Semi-Supervised Learning Approach for Image-Based Corrosion Quantification on Water Resources Infrastructure

HAI D. NGUYEN, SHENGYI WANG, BRIAN EICK
and REBEKAH WILSON

ABSTRACT

Water resources infrastructure, such as locks and dams, are critical to the U.S. economy and public safety. Proper maintenance of these structures is essential to ensure their long-term functionality. Severe corrosion in such infrastructure can lead to significant section loss and, possibly, structural failure. Early detection and quantification of corrosion is therefore crucial to maintaining structural integrity and minimizing maintenance costs. This study explores the application of deep learning techniques, such as convolutional neural networks (CNNs), to segment corrosion areas in high-resolution images. A semi-supervised learning (SSL) approach, which leverages both labeled and unlabeled data, is employed. This method reduces reliance on fully labeled datasets, which are tedious to create and often require subject matter expertise. The proposed SSL approach is compared against baseline CNN models (DeepLabv3+, PSPNet, and UNet++), demonstrating promising results across key evaluation metrics, including mean precision, mean recall, mean F1 score, and mean Intersection-over-Union (IoU). These findings highlight the potential of SSL for rapid corrosion quantification, facilitating diagnosis, prognosis, and improved structural health monitoring of water resources infrastructure.

INTRODUCTION

Corrosion is the chemical or electrochemical deterioration of metals due to environmental reactions [1], posing significant risks to both civilian and military infrastructure. Among practical corrosion mitigation strategies, cathodic protection is one of the most effective and is frequently used in combination with coatings and corrosion inhibitors to protect assets such as underground pipelines and submerged navigation structures like miter and Tainter gates in USACE locks and dams.

Hai D. Nguyen, The U.S. Army Corps of Engineers, Champaign, IL 61822
Shengyi Wang, University of Illinois Urbana-Champaign, Urbana, IL 61801
Brian Eick, The U.S. Army Corps of Engineers, 2902 Newmark Dr, Champaign, IL 61822
Rebekah Wilson, The U.S. Army Corps of Engineers, Champaign, IL 61822

This cathodic technique works by converting the metal surface into the cathode of an electrochemical cell, thereby significantly slowing the corrosion process. However, it does not completely eliminate corrosion. As infrastructure continues to age, particularly USACE locks and dams, many of which exceed 50 years in service, corrosion remains a significant threat to structural integrity, leading to increased maintenance and repair costs. Traditional corrosion inspection methods are labor-intensive and subjective, highlighting the need for more efficient, data-driven, and image-based solutions.

Advancements in artificial intelligence (AI) and machine learning (ML) are transforming corrosion detection by enabling more accurate classification and segmentation of corrosion images. However, key challenges remain, particularly the time-consuming and often subjective process of annotating images required for effective model training. Additionally, high-resolution images, such as those captured by digital microscopes, present computational challenges due to their large size and high level of details.

To address these issues, this paper proposes a semi-supervised convolutional neural network (CNN)-based approach that leverages both labeled and unlabeled data. The developed model was successfully tested on a high-resolution corrosion image dataset of coated steel panels captured in a lab setting and shows potential for application to real-world corrosion scenarios. This research supports the development of a robust platform for accurate corrosion segmentation and quantification in critical infrastructure, enabling early warnings and more effective maintenance planning, thereby reducing maintenance costs and improving operational efficiency.

DATASET DEVELOPMENT: DATA COLLECTION & DATA ANNOTATION AND PROCESSING

DATA COLLECTION

In this study, hundreds of painted steel panels were intentionally damaged using a scribe tool to create standardized defects, including straight-line and X-shaped cuts that penetrated through the coating layer to the bare metal substrate. These scribed panels were then subjected to a 12-week accelerated corrosion testing protocol in accordance with ISO 12944 Part 5 [2], which provides guidelines for protective paint systems and exposure conditions suitable for high-corrosion environments. The performance of the coatings was subsequently evaluated using ASTM D1654 [3], which assesses corrosion-related failures such as rust creep, blistering, delamination, and other forms of coating degradation on intentionally damaged coated metal surfaces. Detailed procedures for panel preparation and testing can be found in Nguyen et al. (2023) [4] and Wang et al. (2025) [5]. After exposure, the corroded areas on the scribed panels were photographed using high-resolution digital microscopy, generating a comprehensive image dataset for deep learning model development.

DATA ANNOTATION AND PROCESSING

A total of 354 high-resolution images (ranging from 2880 × 2160 to 11,020 × 9295 pixels) were captured using a digital microscope. Among these, 262 images were annotated using ImageJ, a Java-based image processing program developed by the National Institutes of Health, together with the Labkit plugin (the segmentation of microscopy image data). Figure 1 presents representative images along with their corresponding masks. The annotated dataset was divided into training, validation, and test sets in an 8:1:1 ratio, resulting in 209 training images, 26 validation images, and 27 test images. An additional 92 unlabeled images were incorporated into the training set for semi-supervised learning, expanding it to 301 images.

To reduce computational costs and ensure compatibility with GPU processing, all images were downsampled to a maximum dimension of 2048 pixels using an anti-aliasing technique. This process helps prevent aliasing artifacts, such as jagged edges or “staircasing” effects [6], that can occur when images are sampled below the Nyquist frequency. According to Shannon’s sampling theorem [7], a band-limited signal (a signal with frequency content restricted to a specific range or bandwidth) can be perfectly reconstructed if sampled at a rate equal to or greater than twice its highest frequency component. Although initially developed for one-dimensional signals, this principle extends to two-dimensional images, where undersampling results in irreversible distortion or loss of detail. Shannon [7] also noted a threshold effect, where performance degrades rapidly if the sampling rate falls below the critical level. To avoid these issues, anti-aliasing is applied before downsampling to filter out high-frequency image details, such as fine textures and sharp edges, that cannot be reliably preserved at lower resolutions. This minimizes visual distortion and results in smoother, more coherent images while maintaining the original aspect ratio.

After downsampling, images were split into 256 × 256-pixel patches to accommodate GPU memory limits. To enhance generalization and mitigate overfitting, horizontal flipping was applied as data augmentation, effectively doubling the training set with mirrored image variants to increase pattern diversity and model robustness.

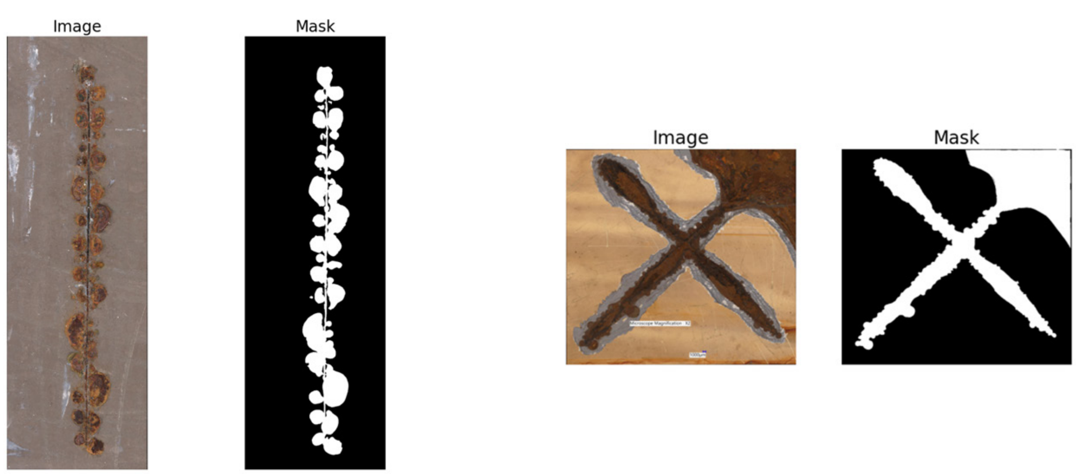


Figure 1. Examples of annotated and preprocessed images with corresponding masks (left: line-scribed; right: X-scribed).

SEMI-SUPERVISED CORROSION SEGMENTATION MODEL

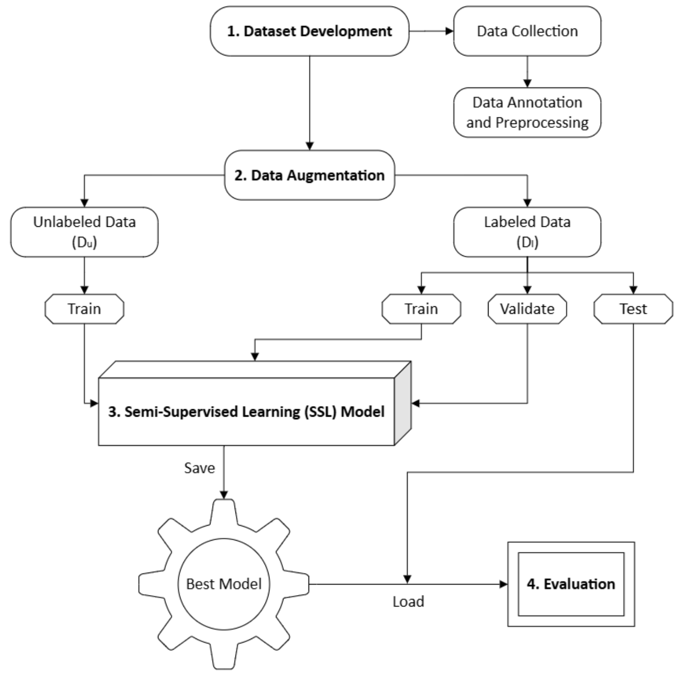


Figure 2. Flowchart of proposed corrosion segmentation method.

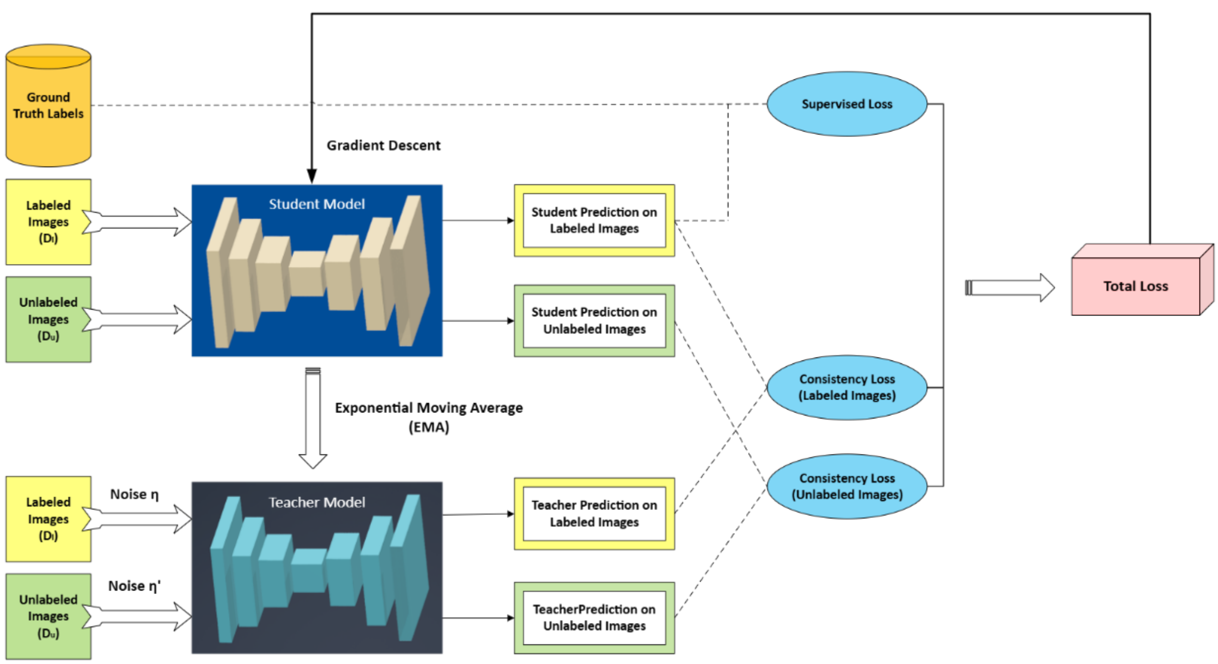


Figure 3. Mean teacher model.

This paper proposes a deep learning-based image segmentation method for accurately quantifying corrosion in high-resolution images, using both labeled and unlabeled data through semi-supervised learning (SSL). A subset of annotated corrosion images of steel panels was used to train CNN-based models, with image patches serving as input. The method includes two main components. First, it implements a mean teacher framework [8], consisting of student and teacher models that share the DeepLabv3+ architecture with a ResNet-34 backbone but operate under different training objectives. Transfer learning from ImageNet is employed to enhance feature extraction. Additional CNN-based models, including UNet++ and PSPNet, were tested for comparison. Figure 2 presents the flowchart of the proposed method, which includes: (1) Dataset development, (2) Data augmentation, (3) SSL model, and (4) Evaluation. An overview of the mean teacher model is shown in Figure 3, with further details on the model framework and architecture available in Wang et al. (2025) [5].

BASELINE CNN MODELS

DEEPLABV3+

The DeepLab system, proposed by Chen et al. (2018) [9], addresses semantic image segmentation challenges by introducing convolution with upscaled filters (also known as atrous or dilated convolution) to control feature resolution and expand the field of view without increasing parameters or computational cost. It employs atrous spatial pyramid pooling (ASPP) to capture multi-scale context and a fully connected Conditional Random Field (CRF) to refine object boundaries. Initially built on VGG-16 and later adapted to ResNet-101 for improved performance, DeepLab achieved great evaluation results on multiple datasets, including PASCAL VOC 2012, PASCAL-Context, Cityscapes, and PASCAL-Person-Part. DeepLabv3 improves upon DeepLabv2 by more effectively addressing the challenge of segmenting objects at multiple scales. It introduces an enhanced ASPP module, which applies parallel atrous convolutions with different dilation rates to capture rich multi-scale contextual information.

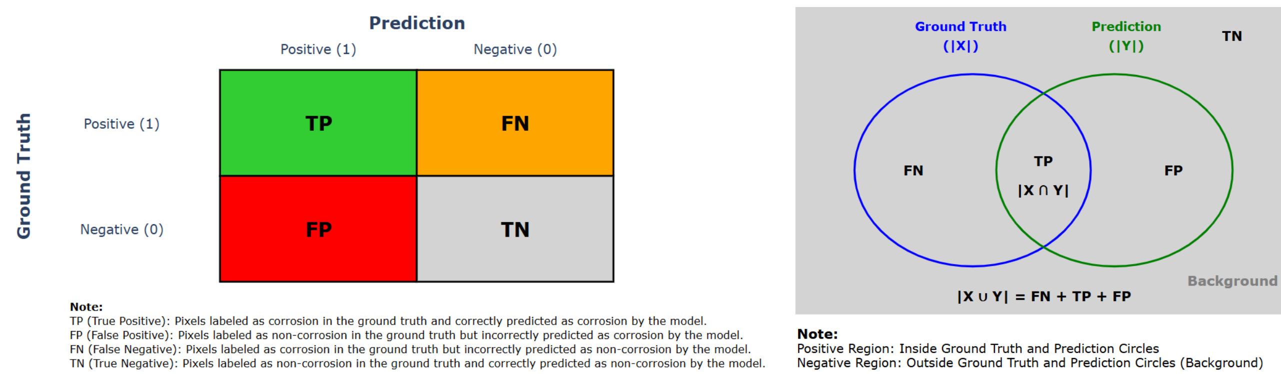
PSPNET

The Pyramid Scene Parsing Network (PSPNet), proposed by Zhao et al. (2017) [10], enhances pixel-level semantic segmentation by leveraging a dilated ResNet backbone with auxiliary loss and introducing a Pyramid Pooling Module to capture both global and local context information. By addressing key limitations of traditional fully convolutional networks (FCNs), such as poor global scene understanding and confusion between similar objects, PSPNet significantly improves segmentation accuracy. It achieved impressive accuracy, winning the ImageNet Scene Parsing Challenge 2016 and setting new benchmarks on PASCAL VOC 2012 and Cityscapes dataset, establishing itself as a major contribution to scene parsing research.

UNET++

UNet++ is a nested U-Net architecture proposed by Zhou et al. (2018) [11] to improve medical image segmentation by addressing the semantic gap between encoder and decoder features found in traditional U-Net models. It introduces re-designed skip pathways using nested dense convolutional blocks, which gradually refine encoder features before merging with decoder features, making the learning task easier and improving segmentation accuracy. Additionally, deep supervision is employed to enhance training and facilitate model pruning at inference time (a technique that selectively eliminates less important parts of the neural network, such as intermediate layers or branches, to reduce model size, complexity, and computational load) for faster processing. Evaluated on multiple medical imaging tasks (e.g., lung nodules, liver, polyp, cell nuclei segmentation, etc.), UNet++ consistently outperforms both the original U-Net and wide U-Net variants.

EVALUATION METRICS



(a) Confusion Matrix

(b) Ground truth vs. predicted segmentation masks

Figure 4. Confusion matrix and segmentation masks.

The evaluation of corrosion segmentation was conducted by comparing the segmentation results to the ground truth data using four metrics: precision, recall, F1-score (aka the Dice coefficient), and Intersection-over-Union (IoU). These metrics are defined mathematically in Equations 1-4 and illustrated both intuitively and visually through the confusion matrix and segmentation masks shown in Figure 4. Precision (aka positive predictive value, Equation 1) measures the proportion of pixels predicted as corrosion that are actually corrosion in the ground truth, indicating the accuracy of positive predictions. Recall (aka the true positive rate or sensitivity, Equation 2) measures the proportion of actual corrosion pixels in the ground truth that were correctly identified by the model, reflecting the model's ability to detect all relevant corrosion pixels. The F1 score or Dice coefficient (Equation 3) is the harmonic mean of precision and recall, balancing the two by measuring the similarity between the predicted corrosion area and the ground truth corrosion area. A high F1 score, close to one, indicates a strong similarity between prediction and ground truth. Finally, IoU (Equation 4) measures the extent of overlap between the predicted corrosion region and the ground truth, relative to their combined area, providing a direct assessment of the segmentation accuracy.

$$\text{Precision} = \frac{\text{number of pixels correctly predicted as corrosion by the model}}{\text{total number of pixels predicted as corrosion by the model}} = \frac{TP}{TP+FP} \tag{1}$$

$$\text{Recall} = \frac{\text{number of pixels correctly predicted as corrosion by the model}}{\text{total number of pixels actually labeled as corrosion in the ground truth}} = \frac{TP}{TP+FN} \tag{2}$$

$$\text{F-1 (Dice)} = \frac{2\times\text{Precision}\times\text{Recall}}{\text{Precision} + \text{Recall}} = \frac{2\times|X \cap Y|}{|X|+|Y|} = \frac{2TP}{FN+2TP+FP} \quad (3)$$

$$\text{IoU} == \frac{|X \cap Y|}{|X \cup Y|} = \frac{|X \cap Y|}{|X|+|Y|-|X \cap Y|} = \frac{TP}{FN+TP+FP} \quad (4)$$

where $|X|$ = the set of pixels labeled as corrosion the ground truth and $|Y|$ = the set of pixels predicted as corrosion by the model.

RESULTS AND DISCUSSION

TABLE I. PERFORMANCE OF PROPOSED CORROSION SEGMENTATION METHOD COMPARED TO BASELINE CNN-BASED MODELS

Model	SSL	Precision	Recall	F-1 (Dice)	IoU
DeepLabv3+ (SSL)	***Yes***	***87.9%***	***95.2%***	***91.1%***	***84.4%***
DeepLabv3+	No	82.5%	91.2%	85.9%	77.0%
PSPNet	No	86.9%	94.7%	90.0%	83.3%
UNet++	No	83.4%	93.6%	87.3%	79.2%

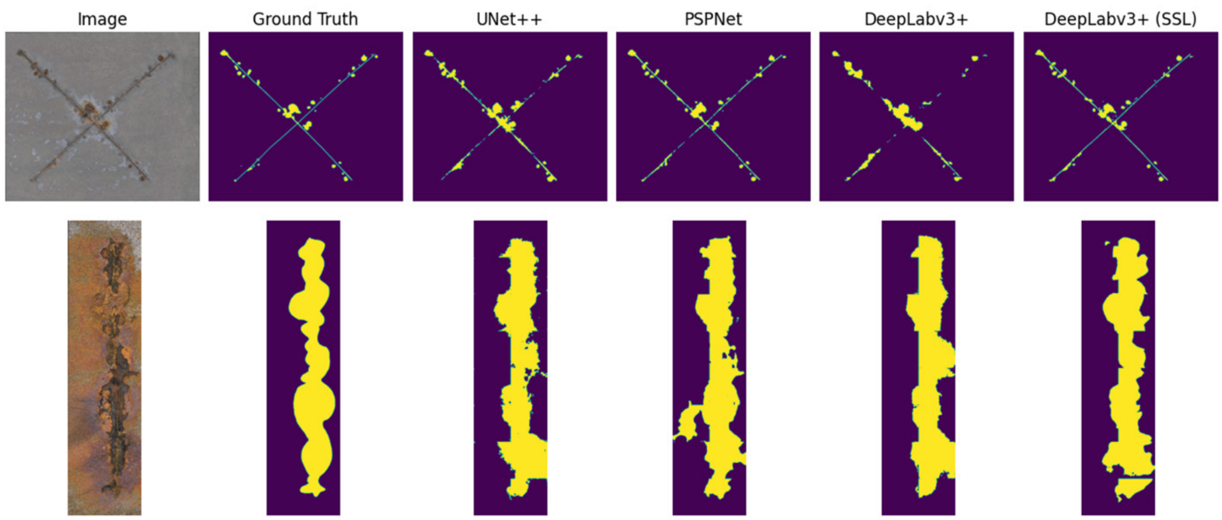


Figure 5. Examples of original image, corresponding ground truth masks, predicted mask for UNet++, PSPNet, DeepLabv3+, and DeepLabv3+ (SSL) models.

Table 1 presents the performance of the proposed corrosion segmentation method, DeepLabv3+ (SSL), in comparison with baseline CNN-based models, including DeepLabv3+, PSPNet, and UNet++. It is important to note that semi-supervised learning (SSL) was applied only to the enhanced version of DeepLabv3+ (i.e., DeepLabv3+ (SSL)) and not to the baseline models. The standard DeepLabv3+ model (without SSL) achieved a precision of 82.5%, recall of 91.2%, F1-score of 85.9%, and IoU of 77.0%. With the integration of SSL, DeepLabv3+ (SSL) achieved the highest performance across all metrics with a precision of 87.9%, recall of 95.2%, F1-score of 91.1%, and IoU of 84.4%. This demonstrates the effectiveness of SSL in enhancing segmentation accuracy. These improvements highlight SSL's impact in refining the model's ability to precisely segment and quantify corrosion in high-resolution images.

Figure 5 presents visual comparisons of the original images, ground truth masks, and predicted masks generated by the baseline models (UNet++, PSPNet, and

DeepLabv3+) and the enhanced DeepLabv3+ (SSL). As illustrated, the baseline models tend to under-segment corroded regions. While DeepLabv3+ (without SSL) outperforms the other baselines, it still exhibits minor segmentation errors. In contrast, DeepLabv3+ (SSL) produces more accurate and complete segmentations, highlighting the effectiveness of integrating semi-supervised learning into the segmentation process.

CONCLUSION

This paper presents a novel semi-supervised deep learning approach for automatic corrosion segmentation on coated steel panels. The proposed method, DeepLabv3+ (SSL), demonstrates strong segmentation performance, achieving mean precision, recall, F1-score, and IoU of 87.9%, 95.2%, 91.1%, and 84.4%, respectively. It outperforms baseline CNN models, including UNet++, PSPNet, and DeepLabv3+. As one of the few studies to apply SSL to corrosion segmentation, the results highlight its potential to improve the precision of corrosion detection and quantification. By enhancing segmentation accuracy, the proposed method may contribute to enhance infrastructure safety, extended service life, and reduced maintenance costs.

REFERENCES

1. Revie, R. W., & Uhlig, H. H. (2008). *Corrosion and corrosion control: an introduction to corrosion science and engineering*. John Wiley & Sons.
2. ISO 12944-5. (2018). Paints and Varnishes—Corrosion Protection of Steel Structures by Protective Paint Systems—Part 5: Protective Paint Systems.
3. ASTM Committee D-1 on Paint and Related Coatings, Materials, and Applications. (2008). *Standard Test Method for Evaluation of Painted or Coated Specimens Subjected to Corrosive Environments*. ASTM International.
4. NGUYEN, H. D., WANG, S., WILSON, R., EICK, B., & BECERRA-STASIEWICZ, N. A. T. A. L. I. E. (2023). Accelerating corrosion surface-area measurements with computer vision and deep learning: an ensemble approach. *STRUCTURAL HEALTH MONITORING 2023*.
5. Wang, S., Nguyen, H., Wilson, R., Eick, B., El-Gohary, N., & Ortiz, C. (2025). Deep CNN-based semi-supervised learning approach for identifying and segmenting corrosion in hydraulic steel and water resources infrastructure. *Structural Health Monitoring*, 14759217241305039.
6. Wu, X. (1991). An efficient antialiasing technique. *Acm Siggraph Computer Graphics*, *25*(4), 143-152.
7. Shannon, C. E. (1949). Communication in the presence of noise. *Proceedings of the IRE*, *37*(1), 10-21.
8. Tarvainen, A., & Valpola, H. (2017). Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results. *Advances in neural information processing systems*, *30*.
9. Chen, L. C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2017). Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE transactions on pattern analysis and machine intelligence*, *40*(4), 834-848.
10. Zhao, H., Shi, J., Qi, X., Wang, X., & Jia, J. (2017). Pyramid scene parsing network. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2881-2890).
11. Zhou, Z., Rahman Siddiquee, M. M., Tajbakhsh, N., & Liang, J. (2018). Unet++: A nested u-net architecture for medical image segmentation. In *Deep learning in medical image analysis and multimodal learning for clinical decision support: 4th international workshop, DLMIA 2018, and 8th international workshop, ML-CDS 2018, held in conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, proceedings 4* (pp. 3-11). Springer International Publishing.