Proceedings of CONF-MLA 2026 Symposium: Intelligent Systems and Automation: AI Models, IoT, and Robotic Algorithms
DOI: 10.54254/2755-2721/2026.LD38191

A Review of Post-weld Inspection Technology

Siqi Liu

Shanghai Hongwen School, Shanghai, China
sigridliu0903@163.com

Abstract. Welding robots are now extensively used in shipbuilding, automobile, and aviation industries. Defects are prone to occur after welding, and traditional non-destructive testing like ultrasonic and radiographic testing, depend on manual judgement, resulting in detection lag and difficulty in achieving online closed-loop control on production lines. This paper focuses on robot-mounted vision-based post-weld inspection, dividing the techniques into non-destructive and destructive testing. It outlines a three-level processing paradigm for vision-based inspection: low-level preprocessing, mid-level defect recognition and segmentation, and high-level quality diagnosis. The paper systematically summarizes three deep-learning technical routes: defect recognition and classification, 3D geometric measurement of weld seams, and multimodal temporal intelligent diagnosis. It also compares the application characteristics and applicable scenarios of mainstream models such as Convolutional Neural Networks (CNNs), You Only Look Once (YOLO), U-Net, and Detection Transformer (DETR). Various interference factors in industrial environments are analyzed. Finally, future technical directions are discussed in four key aspects, providing a comprehensive theoretical reference for the development of intelligent welding inspection algorithms and the engineering integration of robotic systems.

Keywords: post-weld inspection, welding, defect detection

1. Introduction

With the advancement of intelligent manufacturing, welding robots have been widely adopted. Welding processes are prone to defects such as pores and cracks. Traditional inspection methods including ultrasonic and radiographic testing suffer from low efficiency and reliance on manual judgement. Robotic post-weld inspection integrating laser vision sensors and deep learning models such as Convolutional Neural Networks (CNNs), You Only Look Once (YOLO) and ResNet enables the transition from offline sampling qualitative inspection to online quantitative intelligent diagnosis. CNNs and Transformer networks can extract multi-scale features to boost defect detection performance. Nevertheless, disturbances such as strong ambient light and welding fumes remain prominent, and challenges persist in detecting tiny, low-contrast defects. This paper reviews recent advances, technical challenges and future trends of vision-based robotic post-weld inspection driven by deep learning.

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2026.LD38191

2. Technical landscape of vision-based inspection

2.1. Classification by detection principle

Post-weld quality inspection can be broadly divided into destructive testing and non-destructive testing (NDT). Destructive testing evaluates the mechanical properties of welded joints by destroying workpiece specimens, including tensile, bending, and metallographic corrosion tests. It is only used for process qualification sampling and cannot be applied to batch inspection of finished products. Non-destructive testing assesses defects without damaging the workpiece; mainstream solutions include vision-based inspection, ultrasonic testing, and radiographic testing. The performance comparison of various sensors is shown in Table 1.

Table 1. Comparison of major external sensor perception characteristics

Sensor Modality	Data Output	Typical Defects	Advantages	Limitations
2D Vision	RGB / grayscale	Cracks, undercut, spatter, poor bead	Fast, low-cost	No depth, light-sensitive
3D Vision	Point cloud	Reinforcement, width, undercut, misalignment	Accurate 3D	Expensive, reflective surfaces
Radiographic testing	X-ray / CT images	Porosity, inclusions, lack of penetration	Internal defects	Radiation, high cost
Ultrasonic testing	A/B/C-scan	Cracks, LOF, penetration, depth	Deep inspection	Couplant, operator-dependent

Vision sensors offer fast response, no radiation, and easy integration onto robot end-effectors, making them well suited for closed-loop integration of welding, inspection, and process adjustment. Traditional 2D vision captures only planar images and is mostly used for surface defect screening. In contrast, 3D structured-light vision reconstructs the 3D weld profile via laser triangulation, enabling precise quantification of geometric parameters.

Traditional vision relies on manually designed features for defect recognition, requiring preprocessing steps such as image filtering, binarisation, and edge extraction, and uses pixel-based metrics like mean-squared error to judge defects – resulting in poor feature adaptability. Deep learning, by contrast, can automatically learn shallow textures and deep semantic features without manual design, significantly improving the stability of defect recognition under complex working conditions, and has now become the core analytical tool for vision-based inspection.

2.2. Hierarchical classification of inspection tasks

In intelligent industrial vision-based inspection, a complete information processing loop is established from raw sensor data to final decision-making. This loop comprises a hierarchy of low-level processing, mid-level processing, and high-level decision-making, each with distinct objectives, core technologies, and expected outputs.

Low-level processing: The goal is signal enhancement and feature extraction. Taking laser scanning for 3D information acquisition as an example, raw laser scan data often contain noise and local data loss due to specular reflection and insufficient sensitivity at the laser line ends. Existing studies employ data cropping and weighted regression filtering to remove outliers. For image over-

DOI: 10.54254/2755-2721/2026.LD38191

exposure and texture blurring caused by arc light and smoke, adaptive image enhancement and temporal background modelling are used to suppress transient disturbances.

Mid-level processing: This stage determines whether defects exist in images or point clouds and provides their approximate locations. Common methods include sliding-window classifiers like Support Vector Machines and traditional CNNs, and the widely used end-to-end object detection networks YOLO and Faster R-CNN. For tasks requiring precise characterization of defect shape and extent, such as calculating crack length and pore area, semantic segmentation models such as U-Net and the DeepLab series are employed to delineate defect regions at the pixel or point-cloud level. Based on localization and segmentation, precise measurements of defect size, depth, area, volume, etc., are performed, transitioning from qualitative to quantitative assessment and providing numerical support for subsequent quality evaluation.

High-level processing: The objective is semantic interpretation of the collected information to provide useful and reasonable feedback. According to defect type, size, quantity, and distribution, the weld is classified into acceptable, conditionally acceptable, or unacceptable grades. This requires the system not only to identify defects but also to perform quality grading based on their type, size, and distribution.

3. Detailed technical explanation of deep-learning methods

3.1. Defect recognition and classification

Post-weld defects exhibit great diversity in morphology and scale, ranging from micrometer-sized pores to centimeter-long cracks, with vast differences in shape and size. Weld images are often subject to strong noise from residual arc light, smoke, spatter, and oxide scale texture. Moreover, the sample distribution is extremely imbalanced – in actual production, qualified welds far outnumber defective ones, and the quantities of different defect types vary significantly. As early as 2014, researchers explored combining CNNs with RBF networks for defect recognition in radiographic weld images, achieving over 91% accuracy in classifying six defect types, preliminarily verifying the effectiveness of automatic feature extraction via deep learning in post-weld inspection [1].

Current deep-learning methods for defect recognition and classification have developed into a highly systematized technical framework, mainly divided into three directions: high-precision detection, pixel-level segmentation, and robustness enhancement for real-world challenges. A representative recent study in real-time defect detection is DETR, which, based on the Transformer encoder-decoder architecture, achieves true end-to-end detection, maintaining speed while improving the detection rate of small defects and eliminating reliance on manually designed anchor boxes [2].

Beyond traditional geometric and color transformations, Generative Adversarial Networks (GANs) offer a solution to the scarcity of defect samples. Defect generators based on StyleGAN or conditional GANs can synthesize virtual defect images with varying morphologies, sizes, and textures, thereby constructing more balanced and diverse training datasets and significantly improving the model's ability to recognize rare defects [3]. Alternatively, pre-training on large generic datasets such as ImageNet or large-scale industrial defect datasets enables the model to acquire powerful general feature extraction capabilities, followed by fine-tuning with a small number of welding defect samples, allowing the model to recognize new welding defects with only a few examples. Transfer learning bypasses data scarcity by directly extracting high-level semantic features, effectively narrowing the semantic gap and greatly enhancing the precision and practicality of welding defect detection [4] .

DOI: 10.54254/2755-2721/2026.LD38191

3.2. Precise measurement of weld geometric parameters

The goal of weld geometric measurement is to output sub-millimeter-accuracy indicators such as weld width, reinforcement, misalignment, and leg length. Core challenges include strong reflections from workpieces causing point-cloud voids, spatter interfering with laser stripe extraction, and the difficulty of fitting geometric contours for fillet welds.

The current mainstream method is active-vision-based 3D reconstruction and measurement. A line laser or multi-line structured light is projected onto the weld surface, and a binocular camera or a calibrated monocular camera captures the deformation of the light stripe on the weld surface. Using the triangulation principle, the 3D coordinates of each point on the stripe are calculated, yielding the cross-sectional profile or 3D point cloud of the weld. This approach is insensitive to ambient light interference and adapts well to dim or strong residual arc-light environments in welding sites [5]. For example, Gaussian background modelling is used for temporal filtering to suppress transient noise. Then, a lightweight segmentation network, which incorporates uncertainty estimation and self-attention mechanisms, is applied to accurately extract laser stripes. Subsequently, through 3D reconstruction and optimal viewpoint projection, combined with temporal filtering closed-loop correction of corner points, this scheme achieved a reconstruction error of ≤0.1 mm in real welding scenarios, with corner recognition success rate improved by about 13.3% over traditional methods [6]. Gong et al. (2025) proposed a deep-learning-based point cloud processing method that first semantically segments the point cloud to distinguish weld areas, base metal areas, etc., and then applies geometric algorithms such as Random Sample Consensus (RANSAC) and least-squares fitting to specific regions to fit lines or curves for dimension calculation. This method combines the robustness of deep learning with the precision of traditional geometry [7].

In addition, multi-view stitching is employed for large workpieces. Robots move to scan multiple weld segments, and preliminary stitching is performed using encoder poses, followed by fine adjustment with point-cloud registration algorithms like Iterative Closest Point (ICP), forming a complete 3D digital workpiece. This seamlessly integrates locally measured data acquired from different positions and perspectives into a unified global 3D coordinate system, eliminating cumulative errors [8].

4. Conclusion

At present, robot-based post-weld vision-based inspection has evolved from manual sampling inspection to deep-earning-driven intelligent online inspection. Leveraging networks such as CNNs and Transformers, combined with active structured light and multi-scale feature fusion, it addresses challenges including imaging noise, sample scarcity, and real-time detection requirements. At the same time, it achieves closed-loop integration of perception and welding control, enabling automatic path correction and process parameter optimization, and, through multi-view autonomous robot scanning, completes full-component defect quantification. The inspection data can be interfaced with digital twin systems to support process iteration.

However, several limitations remain. Multimodal fusion is still largely based on simple data concatenation and lacks unified mechanisms for cross-modal semantic and physical-causal reasoning. End-to-end deep-learning models also suffer from the black-box problem, which limits the interpretability and reliability of their decisions in industrial applications. In addition, simulated data still exhibit a physical and semantic gap from real welding scenes.

Future developments will focus on three breakthroughs: developing cross-modal Transformer and knowledge-graph architectures to achieve deep causal correlation of multi-sensor information;

DOI: 10.54254/2755-2721/2026.LD38191

building virtual-real integration systems that calibrate simulation environments with a small amount of real data to narrow the simulation-to-reality gap; and combining causal reasoning, digital twins, and embedded explainable AI to construct a fully autonomous, predictive, and deeply integrated flexible intelligent welding inspection system.

References

[1] Yu, Y., et al. (2014). Defect recognition method for radiographic images based on deep learning network. *Journal of Scientific Instrument, 35*(9), 2012–2019.
[2] Liu, X., et al. (2023). Anchor-free object detection methods and their application progress in complex scenes. *Acta Automatica Sinica, 49*(7), 1369–1392.
[3] Liu, L., et al. (2019). Defective samples simulation through adversarial training for automatic surface inspection. *Neurocomputing, 360*, 230–241.
[4] Ajmi, C., Zapata, J., Elferchichi, S., Zaafouri, A., & Laabidi, K. (2020). Deep learning technology for weld defects classification based on transfer learning and activation features. *Advances in Materials Science and Engineering, 2020*, 1–16.
[5] Shi, Y., et al. (2014). Geometric parameter measurement system for overhead contact lines based on binocular vision. *Infrared and Laser Engineering, 43*(6), 1936–1942.
[6] Feng, X., et al. (2025). 3D reconstruction and corner recognition of welds in complex environments. *Journal of Tsinghua University (Science and Technology), 65*(10), 1980–1991.
[7] Gong, Y., et al. (2025). Robot weld milling guidance based on surface structured-light visual inspection. *Computer Integrated Manufacturing Systems, 31*(8), 2816–2828.
[8] Yin, X., et al. (2021). A survey of medical image segmentation techniques based on improved U-Net structure. *Journal of Software, 32*(2), 519–550.