

Deep-learning-based object detection for automated IQI evaluation in analog film radiographic images of piping welds

Seunghun Lim¹ · Shinhyo Kim² · Jinkyu Park[†]

(Received August 5, 2025 : Revised August 15, 2025 : Accepted September 10, 2025)

Abstract: Radiographic testing (RT) is a nondestructive testing method for evaluating internal defects in welds, and image quality is assessed using image quality indicators (IQIs). In this study, two deep-learning-based object detection models, YOLOv8x and Faster R-CNN, were applied to detect IQI wires in analog RT images automatically. YOLOv8x demonstrated excellent performance, achieving a precision of 0.987, recall of 0.972, and mAP@50 of 0.992. The wire count match rate in the low-contrast images was 55.0% for YOLOv8x and 50% for Faster R-CNN. This study demonstrates the feasibility of automating IQI wire detection in analog RT images, and provides a foundation for future digital RT systems.

Keywords: Radiographic testing, Image quality indicator, Object detection, Deep learning, Non-destructive testing

1. Introduction

Nondestructive testing (NDT) is a technique used to identify internal defects in materials without causing physical or chemical damage. It is a testing method that plays a key role in ensuring the structural stability and reliability of components in various industries, including construction, shipbuilding, and heavy industries [1]. Among the available NDT methods, radiographic testing (RT) is the most widely employed, and it is particularly useful for inspecting the welding joints of LNG cargo transfer pipelines on special vessels such as LNG carriers in the shipbuilding industry. In RT, the image quality directly affects the accuracy and reliability of defect detection results, and an image quality indicator (IQI) is used to quantitatively verify this. An IQI evaluates the sensitivity and resolution of radiographic images by determining the visibility of wires with specific diameters. Inspection and verification procedures are typically performed in accordance with international standards such as ASTM E747. In flat structures, RT has been digitalized, allowing the implementation of automated inspection systems. However, in curved structures such as pipes, the application of curved detectors is challenging; therefore, analog film-based RT is still used. In particular, shipyards, which deal with many curved structures, rely

on analog-film-based methods, and IQI readings are performed only by skilled technicians. This process relies entirely on individual capabilities and visual judgment, leading to a high likelihood of errors because of subjective judgment, significant time and labor consumption, and insufficient reproducibility, making it inadequate for implementing an efficient process environment. Wang *et al.* (2005) mentioned the inefficiency of conventional analog-image analysis methods [2], whereas Shafeek *et al.* (2004) pointed out the lack of inter-rater reliability in manual interpretation [3].

Digital radiography (DR) and computed radiography (CR) are being increasingly used to overcome these limitations. In particular, deep-learning-based computer vision techniques have been actively explored and developed for the effective detection of defects. However, automated defect detection and IQI judgment technologies are necessary to introduce an automated system for inspecting defects in pipe welds in shipyards.

Several studies have been conducted on image-based object detection, leading to the adoption of various advanced technologies. Conventionally, image analysis has been performed using artificial neural networks, such as convolutional neural networks (CNNs), in several fields including polyp detection in medicine,

[†] Corresponding Author (ORCID: <http://orcid.org/0000-0002-6483-8457>): Professor, Division of Marine System Engineering, Mokpo National Maritime University, 91, Haeyangdaehak-ro, Mokpo-si, Jeollanam-do, 58628, Korea, E-mail: pjk2019@mmu.ac.kr, Tel: +82-61-240-7219

1 Ph. D. Candidate, Department of Marine Engineering, Mokpo National Maritime University, E-mail: seunghun3902@naver.com, Tel: +82-61-240-7219

2 M. S., Department of Marine Engineering, Mokpo National Maritime University, E-mail: rainbowfn1@naver.com, Tel: +82-61-240-7219

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

pest recognition in agriculture, and crack detection in construction, demonstrating the usefulness of deep learning in image processing tasks [4][5].

YOLO (You only look once) is an algorithm that utilizes CNN architecture and is effective in detecting objects in images. Compared with conventional approaches, such as R-CNN and deformable parts models (DPM), YOLO has demonstrated a very high performance. It was first introduced by Redmon et al. in 2015; since then, numerous versions have been developed and used in various fields [6]-[8].

Previous studies have generally focused on the detection or classification of weld defects, leaving a gap in the research on automatic IQI judgment, which serves as the basis for defect analysis. This study proposes a framework based on the YOLOv8x and Faster R-CNN models to automatically detect IQI locations in RT images, automatically determine the compliance with international standards, and compare their performances. Because various studies on YOLO have been conducted, new versions of YOLO (e.g., YOLOv12) have been developed. However, Choi et al. (2024) confirmed that among the different YOLO models, the YOLOv8 model demonstrated the best performance for small-object detection [9]. Therefore, YOLOv8 was adopted as the primary model in this study. The existing YOLO-based object detection models are optimized for high-contrast, single-object-centric learning. Therefore, we applied a separate refined training dataset and a dedicated labeling strategy for detecting low-contrast, small-scale IQI structures.

In the remainder of this paper, Section 2 reviews previous studies related to this research, and Section 3 describes the dataset and overall experimental methods for YOLOv8x and Faster R-CNN modeling. Section 4 evaluates the validity of the model using a test dataset generated based on the experimental environment. Finally, Section 5 summarizes the results and discusses directions for future research.

2. Related Studies

Zhang *et al.* (2023) proposed the S-YOLO model, an enhanced architecture based on YOLOv8-nano, to detect weld defects. To address challenges such as small defect detection and object overlap, the model incorporated omnidimensional dynamic convolution and the attention mechanism of the Normalization-based attentions module (NAM), while introducing a context augmentation module instead of the Spatial pyramid pooling fast (SPPF) module to efficiently integrate multiresolution information. The experimental results showed an 8.9% improvement

in mAP@50 compared with the existing YOLOv8, demonstrating superior performance in terms of both inference speed and accuracy [10].

Lalinia *et al.* (2023) utilized the YOLOv8 model for automatic detection of colon polyps. Approximately 1,890 endoscopic images were used to train the model through data augmentation. The proposed approach demonstrated high detection performance across polyps of various sizes and shapes, achieving a precision of 95.6%, recall rate of 91.7%, and F1-score of 92.4% [11].

Roussel *et al.* (2025) compared the object detection performance of YOLOv8 and Faster R-CNN in pipe welds. They trained each model using approximately 5,000 images, and the results of YOLOv8 showed a high quantitative performance with a mAP@50 of 0.89, precision of 83%, and recall of 87%. YOLOv8 achieved a good balance between detection speed and accuracy, indicating the suitability of the model for IQI identification and the detection of large-scale objects [12].

Pan *et al.* (2023) proposed a WD-YOLO architecture for X-ray-based weld defect detection and introduced a gray-value curve enhancement (GCE) module and a dual attention mechanism to solve problems related to low contrast and varying defect sizes. The experimental results showed that the proposed method achieved a mAP@50 of 92.6% and an F1-score of 87.6%, demonstrating a high detection performance [13].

3. Experimental Materials and Methods

3.1 Experimental Materials

The material used for the pipes in the experiment was SUS 316 L stainless steel—primarily used for transporting cryogenic fluids such as LNG and LPG. The test pipes had various specifications ranging from 400 to 750 mm, and 272 RT images were used, as detailed in **Table 1**. Each image included annotations for welded joints, IQIs, and pipe specifications.

Of the collected images, 70% were used as the training set, 20% as the validation set for evaluating the proposed model, and the remaining 10% were used as the test set. To address the issue of degraded learning performance caused by the limited number

Table 1: Summary of RT image counts for model training and evaluation

Item	Value				
	400	600	650	700	750
Diameter (mm)	400	600	650	700	750
Thickness (mm)	12.7	12.7	7.9	7.9	7.9
Number (pcs)	102	101	7	14	48
Total number (pcs)	272				

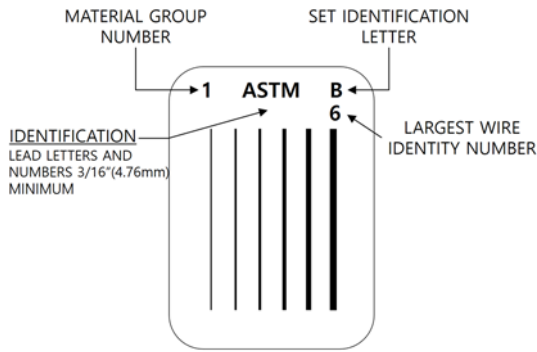


Figure 1: Configuration and labeling format of wire-type image quality indicator (IQI) in accordance with the ASTM E747 standard

Table 2: Wire sizes and corresponding identity numbers for the IQI (ASTM Class B)

Wire diameter in. (mm)	Wire identity
0.010 (0.25)	6
0.013 (0.33)	7
0.016 (0.41)	8
0.020 (0.51)	9
0.025 (0.64)	10
0.032 (0.81)	11

of images, data augmentation was applied, resulting in 570, 150, and 96 images for the training, validation, and test sets, respectively. In addition, augmentation was designed to alleviate class imbalance by ensuring a more uniform distribution of IQI wire classes across the datasets.

As shown in **Figure 1**, the IQI inspection criteria for evaluating the radiographic image quality were selected in accordance with the ASTM E747 Class B specifications, and a wire-type IQI was applied. The IQI wires are made of steel, and as presented in **Table 2**, the thickest wire is 0.032 inches (0.84 mm) in diameter. The number of wires that must be identified in the image varies depending on the classification society or shipowner requirements, as well as individual shipyard regulations, all of which influence the evaluation criteria for RT image quality. The more clearly the IQI wire can be identified, the higher is the sensitivity of the image and the better is the quality, which is considered an indicator of the reliability of weld defect detection.

The experimental environment for the model development consisted of an Intel(R) Core™ i7-14700k CPU (20 cores, 28 threads, 3.40 GHz), an NVIDIA GeForce RTX 4060 Ti GPU (8 GB), and 64 GB DDR5 RAM. The implementation was performed using Python (version 3.12.7), with a Jupyter notebook used as the development environment.

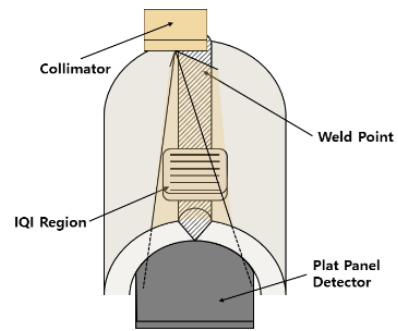


Figure 2: Schematic of the RT imaging setup at weld joint with IQI positioning and detector arrangement

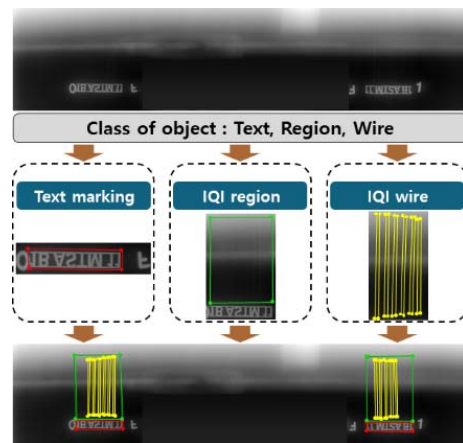


Figure 3: Definition of object classes (text, IQI region, and IQI wire) annotated for training image set

3.2 Experimental Methods

The method used to collect film-based radiographic images for the experiment is shown in **Figure 2**. Typically, the collimator was positioned 20 mm away from the weld joint with a diameter ranging from 20 to 30 mm, and the average radiation output was set to 15 curies. Depending on the pipe size, each RT image was acquired by exposing the specimen for 30–90 s and the resulting images were collected for use in the training, validation, and testing datasets.

The IQI labeling appearing on both the base materials and weld line was performed using the LabelMe tool (version 5.2.1). To accurately identify the positions of the IQI wires, three object classes—test annotations, wire regions, and the wires themselves—were labeled within the commonly identifiable wire areas of the images. As shown in **Figure 3**, each label was assigned to the training set images based on the coordinates, and the two images were combined to generate the final training image.

In this study, we conducted a comparative analysis of two representative object detection models, YOLOv8x and Faster R-

CNN, for automatic IQI reading. YOLOv8x, developed by Ultralytics, is the latest version of a real-time object detection model, offering high adaptability to targets of various sizes and shapes, along with rapid convergence and outstanding computational efficiency [14][15]. By contrast, Faster R-CNN is a two-stage detection model based on a region proposal network (RPN), and is well known for its precise bounding box prediction and high detection accuracy, making it highly effective in analyzing complex images [16].

Figure 4 shows the workflow for the automated IQI detection proposed in this study. The entire set of RT images was divided into 70%, 20%, and 10% for training, validation, and testing, respectively. Data augmentation techniques were applied to the training data to enhance object detection accuracy, with vertical and horizontal flipping as the primary augmentation methods. Alomar *et al.* (2023) and Shorten *et al.* (2019) demonstrated that flipping-based image augmentation methods are effective for images that, unlike characters and digits, are not asymmetric or direction-sensitive [17][18]. This improves the generalization performance of the model, enabling stable IQI recognition under various configurations and orientations. Both models were trained separately based on this procedure. The performance of each trained model was quantitatively evaluated using mAP@50, mAP@50-95, precision, and recall to evaluate detection accuracy and robustness. The validation dataset was used for model training, whereas the test dataset was used for the final evaluation.

The major hyperparameters of each model used in this study are summarized in Table 3. In the case of YOLOv8x, owing to the high resolution of the training images and complexity of the model architecture, the computational resource demands were relatively high. Accordingly, considering the memory capacity of the available GPU, the batch size was set to a maximum feasible

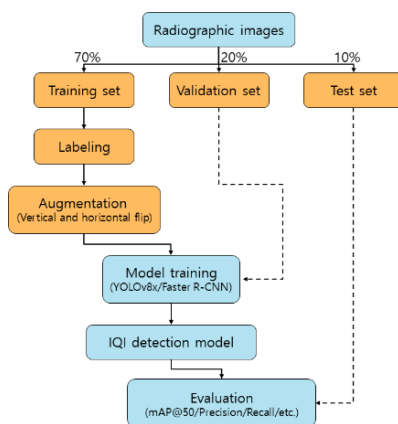


Figure 4: Workflow for detection and IQI wire detection model

Table 3: Hyper parameters of the proposed model

Parameter	Value	
	YOLOv8x	Faster R-CNN
Model	YOLOv8x	Faster R-CNN
Batch size	2	8
Epoch	100	23
Patience	10	10
Learning rate(0)	0.01	0.0001
Learning rate(f)	0.01	0.0001
Image size	1920	
IoU	0.5	

value of two, and the learning rate was fixed at 0.01. For the Faster R-CNN model, a batch size of eight achieved the best performance among the tested values of two, four, and eight, whereas sizes larger than eight could not be applied owing to hardware limitations. To prevent overfitting and underfitting, an early stopping strategy with a patience value of 10 was employed with the stopping point determined by the epoch at which it was triggered for each model. Because many studies regard an intersection over union (IoU) threshold of 0.5, or higher, as a true positive, this study also adopted 0.5 as the criterion.

4. Results and Discussion

When the IQI was positioned on the pipe weld and an RT image was captured, the grey scale intensity varied according to the brightness distribution at each location, as shown in Figure 5. Although a thicker IQI wire generally results in greater sensitivity, the image contrast and shape displayed on the detector may vary depending on the exposure conditions and external environment.

Consequently, even under consistent process parameters and conditions, the resulting RT images exhibit significant variability [19]. As shown in Figure 5(b), when the contrast ratio of the wire indicator was extremely low, it became difficult to distinguish, making detection highly challenging. This resulted in significant time consumption during the production process and reduced the overall process efficiency.

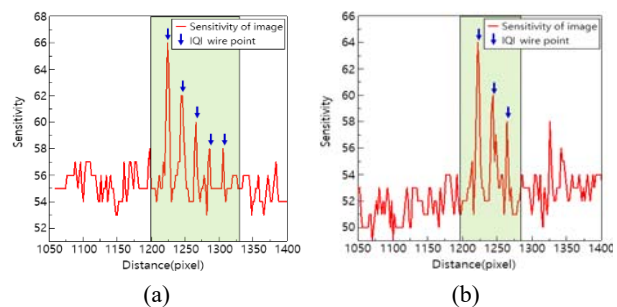


Figure 5: Sensitivity of the IQI wire on welded pipe

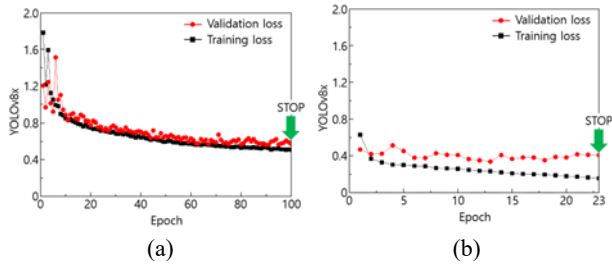


Figure 6: Train loss and validation loss trend according to epoch step. (a) YOLOv8x; (b) Faster R-CNN

The variations in training and validation losses for the models trained based on YOLOv8x and Faster R-CNN are shown in **Figure 6**. Generally, the validation loss tends to be higher than the training loss because it reflects the model’s generalization ability for data not included in the training process. For both models, the training loss gradually decreased as the number of epochs increased, forming a downward-sloping convergence curve, indicating that the training process proceeded appropriately. In terms of validation loss, the YOLOv8x model converged to approximately 0.5705 around epoch 100, whereas Faster R-CNN converged to 0.4076 at epoch 23. Therefore, early stopping was applied at epoch 100 for YOLOv8x and epoch 23 for Faster R-CNN. In these epochs, the final training losses were 0.5067 and 0.155, respectively. Although the validation loss was slightly higher than the training loss, both values converged stably at a certain level, which can be considered as a typical generalization pattern. The presence of a difference between the two values does not necessarily imply under-fitting or overfitting [20]. Rather, it is important to interpret the model performance based on the overall trend of the learning curve rather than relying solely on the absolute level of the validation loss value [21]. The training time for YOLOv8x was 239 min, whereas that for Faster R-CNN, 24 min, was significantly shorter. The difference in training time was due to the structural differences between the two models and the variation in the early stopping points (i.e., the number of epochs) applied during training.

Accordingly, in this study, early stopping was applied at the point where the validation loss stabilized, thereby marking the optimal training endpoint for each model. In particular, the fact that the validation loss of YOLOv8x was slightly higher than that of Faster R-CNN can be interpreted as a result of the more complex neural network architecture, which was more sensitive to the lack of diversity and quantity in the validation image set. In future studies, we plan to obtain additional datasets that reflect the number and shape diversity of the wire classes, retrain the models, and evaluate their generalization performance more precisely.

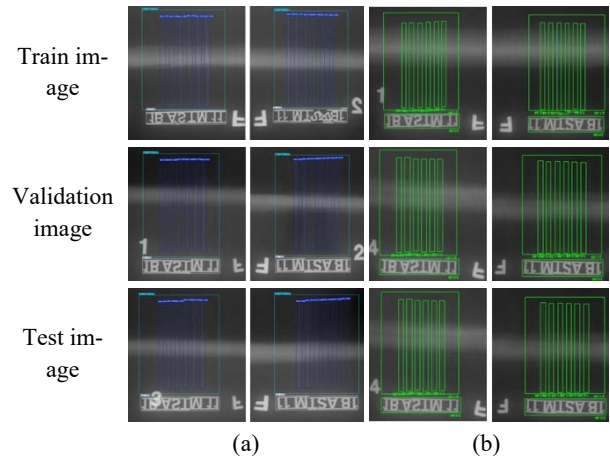


Figure 7: Model performance results of each image set. (a) YOLOv8x; (b) Faster R-CNN

Figure 7 shows the predicted object detection for the training, validation, and test images using the two trained models. In all cases, both the text annotations and IQI regions were accurately detected, satisfying the ASTM Class B standard. **Figure 7(a)** illustrates the detection results of the trained YOLOv8x model, and **Figure 7(b)** presents the results of the Faster R-CNN model. Overall, both models successfully detected a total of twelve wires, with six wires on each side. In particular, the models accurately detected the sixth wire, which had the lowest clarity and was the most susceptible to subjective interpretation.

However, the YOLOv8x model successfully identified the entire wire region even when the object boundaries were not clearly defined. By contrast, Faster R-CNN was able to recognize the boundaries precisely when they were clear; however, some distortion occurred at the wire region boundaries when they were unclear. Furthermore, YOLOv8x was capable of recognizing the entire wire pattern and consistently produced bounding boxes of uniform size even in blurry images, whereas Faster R-CNN was more sensitive to noise. YOLOv8x produced stable bounding boxes across various brightness and noise conditions, whereas Faster R-CNN demonstrated a highly accurate detection performance under high-quality imaging conditions.

The difference in performance between the two models may be attributed to their structural differences: YOLO is a single-stage detector model, whereas Faster R-CNN is a two-stage model. In particular, the YOLO model, which learns by leveraging global features to recognize consistent patterns, demonstrated a distinct advantage in identifying the IQI wire regions in blurred and low-contrast images. Several studies have confirmed that YOLO outperforms conventional CNN-based models in terms of

Table 4: Comparison of object detection performance metrics for each training model

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

accuracy under low-quality imaging conditions. Therefore, when applying the existing analog film method for the automated RT inspection of pipe welds, the YOLOv8x model is more suitable, whereas the Faster R-CNN model is expected to perform more effectively in digital RT environments with high image quality.

Using a set of 150 validation images, the performance of each model was evaluated based on the precision, recall, mAP@50, mAP@50-95, and F1-score metrics, as presented in **Table 4**. YOLOv8x demonstrated excellent performance in the overall metrics with a precision of 0.987, recall of 0.972, mAP@50 of 0.992, and mAP@50-95 of 0.784, whereas Faster R-CNN recorded relatively lower values. The performance gap was particularly evident in the recall and mAP metrics, suggesting that YOLOv8x is more robust in detecting IQI wires of varying quality and shape. Additionally, the lower evaluation scores of Faster R-CNN were likely due to over-detection issues, where the model produced redundant bounding boxes, thereby reducing metric accuracy.

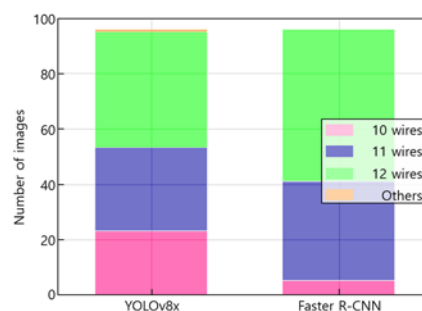
Many studies have compared YOLO, which is a single-stage detector, with CNN-based models classified as two-stage detectors. In general, two-stage detectors demonstrate superior performance in detecting objects within a defined region of interest (ROI). However, performance often varies depending on the image structure and object characteristics. In this study, Faster R-CNN accurately identified IQI regions containing IQI wires, but its wire detection capability was inadequate. This limitation may be attributed to the diverse morphologies of the wires and the susceptibility of conventional detection models or complex two-stage detectors to performance degradation under low brightness and contrast [19][22]. To improve the IQI wire detection performance of two-stage detector-based models, it is necessary to construct an image dataset that incorporates the wire-range signal-to-noise ratio (SNR) and contrast-to-noise ratio (CNR) conditions to enhance the generalization ability of such models [23]. Furthermore, applying image enhancement techniques such as retinex-based contrast enhancement, as proposed by Mirzapour

et al. (2025), could further improve the visualization quality of IQI regions and consequently enhance the detection performance in two-stage detectors [24][25].

In a further analysis of the validation images, the agreement between the ground-truth number of labeled wires and the prediction results of each model was examined. YOLOv8x matched the correct wire count in 82 images (55.7%), whereas Faster R-CNN did so in 39 images (26%). This indicates that YOLOv8x has a strong capability for accurately detecting the number of objects in low-quality RT images.

Additionally, differences in computational efficiency were observed between the two models: YOLOv8x achieved an inference time of 46.0 ms and a processing time of 0.8 ms, whereas Faster R-CNN required 34.79 ms for inference and 0.31 ms for processing. However, because the results of this study were obtained using low-quality analog film RT images, additional comparative analyses are necessary when applying these models to digital RT imaging environments in the future, highlighting the need to validate the model performance across diverse imaging modalities.

Figure 8 shows the IQI wire detection results for a set of 96 test images using both models. As there are two IQI regions in each image, 12 wires must be detected per image to satisfy the ASTM E747 Class B requirements. The YOLOv8x model produces stable detection boxes even under blurred, low-contrast, and poorly defined conditions. However, the boundaries between the individual wires were sometimes unclear, resulting in an accuracy rate of approximately 55% for the exact detection of 12 wires. Faster R-CNN demonstrated lower precision, correctly detecting all 12 wires in 48% of the images, although over-detection occurred in 41 cases. This suggests that the RPN-based region proposal mechanism in a Faster R-CNN is susceptible to bounding-box duplication. Overall, the two models offer complementary strengths, depending on the image quality and inspection objectives.

**Figure 8:** Comparison of wire detection counts by model

This study compared the performance of two object detection models for the automatic detection of IQI wires; however, several limitations should be noted. First, the number and quality of the RT images used were limited, and various imaging conditions and pipe configurations were not sufficiently represented; therefore, the generalization performance could not be fully guaranteed. Second, labeling was conducted by nonexperts, and because of the vagueness of boundaries in blurred images, subjective interpretation may have been introduced into the ground-truth annotations, potentially affecting the accuracy and limiting the objectivity of the model evaluation. Third, owing to hardware constraints, the training conditions (batch size, number of epochs, learning rate, etc.) for YOLOv8x and Faster R-CNN could not be fully standardized, resulting in insufficient fairness in the quantitative comparisons. Therefore, future studies should aim to overcome these limitations by establishing more rigorous experimental protocols to enable comprehensive performance evaluations.

5. Conclusion

In this study, we applied two representative object detection models, YOLOv8x and Faster R-CNN, to detect IQI wires in RT images of pipe welds automatically and conducted a comprehensive performance analysis. In total, 272 RT images were used to train the models, and their qualitative and quantitative detection performances were analyzed using 96 test images under identical conditions.

In the experiments, YOLOv8x demonstrated a robust detection performance despite low-quality and low-contrast images, consistently outputting bounding boxes regardless of image degradation. By contrast, Faster R-CNN exhibited excellent performance in detecting wire boundaries and accurately determining the number of wires under high-quality imaging conditions. However, overdetection was observed with Faster R-CNN, which can be attributed to structural differences between the two models as well as variations in training configurations.

Based on the performance evaluation of the two trained models, YOLOv8x achieved superior results across all quantitative metrics, with a precision of 0.987, recall of 0.972, mAP@50 of 0.992, and mAP@50-95 of 0.784, whereas Faster R-CNN recorded relatively lower values for all categories. These results indicate that YOLOv8x is more robust in handling variability in terms of the image form and quality. The lower performance metrics observed for Faster R-CNN are likely attributable to redundant detections, which reduce the accuracy of the model.

Moreover, the number of labeled wires in the validation images and the prediction results of YOLOv8x and Faster R-CNN exhibited matching rates of 55% and 50%, respectively. This further demonstrates that YOLOv8x exhibits superior performance not only in detection accuracy but also in the quantitative consistency of the predicted wire counts, particularly under low-quality imaging conditions.

In conclusion, YOLOv8x shows excellent robustness with RT images of varying quality, whereas Faster R-CNN exhibits characteristics that are suitable for high-precision quality inspection. Therefore, in the implementation of automated NDT systems, it is recommended to strategically employ the two models complementarily, depending on the radiographic imaging modality (analog or digital).

However, because this study was conducted exclusively using low-quality analog-based RT images, additional validation and supplementary studies are required before these models can be applied to digital RT environments. Future research will focus on improving the generalization and precision of IQI detection by expanding the dataset to include various imaging conditions and IQI wire geometries as well as incorporating expert-driven labeling.

Acknowledgement

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) funded By the Ministry of Trade Industry & Energy(MOTIE, Korea).

Author Contributions

Conceptualization, S. H. Lim and J. K. Park; Methodology, S. H. Lim; Software, S. H. Lim and S. H. Kim; Formal Analysis, S. H. Lim and S. H. Kim; Investigation, S. H. Lim; Resources, S. H. Lim and S. H. Kim; Data Curation S. H. Kim; Writing-Original Draft Preparation, S. H. Lim; Writing-Review & Editing, J. K. Park; Visualization, S. H. Lim and S. H. Kim; Supervision, J. K. Park; Project Administration, J. K. Park; Funding Acquisition, J. K. Park.

References

- [1] Y. B. Ko, G. B. Kim, and K. C. Park, "Soundness evaluation of friction stir welded A2024 alloy by non-destructive test," *Journal of the Korea Society of Marine Engineering*, vol. 37,

- no. 2, pp. 135-143, 2013.
- [2] X. Wang, U. Zscherpel, P. Tripicchio, S. D'Avella, B. Zhang, J. Wu, Z. Liang, S. Zhou and X. Yu, "A comprehensive review of welding defect recognition from X-ray images," *Journal of Manufacturing Processes*, vol. 140, pp. 161-180, 2025.
- [3] H. I. Shafeek, E. S. Gadelmawla, A. A. Abdel-Shafy, and I. M. Elewa, "Assessment of welding defects for gas pipeline radiographs using computer vision," *NDT & E International*, vol. 37, no. 4, pp. 291-299, 2004.
- [4] S. Kim and J. Ahn, "A YOLO-based crop pests detection mobile application for smart farming," *Journal of the Korea Academia-Industrial*, vol. 25, no. 7, pp. 603-610, 2024.
- [5] C. Y. Zhang and D. H. Kim, "Analysis of impact position based on deep learning CNN algorithm," *Transactions of the Korean Society of Mechanical Engineers - A*, vol. 44, no. 6, pp. 405-412, 2020.
- [6] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 779-788, 2016.
- [7] P. Jiang, D. Ergu, F. Liu, Y. Cai, and B. Ma, "A Review of Yolo algorithm developments," *Procedia Computer Science*, vol. 199, pp. 1066-1073, 2022.
- [8] W. Zhiqiang and L. Jun, "A review of object detection based on convolutional neural network," In *2017 36th Chinese Control Conference (CCC)*, pp. 11104-11109, 2017.
- [9] H. Choi, J. Youn, and S. Yoo, "Comparison of changes in classification accuracy by YOLO version using VisDrone-DET training data," *Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography*, vol. 42, no. 6, pp. 713-720, 2024.
- [10] Y. Zhang and Q. Ni, "A novel weld-seam defect detection algorithm based on the s-yolo model," *Axioms*, vol. 12, no. 7, pp. 697-722, 2023.
- [11] M. Lalinia and A. Sahafi, "Colorectal polyp detection in colonoscopy images using YOLO-V8 network," *Signal, Image and Video Processing*, vol. 18, no. 3, pp. 2047-2058, 2024.
- [12] B. ROUSSEL, A. SLIMANI, H. MANSOUR, S. BRAHIMI, I. H. ALI, and S. MANSOUR, "Smart-RT: End-to-end automatic radiographic images analysis through a multi-stage object detection algorithm," *Journal of Nondestructive Testing*, 2025.
- [13] K. Pan, H. Hu, and P. Gu, "Wd-yolo: A more accurate yolo for defect detection in weld x-ray images," *Sensors*, vol. 23, no. 21, pp. 8677-8692, 2023.
- [14] H. Liu, Y. Hou, J. Zhang, P. Zheng, and S. Hou, "Research on weed reverse detection methods based on improved you only look once (yolo) v8: Preliminary results," *Agronomy*, vol. 14, no. 8, pp. 1667-1682, 2024.
- [15] B. Ma, Z. Hua, Y. Wen, H. Deng, Y. Zhao, L. Pu, and H. Song, "Using an improved lightweight YOLOv8 model for real-time detection of multi-stage apple fruit in complex orchard environments," *Artificial Intelligence in Agriculture*, vol. 11, pp. 70-82, 2024.
- [16] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *Advances in Neural Information Processing Systems*, vol. 28, 2015.
- [17] K. Alomar, H. I. Aysel, and X. Cai, "Data augmentation in classification and segmentation: A survey and new strategie," *Journal of Imaging*, vol. 9, no. 2, pp. 46-71, 2023.
- [18] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *Journal of Big Data*, vol. 6, no. 1, pp. 1-48, 2019.
- [19] P. Baniukiewicz and R. Sikora, "Automatic detection of objects in radiographic images," In *Proceedings of the Joint INDS'11 & ISTET'11*, pp. 1-4, 2011.
- [20] S. Salman and X. Liu, "Overfitting mechanism and avoidance in deep neural networks," *arXiv preprint arXiv:1901.06566*, 2019.
- [21] C. Zhang, S. Bengio, M. Hardt, B. Recht, and O. Vinyals, "Understanding deep learning requires rethinking generalization," *arXiv preprint arXiv:1611.03530*, 2016.
- [22] K. Sharma, A. KUMAR, D. K. Banerjee, V. Yadav, and R. Marvel, "Advancements in digital and computed radiography for pipe weld inspection: A focus on sensitivity checks and innovative iqi placement techniques," *International Journal of Creative Research Thoughts*, vol. 13, no. 2, 2025.
- [23] B. Hena, Z. Wei, C. I. Castanedo, and X. Maldague, "Deep learning neural network performance on NDT digital X-ray radiography images: analyzing the impact of image quality parameters—an experimental study," *Sensors*, vol. 23, no. 9, pp. 4324-4341, 2023.
- [24] M. Mirzapour, A. Keshavarz Nasab, A. Movafeghi, E. Yahaghi, "Retinex theory based automated contrast enhancement of gamma radiographic images of pipe welds," *Journal of Nondestructive Evaluation*, vol. 44, no. 3, 2025.
- [25] H. Cheng, H. Jiang, D. Jing, L. Huang, J. Gao, Y. Zhang,

and B. Meng, "Multiscale welding defect detection method based on image adaptive enhancement," *Knowledge-Based Systems*, vol. 327, 2025.