Discover Applied Sciences

REVIEW

Open Access

Visual and thermal monitoring techniques for Wire Arc Additive Manufacturing: from conventional to machine learning approaches

Maedeh Gourakani[1,2*], Carmelo Mineo[1,2*], Donatella Cerniglia[1], Andrea Bertazzo[3], Federica Tiso[3] and Giovanni Avallone[3]

*Correspondence:
Maedeh Gourakani
maedeh.gourakani@unipa.it
Carmelo Mineo
carmelo.mineo@icar.cnr.it
[1]Department of Engineering, University of Palermo, Viale delle Scienze, 90128 Palermo, Italy
[2]Institute of High Performance Computing and Networking (ICAR), National Research Council (CNR), Via Ugo La Malfa, 153, 90146 Palermo, Italy
[3]CARACOL Innovation Hub, CARACOL, Via Giovanni Amendola, 25, 20037 Paderno Dugnano, Italy

Abstract
Wire Arc Additive Manufacturing (WAAM) offers high deposition rates and low feedstock cost, but process variability can cause defects, residual stresses, and geometric errors. No single sensor captures the full process state, motivating multimodal monitoring. This review traces progress from vision- or thermal-only monitoring to machine-learning approaches for vision–thermal fusion. We summarise sensing hardware and measurable parameters; outline traditional image/thermal analyses; review recent ML methods for unimodal data (classification, detection, transfer learning, physics-informed learning); and discuss fusion architectures (early/mid/late), including their synchronisation assumptions and practical guidance for trigger-based versus learned alignment. We then connect intelligent monitoring to control strategies (inter-layer, intra-layer, and hybrid), highlighting opportunities for model-predictive and adaptive control. Open challenges include robust sensing in arc environments, sim-to-real fidelity, data scarcity, and generalisation across materials and setups. Overall, integrating complementary sensors with ML can improve process interpretation and control, supporting the transition toward more autonomous WAAM.

Keywords WAAM, Process monitoring, Vision, Thermography

1 Introduction

Wire Arc Additive Manufacturing (WAAM) is a technique in which metal wire is continuously fed through a nozzle and melted by an electric arc to build components layer by layer. The first patents for this technology were submitted in 1925, indicating that the fundamental principles had been recognised for almost a century. At that time, the technology was primarily used for conventional industrial applications, and its potential was assessed in terms of manufacturing capabilities. Nevertheless, researchers did not start a thorough investigation into the adaptation and application of this technology to additive manufacturing techniques until the 1990 s. This newfound interest was spurred by advancements in digital manufacturing and materials science, as well as an increasing demand for more efficient, adaptable manufacturing techniques to produce large metal components. Since that time, technology has advanced significantly to meet the specific requirements of additive manufacturing. Using an electric arc as a more efficient fusion source than alternative additive manufacturing methods [1], as well as wire as a cost-effective and readily available feedstock [2–4], represent the primary advantages of this approach. This combination enables significantly higher deposition rates and improves material utilisation ratios compared to powder-based additive manufacturing technologies, making WAAM especially suitable for fabricating large, complex components in sectors such as aerospace and marine [5–7]. A primary factor contributing to WAAM's economic advantage is the feedstock itself. According to [8], welding wire is a widely available industrial commodity that costs significantly less than metal powder, with steel wire priced between €2.36 and €17.70/kg and titanium wire between €114.46 and €283.20/kg. Due to this substantial difference in raw material costs, wire-based additive manufacturing techniques are more cost-effective than powder-based techniques. Furthermore, Powder Bed Fusion (PBF) techniques require stringent, capital-intensive infrastructure for powder handling, storage, and recycling to prevent material degradation over multiple builds [9]. WAAM eliminates many of the operational complexities and safety risks linked to powder management. Taken together, these factors make WAAM a streamlined and economically attractive manufacturing method for producing large-scale components, especially in high-value alloys where material and operational costs are critical considerations [10]. Despite these significant advantages, WAAM requires precise control of the process parameters, and a lack of precise control leads to quality fluctuations between parts and even within a single part. In high-performance industries, such as aerospace, real-time monitoring of the deposition process is crucial to meet strict quality standards. However, one key challenge is that no single sensing method can fully represent the WAAM process's complete state. Therefore, a practical monitoring framework must integrate data from multiple sensing modes to simultaneously capture both the thermal and morphological aspects of the process. Multimodal approaches enable a more holistic understanding of defect-formation mechanisms and provide a basis for closed-loop process control.

Among various monitoring methods, thermal and vision-based sensors stand out for their ability to monitor the two most crucial aspects that determine the quality of WAAM components: thermal history and geometric fidelity. The uniformity of the thermal fields governs the final microstructure and mechanical properties. Thermal sensors, such as IR cameras and pyrometers, are commonly used to directly monitor molten-pool and interpass temperatures, providing a basis for controlling the process thermal fields, which are essential for reducing residual stress and preventing harmful phase transformations [11]. Advanced thermal imaging systems can also be used to extract geometric information, for example, by tracking melt-pool boundaries or estimating bead width and layer height from temperature fields. [12]. Complementing this, vision sensors provide crucial geometric and stability data that thermal sensors cannot provide. Laser profilometers and structured light systems provide real-time measurements of bead geometry, which can be fed back to adjust process parameters such as torch travel speed, wire feed rate, or toolpath offset. Although such feedback strategies have been demonstrated in controlled environments, their robust real-time implementation in WAAM remains challenging due to process latency, sensor occlusion, and computational constraints. Simultaneously, alternative imaging modalities, such as high-speed imaging and spectroscopy, provide detailed insight into arc and melt-pool stability. Specialised high-speed welding cameras, operating at high acquisition rates, capture transient phenomena and enable the early detection of conditions that lead to porosity or contamination [13–15].

Although many studies have demonstrated the effectiveness of either thermal or vision-based monitoring when used separately [11, 16], a clear, rapidly growing trend is toward combining these approaches. Integrating thermal and visual data enables a more comprehensive, multidimensional understanding of the process. This cannot be achieved with a single sensor. This integration can improve defect detection [17, 18], process interpretation [17], and control [18–20]. Moreover, there is a need to synchronise data streams at different temporal and spatial resolutions, which can be technically complex. A further challenge is managing and processing large volumes of heterogeneous data while extracting meaningful, correlated features that accurately reflect the underlying processes. Deep learning techniques have enabled the automatic extraction of features and the recognition of patterns from complex, multimodal datasets. Despite these advances, the current research landscape remains fragmented. Studies focus on isolated aspects rather than offering a cohesive framework for intelligent sensor fusion. To our knowledge, few reviews systematically trace the transition from single-sensor monitoring to deep-learning-enabled fusion in WAAM and its implications for process control. This review synthesises current advances and open challenges. This paper outlines a structured pathway toward more autonomous, intelligent manufacturing systems and clarifies existing data fusion methodologies.

This work is organised as illustrated in Fig. 1. As shown, this paper is structured to provide a comprehensive overview of the field, beginning with the fundamental principles of vision and thermal sensing in WAAM, including their respective hardware, measurable parameters, and traditional analysis techniques. Next, the review explores both image-based and thermal-based machine learning approaches for analysing unimodal

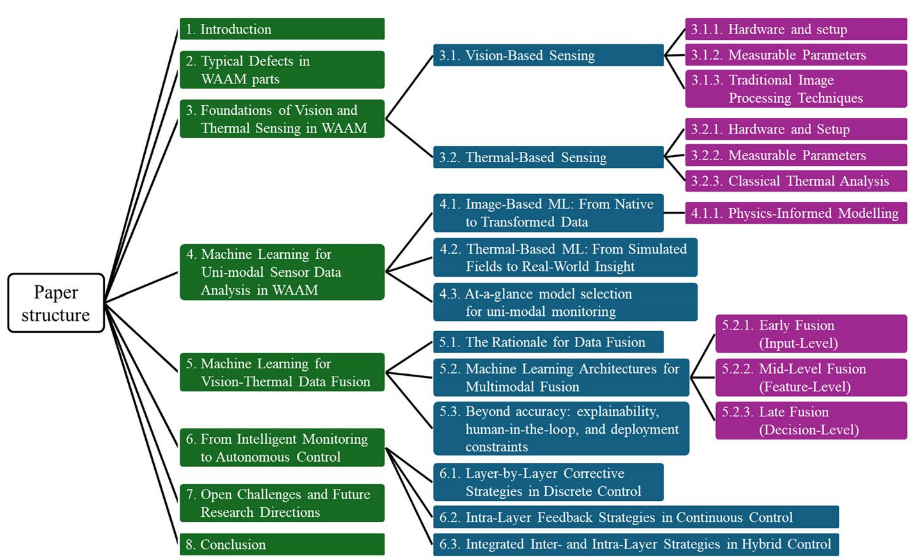

Fig. 1 Organisation of this paper

sensor data. Building on this, the following section delves into the complexities and advantages of deep learning for vision-thermal data fusion, detailing the rationale and various architectural strategies. The paper then progresses from the concept of intelligent monitoring to the goal of autonomous control. Finally, we identify and discuss the significant open challenges, outline promising directions for future research, and present our concluding remarks.

2 Typical defects in WAAM parts

Wire Arc Additive Manufacturing (WAAM) is an emerging technology with significant potential for producing large-scale metallic components due to its high deposition rates and cost-effectiveness. However, the welding process is susceptible to various defects that can compromise the structural integrity of the final part. In the WAAM process, each deposited layer undergoes partial cooling via heat dissipation to its surroundings before subsequent layers are deposited and reheated [21]. In this way, complex and non-steady-state thermal histories are generated [22]. These thermal cycles are a significant contributor to defect formation, as they create non-uniform cooling and reheating patterns throughout the part. The thermal history of the process is characterised by considerable heat buildup, with temperatures in the upper layers potentially increasing significantly while the lower layers continue to cool. This thermal mismatch creates a noticeable temperature gradient across the build height [23], leading to distinct defect types in WAAM parts, as shown in Fig. 2. This section synthesises findings from journal and conference literature to provide an overview of WAAM defects.

One of the most common and critical defects in WAAM components is porosity. These internal voids act as stress concentrators, significantly reducing mechanical properties such as strength and ductility [24]. Porosity in WAAM can originate from multiple sources. A primary cause is gas entrapment within the solidifying metal, which may include shielding gas or, more commonly, hydrogen. During solidification, hydrogen is expelled from the solid, forming gas pores that become trapped within the microstructure. The hydrogen typically originates from contaminants such as moisture or hydrocarbons on the feedstock wire. In addition to process stability, gas entrapment is strongly influenced by arc behaviour and melt pool dynamics.

Cracking is another critical failure mode that can render a WAAM component unusable. Defects initiated in a single layer can propagate to subsequent layers, where they may be magnified or exacerbated, ultimately compromising the component's structural integrity [29]. These defects arise from mechanisms associated with the severe thermal cycles of the process. Solidification cracking (or hot cracking) occurs during the final stages of solidification at high temperatures [30]. As the deposited metal cools and shrinks, tensile stresses develop. If these stresses exceed the strength of the partially liquid metal, cracks can initiate and propagate along grain boundaries [4]. Alloys with a wide solidification temperature range are particularly prone to this type of cracking. Hydrogen-induced cracking (or hydrogen embrittlement) is a delayed cracking phenomenon that can occur near room temperature. Hydrogen absorbed into the molten pool can become trapped within the solidified metal. Even small amounts of hydrogen can migrate to highly stressed regions and accumulate at microstructural features, such as grain boundaries, weakening atomic bonds [31]. This mechanism, known as hydrogen-enhanced cohesion, increases susceptibility to crack initiation and growth under stress.

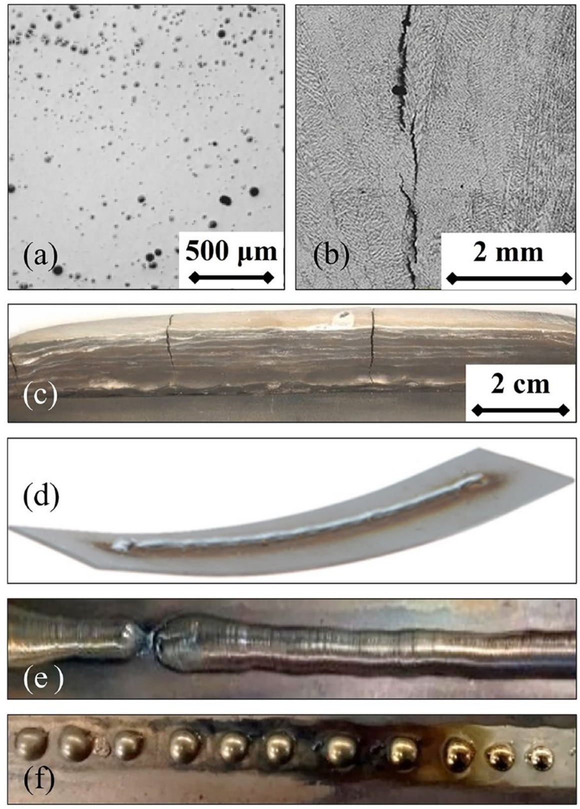

Fig. 2 Typical defects: **a** Porosity in AA5356 WAAM wall built with Argon Q1 shielding gas (adapted from [24], licensed under CC BY 4.0), **b** Cracks found in WAAM material microstructure after heat treatment (adapted from [25], licensed under CC BY 4.0), **c** Cracks observed by unaided eye on Fe-based Shape Memory Alloy (FeSMA) wall deposit (adapted from [26], with permission granted to Springer Nature Limited to publish under CC BY 4.0), **d** Example of weld seam geometry and sample distortion (adapted from [27], licensed under CC BY 4.0), **e** Abnormal bead with bead-cut (adapted from [28], with permission granted to Springer Nature Limited to publish under CC BY 4.0), **f** Abnormal bead with balling (adapted from [28], with permission granted to Springer Nature Limited to publish under CC BY 4.0). Changes limited to image cropping. CC BY 4.0 licence details available at: https://creativecommons.org/licenses/by/4.0/

Consequently, cracks are more prone to initiation and growth under stress because atomic bonds are weakened.

WAAM generates significant residual stresses and distortions during its process [10]. Rapid localised heating followed by cooling, repeated for each layer, produces cyclic thermal strains [32]. Expansion during heating is constrained by surrounding cooler material, resulting in compressive stresses, whereas contraction during cooling induces tensile stresses. The thermal history of this complexity and non-uniformity results in a locked-in internal stress state within the finished part. When these residual stresses exceed the material's yield strength, plastic deformation may occur, leading to macroscopic distortions such as warping, buckling, or dimensional inaccuracies [30]. Differential thermal contraction generates significant residual stress, which mainly contributes to component distortion, cracking, and reduced fatigue life [21]. Lack of fusion is characterised by incomplete bonding between adjacent deposition beads or between a deposited layer and the substrate. This defect acts as a crack-like discontinuity, severely reducing the load-bearing capacity. Insufficient energy input typically leads to insufficient fusion, preventing the underlying layer or substrate from fully melting. Factors such as excessive travel speed or inadequate heat input can prevent a strong metallurgical bond between layers.

In addition to internal defects, WAAM components are prone to geometric and surface irregularities that can affect part shape and surface quality. Balling is the formation of discontinuous, spherical beads due to insufficient energy density, while humping is a common defect characterised by surface roughness or bulging. These defects are typically linked to the stability of metal transfer, the properties of the molten pool, and precise control of deposition parameters.

In summary, the quality and reliability of WAAM components are threatened by a range of defects. These anomalies range from internal porosity and cracks to macroscopic distortions and surface irregularities. They are closely related to complex thermal cycles, the behaviour of the melt pool, and the solidification phenomena that occur during the process [33]. Understanding the root causes of these defects is the first and most critical step toward developing strategies to mitigate them and ensure the successful industrial adoption of this promising manufacturing technology. Table 1 maps the most common defects to their observable indicators, suitable sensing methods, applicable machine-learning (ML) tasks, and relevant literature references.

3 Foundations of vision and thermal sensing in WAAM

3.1 Vision-based sensing

Vision sensors capture and interpret visual information using cameras and advanced image-processing algorithms. These systems can extract features such as patterns, geometries, colour distributions, and surface textures. This capability supports advanced tasks, including object recognition, positioning, dimensional measurement, and quality inspection. Their adaptability and precision make vision sensors valuable tools across a wide range of industrial and commercial applications, including manufacturing, robotics, logistics, and automation. In the WAAM industry, these systems provide an intuitive visualisation of the process state. This capability enables the measurement of critical variables and the detection of surface-level defects. This section will cover the most common hardware and setups used in these processes. It will offer an in-depth discussion

Table 1 Mapping of defects, observable indicators, suitable sensors/metrics, machine-learning task type and relevant references

Defect/feature	Observable indicators	Common sensing method	Common ML task	Ref(s)
Porosity	Surface pores, plume instability, burst-type elastic waves	Vision (external pores), Acoustics (internal)	Classification, anomaly detection	[34]
Lack of fusion	Low penetration, local cold regions	Thermal, electrical	Classification, anomaly detection	[35]
Geometry deviation	Bead width/height variation, layer misalignment	Vision, Geometry mapping	Regression, Segmentation	[36]
Spatter	High-velocity droplet ejection	Vision, Acoustics	Object detection, Anomaly detection	[37]
Cracking	Transient elastic waves; strain concentration	Thermal, Acoustics	Anomaly detection, Classification	[38]
Melt-pool geometry	Pool size, shape, dynamics (oscillations)	Vision, Thermal	Segmentation, Regression	[39, 40]
Process instability	Irregular layer formation or waviness, Deposition fluctuation	Vision, Acoustics	Clustering, Anomaly detection, Time-series forecasting	[41]

of the primary parameters that can be quantified. Additionally, the section will explore conventional, non-AI image processing techniques employed for data extraction, along with their inherent limitations.

3.1.1 Hardware and setup

The camera is central to vision-based monitoring. Selecting the proper imaging hardware is crucial in the harsh environment of WAAM, given the intense welding arc light, high temperatures, and flying spatter. The literature reveals a range of camera technologies employed, each with its own strengths and ideal applications.

A lower-cost option for monitoring the build geometry and detecting large-scale defects is the use of a standard industrial camera, which typically employs a Complementary Metal–Oxide–Semiconductor (CMOS) or Charge-Coupled Device (CCD) sensor [16]. However, the intense arc glare often saturates the sensors, obscuring the critical melt pool region. To overcome this, researchers have turned to more specialised imaging solutions.

High-speed cameras are useful for capturing the rapid dynamics of the melt pool and metal transfer. These cameras enable detailed analysis of droplet detachment, arc stability, and spatter generation, which are crucial for maintaining process stability and ensuring part quality. High-dynamic-range (HDR) imaging helps capture both the bright arc and darker surroundings without saturation. By effectively combining multiple exposures or utilising specialised sensor technologies, HDR cameras prevent pixel saturation in high-intensity zones while preserving detail in shadowed areas. This allows simultaneous observation of the welding arc, melt pool, and background. This results in a more comprehensive and informative view of the welding process [42].

Beyond single-camera, or monocular, systems, stereo vision is gaining traction in WAAM [43]. By employing two cameras with a known spatial relationship, stereo vision systems can reconstruct the three-dimensional (3D) geometry of the deposition area. This allows for precise in-process measurement of bead width, height, and overall layer topography, providing direct feedback for closed-loop control of the robot's path and process parameters.

To achieve even greater 3D measurement accuracy, structured light systems and laser scanners are integrated into WAAM setups [44, 45]. Structured light systems project a known pattern of light onto the workpiece, and a camera captures the pattern's distortion to calculate the surface geometry. Laser scanners, which utilise a laser line and a camera to triangulate 3D points, provide an additional robust method for in-situ metrology. These active vision systems are particularly effective for post-layer inspection and detailed 3D models of the as-built part.

The successful implementation of any vision system in WAAM is critically dependent on the lighting and filtering techniques used to mitigate the arc's overwhelming brightness. A common and highly effective approach is laser illumination in conjunction with narrow-bandpass filters [43, 45, 46]. A laser with a specific wavelength illuminates the area of interest. A corresponding filter is placed in front of the camera lens, allowing only that particular wavelength of light to pass through. This effectively blocks out broadband light from the arc, providing a clear view of the melt pool and surrounding features. Pulsing the laser in synchrony with the camera's shutter further enhances image quality and reduces heat input from the illumination source [47, 48].

The physical setup and protection of vision hardware are crucial for reliable and continuous operation. Sensors are typically mounted on the welding heads or the robotic arm, often with a look-ahead or trailing view of the deposition process [49]. It is crucial that this positioning is carefully considered to ensure an optimal field of view without interfering with the welding torch or shielding gas flow. Spatter shields and cooling systems are essential for protecting sensitive optics from harsh environments. These enclosures often incorporate replaceable windows.

3.1.2 Measurable parameters

In WAAM processes, the energy from the electric arc melts the feedstock wire and the substrate surface, creating a liquid region known as the melt pool. The geometric properties of the molten pool, such as width, length, depth, and area, are directly influenced by heat input, energy distribution, and fluid flow dynamics. Vision-based monitoring enables real-time measurement of these geometric parameters, which serve as indirect indicators of the process's thermal state. The molten pool width is typically used to predict the ultimate weld bead width, while its length can reveal variability in travel speed or heat conduction. Variations in the shape or surface area of the molten pool are direct signs of process instability, which may be caused by factors such as non-uniform shielding gas coverage or arc power variability. A dedicated in-situ optical system by [50] measures melt-pool dimensions in real-time and relates them to process parameters (e.g., voltage, wire-feed speed) via a regression model. Their study demonstrates that it is feasible to accurately measure the dynamic melt pool geometry with the proper hardware setup. Additionally, the arc can be monitored for arc length and positional stability, both of which are crucial for ensuring consistent energy delivery.

The most common application of vision sensing is to measure the geometry of the just-solidified bead in a "trailing view." This approach trades the leading-indicator information from the melt pool for a higher signal-to-noise ratio. The primary geometric characteristics of this individual deposit are its bead width and bead height. For instance, the work in [51] describes a real-time system for measuring bead width and centreline position. To progress beyond mere 2D measurements, more sophisticated hybrid schemes, such as the one proposed in [52], compute the bead height by fitting a geometric model to the directly measured width. These bead-level features matter because they determine layer height, the vertical increment upon which subsequent material is deposited, a critical parameter for the dimensional accuracy of the final part. In Fig. 3, it can be seen visually and graphically that layer height is an important parameter affecting the quality and uniformity of the deposit. The figure shows the first cylinder (C1), produced using a straight-deposition strategy, in which each layer was deposited directly on the previous one. This caused slumping and height fluctuations at the start and end points that increased with the number of layers. In contrast, the second cylinder (C2) employed a staggered deposition strategy, where the start and end points of each layer were slightly shifted relative to those of the previous layer, resulting in nearly uniform layer heights and a flatter top surface. The diagram illustrates these differences through height-variation plots and cumulative layer-height diagrams for both C1 and C2 [52].

Monitoring these parameters in real-time is critical for detecting surface defects. Research by [16] demonstrates the automated vision-based detection of spattering, an instability that compromises surface quality. More subtly, the quality of the interior can

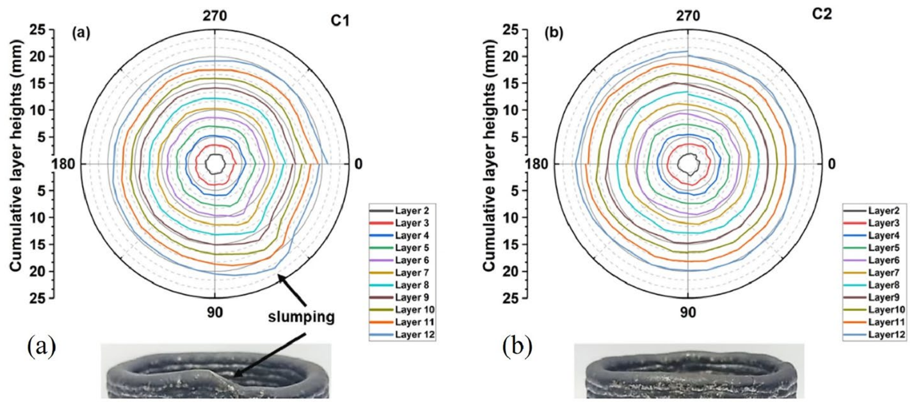

Fig. 3 Circular plots showing the cumulative layer heights for the C1 (**a**) and C2 (**b**) cylindrical deposits. The inset images show the top surface of the deposits. Reproduced from [52], with permission granted to Springer Nature Limited to publish under CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/)

be inferred through visual inspection. The work in [53] demonstrated a key link between surface and internal defectivity by using a process camera to monitor surface cavities on aluminium alloys. These authors established that the empirically evident cavities visible in the melt pool, formed by gas escaping, directly correlate with the amount of undesirable internal porosity, making surface monitoring a viable proxy for volumetric quality control.

While these systems focus on direct detection, more advanced approaches use vision data to train deep learning classifiers for anomaly detection [37] or employ high-fidelity offline metrology to characterise complex geometric deviations of the kind of surface waviness [54] that define the objectives for the next generation of monitoring systems.

Beyond monitoring the deposited material, vision systems also track the position of the deposition tool, as the tool's alignment relative to the workpiece is a critical control parameter. A rudimentary approach does not directly track the tool's position. The bead centreline serves as a proxy to answer: was the material deposited where planned? In one study, Couto et al. [51] developed a system using a classical computer vision pipeline to detect bead edges in a noisy Cold Metal Transfer (CMT) process and compute the centreline as the midpoint between the two detected edges. However, if the bead forms asymmetrically (for example, due to uneven heating or gravity), that midpoint may not correspond to the true centre of mass of the deposited material. Additionally, their method relies on accurate detection of the bead's edge. The Contact Tip to Workpiece Distance (CTWD) primarily governs wire melting rate and heat input. At the same time, the Nozzle-to-Workpiece Distance (NWD) ensures effective gas shielding and, by extension, influences arc stability, making both critical geometric parameters in WAAM.

3.1.3 Traditional image processing techniques

The fundamental pipeline begins with pre-processing to improve quality and highlight relevant features. Common steps include converting images to greyscale, defining a Region of Interest (ROI) to reduce computational load, and applying filters like the Gaussian or median filter, which is a linear filtering method that operates by scanning each pixel in the image with a convolution to mitigate process noise from fumes and spatter [50, 51]. The next key stage is segmentation, in which the bright feature (melt pool or bead) is separated from the darker background. While basic thresholding may be applied, more sophisticated methods have been developed. For example, Couto et al. [51] introduced an adaptive thresholding approach that modifies the threshold based on the local context or neighbourhood of each pixel. The melt pool area in another work is segmented using the Otsu algorithm [55]. This automatic thresholding method measures separability by calculating the inter-class variance between the background and target areas. Afterwards, they used morphological operations, specifically erosion to separate connected objects and dilation to expand the reduced area, approximately restoring its original shape (melt pool). Further, Dong et al. [50] used these morphological operations to clean the binary image and remove noisy artefacts in the melt pool area before contour extraction, resulting in more accurate melt pool boundaries. For the final feature extraction, a standard approach is to apply the Canny edge detector, followed by a Hough transform, to identify linear bead edges [50]. Couto et al. [51] initiate edge detection using the Canny algorithm to find regions of intensity change within the image, which correspond to the boundaries of the beads. A point cloud derived from these detected edges is then processed using a Hough transform to identify potential lines representing the bead edges. To refine these identifications, a filtering mechanism is applied: lines with deviations exceeding a defined threshold relative to the end-effector velocity vector are rejected, as bead edges are expected to be parallel to this vector. This robust filtering ensures that only desired and accurate lines are accepted. Some authors have developed more advanced, hybrid traditional pipelines to extract further information. For instance, Ajay et al. [52] moved beyond simple 2D measurements by creating a hybrid model that fits a semi-elliptic curve to directly measured bead widths to estimate its height. A particularly innovative pipeline was proposed in [55] for monitoring lattice struts. The authors present a novel real-time visual reconstruction method for strut profiles in pulsed WAAM that mitigates challenges like arc light and spatter interference. The system only processes images where the welding current is below 10A (and the arc extinguishes), as arc light at higher currents obscures the strut profile. The core of the reconstruction method involves two main strategies: melt-pool-based and weld-bead-based profile reconstruction. By combining multiple frames of weld-bead segmentation results with the current melt-pool segmentation, a more accurate and robust strut profile is obtained. Finally, key geometric features, such as inclination angle and diameter, are extracted from these reconstructed profiles using centroid-based calculations and camera calibration. This process integrates segmentation, centroid tracking, and geometric projection into a unified workflow for precise feature extraction. The flowchart of the profile reconstruction method is shown in Fig. 4.

The concept of applying 2D image processing to 3D data is demonstrated in [56], where 3D point cloud data from a laser scanner is first converted into a 2D "depth image." Then, the Hue–Saturation–Value (HSV) colour model is used as a data visualisation technique to represent the height of a 3D surface. Height values in the depth image were then mapped to the HSV colour model and converted back to Red–Green–Blue (RGB), producing a colour image in which changes in height appear as shifts from blue to red. This visualisation makes subtle variations in surface topography easier to interpret. Once an image is captured, extracting meaningful data requires a structured pipeline. In traditional, non-AI approaches, this is achieved through a sequence of conventional image processing algorithms, which rely on explicit mathematical rules and operations to convert pixel data into quantitative measurements. These pipelines typically involve three

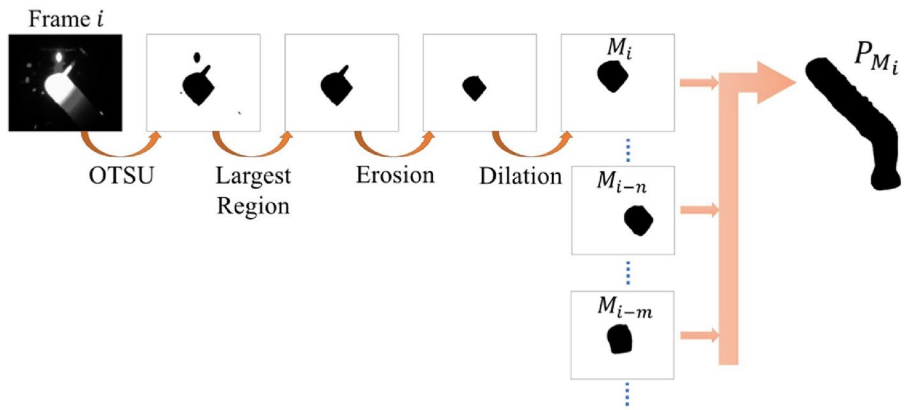

Fig. 4 Flowchart of the melt pool–based profile reconstruction process. Reproduced from [55], licensed under CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/)

stages: pre-processing, segmentation, and feature extraction, which have grown more sophisticated over time. While strong, traditional, rule-based pipelines have notable limitations, these limitations have driven the adoption of machine learning. These pipelines are often brittle and highly sensitive to noise, as their performance depends heavily on manually tuned parameters (e.g., filter sizes, threshold values). These algorithms detect intensity discontinuities and identify the edge between two regions characterised by different properties (e.g., pixel intensities) as the limit. Therefore, because these techniques are based on a locally derived analysis, they are not very effective in the presence of weak object boundaries or noisy patterns [57]. The wire is a thin, highly dynamic object that constantly moves throughout the process. Its visual appearance in the captured images is strongly affected by specular reflections originating from the welding arc and the molten pool. These reflections vary unpredictably from one frame to another. This makes robust identification of the wire tip using classical algorithms, such as edge detection, extremely challenging, as illustrated in Fig. 5.

A pipeline tuned for one condition may fail if the material, process parameters, or lighting change. Traditional algorithms often struggle to capture complex features because they are typically limited to measuring predefined, simple geometric primitives, such as lines and widths. They are ill-equipped to quantify complex, holistic surface defects that truly define part quality.

Another limitation is the lack of a deeper contextual understanding. These systems operate on explicit instructions and cannot learn from experience or infer hidden correlations. They can measure a parameter, but they cannot classify the process state. This is the crucial gap that AI-based methods address. For example, the work by [37] uses deep learning not to measure width, but to analyse the entire melt pool texture to classify the process as "humping" or "normal"; a level of abstraction beyond the scope of traditional algorithms. Many conventional methods address only one part of a larger goal: process control. Processing each frame can introduce delays, making it difficult for them to provide the fast, reliable feedback needed in WAAM. In fact, the need for rapid feedback loops in manufacturing processes can overwhelm conventional image-processing techniques, resulting in delays in defect detection and quality control. These limitations (the narrow geometric focus and the lack of data-driven learning) motivate the development of more robust and adaptive methods. This has strongly motivated the

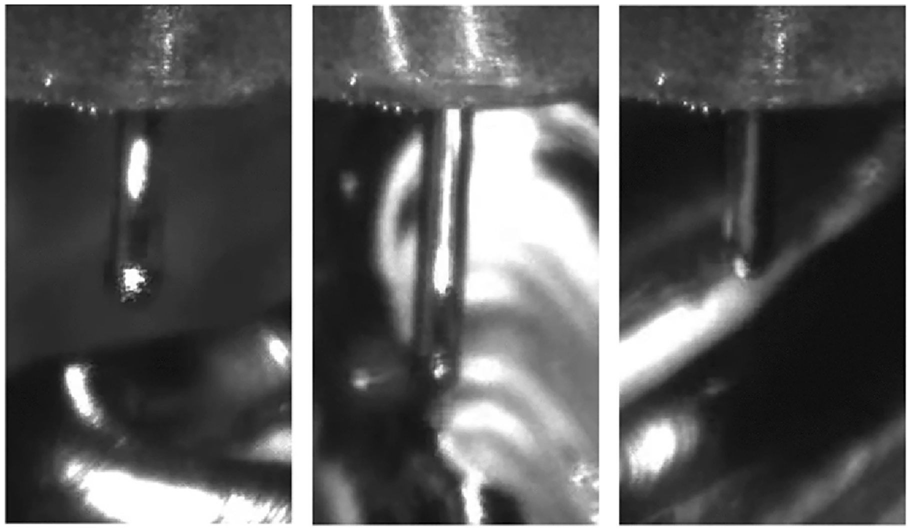

Fig. 5 Processing zone of a test-print at three different times during production. Motion and continuously changing lighting conditions cause unpredictable specular reflections from metal surfaces. It is not easy, even for the human eye, to see the stickout. Reproduced from [16], licensed under CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/)

research community to explore data-driven approaches and deep learning, which will be discussed in the subsequent sections of this review.

3.2 Thermal-based sensing

A thermal sensor detects and measures changes in physical properties or processes by monitoring temperature or heat flow. It works by observing how materials respond to heat, such as changes in expansion, electrical resistance, or the emission of thermal radiation. Thermal sensing is widely used in the industry to monitor processes and ensure product quality through real-time temperature tracking. It helps detect overheating, uneven heat distribution, and abnormal thermal patterns, preventing equipment failure and improving efficiency.

Thermal-based sensing plays a particularly vital role in WAAM. This importance arises because the WAAM process involves rapid, repeated heating and cooling cycles, during which the wire feedstock continuously melts and solidifies to form each layer. These transient thermal cycles significantly impact the shape accuracy, microstructural evolution, mechanical strength, and the formation of defects, such as cracks or porosity, in the final component. To achieve a stable, high-quality build, it is essential to measure, predict, and control the part's thermal behaviour in real-time. This enables both operators and automated systems to dynamically adjust process parameters, maintaining optimal conditions throughout the deposition process. This section describes the hardware and experimental setups used for thermal measurements, identifies the measurable parameters, describes classical analysis methods for data interpretation, and highlights the remaining challenges for future research.

3.2.1 Hardware and setup

Researchers utilise a variety of contact and non-contact sensors to measure these parameters. These sensors are often used in combination to validate results and provide a deeper understanding of the state of thermal systems. IR cameras are the most powerful tool for spatial thermal analysis, providing a real-time 2D map of the surface temperature. They are essential for visualising heat distribution, identifying hot spots, and monitoring the geometry of the melt pool and the deposited layers. Yang et al. [11] utilised an IR camera to effectively study heat accumulation and the impact of inter-layer cooling time and inconsistent bead width in thin-wall parts. Figure 6 presents the schematic layout of the Gas Metal Arc Welding (GMAW)-based additive manufacturing (AM) system, where an IR camera is employed to monitor thermal behaviour during deposition. Also, Fig. 6b illustrates the representative temperature field of the thin-wall part as measured by the IR camera.

To achieve absolute temperature accuracy, which is a significant challenge, a rigorous calibration methodology is required. Baier et al. [58] presented a comprehensive method to calibrate an IR camera for WAAM by determining the in-situ emissivity and transmittance. The authors calibrate the IR camera by first selecting the material emissivity and window transmittance in ambient conditions, based on comparisons between thermographic and thermocouple readings. Then, they repeat the process during WAAM to account for welding fumes and gases, adjusting the values to ensure accurate and

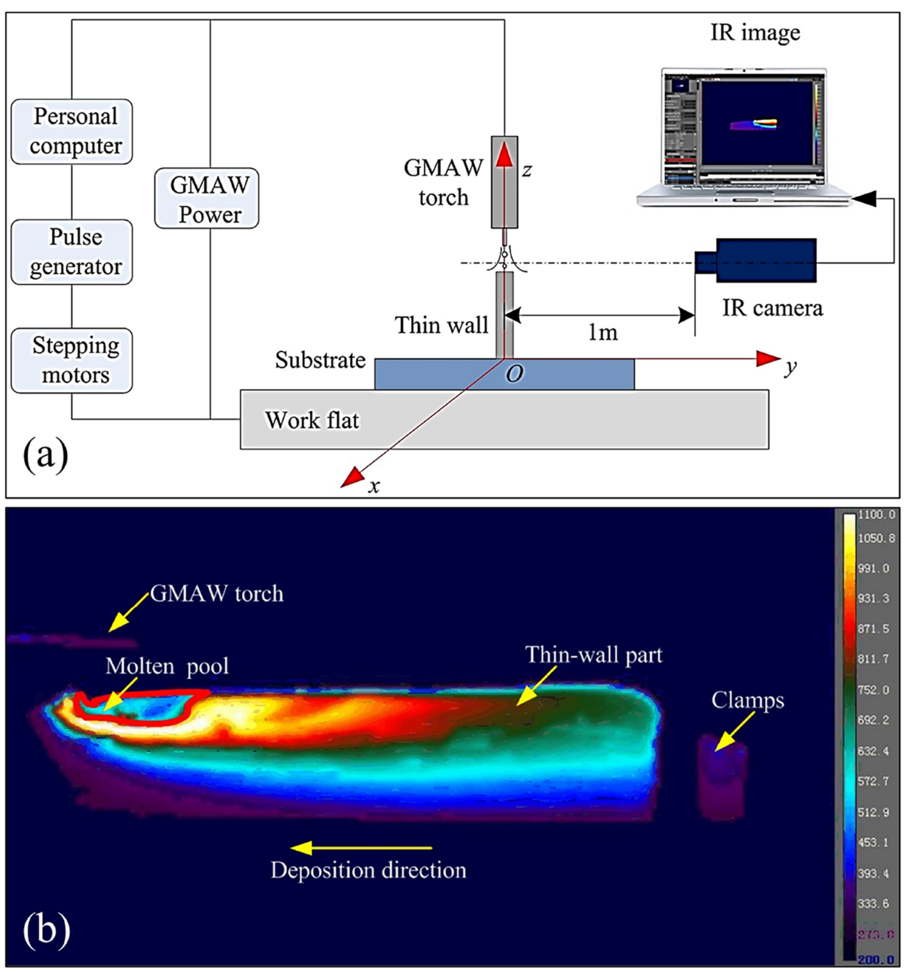

Fig. 6 **a** Schematic diagram of the GMAW-based AM system with the IR camera. **b** The typical temperature field distribution during the deposition process of a thin-wall part, captured by an IR camera (temperature range 200–2000 ° C). Reproduced from [11], with permission granted to Springer Nature Limited to publish under CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/)

reproducible temperature measurements. The valid range of the thermographic signal varies with temperature, and at 200 °C, it was determined to lie between 140 °C and 336 °C. Pyrometers use an optical system to collect and focus this radiation onto a detector, which converts it into an electrical signal. This signal is then processed and converted into a temperature reading based on physical laws, such as Planck's Law. Pyrometers are often easier to implement and calibrate than full-field cameras and are ideal for process control loops that require a single, reliable feedback signal. In a key experimental study, Wu et al. [59] used an IR pyrometer aimed directly at the deposition layer, demonstrating that this in-situ measurement was far more accurate and responsive for tracking the part's thermal state than thermocouples placed on the substrate. As contact-based sensors, thermocouples provide the "ground truth" for thermal measurements. They offer direct, localised temperature readings. During WAAM of AlSi5, substrate temperatures were monitored using thermocouples. Although these measurements underestimate the actual interlayer temperature due to aluminium's high thermal conductivity, they reliably indicate the thermal trend. The substrate temperature increased progressively from 22 $^\circ$ C at the start to 60 $^\circ$ C after the first layer, 160 $^\circ$ C after the fifth layer, and 240 $^\circ$ C after the tenth layer [60]. Thermocouples are also the primary tool for validating the accuracy of thermal-based sensors (IR cameras, pyrometers) and numerical simulations. Nearly all simulation-focused studies, such as those in Srivastava et al. [61], Hosseini et al. [62], and Yildiz et al. [63], rely on thermocouple data for model validation.

3.2.2 Measurable parameters

A WAAM component's thermal state is described by several key parameters, each providing unique insights into the process's health and the quality of the final product. The Inter-layer Temperature (ILT), which is defined as the surface temperature of the previously deposited layer immediately before the start of the next pass, is the most critical measurable parameter for thermal management. In addition, it serves as a direct indicator of heat accumulation. According to [59], the temperature variation between the substrate and the in-situ deposited layers is significantly different due to the accumulation of heat. This heat buildup reduces cooling rates, leading to fluctuations in arc length and unstable metal transfer, which in turn degrade build quality. Baier et al. [58] showed that ILT, dwell time, and geometry are directly related. Uncontrolled, rising ILT can lead to a wider, shorter bead geometry, causing dimensional errors. While ILT provides a snapshot before each layer, the dynamic thermal gradients and cooling rates during and immediately after deposition govern the metallurgical transformations that occur. These parameters are the primary drivers of microstructure formation (e.g., grain size and morphology) and the magnitude of residual stresses. Studies like [63] used simulations to demonstrate that thermal gradients naturally decrease as the build height increases, thereby moving the heat source farther from the primary heat sink (the substrate), underscoring the time-variant nature of the process. It leads to localised heat accumulation in the upper layers, which was discussed in the ILT parameter. The melt pool is where the fundamental physics of deposition occurs. Its temperature and geometry are directly linked to energy input, fusion quality, and defect formation. While direct temperature measurement is challenging due to the arc, its characteristics can be inferred. The authors of [64] reported that the size of the "thermal light region", defined as the high-brightness zone within the weld pool caused by heat radiation and arc reflection, provides an indicator of the instantaneous melt volume and electrode consumption. They monitored this region using a camera and processed the images using dedicated image analysis algorithms. However, because the thermal light region is highly dynamic, exhibiting frame-to-frame variations in both size and geometry, continuous monitoring is necessary to accurately assess momentary welding stability.

3.2.3 Classical thermal analysis

The data acquired from thermal hardware is analysed using a combination of direct experimental methods and predictive numerical models. This synergy allows for both real-time observation and forward-looking process planning. Experimental Thermal Cycle Analysis involves the direct analysis of time-temperature data from sensors. Key metrics are extracted to characterise the process, such as the peak temperature reached during a pass, the interlayer temperature before the next pass, and the cooling rates through critical metallurgical ranges (e.g., Δt 1200-800), as performed by Hosseini et al.[62]. This classical approach directly links process conditions to thermal history. Analytical modelling of the WAAM process often employs three-dimensional transient thermal formulations to predict thermal fields during material deposition. A widely accepted approach is the Goldak double-ellipsoidal heat-source model, which simulates the realistic, non-axisymmetric heat distribution generated by a moving arc [65]. However, conventional single-Gaussian models have also shown limitations in accurately representing the arc's detailed power profile [66]. To improve accuracy, recent studies have focused on developing computationally efficient thermal modelling strategies for WAAM [67]. These advanced analytical heat source models, when integrated into numerical simulations and validated against experimental data, serve as powerful predictive tools for analysing thermal cycles and optimising the WAAM process. The work in [67] introduced a computationally efficient semi-analytical thermal approach that combines an analytical moving Goldak heat source with a numerical boundary correction. It significantly reduces the computational time compared to fully transient finite element analysis (FEA) models in WAAM. The model employs constant thermal properties and neglects phase transitions to maintain linearity, making it well-suited for materials like mild steel, where solid-state transformations have minimal impact on its properties, and was validated against experimental measurements of temperature history and residual stresses, showing good agreement. Similarly, Zani et al. [68] developed a B-spline-based metamodel (or surrogate model) that can rapidly predict thermal responses, making it suitable for optimisation in the WAAM process. Nevertheless, the B-spline entities are not capable of correctly approximating a dataset characterised by strong nonlinearity and/or discontinuities in the first partial derivatives in each direction. Generalising B-splines can extend the methodology. Non-Uniform Rational B-Splines (NURBS) introduce a rational basis by associating a weight with each control point. With this additional degree of freedom, local shape control is enhanced, and functions with strong nonlinearities or discontinuities in their derivatives can be approximated more precisely.

Classic thermal analysis in WAAM provides crucial benefits by establishing a foundational understanding of how thermal cycles influence the final product's quality, including its geometric accuracy and mechanical properties. This is accomplished by monitoring key parameters with sensors and using predictive numerical simulations. However, this traditional approach is fraught with challenges, including high thermal gradients that can cause deformation and cracking, as well as the difficulty of accurately predicting these outcomes using conventional models. The complexity of heat transfer and the need for precise, real-time control often exceed the capabilities of these methods. Machine learning algorithms can analyse vast, complex datasets from in-process monitoring to build robust predictive models that optimise process parameters and detect defects in real-time. This is discussed in Sect. 4.2.

4 Machine learning for uni-modal sensor data analysis in WAAM

4.1 Image-based ML: from native to transformed data

Applying ML to image data advances automated quality monitoring. This section outlines the development of these approaches, beginning with early classification models that interpret raw image pixels to distinguish between different process states. It then examines more advanced architectures capable of accurately detecting and localising defects within a build. Finally, recent progress toward adaptive, material-aware systems is discussed, highlighting their potential to improve model generalisation and predictive reliability. Xia et al. [37] provide a representative example of this shift by investigating the performance of several state-of-the-art CNN architectures, including VGG-16, ResNet, EfficientNet, and GoogleNet, for classifying melt pool images.

To enhance the diversity of the training dataset and improve model generalisation, data augmentation techniques were applied. These included reflection, scaling, rotation, and translation. Their study categorised melt pool states into four classes: humping, spattering, robot suspend, and normal. A total of 19415 melt-pool images were collected under controlled CMT-WAAM conditions, and the high classification accuracies (above 97%) were obtained using a random 70/30 hold-out split rather than cross-validation. This success is partly due to the specific strengths of the networks they tested: ResNet's residual connections allow very deep networks to train effectively, EfficientNet balances depth, width, and resolution to achieve high accuracy efficiently, VGG-16's uniform 16-layer design captures detailed features with a simple structure, and GoogLeNet's Inception modules extract multi-scale information without excessive computation. Despite these capabilities, a limitation of this classification-based approach is that it can detect the presence of an anomaly but does not reveal the exact location of defects within the build or image frame.

The work by Cho et al. [28] on real-time anomaly detection for molybdenum addresses a different, yet critical, practical challenge by focusing on deployment feasibility. Molybdenum is a cost-intensive, difficult-to-machine material, making in-situ quality monitoring particularly crucial to avoid costly scrap. They compare several CNN architectures (MobileNetV2, DenseNet169, etc.) for classifying melt-pool images into "normal" and "abnormal" states (balling and bead-cut). Using 10,685 melt-pool images from a GTAW-based WAAM process with pure molybdenum and an independent hold-out test set (70/30 train–validation split), MobileNetV2 achieved 98% classification accuracy with a processing time of 0.033 s/frame. This demonstrates that selecting an appropriate lightweight architecture enables real-time deployment of deep learning models on accessible hardware. They also compare their classification approach to an object detection model, YOLO (You Only Look Once), and conclude that, for their specific goal of binary state classification, the simpler CNN classifier was both more accurate and faster. Building upon these foundational classification successes, subsequent research sought to address the clear need for defect localisation. To address the challenge of localisation, subsequent research has embraced more advanced object detection frameworks. In another study, authors developed the YOLOv3 architecture to monitor inter-layer deposition quality [35]. Using a small, component-level WAAM image dataset, the authors trained a YOLOv3 detector and selected the best model by validation mAP. The training/validation sets comprised 3 components × 30 layers (51 training images and 17 validation images; 1646 and 602 anomaly boxes, respectively), and evaluation was performed on an independent test set of 3 components × 12 layers (146 images). The parts were fabricated from 4043 Al wire using 100% Ar; CMT was used for the training components, and CMT-pulse for the test components. At confidence 0.4 and NMS 0.05, the best model achieved mAP 0.765, 100% component localisation, and 53% precision for surface-anomaly detection–showing reliable localisation but limited defect identification on this small dataset. The authors of [35] attribute the limitation to the modest dataset and the limited number of defect counts. The issue of data imbalance, a related challenge and a likely contributor to the modest precision observed in localisation, is directly tackled by He et al. [69] in their work on detecting defects in low-carbon steel. In any real-world manufacturing scenario, images of normal, defect-free operation far outnumber images of defects. Traditional CNN models tend to prioritise overall classification accuracy, which can lead to biased predictions towards majority classes and lower accuracy for minority (defect) classes. They propose a novel solution by combining improved remanence/magneto-optical (MO) imaging with a Cost-Sensitive Convolutional Neural Network (CSCNN) trained on 1600 MO images acquired from low-carbon steel CMT-WAAM and evaluated using a 75/25 hold-out split, achieving up to 94.60% classification accuracy. The CSCNN model assigns a higher misclassification cost to minority classes (defects) to improve their recognition rate, thereby focusing more attention on these critical defect types. This is a significant strength, as it enhances detection rates for actual defects without requiring artificial dataset balancing through methods like oversampling, which can sometimes lead to over-fitting.

Finally, the challenge of real-time monitoring and detection advances to a proactive stage by focusing on process indicators rather than the final product. The challenge of real-time monitoring is further explored by Franke et al. [16], who develop a vision-based system for monitoring critical process stability indicators, such as the nozzle-to-work distance and spatter formation. Instead of looking at the final bead, they monitor the state of the welding wire itself using a U-Net for semantic segmentation (3000 labelled frames; CMT–WAAM with AlSi12 wire and Ar shielding, Cavitar C300 at 20 Hz; 20% hold-out validation; U-Net validation accuracy ≈ 99.5%; no independent test set reported). Their spatter algorithm detected 95.5% of spatter with about 10% false positives (evaluated on a consecutive-frame sequence; an auxiliary autoencoder trained on 1000 spatter crops flagged ∼87% of false positives at an MSE threshold of 0.002; 20 Hz at 350 mm/min). This is a proactive approach to quality monitoring; deviations in wire stickout or increases in spatter are often leading indicators of process instability that will later result in defects. Extending this line of research, it is essential to clarify that although previous work primarily relied on image-based features (melt pool, bead, or sensor data), it did not explicitly incorporate material properties into its algorithms. Another investigation focuses on detecting balling defects, a common and problematic anomaly in WAAM [70]. Their work introduces a sophisticated solution to the data scarcity and generalizability problem through a "Material-Adaptive Anomaly Detection" system. Unlike earlier CNN-based approaches that relied solely on process-signal-derived image features, this study explicitly concatenates material properties (e.g., thermal conductivity and melting point) with CNN-extracted image features. The core innovation is a property-concatenated transfer learning approach: recognising that training a model from scratch for every new material is impractical, they pre-train a CNN on a data-rich source material (e.g., low-carbon steel) and then fine-tune it for a data-scarce target material (e.g., Inconel). The novelty lies in concatenating explicit material properties as additional features onto the image features extracted by the CNN, which explicitly informs the model about the material it is analysing, allowing it to adapt its learned patterns more effectively. This represents a powerful fusion of data-driven learning with domain-specific knowledge, addressing the critical weakness of poor model generalizability across different materials. In summary, data collection involves a GTAW-based WAAM system using low-carbon steel, 316 L stainless steel, and Inconel 625 (300 single-bead deposits), and the acquisition of bead-shape observations, numerical voltage time-series, and camera images. These raw data are then preprocessed, converted to image formats, labelled, and fed into CNN-based models, with material properties concatenated to enhance anomaly-detection accuracy, using a random train/test split with an internal 70/30 train–validation split and achieving 81.9–89.4% classification accuracy. The detailed architecture of the proposed method is shown in Fig. 7. This approach allows CNN to extract features and detect balling defects in WAAM processes. This method preserves all original spatial and temporal information while integrating material-specific knowledge, improving both detection performance and generalizability without requiring large datasets for every new material. Extending this material-informed perspective, perhaps the most potent and forward-looking approach to quality monitoring presented in this collection of papers is the use of multisource transfer learning, as detailed by Shin et al. [71]. This work builds upon the previously

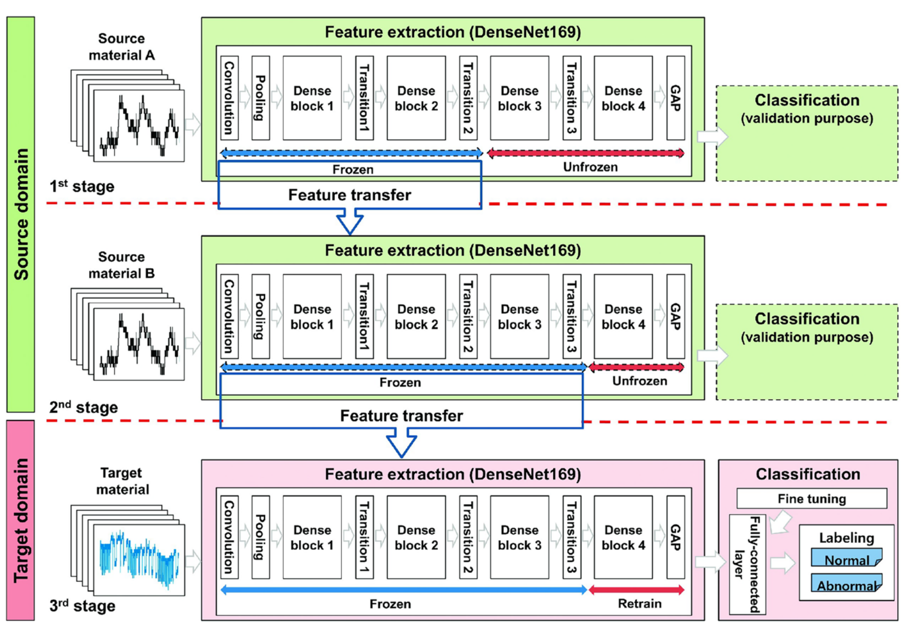

Fig. 7 Proposed architecture of the "Material-Adaptive Anomaly Detection" system for the detection of balling defects. Reproduced from [70], licensed under CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/)

mentioned concept of material-adaptive transfer learning. Instead of transferring knowledge from a single source material, they propose a framework for sequentially learning from multiple source materials (e.g., low-carbon steel, then stainless steel) before fine-tuning the model for a target material (e.g., Inconel). Using a GTAW-based WAAM system based on 300 single-bead deposits with voltage signals converted to voltage-image data. This "stepwise learning" approach has the potential to create a more robust, generalised base model by exposing it to a wider range of melting and solidification behaviours. Their concept of "composite learning," which involves adaptively setting the ratio of frozen (transferred) to unfrozen (newly learned) layers in the CNN, provides a sophisticated mechanism for balancing knowledge retention with adaptation to new data. The network was trained with a 70/30 train–validation split and subsequently evaluated on an independent test set. The resulting accuracies were 82.95%, 89.47%, and 84.22% for the three cross-material transfer cases. This method represents a significant strength and a potential solution to one of WAAM's biggest ML challenges: how to build a single, flexible quality monitoring system that can be efficiently adapted to new materials and processes without requiring a complete restart each time. In summary, the experimental setup utilises specialised equipment for both numerical and visual data collection. The core input to the anomaly detection algorithm is image-type voltage data, derived from raw time-series voltage measurements and validated against bead image labels. This approach also allows CNN to extract features and detect balling defects in WAAM processes. A consolidated overview of the image-based anomaly detection and process monitoring studies discussed in this section is provided in Table 2.

4.1.1 Physics-informed modelling

Rashid et al. [72] introduce a coupled Physics-Informed Neural Network (PINN) framework for bead geometry prediction. This approach represents a notable advancement, as it moves beyond purely data-driven models. A PINN incorporates the governing physical equations (e.g., for heat transfer and fluid dynamics) directly into the neural network's loss function. This integration offers two primary advantages. First, it reduces the amount of experimental data required for training, as known physical laws constrain the model. Second, it ensures that the model's predictions are physically consistent and can generalise better to conditions outside the training data range. The authors present a two-step workflow in which a thermal model first predicts temperature evolution, which then informs a geometry model to predict cross-sectional bead height. They also compare a high-fidelity PINN (computationally intensive but highly accurate) with a low-fidelity version (less accurate but computationally faster), highlighting the trade-off between computational cost and precision, a critical consideration for real-time applications. However, their work focuses on single-layer deposition. Although the same fundamental physics applies to multi-layer deposition, additional complexities arise due to evolving thermal profiles. These complexities also result from changes in substrate geometry, boundary conditions, and initial conditions with each new layer. To address these issues, a group of researchers proposed learning through Fourier Neural Operators (FNO) as a potential solution for future studies [73]. Thanks to the Fast Fourier Transform (FFT), the original FNO can perform computations efficiently even for large input sizes (Fig. 8). However, it requires a regular, fixed, rectangular domain, limiting its applicability to WAAM processes involving irregular, growing part geometries.

Table 2 Summary of studies on anomaly detection and process monitoring[a]

Refs	Title	Data source	Key contribution
[28]	Real-time anomaly detection using CNN in WAAM: Molybdenum material	Melt pool images	Identified MobileNetV2 as a lightweight CNN for fast and accurate real-time defect detection and classification, comparing its performance with other CNNs (DenseNet169, Resnet50V2, InceptionResNetV2, YOLO)
[16]	Vision-based process monitoring in WAAM	Camera images	Enabled closed-loop control to improve part geometry through using U-net architecture and classic image processing for wire and spatter detection and segmentation
[37]	Vision-based melt pool monitoring for WAAM using deep learning method	Melt pool images	Demonstrated that CNNs are effective for classifying melt pool states, comparing multiple CNNs (ResNet, EfficientNet, VGG-16, GoogLeNet)
[35]	Towards intelligent monitoring system in WAAM: a surface anomaly detector on a small dataset	Camera images	Demonstrated feasibility of training a robust defect detector on a small dataset through YOLOv3 with modified anchor settings
[69]	Automatic defects detection and classification of low carbon steel WAAM products using improved remanence/magnetooptical imaging and cost-sensitive CNN	Enhanced Magneto-Optical (MO) images	Combined enhanced NDT imaging with Cost-Sensitive CNN (CSCNN) to detect and classify defects despite class imbalance reliably
[70]	Detecting balling defects using multisource transfer learning in WAAM	Image features concatenated with material properties	Addressed data scarcity for expensive materials by transferring knowledge from cheaper ones through constrained transfer learning
[71]	Material-adaptive anomaly detection using property-concatenated transfer learning in WAAM	Time-series voltage data converted to images	Improved transfer learning reliability using a broader knowledge base from several materials through sequential CNN training

CNN = Convolutional Neural Network; MO = Magneto-Optical; NDT = Non-Destructive Testing; YOLOv3 = You Only Look Once (third version). Performance metrics across studies are not directly comparable. Where available, the body of this article reports the dataset size, material/process conditions, and the evaluation protocol (cross-validation or a single hold-out split)

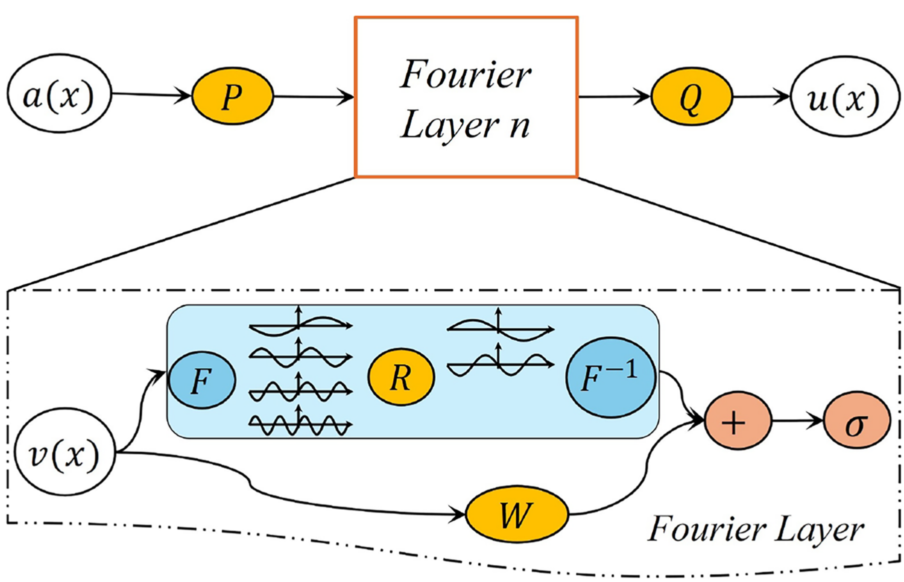

Fig. 8 Architecture of Fourier-based FNO method, performing operator learning on fixed regular domains

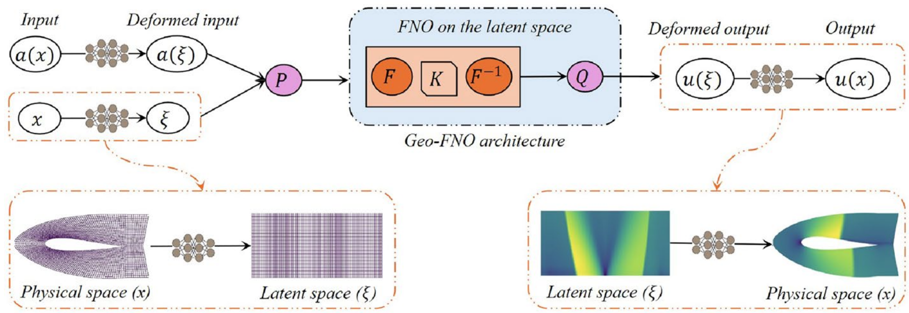

Fig. 9 Architecture of Geo-FNO learning method, capable of handling irregular geometries via learned domain deformation. Adapted from [74], licensed under CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/)

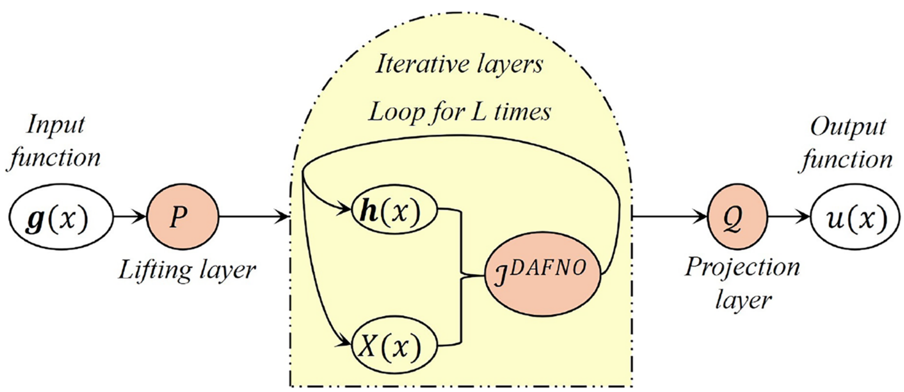

Fig. 10 Architecture of DAFNO learning method, a domain-agnostic approach for irregular and evolving topologies

To improve geometric flexibility, Geo-FNO was introduced [74]. This method utilises an additional learned network to deform the irregular physical domain into a uniform latent grid, enabling the efficient application of the FFT (Fig. 9).

While Geo-FNO retains computational efficiency and can handle fixed irregular shapes, it is more likely to face challenges when applied to scenarios with drastic topology changes, such as the addition of new layers typical in additive manufacturing. The Domain-Agnostic Fourier Neural Operators (DAFNO) explicitly addresses both irregular geometries and evolving topologies [75]. DAFNO incorporates a smoothed characteristic function into the integral layer architecture of the FNO (Fig. 10). This function acts as a binary mask, with a value of 1 inside the material domain and 0 outside. Multiplying the integral kernel by this function ensures that the FNO layer update considers only interactions between points within the material domain, effectively eliminating non-physical influences from outside the part. Although DAFNO's architecture is more complex than the original FNO and may introduce minor approximations at the boundaries, its key advantage lies in its generalizability to unseen geometries and ability to model topology changes. This capability is crucial for accurately simulating the layer-by-layer material deposition and thermal behaviour in WAAM processes. As illustrated in Figs. 8, 9 and 10, the FNO, Geometry-aware FNO, and DAFNO architectures highlight the progression from standard operator learning to geometry-informed and dynamically adaptive formulations.

4.2 Thermal-based ML: from simulated fields to real-world insight

The thermal history of a component during Wire Arc Additive Manufacturing (WAAM) is the fundamental driver of its final microstructural and mechanical properties. While thermal sensing provides a direct window into this history, classical analysis, often limited to simple metrics like peak temperature or cooling rates at discrete points, fails to capture the intricate spatiotemporal dynamics that define the process. Machine Learning (ML) offers a transformative approach, enabling the extraction of deep, actionable insights from the high-dimensional data streams generated by thermal sensors and simulations. This section reviews key studies that leverage ML to analyse unimodal thermal data, following a logical progression from the development of high-fidelity digital models to their application in real-world, in-situ process monitoring. A critical challenge in leveraging simulations for process understanding is bridging the 'sim-to-real' gap to ensure that digital models are physically veracious. Addressing this precise challenge, Hasan et al. [76] present a complementary machine learning framework that focuses not on interpreting thermal data but on calibrating the simulation itself to enhance its physical fidelity. They are attempting to determine the optimal input parameters for a Finite Element (FE) thermal model, specifically thermal efficiency (η), convection coefficient (h), and emissivity (ϵ), to ensure its outputs correspond to actual thermocouple measurements. The entire calibration process is navigated by Bayesian optimisation. This is a probabilistic search technique designed to find the optimal solution with a minimal number of computationally expensive simulations. To make this intelligent search computationally feasible, the authors employ a multi-fidelity strategy, using two distinct models for different stages of the search. A fast Response Surface Model (RSM) is used to quickly map the general relationship between input parameters and thermal outputs during the initial, broad exploration stage. Once promising domains are identified, a high-fidelity Extra Trees Regressor is used for precise evaluation. By applying this hierarchical framework to sparse time-series thermal data, the authors produce a calibrated simulation that exhibits strong experimental concordance, achieving a mean absolute error of only 7.47 °C. This work is significant as it demonstrates a rigorous pathway for creating a calibrated 'digital twin,' providing a solid foundation for generating the physically veracious thermal datasets requisite for subsequent machine learning endeavours.

Once a reliable simulation framework is established, it can be used to uncover the fundamental relationship between the transient thermal patterns of the process and the resultant material microstructure. Fernandez-Zelaia et al. [77] pioneered a framework that interprets the entire spatio-temporal thermal evolution as a "thermal signature". To overcome the prohibitive cost of physical experiments, they generated extensive spatio-temporal thermal data using a computationally efficient, reduced-order physics-based model, which rapidly simulates the thermal evolution history "video" for various printing scan patterns. To implement the study's AI framework, the team utilised Video Vision Transformers (ViViT), a model specifically designed to process spatiotemporal data by treating the thermal signal as a cohesive video sequence. Crucially, the model was trained using a self-supervised learning approach, which avoids the need for a large, pre-labelled dataset. Through this process, the architecture learned to distil the complex, high-dimensional thermal signals into a compact 'signature' that effectively represents the key dynamics of the manufacturing process. Finally, this learned signature was used to directly link to the actual microstructural properties (such as grain size and orientation) that were precisely measured from the physical parts using advanced microscopy. While groundbreaking, this approach also highlights a critical 'sim-to-real' gap, as the model's accuracy in real-world applications depends heavily on how well the reduced-order simulation captures the process's actual physics.

Extending the concept of monitoring from mere observation to predictive process control, some studies leverage validated simulations to explore profound questions of process robustness. The work of Pham and Tran [78] provides a critical proof of concept for how deep learning can bridge this gap. Their study's primary innovation is not the development of a novel neural network, but the strategic application of a well-established architecture, the Feed-forward Neural Network (FFNN), as a computationally lightweight surrogate to enable a full-scale Uncertainty Quantification (UQ) analysis. The UQ statistical approach systematically investigates how minor, real-world fluctuations in WAAM parameters, like current and speed, propagate through the system. Its goal is to precisely quantify the resulting range of variability in critical thermal outputs, such as the final cooling rate. The FFNN, validated against experimental data, successfully emulated thermal physics within the tested input range, enabling a previously computationally prohibitive Monte Carlo simulation. This probabilistic approach yielded a key insight: that inherent, minor fluctuations in current intensity have a more pronounced impact on thermal variability than similar fluctuations in travel velocity. However, it is worth noting that this high-fidelity performance is limited to the specific input parameters selected. The temperature at each point in the WAAM process is inherently influenced by time and previous layers due to heat accumulation during multi-layer deposition. As a point-wise regression model, the FFNN excels at mapping a vector of inputs to a temperature output, but does not inherently capture the rich spatio-temporal context of the thermal field in the same way a more complex architecture, such as a Vision Transformer (ViT) or a Long Short-Term Memory (LSTM) network, might.

While simulation-based approaches provide profound insights, the ultimate goal of monitoring is to detect defects in real-time during a physical build. This transition to in-situ sensor data introduces new challenges, including process noise, physical interference, and the critical bottleneck of data scarcity. Era et al. [79] address this last challenge head-on by deploying a state-of-the-art, transformer-based framework to detect melt pool anomalies linked to porosity. Their approach relies on a two-stage, self-supervised learning strategy that utilises a ViT-based Masked Auto-encoder (MAE). The choice of a ViT is strategically significant; unlike traditional CNNs that focus on local features, the ViT's self-attention mechanism allows it to capture the global spatial context of the melt pool, making it theoretically more adept at identifying subtle anomalies that are only apparent when considering the entire melt pool's shape and temperature distribution. The framework's primary strength is its elegant solution to the data scarcity problem. A masked autoencoder fine-tuned on 7812 unlabelled thermal images enabled data-efficient melt-pool classification on 1447 labelled Inconel 718 samples. Under a 15% hold-out test split with six-fold cross-validation, the method achieved accuracies of 95.44–99.17%. However, this cutting-edge approach comes with an inherent trade-off. Transformer-based models are computationally intensive, and while their performance is high, their "black box" nature can make the learned features difficult to interpret from a metallurgical perspective. Furthermore, by tackling the melt pool directly, the model must still contend with the inherent optical noise and thermal variability of the most chaotic region of the process.

In direct contrast to the head-on approach of Era et al. [79], the work of Yu et al. [80] exemplifies a pragmatic engineering compromise aimed at achieving maximum robustness in a noisy industrial environment. Instead of tackling the optically challenging melt pool, their innovation was to avoid it entirely and use the thermal signature on the recently solidified sidewall to detect cladding-layer offsets. A deep residual network classified ROI temperature-field images from CMT–WAAM with ER316L wire on a 304 substrate. Using a 5-class dataset of 3750 IR sidewall images (750 per class) acquired with a MAG32 camera at 50 Hz, the model was trained with a per-class hold-out split (500 train / 250 test per class) and achieved 99.84% accuracy on the independent test set (ROI sensitivity: left/middle/right = 100%/99.92%/99.92%). In additional generalisation trials with two new offset schemes at the arc extinguishing point (±2 mm), test accuracy reached 100%. This design is practical and robust–by choosing a stable sidewall ROI, it sidesteps arc interference–but it is inherently a lagging indicator: it verifies the just-deposited layer rather than correcting during formation, making it excellent for layer-wise QA/stop decisions yet less suited to immediate intra-layer feedback control.

In conclusion, the reviewed studies demonstrate that machine learning can effectively extract rich information from unimodal thermal data. This enables both predictive process modelling and in-situ defect detection. However, these approaches also highlight the inherent limitations of a single analytical mode. This raises a critical question: Could a more comprehensive and robust monitoring system be achieved by synergistically combining this deep thermal insight with a complementary sensory modality? The following section directly addresses this question, examining the frontier of vision-thermal data fusion, where the goal is to create an integrated perception system that is fundamentally more powerful than the sum of its parts. The thermal- and process-oriented ML studies reviewed in this subsection are summarised in Table 3.

4.3 At-a-glance model selection for uni-modal monitoring

Image- and thermal-based monitoring in WAAM usually falls into three task families: (1) classification of process states such as normal, humping, or spatter, (2) object detection of localised surface anomalies, and (3) semantic segmentation of melt-pool, wire, arc, or recently solidified geometry. In practice, model choice is governed by four constraints: latency budget, annotation cost, data volume and class imbalance, and where the signal is measured (melt pool, sidewall, or just-solidified bead). For vision, mid-depth CNNs such as ResNet-18 and EfficientNet-b0 are strong defaults for state classification with small or medium datasets; MobileNet-V2 is often preferred for on-edge real-time inference when compute is tight. When localising anomalies matters, one-stage detectors such as YOLOv3 provide spatial interpretability at modest annotation cost, while U-Net variants remain the workhorse for pixel-level geometry or wire segmentation in stability monitoring and CTWD/NtW proxies. For thermal sensing, models benefit from global context: transformer-based approaches such as ViT and MAE-ViT have achieved high accuracy in melt-pool characterisation and the detection of porosity-linked anomalies, often with self-supervised pretraining to reduce labelling effort. When the region of interest is a stable sidewall rather than the arc or melt-pool, lightweight CNNs on IR frames can classify layer offsets with high accuracy and robustness for layer-wise quality

Table 3 Recent ML-driven advances for thermal/process data: calibration, microstructure prediction, uncertainty analysis, and defect detection[a]

Refs	Title	Data source	Key contribution
[76]	Precision calibration in wire-arc-directed energy deposition simulations using a machine-learning-based multi-fidelity model	Measured and simulated thermal data	Surrogate-modelling framework that accelerates calibration of physics-based simulations through Bayesian optimisation
[77]	Self-supervised learning of spatiotemporal thermal signatures in additive manufacturing using reduced-order physics models and transformers	Simulated thermal videos and experimental EBSD data	Exploitation of triplet loss ViViT for automating discovery of complex process–structure relationships, eliminating manual feature engineering
[78]	Sensitivity study of process parameters of WAAM using probabilistic deep learning and uncertainty quantification	FE simulation data	Integration of deep learning surrogate into an FFNN probabilistic framework for efficient uncertainty quantification
[79]	In-situ melt pool characterisation via thermal imaging for defect detection in directed energy deposition using vision transformers	Labelled and unlabelled thermal images	Drastic reduction of manual labelling by effective self-supervised learning from unlabelled data and transferring learning
[80]	Identification of cladding layer offset using infrared temperature measurement and deep learning for WAAM	IR images of the weldment sidewall	End-to-end supervised ResNet-based CNN for direct classification of cladding-layer offset from thermal images

CNN = Convolutional Neural Network; DL = Deep Learning; EBSD = Electron Backscatter Diffraction; FE = Finite Element; FFNN = Feed-Forward Neural Network; IR = Infrared; MAE = Masked Autoencoder; MO = Magneto-Optical; ViT = Vision Transformer; ViViT = Video Vision Transformer. Performance metrics across studies are not directly comparable. Where available, the body of this article reports the dataset size, material/process conditions, and the evaluation protocol (cross-validation or a single hold-out split)

assurance, albeit as a lagging indicator. Across both modalities, cost-sensitive losses or thresholding help improve recall for rare defects, and transfer learning with augmentation is important when labelled datasets are small. Table 4 summarises deployment-oriented trade-offs and selection cues, while Sect. 5 explains how these uni-modal insights combine in vision–thermal fusion.

5 Machine learning for vision-thermal data fusion

As discussed in the previous chapter, unimodal sensing approaches that use vision or thermal data have demonstrated considerable success in monitoring specific aspects of the WAAM process. With these single-modality systems, it is possible to build calibrated digital twins, quantify process uncertainty, and detect specific defects in real-time. Thermal sensors, however, possess a unique perspective that makes them inherently limited; they can detect a process anomaly without the geometric context to diagnose its cause, whereas visual sensors may observe the shape of beads but remain unaware of the underlying thermal conditions that contribute to their formation. To overcome these limitations and create a more comprehensive and robust monitoring system, researchers are increasingly utilising multi-modal data fusion, specifically the combination of visual and thermal data. In this chapter, the fundamental rationale for fusion is presented, followed by a description of the deep learning architectures used to implement it, and concluding with a critical analysis of relevant studies.

5.1 The rationale for data fusion

The development of robust monitoring systems for WAAM is fundamentally an effort to gain a more complete understanding of the process. Researchers aim to uncover how different aspects of the process interact and influence the final build quality. Single-sensor systems have provided valuable insights. However, their perspective remains

Table 4 Uni-modal model selection for WAAM monitoring. Task families, representative models discussed in Sect. 4, strengths/limitations, and deployment cues[a]

Modality	Task family	Models	Strengths	Limitations and When to prefer
Vision	State classification	ResNet-18; EfficientNet-b0	Good accuracy vs. latency; robust transfer learning; modest data needs	No localisation; may miss tiny/brief events. Prefer for small/medium datasets and fast pass/fail decisions in process-camera views
Vision	On-edge classification	MobileNet-V2	Very low compute/ memory; real-time on CPU/embedded PCs	Lower ceiling accuracy than larger CNNs. Prefer under strict latency/ compute budgets (industrial PCs, edge)
Vision	Object detection (surface anomalies)	YOLOv3 / YOLO-tiny	Localises defects; interpretable bounding boxes; millisecond-level inference	Needs box annotations; precision can drop for tiny/rare targets. Prefer when you must say *where* the defect is
Vision	Segmentation (melt-pool / wire / geometry)	U-Net family	Pixel-level outputs; geometry-aware control hooks (e.g., stickout, NtW proxies)	Mask annotation cost; higher latency than classifiers. Prefer for geometry/stability monitoring and control inputs
Thermal (melt-pool ROI)	Anomaly classification with global context	ViT / MAE-ViT	Captures global thermal shape; self-supervised pretraining reduces label burden	Heavier compute; careful regularisation needed. Prefer for melt-pool characterisation and porosity-linked anomalies
Thermal (sidewall ROI)	Layer-offset classification (QA)	Mid-depth CNN (e.g., ResNet)	Very high accuracy on stable ROIs; robust to arc glare; simple labelling	Lagging indicator (post-track). Prefer for layer-wise QA/stop decisions with IR sidewall imaging
Either	Imbalance handling (rare defects)	Cost-sensitive losses / thresholding	Improves minority-class recall without oversampling	Requires tuning; possible FP increase. Prefer in heavily imbalanced datasets typical of production

[a]Examples of these families and WAAM-specific results are discussed in the body of this article (e.g., MobileNet-V2 for real-time anomaly classification; YOLOv3 for small-dataset detection; U-Net for wire/spatter segmentation; ViT/MAE for thermal melt-pool analysis; IR-sidewall CNNs for offset detection). This table distils those choices into deployment-oriented guidance and prepares the ground for Sect. 5 on multimodal fusion

inherently narrow and cannot fully capture the complexity of this multi-physics phenomenon. Each sensor type observes only part of the process, leaving other vital factors unseen. For example, a vision camera can accurately track the shape and surface geometry of the build, yet it cannot detect the thermal fluctuations that drive those geometric changes. In other words, it can see the visible outcome but not the invisible causes that led to it. This limitation shows why relying on a single sensing approach often produces an incomplete picture of the process.

To address this, multimodal data fusion appears as a promising solution. By combining multiple sensors, each with its own strengths, one sensor's limitations can be offset by another's capabilities. The goal is not simply to collect more data, but to integrate information from different physical domains in a manner that enables algorithms to utilise it effectively. When data from multiple sources are combined effectively, they provide a more comprehensive and reliable understanding of the WAAM process. This goes beyond what any single sensor could achieve on its own.

The argument for fusing sensor data is based on the interconnected nature of the additive manufacturing process. Chabot et al. [81] describe the geometry and thermal history of a deposited part as "entangled phenomena affecting the part's integrity." In other words, the geometric deviations of a part cannot be fully understood without considering its thermal history, and vice versa. Also, in another study, Shamsaei et al. [82] argue that the interaction between thermal cycles (e.g., heat accumulation, cooling steps) and geometrical features (e.g., layer height, bead width) directly influences critical aspects like residual stresses, thermal distortions, microstructure, and overall mechanical properties of the final part. This implies that attempts to control one aspect while ignoring the other are inherently limited. Focusing on a single factor provides only a partial view of the process. This conceptual framework is strongly supported by experimental evidence. While improvements in absolute accuracy from data fusion might appear modest at first glance, a closer examination reveals a significant increase in overall model reliability. For example, in a study on Laser Directed Energy Deposition (L-DED), Chen et al. [17] reported a confusion matrix to evaluate classification performance. Using 292,200 spatiotemporally registered samples collected during robotic laser DED of maraging steel C300 and an 80/20 stratified train–test split with 10-fold cross-validation, the multisensor fusion model achieved a false alarm rate of 4.4%, outperforming vision (11%), acoustic (8.4%), and thermal (8.4%) approaches. The model also achieved an overall accuracy of 95.6% and an ROC–AUC of approximately 0.99. Similarly, Shen et al. [18] reported that their multimodal WAAM model achieved a competitive prediction error of less than 0.06 mm (2400 paired samples; CMT–WAAM on 304 SS/high-carbon steel wires, 304 SS baseplate; 70/10/20 hold-out; test MAE $\approx$ 0.051 mm, per-layer 0.041–0.062 mm). In comparison, the single-modal model's error exceeded 0.08 mm. However, this absolute difference seems small. In high-value manufacturing, where the cost of a single failure is substantial, such reductions are not merely incremental improvements: they are critical for achieving the reliability required for industrial adoption.

Beyond improving the accuracy of individual predictions, data fusion is crucial for developing a more robust and comprehensive diagnostic system. It leverages the complementary strengths of different sensors. Different sensor types are needed to measure various physical phenomena. For instance, a vision system cannot detect chemical anomalies, just as a spectral sensor cannot track process kinematics. No single sensor can provide a complete picture of a complex manufacturing process. The innovation of multimodal monitoring lies not in recognising this principle, but in designing systems that exploit it effectively. The "cooperative awareness" system developed by Zhao et al. [19] illustrates this approach. In their system, a vision-based model monitored welding speed, while spectral analysis identified chemical anomalies, shielding gas instability, and surface contaminants. By combining both, the system could detect a broader range of process faults. This multi-sensor fusion approach effectively compensated for individual sensor blind spots, enabling more reliable detection of diverse process faults compared to single-sensor methods. Furthermore, multi-sensor fusion can unlock entirely new monitoring capabilities. The work by Hauser et al. on multi-material WAAM provides a compelling example [20]. A combined monitoring method integrated spectral analysis, a welding camera (vision), electrical current and acoustic-emission signals during CMT-WAAM of alternating Al 6060 and Al 5087 layers, enabling real-time identification of the deposited alloy. During the processing of alloy 5087 in CMT mode, welding camera images revealed characteristic 'white lines' caused by magnesium soot, providing a visual cue for alloy identification. Spectral analysis further distinguished alloy 5087 from alloy 6060 by detecting variations in green, blue, and red emissions, which were correlated with the magnesium content. Additionally, shifts in electrical and acoustic signals, linked to differences in thermal conductivity and process behaviour, offered complementary confirmation of the alloy type. This enables in-process material verification, which is not possible with a single modality. At its core, the rationale for sensor fusion can be summarised in three key points. First, it enables higher accuracy by capturing interactions across modalities. Second, it improves robustness by compensating for the limitations of individual sensors. Third, it creates new capabilities by providing a more comprehensive understanding of the process state.

5.2 Machine learning architectures for multimodal fusion

There are three main strategies for fusing multimodal data: early (input-level), mid-level (feature-level), and late (decision-level) fusion. In input-level fusion, raw data from different modalities are combined at the start to form a single input. Feature-level fusion, by contrast, extracts features from each modality separately and then merges them for further analysis. Finally, decision-level fusion combines the outputs or decisions from individual modalities to reach an overall conclusion. Fusion choice balances feature richness, model complexity, and robustness. The following section reviews these fusion strategies, using examples from WAAM and related Directed Energy Deposition (DED) processes to illustrate their implementation and main benefits and limitations. In early fusion, raw vision/thermal streams are stacked, which assumes frame-level temporal synchrony and pixel-accurate spatial registration; mismatches in frame rate or delay immediately inject noise at the pixel level. Mid-level fusion tolerates moderate misalignment because each encoder operates at its native sampling, and the model fuses time-aligned feature batches rather than raw pixels. Late fusion only requires that decision timestamps are comparable, and can aggregate heterogeneous sensors with different frame rates. In practice, the recent literature recommends (1) trigger-based alignment where available (e.g., using a shared hardware trigger/clock and pulsed illumination gated to the camera and calibrating a fixed spatial map once) [46–48, 83] and (2) learned alignment when hardware locking is infeasible (e.g., frame-interpolation to up-sample the slower thermal stream to the vision rate, or sequence models that learn cross-modal lags, with quantitative checks against visible fiducials or planned trajectory markers) [17, 18, 83, 84].

5.2.1 Early fusion (input-level)

In early fusion, also known as data-level fusion, multiple sensor types are combined at the earliest stage of input. Before sending data to a machine learning model, raw or lightly processed data is combined into a single tensor or high-dimensional feature set. For example, in vision–thermal applications, a three-channel visual (RGB) image can be stacked with a single-channel thermal image to create a four-channel (RGB-T) input, which a CNN then analyses. This model learns low-level patterns and relationships between the different data types directly.

However, this approach is extremely sensitive to the spatio-temporal registration challenges detailed previously. While all fusion methods require good registration, early fusion is particularly sensitive to registration errors. If the model is misaligned in space or time, it is challenging for it to learn meaningful patterns. It is a result of noise and contradictory information introduced directly at the pixel level.

One study provides a practical implementation solution in laser welding [84]. They addressed the common issue of asynchronous sensors (a fast vision camera and a slower thermal camera) not by discarding data, but by using a Melt Pool Thermal Video Frame Interpolation (MTVFI) model to up-sample the thermal video to match the vision camera's frame rate. It is a type of VFI model, which increases the temporal resolution (frame rate) of video data by generating new frames between captured frames. This ensures that the dynamic changes in the melt pool are captured effectively, balancing the frame rates and supplementing the infrared video's dynamic information. The important thing about the VFI is that, while it can reveal transient thermal behaviours, its accuracy depends heavily on model training and noise sensitivity. In addition, it risks generating misleading artefacts that distort fundamental melt pool dynamics if not validated against physical data. Once synchronised, a Swin Transformer–based model (Swin Fusion) was used to fuse each pair of frames at the pixel level into a single enhanced image for process analysis. The innovation of Swin Fusion lies in its hierarchical attention framework for combining multiple image sources. Rather than merging features at a single level, it combines low-level textures with high-level semantic details. With window-based attention mechanisms, important regions are adaptively focused on during fusion, improving clarity and detail preservation. Compared to traditional CNN-based or pixel-level fusion methods, this design achieves balanced integration of global and local information. While this method is robust, its real-time implementation in resource-constrained industrial environments remains challenging due to the high computational costs of both interpolation and fusion. Early fusion relies highly on spatio-temporal registration and can be computationally demanding due to the high dimensionality of the initial input [84].

5.2.2 Mid-level fusion (feature-level)

As part of the Mid-Level Fusion paradigm, each sensor modality is initially processed by a separate feature extraction stream, which may be a specialised algorithm or a neural network. This allows each stream to be optimised for its specific data type (e.g., a CNN for images, a recurrent neural network for time-series data). To form a joint representation, high-level features from each modality are typically concatenated. This unified feature vector is then fed into a final set of layers that produce the final prediction [17]. This approach is highly modular, enabling the model to learn relationships between high-level, abstract concepts from each sensor rather than only raw pixel or signal values. Consequently, most recent advanced fusion systems in additive manufacturing, including those by Yang et al. [85] and Shen et al. [18], adopt this architecture due to its balance of flexibility and performance.

The effectiveness of this approach is demonstrated in multiple studies. The digital twin framework introduced by Chen et al. [17] is a prime example. They perform feature-level (mid-level) fusion by first extracting modality-specific features–melt-pool geometric descriptors from the coaxial vision images, Mel-Frequency Cepstral Coefficients (MFCC) and spectral features from the acoustic signal, and statistical temperature features from the Short-Wave Infrared (SWIR) thermal data, and then synchronising these features using their timestamps. The aligned feature vectors are spatiotemporally registered with the robot TCP position and concatenated into a single multimodal dataset, which is used to train the machine-learning model for location-specific defect prediction and virtual 3D quality mapping. In WAAM, Yu et al. [86] implemented this architecture to monitor deviations in robot trajectories. Their hybrid approach combines features from a U-Net (processing melt pool images) with those from a classical HOG algorithm (processing thermal images), illustrating the integration of deep learning and classical computer vision techniques.

Among advanced CNN architectures, the Residual Network (ResNet) enables the construction of networks with unprecedented depth. Its core innovation, the residual block, incorporates a skip connection, providing a direct path that bypasses convolutional layers. This ensures gradients can flow freely and mitigates the vanishing gradient problem. By learning only the residual difference from the input, the network is easier to train and achieves higher accuracy in complex tasks such as image classification. However, this architecture increases model complexity and memory usage relative to shallower networks, creating a trade-off between performance and computational cost.

Within the ResNet family, comprising versions up to ResNet-152, ResNet-18 is the most lightweight and computationally efficient variant, consisting of 18 convolutional layers with trainable weights. Its architecture uses the basic residual block structure, containing two 3×3 convolutional layers. Due to its balance of speed, size, and accuracy, the pre-trained ResNet-18 is widely adopted as a robust baseline for rapid experimentation, transfer learning, and deployment in resource-constrained computer vision applications.

The work presented in [18] demonstrates an advanced methodology by fusing melt-pool images and infrared data using an "implicit fusion" method. They employed two parallel ResNet-18 streams fused via a shared-parameter fully connected layer, forcing the model to learn a shared decision space. Their multimodal model significantly outperformed single-modal counterparts, reinforcing the advantages of fusion. Similarly, Yang et al. [87] developed an intelligent control system that fuses data from an infrared camera and a 3D laser scanner, employing a mid-level architecture with a ResNet stream for thermal images and a Point Cloud Transformer (PCT) stream for geometric data. These fused features are then used in an inverse model to predict optimal process parameters.

A truly comprehensive monitoring strategy, however, extends beyond vision and thermal data. The mid-level fusion architecture is well-suited to incorporating additional modalities. Reisch et al. [88] and Pringle et al. [89] demonstrated that spectral, acoustic, and RF signals provide unique information about process stability and arc characteristics that cameras cannot capture. Similarly, Hauser et al. [53] showed that electrical signals can fingerprint the material itself. Thus, while vision-thermal fusion currently represents the state of the art for combining geometric and physical information, the future of intelligent monitoring likely lies in multimodal systems that combine multiple types of measurements to create a more comprehensive digital twin of the process.

5.2.3 Late fusion (decision-level)

A late fusion method, also known as decision-level fusion, involves integrating data at the end of the process. In this approach, each sensor type is processed by its own independent predictive model. Models are trained separately, and each predicts a different outcome, such as a classification score or probability. The final decision is made by combining these outputs using simple rules, such as averaging, voting, or a weighted sum [83]. The process is similar to a "committee of experts," where each expert offers an opinion, and the final decision is based on the group's input. The main advantages of Late Fusion are its simplicity and flexibility. Each model can be trained or updated independently, and the system remains functional even if one sensor fails. Its main limitation is that it cannot learn direct interactions between features from different sensors [90].

The comprehensive sensor evaluation by Reisch et al. [88] provides a strong rationale for such a system. The welding camera and spectrometer were found to be the most effective in-process nozzle-to-work (NtW) sensors due to their inherent robustness and strong correlation with actual distances. These optical-based sensors consistently delivered reliable data, making them the preferred choices for immediate implementation in monitoring systems. In contrast, while electrical and acoustic sensors showed promise, they require further focused research to enhance their overall reliability. This will reduce their dependence on environmental and process-specific variables. This detailed analysis provides a foundation for designing a multi-modal sensing system that leverages the strengths of different sensors through a weighted voting scheme, prioritising outputs from more reliable sensors, such as the camera and spectrometer.

From a critical perspective, although late fusion is robust, it is fundamentally less capable than mid-level fusion of identifying feature interactions. For instance, it cannot detect that a specific visual feature combined with a specific thermal feature indicates an impending defect. Nevertheless, this simplicity can also be advantageous. In the multi-arc system proposed by [91], the complexity of the setup may make a modular, late-fusion approach more practical, with separate models monitoring each arc and their outputs aggregated by a central controller. Similarly, the open-source "Arc Analyser" from the study by Pringle et al. [89] could be implemented as a late-fusion system, with independent classifiers for sound, light, and RF signals voting on the process quality.

In conclusion, the choice of fusion architecture is a critical design decision with no universally correct answer. Early fusion offers the deepest level of integration but faces significant technical challenges. Late fusion provides robustness and simplicity at the expense of detailed insight. Mid-level fusion currently represents the optimal compromise for research, providing a flexible and robust framework to integrate diverse data streams, as demonstrated by its successful applications across WAAM and related Directed Energy Deposition (DED) processes.

5.3 Beyond accuracy: explainability, human-in-the-loop, and deployment constraints

In WAAM production environments, model selection is governed not only by accuracy and latency but also by explainability, the ability to encode human knowledge in the loop, and practical deployment constraints. Early fusion can deliver the tightest cross-modal coupling, but it requires strict temporal and spatial synchronisation, greater computational resources, and produces decisions that are harder to audit post hoc. Mid-level fusion provides a good balance: each encoder can be examined independently (e.g., saliency in image encoders, attention maps in thermal transformers, or feature importance in tabular channels), while the joint head captures cross-modal interactions. Late fusion prioritises operator trust and maintainability: per-modality scores can be logged, thresholded, and combined with rules or weighted voting, which enables straightforward overrides and clear fault isolation when a sensor becomes unreliable. Human-in-the-loop operation is common in industrial WAAM. Typical patterns include rule-based vetoes or severity thresholds on modality-specific indicators, weight schedules that shift reliance toward more robust channels under noise or occlusion, and operator acknowledgement of rare but high-impact alarms before escalation. Such mechanisms are easiest to implement and audit when modalities remain separable at decision time. Practical constraints often dominate the final design. These include heterogeneous frame rates and intermittent dropouts, air-gapped cells or limited industrial PCs, strict change control, and integration with legacy controllers and safety systems. In particular, brownfield cells (existing WAAM or arc-welding workcells retrofitted with new sensing and machine learning) commonly impose limited compute, mixed networking, and certification or audit requirements. In these conditions, late fusion can be preferable even when raw accuracy is similar across approaches, because it tolerates missing or unsynchronised channels, supports per-sensor auditing, and allows incremental roll-outs. Related perspectives on human-AI collaboration for AM decision-making and auditability are discussed by Papacharalampopoulos et al. [92]. Table 5 summarises these non-accuracy trade-offs and indicates representative WAAM scenarios where each strategy is advantageous.

Beyond accuracy, these considerations often drive architecture choice in production. Early fusion suits tightly controlled research settings; mid-level fusion balances interaction modelling with maintainability; and late fusion is frequently optimal in retrofits and certificated environments where explainability, human oversight, and operational resilience are paramount. This perspective complements the performance-focused comparison and prepares the ground for the control strategies discussed in Sect. 6.

6 From intelligent monitoring to autonomous control

Intelligent monitoring systems are the backbone of modern additive manufacturing. They continuously capture and interpret process behaviour across multiple physical domains. By fusing data from cameras, infrared sensors, and other modalities, these systems provide a detailed view of melt pool dynamics, layer geometry, and thermal gradients. Manufacturing becomes dynamic and adaptive when this integrated information is available, rather than a rigid, pre-programmed procedure. However, monitoring alone represents only the observational aspect of intelligence. Its ultimate value lies not just in identifying irregularities but in enabling an adequate response. The full potential of such systems is realised when their diagnostic insights are integrated within a

Table 5 Non-accuracy factors guiding fusion choice in WAAM deployments

Factor	Early fusion	Mid-level fusion	Late fusion
Explainability and audit	Hardest to attribute decisions due to pixel-level coupling	Encoder-level attributions possible per modality; joint head still inspectable	Easiest: per-modality scores/logs; transparent rules or weighted votes
Human knowledge in the loop	Requires retraining or complex constraints to add rules	Rules can be injected at encoder or head with moderate effort	Natural fit: thresholds, votes, operator overrides without retraining
Synchronisation tolerance	Low (tight temporal and spatial lock)	Medium (time-align feature batches)	High (align decisions, not pixels)
Fault isolation and resilience	Coupled failure modes across streams	Partial isolation via separate encoders	High isolation; modalities can be down-weighted or dropped at runtime
Compute and deployment	Highest GPU and memory; tight networking and triggers	Moderate; modular scaling by encoder	Lowest; runs on mixed hardware; tolerant to missing data
WAAM scenarios where it shines	Research prototypes with rigid triggers and ample GPUs	Balanced R&D and pilot lines that must capture interactions	Brownfield cells (retrofits), certification/audit needs, heterogeneous sensors, staged roll-outs

control architecture that can autonomously execute corrective actions. These systems adjust parameters such as wire feed rate, travel speed, and current in real-time to maintain process stability and part quality. The purpose of this section is to examine how sensor-informed control frameworks can close the detection-response loop by moving from passive observation to active intervention. These closed-loop control strategies vary widely in their operational principles, response times, and degrees of autonomy. Broadly speaking, they can be categorised into three groups. Strategic open-loop control utilises pre-processed sensor data to predict and plan adjustments before deposition, focusing on preventive stability. Discrete inter-layer correction applies periodic updates between successive layers, recalibrating parameters based on accumulated deviations. Finally, continuous intra-layer feedback represents the most advanced approach. It performs on-the-fly adjustments within each layer to directly counter process disturbances as they arise. Together, these approaches illustrate the ongoing evolution of intelligence in WAAM, moving from simple monitoring toward adaptive, self-regulating manufacturing systems.

6.1 Layer-by-layer corrective strategies in discrete control

Inter-layer control represents the first and most straightforward level of in-process feedback. It operates on a discrete "deposit, measure, correct" cycle. In this approach, each layer is fully deposited before the system pauses to assess the resulting geometry. The measurement subsystem, often using vision, laser scanning, or structured-light sensors, quantifies the cumulative deviation from the intended build profile. Based on this diagnostic information, a corrective action is then applied before depositing the next layer. This method is particularly effective for compensating slow, accumulative errors such as geometric drift, layer height inconsistencies, or heat buildup. However, because it functions between layers rather than within them, it cannot respond to short-lived or transient anomalies that occur during a single layer, such as arc instability, momentary feed interruptions, or spatter interference.

Two main approaches for implementing inter-layer correction have been demonstrated in the literature. The first, and most direct, is physical part modification, where the already-deposited layer is treated as a workpiece to be physically reshaped. A canonical example is presented by Tang et al. [56], who developed a hybrid additive–subtractive system. In their setup, a 3D scanner captures the surface topography of each newly deposited layer. Any excess material is then precisely removed using CNC milling, while missing regions are refilled through localised re-deposition. This hybrid approach ensures high geometric fidelity and provides a near-perfect foundation for subsequent layers. However, the trade-offs are substantial. From a broader perspective, this method, despite its advantages, can slow the overall build rate and increase tool-change overhead. It may also disrupt thermal continuity, potentially causing residual stress variations and non-uniform microstructures.

The second, less invasive, and more commonly used approach is process plan modification, in which no physical rework occurs. Instead, the control system adjusts the planned deposition strategy for the following layers based on measured deviations. Baldauf et al. [93] developed an adaptive control system for maintaining a stable Contact Tip to Work Distance (CTWD). Their setup uses a vision-based height measurement system to estimate the average build height after each layer. The controller then modifies the robot's G-code in real-time. If the part is too high, the system skips the deposition path for the next layer. If the part is too low, it commands a repetition of the previous path. The main advantage of this method is its non-invasive nature. It avoids mechanical interaction with the deposited material, thereby preserving the metallurgical integrity and thermal continuity of the build. It can also be implemented with minimal hardware modifications, relying primarily on software control and real-time data processing. Nevertheless, its correction resolution is inherently limited. The available corrective actions, adding or omitting an entire layer, are discrete. This means the system cannot make fine-grained adjustments to minor deviations, such as ±0.5 mm. Consequently, the control acts in an all-or-nothing manner. While this is effective for larger errors, minor deviations may remain uncorrected.

A more precise and adaptive extension of inter-layer control was proposed by [94], as shown in Fig. 11. Instead of altering the toolpath, they adjusted welding parameters, such as wire feed speed and travel speed, based on measured layer errors. This method bridges discrete inter-layer correction and continuous feedback, allowing finer compensation without physically modifying the part. Their work compared two control strategies, the widely used Proportional, Integral, Derivative (PID) controller and the advanced Model Predictive Control (MPC). PID operates reactively. The Proportional term addresses current error, the Integral term eliminates accumulated deviations, and the Derivative term damps oscillations based on the rate of change. MPC, in contrast, is proactive. It uses a mathematical process model to predict future outputs over a defined time horizon, solves an optimisation problem to determine the best sequence of corrective actions, and applies only the first action before repeating the cycle.

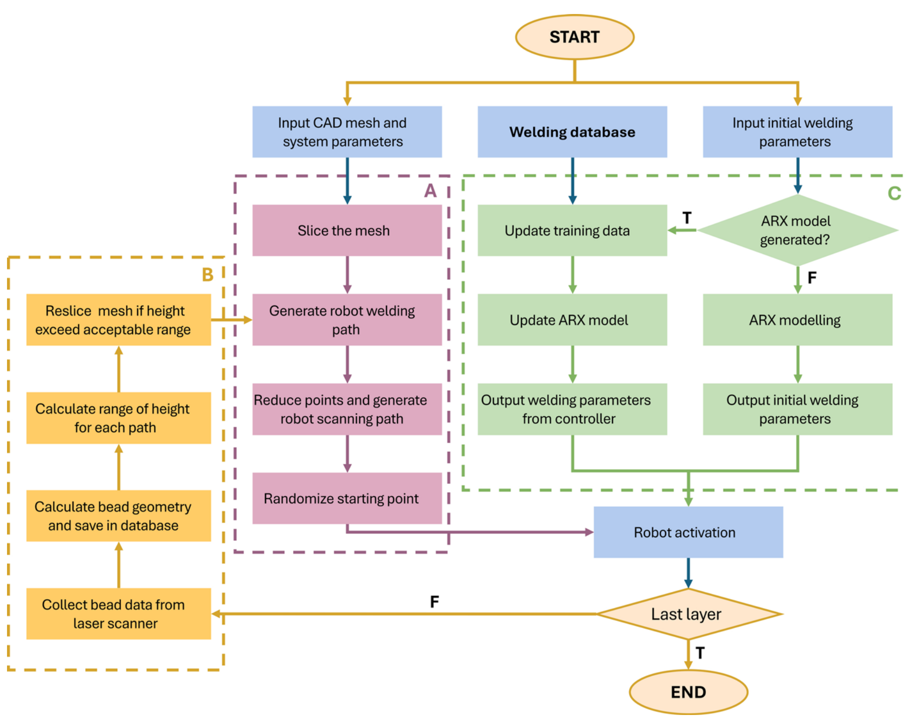

Fig. 11 Flowchart of Layer-by-Layer Corrective Strategy, composed of robot path generation module (**A**), bead geometry identification module (**B**) and control module (**C**). Reproduced from [94], licensed under CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/)

On an inclined tube with 30 layers and 12 control points per layer, MPC achieved markedly lower layer-height variability than PID. The height range was reduced from 4.44 mm with PID to 1.50 mm with MPC, corresponding to a 296% improvement. Likewise, the per-layer standard deviation across the control points decreased from 1.20 to 0.51 mm, representing a 235% reduction. For inclined thin-wall structures (20 layers), both PID and MPC significantly stabilised the process. Wall-height fluctuations were reduced from approximately 2 mm without control to below 0.5 mm under feedback control, while the standard deviation decreased from 0.7 to 0.2 mm. This demonstrates the advantage of predictive, model-based approaches for managing the nonlinear, interdependent thermal and geometric effects of WAAM. Together, these inter-layer strategies illustrate a spectrum of corrective methods, from physical reshaping to plan adaptation to parameter optimisation. Each approach balances fidelity, speed, and thermal consistency differently. While discrete inter-layer control cannot react to transient intra-layer disturbances, it provides a robust framework for addressing accumulated errors. It serves as a practical step toward higher-resolution, continuous feedback architectures in WAAM.

6.2 Intra-layer feedback strategies in continuous control

In contrast to discrete inter-layer methods, intra-layer feedback is the most dynamic and responsive type of process control in WAAM. These systems operate continuously during single-layer deposition. They utilise high-frequency sensor feedback to adjust process parameters in real-time. By reacting instantly to deviations, intra-layer controllers can correct transient disturbances such as arc fluctuations, inconsistent wire feeding, and localised thermal changes. These are issues that discrete control can only handle after they have accumulated into measurable defects. In effect, continuous feedback turns WAAM from a reactive correction process into a self-stabilising system. It prevents errors from forming in the first place. This type of control spans a wide range of methods, from classical industrial controllers to modern data-driven and intelligent systems.

The Proportional–Integral–Derivative (PID) controller is the foundation of real-time feedback and remains one of the most widely used methods in WAAM. PID maintains stability by continuously calculating corrections based on current, past, and predicted error trends. Reisch et al. [88] provided a simple example: they used a Proportional (P) controller to regulate the Nozzle-to-Work Distance (NTWD) by continuously adjusting the welding speed. Kumar et al. [95] demonstrated a more advanced setup, using a complete PID controller to regulate input current based on real-time temperature feedback from a pyrometer. This prevented overheating and maintained a stable melt pool under varying conditions.

While PID controllers are fast, robust, and relatively easy to implement, they are limited by their use of fixed gain parameters (K_p, K_i, K_d). These gains are usually tuned for initial conditions and stay constant throughout the build. However, as the part grows taller, its heat dissipation, surface emissivity, and thermal inertia change significantly. As a result, a PID tuned for early layers may perform poorly or even become unstable later on. This sensitivity to changing process conditions is a significant limitation of classical controllers in WAAM.

To handle these time-varying interactions, Model Predictive Control (MPC) offers a more advanced, proactive solution. MPC uses a process model to predict how the melt pool and bead geometry will evolve over a short time horizon, as illustrated in Fig. 12. This is in response to changes in the parameters. It then solves an optimisation problem to determine the most appropriate control sequence. Typically, adjusting the Wire Feed Speed (WFS) or Travel Speed (TS) minimises deviations from the target layer width or temperature.

The work presented in [96] demonstrates that MPC is a highly effective control strategy for real-time layer width control in WAAM, as shown in Fig. 13a. It offers superior performance in terms of stability, tracking accuracy, and robustness compared to traditional methods. This makes it well-suited to the dynamic and complex nature of additive manufacturing. Beyond PID and MPC, several new intelligent control methods are gaining attention for handling the nonlinear behaviour of WAAM. Fuzzy logic control (FLC) is especially promising. Unlike PID or MPC, which rely on exact mathematical models, FLC mimics human reasoning and intuition through if–then rules. For example, "IF the temperature is too high AND the bead's width is narrow, THEN decrease the current slightly." This allows flexible decision-making under uncertainty. It is beneficial when process dynamics are too complicated or poorly understood to model precisely. By directly encoding operator experience into control rules, FLC can maintain adaptive performance in situations where model-based methods struggle or need extensive calibration. In summary, intra-layer feedback offers higher temporal responsiveness in WAAM control. From PID regulation to predictive MPC and fuzzy-logic intelligence, these systems aim to maintain process stability and geometric fidelity while responding in real-time. As sensor technology and computational power improve, intra-layer feedback systems are expected to evolve into fully autonomous hybrid controllers. These controllers will combine physical modelling, data-driven learning, and rule-based reasoning to achieve precision and high reliability in additive manufacturing. While discrete inter-layer control addresses cumulative errors, it remains limited by its reliance on fixed-gain controllers, which may struggle to handle the nonlinear, time-varying dynamics of WAAM. A more flexible approach is adaptive control, in which the controller continuously updates its internal parameters in real-time. Adaptive controllers are essentially self-tuned systems. They monitor ongoing process behaviour, detect changes in dynamics, and adjust gains or other control parameters to maintain optimal performance throughout the build. This ability enables the system to compensate for effects such as changing thermal conditions, variations in material properties, or slight sensor

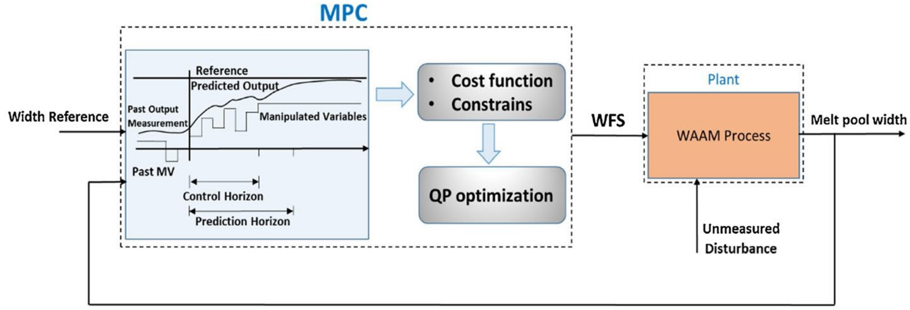

Fig. 12 Principle of MPC for WAAM (reproduced from [96], with permission granted to Springer Nature Limited to publish under CC BY 4.0)

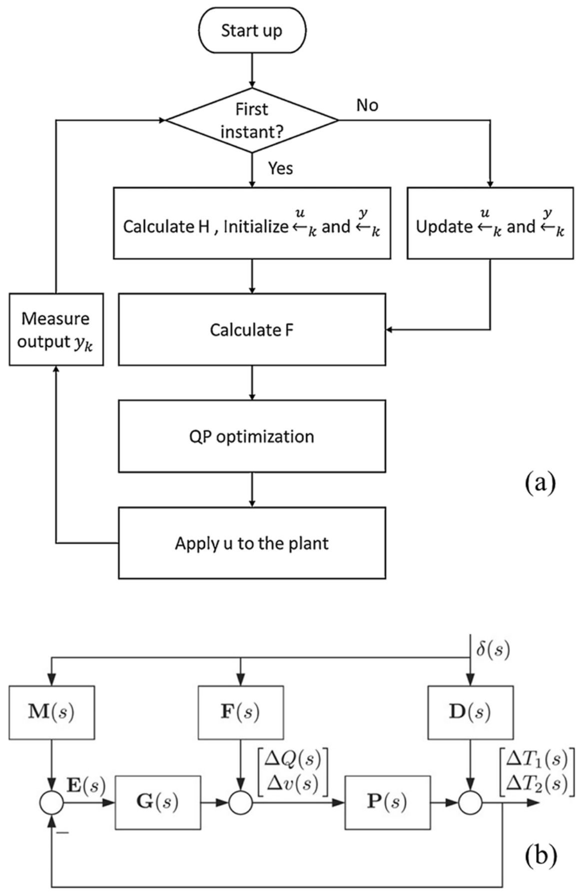

Fig. 13 **a** Algorithm flowchart of MPC in WAAM (reproduced from [96], with permission granted to Springer Nature Limited to publish under CC BY 4.0). **b** Proposed structure of combined feedforward–feedback QFT controller (reproduced from [98], licensed under CC BY 4.0). CC BY 4.0 licence details available at: https://creativecommons.org/licenses/by/4.0/

drifts. This approach maintains control even under fluctuating conditions. One way to implement adaptive control is to improve classical proportional, Integral, and Derivative (PID) controllers using advanced optimisation methods. For example, Li et al. [97] demonstrated a DMC-IEPO-PID algorithm for composite additive manufacturing. This hybrid controller combines three complementary techniques. First, Dynamic Matrix Control (DMC) acts as a forward-looking predictive component. It uses a process model to forecast how the system will respond over a short time horizon. This enables it to make proactive adjustments to prevent future errors, rather than reacting to current ones. Second, Improved Eagle Perching Optimisation (IEPO) works as an intelligent tuner. Inspired by an eagle's intelligent hunting behaviour, this metaheuristic algorithm continually searches for optimal PID parameters in real-time. By carefully testing different settings, the controller can quickly adjust to the current process state. Finally, the PID controller, now guided by DMC's predictions and continuously optimised by IEPO, performs fine-tuned feedback control. According to [97], the traditional PID algorithm is slower to respond across all temperatures, resulting in longer stabilisation times. While this specific algorithm has not yet been used in WAAM, it shows the promise of combining predictive control and adaptive optimisation with PID to manage complex and changing deposition processes. In principle, such a system could provide finer layer control, better process stability, and higher build quality. This is especially true when traditional fixed-gain controllers cannot handle changing process conditions. The outcome is more precise layer regulation and improved process stability.

For situations where robustness against known uncertainties is critical, Quantitative Feedback Theory (QFT) offers a mathematically rigorous design framework. QFT is a frequency-domain approach that designs a controller guaranteed to meet performance specifications across an entire family of possible process models. This makes it inherently resilient to variations in material properties, environmental conditions, or sensor noise. Masenlle et al. [98] applied this to WAAM to tackle the "direction flip problem," where sudden 180-degree toolpath reversals introduce rapid thermal disturbances as shown in Fig. 13b. By modelling these disturbances as predictable impulses, they designed a combined feedforward–feedback QFT controller. It modulates the heat source power and travel speed to stabilise the temperature. The outcome is improved layer consistency and geometric fidelity, showing how classically robust control methods can solve specific, critical WAAM challenges in practice. Together, adaptive and robust control strategies enhance inter-layer correction capabilities by incorporating real-time tuning and ensuring guaranteed performance in the presence of uncertainty. Adaptive control offers flexibility in responding to evolving process dynamics. Robust methods, such as QFT, ensure that the system remains stable and precise even under significant variations or disturbances. These approaches are a crucial step toward achieving fully autonomous, reliable WAAM operations capable of managing both predictable and unpredictable process variability. With these tools in place, operators can rely on the system to maintain a smooth process. It supports the transition to modern WAAM setups.

6.3 Integrated inter- and intra-layer strategies in hybrid control

While inter-layer control effectively handles cumulative errors, and intra-layer strategies excel at reducing short-term disturbances, neither approach alone is sufficient to fully manage the complex dynamics of WAAM. Hybrid control architectures combine the strengths of both methods, offering a more effective approach. This provides a multi-horizon control framework. It enables the system to handle both slow-evolving errors and fast, high-frequency disturbances simultaneously. It is an effective way to maintain stability across all time scales.

A good example is the system developed by Marcotte et al. [99]. Their setup uses two nested control loops. The outer inter-layer feedforward controller first examines the final error profile from the previous layer. It then pre-plans a compensatory torch velocity profile for the next layer. At the same time, an inner intra-layer feedback controller uses live infrared camera measurements to make fast, proportional adjustments to the torch path during deposition. Combining these loops enables the hybrid system to correct known, accumulated deviations from earlier layers while actively addressing unexpected, transient disturbances. This dual-loop design aims to improve robustness. It addresses both long-term and short-term sources of variability in WAAM processes. However, hybrid architectures are more complex. One major challenge is the potential for adverse interactions between the loops. Adjustments from one controller can sometimes amplify oscillations or conflict with the other, as observed in experimental error signals [99]. This highlights the need for careful loop tuning and coordination to maintain stability.

Despite these challenges, hybrid control offers a straightforward approach to combining the benefits of discrete and continuous strategies.

In summary, WAAM control has evolved from static, pre-planned optimisation toward dynamic, feedback-driven strategies. Choosing the right control architecture means balancing resolution of control actions, overall process speed, and the thermal and mechanical effects on the part. Most current research focuses on single-input, single-output (SISO) control of individual variables. The next significant step will be the development of robust multi-input, multi-output (MIMO) systems. These advanced controllers, likely enhanced with machine learning for MPC or adaptive control, will simultaneously adjust multiple process parameters. This will enable the system to control coupled outputs, such as bead height, bead width, and part temperature, in a coordinated manner.

Such all-in-one predictive control systems represent a likely direction for WAAM, moving toward more autonomous, reliable, and high-quality additive manufacturing. They show how far the field has come. Real-time feedback and intelligent control complement each other.

7 Open challenges and future research directions

Despite significant advances in both visual and thermal monitoring for metal additive manufacturing (AM), several challenges will likely guide future research. These challenges are not independent and should be prioritised according to their impact on deployable WAAM systems. Data fidelity and synchronisation represent the primary bottleneck, as unreliable measurements propagate through the entire monitoring, fusion, and control pipeline. Real-time multimodal data fusion and computational efficiency constitute the second critical layer, while generalisation and data availability are longer-term priorities for scalable industrial adoption. One of the main technical obstacles lies in data fidelity and sensor reliability. The harsh and highly dynamic environment of WAAM processes is characterised by intense arc light, metallic vapour (plume), and rapid temperature fluctuations, all of which significantly degrade the quality of both visible and thermal images. Current approaches largely rely on passive optical filtering and laboratory calibration, which are difficult to maintain during long-duration builds and under varying process conditions. Future work should focus on improving sensor-side filtration and hardware design, including multi-spectral and polarisation imaging systems, to minimise interference from arc and plume effects. These developments are crucial for achieving consistent and high-quality data capture under variable process conditions.

Thermal monitoring provides valuable information about melt pool dynamics and cooling rates. However, accurate measurement of key process parameters, such as melt pool dimensions and thermal gradients, strongly depends on precise emissivity estimation. Since emissivity varies with the states of molten and partially solidified metals, real-time emissivity mapping remains a critical challenge. This area could benefit from sensor fusion and ML.

Another significant challenge involves bridging the gap between simulation and reality. While physics-based simulations and digital twins offer powerful tools for understanding thermal behaviour and predicting microstructural evolution, their accuracy depends on how well they represent the actual process. Enhancing these models with more advanced calibration methods and the real-time integration of sensor data will be essential to maintaining accuracy during operation. Developing computationally efficient yet physically realistic models, such as those based on Fourier Neural Operators, will also be necessary to enable real-time, physics-informed control of WAAM processes. Most existing approaches assume constant or pre-calibrated emissivity values, which limits their applicability in dynamic WAAM conditions where surface state and oxidation continuously evolve.

A further challenge is the transition from monitoring to autonomous control. Many current studies have demonstrated the effectiveness of post-process defect detection or simple single-variable closed-loop control, such as layer-width regulation [96] or temperature control using PID or DMC algorithms [97]. The next step is to develop comprehensive, multi-objective control systems capable of simultaneously managing geometry, temperature, and microstructure based on fused, real-time multimodal data (visual, thermal, acoustic, and electrical). Achieving this goal requires fast and efficient algorithms that can merge diverse data streams (data fusion) and convert detected anomalies into immediate, multi-parameter control actions (e.g., adjustments to travel speed, wire feed rate, or current/voltage) with sub-millisecond response times. In practice, this transition is constrained not only by algorithmic speed but also by sensing latency, data-transfer bandwidth, and actuator response time, which together define the achievable control frequency. Future artificial intelligence (AI) research should prioritise deep reinforcement learning (DRL) approaches trained on large, synthetic, and experimental defect datasets to create predictive control systems that can anticipate and prevent defects rather than respond to them. However, DRL-based control is currently limited by the scarcity of temporally aligned multimodal datasets and the difficulty of ensuring stable exploration in safety-critical manufacturing environments.

Beyond fully autonomous systems, recent research is moving toward Human–AI Collaboration (HAIC) in line with Industry 5.0. This approach aims to improve process transparency and safety. However, several challenges still exist [92]. One key issue is scalability across multiple machines. Another is the high cognitive load on operators when they must interpret large volumes of sensor data. In addition, Explainable AI is needed to clarify how AI systems make decisions. Additionally, managing uncertainty propagation from visual and thermal data through to the final decision-making stage remains an open and critical challenge in this field. Finally, one of the most persistent challenges concerns data scarcity and generalisation. As seen in vision-based machine learning studies, model performance often depends on dataset size and diversity. Models trained on small datasets, especially for rare but critical defect types, fail to generalise across different materials, geometries, or process conditions. Future efforts should aim to develop large, open-access, multimodal datasets for WAAM. Establishing shared metadata standards, benchmark metrics, and open-source repositories for various AM modalities will be essential for accelerating research and ensuring fair comparison across algorithms. These efforts will help move the field from isolated, machine-specific studies toward general, deployable industrial solutions. Moreover, advanced learning methods that perform well with limited data, such as material-adaptive and multi-source transfer learning, require further refinement to enable reliable, universal monitoring and control frameworks. From a research roadmap perspective, short-term efforts should focus on robust multimodal sensing, synchronisation strategies, and benchmark datasets. Mid-term research should address real-time physics-informed data fusion and multi-parameter closed-loop control. Long-term progress will depend on generalisable, self-adaptive learning frameworks that can transfer across materials, geometries, and machine platforms.

8 Conclusions

Multimodal sensing combined with ML delivers a fuller view of WAAM than any single sensor alone. In particular, vision–thermal fusion improves both defect detectability and the interpretation of process dynamics, and it provides actionable inputs for closed-loop control. Choosing among early-, mid-, and late-fusion strategies requires balancing synchronisation effort, computational cost, and explainability in the intended deployment. Recent progress in inter-layer, intra-layer, and hybrid control confirms that sensor-informed feedback can make WAAM more stable and repeatable.

This review contributes three takeaways for practitioners. First, it consolidates the foundations of vision and thermal sensing for WAAM, detailing measurable variables and traditional pipelines, and it summarises the most effective unimodal ML models by task and deployment constraint. Second, it organises the state of the art in vision–thermal fusion–early, mid, and late–and relates each choice to the realities of synchronisation, robustness to occlusions, fault isolation, and operator trust. Third, it links intelligent monitoring to control by comparing inter-layer, intra-layer, and hybrid strategies and by indicating where model-predictive, adaptive, and data-driven controllers are most useful.

Looking ahead, advanced control (including deep reinforcement learning, DRL) should be framed for WAAM as a concrete recipe rather than as a generic promise. In practice: (1) the multi-objective vector groups geometry targets (e.g., bead width/height error, layer-height drift), thermal targets (e.g., inter-layer temperature and melt-pool proxies), and quality proxies (e.g., porosity/balling indicators); (2) the action space includes wire-feed speed, travel speed, current/voltage waveform parameters and weaving amplitude/frequency; (3) real-time viability is set by the end-to-end latency budget (sensor acquisition, pre-processing, inference, command dispatch) relative to sensor framerates and actuator bandwidth; (4) safety and reliability are enforced via action filters and envelope protection, with MPC/PID fallbacks; (5) training is enabled by calibrated digital twins and domain randomisation for sim-to-real transfer, complemented by offline learning from logged runs; and (6) evaluation should report both control metrics (overshoot, settling, steady-state error) and product metrics (geometry MAE, defect rate), not accuracy alone.

In short, the near-term priorities are robust data capture under arc/plume conditions, physics-consistent sim-to-real methods, and open, diverse multimodal datasets to improve generalisation. Framing advanced control as a deployable recipe–objectives, actions, latency, safety guards, training, and evaluation–can accelerate the transition from intelligent monitoring to reliable, closed-loop WAAM on the shop floor.

Author contributions

MG conducted the comprehensive literature search, designed the review structure, and drafted the main manuscript text and figures. CM and DC supervised the work and provided methodological guidance. CM secured the PhD funding through CARACOL, coordinated the submission, and led major revisions, including figure editing. AB, FT, and GA contributed industrial WAAM expertise and provided application examples and technical inputs. All authors reviewed and approved the final manuscript.

Funding
This work was supported by CARACOL, which funded Maedeh Gourakani's PhD program.

Data availability
No datasets were generated or analysed during the current study.

Materials availability
Not applicable.

Code availability
Not applicable.

Declarations

Ethics approval and consent to participate
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 15 January 2026 / Accepted: 16 April 2026
Published online: 28 April 2026

References

1. Jackson MA, Asten A, Morrow JD, Min S, Pfefferkorn FE. A comparison of energy consumption in wire-based and powder-based additive-subtractive manufacturing. In: Procedia Manufacturing, pp. 989–1005. Elsevier B.V. https://doi.org/10.1016/j.promfg.2016.08.087
2. Tang HP, Qian M, Liu N, Zhang XZ, Yang GY, Wang J. Effect of powder reuse times on additive manufacturing of Ti-6Al-4V by selective electron beam melting. JOM. 2015;67(3):555–63. https://doi.org/10.1007/s11837-015-1300-4.
3. Cunningham CR, Flynn JM, Shokrani A, Dhokia V, Newman ST. Invited review article: strategies and processes for high quality wire arc additive manufacturing. Addit Manuf. 2018. https://doi.org/10.1016/j.addma.2018.06.020.
4. Singh SR, Khanna P. Wire arc additive manufacturing (waam): a new process to shape engineering materials. In: Materials today: proceedings. Elsevier Ltd. pp. 118–128. https://doi.org/10.1016/j.matpr.2020.08.030
5. Wu B. A review of the wire arc additive manufacturing of metals: properties, defects and quality improvement. J Manuf Process. 2018. https://doi.org/10.1016/j.jmapro.2018.08.001.
6. Gisario A, Kazarian M, Martina F, Mehrpouya M. Metal additive manufacturing in the commercial aviation industry: a review. J Manuf Syst. 2019. https://doi.org/10.1016/j.jmsy.2019.08.005.
7. Ding D, Pan Z, Cuiuri D, Li H. Wire-feed additive manufacturing of metal components: technologies, developments and future interests. Int J Adv Manuf Technol. 2015. https://doi.org/10.1007/s00170-015-7077-3.
8. Derekar K. A review of wire arc additive manufacturing and advances in wire arc additive manufacturing of aluminium. Mater Sci Technol. 2018;34(8):895–916. https://doi.org/10.1080/02670836.2018.1455012.
9. Soundarapandiyan G, Johnston C, Khan RHU, Chen B, Fitzpatrick ME. A technical review of the challenges of powder recycling in the laser powder bed fusion additive manufacturing process. J Eng. 2021;2:97–103. https://doi.org/10.1049/tje2.12013.
10. Knezović N, Topić A. Wire and arc additive manufacturing (waam)—a new advance in manufacturing. In: Lecture notes in networks and systems 42; pp. 65–71. https://doi.org/10.1007/978-3-319-90893-9_7
11. Yang D, Wang G, Zhang G. Thermal analysis for single-pass multi-layer GMAW based additive manufacturing using infrared thermography. J Mater Process Technol. 2017;244:215–24. https://doi.org/10.1016/j.jmatprotec.2017.01.024.
12. Chandrasekhar N, Vasudevan M, Bhaduri AK, Jayakumar T. Intelligent modeling for estimating weld bead width and depth of penetration from infra-red thermal images of the weld pool. J Intell Manuf. 2015;26(1):59–71. https://doi.org/10.1007/s10845-013-0762-x.
13. Repossini G, Laguzza V, Grasso M, Colosimo BM. On the use of spatter signature for in-situ monitoring of laser powder bed fusion. Addit Manuf. 2017;16:35–48. https://doi.org/10.1016/j.addma.2017.05.004.
14. Huang Y, Wu D, Zhang Z, Chen H, Chen S. EMD-based pulsed TIG welding process porosity defect detection and defect diagnosis using GA-SVM. J Mater Process Technol. 2017;239:92–102. https://doi.org/10.1016/j.jmatprotec.2016.07.015.
15. Zhang Z, Yu H, Lv N, Chen S. Real-time defect detection in pulsed GTAW of al alloys through on-line spectroscopy. J Mater Process Technol. 2013;213(7):1146–56. https://doi.org/10.1016/j.jmatprotec.2013.01.012.
16. Franke J, Heinrich F, Reisch RT. Vision based process monitoring in wire arc additive manufacturing (WAAM). J Intell Manuf. 2025;36:1711–21. https://doi.org/10.1007/s10845-023-02287-x.
17. Chen L, Bi G, Yao X, Tan C, Su J, Ng NPH, et al. Multisensor fusion-based digital twin for localized quality prediction in robotic laser-directed energy deposition. Robot Comput-Integr Manuf. 2023;84:102581. https://doi.org/10.1016/j.rcim.2023.102581.
18. Shen B, Lu J, Wang Y, Chen D, Han J, Zhang Y, et al. Multimodal-based weld reinforcement monitoring system for wire arc additive manufacturing. J Market Res. 2022;20:561–71. https://doi.org/10.1016/j.jmrt.2022.07.086.
19. Zhao Z, Guo Y, Bai L, Wang K, Han J. Quality monitoring in wire-arc additive manufacturing based on cooperative awareness of spectrum and vision. Optik. 2019;181:351–60. https://doi.org/10.1016/j.ijleo.2018.12.071.
20. Hauser T, Reisch RT, Seebauer S, Parasar A, Kamps T, Casati R, et al. Multi-material wire arc additive manufacturing of low and high alloyed aluminium alloys with in-situ material analysis. J Manuf Process. 2021;69:378–90. https://doi.org/10.1016/j.jmapro.2021.08.005.
21. Mahato DK, Chandra M, Tanmay, Rajak S. Layer-by-layer thermal modelling for wire-arc additive manufacturing of aluminum alloy wall. Mater Today Proc. 2024;113:120–5. https://doi.org/10.1016/j.matpr.2023.08.046.

22. Chen X, Kong F, Fu Y, Zhao X, Li R, Wang G, et al. A review on wire-arc additive manufacturing: typical defects, detection approaches, and multisensor data fusion-based model. Int J Adv Manuf Technol. 2021;117(3):707–27. https://doi.org/10.1007/s00170-021-07807-8.
23. Cai X, Xia Y, Dong B, Zhou T, Lin S. Effects of deposition paramaters on the microstructure evolution of wire arc additive manufactured Al–Zn–Mg–Cu alloy. J Market Res. 2023;26:1572–83. https://doi.org/10.1016/j.jmrt.2023.07.282.
24. Arana M, Ukar E, Rodriguez I, Iturrioz A, Alvarez P. Strategies to reduce porosity in Al-Mg WAAM parts and their impact on mechanical properties. Metals. 2021;11(3):524. https://doi.org/10.3390/met11030524.
25. James WS, Ganguly S, Pardal G. High temperature performance of wire-arc additive manufactured Inconel 718. Sci Rep. 2023;13(1):4541. https://doi.org/10.1038/s41598-023-29026-9.
26. Koley S, Ganguly S. Prevention of hot cracking in wire-arc additive manufacturing of iron-based shape memory alloy using a novel inter-layer machine hammer peening process. Mater Sci Eng, A. 2025;937:148453. https://doi.org/10.1016/j.msea.2025.148453.
27. Wacker C, Köhler M, David M, Aschersleben F, Gabriel F, Hensel J, et al. Geometry and distortion prediction of multiple layers for wire arc additive manufacturing with artificial neural networks. Appl Sci. 2021;11(10):4694. https://doi.org/10.3390/app11104694.
28. Cho H-W, Shin S-J, Seo G-J, Kim DB, Lee D-H. Real-time anomaly detection using convolutional neural network in wire arc additive manufacturing: molybdenum material. J Mater Process Technol. 2022;302:117495. https://doi.org/10.1016/j.jmatprotec.2022.117495.
29. Jafari D, Vaneker THJ, Gibson I. Wire and arc additive manufacturing: opportunities and challenges to control the quality and accuracy of manufactured parts. Mater Des. 2021;202:109471. https://doi.org/10.1016/j.matdes.2021.109471.
30. Albannai AI. A brief review on the common defects in wire arc additive manufacturing. Int J Curr Sci Res Rev. 2022;05(12):4556–76. https://doi.org/10.47191/ijcsrr/v5-i12-19.
31. Safyari M, Haunreiter F, Furuta S, Khoo PL, Mayrhofer F, Kobayashi M, et al. Manipulating nanostructure during wire arc additive manufacturing defeats hydrogen embrittlement. Corros Sci. 2024;233:112088. https://doi.org/10.1016/j.corsci.2024.112088.
32. Gurmesa F, Lemu H, Adugna Y, Harsibo M. Residual stresses in wire arc additive manufacturing products and their measurement techniques: a systematic review. Appl Mech. 2024;5(3):420–49. https://doi.org/10.3390/applmech5030025.
33. Liu N, Jafarzadeh S, Yu Y. Domain agnostic fourier neural operators. arXiv (2023). https://doi.org/10.48550/ARXIV.2305.00478. arXiv:2305.00478
34. Alcaraz JY, Sharma A, Tjahjowidodo T. Predicting porosity in wire arc additive manufacturing (WAAM) using wavelet scattering networks and sparse principal component analysis. Weld World. 2024;68(4):843–53.
35. Li Y, Mu H, Polden J, Li H, Wang L, Xia C, et al. Towards intelligent monitoring system in wire arc additive manufacturing: a surface anomaly detector on a small dataset. Int J Adv Manuf Technol. 2022;120(7–8):5225–42. https://doi.org/10.1007/s00170-022-09076-5.
36. Šket K, Brezočnik M, Karner T, Belšak R, Ficko M, Vuherer T, et al. Predictive modelling of weld bead geometry in wire arc additive manufacturing. J Manuf Mater Process. 2025;9(2):67.
37. Xia C, Pan Z, Li Y, Chen J, Li H. Vision-based melt pool monitoring for wire-arc additive manufacturing using deep learning method. Int J Adv Manuf Technol. 2022;120(1–2):551–62. https://doi.org/10.1007/s00170-022-08811-2.
38. Afolabi SO, Akinsooto O. Conceptual framework for mitigating cracking in superalloy structures during wire arc additive manufacturing (WAAM). Int J Multidiscip Compreh Res. 2023;1(1):1–15.
39. Zhu Q, Huang Z, Li H. Research progress of acoustic monitoring technology in welding and additive manufacturing processes. Micromachines. 2026;17(2):246.
40. Kim KW, Kamerkar A, Chiu T-E, Abdi I, Qin J, Suder W, et al. Waam-vid: towards universal vision-based monitoring for wire arc additive manufacturing. Front Manuf Technol. 2025;5:1676365.
41. Rahman MA, Jamal S, Cruz MV, Silwal B, Taheri H. In situ process monitoring of multi-layer deposition in wire arc additive manufacturing (WAAM) process with acoustic data analysis and machine learning. Int J Adv Manuf Technol. 2024;132(9):5087–101.
42. MeltTools: cameras for monitoring automated welding. 2025. https://melttools.com/blogs/insight/cameras-for-monitoring-automated-welding. Accessed 20 Oct 2025.
43. Yin Z, Xiong J. Stereovision measurement of layer geometry in wire and arc additive manufacturing with various stereo matching algorithms. J Manuf Process. 2020;56:428–38. https://doi.org/10.1016/j.jmapro.2020.05.001
44. Yu R, Li W, Wu M, Wang J, Han Q, Wang J, et al. Advanced visual sensing and control in CMT-based WAAM processes. Front Mater. 2025. https://doi.org/10.3389/fmats.2025.1499635.
45. Shen H, Jin J, Liu B, Zhou Z. Measurement and evaluation of laser-scanned 3D profiles in wire arc hybrid manufacturing processes. Measurement (Lond). 2021;176(109089):109089.
46. CaviAdn: common welding methods and the use of laser illumination in welding imaging. 2020. https://www.cavitar.com/library/common-welding-methods-and-the-use-of-laser-illumination-in-welding-imaging/. Accessed 20 Oct 2025.
47. Monitoring weld deposition in additive manufacturing. https://andor.oxinst.com/learning/view/article/monitoring-weld-deposition-in-additive-manufacturing. Accessed 20 Oct 2025
48. A job weld done: How pulse lasers and high-speed cameras visualize extremely bright events. https://www.phantomhighspeed.com/news/newsarticles/2024/april/a-job-weld-done-how-pulse-lasers-and-high-speed-cameras-visualize-extremely-bright-events. Accessed 20 Oct 2025
49. https://www.lincolnelectric.com/en/welding-and-cutting-resource-center/process-and-theory/intelligent-robotic-arc-sensing. Accessed 20 Oct 2025
50. Dong K, Wu Q, Qin X, Hu Z, Hua L. In-situ optical monitoring and analysis of weld pool based on machine vision for wire and arc additive manufacturing. Int J Adv Manuf Technol. 2024;133(9–10):4865–78. https://doi.org/10.1007/s00170-024-14003-x.
51. Couto MO. Mapping of bead geometry in wire arc additive manufacturing systems using passive vision. J Control Autom Electr Syst. 2022;33(4):1136–47. https://doi.org/10.1007/s40313-021-00880-0.
52. Ajay V, Shrivastava A. A novel single-camera-based vision system for real-time measurement of bead width and height during wire arc additive manufacturing. Progress Addit Manuf. 2025;10:5923–35. https://doi.org/10.1007/s40964-024-00943-z.

53. Hauser T. Porosity in wire arc additive manufacturing of aluminium alloys. Addit Manuf. 2021;41:101993. https://doi.org/10.1016/j.addma.2021.101993.
54. Lettori J, Esposto C, Peruzzini M, Pellicciari M, Raffaeli R. Geometrical characterization of circular multi-layered CMT WAAM specimens by 3d structured light scanning. Int J Adv Manuf Technol. 2025;136(11):5305–34. https://doi.org/10.1007/s00170-025-15107-8.
55. Pan J. In-situ visual reconstruction of strut profiles in pulsed wire arc additive manufacturing of lattice structures. Virtual Phys Prototyp. 2024;19(1):e2425822. https://doi.org/10.1080/17452759.2024.2425822.
56. Tang S, Wang G, Zhang H. In situ 3d monitoring and control of geometric signatures in wire and arc additive manufacturing. Surf Topogr. 2019;7(2):025013. https://doi.org/10.1088/2051-672X/ab1c98.
57. Errico V, Campanelli SL, Angelastro A, Dassisti M, Mazzarisi M, Bonserio C. Coaxial monitoring of AISI 316l thin walls fabricated by direct metal laser deposition. Materials. 2021;14(3):1–17. https://doi.org/10.3390/ma14030673.
58. Baier D, Weckenmann T, Wolf F, Wimmer A, Zaeh MF. Underlying methodology for a thermal process monitoring system for wire and arc additive manufacturing. J Manuf Mater Process. 2022;7(1):10. https://doi.org/10.3390/jmmp7010010.
59. Wu B, Ding D, Pan Z, Cuiuri D, Li H, Han J, et al. Effects of heat accumulation on the arc characteristics and metal transfer behavior in wire arc additive manufacturing of Ti6Al4V. J Mater Process Technol. 2017;250:304–12. https://doi.org/10.1016/j.jmatprotec.2017.07.037.
60. Hauser T, Silva AD, Reisch RT, Volpp J, Kamps T, Kaplan AFH. Fluctuation effects in wire arc additive manufacturing of aluminium analysed by high-speed imaging. J Manuf Process. 2020;56:1088–98. https://doi.org/10.1016/j.jmapro.2020.05.030.
61. Srivastava S, Garg RK, Sachdeva A, Sharma VS. A multi-tier layer-wise thermal management study for long-scale wire-arc additive manufacturing. J Mater Process Technol. 2022;306:117651. https://doi.org/10.1016/j.jmatprotec.2022.117651.
62. A Hosseini V, Högström M, Hurtig K, Valiente Bermejo MA, Stridh L-E, Karlsson L. Wire-arc additive manufacturing of a duplex stainless steel: thermal cycle analysis and microstructure characterization. Weld World. 2019;63(4):975–87. https://doi.org/10.1007/s40194-019-00735-y.
63. Yildiz AS, Koc B, Yilmaz O. Thermal behavior determination for wire arc additive manufacturing process. Proc Manuf. 2021;54:233–7. https://doi.org/10.1016/j.promfg.2021.07.036.
64. Güler O, Dinçer E. Measuring and qualification on camera recordings in wire arc additive manufacturing monitoring. In: 2023 5th International Congress on human-computer interaction, optimization and robotic applications (HORA), pp. 1–2. IEEE, (2023). https://doi.org/10.1109/hora58378.2023.10156704.
65. Belhadj M, Werda S, Belhadj A, Kromer R, Darnis P. Thermal analysis of wire arc additive manufacturing process. ESAFORM 2021. 2021. https://doi.org/10.25518/esaform21.4095
66. Chen G, Sun Y, Wang C, Ding J, Suder W, Li Z, et al. Characterising and modelling plasma transferred arc for additive manufacturing. Int J Heat Mass Transf. 2025;241:126735. https://doi.org/10.1016/j.ijheatmasstransfer.2025.126735.
67. Yang Y, Zhou X, Li Q, Ayas C. A computationally efficient thermo-mechanical model for wire arc additive manufacturing. Addit Manuf. 2021;46:102090. https://doi.org/10.1016/j.addma.2021.102090.
68. Zani M, Montemurro M, Panettieri E, Marin P. B-spline based metamodel of the thermal analysis of the wire arc additive manufacturing process. Proc Des Soc. 2023;3:807–16. https://doi.org/10.1017/pds.2023.81.
69. He X, Wang T, Wu K, Liu H. Automatic defects detection and classification of low carbon steel WAAM products using improved remanence/magneto-optical imaging and cost-sensitive convolutional neural network. Measurement (Lond). 2021;173:108633. https://doi.org/10.1016/j.measurement.2020.108633.
70. Shin S-J, Hong S-H, Jadhav S, Kim DB. Detecting balling defects using multisource transfer learning in wire arc additive manufacturing. J Comput Des Eng. 2023;10(4):1423–42. https://doi.org/10.1093/jcde/qwad067.
71. Shin S-J, Lee J-H, Jadhav S, Kim DB. Material-adaptive anomaly detection using property-concatenated transfer learning in wire arc additive manufacturing. Int J Precis Eng Manuf. 2023;25(2):383–408. https://doi.org/10.1007/s12541-023-00924-2.
72. Rashid A, Vatandoust F, Kota A, Melkote SN. Bead geometry prediction in wire arc directed energy deposition using physics-informed machine learning and low-fidelity data. Addit Manuf. 2025;109:104881. https://doi.org/10.1016/j.addma.2025.104881.
73. Li Z, Kovachki N, Azizzadenesheli K, Liu B, Bhattacharya K, Stuart A, Anandkumar A. Fourier neural operator for parametric partial differential equations. arXiv .2020. https://doi.org/10.48550/ARXIV.2010.08895.
74. Li Z, Huang DZ, Liu B, Anandkumar A. Fourier neural operator with learned deformations for pdes on general geometries. 2023. http://jmlr.org/papers/v24/23-0064.html.
75. Liu M, Bai X, Xi S, Dong H, Li R, Zhang H, et al. Detection and quantitative evaluation of surface defects in wire and arc additive manufacturing based on 3d point cloud. Virtual Phys Prototyp. 2023. https://doi.org/10.1080/17452759.2023.2294336.
76. Hasan F, Hamrani A, Rayhan MM, Dolmetsch T, McDaniel D, Agarwal A. Precision calibration in wire-arc-directed energy deposition simulations using a machine-learning-based multi-fidelity model. J Manuf Mater Process. 2024;8(5):222. https://doi.org/10.3390/jmmp8050222.
77. Fernandez-Zelaia P, Dryepondt SN, Ziabari AK, Kirka MM. Self-supervised learning of spatiotemporal thermal signatures in additive manufacturing using reduced order physics models and transformers. Comput Mater Sci. 2024;232:112603. https://doi.org/10.1016/j.commatsci.2023.112603.
78. Pham TQ-D, Tran V-X. Sensitivity study of process parameters of wire arc additive manufacturing using probabilistic deep learning and uncertainty quantification. Concurr Eng. 2024;32(1–4):20–33. https://doi.org/10.1177/1063293x241277189.
79. Era IZ, Zhou F, Raihan AS, Ahmed I, Abul-Haj A, Craig J, et al. In-situ melt pool characterization via thermal imaging for defect detection in directed energy deposition using vision transformers. J Manuf Process. 2025;145:11–21. https://doi.org/10.1016/j.jmapro.2025.03.123.
80. Yu R, He S, Yang D, Zhang X, Tan X, Xing Y, et al. Identification of cladding layer offset using infrared temperature measurement and deep learning for WAAM. Opt Laser Technol. 2024;170:110243. https://doi.org/10.1016/j.optlastec.2023.110243.
81. Chabot A, Rauch M, Hascoët J-Y. Towards a multi-sensor monitoring methodology for am metallic processes. Weld World. 2019;63(3):759–69. https://doi.org/10.1007/s40194-019-00705-4.
82. Shamsaei N, Yadollahi A, Bian L, Thompson SM. An overview of direct laser deposition for additive manufacturing; part ii: mechanical behavior, process parameter optimization and control. Addit Manuf. 2015;8:12–35. https://doi.org/10.1016/j.addma.2015.07.002.

83. Zhang H, Vallabh CKP, Zhao X. Registration and fusion of large-scale melt pool temperature and morphology monitoring data demonstrated for surface topography prediction in LPBF. Addit Manuf. 2022;58:103075. https://doi.org/10.1016/j.addma.2022.103075.
84. Ren H, Zhang Y, Li H, Long Y. Video interpolation-based multi-source data fusion method for laser processing melt pool. Appl Sci. 2025;15(9):4850. https://doi.org/10.3390/app15094850.
85. Yang X, Li R, Hu K, Sun A, Ma X, Liu M, et al. Intelligent control of additively manufactured beads on the basis of target morphologies and thermal field. Front Mech Eng. 2024. https://doi.org/10.1007/s11465-024-0813-3.
86. Yu R, Tan X, He S, Huang Y, Wang L, Peng Y, et al. Monitoring of robot trajectory deviation based on multimodal fusion perception in WAAM process. Measurement. 2024;224:113933. https://doi.org/10.1016/j.measurement.2023.113933.
87. Yang S, Li H, Tang Z, Liu W, Wang F, Peng S. Data fusion and swin transformer architecture-based model for monitoring geometric morphology of blue laser directed energy deposited Al-Mg alloy. Measurement. 2025;248:116992. https://doi.org/10.1016/j.measurement.2025.116992.
88. Reisch RT, Hauser T, Franke J, Heinrich F, Theodorou K, Kamps T, Knoll A. Nozzle-to-work distance measurement and control in wire arc additive manufacturing. In: 2021 2nd European symposium on software engineering. ESSE 2021, pp. 163–172. ACM, 2021. https://doi.org/10.1145/3501774.3501798.
89. Pringle AM, Oberloier S, Petsiuk AL, Sanders PG, Pearce JM. Open source arc analyzer: multi-sensor monitoring of wire arc additive manufacturing. HardwareX. 2020;8:00137. https://doi.org/10.1016/j.ohx.2020.e00137.
90. Baltrusaitis T, Ahuja C, Morency L-P. Multimodal machine learning: a survey and taxonomy. IEEE Trans Pattern Anal Mach Intell. 2019;41(2):423–43. https://doi.org/10.1109/tpami.2018.2798607.
91. Tianying H, Shengfu Y, Yi L, Wuxin Y. Design of multi-arc collaborative additive manufacturing system and forming performance research. J Manuf Process. 2024;109:313–25. https://doi.org/10.1016/j.jmapro.2023.12.018.
92. Papacharalampopoulos A, Foteinopoulos P, Karagianni OM, Stavropoulos P. A human-AI collaborative framework for additive manufacturing modeling and decision-making. Processes. 2025;13(12):3877. https://doi.org/10.3390/pr13123877.
93. Baldauf M, Lohrer P, Hauser T, Jauer L, Schleifenbaum JH. Camera-based measurement and control of the contact tip to work distance in wire arc additive manufacturing. Progress Addit Manuf. 2024;9(3):565–74. https://doi.org/10.1007/s40964-024-00655-4.
94. Mu H, Polden J, Li Y, He F, Xia C, Pan Z. Layer-by-layer model-based adaptive control for wire arc additive manufacturing of thin-wall structures. J Intell Manuf. 2022;33(4):1165–80. https://doi.org/10.1007/s10845-022-01920-5.
95. Kumar A, Sadhya S, Khan AU, Madhukar YK. Improvement in process efficiency of WAAM-TIG by in-situ voltage–current–temperature monitoring and feedback control system. Int J Precis Eng Manuf. 2025;26(7):1673–81. https://doi.org/10.1007/s12541-024-01208-z.
96. Xia C, Pan Z, Zhang S, Polden J, Wang L, Li H, et al. Model predictive control of layer width in wire arc additive manufacturing. J Manuf Process. 2020;58:179–86. https://doi.org/10.1016/j.jmapro.2020.07.060.
97. Li F, Feng X, Wang D. Research on temperature control in additive manufacturing of thermosetting carbon fiber composites using DMC-IEPO-PID algorithm. J Phys: Conf Ser. 2024;2902(1):012013. https://doi.org/10.1088/1742-6596/2902/1/012013.
98. Masenlle M, Elso J, Ostolaza JX. Multivariable QFT control of the direction flip problem in wire arc additive manufacturing. IET Control Theory Appl. 2025. https://doi.org/10.1049/cth2.12765.
99. Marcotte J, Mishra S, Wen JT. Combining layer-to-layer and in-layer feedback control for wire arc additive manufacturing. In: 2025 IEEE 21st international conference on automation science and engineering (CASE), pp. 1505–1510. 2025. https://doi.org/10.1109/case58245.2025.11164035.