Automated Detection of IQI Wires and Weld Defects in Ship Pipe Radiographic Images Using Deep Object Detection Model

Shinhyo Kim
Department of Marine Engineering
Mokpo National Marintime University
Mokpo, Republic of Korea
rainbowfin1@naver.com

Seunghun Lim
Department of Marine Engineering
Mokpo National Marintime University
Mokpo, Republic of Korea
seunghun3902@naver.com

Sungho Jo
Department of Marine Engineering
Mokpo National Marintime University
Mokpo, Republic of Korea
jang5938@naver.com

Jungmo Oh
Division of Marine System Engineering
Mokpo National Marintime University
Mokpo, Republic of Korea
jmoh@mmu.ac.kr

Jinkyu Park
Division of Marine System Engineering
Mokpo National Marintime University
Mokpo, Republic of Korea
pjk2019@mmu.ac.kr

***Abstract*—**Nondestructive Testing(NDT) is used to ensure the reliability of ship pipe welds, and RT is the most widely used method. However, RT interpretation relies on experimented technicians. Thus IQI interpretation and defect detection automatic systems are in demand. Accordingly, we analyzed and evaluated the performance of object detection models for IQI and weld defects. The dataset was constructed using YOLOv8x and Faster R-CNN models for IQI wire detection, and YOLOv8s for defect detection. The result of experiment showed that YOLOv8x had high robustness in detection IQI wires in RT images and demonstrated excellent performance overall. Faster R-CNN showed high detection rates in high-quality images but had some over detection. In defect detection, the YOLOv8s model showed high precision but low recall due to missing small defects. The results of this study demonstrate that object detection models can effectively be applied to the automation of RT image-based pipe welding defect inspection. In the future work, we include expanding the dataset to enhance model performance.

Keywords—radiographic testing, image quality indicator, object detection, weld defect, machine learning

I. Introduction

To ensure the reliability and integrity of welded or fabricated components, various inspection techniques are employed[1]. In ship pipe welds, non-destructive testing(NDT) in utilized to detect internal defects without causing any physical or chemical damage to the material. Among the available NDT methods, radiographic testing(RT) is one of the most widely used techniques, and it becomes particularly indispensable for structures requiring exceptionally highly reliability, such as those used in LNG vessels.

In shipyard, RT inspections are primarily interpreted by qualified and experienced technicians. However, this manual interpretation process is prone to human error, as the results can vary depending on the inspector's individual skill and judgement, leading to limited reproducibility. Wang(2005) and Shafeek(2004) et. al., highlighted the inefficiency of conventional analog radiographic image assessment method and the low inter-rater reliability among inspectors [2, 3].

RT inspection procedures are generally divided into image quality indicator(IQI) evaluation and defect inspection. The IQI evaluation assesses the sensitivity and resolution of radiographic images in accordance with ASTM E747., ensuring that the captured images meet eh required quality standards. Once image quality has been verified, defect interpretation is performed. The presence or absence of defect is the determined based on factors such as pipe size and defect characteristics.

Although automated image-based defect identification technologies have been actively developed in fields such as construction and medicine, their adoption in the shipbuilding industry has been relatively slow due to concerns regarding reliability and consistency. Recent advances in image-based object detection have demonstrated the effectiveness of machine learning in diverse application domains, including medicine, agriculture, and construction[4-6]. Motivated by these developments, this study proposes an automated image-based inspection approach for detecting defects in pipe welds.

We propose an algorithm for automatically reading IQI, which serves as a benchmark for defect analysis, and for detecting various types of defects. To this end, we train multiple Faster R-CNN variants with different backbones, together with the YOLOv8 model–renowned for its strong small-object detection capability–to develop models tailored for pipe weld defect inspection, and we rigorously compare their performance.

II. EXPERIMENTAL PROCEDURE

A. Experimental materials

The dataset for this study is divided into two parts: one for IQI detection and another for weld defect detection.

As shown Table I, IQI dataset consisted of 469 original RT images of SUS316. To enhance the model's robustness and prevent overfitting, horizontal and vertical flipping techniques were applied, resulting in a total of 1,407 images.

TABLE I. SPECIFICATIONS AND NUMBER OF RT IMAGES FOR IQI DETECTION ALGORITHM

Specification	Value					
Diameter (mm)	400	500	600	650	700	750
Thickness (mm)	12.7	5.5	12.7	7.9	7.9	7.9
Number (pcs)	108	7	101	54	151	48
Total number (pcs)	469					

As shown table II, the weld defect detection dataset was constructed using a total of 3,571 images obtained from two methods. Due to the limited defect images, the dataset was composed of original RT images and AI-Hub images, an artificial intelligence data platform.

TABLE II. DATASET SPLIT FOR WELD DEFECT DETECTION

Dataset	Original images	AI-Hub images	Total
Training	1,611	1,250	2,861
Validation	350	290	640
Test	70	-	70
Total	2,031	1,540	3,571

To ensure robust feature extraction and address potential class imbalance, a stratified data splitting strategy was employed rather than simple random sampling. The dataset was partitioned so that all categories of welding defects were represented in the training set. Specifically, both datasets were split into training, validation, and test sets at a ratio of 7:2:1, maintaining a consistent distribution of defect types to enhance the model's generalization performance.

The hardware and software specifications are shown in Table III.

TABLE III. HARDWARE AND SOFTWARE ENVIRONMENT

CPU	Intel(R) Core(TM) i7-14700k	**GPU**	NVIDIA GeForce RTX 4060 Ti(8.0GB)
Mainboard	Gigabyte b760M AORUS ELITE, Intel B760M Chipset		
Storage	2.0 TB HDD ST2000DM008-2UB102, 1.0 TB SSD Samsung SSD 990 pro 1 TB		
OS	Windows 11 pro 64 bit	**RAM**	64.0GB DDR5

B. Experimental methods

For model training, we used the Labelme tool(5.2.1) to label the IQI and defect areas appearing in the image. To accurately designate the location of the IQI wire, labeling was performed into three classes:Text, wire, wire region, which are common to the wire area in the image. In addition, the weld defect area was labeled with in 'defect' class. Although the two results are combined and visualized as a single example in Fig.1 for convenience of explanation, the IQI and defect area were actually labeled and trained independently. A visual example of the labeling result is shown in Fig.I.

First, we compared the performance of YOLOv8x, a representative object detection model, which are representative object detection models, for IQI wire detection. YOLOv8x is a model that was released in January 10, 2023 and is based on the development of YOLO versions. It uses cutting-edge backbone and neck architectures to improve feature extraction and object detection performance. Faster R-CNN introduces a Region Proposal Network(RPN) that performs region proposal and object classification within a single algorithm. It shares feature maps to achieve high accuracy and speed. Also, we used the YOLOv8s object detection model to detect welding defects in RT images. This model's structural features follow the YOLOv8 version mentioned earlier.

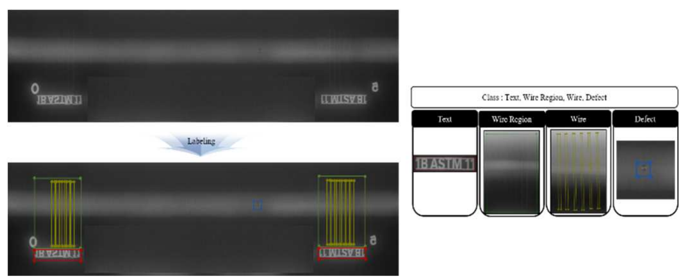


FIGURE I. LABELING RESULT FOR IQI AND WELD DEFECTS

The main hyperparameters of each model used in the IQI Wire detection study are shown in Table 3. In the case of YOLOv8x, the high-resolution training images and model structure lead to large memory usage. Therefore, the batch size was set to 2 considering the GPU available in the experimental environment. The learning rate was set to 0.01 for learning stability. In addition, the main hyperparameters of the models used in the pipe weld defect detection study are shown in Table IV. In the defect detection model training, some parameters, including the batch size, were set to accommodate hardware constraints because the GPU usage available in the experimental environment was limited. In addition, the main hyperparameters of the models used in the pipe weld defect detection study are shown in Table IV. In the defect detection model training, some parameters, including the batch size, were set to accommodate hardware constraints because the GPU usage available in the experimental environment was limited.

TABLE IV. HYPERPARAMETERS OF THE DETECTION MODELS

	IQI Wire model		Defect model
Parameter	**YOLOv8x**	**Faster R-CNN**	**YOLOv8s**
Batch size	2	8	2

Epoch	180	35	286
Learning rate(0)	0.01	0.0001	0.01
Learning rate(f)	0.01	0.0001	0.01
Image size	1920		1920
IoU	0.5		0.3

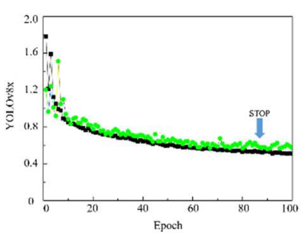
The model that was trained to detect IQI wires was evaluated using precision, recall, mAP@50, mAP@50-95 and F1 score. The location accuracy of the IQI wires was determined when the IoU was greater than 0.5. The model for defect detection was evaluated using the same quantitative evaluation criteria, considering the defect location accuracy to be 0.3 or higher.

III. RESULTS AND DICUSSION

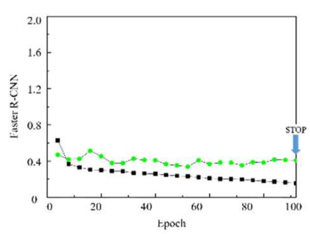
A. *Analysis of Training Loss Trend*

Fig.II showed the loss changes of YOLOv8x and Faster R-CNN for IQI wire detection and YOLOv8s for weld joint defect detection. First, as shown in Fig.2 YOLOv8x converged at epoch 180 and Faster R-CNN at epoch 35, and the models were terminated early at the points. Although the validation loss was relatively higher than the training loss, it showed a tendency to converge at a certain point, indicating a stabilized learning pattern. In the future, we plan to use a model re-learning process with images of various IQI shapes to improve generalization performance.

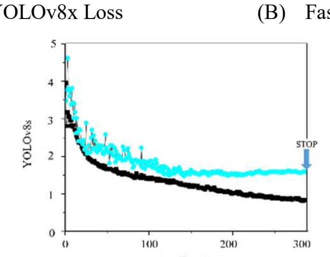
YOLOv8s is trained for a total of 1,000 epochs, with an early stopping patience of 80. Training loss decreased as the Epochs increased, and validation loss initially fluctuated but then decreased. This indicates that the model has the ability to generalize to defect detection without overfitting. However, the validation loss didn't decrease enough, so in the future work, we will need to expand the diversity of the defect data and use additional augmentation techniques to improve the model's detection performance



(A) YOLOv8x Loss



(B) Faster R-CNN Loss



(c) YOLOv8s Loss

FIGURE II. TRAINING AND VALIDATION LOSS FOR EACH DETECTION MODELS

B. *Analysis of Quantitative Metrics for the models*

As shown in Table V, the YOLOv8x model achieved excellent performance in IQI wire detection using a total of 150 validation images, demonstrating consistent detection rates across various RT image qualities. Conversely, the Faster R-CNN model showed relatively lower performance. This can be interpreted as a susceptibility to duplicate and low-quality RT images.

Meanwhile, the YOLOv8s model, evaluated using 350 validation images, recorded high precision values. This indicates a relatively low rate of over-detection, indicating that the learning model achieves a certain level of reliability in defect detection. However, some small sized defects were missed, resulting in a somewhat low recall value. This suggests a lack of small sized defect data. Therefore, in the future work, we plan to secure defect data of various sizes to enhance model reliability

TABLE V. QUANTITATIVE METRICS FOR THE MODELS

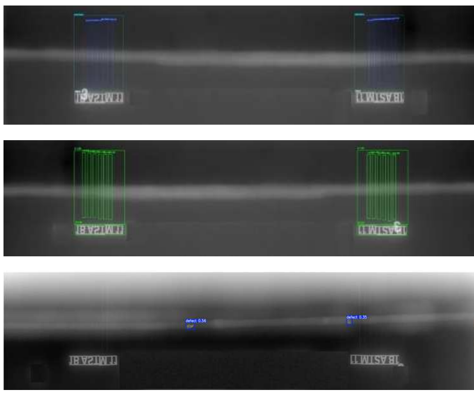
Model	Precision	Recall	mAP@50	mAP@50-95	F1 score
YOLOv8x	0.987	0.972	0.992	0.784	0.979
Faster R-CNN	0.338	0.528	0.483	0.338	0.206
YOLOv8s	0.783	0.609	0.668	0.243	0.685

C. *Test image Detection results*

Table III provides a visual comparison of IQI and defect prediction results using test images.

First, in the IQI wire detection performance evaluation using 96 test images, the YOLOv8x model showed relatively stable detection performance even in low quality images, with an accuracy rate of approximately 55% in recognizing wires. Meanwhile, the Faster R-CNN model detected wires in only about 48%, which is lower than YOLOv8x. Additionally, a total of 41 cases of duplicate detection were identified. These performance differences are due not only to the structural characteristics of the models, but also to the influence of the labeling strategy performed by non-experts. Wire patterns that had low contrast were difficult to identify even with the eyes and were excluded form the manual labeling process. Under these conditions, YOLOv8x learned the labeled wire patterns consistently and showed high detection accuracy that aligned well with the given correct data. Meanwhile, Faster R-CNN showed a tendency to detect structural patterns in unlabeled areas as wires, which were counted as duplicate of false detections in the evaluation process. This characteristic of Faster R-CNN suggests that it can identify the boundaries of individual wires more precisely when the image quality is high enough. Therefore, the performance differences observed in this study do not necessarily indicate the superiority of a specific model, but rather, the different result of each model to varying image quality.

Second, the performance evaluation of pipe welding defect detection using 61 test images revealed that 41 images accurately detected defects, demonstrating a certain level of reliability. However, in some images, small defects were missed, resulting in a lower detection accuracy. This is likely due to the insufficient representation of small defects in the training data. Additionally, due to the limitations of the GPU memory used in the experiment, other hyperparameters had to be set to fixed values, which have prevented the optimal detection of defects. Future studies will include additional data with a variety of defect sizes, as well as hyperparameter optimization in a high quality computing environment, to improve the model's generalization performance and detection accuracy.



(a) IQI prediction results for YOLOv8x

(b) IQI prediction results for Faster R-CNN

(c) Defect prediction results for YOLOv8s

FIGURE III. VISUALIZATION OF IQI AND DEFECT PREDICTIONS

IV. CONCLUSION

The objective of this study was to automate RT image-based pipe welding joint inspection. We compared and evaluated the performance of object detection models for IQI wire detection and weld joint defect detection. A dataset was built using images that included various IQI shapes and defects and the models' performance metrics were analyzed.

In the IQI detection results, YOLOv8x showed excellent performance in most indicators and high robustness. Faster R-CNN showed a tendency to detect more precisely in high-quality images, but performance degradation was confirmed due to some images with duplicate detections.

Additionally, YOLOv8x demonstrated consistent detection performance when evaluated using test images. In the result of detecting welding defects, the YOLOv8s model recorded a high accuracy, but due to the absence of small defects, it recorded a low recall value. This is likely due to the limited size of the data set. However, it accurately detected defects in 41 of the 61 test images, achieving an approximate success rate of 67%. The results of this study demonstrate the effectiveness of developing automated pipe welding defect inspection technology using object detection models and suggest the possibility of using models complementarily depending on the quality of the images.

The results for this study demonstrate the feasibility and effectiveness of applying object detection models to the automated inspection of ship pipe IQI and defect detection. In particular, we confirmed that the detection characteristics of the models differ depending on the image quality. Future research will involve building a dataset and performing hyperparameter optimization in an upgraded computer environment to further improve detection accuracy and generalization performance.

ACKNOWLEDGMENT

This work was supported by the Industrial Strategic Technology Development Program (RS-2024-00421155, Development of an Automated System for High-Efficiency Radiographic Testing of 3D Pipe Spool Welds)

REFERENCES

[1] Boogaard, J., & van Dijk, G. M. (Eds.). (2012). Non-destructive testing (Vol. 1). Elsevier.

[2] Wang, X., Zscherpel, U., Tripicchio, P., D'Avella, S., Zhang, B., Wu, J., Liang, Z., Zhou, S. and Yu, X. (2025). A comprehensive review of welding defect recognition from X-ray images. Journal of Manufacturing Processes, Vol. 140, pp. 161-180.

[3] Shafeek, H. I., Gadelmawla, E. S., Abdel-Shafy, A. A. and Elewa, I. M. (2004). Assessment of welding defects for gas pipeline radiographs using computer vision. NDT & e International, Vol. 37, No. 4, pp. 291-299.

[4] Kim, S. and Ahn, J. (2024). A YOLO-Based Crop Pests Detection Mobile Application for Smart Farming. Journal of the Korea Academia-Industrial, Vol. 25, No. 7, pp. 603-610.

[5] Zhang, C. Y., & Kim, D. H. (2020). Analysis of impact position based on deep learning CNN algorithm. Trans. Korean Soc. of Mech. Eng. A, vol. 44, no. 6, pp. 405-412.

[6] Lalinia, M. and Sahafi, A. (2024). Colorectal polyp detection in colonoscopy images using YOLO-V8 network. Signal, Image and Video Processing, vol. 18, no. 3, pp. 2047-2058.