

Detection of Weld Defects in Power Components Using the Attention-Enhanced YOLOv8 Model

어텐션 강화 YOLOv8 모델을 활용한 전력부품 용접 결함 탐지

Ugur Ercelik · Keo Sik Kim · Hyoung-Jun Park · Kyungbaek Kim

에르셀릭 우우르** · 김거식* · 박형준* · 김경백†

Abstract

In transformer power equipment, major welding defects often cause oil leakage due to flaws such as porosity, spatter, undercuts, and overlaps, as well as cracks and leakage resulting from frequent electromagnetic vibrations. These issues typically arise from reduced weld quality, component fatigue, thermal overload, and long-term operational conditions, ultimately leading to shorter equipment lifespan and serious safety risks. To address these challenges, in this paper, we propose an improved YOLOv8 model with a Convolutional Block Attention Module(CBAM) to detect and resolve multiple types of small defects on weld beads. The CBAM module is integrated into the neck part of the YOLOv8 network to enhance target features from both channel and spatial dimensions. The proposed YOLOv8-CBAM model demonstrated the most balanced performance across all metrics, reaching the highest precision (75.2%) and recall (77.6%). Although its mAP@50 (81.0%) was slightly lower than YOLOv5, it still outperformed YOLOv10 and YOLOv11, and its competitive mAP@50-95 (49.0%) confirmed its robustness in detecting defects of varying sizes.

Key Words

Computer Vision, Weld Defect Detection, Industrial AI, Channel Attention, Spatial Attention, YOLOv8

1. Introduction

Welding technology is common in various industrial applications such as manufacturing, infrastructure, aerospace and shipbuilding. Therefore, the quality and strength of weld is vitally important part of industry. This technology ensure strong structural integrity by combining multiple steel or metal using heat and pressure in industrial ares [1]-[2]. During the use of these methods, several defects arise such as porosity, spatter, overlap, and undercut, due to irregularities and complexities in the weld surface and beads and these may cause quite dangerous results depending on the fields of usage [3]. In transformer power equipment, major welding defects include oil leakage caused by poor weld quality such as pinholes, porosity, and lack of fusion, as well as cracks and leakage resulting from frequent electromagnetic vibrations. These defects generally originate from various factors, including reduced weld

quality, material fatigue, thermal overload, and prolonged operational conditions. Such issues not only degrade the reliability of but also shorten its service life and pose significant safety concerns. In these cases, non-destructive testing(NDT), which obstruct damage to the weld material and product, is significantly important. Common non-destructive detection(NDT) of welds contains ultrasonic testing, radiographic, penetration and laser [2][6][8]. However, these methods have some challenges such as ultrasonic testing results visual lackness, radiographic testing cause a health problems, and penetration testing is time-consuming. Also, Laser-based NDT require high equipment costs and high skilled personel and environmental sensitivity [3]. For these purposes, we propose improved YOLOv8 model with Convolutional Block Attention Module for automatic small defect detection model to point out defects on weld beads. The CBAM module handles the defects in terms of enhancing and capturing important

[†] Corresponding Author : Dept. of Artificial Intelligence Convergence, Chonnam National University, Gwangju, South Korea.

E-mail : kyungbaekkim@jnu.ac.kr <https://orcid.org/0000-0001-9985-3051>
Electronics and Telecommunications Research Institute (ETRI), South Korea.
Email : keosikis@etri.re.kr <https://orcid.org/0000-0002-3234-1008>

Email : spacegon@etri.re.kr <https://orcid.org/0000-0002-8258-3224>

^{**} Dept. of Artificial Intelligence Convergence, Chonnam National University
Email: ugurerelik@hotmail.com <https://orcid.org/0009-0007-0235-6499>

Received : Sep. 23, 2025 Revised : Dec. 04, 2025 Accepted : Jan. 02, 2026

Copyright © The Korean Institute of Electrical Engineers

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

features of defects and emphasize. The proposed system aims to provide a fast, accurate and automated detection model. In summary, the main contributions consist of the following:

- 1) A novel small defect detection model (YOLOv8-CBAM) is designed based on YOLOv8 for detecting surface defects on weld beads. By integrating CBAM to neck part of YOLOv8, adaptive attention mechanism provided. Moreover, feature discrimination and spatial focus is enhanced.
- 2) Extensive experiments are conducted on a weld bead defect dataset, and the proposed model is compared with the baseline YOLOv8 and other state-of-the-art detection methods. The results demonstrate that our model achieves superior accuracy, particularly in small defect detection.
- 3) An ablation study is performed to validate the effectiveness of the CBAM integration, showing the contribution of channel and spatial attention to feature representation.
- 4) The study provides practical insights into applying attention-enhanced deep learning models for industrial quality inspection, which can be further extended to other manufacturing processes.

The remainder of this paper is structured as follows: Section 2 provides review of related works, section 3 presents the proposed methodology and contributions, section 4 provides experiments, results and evaluations. Finally, section 5 concludes the study and outlines future research directions.

2. Related Works

2.1 Defect Detection in Welding

Due to the inability of traditional and human-based methods, the improvement of AI-based methodologies has been targeted. Liu et al. [1] proposed lighter and faster YOLOv5 model to point out the differences of shape and multiscale problem on multi defect types. RMF module is integrated to extract parameter-based and parameter-free multiscale information and EFE module to improve the performance of detection network. Block et al. [2] proposed a novel dataset and experiments in welding. The LoHi-WELD dataset solve shortage of vast and reliable datasets by utilizing augmentation methods and CNN, YOLO-based models. A important contribution was made by Wang et al. [3], who improved Yolo with MSAPF module to improve low accuracy caused by interference information on weld defect detection. The proposed model aimed to eliminate interference information and enhance necessary features in each

scale in the same time by integrate multiscale alignment fusion(MSAF) with parallel features filtering (PFF) modules.

2.2 YOLOv8 Based Defect Detection

The development of YOLOv8 has made it easier to identify objects which have complexity in surfaces. Chen et al. [10] proposed a lightweight YOLOv8-OCHD wood surface defect detection method to handle the drawbacks of manual defect detection and the challenges of traditional visual inspection algorithms such as high missed detection rates, slow detection speeds. Mao et al. [11] presents an enhanced YOLOv8 based framework which integrated novel attention mechanisms and advanced architectural modules to improve detection accuracy and robustness to point out fabric defect detection in textile industry. Their framework incorporate the SimAM attention mechanism with the SPPF module and adopts an optimized Dilation-wise Residual structure in the backbone. Another research in scope of defect detection with YOLOv8 proposed by Zhang et al. [12]. They designed a novel GDM-YOLO model, tailored for steel surface defect detection tasks. In the backbone network, they use Space-to-Depth- Ghost Convolution(SPDG) downsampling module to aim at minimizing information loss during downsampling operations. Secondly, their research introduces the C2f-Dilated-Reparam-Block(C2f-DRB) module to leverage reparameterization and larger kernel convolutions for enhancing feature extraction capabilities without inference costs.

3. Proposed Attention-Enhanced YOLOv8 with CBAM

3.1 CBAM

CBAM is a lightweight attention module that refines feature maps by sequentially applying channel attention and spatial attention. The channel attention mechanism helps the model emphasize ‘what’ is important, while spatial attention guides the model to focus on ‘where’ the relevant information is located. The channel attention module exploits global average pooling and max pooling operations, followed by shared multilayer perceptrons(MLPs), to compute an attention map that highlights the most discriminative feature channels. Subsequently, the spatial attention module utilizes convolution operations over concatenated average-pooled and max-pooled feature maps along the channel axis to generate a spatial attention map. This process allows the network to selectively focus on relevant

Table 1. Detection results of each defect types (P: Precision, R: Recall)

Model	Overlap				Porosity				Spatter				Undercut			
	P	R	mAP 50	mAP 50-95	P	R	mAP 50	mAP 50-95	P	R	mAP 50	mAP 50-95	P	R	mAP 50	mAP 50-95
Yolov5	83.3	90.0	93.9	74.2	69.2	70.6	77.6	40.8	70.4	68.7	76.2	43.5	73.4	74.5	80.2	42.1
Yolov8	80.8	89.8	94.1	73.3	77.5	69.9	78.2	40.5	70.6	66.9	74.8	42.2	67.8	70.7	75.8	39.1
Yolov10	84.8	89.3	93.8	73.5	67.3	67.8	73.6	39.1	70.9	68.9	74.3	44.0	75.0	65.5	74.6	37.7
Yolov11	85.6	87.3	94.0	74.4	72.4	62.9	73.2	39.7	65.9	73.9	76.0	43.6	69.3	74.9	75.4	38.5
Yolov8-CBAM	93.0	89.4	95.1	75.0	72.7	77.3	77.5	40.0	71.6	74.0	76.2	44.1	72.2	70.7	75.9	37.9

different pyramid levels, CBAM modules were applied before feeding the features into subsequent C2f blocks. This integration enables the network to adaptively recalibrate multiscale features during the feature aggregation process. By doing so, redundant or noisy features are suppressed, while important spatial cues are preserved across pyramid levels (P3, P4, P5).

3.2.3 Overall Benefits

The dual integration of CBAM in both the backbone and neck ensures that attention is applied not only during the feature extraction stage but also during the feature fusion stage. This comprehensive attention enhancement allows the network to capture subtle defect regions more effectively, while maintaining robustness against complex backgrounds typically observed in weld bead surfaces.

3.2.4 Data Collection and Preprocessing

Weld images which contains defects were obtained by recording videos of welding plates. The recorded videos were then cropped to extract image sequences corresponding to weld regions by using a custom OpenCV-based method (Figure 3). The final weld defect dataset consists of 1,539 images, divided into 1,074 training images, 309 validation images, and 156 testing images. The dataset covers four major defect categories: overlap, spatter, porosity, and undercut. The defects consists of small-scale, low-contrast defect regions that often occupy less than 1-3% of the image area. Figure 4 shows representative examples of four types of small defects, highlighting the challenges related to scale variation and subtle texture differences. Also, to improve the model's robustness and generalization, data augmentation techniques were employed. Basic augmentations included horizontal flipping (p=0.5), hue adjustment (± 0.015), saturation adjustment (± 0.7), and value adjustment (± 0.4). In addition, advanced augmentation strategies

such as MixUp (p=0.2) and Mosaic (p=1.0) were applied to further enhance the diversity of training samples. Each model was trained with parameters of 300 epoch, 20 patience, and 1024 image size.

4. Experiments

4.1 Metrics

To evaluate the performance of the proposed and baseline models, standard object detection metrics were employed, including Precision, Recall, mAP50, and mAP50-95. Precision measures the fraction of correctly detected objects among all actual objects. It evaluates the ability of the model to find all relevant instances. Formally:

$$Precision = \frac{TP}{TP + FP} \quad (6)$$

where TP is True Positive (correctly detected objects), FP is False Positive (incorrect detections).

Recall measures the fraction of correctly detected objects among all actual objects. It evaluates the ability of the model to find all relevant instances. Formally:

$$Recall = \frac{TP}{TP + FN} \quad (7)$$

where FN is False Negatives (missed objects). High recall indicates that the model successfully detects most of the objects in the dataset. Average Precision (AP) summarizes the precision-recall curve for a given Intersection over Union (IoU) threshold. The mean Average Precision (mAP) is the average of AP over all classes. IoU between a predicted bounding box B_p and ground truth B_{gt} is defined as:



Fig. 5 Visualization Results.

Table 2. Detection results of overall ('all' class)

Model	Precision	Recall	mAP50	mAP50-95
Yolov5	74.1	76.0	82.0	50.0
Yolov8	74.2	74.3	80.8	48.8
Yolov10	74.5	72.9	79.1	48.6
Yolov11	73.3	74.8	79.7	49.0
Yolov8-CBAM	75.2	77.6	81.0	49.0

Table 3. Ablation study

Model	Precision	Recall	mAP50	mAP50-95
Yolov8	74.2	74.3	80.8	48.8
Yolov8-C2f_CBAM (Backbone only)	72.0	74.6	80.3	47.6
Yolov8-CBAM	75.2	77.6	81.0	49.0

$$IoU = \frac{B_p \cap B_{gt}}{B_p \cup B_{gt}} \quad (8)$$

AP at given IoU threshold (e.g., 0.5) is:

$$AP = \int_0^1 P(R) dR \quad (9)$$

where $P(R)$ is precision at recall R .

mAP50 is the mean AP across all classes at IoU threshold 0.5:

$$mAP_{50} = \frac{1}{N} \sum_{i=1}^N AP_i^{IoU=0.5} \quad (10)$$

mAP50-95 is the mean AP averaged over multiple IoU thresholds from 0.5 to 0.95 in steps of 0.05:

$$mAP_{50:95} = \frac{1}{N} \sum_{i=1}^N \frac{1}{10} \sum_{t=0.5}^{0.95} AP_i^{IoU=t} \quad (11)$$

where N is the number of classes.

Table 4. Inference speed analysis

Model	FPS	Inference Time(ms)
Yolov8	38.24	6-9 ms
Yolov8-CBAM	32.41	11-14 ms

4.2 Experimental Results

To validate the effectiveness of the proposed YOLOv8-CBAM model, extensive experiments were conducted. In Table 1, the results reveal that the integration of CBAM significantly enhances the detection capability of YOLOv8, especially for complex defect types. For the overlap, YOLOv8-CBAM achieved the highest precision (93.0%) and mAP@50-95 (75.0%), outperforming all baseline models. Similarly, for spatter detection, YOLOv8-CBAM reached superior mAP@50 (74.0%) and mAP@50-95 (44.1%), demonstrating improved small-defect detection. In the case of porosity, YOLOv8 and YOLOv8-CBAM yielded comparable results, while YOLOv8-CBAM still offered competitive precision and recall. For undercut, YOLOv5 achieved slightly higher mAP@50 (80.2%), but YOLOv8-CBAM maintained consistent performance across all metrics. Also, Table 2 presents the overall detection performance of the evaluated models. YOLOv5 achieved the highest mAP@50 (82.0%), while YOLOv8 closely followed with 80.8%. However, the proposed YOLOv8-CBAM model demonstrated the most balanced performance across all metrics, reaching the highest precision (75.2%) and recall (77.6%). Although its mAP@50(81.0%) was slightly lower than YOLOv5, it still outperformed YOLOv10 and YOLOv11, and its competitive mAP@50-95 (49.0%).

4.3 Ablation Study

To further investigate the contribution of the CBAM modules in the proposed YOLOv8-CBAM model, an ablation study was conducted by systematically enabling and disabling the attention components:

1. YOLOv8 (baseline): The original YOLOv8 architecture without any CBAM integration.

2. YOLOv8-C2f_CBAM (Backbone only): CBAM modules embedded within the backbone by replacing standard C2f blocks with C2f_CBAM blocks, while the neck remains unchanged.
3. YOLOv8-CBAM (Backbone+Neck): The full proposed model.

In Table 3, The experimental results present that when CBAM was integrated only into the backbone, the performance slightly decreased, indicating that the backbone alone is not sufficient to fully exploit the benefits of the attention mechanism. In contrast, the full proposed model achieved the best results with 75.2% precision, 77.6% recall, 81.0% mAP50, and 49.0% mAP50-95. Although both YOLOv8 and YOLOv8-CBAM achieved comparable precision and mAP scores, YOLOv8-CBAM demonstrated a higher recall. This result implies that YOLOv8-CBAM is able to detect a greater proportion of true weld defects, thus reducing the likelihood of missing potential defects. In safety-critical domains such as weld quality assurance, higher recall is essential to minimize the risk of undetected defects, thereby enhancing the overall reliability and safety of the inspection process. Sample detection results on test images are presented in Figure 5.

4.4 Inference Speed Analysis

To evaluate the impact of the CBAM integration on real-time usability, we measured the inference speed(frames per second, FPS) of both the YOLOv8 and the proposed YOLOv8-CBAM model in Table 4. The standard YOLOv8 model achieved 38.24 FPS, and after integrating the CBAM module, the inference speed slightly decreased to 32.41 FPS due to additional channel and spatial attention computations. To measure this, we performed a standardized FPS test by running each model on a fixed 1024x1024 input. FPS was then computed as the number of processed frames divided by the elapsed time. Despite increase in computational cost, the proposed model still maintains real-time performance(≥ 30 FPS), confirming that the enhancement in detection accuracy does not compromise practical deployability.

5. Conclusion and Future Works

In this study, we proposed YOLOv8-CBAM model by integrating the Convolutional Block Attention Module (CBAM) into both the backbone and neck of the YOLOv8 architecture to improve feature discrimination and spatial focus, therefore achieving more reliable detection of weld surface defects. The experimental results demonstrated that YOLOv8-CBAM achieved superior performance compared to baseline YOLO

models, particularly in terms of precision and recall, while also maintaining competitive mAP values across different types. Despite these promising results, there are still directions for future research. First, additional advanced modules such as Transformer-based or lightweight adaptive attention blocks can be explored to further enhance the representation power of the model. Second, extending the detection framework to multi-modal approaches(e.g., combining ultrasonic signals, infrared, or X-ray images with visual data) may provide more comprehensive insights for weld quality assessment.

Acknowledgements

This work was supported by an Electronics and Telecommunications Research Institute (ETRI) grant funded by the Korean government [26ZK1100, Honam region regional industry-based ICT convergence technology advancement support project].

This work was supported by the Institute of Information & Communications Technology Planning & Evaluation(IITP)-Innovative Human Resource Development for Local Intellectualization program grant funded by the Korea government(MSIT)(IITP-2025-RS-2022-00156287)

References

- [1] M. Liu, Y. Chen, J. Xie, L. He and Y. Zhang, "LF-YOLO: A Lighter and Faster YOLO for Weld Defect Detection of X-Ray Image," in *IEEE Sensors Journal*, vol. 23, no. 7, pp. 7430-7439, 1 April, 2023.
- [2] S. Biasuz Block, R. Dutra da Silva, A. Eugenio Lazzaretti and R. Minetto, "LoHi-WELD: A Novel Industrial Dataset for Weld Defect Detection and Classification, a Deep Learning Study, and Future Perspectives," in *IEEE Access*, vol. 12, pp. 77442-77453, 2024.
- [3] G. -Q. Wang et al., "Yolo-MSAPF: Multiscale Alignment Fusion With Parallel Feature Filtering Model for High Accuracy Weld Defect Detection," in *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1-14, 2023.
- [4] Y. Gao, P. Zhong, X. Tang, H. Hu and P. Xu, "Feature Extraction of Laser Welding Pool Image and Application in Welding Quality Identification," in *IEEE Access*, vol. 9, pp. 120193-120202, 2021.
- [5] Y. Zhang, J. Xiao, Z. Zhang and H. Dong, "Intelligent Design of Robotic Welding Process Parameters Using Learning-Based Methods," in *IEEE Access*, vol. 10, pp. 13442-13450, 2022.
- [6] K. Kim, K. S. Kim and H. -J. Park, "Multi-Branch Deep Fusion Network-Based Automatic Detection of Weld Defects Using Non-Destructive Ultrasonic Test," in *IEEE Access*, vol. 11, pp. 114489-114496, 2023.
- [7] K. Tian, J. Peng, X. Zhang, Q. Zhang, T. Wang and J. Lee,

“Multifeature Fusion Imaging Based on Machine Learning for Weld Defect Detection Using Induction Scanning Thermography,” in *IEEE Sensors Journal*, vol. 24, no. 5, pp. 6369-6379, 1 March1, 2024.

- [8] A. Mohammed and M. Hussain, “Advances and Challenges in Deep Learning for Automated Welding Defect Detection: A Technical Survey,” in *IEEE Access*, vol. 13, pp. 94553-94569, 2025.
DOI: 10.1109/ACCESS.2025.3574083.
- [9] S. Yu, J. Hu, J. Hong, H. Zhang, Y. Guan and T. Zhang, “Optimal Imaging Band Selection for Laser-Vision System Based on Welding Arc Spectrum Analysis,” in *IEEE Sensors Journal*, vol. 25, no. 2, pp. 2534-2546, 15 Jan.15, 2025.
- [10] Z. Chen, J. Feng, X. Zhu and B. Wang, “YOLOv8-OCHD: A Lightweight Wood Surface Defect Detection Method Based on Improved YOLOv8,” in *IEEE Access*, vol. 13, pp. 84435-84450, 2025.
- [11] Y. Mao, G. Wang, Y. Ma and X. Gui, “Enhancing Fabric Defect Detection With Attention Mechanisms and Optimized YOLOv8 Framework,” in *IEEE Access*, vol. 13, pp. 96767-96781, 2025.
- [12] T. Zhang, H. Pang and C. Jiang, “GDM-YOLO: A Model for Steel Surface Defect Detection Based on YOLOv8s,” in *IEEE Access*, vol. 12, pp. 148817-148825, 2024.
- [13] Ji-Soo Kim and Kyungyong Chung, “YOLO Based Crack Detection of Structures with Edge Detection Process.” *The Transactions of the Korean Institute of Electrical Engineers*, 73(12), 2391-2397, 2024.

저자소개



Ugur Ercelik

He received his B.S. degree in Information Systems Engineering from Kocaeli University, Kocaeli, Republic of Turkiye, in 2022. He is currently pursuing the M.S. degree in Artificial Intelligence Convergence at Chonnam National University and working as a student (part-time) researcher in Electronics and Telecommunications Research Institute (ETRI). His research interests include computer vision based models and systems on industry and autonomous vehicles, 2d-3d object detection, image classification, neural networks and deep learning.



김거식(Kim Keo Sik)

He received the B.S, M.S., and Ph.D. degrees in Department of Electronics Engineering from Chonbuk National University in 2004, 2006 and 2011, respectively. He is presently the principal researcher of Electronics and Telecommunications Research Institute (ETRI). His research interests include machine learning, agentic AI, optical imaging system,



박형준(Park Hyoung-Jun)

He received the B.S, M.S., and Ph.D. degrees in Department of Electronics Engineering from Chonbuk National University in 2002, 2004 and 2009, respectively. Since 2013, he has been a Principal Researcher with the Optical ICT Convergence Research Section, Electronics and Telecommunications Research Institute (ETRI), South Korea, and he is currently serving as the Director of the Optical ICT Convergence Research Section. His current research interests include AI-optical engine convergence, intelligent optical sensor systems, and image analysis.



김경백(Kim Kyungbaek)

Kyungbaek Kim received the B.S., M.S., and Ph.D. degrees in electrical engineering and computer science from Korea Advanced Institute of Science and Technology (KAIST), South Korea, in 1999, 2001, and 2007, respectively. He is currently a Professor with the Department of Artificial Intelligence Convergence, Chonnam National University. Previously, he was a Postdoctoral Researcher with the Department of Computer Sciences, University of California, Irvine, CA, USA. His research interests include intelligent distributed systems, software-defined networks/infrastructure, big data platforms, GRID/cloud systems, social networking systems, AI applied to cyber-physical systems, blockchain, and other distributed systems issues. He is a member of ACM, IEICE, KIISE, KIPS, KICS, KIISC, and KISM.