Welding in the World
https://doi.org/10.1007/s40194-025-01968-w

RESEARCH PAPER

Vision-driven adaptive welding solutions for the top three challenges in welding fabrication

Mahyar Asadi[1] · Ahmad Ashoori[1] · Mehrnoosh Afshar[1] · Ali Sheikhshab[1] · Todd Scheerer[1] · Austin Kaspardlov[1] · Soroush Bagheri[1] · Sina Firouz[1] · Soroush Karimzadeh[1]

Received: 6 November 2024 / Accepted: 2 February 2025
© International Institute of Welding 2025

Abstract

With experience in more than over 100 robotic deployments in pipe prefabrication and a decade-long dedication to welding automation, we have pinpointed the key challenges, notably fit-up variation, tack adaptation, and live seam tracking. We engineered an innovative adaptive welding solution that integrates the perceptual and cognitive abilities of welders into articulated robots. This system dynamically responds to real-time welding scenarios, effectively tackling associated challenges. Unlike existing methods reliant on pre-scanning or laser readings before welding, our vision-based adaptive welding technology operates instantaneously, replicating the expertise of proficient human welders. The outcome is a consistent delivery of high-quality welds. Given the widespread advancement of AI, the heart of the adaptive welding system must skillfully manage diverse welding conditions, covering different joint preparations, types, positions, thicknesses, materials, and beyond. Addressing the necessity of training the AI core requires navigating through diverse practical challenges in deployments. Leveraging our expertise in deploying various methodologies, we ultimately provide an efficient solution for training the welding AI, primed for widespread deployment across high-mix low-volume applications. This solution incorporates a data tracing and monitoring platform across deployments, enhancing ERP (Enterprise Resource Planning) functionality, and providing insights into welding operations, historical performance analytics, and problem tracking with proactive improvements.

Keywords Adaptive welding · Fitup variation · Tack adaptation · Seam tracking · Vision welding · Gas metal arc welding · Artificial intelligence

1 Introduction

In the age of fast-growing smart technologies and devices, the art of connecting intelligent solutions to industrial setbacks will be the main drive to the sustainability of those solutions, perhaps an essential factor in self-fostering them at large. Like many other industries, welding engineering is fast transforming toward smart systems from a field that traditionally relies on the experience, perception, and cognition of welders, supervisors, and engineers.

Smartization in adaptive welding [1] covers a broad spectrum of techniques, each tailored to meet specific needs in various industries, but they all share a standard set of objectives. At the core, these approaches try to improve consistency in weld quality, boost production throughput, and elevate the overall craftsmanship in manufacturing and fabrication processes [2].

Consistent quality is one of the primary drivers behind adaptive welding technologies. Adaptive welding systems use real-time monitoring and control mechanisms, sensors, and machine learning algorithms to adjust the welding process and motion parameters like current, voltage, wire feed speed, travel speed, angles, and more, to compensate for any fluctuation from the set quality on the fly [3]. This real-time adaptability ensures uniformity over weldment inconsistency and imperfect fit-ups.

Increasing throughput is another critical drive. In today's competitive industrial landscape, manufacturers are constantly required to optimize efficiency and reduce cycle times. Adaptive welding systems tackle this challenge by automating tasks that traditionally require human intervention. For

✉ Mahyar Asadi
mahyar@novarctech.com

[1] Novarc Technologies, 1225 Keith Rd E Unit #6, North Vancouver, BC V7J 1J3, Canada

instance, they eliminate the need for robot programmers to manually touch up programming to accommodate part variations, allowing for relaxed part tolerances and enabling faster, more efficient welding in high-mix environments. Adaptive systems can make complex welds more accessible, even for less experienced operators, while ensuring the final product meets stringent quality standards [4].

Eventually, the costs of quality workmanship are critical, especially in large-scale manufacturing, where repairs can lead to significant losses in both time and money. Historically, overwelding has been a conservative approach to minimize the risk of repairs. However, the costs associated with overwelding can quickly amount to millions of dollars.

Current welding technologies are primarily based on prescriptive logic, which depends on pre-measured input data to function effectively [5]. These inputs primarily include detailed information about joint configurations, such as the shape, size, and orientation of the weld joint; robot configurations, encompassing the positioning, motion paths, and operational constraints of the robotic system; and welding parameters. These inputs are essential to generate adaptive scenarios that guide the welding process. Our exploration stage with customers concluded that this approach limits the system's flexibility and responsiveness, as it relies heavily on accurate, upfront data and predefined parameters rather than real-time adjustments or dynamic learning. This often poses challenges in welding environments where variability or unexpected conditions occur, highlighting the need for more advanced, real-time adaptive solutions. To address this, we have developed a system capable of responding within a fraction of a second—significantly faster than the response time of welding physics phenomena. Our solution operates directly in the live welding environment, eliminating the need for pre-scanning or pre-measured inputs to the AI, enabling unparalleled adaptability and precision.

Our approach, which reacts to the live welding scene without requiring prior information, eliminates the need for robot program corrections caused by welding distortion during the welding. This innovation effectively removes the necessity for touch-up robot programming to accommodate part variations. Through customer feedback, we have learned that this addresses a significant challenge for many users, as they heavily depend on shop-floor welding robot programmers to perform these frequent adjustments. Finding skilled personnel for such tasks is increasingly difficult, with this resource scarcity surpassing even the challenge of hiring qualified welders.

2 Problem statement

The author's extensive experience in the welding industry, including hundreds of welding robot deployments, has summarized the top three challenges that offer the most outstanding value proposition for industrial applications and customers:

1. Adaptation of welding to fit-up variation. Welding is highly sensitive to submillimeter variations in features like gap opening, root opening, bevel angle, and high-low misalignment [6]. Consistently controlling these variations within welding's expected tolerances is challenging, if not practically impossible. The need for intelligent adaptation to these fit-up variations has long been recognized as a significant pain point on the welding floor. Our solution operates independently of joint information, making real-time decisions based on the live interaction between the weld puddle and joint variations. This approach ensures that the puddle remains consistent throughout the welding process.
2. Adapting welding to tack detection and fusion. Tack welding is an indispensable step in weld preparation before robotic welding begins [7]. However, the manual nature of tack welding introduces significant variability in aspects like pattern, size, shape, profile, and spacing between tacks from part to part. As a result, robots must account for these variations each time a new part is introduced. Pre-scanning methods and techniques have proven counterproductive, and more importantly, developing prescriptive weld logic for robots to fuse each tack differently is an impractically unproductive and costly solution. Without proper adaptation and fusion of tacks, both weld aesthetics and quality suffer—an increasingly critical concern for top fabrication brands and their customers. Current solutions typically rely on a special tack feature that is familiar to the sensor. However, we have trained our model on a wide range of variations and enhanced it through generative AI, eliminating the need for any specific tach features requirement. This broadened training enables our AI system to detect tacks reliably without prior feature requirements.
3. Real-Time Seam Tracking. Seam tracking ensures the weld stays in the correct location, regardless of part movement or misalignment during the welding process [8]. Defects like lack of fusion can quickly occur with submillimeter deviations in seam alignment. Current seam tracking methods rely on pre-scanning, therefore becoming blind to the distortion of seam during welding and misalignments occurring after pre-scanning. Currently, seam tracking is commonly implemented Through Arc Seam Tracking (TAST), with laser tracking being an alternative option. While there are numerous references to these approaches, there are also well-known limitations that restrict their broader application. By integrating seam tracking directly into our system, we demonstrate that vision-based seam tracking offers a more widely adopted solution. Additionally, the integra-

tion of seam tracking analysis on a single sensor input, which also handles other control functions, makes it highly advantageous in terms of space and weight constraints around welding torches and end-of-arm robotic systems.

While there are additional challenges that demand intelligent adaptive welding, the authors have chosen to focus on these top three challenges and develop solutions specifically to address them.

This technical solution is designed to integrate the perception and cognition of welders into robots. It enables live adaptation without relying on pre-scan information, allowing for immediate adjustments to fit-up variations without time-consuming reprogramming. Additionally, it adapts to tacks in real time and accounts for seam and welding distortions without altering the initial path planning.

3 Solution approach

The solution architecture for each feature is outlined below. These solutions form the foundation for developing the main hardware and software pipeline, ensuring the functionality needed to support these strategies.

Without adaptation to fit-up variations, the pre-loaded weld procedure for a nominal fit-up delivers a fixed weld volume at a constant travel speed. This approach, while straightforward, fails to account for variations in the bevel geometry, such as changes in gap size, root opening, or bevel angle.

By looking at the fit-up geometry in the welding vicinity and tracking real-time bevel dimensions before entering the welding location, the system can dynamically adjust the welding parameters to tailor the weld volume to the actual bevel profile, ensuring proper fill and avoiding excessive material deposition. Such an adaptive approach optimizes material usage, improves weld consistency, and enhances overall quality by responding in real-time to part variations, eliminating the need for time-consuming manual adjustments.

This method represents a significant advancement in automated welding systems, reducing the risk of defects while increasing efficiency in high-mix, low-volume production environments (see Fig. 1).

The adaptive system must receive real-time signals from sensors to detect the approach of a tack and precisely define where the weld should enter and exit. Tack welding is a highly skilled process, with significant variation in the techniques used by different welders. Consequently, tacks can vary widely in size, shape, profile, and even spacing, creating a challenge for robotic systems to handle tack fusion consistently. Therefore, relying on predefined lookup tables for each tack style is impractical. Creating a specific fusion strategy for every possible tack variation would be time-consuming and inefficient. Instead, the system requires intelligent, adaptive logic capable of responding in real-time to each tack, regardless of its specific dimensions or characteristics (see Fig. 2).

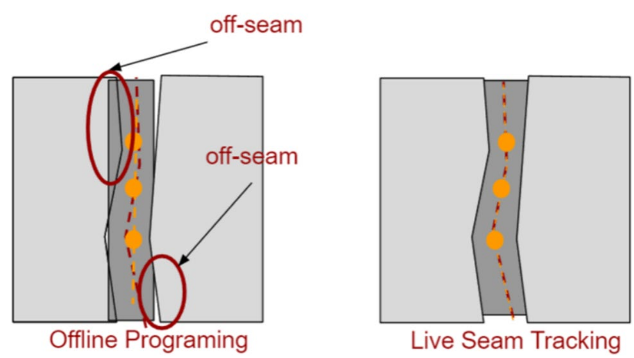


Fig. 1 Vision-based seam tracking

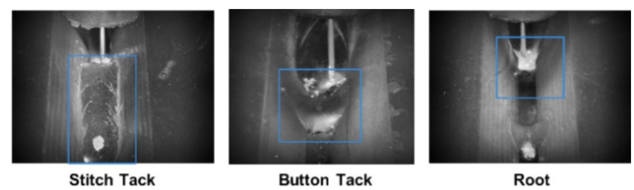


Fig. 2 Adapting robotic welding to tack detection and fusion

Seam tracking must correct the pre-loaded weld path by adjusting to real-time part variations. Using vision to detect the seam's location, the system compensates for shifts caused by misalignment, distortion, or inconsistent fit-up, ensuring precise weld placement. The solution must continuously monitor and compare the position of the joint center with the arc center to ensure proper alignment. If a deviation is detected during welding, the system instantly calculates and applies an offset to correct the seam on the fly (See Fig. 3).

4 System modules and integration

To replicate a welder's perception and cognitive abilities in robotic systems, a solution pipeline requires three key modules (see Fig. 4):

1. Vision System: This module acts as the welder's eyes, consistently monitoring the arc and weldment to provide real-time visual feedback.
2. Perception and Control System: Functioning as the welder's mind and skill center, this module cognitively

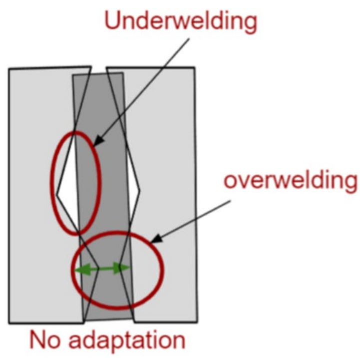


Fig. 3 Adaptation of welding to fit-up variation

Fig. 4 A solution pipeline for perception and cognition of welders in robots

processes the visual data and generates precise execution signals for the welding task.

3. Execution Module: This component replicates the dexterity of the welder's hands, managing the fine motor movements of the welding torch through the robot and adaptive parameter control through the power source.

By replicating the welder's eyes, mind, and hands through a vision system, cognitive control, and execution module, this solution pipeline provides a robust foundation for robotic adaptive welding. The proper integration of these modules not only facilitates precise and consistent welds but also creates a self-improving system that learns from each welding operation. As a result, robotic welding systems can evolve continuously, becoming more capable of handling complex and varied tasks with the same level of craftsmanship as a human welder, while offering greater efficiency and scalability in manufacturing environments.

5 Vision sensor

The vision sensor utilized in this application is an industrial-grade monochromatic camera designed specifically for real-time weld puddle monitoring in demanding industrial environments (see Fig. 5). Its sophisticated engineering combines advanced optical filtering, laser illumination techniques, and robust hardware to overcome the challenges posed by intense welding arcs. These arcs often produce overwhelming glare, but the camera's unique filtering and illumination methods effectively suppress this interference. The result is a vivid, clear visualization of the weld puddle, its boundaries, and the surrounding features critical for process monitoring and control.

The camera is positioned at a viewing angle of 30 cm away from the wire, directed toward the weld pool, while pointing in the opposite direction of the travel path. This configuration provides an optimal perspective for capturing critical features of the weld puddle and surrounding area. The optical axis of the camera is pitched at an angle between 45 and 55 degrees relative to the torch, with a yaw angle of 0 degrees, to keep the torch at the center of the view. A roll angle of 0 degrees ensures symmetry in the captured imagery, enabling consistent monitoring of the welding process across different scenarios. This carefully designed camera orientation maximizes the visibility of the weld pool while maintaining a stable and geometrically accurate representation of the scene.

To achieve precise spatial measurements, planar scaling factors are employed to convert pixel dimensions into millimeters. A reference plane is fitted at the position of the wire, with the horizontal axis defined as perpendicular to the optical axis and the vertical axis aligned parallel to the wire. This alignment compensates for the foreshortening effect caused by the pitch angle of the optical axis relative to the weld puddle. By correcting this geometric distortion, the system ensures accurate measurements of puddle dimensions and associated features, which are critical for effective process monitoring and control.

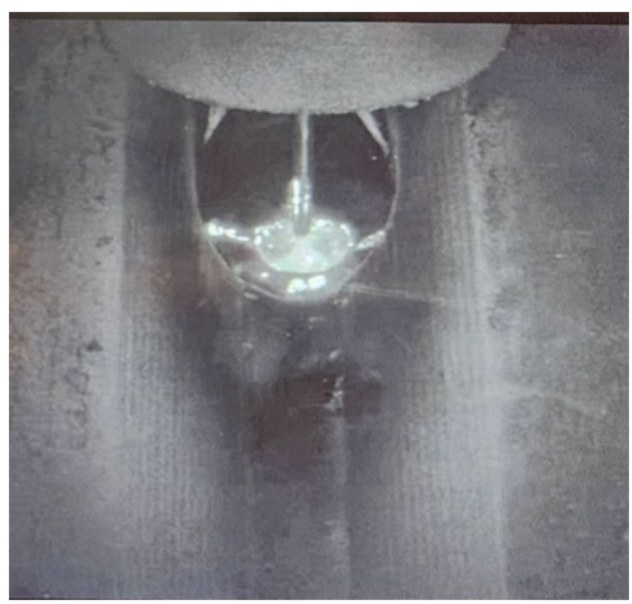


Fig. 5 Image taken with the welding camera from GMAW on a groove weld used in this work

With a resolution of 1440 × 1080 pixels and 8-bit sampling, the camera captures high-fidelity images that preserve intricate details of the weld puddle and its geometry. This level of precision is crucial for monitoring subtle variations and ensuring the detection of defects or irregularities. The field of view, spanning 58 × 43.5 mm, provides an optimal balance between capturing a detailed close-up of the weld puddle and maintaining sufficient spatial coverage for the surrounding area.

Operating at a frame rate of over 30 frames per second, the vision sensor ensures real-time responsiveness to dynamic changes in the welding process. Captured frames undergo onboard pre-processing to enhance their quality and are subsequently sub-sampled downstream to optimize computational efficiency for AI algorithms. This streamlined workflow enables accurate and efficient analysis without overloading processing resources, allowing the system to function seamlessly in real-time industrial applications.

The camera's ability to handle high-contrast scenes is supported by its 70 dB dynamic range, which ensures that both bright and dark regions within the weld zone are rendered with exceptional clarity. To further enhance image quality, the system employs Contrast-Limited Adaptive Histogram Equalization (CLAHE), a sophisticated image enhancement technique. CLAHE works by dividing the image into small, non-overlapping tiles and applying localized histogram equalization within each tile. This process enhances local contrast, making finer details more visible, while the contrast-limiting aspect prevents the over-amplification of noise in uniform areas. This capability is particularly beneficial in welding environments where intense glare and varying brightness levels can obscure critical features.

Designed to operate reliably in extreme industrial conditions, the camera incorporates a liquid cooling system that ensures consistent performance even in ambient temperatures exceeding 80 °C. This thermal management solution protects sensitive components and maintains optimal functionality during extended operations in high-temperature environments. Additionally, the camera's Gigabit Ethernet (Gig-E) interface supports high-speed data transmission, ensuring low-latency communication and seamless integration into industrial networks.

The combination of these advanced features positions this vision sensor as a pivotal technology in modern welding applications. Its high-resolution imaging, dynamic range capabilities, and robust real-time processing enable precise monitoring and control of welding parameters. By facilitating early detection of defects and enabling real-time feedback for process optimization, the sensor enhances both the quality and efficiency of automated welding systems. Furthermore, the incorporation of CLAHE and the ability to perform in extreme environments underscore its adaptability and reliability in diverse industrial settings.

This vision sensor exemplifies the integration of cutting-edge optical and computational technologies to address the unique challenges of welding applications. Its ability to capture high-quality images in real time under adverse conditions supports the development of intelligent automation, fostering advancements in manufacturing precision and productivity.

6 Vision perception

In analyzing Fig. 5, it is important to note that classifying each pixel in the image is crucial for generating useful intention signals for the Programmable Logic Controller (PLC). This process, known as image segmentation, focuses on distinguishing specific classes within each frame, including puddle, torch, wire, gap, tacks, and the edge of the bevel and background (7 classes).

6.1 Network architecture

To achieve this segmentation, we have chosen the U-Net architecture due to its strong ability to maintain image resolution. This is especially important for achieving sub-millimeter accuracy, as we need to define the widths of the root opening and puddle based on the segmentation masks. U-Net effectively meets our needs by preserving resolution while also allowing for real-time processing.

The U-Net is a fully convolutional neural network characterized by its U-shaped architecture, which gives rise to its name [11]. This structure comprises an encoding path and a decoding path. The encoding path, also referred to as the analysis path or contracting path, is designed to shape the typical convolutional neural network (CNN) by applying convolutions repeatedly, each followed by a rectified linear unit (ReLU) and a max pooling operation. Its primary function is to reduce the dimensionality of the input layers while increasing the number of feature channels. In this path, a 3 × 3 convolution is followed by ReLU and a 2 × 2 max pooling operation, which downsamples the input while doubling the number of feature channels. Conversely, the decoding path serves the opposite purpose of the encoding path; it reduces the number of channels while increasing the spatial dimensions of the layers. This path begins with an upsampling process that halves the number of channels using a 2 × 2 convolution. Subsequently, 3 × 3 convolution layers followed by ReLU are employed. Skip connections are utilized to concatenate the corresponding feature layers from the encoding path, thereby recovering information lost during the downsampling process. Ultimately, the dimensions of the layers are restored using a 1 × 1 convolution to produce a pixel-wise classified predicted map [11].

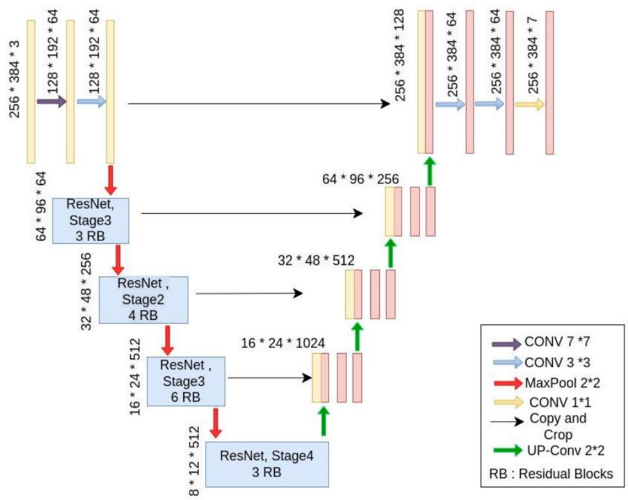


Fig. 6 Unet architecture with Resnet50 backbone neural network

To address the challenge of limited training data and to leverage the knowledge representations of low-level features acquired by networks pre-trained on extensive datasets for other computer vision tasks, the encoding path of the U-Net has been replaced with a ResNet-50 backbone pre-trained on the ImageNet dataset. ResNet [12], or the residual neural network, employs identity skip connections to mitigate the degradation problem that occurs when accuracy plateaus and then declines rapidly as network depth increases. This architecture is constructed by stacking multiple bottleneck residual blocks, each consisting of a series of 1×1, 3×3, and 1×1 convolutions. To preserve the information contained within the layers of the ResNet backbone during training, this backbone is kept frozen. In contrast, the U-Net decoder is left unfrozen and trainable, allowing it to adapt to the parameters of the pre-trained layers. Figure 6 illustrates the architecture of our U-Net model incorporating the ResNet-50 backbone. Figure 6 depicts the architecture of our U-Net model with a ResNet-50 backbone. Model parameters are 34 M and FLOPS of the model is 74.

Table 1 Per-class recall, precision, and F1 score

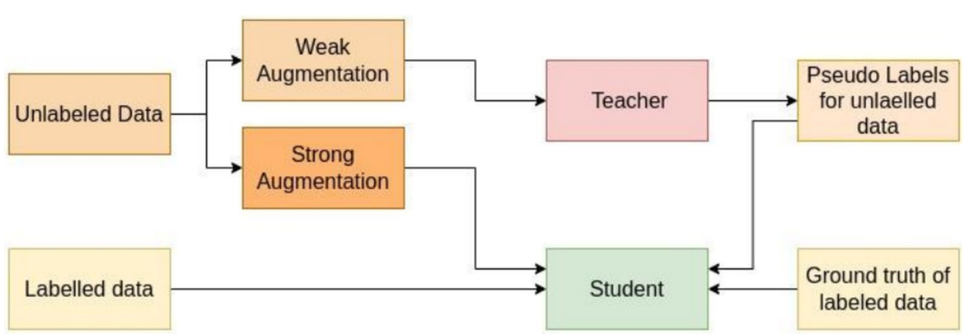
Classes	Recall	Precision	F1
Puddle	0.997	0.997	0.997
Gap	0.995	0.995	0.998
Tack	0.999	0.999	0.999
Torch	0.999	0.999	0.999
Wire	0.991	0.990	0.991
Bevel	0.857	0.857	0.856

6.2 Training procedure and results

Supervised learning is a machine learning paradigm in which a model is trained on labeled datasets, allowing it to learn direct mappings between inputs and their corresponding outputs. This approach is generally considered straightforward due to the explicit nature of the training data. However, a significant limitation arises from the necessity for extensive human annotation, which can be resource-intensive and time-consuming. Consequently, this dependence on labeled data restricts the scalability of supervised learning. To train the model in order to take advantage of both labeled and unlabeled data we adopt a semi-supervised approach for training. Semi-supervised is typically addressed in three ways, explicit consistency regularization [13], using a teacher-student framework [14], and recently, combining teacher-student with contrastive embeddings [15].

Teacher-student training pipelines utilize the mean-teacher framework, which maintains an exponentially moving average of the student model to generate smoother pseudo-labels. This methodology involves passing weakly augmented copies of images to the teacher model, which then produces pseudo-labels. Subsequently, the student model is trained on strongly augmented versions of the same image set. This approach enhances the quality of the learning process by leveraging the stability of the teacher model in providing consistent labels. [16]. We adopt the mean-teacher framework to generate pseudo-labels from the unlabeled data pool and train the model on these labels. In this

Fig. 7 Semi-supervised framework



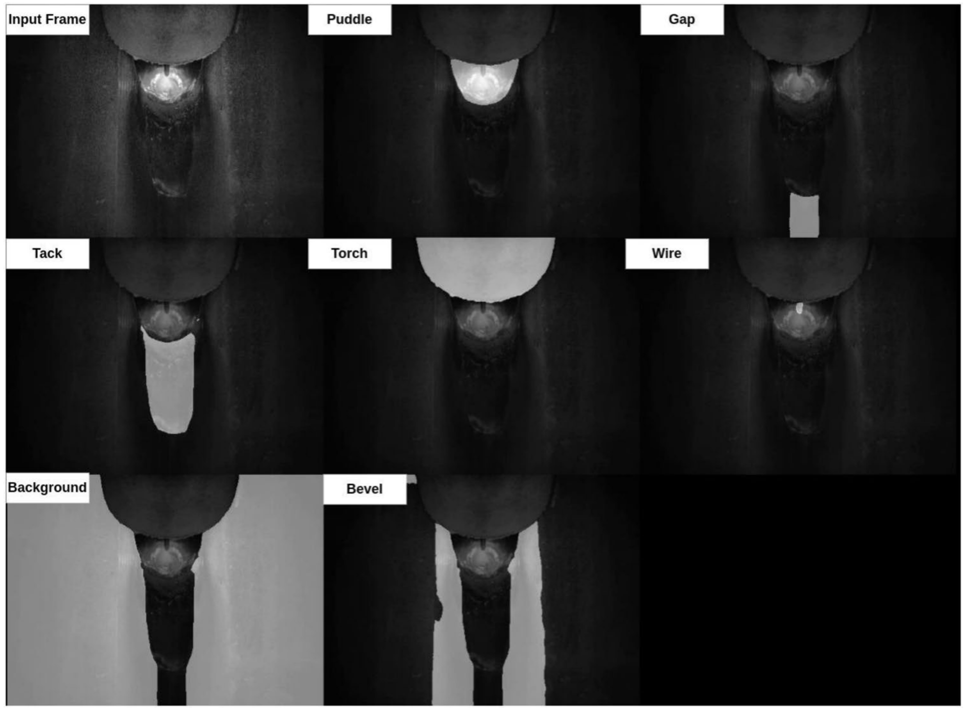


Fig. 8 Sample of model interference applied to an input frame, illustrating all seven classes represented in the image

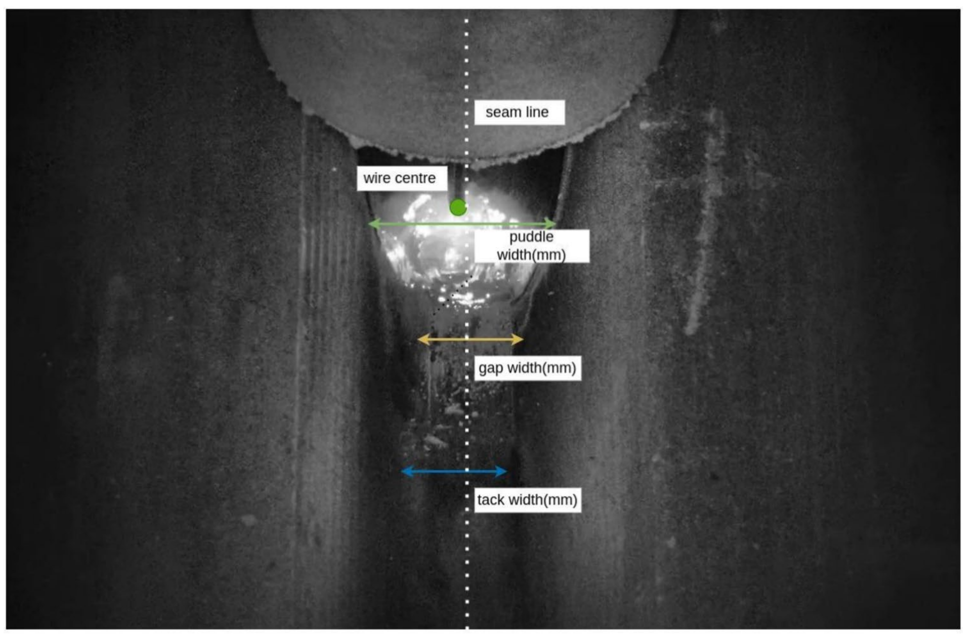


Fig. 9 Post-processing measurements derived from masks, which will be transmitted as intention signals to the Programmable Logic Controller (PLC)

Table 2 Welding and motion parameters for each state

	Recipe	*Wide Gap*	*Tight Gap*	*Tack*
Wire Feed Speed (ipm)	140–150	120–130	150–160	145–150
Peak (amps)	280–290	260–275	290–300	295–300
Background (amps)	80–90	60–75	90–100	95–100
Travel Speed (ipm)	5–6	3–4	6–7	7–7.5
Weave Amplitude (mm)	2–3	3–4	1–2	3–4

setup, both the student and teacher share the same network architecture, with the teacher updated using an exponential moving average (EMA) of the student's parameters. For the labelled images, the student is trained using a supervised loss, specifically cross-entropy, based on the ground truth information. The continuously updated teacher model utilizes EMA to produce pseudo-labels for the unlabeled data, which are then used to train the student network. The teacher

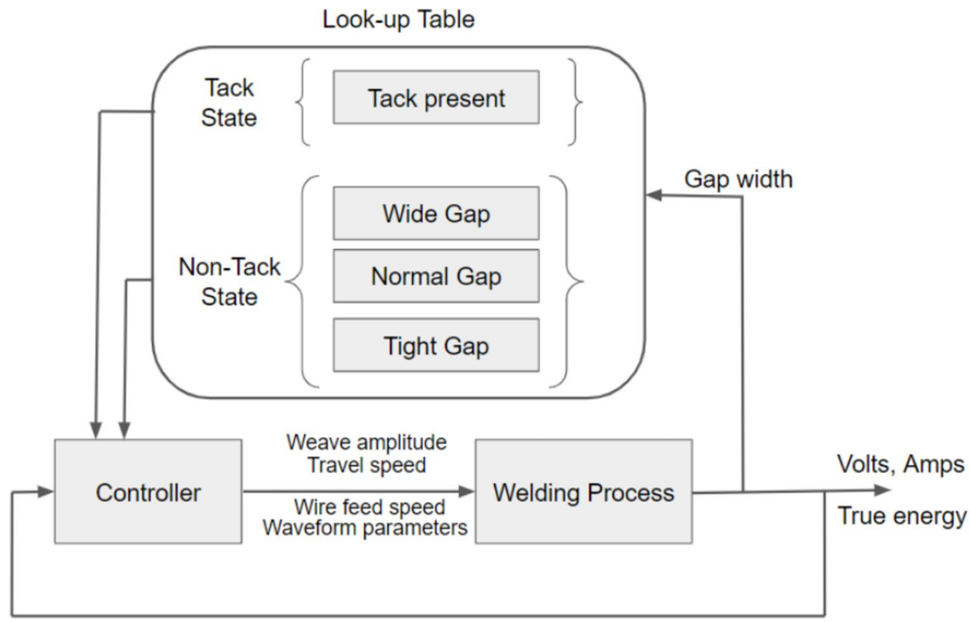


Fig. 10 Gain-scheduling adaptive control strategy

processes a weakly augmented version of the images to generate predictions, while the student model is trained with strong augmentations of the same images, treating the generated pseudo-labels as ground truths. The student–teacher framework has been represented in Fig. 7. Student loss follows the blew equation:

$$l_{total} = l_{supervised} + l_{unsupervised}$$

For the labeled images, the student is learned using a supervised loss (cross-entropy) with the ground truth information:

$$l_{supervised} = l_{ce}(\theta_S(x), y)$$

where $x$are the inputs to the network and y are the corresponding ground truth labels and θs is the student network. The teacher receives weakly-augmented version of the images Xweak to generate predictions and the student model is trained using the generated pseudo labels as ground truths following

$$l_{unsupervised} = l_{ce}(\theta_s(x_{weak}), \theta_t(x_{strong}))$$

where Xweak is the weak-augmented version as input to the teacher θt and Xstrong is the strong-augmented version as input to the student network.

For the semantic segmentation task on Labelled data for the student network, 80% of the annotated data is used for training, ensuring the model learns robust patterns from a broad set of examples. The remaining 20% is reserved for validation, providing a reliable measure of the model's generalization to unseen data. The accuracy of each class detection performed by the model was assessed through standard semantic segmentation metrics: Recall, Precision, and F1 score.

$$Recall = \frac{Truepositives}{Truepositives + FalseNegatives}$$
$$Precision = \frac{Truepositives}{Truepositives + FalsePositives}$$
$$F1 = 2 \times \frac{Recall \times Precision}{Recall + Precision}$$

The results of the evaluation on the test set of the student network are reported in Table 1.

6.3 Model inference and post processing

Figure 8 represents the view of model inference on an input frame and all 7 classes detected by the model. As it is shown in the figure, the model is very powerful in distinguishing the boundaries of Gap from Tack and puddle and torch.

Following the generation of binary masks for each class, it is essential to derive the intention signals necessary for the effective operation of the Programmable Logic Controller (PLC). One critical signal is the seam tracking offset, which is determined by measuring the distance between the wire tip and the seam center, as illustrated in Fig. 9. In addition, the widths of the gap and puddle are transmitted to the PLC, accompanied by a Boolean value indicating whether the puddle is positioned on the gap or in contact with a tack. All intention signals are depicted in Fig. 9.

7 Controller

Once the AI-based perception module has processed the frames and identified key features, the perception module generates three outcomes. One Boolean outcome of tack or non-tack status, one seam tracking offset and one gap width in millimeters. These signals are sent to a Programmable Logic Controller (PLC), which serves as the central hub for executing the welding instructions. The PLC maps the AI-generated signals into different states such as tack, wide gap, normal gap, and tight gap. For each state, adjustments to the welding parameters and motion parameters based on the specific needs of the weld are sent to the robot. Also, move left and right commands are issued to maintain the torch in the center of the groove according to the seam tracking offset.

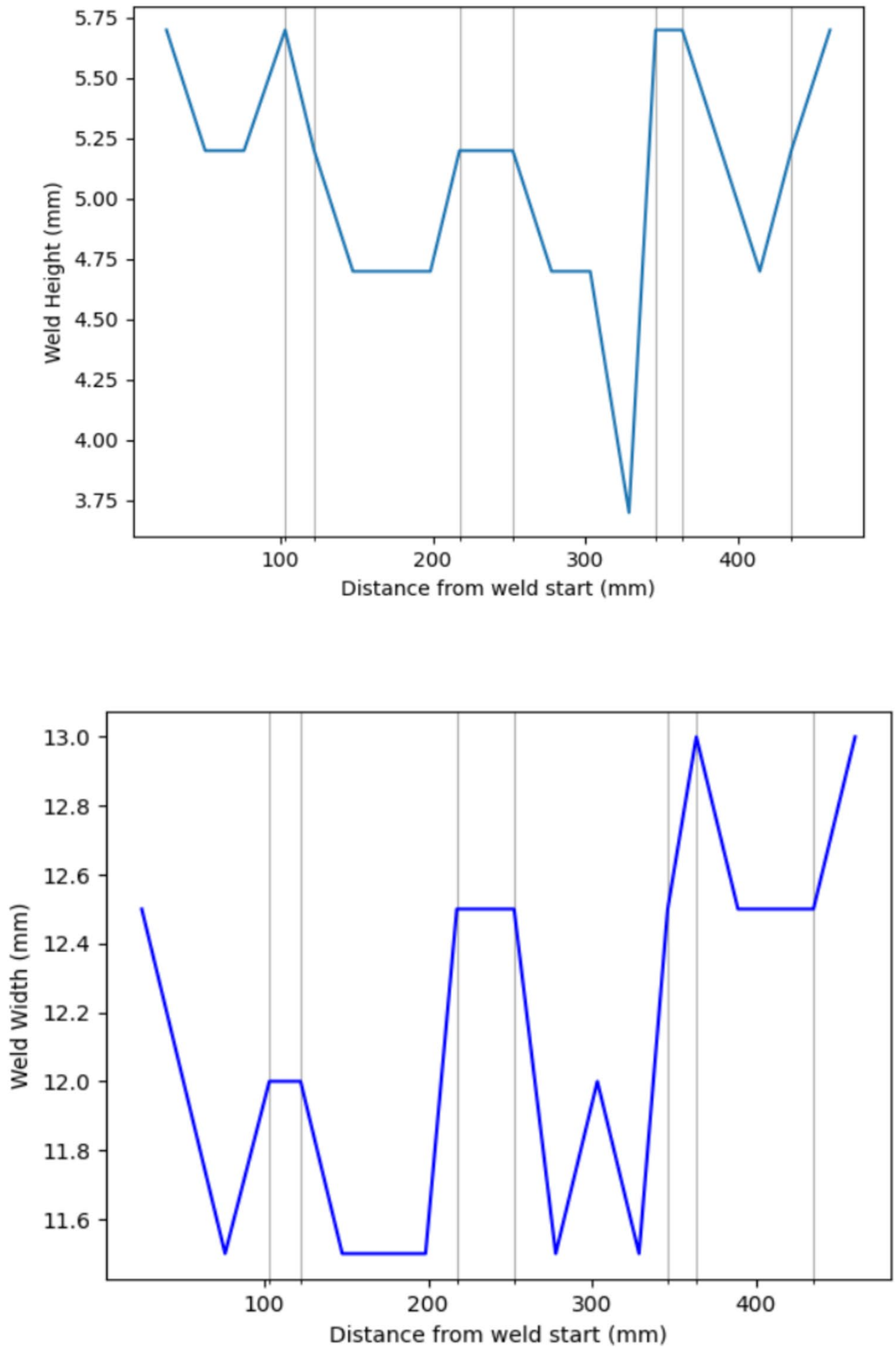
It is important to note that a sensitivity analysis was conducted to assess the impact of changes in each welding and motion parameter on factors such as amps, puddle height and width, and heat input. Based on the requirements of each state, only a subset of the most influential welding parameters needed to be adjusted to achieve the desired heat level, puddle width, and height. Our welding-tailoring approach to the general control method is similar to singular value decomposition (SVD), where only the parameters with the greatest impact on the outcome are considered. Table 2 summarizes the key parameters used for each state.

The embedded weld logic within the PLC ensures that these commands are not just static instructions but are dynamically adjusted based on real-time feedback from the vision system. This continuous feedback loop enables fine-tuning of the welding process, ensuring high-quality welds even under varying conditions, such as changes in fit-up.

The provided solution employs an approach similar to a gain-scheduling adaptive control strategy (see Fig. 10), tailored for welding processes. At its core, the vision-based system continuously monitors and measures the weld's features.

Fig. 11 A consistent, uniform weld bead (left) achieved across varying root openings, gaps, and diverse tacks within the joint (right), without prior information or pre-scanning provided to the robot

Fig. 12 The height and width variation (mm) along the weld. Gray lines show the tacks



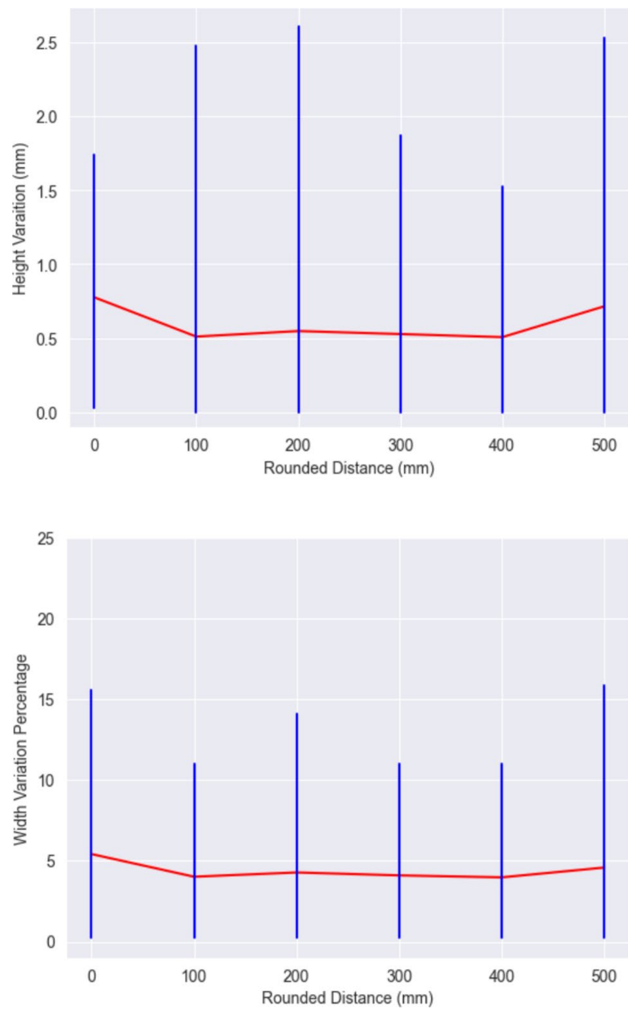


Fig. 13 Blue lines: Spread of deviations (from the average all along the part) in each 100 mm subsection. Red line: Median of deviations in each subsection

The controller makes decisions based on the weldment geometry. This control structure makes the weld bead more consistent in different geometrical variations.

8 Execution

Once the weld joint geometry is defined, the appropriate set of motion parameters (such as torch speed and angle) and welding waveform parameters (such as voltage, amperage, and waveform shape) is applied. This adaptive adjustment ensures that the welding process is fine-tuned to the specific conditions of the weld joint, leading to more precise and consistent weld quality. The lookup table functions as a decision-making tool that allows the controller to switch between different predefined parameter sets in real time.

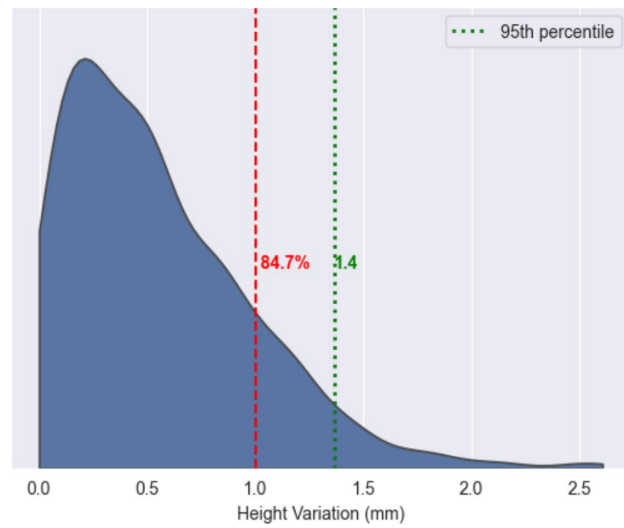


Fig. 14 Weld bead height magnitude variation (mm) for our pipeline performance and repeatability

This control structure ultimately leads to more consistent weld bead quality, even in situations where the weld joint and fit-up vary along the weld path. By dynamically adjusting both motion and process parameters in response to real-time weld joint measurements, the system adapts to changing conditions and reduces the need for manual intervention, making the welding process more efficient and reliable.

When designing adaptive robotic welding systems, traditional control-based systems, such as those using PLCs, offer a structured approach to managing welding parameters and adapting to joint variations in real-time. However, these systems have several inherent limitations that can restrict flexibility and overall adaptability. Transitioning to a learning-based policy, such as one enabled by imitation learning, has the potential to overcome these challenges by offering more dynamic and context-aware adjustments.

9 Adaptive welding results

Our AI model achieved a high confidence level of 99% in detecting features relevant to the welding process, such as joint position, gap width, and torch alignment. However, despite this high confidence, ensuring the robustness of the overall pipeline requires a thorough analysis of the potential impact of both false positives and false negatives on the final weldment quality.

A false positive occurs when the model mistakenly identifies a feature, such as a gap or tack, that is not actually present. For instance, the AI model might signal that a tack adjustment is necessary when the joint has no tack to fuse. This type of error can result in unnecessary adjustments to motion

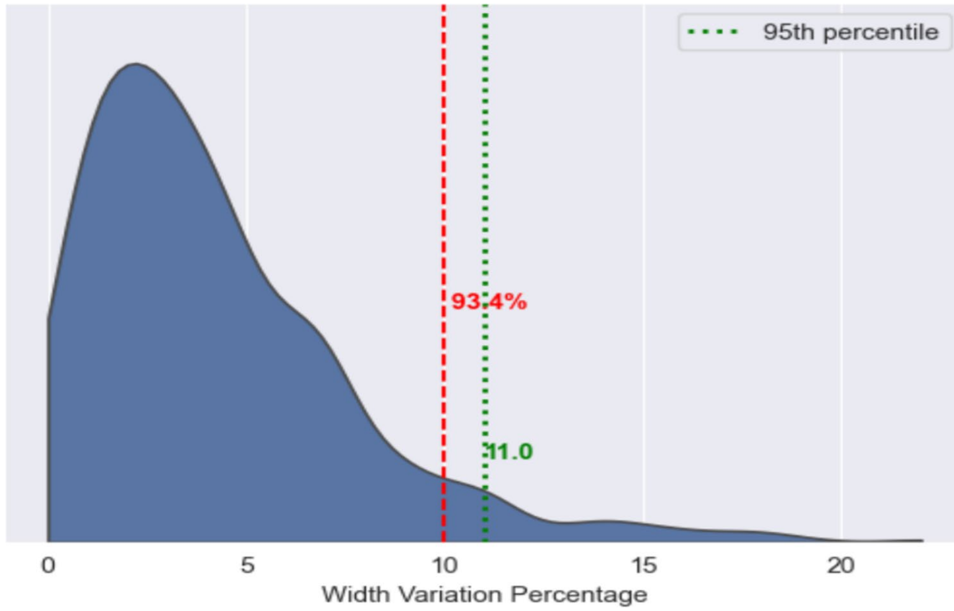


Fig. 15 Weld bead width % variation for our pipeline performance and repeatability

or welding parameters, potentially causing an excessive weld bead profile.

Conversely, a false negative occurs when the model fails to detect an actual feature, such as a joint misalignment or an abrupt change in gap width. Missing these critical signals could lead to improper weld bead formation, poor penetration, or insufficient fusion.

Quantifying false positive and false negative rates will allow the development of thresholds and tolerances within the system, leading to improved system reliability.

Our solution is robot-agnostic, having been successfully integrated and tested with UR, Vectis, ABB IRC5, and Yaskawa YRC1000 controllers. The results presented in this paper were achieved using a Vectis UR system, with similar outcomes observed across other robots, demonstrating the versatility and adaptability of our robot-agnostic approach.

Our solution delivers a consistent uniform weld bead across a joint that presents various challenges, including differing root openings, gaps, and an assortment of tack welds (example in Fig. 11). This was welded without any prior information or pre-scanning input to the robotic system. The robot's adaptive capability allowed it to compensate for these variations in real-time, ensuring a high-quality weld despite the inconsistencies typically found in the joint preparation. This method offers an advanced AI-driven welding systems to operate effectively in dynamic environments without the need for pre-weld data collection.

We defined the variation in puddle height and width along the weld bead profile as a quantifiable metric for assessing the consistency of the pipeline solution. A series of 36 welds in a flat groove joint configuration were welded, incorporating a range of varying weldment features, such as joint configurations, tack patterns, and geometries. Figures 12, 13, 14, and 15 present our statistical analysis of the pipeline's capability to achieve a consistent weld profile across these variations in joint and weldment features, as well as the repeatability of its performance. We measured the puddle height and width every 2.5 cm along the 50 Cm long weldment. Among over 700 measurements, the 95 percentile shows a variation in bead height of less than 1.4 mm and a variation in bead width of less than 11%.

Our AI model's high confidence level is promising, but fully assessing the impact of false positives and negatives allows the welding pipeline to achieve optimal robustness. Quantifying these errors ensures the system is both reliable and capable of maintaining the highest standards for final weldment quality, aligning with the demands of production.

10 Conclusion

We built a solution for key challenges in robotic welding: adapting to fit-up variations, handling tack weld fusion, and ensuring real-time seam tracking. Our vision-driven solutions enable precise, adaptive welding for consistent weld by combining a vision system, perception module, and execution component to adjust in real-time, closely replicating the perception and cognition of human welders in the robots.

The integration of AI learning strategies, including unsupervised, federated, and imitation learning, is pivotal to system scalability. These strategies reduce the need for exhaustive data annotation by enhancing the system's ability to generalize and adapt to new tasks autonomously.

Our results show high confidence in feature detection, with quantifiable metrics for assessing weld consistency. The system's ability to maintain a uniform bead profile across varying joint configurations and tack styles, alongside low false positive and false negative rates, highlights its robustness in meeting production-quality demands.

Authors' contributions Mahyar Asadi.
Solution architect and welding technical contribution.
Ahmad Ashoori.
Robotic and control technical contribution.
Mehrnoosh Afshar.
AI and machine learning technical contribution.
Ali Sheikhshab.
Data and measurements technical contribution.
Todd Scheerer.
Welding technologist for performing welds.
Austin Kaspardlov.
Robotic technologist for robot programming.
Soroush Bagheri.
Vision and sensor technical contribution.
Sina Firouz.
Chief technical officer and contribution.
Soroush Karimzadeh.
Chief execution officer and contribution.

Data availability The data supporting the findings of this study are available from the corresponding author upon reasonable request. Access to certain datasets may be restricted to comply with Novarc Technologies' policies on intellectual property protection and confidentiality. All requests will be reviewed to ensure adherence to these policies.

Competing interests The authors declare no competing interests.

References

1. Li X, Tao F, Cheng Y, Zhao L (2022) Digital twin-driven smart manufacturing: Connotation, reference model applications and research issues. Robot Comput-Integr Manuf 72:102260. https://doi.org/10.1016/j.rcim.2021.102260
2. Kumar S, Gaur V, Wu C (2022) Machine learning for intelligent welding and manufacturing systems: research progress and perspective review. Int J Adv Manuf Technol 123:3737–3765. https://doi.org/10.1007/s00170-022-10403-z
3. Zhang Y, Chen S, Yan W (2021) Intelligent robotic welding technology and application: Current status and future perspectives. J Manuf Process 62:508–523. https://doi.org/10.1016/j.jmapro.2021.01.004
4. Zhao Y, Li J, Gu H (2020) A review on research and application of intelligent welding technologies. Procedia CIRP 88:167–173. https://doi.org/10.1016/j.procir.2020.05.029
5. Biber A, Sharma R, Reisgen U (2024) Robotic welding system for adaptive process control in gas metal arc welding. Weld World 68:2311–2320. https://doi.org/10.1007/s40194-024-01756-y
6. Ding H, Zhang L, Zhang Y (2018) Adaptive control and optimization in robotic welding: A review of recent advances. Robot Comput-Integr Manuf 54:16–27. https://doi.org/10.1016/j.rcim.2018.05.001
7. Xu H, Yang Z (2020) Tack detection and adaptive fusion in robotic welding systems. J Manuf Process 50:650–657. https://doi.org/10.1016/j.jmapro.2020.10.007
8. Gao R, Wang H (2021) Real-time seam tracking and adaptive welding: A comprehensive review. J Intell Manuf 32(6):1593–1606. https://doi.org/10.1007/s10845-020-01641-8
9. Cavitar Welding vision systems. from https://www.cavitar.com/welding-vision-systems/
10. He B, Liu S, Wang F (2021) A study on the use of high-speed imaging for weld pool monitoring and control in gas metal arc welding (GMAW). Int J Adv Manuf Technol 114:1795–1804. https://doi.org/10.1007/s00170-021-06523-9
11. Ronneberger O, Fischer P, Brox T (2015) U-Net: Convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells WM, Frangi AF (eds) Medical Image Computing and Computer-Assisted Intervention 2015. Springer International Publishing, Cham, Switzerland, pp 234–241
12. He K, Zhang X, Ren S, Sun J. 2016 Deep residual learning for image recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Las Vegas NV USA. 770–778
13. Zhengyang Feng, Qianyu Zhou, Qiqi Gu, Xin Tan, Guan-gliang Cheng, Xuequan Lu, Jianping Shi, and Lizhuang Ma.Dmt: Dynamic mutual training for semi-supervised learning. arXiv preprint arXiv:2004.08514 2020
14. Sudhanshu Mittal, Maxim Tatarchenko, and Thomas Brox. Semi-supervised semantic segmentation with high-and low- level consistency. IEEE transactions on pattern analysis and machine intelligence 2019
15. Yanning Zhou, Hang Xu, Wei Zhang, Bin Gao, and Pheng-Ann Heng. C3-semiseg: Contrastive semi-supervised segmentation via cross-set learning and dynamic class- balancing. In Proceedings of the IEEE/CVF International Conference on Computer Vision pages 7036–7045 2021
16. Antti Tarvainen and Harri Valpola. Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results. In Proceedings of the 31st International Conference on Neural Information Processing Systems NIPS'17 page 1195–1204 Red Hook NY USA 2017. Curran Associates Inc

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.