reservoir

a Heiquan Reservoir site

b Experimental model of the gate

platform takes the form of a gate model scaled at a ratio of 1:10, embodying a gate structure akin to the in-service gate. A segment of the image data is procured from this downscaled gate model. Within the dataset, a sum total of 140 corrosion images is encompassed. Given the relatively modest size of the original dataset, precautionary measures are undertaken to avert network overfitting. To this end, data augmentation techniques are employed in this study to amplify the original dataset. Post-capturing, the images are subjected to horizontal and vertical flipping, followed by illumination adjustments. Consequently, a cumulative tally of 607 augmented images are generated.

3.2 Corrosion segmentation test results

In order to verify the effect of the corroded segmentation model proposed in this paper, the weakly supervised segmentation method based on the research in this paper is compared with the fully supervised semantic segmentation networks U-Net and DeepLabV3, and four evaluation metrics are mainly used, Precision (Pre), Recall (Rec), Accuracy (Acc), and Intersection over Union (IoU). The evaluation metrics for corrosion segmentation are as follows: True Positives(T_P), where corroded regions were accurately identified; True Negatives(T_N), correctly pinpointing non-corroded areas; False Positives(F_P), indicating misclassification of some non-corroded regions; and False Negatives(F_N), highlighting instances where corroded areas were mistakenly predicted as non-corroded. The formulae are as follows:

$$P_{re} = \frac{T_P}{T_P + F_P} \times 100, \tag{6}$$

$$R_{ec} = \frac{T_P}{T_P + F_N} \times 100, \tag{7}$$

$$A_{cc} = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \times 100, \tag{8}$$

$$I_{oU} = \frac{T_P}{T_P + F_P + F_N} \times 100. \tag{9}$$

The outcomes of the algorithm testing are shown in Table 2. For the training of U-Net and DeepLabV3 + using fully supervised data, the source is the Heiquan Reservoir field dataset. To ensure a fair and comprehensive evaluation of the proposed approach, all models underwent training and testing within environments consisting of Windows 10, Python 3.6, and Pytorch 1.12. The computational acceleration was facilitated by an RTX3060 graphics card with 6 GB of RAM. In the corrosion segmentation part

mainly contains two neural networks, when training the first classification network for generating Grad-CAM, VGG16 is used as the backbone, and the stochastic gradient descent method is used for network optimization, the initial learning rate is set to 0.1, and polynomial decay is used to reduce the learning rate for each iteration with a weight decay coefficient of 0.0005, and 85 epochs are iterated during the training process. In training the second pixel affinity network for extracting high-dimensional semantic features, the same backbone and optimization approach as the first stage network is used. The default value of α in Eq. 2 and Eq. 3 has a default value of 32, which was changed to 16, and 64 for validation in the experiments, respectively.

To validate the efficacy of the network architecture proposed in this paper, a set of corrosion images captured from gate conditions was chosen for segmentation testing. The outcomes of this test are illustrated in Fig. 12. Figure 12a and 10b show partial portions of an in-service gate leaf, respectively. Figure 12c and d localizes the gate experimental platform. In the individual algorithm segmentation result plots in Fig. 12, black represents the background that was correctly segmented (T_N), red represents the corrosion that was correctly segmented (T_P), green represents the background that was misidentified as corrosion (F_P), and white represents the corrosion that was misidentified as background (F_N).

Based on the aforementioned data comparison and visual examples, as shown in Fig. 12a and b, it becomes evident that both the approach proposed in this paper and U-Net exhibit commendable resistance to interference. The encoding–decoding structure of U-Net establishes a link between high-level features and low-level details, thereby enhancing its anti-interference ability. The technique proposed in this paper uses a semantic affinity network based on seed regions. It can be seen in all the segmentation results in Fig. 12 that the proposed method in this paper has the least number of T_P regions in the segmentation results and the method in this paper has a better ability to recognize the regions that look like corrosion but are not actually corroded. Nonetheless, a comprehensive comparison of the segmentations in Fig. 12c and d indicates that the proposed method in this paper is more susceptible to the influence of the seed region. In Fig. 12d, instances arise where dark corrosion is incorrectly identified as background due to the stark disparities in the features of distinct corrosion regions.

Table 2 Quantitative evaluation of the effect of each model

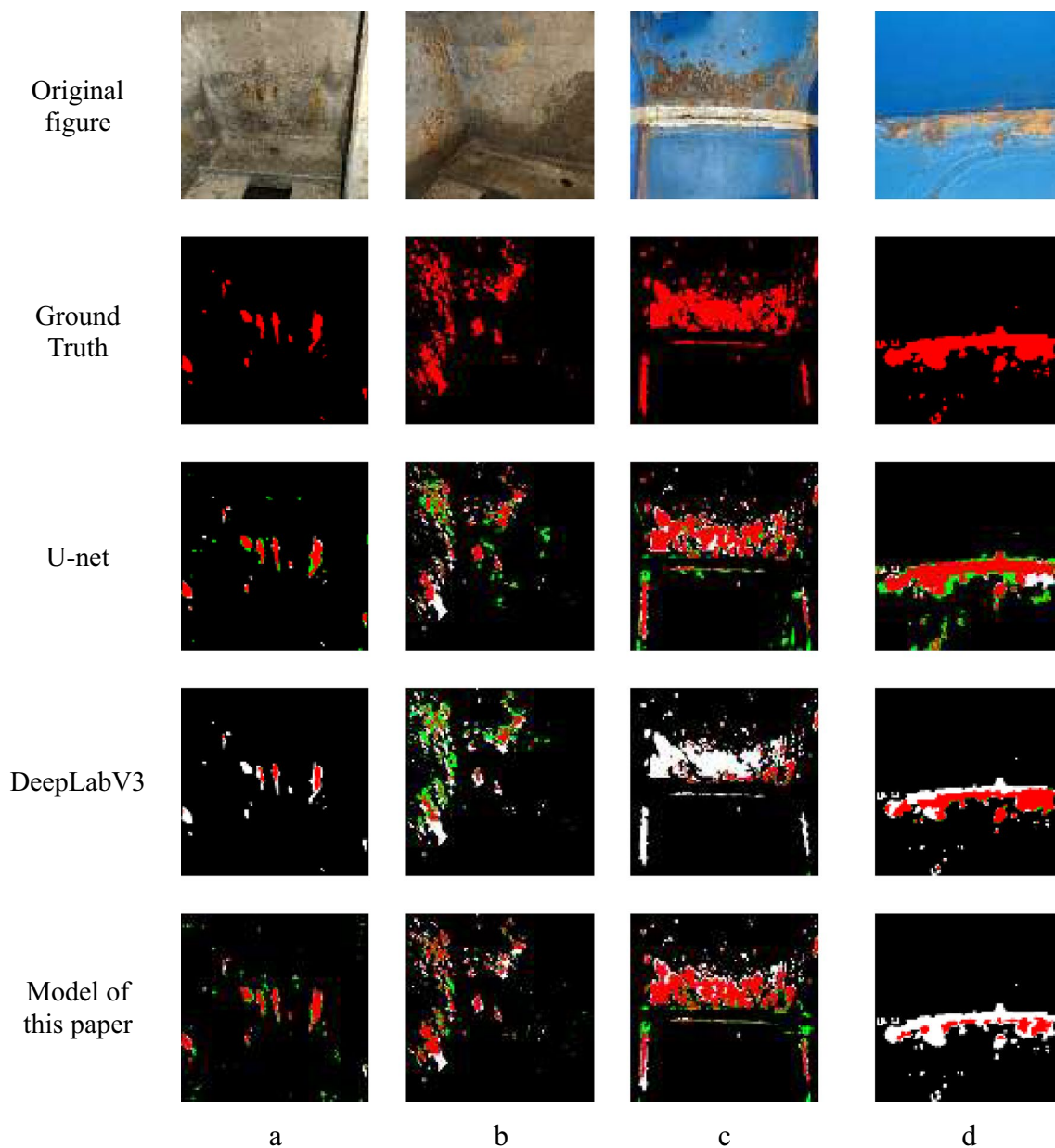|	Accuracy (%)	Recall rate (%)	Precision (%)	IOU(%)
Ours	85.78	69.56	91.65	62.37
U-Net	78.91	75.91	97.22	63.11
DeepLabV3	78.43	46.58	89.10	38.94

Fig. 12 Project site image corrosion segmentation results

In contrast, Fig. 12c illustrates that regions with slight feature differences in corrosion are more accurately classified. Furthermore, DeepLabV3 + incorporates features like dilation convolution that can lead to the blurring of boundary information, thus constraining its ability to perceive corrosion boundaries effectively. As a result, during testing, DeepLabV3 + exhibits a generally conservative prediction of corrosion, emphasizing caution when employing it for corrosion segmentation.

In Eq. 3 and Eq. 4, we used the parameter α. Here, we discuss the effect of parameter α on the segmentation results. The results are shown in Fig. 13. From the figure, we can

see that the best segmentation results can be achieved when α is defined as 32. The reason is that α can only affect the selection of the confidence region; in the refinement segmentation process we use the average feature vector of the confidence region to represent the confidence region, so when α becomes larger the smaller the confidence region, the smaller the impact on the average feature vector.

3.3 Corrosion grade classification test results

To validate the efficacy of the corrosion grade classification method, ResNet50, MobileNetV2, VGG16, and VGG19
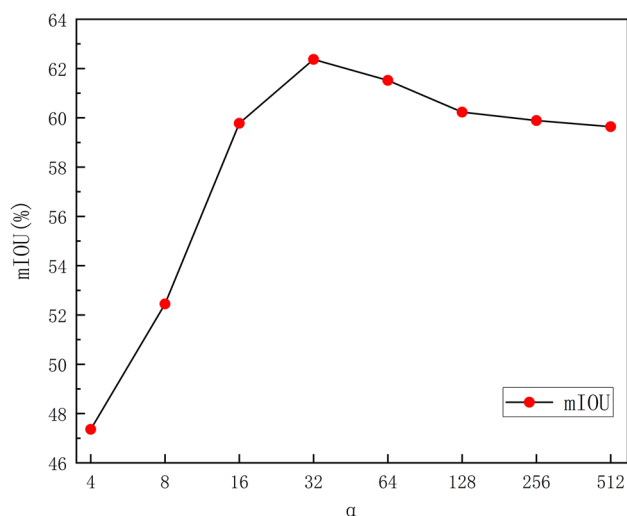
Fig. 13 Effect of parameter α on IOU

were chosen as comparative testing approaches. These networks were trained using the image dataset derived from the salt spray corrosion experiments. Subsequently, the trained models were evaluated using images of corrosion areas acquired through segmentation under the field conditions of the Heiquan Reservoir. The loss function of the model is a multicategorical cross-entropy function. The Adam optimizer is used to train the network model and update the weights. The exponential decay rate of the first-order moment estimation is set to 0.9, the exponential decay rate of the second-order moment estimation is set to 0.999, and the weight decay coefficient is set to 0.000001. The batch size of the dataset is set to 16, and the number of training rounds is set to 80 epochs. In the later stage, when the network is closer to the optimal solution, the loss function converges slower and needs to be adjusted with a smaller learning rate in a longer number of iterations. Therefore, in this paper, we use the MultiStepLR dynamic learning rate adjustment strategy to realize the automatic adjustment of the learning rate, set the initial learning rate to 0.001, and reduce the learning rate to 20% of the original learning rate at the 5, 10, 30, 40, and 70th epochs of training. The mean values of the evaluation metrics are presented in Table 3.

The outcomes demonstrate that VGG19, VGG16, and ResNet50 yield superior classification test accuracies, with VGG19 achieving the highest accuracy of up to 89.17%. Conversely, MobileNetV2 is deemed unsuitable for corrosion grade classification tasks. This discrepancy arises from the fact that both VGG and ResNet-50 possess relatively deep network architectures, incorporating multiple convolutional and pooling layers. Deep networks typically deliver enhanced performance when dealing with diverse scales, textures, and shapes. In contrast, MobileNetV2

exhibits a shallower structure and lacks the capacity for robust feature extraction, particularly when intricate texture and shape features are involved.

Drawing from the outcomes of the individual neural network experiments, this study opts for VGG19, VGG16, and ResNet50 as the foundational models for constructing a bilinear classification network. The objective is to enhance the neural network's proficiency in extracting intricate corrosion details. The test outcomes, illustrated in Table 4, demonstrate that the bilinear network combining VGG19 and VGG16 exhibits superior feature extraction capability. Consequently, the test accuracy experiences a notable enhancement in comparison to that of the single neural network, achieving a remarkable accuracy of 95.77%.

4 Conclusion

1) In the corrosion region segmentation aspect, this paper presents a method that achieves corrosion segmentation and evaluation in field condition images of gates under the constraint of image category labeling only. This method effectively addresses the issue of insufficient high-quality labeled data in the gate corrosion detection. Experimental results demonstrate the competitiveness of this paper's approach with fully supervised U-Net and DeepLabV3 + . However, it is important to note that the pixel-level labeling of corrosion regions can have some subjective errors, making accurate labeling of all corrosion regions in an image challenging. The test results on actual working conditions depicted in Fig. 9 indicate that this method possesses better corrosion region identification capabilities.

2) Regarding corrosion degree assessment, this paper conducts comparison experiments involving single CNN and BCNN. VGG16, VGG19, ResNet50, and MobileNetV2 are compared for their classification

Table 3 Comparison of different single neural network model metrics

Backbone	Precision	Recall	F1-score	Accuracy
ResNet50	0.8430	0.7656	0.7551	0.7656
MobileNetV2	0.7509	0.6984	0.6636	0.6984
VGG16	0.8956	0.8906	0.8907	0.8906
VGG19	0.9077	0.8906	0.8893	0.8917

Table 4 Comparison of bilinear neural network model metrics

Backbone	Precision	Recall	F1-score	Accuracy
VGG19&VGG16	0.9601	0.9577	0.9575	0.9577
VGG19&esnet50	0.9279	0.9218	0.9224	0.9218

effects. Results reveal that MobileNetV2 yields the least effective performance, achieving only a converged test accuracy of 67.3%. Among the remaining three models, their test accuracy is approximately 90%, with VGG19 achieving the highest accuracy, rendering the use of MobileNetV2 in a bilinear network unfavorable. Through the combination of different CNN structures, it is concluded that the VGG19 + VGG16 combination proves more effective, attaining a classification accuracy of 96.54%. This offers a dependable solution for assessing the degree of corrosion in hydraulic steel gates.

3) The segmentation algorithm proposed in this study utilizes a region growing approach. While steel gate corrosion is a gradual process with relatively lenient real-time requirements, the algorithm's slower speed poses a limitation, especially when dealing with large-scale images. Although real-time processing is not a primary concern in engineering scenarios, the algorithm's speed remains an area for optimization, particularly in situations where a significant number of images may need to be processed.

Data availability The data that have been used are confidential.

Declarations

Conflict of interests 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.

References

1. Verma C, Ebenso EE, Quraishi M, Hussain CM (2021) Recent developments in sustainable corrosion inhibitors: design, performance and industrial scale applications. Mater Adv 2:3806–3850. https://doi.org/10.1039/D0MA00681E

2. Anami K, Ishii N, Knisely CW (2012) Added mass and wave radiation damping for flow-induced rotational vibrations of skin-plates of hydraulic gates. J Fluids Struct 35:213–228. https://doi.org/10.1016/j.jfluidstructs.2012.07.008

3. Agdas D, Rice JA, Martinez JR, Lasa IR (2016) Comparison of visual inspection and structural-health monitoring as bridge condition assessment methods. J Perform Constr Facil 30:04015049. https://doi.org/10.1061/(ASCE)CF.1943-5509.0000802

4. Ahuja SK, Shukla MK (2018) A survey of computer vision based corrosion detection approaches. Inform Commun Technol Intell Syst (ICTIS 2017) 2:55–63. https://doi.org/10.1007/978-3-319-63645-0_6

5. Nash WT, Powell C, Drummond T, Birbilis N (2020) Automated corrosion detection using crowdsourced training for deep learning. Corrosion 76:135–141. https://doi.org/10.5006/3397

6. Bastian BT, Ranjith N, KumarJiji S (2019) Visual inspection and characterization of external corrosion in pipelines using deep neural network. NDT E Intern: Indep Nondestruct Test Eval 107:102134. https://doi.org/10.1016/j.ndteint.2019.102134

7. Chen Q, Wen X, Lu S and Sun D (2019) Corrosion detection for large steel structure base on uav integrated with image processing system. In 1. (IOP Publishing). pp. 012020. https://doi.org/10.1088/1757-899X/608/1/012020.

8. Yu L, Yang E, Luo C, Ren P (2021) AMCD: an accurate deep learning-based metallic corrosion detector for MAV-based real-time visual inspection. J Ambient Intelli Humaniz Comput 14:8087–8098. https://doi.org/10.1007/s12652-021-03580-4

9. Atha DJ, Jahanshahi MR (2018) Evaluation of deep learning approaches based on convolutional neural networks for corrosion detection. Struct Health Monit 17:1110–1128. https://doi.org/10.1177/1475921717737051

10. Liu L, Tan E, Yin XJ, Zhen Y and Cai ZQ (2019). Deep learning for coating condition assessment with active perception. pp. 75–80. https://doi.org/10.1145/3341069.3342966

11. Wang D, Peng B, Pan Y, Chen A (2018) Segmentation and quantitative analysis of corrosion images based on deep neural networks. J. South China Univ. Technol. (Nat. Sci. Ed.) 46:127–133. https://doi.org/10.3969/j.issn.1000-565X.2018.12.015

12. Baboian R (2005) Corrosion tests and standards: application and interpretation (ASTM international). https://doi.org/10.1007/s00371-023-02846-6

13. Zhu T, Zhu S, Zheng T, Ding H, Song W, Li C (2023) HEU-Net: hybrid attention residual block-based network with external skip connections for metal corrosion semantic segmentation. The Visual Comput 40:1273–1287. https://doi.org/10.1007/s00371-023-02846-6

14. Dai J, He K and Sun J (2015). Boxsup: Exploiting bounding boxes to supervise convolutional networks for semantic segmentation. pp. 1635–1643. https://doi.org/10.1109/ICCV.2015.191.

15. Lin D, Dai J, Jia J, He K and Sun J (2016). Scribblesup: Scribble-supervised convolutional networks for semantic segmentation. pp. 3159–3167. https://doi.org/10.1109/CVPR.2016.344.

16. Papandreou G, Chen LC, Murphy KP and Yuille AL (2015). Weakly-and semi-supervised learning of a deep convolutional network for semantic image segmentation. pp. 1742–1750. https://doi.org/10.1109/ICCV.2015.203.

17. Jing L, Chen Y, Tian Y (2019) Coarse-to-fine semantic segmentation from image-level labels. IEEE Trans Image Process 29:225–236. https://doi.org/10.1109/TIP.2019.2926748

18. Zhou B, Khosla A, Lapedriza A, Oliva A and Torralba A (2016). Learning deep features for discriminative localization. pp. 2921–2929. https://doi.org/10.1109/CVPR.2016.319.

19. Ahn J and Kwak S (2018). Learning pixel-level semantic affinity with image-level supervision for weakly supervised semantic segmentation. pp. 4981–4990. https://doi.org/10.1109/CVPR.2018.00523.

20. Grady L (2006) Random walks for image segmentation. IEEE Trans Pattern Anal Mach Intell 28(11):1768–1783. https://doi.org/10.1109/TPAMI.2006.233

21. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D and Batra D (2017). Grad-cam: Visual explanations from deep networks via gradient-based localization. pp. 618–626. https://doi.org/10.1109/ICCV.2017.74.

22. Tang J (2010). A color image segmentation algorithm based on region growing. (IEEE), pp. V6–634-V636–637. https://doi.org/10.1109/ICCET.2010.5486012.

23. Wang Y, Shen X, Wu K, Huang M (2022) Corrosion grade recognition for weathering steel plate based on a convolutional neural network. Measurem Sci Technol 33:095014. https://doi.org/10.1088/1361-6501/ac7034

24. Organization IS (2007) Preparation of steel substrates before application of paints and related products — Visual assessment of surface cleanliness — Part 1: Rust grades and preparation grades of uncoated steel substrates and of steel substrates after overall removal of previous coatings.

25. Lin TY, RoyChowdhury A and Maji S (2015). Bilinear CNN models for fine-grained visual recognition. (IEEE), pp. 1449–1457. https://doi.org/10.1109/ICCV.2015.170.

26. Khayatazad M, De Pue L, De Waele W (2020) Detection of corrosion on steel structures using automated image processing. Develop Built Environ 3:100022. https://doi.org/10.1016/j.dibe.2020.100022

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