Available Online: 19 August 2026 DOI: 10.54254/2977-3903/2026.36215

Key technologies and recent advances in online monitoring of welding quality

Yazheng Cheng

Faculty of Innovation Engineering, Macau University of Science and Technology, Macau, China

3545280357@qq.com

Abstract. Online monitoring of welding quality is a key technology in intelligent manufacturing and process closed-loop control. Its core essence is the paradigm shift from passive post-weld inspection to in-process active perception, real-time diagnosis, and closed-loop feedback. Currently, technological development in this field advances along three main paths: first, vision-based detection technology, which provides process information by analyzing weld pool morphology, weld geometry, and surface defects; second, multi-modal sensor-based detection technology, which utilizes acoustic emission, arc/spectrum, thermal imaging, and ultrasonic methods to reflect the physical state and internal changes of the welding process; third, multi-source information fusion technology, which integrates the complementary advantages of multi-source heterogeneous sensing data to enhance diagnostic accuracy and robustness. This paper systematically reviews online monitoring technologies for welding quality: first, it analyzes single-modal detection methods based on vision and multi-sensing; then, taking fusion hierarchy as the dimension, it organizes the technical framework of data-level, feature-level, and decision-level fusion, with emphasis on the frontier progress of deep learning-driven feature-level fusion; finally, it envisions future development trends and key challenges from three dimensions—advanced sensing innovation, large-model enablement, and digital twin.

Keywords: machine vision, multi-modal sensing, information fusion, deep learning, online monitoring

1. Introduction

Welding is a key joining technology in equipment manufacturing, and its quality directly determines the structural safety and service life of final products. In high-end manufacturing fields such as aerospace, rail transit, shipbuilding, and new energy vehicles, the requirements for weld quality are increasingly stringent, and any minor welding defect may lead to catastrophic consequences. The traditional quality control mode primarily relies on post-weld offline inspection methods such as radiographic testing, ultrasonic testing, and penetrant testing. Such post-weld inspection approaches are not only inefficient but also unable to detect and prevent the production of batch defects in a timely manner, making it difficult to meet the demands of modern intelligent manufacturing for high efficiency and zero defects [1].

Against this background, the welding quality control paradigm is undergoing a profound transformation: evolving from "post-weld inspection" toward a closed-loop monitoring paradigm of "in-process perception, online diagnosis, and real-time regulation" [2]. Online welding quality monitoring technology acquires and analyzes multi-dimensional signals during the welding process in real time, enabling instant assessment of weld pool dynamics, weld formation, and defect initiation, thereby providing decision support for process adaptive optimization and real-time intervention. Here, "monitoring" emphasizes the passive perception and recording of process states, whereas "control" emphasizes active diagnosis, decision-making, and feedback control based on perceptual information. This transition from monitoring to control marks the shift of the welding quality assurance system from open-loop to closed-loop [3].

Early online monitoring research primarily drew on the ideas of offline inspection, attempting to introduce a single sensor into the welding process. This laid the technical foundation for the field and gave rise to two dominant signal sources. The first category is visual signals, which acquire image information of the weld pool, arc, and weld surface through industrial cameras, high-speed cameras, and structured light sensors, with the advantages of high information dimensionality and high intuitiveness [4]. The second category is the physical signals inherent to the welding process, where sensors acquire multi-modal physical quantities such as acoustic emission, arc electrical parameters, spectra, and temperature fields, directly reflecting the physical state of the welding process and being more sensitive to sub-surface and internal defects [5, 6]. However, as research deepened, the limitations of single sensing sources gradually emerged: visual signals are severely affected by intense arc light and fumess interference, and it is difficult to perceive internal defects; although physical signals can reflect the essential characteristics of the process, they suffer from low signal-to-noise ratio and difficult signal decoupling. This fundamental contradiction has driven the rise of multi-source information fusion technology, which integrates the complementary advantages of multi-source heterogeneous sensing information to overcome the inherent limitations of single sensing [7].

Meanwhile, the introduction of deep learning technology has provided a powerful tool for analyzing complex signals and recognizing quality in welding processes, driving monitoring systems from reliance on manual features and threshold-based discrimination toward data-driven intelligent diagnosis [8, 9]. Unlike previous reviews that mostly focus on a single technical path or specific scenario, this paper aims to systematically and multidimensionally review robotic welding quality online monitoring technology. As shown in Figure 1, the paper is organized as follows: Chapter 2 systematically reviews vision-based detection technologies, covering two major branches of 2D object detection and 3D structured light measurement; Chapter 3 explores multi-sensor-based detection technologies in depth, discussing three main methods including acoustic emission, arc spectroscopy, and thermal imaging; Chapter 4, taking information fusion hierarchy as the main thread, organizes the technical framework of data-level, feature-level, and decision-level fusion, with a focus on the frontier of deep learning-driven end-to-end feature-level fusion; Chapter 5 envisions technological development trends and key challenges from three dimensions—advanced sensing innovation, welding large-model enablement, and digital twin integration.

The main contributions of this paper are reflected in the following three aspects: (1) It covers the latest research progress in the field of online welding quality monitoring in recent years, especially frontier work such as OCT depth measurement, large-model-driven defect analysis, and digital twin closed loops; (2) From the perspective of fusion hierarchy (data-level/feature-level/decision-level), it establishes a unified analytical framework for multi-source information fusion and systematically organizes the technical evolution path from non-end-to-end to attention-based fusion; (3) Based on current technical bottlenecks, it distills future key technological development directions from three dimensions: advanced sensing, large-model enablement, and digital twin.

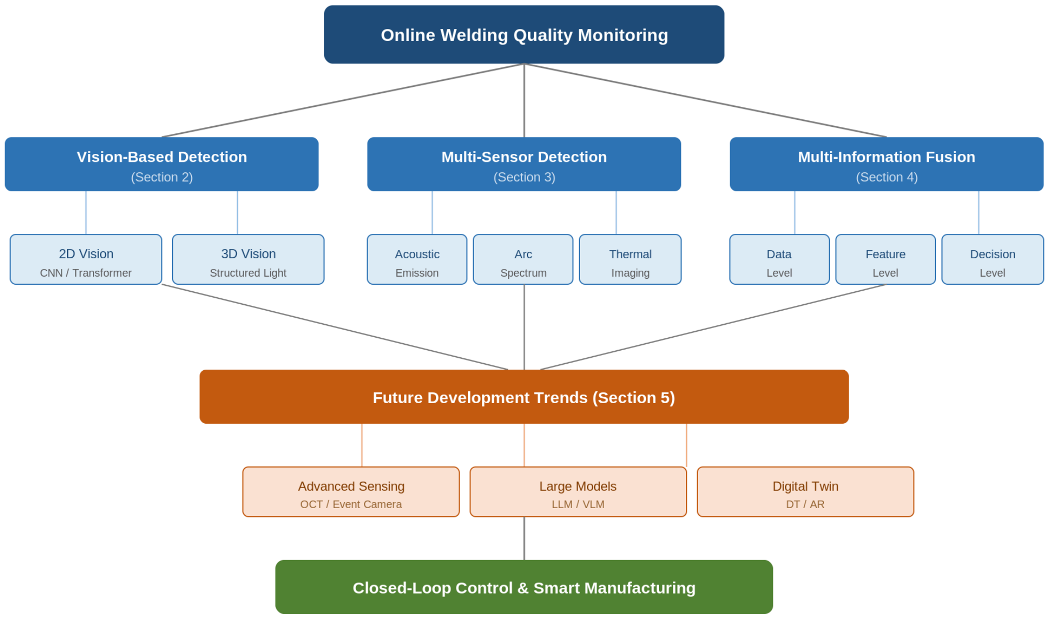

Figure 1. Framework of the review on online welding quality monitoring technologies

2. Vision-based detection technologies

Vision-based detection technology acquires image information of the welding area and is the most widely applied and most actively researched direction in online monitoring of welding quality. Its technical principle is to use optical devices such as industrial cameras, high-speed cameras, or structured light sensors to acquire weld pool and weld images under complex interference conditions such as arc light and fumes, extracting features including weld pool morphology, weld geometric dimensions, and surface defects, thereby evaluating welding quality. According to the imaging dimension, it can be mainly classified into two-dimensional (2D) vision detection technology and three-dimensional (3D) vision detection technology.

2.1. 2D vision detection technology

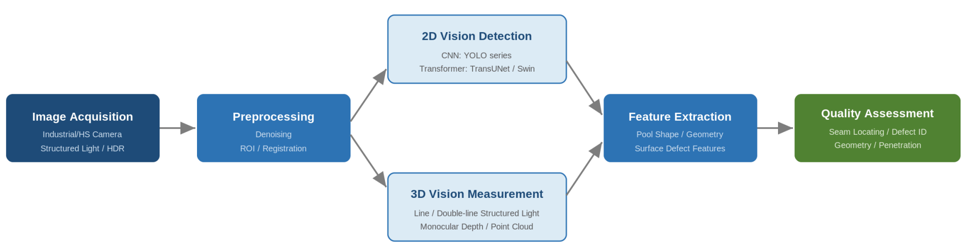

Figure 2. Roadmap of vision-based welding quality inspection technologies

In terms of 2D vision detection, research mainly focuses on weld localization and surface defect recognition based on deep learning object detection networks. The overall roadmap of vision-based welding quality detection technology is shown in Figure 2, covering image acquisition, preprocessing, 2D/3D detection, and quality assessment. According to the evolution of network architectures, existing methods can be summarized into two paradigms: convolutional neural networks and Transformers.

The first category uses Convolutional Neural Networks (CNNs) as the basic architecture, with typical methods mostly adopting YOLO series models, enhancing the perception of small-scale defects through structural improvements. Yang et al. [10] proposed an improved YOLOv7 algorithm for the high miss rate of surface defects in pipeline welds, improving the detection accuracy of small target defects; Feng et al. [11] designed a lightweight YOLO-Weld model combined with targeted data augmentation strategies, reducing model complexity while maintaining accuracy, adapting to the online detection needs of laser weld spots. Zhu et al. [12] proposed an improved YOLOv5s model for steel structure weld recognition, introducing a coordinate attention mechanism with mAP@0.5 reaching 93.79%, an improvement of 2.48% over the original model. In addition, to address the difficulty of scarce labeled data in actual production, researchers have introduced few-shot and transfer learning strategies: an improved generative adversarial network was used for few-shot welding defect image generation and data augmentation, effectively improving the performance of GMAW defect detection models [13]; a method based on VGG16 transfer learning with adaptive fine-tuning achieved approximately 90% average classification accuracy for multiple defect types on a small-scale X-ray weld dataset [14].

The second category of methods introduces the Transformer architecture to overcome the inherent limitations of CNNs in global context modeling. Eren et al. [15] transformed weld joint detection into an image segmentation problem, adopting TransUNet, which fuses convolution and Transformer, to achieve accurate semantic segmentation of welds. Comparative studies have shown that the segmentation model based on Swin Transformer achieved an F1 score of 96.31 in the weld segmentation task, outperforming traditional algorithms [16]. Furthermore, hybrid architectures integrating CNN local feature extraction and Transformer global modeling capabilities have been applied to weld pool segmentation, achieving a balance between local details and global context [17], representing the frontier direction of 2D vision detection.

However, the robustness of 2D vision methods against intense arc light and spatter interference during welding remains limited, and they cannot acquire three-dimensional geometric information of the weld, such as key quality parameters including reinforcement, weld width, and penetration depth. This fundamentally limits their depth of application in precision welding scenarios [18].

2.2. 3D vision detection technology

To address the fundamental inadequacy of 2D vision in acquiring three-dimensional geometric information, 3D vision detection technology has gradually become a research hotspot in the field of online weld monitoring. Among them, active vision methods based on line structured light have become the mainstream technical route for online inspection of weld geometric dimensions, owing to their significant advantages such as high measurement accuracy, strong real-time performance, and good environmental adaptability. Relevant reviews have systematically organized key aspects including structured light sensor design, laser stripe processing, and weld seam tracking [19, 20].

The accuracy and speed of laser stripe center extraction directly determine the measurement performance. Li et al. [21] proposed the interior Propagating Center Extraction (IPCE) algorithm, which directly traverses interference regions by propagating along the interior of the laser stripe center, enhancing robustness against intense arc light and spatter while optimizing processing efficiency; for the problems of strong noise and complex weld pass contours in multi-layer multi-pass welding, an automatic weld seam feature point extraction method based on laser vision further ensured high-precision tracking [22]. In terms of 3D reconstruction, a 3D vision measurement method based on double-line combined structured light experimentally verified small repeatability measurement errors [23]; for weld defect recognition, a method based on 3D laser scanning achieved a detection classification accuracy of over 97.1% for typical defects on the welding surface of aluminum reservoirs [24].

In addition to line structured light methods, emerging 3D vision technologies are also accelerating their integration into the welding monitoring field: an interpretable neural network detection system fusing 2D images and 3D point clouds overcame the limitations of single-dimension methods [25]; a method based on monocular depth estimation and point cloud networks explored a new approach to recover 3D weld information from a single frame image [26]; progress in deep learning-driven optical metrology has provided new tools for structured light encoding and phase measurement [27]. High Dynamic Range (HDR) cameras and Near-Infrared (NIR) imaging technologies have provided new hardware foundations for high-quality image acquisition under intense arc light interference.

3. Multi-sensor-based detection technologies

Multi-sensor detection technology indirectly evaluates weld quality by capturing the dynamic evolution of multi-dimensional physical quantities such as sound, light, electricity, and heat during the welding process, starting from the essential characteristics of the process, forming a complementary system with vision detection technology. Among them, acoustic emission is most sensitive to the initiation and propagation of internal microcracks in materials, arc/spectrum sensing can reflect arc stability and metallurgical reaction progress in real time, and thermal imaging can visually present the spatiotemporal distribution of the welding temperature field. The three methods each have unique advantages, but a single physical quantity is often difficult to comprehensively characterize diverse welding defects, requiring multi-sensor collaboration or multi-modal information fusion to enhance the comprehensiveness and reliability of diagnosis. The basic principles of the three typical multi-sensor detection methods are shown in Figure 3.

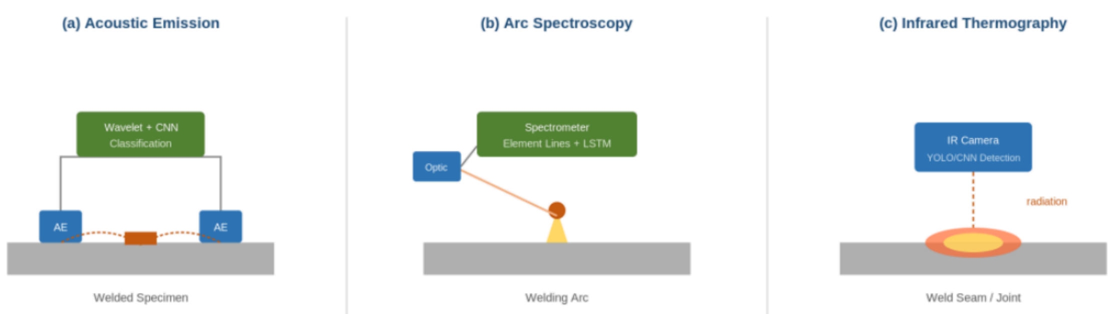

Figure 3. Principles of three typical multi-sensor detection methods

3.1. Acoustic emission detection technology

In acoustic emission detection, research mainly focuses on the dynamic evolution characteristics of internal material damage during welding. Early methods mostly relied on time-domain statistical parameters for threshold-based discrimination, making it difficult to capture precursor signals of microcracks. In recent years, automatic analysis technology for acoustic emission signals based on deep learning has developed rapidly. Ma et al. [28] combined wavelet transform to extract time-frequency features and adopted a convolutional neural network to intelligently classify weld pass damage states for stainless steel welded specimens, effectively overcoming the limitations of traditional parameter analysis methods in damage evolution recognition. In terms of method comparison, a study systematically compared the classification performance of shallow machine learning and deep learning in acoustic emission monitoring of aluminum alloy pulsed laser welding, demonstrating that deep learning has significant advantages in acoustic emission signal defect classification [29]. For engineering structures, targeting the problem of weld crack leakage in pressure pipelines, a method based on residual Swin Transformer encoded acoustic emission signals into images for classification, solving the recognition difficulty caused by the diversity of weld cracks [30].

3.2. Arc spectroscopy detection technology

In arc spectroscopy detection, traditional methods mostly judge welding process stability through statistical analysis of arc voltage and current waveforms, but such electrical parameters lack direct physical correlation with the formation of internal defects such as porosity. Emission spectroscopy can reflect the elemental composition and metallurgical reaction progress of arc plasma, providing a new approach for internal defect monitoring. Rodriguez-Cobo et al. [31] developed a low-cost online spectral monitoring method that achieves real-time recognition of arc welding defects in tanker trucks by analyzing characteristic spectral lines of elements such as Fe and Cr in the plasma. For high-demand aerospace scenarios, a method based on optical emission spectroscopy was used for online monitoring of internal porosity in GTAW, focusing on the internal microstructural integrity beyond the surface [32]. In terms of multi-sensor collaboration, Geiger et al. [33] combined acoustic and spectral emission signals for online defect detection in laser welding of battery cells, verifying the effectiveness of acoustic-optical multi-information collaboration in enhancing monitoring reliability.

3.3. Thermal imaging detection technology

In thermal imaging detection, traditional infrared thermometry is mostly used for post-weld analysis or qualitative early warning. In recent years, with the introduction of deep learning technology, thermal imaging is evolving from an auxiliary monitoring tool to a quantifiable automatic defect detection method. Bevilacqua et al. [34] developed a fully automatic laser weld defect detection system based on infrared thermal image sequences, using a custom CNN architecture for defect classification with an average accuracy of 99%. Fang et al. [35] further investigated the capabilities of deep learning models such as object detection and semantic segmentation in automatically detecting and recognizing defects in pulsed thermography data, promoting the integration of thermal imaging nondestructive testing into intelligent workflows. For traditional arc welding, Palumbo et al. [36] applied infrared thermography to online defect detection of SMAW and GMAW steel welded joints, achieving full-field non-contact inspection in both real-time and offline configurations.

3.4. Ultrasonic and acoustic emerging sensing technologies

In addition to the three main methods above, ultrasonic and welding acoustic signal analysis also demonstrate unique value in welding process monitoring. In terms of ultrasonic detection, online intervention of Phased Array Ultrasonic (PAUT) was used to evaluate signal disturbance and recovery during welding [37], and ultrasonic metal welding research showed that multi-sensor signal combinations can significantly enhance quality prediction robustness [38]. In terms of acoustic signal analysis, a method combining multi-spectral emission sensor features and machine learning was used for quantitative detection of defects such as porosity in laser welding [39]; for short-time welding processes such as resistance projection welding, airborne sound analysis combined with machine learning achieved online acoustic evaluation of welding quality [40]. The above progress indicates that ultrasonic and acoustic sensing are evolving from auxiliary means to important components of welding process monitoring.

In summary, multi-sensor detection technology can effectively compensate for the inherent inadequacies of vision methods in internal defect recognition, starting from the physical essence of the welding process. Current research is rapidly evolving from threshold-based judgment of single physical quantities toward multi-physical-quantity collaborative perception and deep learning-driven intelligent diagnosis. However, in complex industrial environments, the strong coupling of sensing signals, effective decoupling of environmental noise, real-time processing of multi-source data, and engineering robustness of models under variable operating conditions remain key bottlenecks that urgently need breakthroughs in this field.

4. Hierarchy and methods of multi-source information fusion

Multi-source information fusion aims to systematically integrate multi-dimensional information from heterogeneous sensors to obtain more accurate, reliable, and comprehensive welding quality diagnostic conclusions than any single information source [41]. According to the abstraction level at which information processing occurs, existing fusion paradigms can be categorized into three levels: data-level fusion, feature-level fusion, and decision-level fusion. The hierarchical architecture of multi-source information fusion is shown in Figure 4.

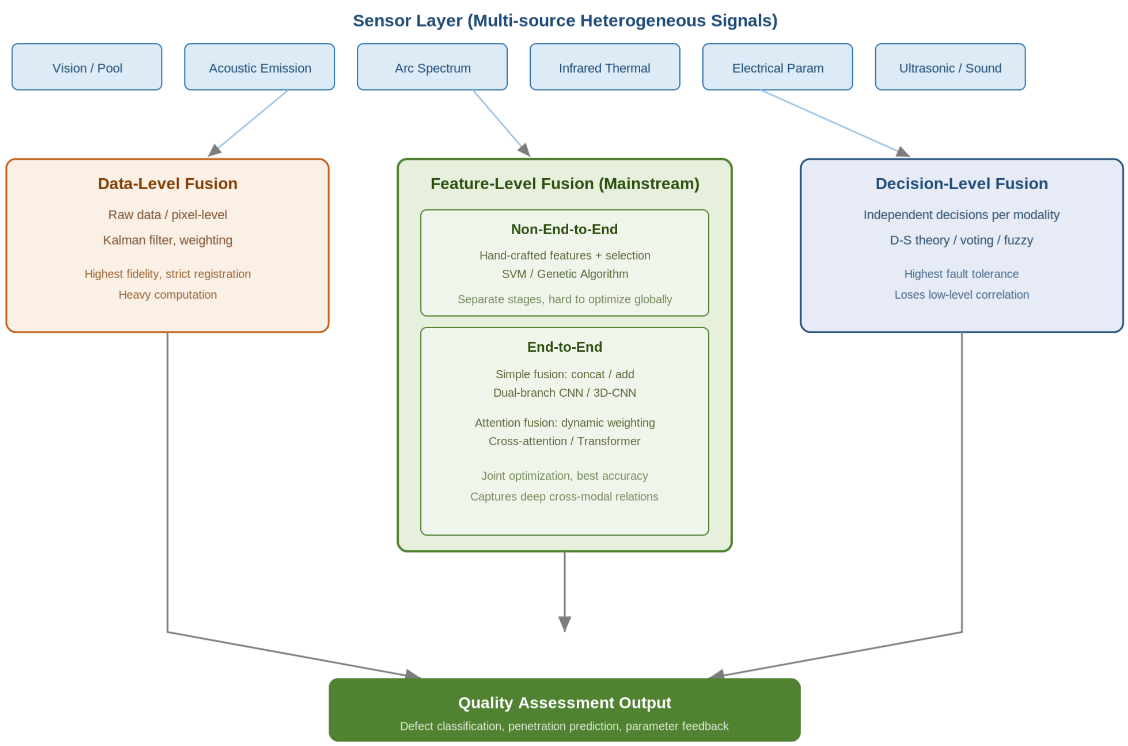

Figure 4. Hierarchical architecture of multi-source information fusion

4.1. Data-level fusion

Data-level fusion directly integrates at the raw data level, preserving the most complete information, but imposes extremely high requirements on the strict spatiotemporal registration of multi-source data, and has high computational cost, posing challenges to real-time performance. Cai et al. [42] performed adaptive fusion of multi-source coaxial images for laser keyhole welding processes, inputting them into a convolutional neural network to determine the penetration state, achieving an effective combination of raw image fusion and deep learning. In the multi-physics field dimension, Marko et al. [43] adopted multi-axis infrared thermography sensing during laser directed energy deposition, fusing multi-channel raw infrared data and developing automatic algorithms to identify defect formation, demonstrating the unique advantages of pixel-level fusion in weak defect signal detection. Given the stringent requirements of data-level fusion on computational resources and spatiotemporal registration, its application in real-time welding monitoring remains relatively limited.

4.2. Feature-level fusion

Feature-level fusion is the most mainstream fusion paradigm in current research, with its core advantage being that it both preserves the effective discriminative information of each modality and significantly reduces the computational burden through feature dimensionality reduction. This method first independently extracts representative feature vectors from each modality, then combines the features of each modality in the feature space, and finally inputs them into classification or regression models for decision-making. According to whether feature extraction and the decision process are jointly optimized within a unified framework, feature-level fusion can be divided into two paradigms: non-end-to-end and end-to-end.

4.2.1. Non-end-to-end feature fusion

Non-end-to-end fusion separately executes the three stages of feature extraction, feature concatenation, and classifier training, with each stage independently optimized. In the friction stir welding process, researchers acquired multi-sensor signals such as force and temperature, extracted time-frequency features, and input them into an SVM regression model after feature fusion to predict joint quality, verifying that the fused feature model outperforms single-sensor models [44]. Li et al. [45] established a multi-sensor monitoring system containing keyhole/weld pool images and laser-induced plasma spectra for TC4 titanium alloy laser welding, achieving online detection of penetration depth through multi-sensor feature fusion. In addition, feature fusion of images and multi-sensor data has also been used for weld formation quality inspection [46]. However, in such methods, feature extraction relies on manual design, and classifiers are trained separately, making the overall process difficult to achieve global optimality, and they are gradually being replaced by end-to-end methods.

4.2.2. End-to-end feature fusion

End-to-end fusion directly inputs raw data into the network, with feature extraction, cross-modal fusion, and final decision jointly learned and optimized within a unified framework, significantly enhancing real-time performance and diagnostic accuracy. According to the different cross-modal interaction mechanisms, end-to-end feature-level fusion can be further subdivided into two categories: simple fusion and attention-based fusion.

Simple fusion mainly combines feature vectors of different modalities in the feature space through concatenation or element-wise addition. Wang et al. [47] proposed a multi-scale deep feature fusion method based on weld pool videos for online monitoring of welding defects, with an end-to-end architecture directly predicting defects from video data. The SIMNet multi-modal fusion network proposed by Jiang et al. [48] adopts short-time Fourier transform to extract time-frequency features of sound signals while designing a visual feature extractor to process weld pool images, achieving penetration state monitoring through feature space alignment and cross-modal interaction. Simple fusion methods are straightforward to implement and computationally efficient, but their fundamental limitation lies in the inability to adaptively evaluate the differentiated importance of different modal features for the current task.

Attention-based fusion aims to overcome the aforementioned limitations of simple fusion. The attention mechanism enables the network to adaptively evaluate the importance of each modality's information for the current task and dynamically assign differentiated weights, and the Transformer architecture, with its global context modeling capability, has further advanced this direction [49, 50]. A review systematically organized the applications of mechanisms such as self-attention, channel attention, and cross-attention in sensing information processing for dynamic welding processes [49]. In specific practice, Chang et al. [51] proposed an end-to-end model combining CNN-LSTM with attention mechanism, performing attention-weighted fusion of multi-dimensional time-series signals of resistance spot welding to achieve online detection of spot welding quality for automotive bodies; for high-power laser-arc hybrid welding, an improved channel attention CNN enhanced defect recognition capability by adaptively focusing on key feature channels [52]. Compared with simple fusion, attention-based fusion can effectively capture the deep semantic correlations between modalities, representing the current technological frontier of feature-level fusion.

4.3. Decision-level fusion

Decision-level fusion is the highest-level fusion method, where each modality's data first independently completes a preliminary decision, and then final fusion of multiple decisions is performed through voting mechanisms, weighted average, Bayesian inference, or Dempster-Shafer (D-S) evidence theory [53]. The significant advantage of this method is high system fault tolerance—even if one modal sensor completely fails, the remaining modalities can still independently support decisions. Researchers proposed a weld defect classification method that uses analytic hierarchy process to determine weights and D-S evidence theory to fuse decisions of multiple classifiers, improving recognition accuracy under limited samples [54]; furthermore, a modified maximum likelihood estimation D-S model based on metal magnetic memory signals was used for quantitative multi-source decision fusion of weld defect levels [55]. However, decision-level fusion may lose the complementary correlations of low-level information between modalities, and its diagnostic accuracy is typically lower than that of feature-level fusion.

4.4. From fusion diagnosis to closed-loop control

The above fusion methods mainly focus on the quality diagnosis stage, while the complete connotation of "monitoring and control" requires feeding diagnostic results back into the control loop. In terms of visual servoing, Luo and Yamane et al. [56] proposed a vision-based closed-loop control system for pulsed MAG welding using Otsu segmentation, regulating welding parameters in real time through weld pool features; Lu and Yamane et al. [57] further introduced the U-Net deep learning model into weld pool boundary detection for plasma arc welding, driving real-time weld seam tracking control. In terms of process optimization, reinforcement learning as a data-driven optimization technique was used for GMAW process parameter optimization, reducing the number of experiments required for traditional welding procedure formulation [58]; a method based on multi-modal online data fusion achieved monitoring and control of penetration depth in complex groove welding [59]. The above work indicates that the deep coupling of multi-source information fusion diagnosis and adaptive control strategies is key to achieving true unification of "monitoring" and "control".

Table 1 summarizes the comparison of the three fusion levels in terms of information fidelity, computational complexity, fault tolerance, typical algorithms, and representative references. In summary, data-level fusion preserves the most complete information but imposes extremely high requirements on spatiotemporal registration, feature-level fusion (especially the end-to-end paradigm driven by attention mechanisms) is the current research mainstream, and decision-level fusion has unique advantages in terms of fault tolerance. Current fusion technologies still face three major limitations—limited sensing dimensions, insufficient model generalization, and closed loops not yet established. The following Chapter 5 presents an outlook on these bottlenecks.

Table 1. Comparison of three fusion levels

Fusion level	Information fidelity	Computational complexity	Fault tolerance	Typical algorithms	Representative references
Data-level fusion	Highest	High	Low	Kalman filter, pixel-level weighting	[42, 43]
Feature-level (non-end-to-end)	Medium	Medium	Medium	SVM, genetic algorithm	[44-46]
Feature-level (end-to-end-simple)	Relatively high	Relatively low	Medium	Dual-branch CNN, 3D-CNN	[47, 48]
Feature-level (end-to-end-attention)	High	Relatively high	Relatively high	Cross-attention, Transformer	[49-52]
Decision-level fusion	Relatively low	Low	Highest	D-S evidence theory, fuzzy inference	[53-55]

5. Future development trends

The field of online welding quality monitoring is rapidly evolving from "single-signal perception" to "multi-modal collaborative cognition". Despite rapid development, industrial deployment still faces numerous challenges: complex interferences such as intense arc light, fumes, and spatter at welding sites severely affect sensor stability and data quality; high-quality labeled data, especially defect samples, are difficult to acquire and costly, limiting the training effectiveness and generalization ability of deep learning models; the computational cost of multi-modal data processing and complex model inference is high, making it difficult to meet the stringent requirements of millisecond-level real-time response in welding processes; and the insufficient interpretability of deep learning models also constrains their promotion and application in high-safety-level fields. Looking to the future, this field will rapidly evolve along the following three main dimensions.

5.1. Innovation of advanced sensing technologies

On the hardware sensing side, a series of new sensing technologies are breaking through the perception dimension bottlenecks of traditional means. Among them, Optical Coherence Tomography (OCT) technology has attracted the most attention as a new means for weld penetration measurement in recent years; utilizing the low-coherence interference principle, it can directly measure the keyhole depth during laser welding with micrometer-level precision. Researchers optically coupled OCT depth measurement with coaxial high-speed camera imaging and fused the two data streams, achieving online monitoring of weld penetration depth and contact width in laser lap welding [60], marking the maturation of OCT from laboratory principle verification to industrial-grade online measurement. Its measurement principle and data processing flow are shown in Figure 5.

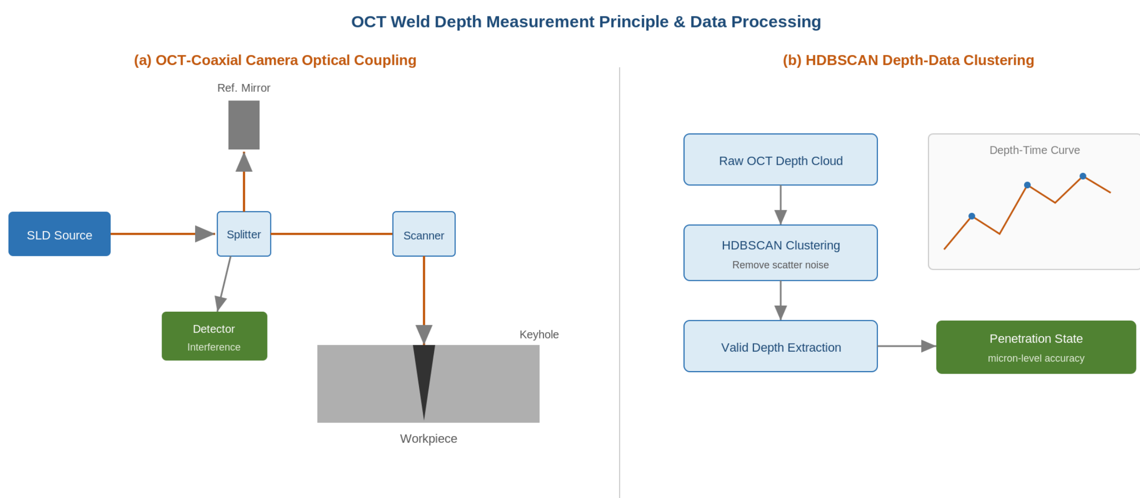

Figure 5. OCT weld depth measurement principle and data processing

Meanwhile, High Dynamic Range (HDR) cameras and Near-Infrared (NIR) imaging technologies provide new hardware foundations for high-quality imaging under intense arc light interference. A low-cost near-infrared imaging system achieved real-time monitoring of GMAW by filtering to isolate weld pool thermal radiation and suppress arc interference, combined with machine learning [61]. In addition, as a transformative sensing paradigm, event cameras asynchronously sense pixel-level brightness changes, capturing transient events during the welding process with sub-millisecond temporal resolution while maintaining high dynamic range. Event cameras have been first applied to laser welding process monitoring [62], and have been verified to be applicable to online monitoring of weld pools in TIG and laser welding [63]. The synergistic introduction of the above emerging sensing technologies will lay the hardware foundation for building a comprehensive, multi-dimensional, high-precision welding process perception system.

5.2. Enablement prospects of welding large models

The rapid development of Large Language Models (LLM) and Vision Language Models (VLM) is injecting new momentum into welding intelligence. The welding field has long faced the dilemma of "data scarcity, tacit knowledge, and variable scenarios"—high-level welding process knowledge relies heavily on the individual experience of welders, making it difficult to explicitly express and scale for reuse. Large model technology provides a new paradigm for addressing this fundamental contradiction.

In terms of process modeling and defect prediction, a research team introduced LLMs into the welding field, proposing a modeling paradigm that automatically derives mathematical equations describing physical phenomena from scientific literature and a small number of experiments, successfully applied to quantitative prediction of humping defects in high-speed laser welding [64], marking the transition of LLMs from "literature retrieval assistance tools" to "core components of welding process modeling". In terms of intelligent defect classification and mechanism analysis, a study utilized large language models to achieve microstructure optimization and defect classification of welded titanium alloys, verifying the application potential of large models in welding materials science [65]. From a broader perspective, the application of large-scale foundation models in intelligent manufacturing is becoming a research hotspot [66].

Looking forward, the construction of welding vertical large models has become a consensus frontier direction in this field. Its core idea is to deeply fuse welding process knowledge graphs, multi-source sensing databases, and pre-trained large models, enabling models not only to "understand" the multi-modal data of the welding process but also to "reason" the causal relationships between process parameters and weld quality, thereby driving the transformation of welding process development from "experience-driven" to "knowledge-driven plus data-driven". However, in-depth exploration is still urgently needed in terms of computational efficiency, data security, and industrial deployment feasibility.

5.3. Integration of digital twin and immersive technologies

Digital Twin (DT) technology, by constructing high-fidelity mirrors of physical welding processes in virtual space, is expected to achieve full life-cycle monitoring, predictive quality assessment, and closed-loop optimal control of welding processes [67]. In terms of system architecture, Dai et al. [68] proposed an AI-Empowered Digital Twin (AI-DT) framework, integrating two core modules—generative AI and predictive AI—providing a systematic framework for real-time monitoring and intelligent decision-making of welding processes. In terms of quality prediction, Thangavel et al. [69] proposed a digital twin-based TIG welding quality prediction system, combining thermal imaging and process parameters to achieve online quality prediction; a deep learning digital twin system based on acoustic signals achieved online weld defect diagnosis [70]. In terms of lightweight and engineering deployment, a lightweight digital twin model for CMT welding based on twin data and improved generative adversarial networks was proposed [71].

Augmented Reality (AR) and Virtual Reality (VR) technologies provide human-machine interaction interfaces for the engineering deployment of digital twins, and relevant systematic reviews have organized the application status of VR/AR in welding skill training and immersive learning [72]. AR-assisted welding can overlay key process parameters as augmented information onto the welder's field of view, bridging the information gap between "offline simulation" and "online execution", thereby constructing a complete closed loop from "perception → simulation → prediction → intervention". However, high-fidelity multi-physics real-time simulation, standardization of heterogeneous system interoperability, and industrial-level deployment costs remain major challenges in this field.

6. Conclusion

This paper systematically reviews the online monitoring technologies for welding quality, covering three main technical paths: vision-based detection, multi-sensor-based detection, and multi-source information fusion. Vision-based detection provides rich geometric and morphological information of the weld pool and weld surface, multi-sensor detection complements the perception of internal defects from the physical essence of the welding process, and multi-source information fusion integrates the complementary advantages of heterogeneous sensing data. Among them, the feature-level fusion paradigm driven by deep learning and attention mechanisms has become the current research mainstream and represents the most active frontier of this field.

Despite the significant progress achieved, the industrial deployment of online welding quality monitoring still faces challenges in several aspects: the robustness of sensing systems under intense arc light, fume, and spatter interference; the acquisition of high-quality labeled defect data; the real-time computational cost of multi-modal data processing; the interpretability of deep learning models; and the establishment of complete closed-loop control. Future development will focus on three main dimensions: advanced sensing innovation (including OCT, HDR/NIR imaging, and event cameras), the enablement of welding large models, and the integration of digital twins and immersive technologies.

In summary, online welding quality monitoring is evolving from single-signal perception toward multi-modal collaborative cognition and closed-loop intelligent control. It is expected to play an increasingly important role in intelligent manufacturing and to provide a solid foundation for achieving high-efficiency, zero-defect welding.

References

[1] Shaloo, M., Schnall, M., Klein, T., Huber, N., & Reitinger, B. (2022). A review of non-destructive testing (NDT) techniques for defect detection: Application to fusion welding and future wire arc additive manufacturing processes. *Materials, 15*(10), Article 3697.

[2] Knaak, C., Kolter, G., Schulze, F., Kröger, M., & Abels, P. (2022). In situ detection of welding defects: A review. *Welding in the World, 66*(4), 611–633.

[3] Stavridis, J., Papacharalampopoulos, A., & Stavropoulos, P. (2018). Quality assessment in laser welding: A critical review. *The International Journal of Advanced Manufacturing Technology, 94*(5–8), 1825–1847.

[4] Fan, X., Gao, X., Liu, G., Ma, N., & Zhang, Y. (2021). Research and prospect of welding monitoring technology based on machine vision. *The International Journal of Advanced Manufacturing Technology, 115*(11–12), 3365–3391.

[5] Lv, N., Xu, Y., Li, S., Yu, X., & Chen, S. (2020). Automated control of welding penetration based on audio sensing with deep learning approach. *IEEE Transactions on Industrial Informatics, 17*(1), 524–534.

[6] Huang, Y.-W., Gao, X.-D., Gao, P. P., Ma, B., & Zhang, Y.-X. (2024). Laser welding monitoring techniques based on optical diagnosis and artificial intelligence: A review. *Advances in Manufacturing, 12*(3), 507–535.

[7] Kong, L., Peng, X., Chen, Y., Wang, P., & Xu, M. (2020). Multi-sensor measurement and data fusion technology for manufacturing process monitoring: A literature review. *International Journal of Extreme Manufacturing, 2*(2), Article 022001.

[8] Wang, Z., Li, L., Chen, H., Wu, X., Dong, Y., Tian, J., & Zhang, Q. (2023). Penetration recognition based on machine learning in arc welding: A review. *The International Journal of Advanced Manufacturing Technology, 125*(9), 3899–3923.

[9] Singh, G., & Vasudev, H. (2025). Advancements in sensor technologies and AI-based automation for intelligent welding systems: A comprehensive review. *The International Journal of Advanced Manufacturing Technology*. Advance online publication. https://doi.org/10.1007/s00170-025-16745-8

[10] Xu, X., & Li, X. (2024). Research on surface defect detection algorithm of pipeline weld based on YOLOv7. *Scientific Reports, 14*, Article 1881. https://doi.org/10.1038/s41598-024-52451-3

[11] Feng, J., Wang, J., Zhao, X., Liu, Z., & Ding, Y. (2024). Laser weld spot detection based on YOLO-weld. *Scientific Reports, 14*, Article 28098.

[12] Zhu, S., Zhang, L., Zhao, J., Hu, X., & Luo, X. (2025). Research on steel structure weld seam recognition algorithm based on improved YOLOv5. *Scientific Reports, 15*(1), Article 34276.

[13] Zhu, K., Chen, W., Hou, Z., Wang, Q., & Chen, H. (2023). Modified fusing-and-filling generative adversarial network-based few-shot image generation for GMAW defect detection using multi-sensor monitoring system. *The International Journal of Advanced Manufacturing Technology, 128*(5), 2753–2762.

[14] Kumaresan, S., Jai Aultrin, K. S., Kumar, S. S., & Dev Anand, M. (2023). Deep learning-based weld defect classification using VGG16 transfer learning adaptive fine-tuning. *International Journal on Interactive Design and Manufacturing, 17*(6), 2999–3010.

[15] Eren, B., Demir, M. H., & Mistikoglu, S. (2024). A new approach for detection of weld joint by image segmentation with deep learning-based TransUNet. *The International Journal of Advanced Manufacturing Technology, 134*(11), 5225–5240.

[16] Zhang, B. X., Wang, X. P., Cui, J. H., Wu, J. T., Xiong, Z., & Zhang, W. P. (2024). Enhancing weld inspection through comparative analysis of traditional algorithms and deep learning approaches. *Journal of Nondestructive Evaluation, 43*, Article 47.

[17] Yang, L., Wang, H., Meng, W., & Pan, H. (2024). CvT-UNet: A weld pool segmentation method integrating a CNN and a transformer. *Heliyon, 10*(15), Article e35738.

[18] Zhang, M., Feng, M., Chen, C., Yu, X., & Lian, G. (2026). Weld defect detection: Deep learning-based image processing and the mechanisms of defect formation. *Archives of Computational Methods in Engineering, 33*(3), 3525–3563.

[19] Zou, Y., & Chen, J. (2020). Advanced techniques of the structured light sensing in intelligent welding robots: A review. *The International Journal of Advanced Manufacturing Technology, 107*(9–10), 4089–4103.

[20] Xu, Y., Lv, N., Fang, G., Du, S., Zhao, W., Ye, Z., & Chen, S. (2017). Welding seam tracking in robotic gas metal arc welding. *Journal of Materials Processing Technology, 248*, 92–100.

[21] Li, W., Mei, F., Hu, Z., Gao, X., Yu, H., Housein, A. A., & Wei, C. (2022). Multiple weld seam laser vision recognition method based on the IPCE algorithm. *Optics & Laser Technology, 155*, Article 108388.

[22] Xu, F., He, L., Hou, Z., Xiao, R., Zuo, T., Li, J., & Zhang, H. (2024). An automatic feature point extraction method based on laser vision for robotic multi-layer multi-pass weld seam tracking. *The International Journal of Advanced Manufacturing Technology, 131*(12), 5941–5960.

[23] Wang, M., Sun, Q., Gao, C., Ren, Z., & Dai, W. (2023). A three-dimensional vision measurement method based on double-line combined structured light. *Scientific Reports, 13*, Article 18660.

[24] Huang, H., Zhou, B., Cao, S., Song, T., Xu, Z., & Jiang, Q. (2025). Aluminum reservoir welding surface defect detection method based on three-dimensional vision. *Sensors, 25*(3), Article 664.

[25] Wang, S., Li, Z., Chen, G., & Yue, Y. (2024). Weld seam object detection system based on the fusion of 2D images and 3D point clouds using interpretable neural networks. *Scientific Reports, 14*(1), Article 21137.

[26] Zhao, N., & Chen, S. (2026). Industrial weld defect detection based on monocular depth estimation and dual-attention point cloud network. *Sensors, 26*(11), Article 3321.

[27] Zuo, C., Qian, J., Feng, S., Yin, W., Li, Y., Fan, P., Han, J., Qian, K., & Chen, Q. (2022). Deep learning in optical metrology: A review. *Light: Science & Applications, 11*(1), Article 39.

[28] Ma, Y., Liu, M., Yang, L., Fan, J., & Tsangouri, E. (2025). Monitoring of weld damage based on deep learning and acoustic emission methods. *Journal of Constructional Steel Research, 231*, Article 109583.

[29] Zeng, D., Wu, D., Luo, Z., Dong, J., Huang, H., Yang, F., & Ye, X. (2024). A performance comparison of deep learning and shallow machine learning in acoustic emission monitoring of aluminium alloy pulsed laser welding. *Soft Computing, 28*(17), 10263–10279.

[30] Huang, J., Zhang, Z., Qin, R., Yu, Y., Li, Y., Wen, G., Cheng, W., & Chen, X. (2024). Residual Swin transformer-based weld crack leakage monitoring of pressure pipeline. *Welding in the World, 68*(4), 879–891.

[31] Rodriguez-Cobo, L., Mirapeix, J., Cobo, A., & Lopez-Higuera, J. M. (2022). Spectroscopic approach for the on-line monitoring of welding of tanker trucks. *Applied Sciences, 12*(10), Article 5022.

[32] Zhang, Z., Wen, G., & Ren, W. (2025). Optical spectroscopy-based inner porosity monitoring for GTAW. In *Multiple sensing-driven feature extraction and defect detection* (pp. 15–31). Springer.

[33] Heilmeier, J., Kick, M. K., Grabmann, S., Muschol, T., Schlicht, F., von Hundelshausen, F., von Ribbeck, H.-G., Weiss, T., & Zaeh, M. F. (2024). Inline failure detection in laser beam welding of battery cells: Acoustic and spectral emission analysis for quality monitoring. *Journal of Laser Applications, 36*(2), Article 022007.

[34] Buongiorno, D., Prunella, M., Grossi, S., Hussain, S. M., Rennola, A., Longo, N., Di Stefano, G., Bevilacqua, V., & Brunetti, A. (2022). Inline defective laser weld identification by processing thermal image sequences with machine and deep learning techniques. *Applied Sciences, 12*(13), Article 6455. https://doi.org/10.3390/app12136455

[35] Fang, Q., Ibarra-Castanedo, C., Garrido, I., Duan, Y., & Maldague, X. (2023). Automatic detection and identification of defects by deep learning algorithms from pulsed thermography data. *Sensors, 23*(9), Article

4444.

[36] Santoro, L., Sesana, R., Molica Nardo, R., & Curá, F. (2023). Infrared in-line monitoring of flaws in steel welded joints: A preliminary approach with SMAW and GMAW processes. *The International Journal of Advanced Manufacturing Technology, 128*(5–6), 2655–2670.

[37] Dimakos, A., Sweeney, N. E., Loukas, C., Thompson, C. D., Serjeant, S., MacLeod, C. N., Mohseni, E., Lines, D., & Sibson, J. (2025). Evaluation of signal disturbance and recovery in phased array ultrasonic inspection during welding. *Welding in the World, 69*(11), 3021–3034.

[38] Müller, F. W., Mirz, C., Schiebahn, A., & Reisgen, U. (2025). Influence of quality features, disturbances, sensor data, and measurement time on quality prediction for ultrasonic metal welding. *Welding in the World, 69*(3), 615–629.

[39] Darwish, A., Persson, M., Ericson, S., Ghasemi, R., & Salomonsson, K. (2025). Weld defect detection in laser beam welding using multispectral emission sensor features and machine learning. *Sensors, 25*(16), Article 5120.

[40] Koal, J., Baumgarten, M., Nikolov, C., Ramakrishnan, S., & Mathiszik, C. (2025). Acoustic process monitoring during projection welding using airborne sound analysis and machine learning. *Welding in the World, 69*(2), 459–470.

[41] Gravina, R., Alinia, P., Ghasemzadeh, H., & Fortino, G. (2017). Multi-sensor fusion in body sensor networks: State-of-the-art and research challenges. *Information Fusion, 35*, 68–80.

[42] Cai, W., Jiang, P., Shu, L., Geng, S., & Zhou, Q. (2022). Real-time laser keyhole welding penetration state monitoring based on adaptive fusion images using convolutional neural networks. *Journal of Intelligent Manufacturing, 33*(5), 1259–1272.

[43] Herzog, T., Brandt, M., Trinchi, A., Sola, A., Hagenlocher, C., & Molotnikov, A. (2024). Defect detection by multi-axis infrared process monitoring of laser beam directed energy deposition. *Scientific Reports, 14*, Article 3234.

[44] Das, B., & Ramon, J. (2023). Sensor fusion model for weld quality monitoring in friction stir welding process using machine learning technique. *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering, 237*(6), 2547–2556.

[45] Li, D., Yang, K., Li, M., Wu, B., & Yang, L. (2024). Online detection of laser welding penetration depth based on multi-sensor features. *Materials, 17*(7), Article 1580.

[46] Lee, J., Choi, H., & Kim, J. (2023). Welding bead inspection using image and multi-sensor fusion. *Applied Sciences, 13*(20), Article 11497. https://doi.org/10.3390/app132011497

[47] Wang, L., Zhu, C., Han, T., & Zhang, X. (2024). Welding defect monitoring based on multi-scale feature fusion of molten pool videos. *Sensors, 24*(20), Article 6561.

[48] Jiang, Q., Wang, Y., Kong, Y., Liu, Y., & Ma, C. (2025). A monitoring network SIMNet for weld penetration status based on multimodal fusion. *Scientific Reports, 15*(1), Article 23355.

[49] Xu, J., Liu, Q., Xu, Y., Xiao, R., Hou, Z., & Chen, S. (2024). Review on the application of the attention mechanism in sensing information processing for dynamic welding processes. *Journal of Manufacturing and Materials Processing, 8*(1), Article 22. https://doi.org/10.3390/jmmp8010022

[50] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems (Vol. 30*, pp. 5998–6008). Curran Associates.

[51] Chang, F., Zhou, G., Ding, K., Li, J., Jing, Y., Hui, J., & Zhang, C. (2023). A CNN-LSTM and attention-mechanism-based resistance spot welding quality online detection method for automotive bodies. *Mathematics, 11*(22), Article 4570.

[52] Qiu, Y., Ping, J., Shu, L., Song, M., Ma, D., Yan, X., & Li, S. (2025). Defect monitoring of high-power laser-arc hybrid welding process based on an improved channel attention convolutional neural network. *Journal of Intelligent Manufacturing, 36*(4), 2657–2676.

[53] Sentz, K., & Ferson, S. (2002). Combination of evidence in Dempster-Shafer theory (SAND2002-0835). Sandia National Laboratories.

[54] Jiang, H., Wang, R., Gao, Z., Gao, J., & Wang, H. (2019). Classification of weld defects based on the analytical hierarchy process and Dempster-Shafer evidence theory. *Journal of Intelligent Manufacturing, 30*(4), 2013–2024.

[55] Xing, H., Xu, C., Yi, M., Gao, S., & Liu, W. (2023). Modified maximum likelihood estimation metal magnetic memory quantitative identifying model of weld defect levels based on Dempster-Shafer theory. *Applied Sciences, 13*(13), Article 7959.

[56] Luo, Y., Yamane, S., Wang, W., Tsumori, R., Ochiai, K., Lu, J., & Xia, Y. (2025). Vision-based closed-loop control of pulsed MAG welding using Otsu-segmented arc features. *Applied Sciences, 15*(16), Article 8950.

[57] Lu, J., Yamane, S., Wang, W., Li, N., Wang, S., & Xia, Y. (2025). Weld pool boundary detection based on the U-Net algorithm and weld seam tracking in plasma arc welding. *Applied Sciences, 15*(4), Article 1814.

[58] Mattera, G., Caggiano, A., & Nele, L. (2024). Reinforcement learning as data-driven optimization technique for GMAW process. *Welding in the World, 68*(4), 805–817.

[59] Gao, P., Wu, Z., Wang, Y., Lu, J., & Zhao, Z. (2024). Method for monitoring and controlling penetration of complex groove welding based on online multi-modal data. *Journal of Intelligent Manufacturing, 35*(3), 1247–1265.

[60] Li, J., Zhao, J., Zhang, C., Ding, Y., Bai, L., & Cheng, Y. (2024). In-process monitoring of penetration depth and contact width during laser lap welding using fused OCT depth and coaxial high speed image. In *2024 IEEE International Conference on Mechatronics and Automation (ICMA)* (pp. 1586–1591). IEEE.

[61] Muñoz Chávez, J. J., Lira, M. N. d. S., Pizo, G. A. I., Payão Filho, J. d. C., Alfaro, S. C. A., & Motta, J. M. S. T. d. (2025). Development of a low-cost infrared imaging system for real-time analysis and machine learning-based monitoring of GMAW. *Sensors, 25*(22), Article 6858.

[62] Dold, P. M., Nadkarni, P., Boley, M., Schorb, V., Wu, L., Steinberg, F., Burggräf, P., & Mikut, R. (2025). Event-based vision in laser welding: An approach for process monitoring. *Journal of Laser Applications, 37*(1), Article 012040.

[63] Mascareñas, D., Green, A., Liao, A., Torrez, M., Cattaneo, A., Black, A., Bernardin, J., & Kenyon, G. (2025). Demonstration of neuromorphic event-based imagers for optical measurement of melt pools for additive manufacturing and welding diagnostics. In *Conference Proceedings of the Society for Experimental Mechanics Series* (pp. 57–67). Springer.

[64] Yu, Z., Lai, Z.-H., Geng, P., Lee, K., Rinker, T. J., Tan, C., Carlson, B., Xu, S., & Li, J. (2025). Derivation of physical equations for high-speed laser welding using large language models. *International Journal of Machine Tools and Manufacture, 208*, Article 104255.

[65] Zhang, S., Wang, W. Y., Wang, X., Li, G., Ren, Y., Gao, X., Sun, F., Tang, B., Song, H., & Li, J. (2024). Large language models enabled intelligent microstructure optimization and defects classification of welded titanium alloys. *Journal of Materials Informatics, 4*, Article 34.

[66] Zhang, H., Wang, Z., Lv, X., Xu, K., Wu, L., Dereck, S. S., Jia, Y., Wu, J., Long, Z., Liang, W., & Ma, X. G. (2026). Large scale foundation models for intelligent manufacturing applications: A survey. *Journal of Intelligent Manufacturing, 37*(1), 119–170.

[67] Lu, Y., Liu, C., Wang, K. I. K., Huang, H., & Xu, X. (2020). Digital twin-driven smart manufacturing: Connotation, reference model, applications and research issues. *Robotics and Computer-Integrated Manufacturing, 61*, Article 101837.

[68] Dai, D., Zhao, B., Yu, Z., Franciosa, P., & Ceglarek, D. (2025). Generative and predictive AI for digital twin systems in manufacturing. *Frontiers in Artificial Intelligence, 8*, Article 1655470. https://doi.org/10.3389/frai.2025.1655470

[69] Thangavel, S., Mohanraj, T., & Sakthivel, M. (2024). Digital twin-based TIG welding quality prediction using electrode tip angle degradation. *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of*

Process Mechanical Engineering, 238(4), 1926–1940.

[70] Ji, T., & Mohamad Nor, N. (2023). Deep learning-empowered digital twin using acoustic signal for welding quality inspection. *Sensors, 23*(5), Article 2643.

[71] Li, X., Yang, K., Huang, H., Chen, J., & Yang, J. (2026). A lightweight digital twin of cold metal transfer welding based on twin data and improved C-DCGAN algorithm. *Applied Intelligence, 56*(1), Article 45.

[72] Chan, V. S., Haron, H., Mohamed Sunar, M. S., & Ismail, M. H. (2022). VR and AR virtual welding for psychomotor skills: A systematic review. Multimedia Tools and Applications, 81(9), 12459–12493.