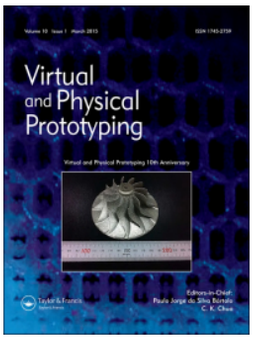

Virtual and Physical Prototyping

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/nvpp20

Multimodal deep learning for enhanced temperature prediction with uncertainty quantification in directed energy deposition (DED) process

Adrian Matias Chung Baek, Taehwan Kim, Minkyu Seong, Seungjae Lee, Hogyeong Kang, Eunju Park, Im Doo Jung & Namhun Kim

To cite this article: Adrian Matias Chung Baek, Taehwan Kim, Minkyu Seong, Seungjae Lee, Hogyeong Kang, Eunju Park, Im Doo Jung & Namhun Kim (2025) Multimodal deep learning for enhanced temperature prediction with uncertainty quantification in directed energy deposition (DED) process, Virtual and Physical Prototyping, 20:1, e2474532, DOI: 10.1080/17452759.2025.2474532

To link to this article: https://doi.org/10.1080/17452759.2025.2474532

VIRTUAL AND PHYSICAL PROTOTYPING
2025, VOL. 20, NO. 1, e2474532 (30 pages)
https://doi.org/10.1080/17452759.2025.2474532

Taylor & Francis
Taylor & Francis Group

Multimodal deep learning for enhanced temperature prediction with uncertainty quantification in directed energy deposition (DED) process

Adrian Matias Chung Baek, Taehwan Kim, Minkyu Seong, Seungjae Lee, Hogyeong Kang, Eunju Park, Im Doo Jung and Namhun Kim

Department of Mechanical Engineering, Ulsan National Institute of Science and Technology (UNIST), Ulsan, Republic of Korea

ABSTRACT

The accurate and reliable prediction of temperature history is crucial in meeting the ever-increasing demands for part quality and process reliability in metal additive manufacturing (AM). While many recent studies based on deep learning approaches have shown promise, they are subject to major limitations: inadequate handling of deposition strategy and insufficient consideration of uncertainty, both of which impact the prediction model performance. This work proposes a novel multimodal deep learning approach for temperature prediction with uncertainty quantification in directed energy deposition (DED) process. The proposed methodology implements multimodal data fusion, combining reproduced grayscale images of deposition strategy with numerical process variables, including process parameters, geometrical features, and printing process status. Furthermore, a novel approach for direct uncertainty estimation is introduced, inspired by object detection methods in computer vision. Through extensive comparative analyses, the proposed method outperforms conventional deterministic and probabilistic deep learning approaches, as well as state-of-the-art methods, in both temperature prediction and confidence interval estimation. An ablation study further validates the effectiveness of the proposed architecture, and quality inference based on predicted temperature distributions proves the feasibility and applicability of the proposed method for advancing process and quality control in metal AM processes.

ARTICLE HISTORY
Received 24 December 2024
Accepted 26 February 2025

KEYWORDS
Metal additive manufacturing; deep learning; multimodal data fusion; uncertainty quantification; temperature prediction; directed energy deposition

Nomenclature

AM	additive manufacturing
DED	directed energy deposition
LPBF	laser powder bed fusion
WAAM	wire-arc additive manufacturing
D10, D50, D90	characterised particle sizes
z	latent space representation
μ	mean
σ	standard deviation
ϵ	standard normal noise sampled from N(0,1)
λ	balancing parameter
MSE	mean squared error
$smooth_{L_1}$	smooth L1 loss
Δt	time change
LP	laser power
d_e	Euclidean distance between the point of interest and the heat source
d_s^1, d_s^2, d_s^3, d_s^4	Euclidean distance between the heat source and cooling surfaces
L	layer index
LP_s	laser state
t	total printing time
t_p	laser on time per layer
t_l^{intra}	intralayer dwell time
t_l^{inter}	interlayer dwell time
l	dwell index
D	dwell sequence
SOTA	state-of-the-art
MAE	mean absolute error
RMSE	root mean squared error
MAPE	mean absolute percentage error
R^2	coefficient of determination
CP	coverage probability
IoU	intersection over union
ML	machine learning
DL	deep learning
DT	decision trees
ANN	artificial neural network

CONTACT Namhun Kim nhkim@unist.ac.kr Department of Mechanical Engineering, Ulsan National Institute of Science and Technology (UNIST), 50 UNIST-gil, Ulju-gun, Ulsan 44919, Republic of Korea

Supplemental data for this article can be accessed online at https://doi.org/10.1080/17452759.2025.2474532.

FFNN	feedforward neural network
GRU	gated recurrent unit
RNN	recurrent neural network
DNN	deep neural network
GNN	graph neural network
PINN	physics-informed neural network
XGB	extreme gradient boosting
RF	random forest
GP	gaussian process
CNN	convolutional neural network
VAE	variational autoencoder
QRNN	quantile regression neural network
BNN	Bayesian neural network
GAN	generative adversarial network
Grad-CAM	gradient-weighted class activation map

1. Introduction

Metal additive manufacturing (AM) has transformed the manufacturing landscape by offering layer-by-layer production of complex, high-performing metal components using a variety of materials [1,2]. Laser Powder Bed Fusion (LPBF) and Directed Energy Deposition (DED) have established themselves as dominant metal AM processes; LPBF involves the selective melting of powder layers on a build plate under the controlled environment to produce near-net-shape components [3,4], whereas DED employs a high-density heat source (e.g. laser, electron beam, plasma arc) to process continuously deposited powder or wire feedstock along predetermined paths on a substrate [5]. Given their capability to produce functional components with tailored properties, metal AM processes have attracted substantial interest from both research communities and high-value industries, including automotive, aerospace, and nuclear sectors [6].

Despite its potential, the widespread industrial adoption of metal AM is hindered by inconsistencies in quality and performance. Employing high-energy heat sources during printing induces intense thermal conditions, predominantly manifested by steep thermal gradients and high cooling rates from repetitive melting and solidification cycles [7,8]. The directional nature of material deposition further intensifies thermal instability, profoundly influencing microstructural evolution [9,10]. These thermal phenomena often give rise to detrimental structural anomalies including defects, porosity, residual stress accumulation, and dimensional distortions, all of which collectively degrade the physical and mechanical properties of the manufactured components [11–13]. Therefore, a thorough understanding of the transient thermal behaviour and its intricate interactions with key process factors is of utmost importance for implementing effective process control and optimisation strategies that assure the production of robust, reliable, and defect-free components with superior performance [14,15].

Thermal modelling has been a central focus of research, with diverse approaches developed over the years [5,16–25]. Physics-based computational modelling, such as finite element method and computational fluid dynamics [26–30], are commonly used to simulate thermal behaviours with high accuracy and stability, grounded in established principles of thermal-fluid dynamics and heat transfer. Unfortunately, their massive computational requirements constrain their practicality for iterative and time-sensitive operations, such as process optimisation, real process control, and uncertainty quantification [16,18,19,22,31]. Analytical and semi-analytical methods offer a more computationally efficient alternative at the expense of prediction accuracy [32–35], as they rely on broader assumptions and simplifications compared to physics-based models [17,22,36]. This trade-off results in discrepancies between simulated results and actual thermal behaviours, diminishing their reliability. Notably, discrepancies persist even in high-fidelity physics-based models, underscoring the ongoing challenges in fully capturing the complexities and uncertainties of metal AM processes through simulation.

Data-driven approaches, particularly machine learning (ML) and deep learning (DL), have become a trending practice for thermal prediction due to their advantage in uncovering nonlinear interdependencies between input and output variables [5,16,18,20,37–44]. Initial research focused on simple ML and DL models. Paul et al. [38] implemented extreme randomised trees (ERT) and bagged decision trees (DT) to predict the thermal behaviours in the DED process. Lo et al. [39] developed a shallow neural network to estimate peak temperatures, melt pool depths, and track overlap for specific process parameters. Similarly, Stathatos and Vosniakos [40] constructed an artificial neural network (ANN) to model thermal evolution along custom long-track paths. Recent research has advanced toward sophisticated DL models to capture the complex sequential and spatial dynamics of temperature data. Mozaffar et al. [17] employed a gated recurrent unit (GRU) for the temperature modelling. Ren et al. [5] developed a hybrid solution combining recurrent neural network (RNN) and deep neural network (DNN) to predict temperature fields under arbitrary deposition patterns. Mozaffar et al. [16] also presented a geometry-agnostic thermal model using graph neural network (GNN) to enable the prediction

of temperature across different part geometries. More recently, there is a growing inclination towards leveraging physics-informed neural networks (PINNs), which integrate physics laws directly into the neural network architecture, to improve model accuracy and interpretability [19,23,24,45]. A summary of relevant literature is available in Table S1, included in the Supplementary Material section.

While prior works have yielded notable progress in temperature prediction, they fall short of adequately addressing a critical process factor, the deposition strategy, also known as the laser scanning strategy or toolpath. The deposition strategy defines the specific pattern and sequence of energy source movement, exerting significant influence on the temperature distributions, and thus serving as an important means of controlling the part quality [19,46,47]. However, it is either neglected or oversimplified, often into the Euclidean distance between the point of interest and the energy source's position. Although some have attempted to represent the deposition strategy into specialised data structure [5,22], they have yet to integrate other key process variables (e.g. laser power, and powder flow rate) when training the prediction models, thereby hampering their accuracy. The primary challenge stems from differences in data modality: unlike other process variables, which are represented as single numerical values, the deposition strategy involves the spatiotemporal evolution of position coordinates, making its integration with static process parameters particularly difficult.

In light of this issue, multimodal data fusion approaches offer a promising solution. The importance and the necessity of multimodal data fusion in AM has been increasingly emphasised [48–52], as evidenced by successful implementations of multimodal learning approaches in metal AM [49,53–71]. Perani et al. [64] proposed a DL architecture that processed coaxial melt pool images alongside scalar process parameters (e.g. laser power, laser status, speed command) to predict track geometries in Laser-DED. Using a combination of data collected from acoustic emission, accelerometer, and thermal sensor, Karthikeyan et al. [60] applied convolutional neural network (CNN) models to identify surface pore formation in printed stainless steel 316L parts. In the LPBF domain, Wu et al. [59] constructed a triple-sensor fusion monitoring system using an enhanced LeNet-5 deep CNN architecture to classify melt pool quality, transforming time-domain signals from sensors into two-dimensional (2D) grayscale images for improved feature capture. It is evident that, despite the keen interest in multimodal data fusion, there has been no emphasis on harnessing this approach to integrate deposition strategy with other process variables for accurate temperature prediction, highlighting the necessity to address this research gap.

On the other hand, uncertainty is another concern that must be addressed in the context of temperature prediction. Metal AM processes are inherently susceptible to multiple sources of uncertainty due to the complex physical phenomena occurring during the printing [72]. Even under controlled process conditions, such uncertainties considerably impact the process repeatability and reproducibility [73]. Given that high levels of uncertainty can result in large deviations from expected behaviour, assessing prediction confidence is beneficial. Indeed, several efforts have been made to quantify uncertainty in various aspects of metal AM, including melt pool depth [69], material properties (e.g. density [70], amorphicity [71]), mechanical [72,73] and thermal properties [74]. Likewise, quantifying uncertainty in temperature prediction is essential for establishing a reliable process and quality control systems [75]. For instance, Sideris et al. [21] presented a modified Gaussian process (GP) model for temperature prediction with uncertainty estimation in a wire-arc additive manufacturing (WAAM) process. However, the majority of uncertainty quantification approaches are constructed upon conventional probabilistic models, which are characterised by high computational costs and prolonged response times in comparison to deterministic models. These constraints render them infeasible for time-critical real-time applications and restrict their ubiquitous implementation. Consequently, an efficient approach for accurate and reliable temperature predictions with quantified uncertainties must be developed.

This study proposes a novel multimodal DL approach for enhanced temperature prediction with uncertainty quantification in metal AM process, specifically focusing on the DED process. The proposed methodology encompasses (1) a multimodal architecture that integrates 2D representations of deposition strategy with other key process variables and (2) an uncertainty estimation method inspired by object detection models. To validate the effectiveness of the proposed method, comprehensive experiments were conducted, in which the temperature predictions were compared against established DL benchmarks. The results demonstrate the superior performance of the proposed method in both temperature profile and confidence interval prediction accuracy. Additionally, a systematic feature importance analysis was performed to elucidate the relative influence of input variables on the temperature predictions. The effectiveness of the proposed architecture was further validated through an ablation study.

Finally, the practical utility of the proposed method was demonstrated through the successful application of predicted temperature distributions for quality assessment, establishing its potential for advancing process monitoring and quality control in metal AM processes.

2. Methodology

This section details the proposed methodology for enhancing temperature prediction and uncertainty estimation through data fusion between image-based deposition strategy and numerical process variables and defines the input and output of the prediction models. Of particular importance is the treatment of deposition strategy data, initially existing in the form of time-series implying the position of the energy source throughout the process. Rather than processing this data directly, which results in the loss of spatial information, an image representation is introduced to preserve both the temporal progression and spatial characteristics of the deposition state. In this work, the use of empirical data is proposed over synthetic data as a means of addressing the two prevalent uncertainties in AM processes: aleatoric and epistemic uncertainty. Aleatoric uncertainty pertains to inherent process variability that cannot be reduced, whereas epistemic uncertainty arises from incomplete knowledge of model inputs, parameters, or modelling assumptions [73].

2.1. Proposed architecture

The proposed multimodal DL model consists of three main components as shown in Figure 1: (1) a multimodal data fusion module, including an encoder of variational autoencoder (VAE) for extracting representative features from the deposition strategy image, (2) a direct uncertainty estimation module, and (3) a multi-task loss function.

2.1.1. Multimodal data fusion module

The proposed multimodal architecture requires careful design to effectively process both numerical and image data. While numerical data can be processed directly through conventional dense layers, applying similar processing to flattened images presents significant limitations in both performance and computational efficiency. Thus, convolutional-based approaches are preferred.

CNN is the most widely adopted convolutional-based approach for image processing. It is typically composed of convolutional layers, pooling layers, and activation functions arranged sequentially to learn features from the input image [74]. However, as the input dimension (i.e. the image size) increases, deeper architecture becomes necessary, leading to increased computational complexity and the risk of overfitting if not properly regularised [75]. Furthermore, CNN tends to learn discriminative features that do not fully represent the underlying image data distribution, resulting in poor generalisation to out-of-distribution samples [76].

Recently, deep convolutional generative models such as VAE and generative adversarial network (GAN), which excel at reconstructing images from learned latent distributions, have gained traction for demonstrating superior image feature extraction capabilities than traditional CNN. Unfortunately, GANs are known for their training instability and high computational demands, which limit their practicality for simple feature extraction tasks [77]. In contrast, VAE offers a more balanced approach with efficient representation learning, fewer model parameters, and faster inference times [78].

VAE excels at learning compact latent representations from images, capturing spatial and temporal patterns [79,80]. Its structure comprises an encoder and a decoder: the encoder compresses the input data into a latent space $z = \mu + \sigma^2 \cdot \epsilon$, represented as a probability distribution rather than a single deterministic point. μ, σ^2, ϵ denote the learned mean, variance, and a noise term that is sampled from a standard normal distribution, respectively. The decoder then reconstructs the original input by minimising the reconstruction loss and the regularisation loss (i.e. the Kullback-Leibler divergence) to ensure that the latent space remains both informative for reconstruction and structured for generating new samples [81].

Another advantage of VAE is that, once trained, the encoder can be decoupled and used independently for feature extraction without requiring further training [79,80,82,83]. This characteristic is particularly beneficial in the context of the proposed study, as it enables efficient processing of deposition strategy images: the feature extraction can be performed in advance using a pre-trained encoder, eliminating the need for image processing during the actual training of the temperature prediction and uncertainty quantification model. Accordingly, the proposed methodology employs a VAE to derive features from the reproduced deposition strategy images.

Following VAE training on deposition strategy images, the encoder is separated and integrated into the multimodal architecture for image processing during the prediction. The parameters of the encoder are frozen, preventing further training adjustments. The latent space representations z obtained from the encoder are concatenated with numerical input data. The combined input passes through multiple dense layers to perform the regression tasks: temperature prediction and uncertainty quantification.

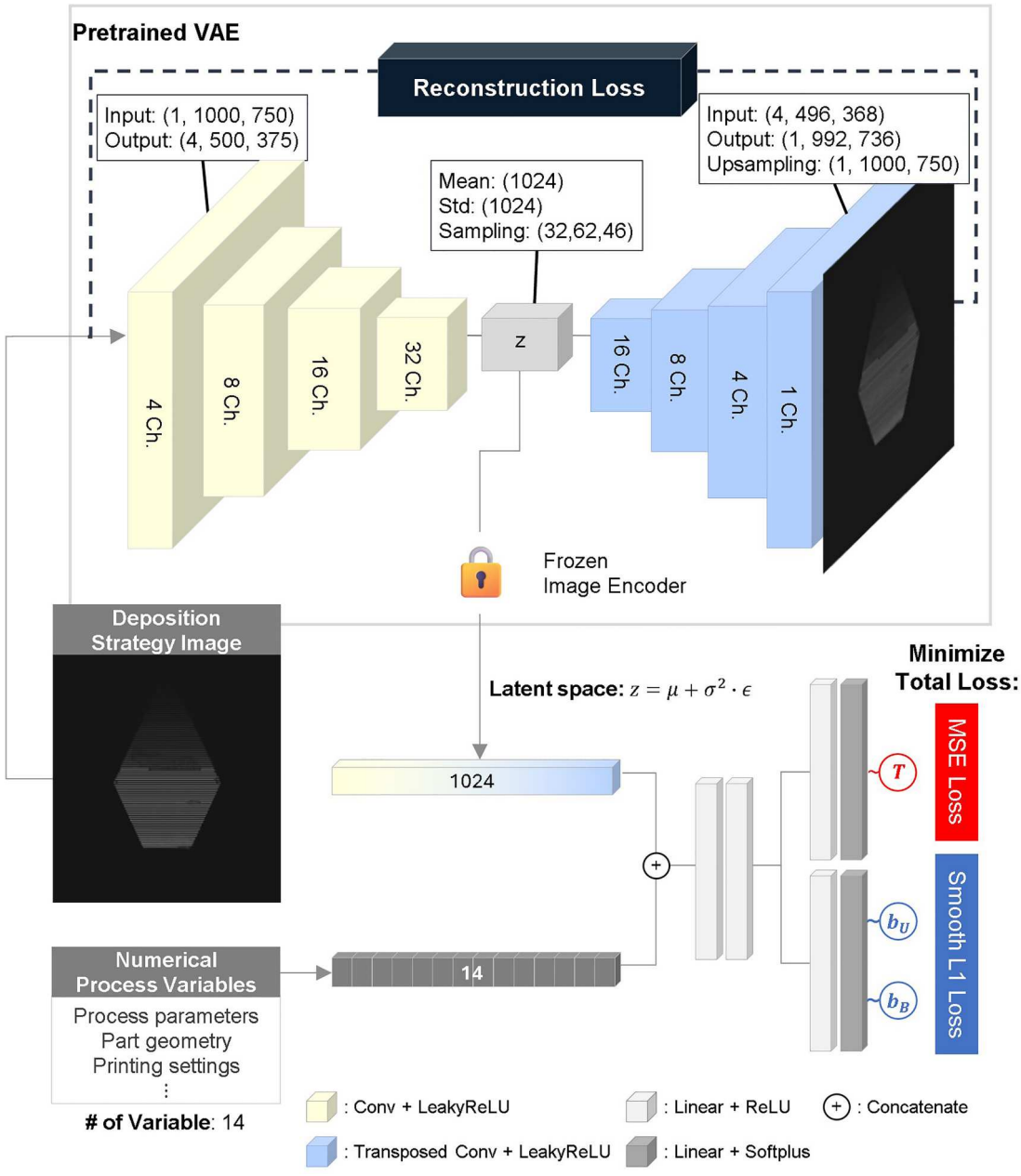

Figure 1. Schematic representation of the proposed multimodal DL architecture for temperature prediction with uncertainty quantification. The framework combines deposition strategy image data processed through a VAE encoder with numerical process variables.

2.1.2. Direct uncertainty estimation module

Probabilistic DL models are widely used for uncertainty quantification, as they estimate full probability distributions of predicted values (e.g. mean and variance) rather than single point predictions [84]. While theoretically sound, these approaches impose substantial computational burdens compared to deterministic approaches due to iterative sampling-based inference or Monte Carlo approximations. This is especially acute when processing multimodal data, where the inherently elevated computational requirements hinder real-time applications. Furthermore, probabilistic models often rely on distributional assumptions (e.g. Gaussian noise) or Bayesian priors [85], potentially introducing model uncertainty (i.e. epistemic uncertainty) that may misrepresent aleatoric uncertainty.

To address these challenges, a direct uncertainty estimation inspired by the concept of bounding box in object detection is proposed. In object detection, a bounding box, a rectangle surrounding an object defined by the x and y coordinates of its corners, is regressed alongside the object class labels [86]. In an analogous fashion, we apply this principle to temperature prediction by treating confidence intervals as 'bounding boxes' around the true temperature value. That is, the proposed method predicts upper and

lower temperature bounds directly, enabling efficient localisation without probabilistic sampling.

Direct prediction of upper and lower bounds is grounded by quantile regression, which estimates the conditional quantiles or percentiles without assuming parametric distributions [87]. In typical quantile regression, the asymmetry in the loss function drives the model to learn the assumed quantile functions rather than true interval labels, primarily because explicit interval labels are rarely available in practice. However, AM processes permit multiple experimental runs under identical settings, thus enabling direct observation of temperature variability and the derivation of ground truth bounds from experimental data.

The direct uncertainty estimation approach streamlines uncertainty quantification by treating it as a regression task. By doing so, the method avoids the computational burdens associated with probabilistic sampling and distributional assumptions in probabilistic models. Bounds defined experimentally inherently encode aleatoric uncertainty, such as the inherent randomness, sensor noise, and process stochasticity arising from uncontrollable factors in the AM process. By learning directly from these bounds, the model prioritises capturing data-driven variability rather than epistemic uncertainty, which is minimised through deterministic architecture design.

Figure 1 illustrates the branched structure of the network, separating the prediction tasks into distinct output layers – one dedicated to temperature prediction and another to confidence interval estimation, thereby reinforcing the conceptual link between the proposed approach and bounding box regression in object detection.

2.1.3. Multi-task loss function

In regression models, the mean squared error (MSE) loss function is commonly preferred due to its simplicity and effectiveness. However, its equal weighting for all predictions presents an undesired prediction in our context, as we prioritise the accurate prediction of temperature over capturing uncertainty. To address this, the multi-task loss function employed in object detection tasks is adapted [88]. The multi-task loss function is expressed as follows in Equation 1.

$$L(\{p_i^*\}, \{b_i^*\}) = L_{cls}(p_i^*, p_i) + \lambda p_i L_{reg}(b_i^*, b_i) \quad (1)$$

Where p_i^* is the predicted class probability, p_i is the ground truth class label, b_i^* is the predicted bounding box regression offsets, b_i is the ground truth bounding box regression offsets, and λ is the balancing parameter to weigh the different tasks. $L_{cls}(p_i^*, p_i)$ is the log loss function for the classification task and $L_{reg}(b_i^*, b_i)$, also expressed as $L_{loc}(b_i^*, b_i)$, is the localisation loss for the regression task to predict the coordinates of the bounding boxes. As the proposed method does not involve a classification task but instead focuses on two separate regression tasks – predicting the temperature and estimating the confidence interval – $L_{cls}(p_i^*, p_i)$ term is replaced with $L_{reg}(y_i^*, y_i)$, a MSE loss function to measure the error between the predicted temperature y_i^* and true temperature y_i.

Additionally, the term $p_i L_{reg}(b_i^*, b_i)$ is expressed as $L_{loc}(c_i^*, c_i)$ where c_i^* and c_i stands for predicted and ground truth value of the upper and lower bound of the confidence interval. For the localisation loss $L_{loc}(c_i^*, c_i)$, Smooth L1 loss is preferred because it balances robustness and sensitivity to outliers [89]. Specifically, the Smooth L1 loss function adaptively switches between L2 and L1 losses according to the error magnitude. This enables rapid convergence for small errors while preventing gradient explosion for large errors. As the proposed uncertainty estimation approach can be sensitive to initial errors, implementing Smooth L1 loss can enhance robustness against outliers while maintaining appropriate predictive sensitivity. Smooth L1 loss is defined as follows:

$$smooth_{L_1}(x) = \begin{cases} 0.5x^2 \ if \ |x| < 1 \\ |x| - 0.5 \ otherwise \end{cases} \quad (2)$$

Thus, $L_{loc}(c_i^*, c_i)$ is expressed as Equation (3):

$$L_{loc}(c_i^*, c_i) = smooth_{L_1}(c_i - c_i^*) \quad (3)$$

Combining Smooth L1 loss with the balancing parameter λ to the MSE loss, we obtain a multi-task loss function that places greater emphasis on temperature prediction than uncertainty estimation. The complete form of the proposed multi-task loss is given by

$$L(\{y_i^*\}, \{c_i^*\}) = L_{reg}(y_i^*, y_i) + \lambda L_{loc}(c_i^*, c_i) \quad (4)$$

In this study, the balancing parameter λ was empirically set to 20.

2.2. Definition of input and output variables

As a supervised learning task, the proposed method maps input parameters to temperature predictions, where the inputs include both deposition strategy image and numerical process variables, and the output corresponds to the temperature at specified temporal instances.

2.2.1. Representation of deposition strategy

A representation of the deposition strategy that captures spatial and temporal information of the heat source is generated for accurate temperature prediction, motivated by academic pursuits [5,22,90]. This representation incorporates details such as the deposition sequence at a given time, prior depositions, deposition patterns, and component geometry, relying on G-code instruction obtained from slicing software that specify the laser head movement along X, Y, and Z axes and laser status. Figure 2 illustrates the step-by-step procedure of generating the deposition strategy representation.

As depicted in Figure 2(a), a matrix is initially constructed to match the dimensions of the substrate ($L_x \times L_y$). To retain fine details of the deposition strategy, the dimensions are scaled by a factor of 100. Initially, the matrix contains elements with zero value by default, representing the initial state of the substrate surface before any laser processing. Each element in the matrix corresponds to a pixel in the final image. The laser deposition sequence for a specific layer at time t is then recorded by updating the matrix elements. As the laser traverses across the substrate, the matrix elements at the laser's coordinates are assigned a time index. The index advances by $\Delta t = 1$ s in our study and resets at the completion of each layer, clearly distinguishing between successive layer depositions. During laser idle periods (i.e. dwell time), the time index continues to increment, maintaining temporal continuity even in absence of spatial movement. After the deposition path recording, the matrix undergoes normalisation to a grayscale range of 0 (black) to 255 (white), thus the pixel intensity correlates with the

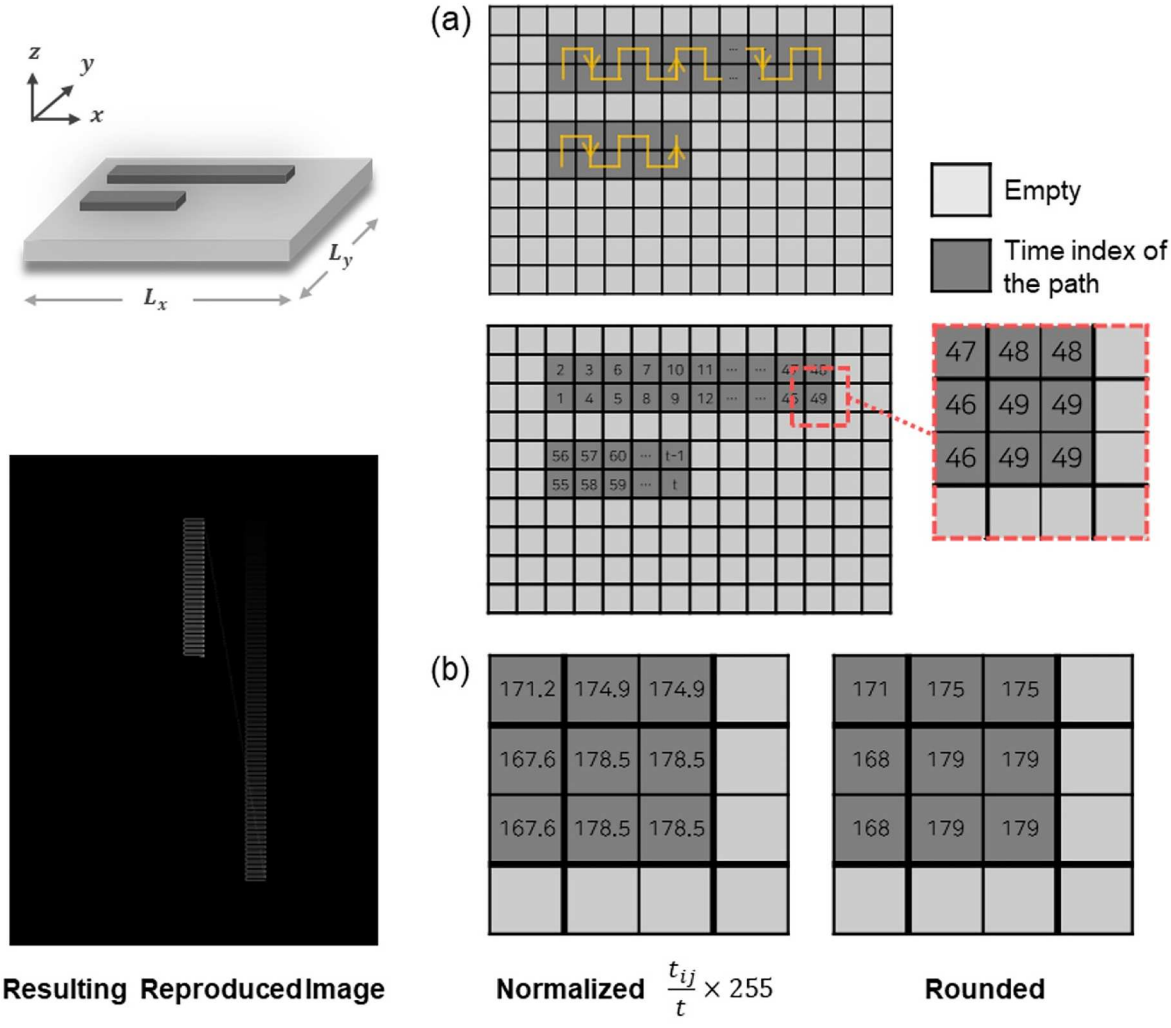

Figure 2. The illustration of step-by-step procedure for generating the deposition strategy image: (a) A 2D matrix records the time index for each portion of the deposition path at each time step. Note that the number of updated elements depends on both the nozzle speed and the chosen matrix resolution; the example of four element updates shown here is for demonstration only. (b) Normalisation and rounding of the matrix elements.

temporal sequence of laser-substrate interactions, with more recent interactions appearing brighter in the image. The normalised values are then rounded to the nearest integer to ensure temporal mapping of each pixel. The resulting grayscale image effectively represents the cross-section view of the deposition strategy at a specific time t, as shown in Figure 2.

2.2.2. Selection of numerical input data

In accordance with the extant literature, the numerical inputs related to process parameters, deposition strategy, part geometry, and process status information have been selected for the purpose of temperature prediction.

Laser power *LP* represents a fundamental process parameter with a direct influence on thermal behaviour. Due to its dominant impact on temperature evolution, laser power has been consistently identified as an indispensable variable [17–21,43,91]. In place of the detailed deposition strategy, the shortest Euclidean distance between the point of interest and the heat source d_e is frequently chosen to approximate the effect of the deposition path on temperature [16,17,20,21,37,43,91]. Geometric considerations are captured through the following variables: the layer index *L*, which reflects the printing stage and cumulative thermal effects [16,17,21,91], and the Euclidean distances between the heat source and cooling surfaces d_s^1, d_s^2, d_s^3, d_s^4, which account for thermal dissipation [16,17,37,43]. In this study, d_s^1, d_s^2, d_s^3, d_s^4 measure the distances (in millimetres) from the heat source to each of the four edges of the substrate. d_s^1 and d_s^2 correspond the distance to the left and right edges, while d_s^3 and d_s^4 denote the distance to the top and bottom edges, respectively, as illustrated in Figure S1.

The remaining variables pertain to the printing status. One such variable is the laser state LP_s, where the laser is represented as 1 when active and 0 when inactive [17]. Temporal variables include the total printing time *t* [17,37,43,91] and the time the laser was active per layer t_p [20,37]. The variable t_p resets at the start of each new layer. These variables capture the temporal dynamics of the laser, allowing for a detailed characterisation of the thermal flow throughout the manufacturing.

The significance of dwell-related variables has become increasingly recognised in temperature modelling, reflecting their fundamental influence on thermal dynamics. The intralayer dwell time t_l^{intra} is the dwell time within a layer and the interlayer dwell time t_l^{inter} is the dwell time between a layer. Both variables are reset at the beginning of each new layer [18].

Table 1. Details of the numerical input data.

Category	Variables	Expression
Process parameter	Laser power (W)	LP
Deposition strategy	Euclidean distance between the point of interest and the heat source	d_e
Part geometry	Euclidean distance between the heat source and cooling surfaces	d_s^1, d_s^2, d_s^3, d_s^4
	Layer index	L
Process status	Laser state (on or off)	LP_s
	Dwell index	I
	Total printing time (global time) (s)	t
	Laser on time per layer (local time) (s)	t_p
	Intralayer dwell time (local time) (s)	t_l^{intra}
	Interlayer dwell time (local time) (s)	t_l^{inter}
	Dwell sequence	D

Two other dwell-related variables are introduced that have not been previously considered: Dwell index *I*, which tracks the global occurrence number of dwell events throughout the printing process, and dwell sequence *D*, which monitors the sequential order within each dwell event. These new variables are designed to capture the temporal evolution of cooling patterns, as shown in Figure S1.

In aggregate, the numerical inputs encompass 14 process variables that provide a comprehensive characterisation of the AM process, as summarised in Table 1.

2.2.3. Ground truth for temperature and confidence bounds

The target variables include not only the temperature for a specific point at a global time *t*, but also its associated confidence bounds, upper and lower bound, for uncertainty quantification. In addition, experimental data is utilised rather than simulation-based synthetic data to minimise epistemic uncertainty. Most research mentioned in Table S1 is prone to epistemic uncertainty as they rely on synthetic data for model training, inheriting the same simulation-to-reality gap that plagues traditional physics-based models. This gap between simulated and actual data significantly hinders the generalizability and robustness of these models, limiting their potential for efficient and reliable in-situ process monitoring and control.

The temperature at a specific point of interest of the printed part or the substrate at a global time t, which is the primary focus, can be measured using various thermal sensors (e.g. thermocouples, pyrometers, or infrared (IR) cameras) [92]. Ground truth values of the upper and lower bound for the confidence interval are derived from experimental replicates, multiple experimental runs with the same factor settings (levels). This approach of finding the ground truth of the confidence interval, to the best of the authors' knowledge, has not been previously documented in the literature.

Three explicit assumptions underlying the proposed method are established [10]:

(1) The temperature measurements at a single point obtained from replicates or multiple experimental runs under identical conditions represent the population samples and, therefore, characterise the overall population behaviour in terms of temperature variation.
(2) These temperature measurements at time t are independent and identically distributed (i.i.d.) instances.
(3) These temperature measurements at time t from replicates follow a Gaussian (normal) distribution.

Based on these assumptions, the upper and lower bounds are computed using the mean μ and standard deviation σ derived from the temperature measurements of the replicates at time t. For a 95% confidence interval, the lower and upper bounds correspond to $\mu \pm 1.96\sigma$, respectively, implying a high probability that the actual temperature value at the given point and time falls within this range.

Finally, each set of input features (i.e. deposition strategy image data and numerical data) is paired with corresponding target data, creating a comprehensive dataset for training and testing the proposed multimodal DL method.

3. Experiments

3.1. Experimental setup

Experiments were conducted on a powder-fed laser DED system, MX 600 (Insstek Co, Korea), equipped with a 1 kW Ytterbium fibre laser with a beam diameter of 0.8 mm and a build volume of 450×600×380 mm. High purity argon gas (99.999%, grade 5.0) served as both a powder carrier gas and as a shielding gas that prevents oxidation. The deposited material was a commercially available gas-atomized stainless steel 316L powder (KoreaPowder Co., Korea) with a particle size distribution of 45–150 μm, characterised by 3.1% of D10, 94.9% of D50, and 2% of D90. Table S2 provides details on the chemical composition of the powder. Stainless steel 304L plates with dimensions of 100×75×10 mm^3 were used as the substrate. The initial working distance between the powder nozzle and the substrate was set at 9 mm. Layers were deposited vertically with a consistent layer thickness of 0.25 mm, while the deposition head moved in the x-y plane with a hatch spacing of 4 mm between adjacent tracks to create the 2D geometry of each layer as dictated in the pre-programmed G-code.

Following process parameters were held constant across all experiments: coaxial gas rate at 2.5 L/min, carrier gas rate at 4.0 L/min, powder flow rate at 3.50 ± 0.1 g/min, and nozzle speed at 0.85 m/min. By fixing these parameters, any observed variations in the temperature would be a function of scanning strategy and heat input, thereby allowing the evaluation of the proposed method when predicting the process temperature based on corresponding variables. The specific values for the parameters were carefully chosen based on the authors' prior research [93,94], seeking to maximise the relative density of the deposited material while ensuring superior mechanical properties.

For the temperature measurement, 3 K-type thermocouples (Omega 5TC-GG-K-36-72, rated up to 482°C) insulated with high-temperature glass-fibre, with a diameter of 0.127 mm and an accuracy of $\pm 0.4^{\circ}C$ were attached to the bottom surface of the substrate at locations designated as TC1, TC2, and TC3 in Figure 3(a), following the methodology described by Li et al [93]. Thermal profiles were recorded at a sampling rate of 1 Hz using a Keysight 34970A data acquisition logger with the 34970A plug-in module connected to a laptop computer. To mitigate measurement errors and noise caused by the contact between the bed platform and its movement, the substrate was elevated by 5 mm and restrained along the x- and y-axes using jigs. Although thermocouples were utilised for temperature data collection in this study, alternative sensors such as thermal cameras or pyrometers can be used. Figure 3 presents a detailed representation of the experimental setup.

3.2. Data description

To build a comprehensive dataset containing diverse substrate temperature profiles for training the DL models, a series of experiments were conducted using various combinations of deposition strategies and laser power levels. The deposition strategies included bidirectional patterns, characterised by continuous parallel tracks in the longitudinal and transverse directions, and a spiral pattern in the outward direction, as illustrated in Figure 4(a). These strategies, referred to as zigzag X, zigzag Y, and spiral out in this study, are the most widely adopted in the DED process [16,95–97]. The laser power was varied from 250 W to 450 W in increments of 50 W. Furthermore, six distinct geometries in Figure 4(b) and (c) were selected, varying in size, number of layers, shapes, and geometric features. Note that the difference in part geometries introduces additional variations in deposition paths. Each combination was replicated three times for deriving the

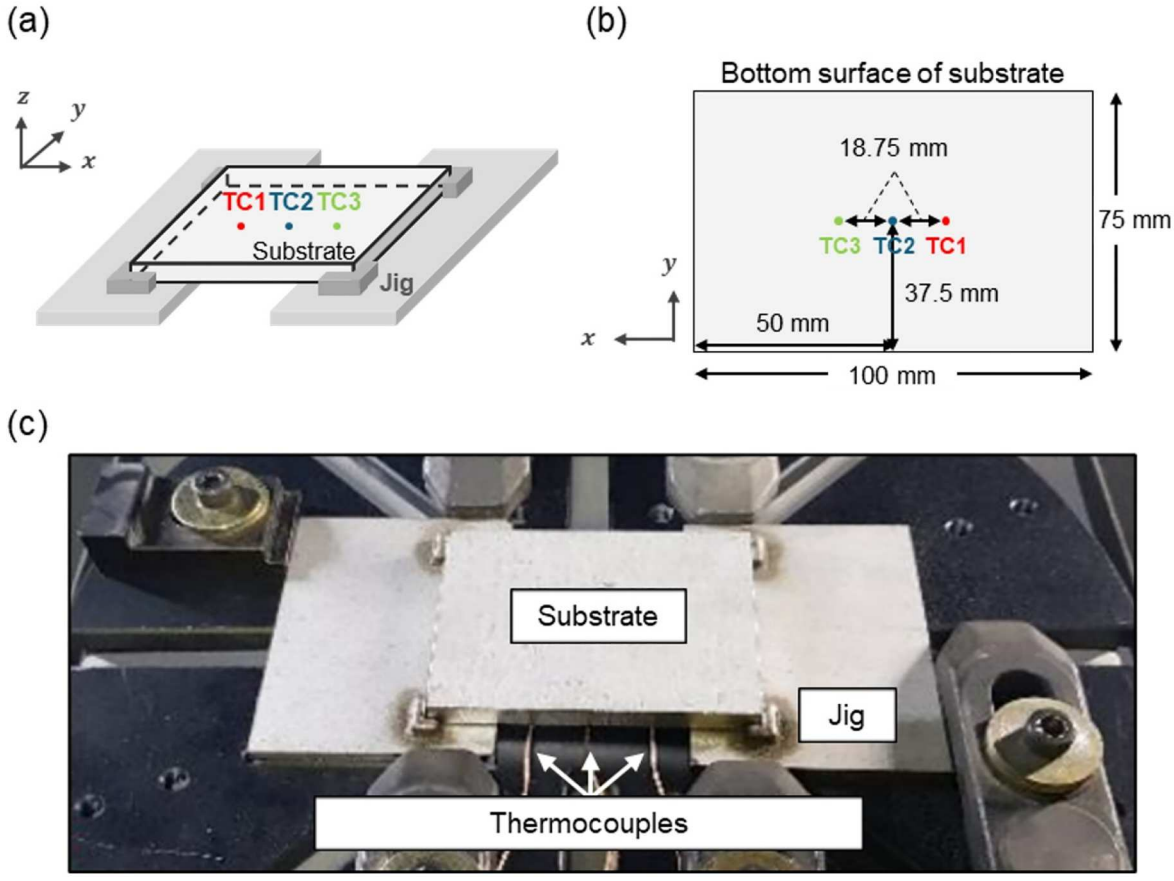

Figure 3. (a) An illustration depicting the substrate secured on jig designed to restrict motion and provide elevation, (b) The precise positions (TC1, TC2, TC3) of the thermocouples attached on the bottom surface for data acquisition, (c) The actual experimental setup in the DED process.

ground truths of the upper and lower bound, yielding a total of 84 printings across 28 combinations. Figure S2 visualises exemplary temperature measurements, highlighting the inherent uncertainty in process temperature, even under identical process settings and measurement locations.

Temperature prediction was performed at specific time instances; thus, the collected temperatures at time t, along with the deposition strategy image data and numerical data, constitute a data sample in the dataset. To ensure independence between the training and test datasets, geometries 1–4 were used for training and validation and the remaining two geometries were reserved for testing. The training dataset was further split into training and validation subsets in an 80:20 ratio, and the test dataset was subdivided into 5 cases according to the combinations used. The dataset resulted in 997,713 samples for training and 140,526 samples for testing. Table 2 provides the combinations used for data collection, corresponding geometries, and dataset details.

A min–max normalisation approach, defined in Equation 5, was applied to scale the input feature set between 0 and 1, with the exception of the image data, to prevent any potential bias in model training that may arise due to scale differences of the features.

$$x_{normalized} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{5}$$

Where x is the actual value, x_{min} denotes the minimum value of the variable within the dataset, and x_{max} denotes the maximum value of the variable.

3.3. Evaluation metrics

To assess the performance of the developed models outputting continuous scalar values, common regression metrics were computed, including root mean squared error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and coefficient of determination (R^2). MAE quantifies the average absolute difference between the predicted and observed values, RMSE is the standard deviation of the residuals, and MAPE is the prediction errors in a percentage, providing a scale-independent measure that enables performance comparisons across different datasets. For all three

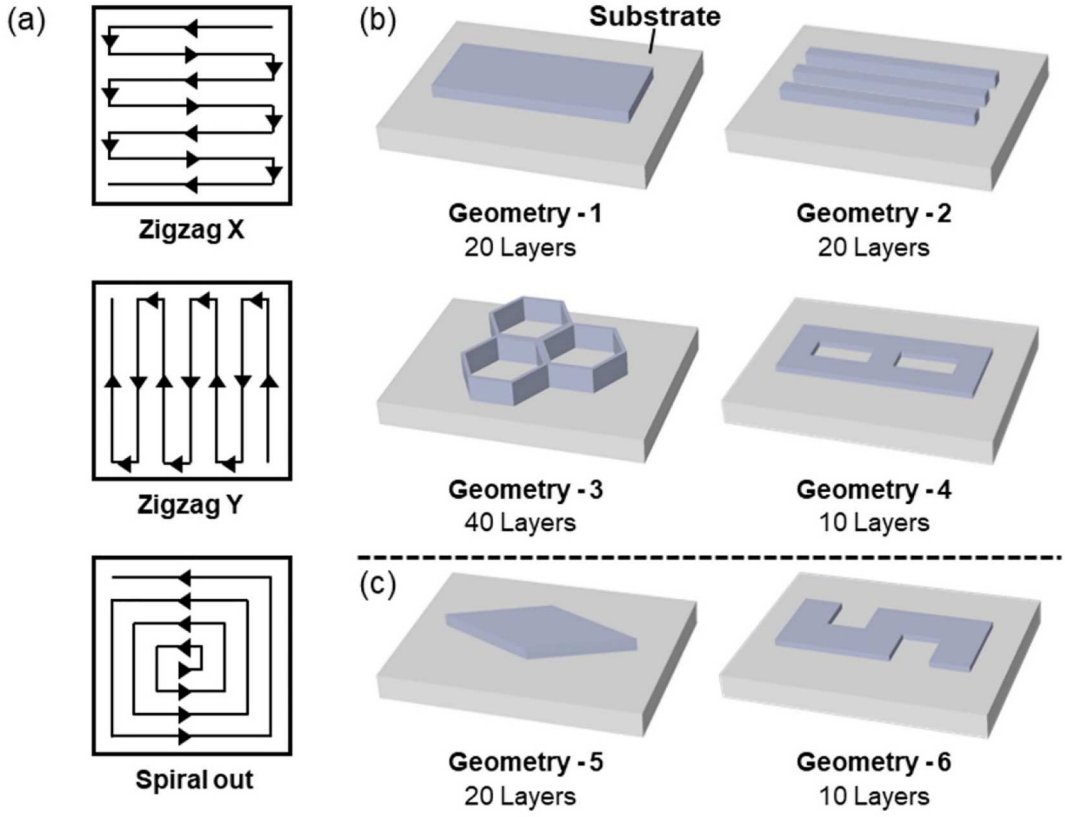

Figure 4. (a) Representation of the selected deposition strategies: zigzag X, zigzag Y, spiral out, (b) Geometries corresponding to the training dataset, (c) Geometries corresponding to the test dataset.

Table 2. A summary of the datasets along with their parameter configurations.

Dataset	Case	Geometry	Deposition strategies	Laser power (W)	Replicate	Data sample	Data sample in dataset
Train & Validation	–	1	Zigzag X & Y, Spiral out	250, 300, 350, 400, 450	3	188,559	997,713
		2	Zigzag X & Y	250, 450		36,852	
		3	Spiral out	250, 350, 450		98,955	
		4	Spiral out	350		8205	
Test	1	5	Zigzag X	250	3	9597	140,526
	2		Zigzag Y	450		9594	
	3		Zigzag X	250		9723	
	4		Zigzag Y	450		9723	
	5	6	Spiral out	350		8205	

metrics, lower values closer to 0 indicate better model performance. In contrast, R^2 evaluates the proportion of variance in the dependent variable that is predictable from the independent variables, and a score of 1 indicates a perfect correlation. The expressions for these metrics are given from Equation 6–9.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(t_i - \hat{t}_i)^2} \tag{6}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{N}|t_i - \hat{t}_i| \tag{7}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|t_i - \hat{t}_i|}{|t_i|} \times 100 \tag{8}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(t_i - \hat{t}_i)^2}{\sum_{i=1}^{n}(t_i - t_{mean})^2} \tag{9}$$

Where n is the number of samples, $\hat{t}_i$ the predicted temperature, and t_i the ground truth temperature.

To quantitatively evaluate the confidence intervals, coverage probability (*CP*) is chosen, a conventional metric that indicates the likelihood of a constructed confidence interval containing the true value [98,99].

CP is defined as:

$$CP = \frac{1}{n}\sum_{i=1}^{n} C_i \times 100 \tag{10}$$

n represents the total number of samples, C_i is a binary indicator that equal 1 if true value falls within the upper and lower bound, and 0 otherwise. However, *CP* alone can be misleading since overly wide confidence intervals can inflate the probability. Hence, a novel evaluation approach by adapting the Intersection over Union (*IoU*) metric from object detection applications [100]. This metric, not previously used for confidence interval evaluation, quantifies the degree of overlap between the predicted and actual confidence intervals, producing a normalised score between 0 and 1, where 1 represents perfect alignment.

$$IoU = \frac{Overlap}{Union} = \frac{Length(A \cap B)}{Length(A) + Length(B) - Overlap} \tag{11}$$

3.4. Hyperparameter tuning and implementation details

The performance of DL models depends on the appropriate fine tuning of hyperparameters. Given the exponential growth in possible configurations with an increasing number of hyperparameters, we adopted a random search strategy to systematically navigate the hyperparameter space [101]. Table S3 lists the hyperparameters investigated in this work, along with their respective search ranges. The model was trained for a maximum of 150 epochs, with early stopping employed to prevent overfitting. The optimal hyperparameter configuration, determined through tuning, is presented in Table 3. The models were implemented using the PyTorch library in Python 3.9.16 version. All training and testing were performed using Intel® Core i7-12700F CPU with 48 GB RAM and an NVIDIA GeForce RTX 3080 Ti.

4. Results and discussions

In this section, the effectiveness of the proposed multimodal DL model for temperature prediction with uncertainty quantification was demonstrated using real-experimental data. The performance was compared against existing DL approaches, including standard neural network, convolutional network, and state-of-the-art (SOTA) multimodal methods. Comparison with existing DL-based probabilistic methods was also made. To ensure fairness, all comparing models were trained following the same procedure as the proposed method. The prediction performance was evaluated across the five cases in the test datasets. Furthermore, a sensitivity analysis was performed to examine the influence of input features on the prediction. An ablation study was undertaken to validate the efficacy of the proposed architecture. Lastly, predicted temperature distributions were analyzed to establish their relationship with quality performance.

4.1. Performance comparison with existing methods

The temperature prediction performance of the proposed method is listed in Table 4. As shown in Table 4, the proposed method achieved strong prediction accuracy with an RMSE of 19.9069, an MAE of 16.0029, a MAPE of 7.1527, and an R2 of 0.9131. The RMSE value indicates an average prediction error of approximately

Table 3. The detailed architecture of the proposed method.

Model			Layer	Kernel size	Activation	Output size
VAE	Encoder	1	Conv2D	4 × 4	LeakyReLU	4 × 500 × 375
		2	Conv2D	4 × 4	LeakyReLU	8 × 250 × 188
		3	Conv2D	4 × 4	LeakyReLU	16 × 125 × 94
		4	Conv2D	4 × 4	LeakyReLU	32 × 62 ×46
	z sampling	Mean	Linear			1024
		Std	Linear			1024
		z ~ q(z|x)	Linear			32 × 62 ×46
	Decoder	1	Transposed Conv2D	4 × 4	LeakyReLU	16 × 124 ×92
		2	Transposed Conv2D	4 × 4	LeakyReLU	8 × 248 ×184
		3	Transposed Conv2D	4 × 4	LeakyReLU	4 × 496 ×368
		4	Transposed Conv2D	4 × 4	LeakyReLU, Sigmoid	1 × 992 ×736
		5	Upsampling	–	–	1 × 1000 ×750
Multimodal architecture	Input concatenation	Concatenate	VAE output + Numerical		–	1038
		1	Linear		ReLU	1024
		2	Linear		ReLU	1024
	Temperature prediction	1	Linear		ReLU	1024
		2	Linear		Softplus	1
	Confidence intervalestimation	1	Linear		ReLU	1024
		2	Linear		Softplus	2

Table 4. Temperature prediction performance comparison with existing deterministic models and SOTA multimodal methods. The best result is highlighted in bold.

Model	Input	Metric			
		RMSE (↓)	MAE(↓)	MAPE(↓)	R^2(↑)
ANN	Numerical	24.8833	20.1473	9.1372	0.8641
ResNet18	Image	67.5573	51.7360	28.9373	−0.0017
EfficientNet	Image	68.0951	52.3924	28.5333	−0.0177
ANN	Combined	23.7021	19.1028	8.5102	0.8768
ResNet		21.5835	17.4681	7.4460	0.8978
Perani et al. [64]		21.5020	17.0768	7.6375	0.8985
Khusheef et al. [62]		23.1072	18.2327	9.4663	0.8828
Huang [104]		22.0235	17.6038	7.6757	0.8935
Proposed		**19.9069**	**16.0029**	**7.1527**	**0.9131**

20°C, while the higher RMSE relative to MAE suggests the presence of outliers, as RMSE penalises larger errors more heavily. The low MAPE confirms that errors remain small relative to the actual values, and the high R2 indicates a strong fit, meaning that the model can explain a substantial portion of the variance in the data.

The prediction errors were analyzed through multiple residual visualisation techniques as shown in Figure 5. According to the histogram and density plot in Figure 5(a), residuals predominantly fall within −50°C and +50°C and approximate a normal distribution with a slight negative skew. The Q-Q plot in Figure 5(b)

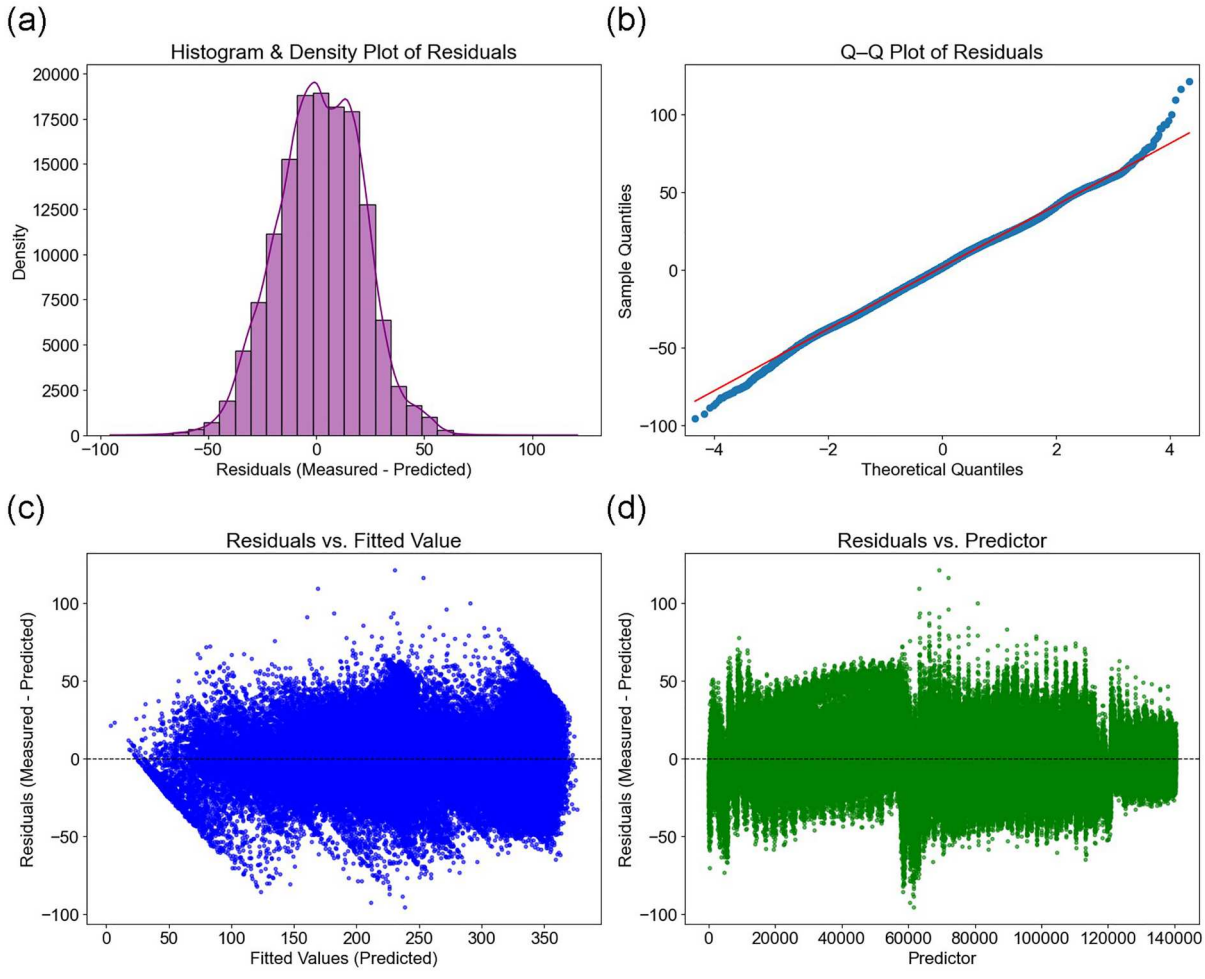

Figure 5. Residual analysis for the proposed method, including (a) a histogram and density plot, (b) Q–Q plot, (c) residuals vs. fitted plot, and (d) residuals vs. predictor plots. The results suggest approximate normality of errors and a lack of strong nonlinear patterns or heteroskedasticity.

shows that the residuals closely follow the diagonal line well in the central region, with minor deviations at the tails, particularly in the upper range, indicating slight non-normality in extreme values. No distinctive patterns were observed in the residuals vs. fitted and residuals vs. predictor plots (Figures 5(c) and (d)), suggesting the absence of significant nonlinear relationships or heteroskedasticity in the predictions. As the results do not violate the assumptions of linearity, normality, and homoscedasticity, the predictions of the proposed method can be considered statistically reliable.

The prediction performance of the proposed method was compared with existing deterministic DL models, including ANN and CNNs such as ResNet18 and EfficientNet-b0 [102]. These models underwent training using three distinct input configurations: numerical input data alone, reproduced deposition strategy image data alone, and a combination of both. For ANN training with combined data, the reproduced image was flattened and concatenated with numerical input data prior to introduction into the first dense layer. For ResNet training with combined data, each numerical input feature was transformed into individual channels matching the reproduced image dimensions (1000×750 pixels) and stacked to the corresponding image, producing a 15-channel input image. An optimised version of ResNet18, featuring reduced kernel sizes and basic blocks, was implemented to mitigate overfitting, as the standard ResNet architecture's depth is unsuitable for processing the grayscale image. The above methods for handling combined data represent alternative multimodal fusion approaches documented in the literature [103], but they were not selected as the primary methodology in this paper.

The proposed method outperformed both unimodal and multimodal benchmarks in temperature prediction. Notable improvements were observed compared to the unimodal ANN trained solely on numerical input data. Specifically, the proposed method achieved improvements of 20.00% in RMSE, 20.57% in MAE, 21.72% in MAPE, and 5.67% in R^2. Even when ANN model was trained again with combined data, the performance gains (i.e. 4.75% in RMSE, 5.18% in MAE, 6.86% in MAPE, and 1.47% in R^2) remained below those achieved by the proposed method. When compared with ResNet trained on combined input data, the improvements were 7.77% in RMSE, 8.39% in MAE, 3.94% in MAPE, and 1.70% in R^2. ResNet, when trained exclusively on deposition strategy images, produced predictions that deviated significantly from ground truth temperature, underperforming relative to other models. The underperformance of EfficientNet-b0 indicates that the underperformance of ResNet model was not due to insufficient training or improper hyperparameter tuning. Rather, the results suggest that while incorporating deposition strategies enhances temperature prediction capabilities, relying only on the image data is insufficient for precise temperature prediction, as it overlooks critical variables such as printing time, laser power, and others.

To further validate the effectiveness of the proposed methods, the uncertainty estimation results were evaluated using the four regression metrics by assessing the upper and lower bound of the confidence intervals. Additionally, CP and IoU, as detailed in Section 3.3, were employed to provide a single-value performance evaluation measure. Table 5 shows that the proposed method achieved the lowest RMSE, MAE, MAPE, and highest R^2 in predicting the lower bound. It also demonstrated competitive performance in upper bound predictions; however, ResNet trained with combined input performed better in this aspect. This outcome can be attributed to prioritising temperature prediction – the primary objective – by incorporating a balancing parameter $\boldsymbol{\lambda}$ into the training loss, as discussed in Section 2.1.3. These seemingly contradictory results complicated

Table 5. Uncertainty quantification performance comparison with existing deterministic models and SOTA multimodal methods. The best result is highlighted in bold.

		Confidence interval									
		Upper bound				Lower bound					
Model	Input	RMSE (↓)	MAE (↓)	MAPE (↓)	R^2 (↑)	RMSE (↓)	MAE (↓)	MAPE (↓)	R^2 (↑)	CP (↑)	IoU (↑)
ANN	Numerical	27.8504	21.6558	8.8592	0.8597	34.8883	29.5536	13.4207	0.6688	74.8168	0.3426
ResNet18	Image	74.4816	58.1609	30.3771	−0.0037	61.9222	48.3858	28.6177	−0.0432	37.6485	0.1869
EfficientNet	Image	74.3439	57.8446	29.6757	−0.00003	63.0986	50.0596	28.7418	−0.0832	38.0634	0.1822
ANN	Combined	24.2589	18.6872	7.8539	0.8936	29.9259	24.7732	11.3175	0.7566	73.7771	0.3736
ResNet		**20.6155**	**16.2467**	**6.5643**	**0.9231**	26.8480	22.2127	9.9640	0.8039	73.6632	0.3953
Perani et al. [64]		21.4878	16.6921	7.1050	0.9165	28.2409	22.3310	10.2432	0.7830	79.1633	0.4139
Khusheef et al. [62]		23.3926	18.0367	8.9932	0.9010	27.1051	22.2595	11.6133	0.8001	75.9148	0.4035
Huang [104]		57.4788	52.1047	18.5540	0.4022	42.9111	35.5732	15.5404	0.4990	70.5414	0.3996
Proposed		23.1049	18.4099	7.5480	0.9035	**24.9158**	**20.1947**	**9.2953**	**0.8312**	**79.7429**	**0.4156**

the validation of the proposed method, underscoring the need for unified evaluation metrics such as CP and IoU. When evaluated using these metrics, the proposed method surpassed all other unimodal and multimodal models.

A performance comparison was carried out with several SOTA multimodal methods published in peer-reviewed journals. The benchmarks include a CNN-DNN-based method [64], a hybrid CNN-LSTM-based method [62], and an image-tabular contrastive learning-based method [104], each demonstrating proven capability in processing both image and numerical data within their respective domains. The results for temperature prediction and uncertainty quantification are presented in Tables 4 and 5, respectively. In temperature prediction, the proposed method significantly exceeded existing SOTA methods across all evaluation metrics. In uncertainty quantification, the results revealed more nuanced performance characteristics. While the proposed method excelled in the lower bound predictions, the CNN-DNN-based method [64] achieved better results in upper bound predictions. However, the proposed method attained higher CP and IoU scores compared to [64] Meanwhile, the image-tabular contrastive learning-based method [104], despite its sophisticated architecture, exhibited poor performance in estimating uncertainty, likely due to overfitting. These comparisons, supported by quantitative metrics, demonstrate the effectiveness of the proposed method for both temperature prediction and uncertainty quantification.

As the proposed method addresses uncertainty quantification, comparisons with conventional non-parametric and parametric probabilistic methods were also conducted, namely quantile regression neural network (QRNN) and Bayesian neural network (BNN). QRNN is a neural network trained using the pinball loss function for quantile regression to capture the median as well as other conditional quantiles τ [87,105], with τ set to {0.025, 0.5, 0.975} in this experiment to obtain a 95% confidence interval. BNN is the most adopted method for probabilistic prediction tasks, in which introduces Bayesian inference into neural network to model uncertainty in predictions and provides probabilistic outputs [99]. According to Table 6, our proposed method exceeded QRNN across most metrics, with the exception of CP value. Quantile regression methods are susceptible to overestimating the uncertainty levels in data [99,106]. QRNN demonstrated this tendency by producing wide confidence intervals, which, while increasing the likelihood of containing the true value as mentioned in Section 3.3, resulted in diminished performance across other evaluation metrics.

Table 6. Prediction performances comparison with existing probabilistic models. The best result is highlighted in bold.

	Temperature				Confidence interval								
					Upper bound				Lower bound				
Model	RMSE (↓)	MAE (↓)	MAPE (↓)	($\uparrow$)	RMSE (↓)	MAE (↓)	MAPE (↓)	R^2 (↑)	RMSE (↓)	MAE (↓)	MAPE (↓)	R^2 (↑)	CP (↑)	IoU (↑).
QRNN	23.5959	18.9923	8.4453	0.8778	41.2028	33.0521	15.001	0.6928	39.4785	33.2821	14.8888	0.5760	**96.5658**	0.3378
BNN	27.1979	21.8071	10.3012	0.8377	37.9439	31.6045	12.7006	0.7395	**24.0711**	**19.1088**	10.1211	**0.8424**	10.7283	0.0803
Proposed	**19.9069**	**16.0029**	**7.1527**	**0.9131**	**23.1049**	**18.4099**	**7.5480**	**0.9035**	24.9158	20.1947	**9.2953**	0.8312	79.7429	**0.4156**

When compared with BNN, the proposed method made better predictions of the upper bound, while BNN performed better for the lower bounds. However, BNN's notably low IoU value reveals limited alignment between predicted and actual intervals. Although the proposed method did not uniformly yield the highest performance over all interval predictions, it demonstrated markedly superior accuracy in the temperature prediction.

Finally, the computational efficiency of the models was assessed by computing the average inference time using a batch size of 1. As presented in Table 7, the proposed method achieved the third-fastest inference time among all models and a comparable number of parameters with EfficientNet. ANN recorded the fastest time and lowest number of trainable parameters, owing to its simplicity in handling only 14 numerical input variables. It is seen that considering the image data led to a drastic increase in both inference time and trainable parameters, reflecting the additional computational complexity required for image processing [107]. EfficientNet exhibited a lower number of trainable parameters, attributable to its architecture being specifically designed for balancing model size, accuracy, and computational cost [102]. In the case of the SOTA CNN-DNN-based multimodal method [64] it showed a short inference time and a relatively small number of trainable parameters, as it processes a smaller input image ($1\times224\times224$) and employs a CNN model comprising only 4 convolutional layers for the image feature extraction. BNN incurred the highest latency with an inference time of 497 ms, attributed to its sampling and iterations to compute posterior distribution-based probabilistic output [99]. While QRDNN showed lower inference times than BNN, its latency remained twice that of the proposed method and required a large number of parameters. These results prove the unsuitability of probabilistic models for time-sensitive iterative processes.

Table 7. Computational efficiency comparison of mentioned models.

Model	Input	Number of trainable parameters	Inference time (ms)
ANN	Numerical	19,977	0.0500
ResNet18	Image	11,171,779	4.2998
EfficientNet	Image	4,010,815	7.8100
ANN	Combined	44,952,787	3.6546
ResNet		771,190,787	2.1268
Perani et al. [64]		641,238	0.2642
Khusheef et al. [62]		75,096,695	3.1447
Huang [104]		13,612,007	2.0460
QRDNN		770,121,731	3.6188
BNN		192,541,698	497.0926
Proposed		**4,744,195**	**1.6998**

In general, the proposed method exhibited the most favourable performance across the majority of metrics and commendable computational efficiency, suggesting that handling the deposition strategy image in the proposed multimodal fusion manner markedly improves the predictions.

4.2. Analysis and interpretation of input variables

To investigate the influence of the input variables, particularly the deposition strategy image on the predictive modelling process, a comparative analysis was conducted between the proposed method and ANN trained with only numerical input data for all test cases detailed in Table 2. The proposed method consistently delivered superior accuracy in temperature prediction across all cases as shown in Table 8, maintaining MAPE value below 10% and achieving average R^2 values exceeding 0.862. It is noteworthy that both models demonstrated robust prediction capabilities, reflecting a strong correlation between the selected input variables and the temperature. In Case 1, both models exhibited comparatively lower performances, pointing to potential shortcomings in the dataset (e.g. a lack of representativeness or insufficient variability). This becomes particularly evident when compared with Case 2, which differed only in LP (450W). This comparison emphasises that the variability in laser power is crucial for capturing temperature profiles in the AM process, thus indicating the dataset enrichment by considering a wider range of LP settings.

In terms of confidence intervals, the proposed method surpassed ANN in both upper and lower bounds across most cases. However, some exceptions were noted in the upper bounds of Case 3 and 5, though the differences were marginal. Again, the conflicting results stem from the balancing parameter λ in the proposed loss function, which places greater emphasis on the prediction of the temperature. The balancing parameter λ plays a pivotal role in model performance and requires careful calibration to align with the desired prediction objectives. The CP values for Cases 2 and 4 were comparatively lower, primarily due to the limitation of the metric, as previously mentioned.

The enhanced temperature prediction performance is clearly demonstrated in the regression plots in Figure 6, where the predictions from the proposed method better align with the ideal diagonal line, representing perfect accuracy. In contrast, the predictions of ANN display wider dispersion from the diagonal line, especially in temperature ranges below $120^{\circ}C$ and above $330^{\circ}C$.

Table 8. Performance comparison between ANN trained with numerical data and the proposed method per test case. Higher performance in each test case is highlighted in bold.

		Temperature				Confidence interval									
						Upper bound				Lower bound					
Model	Test Case	RMSE (↓)	MAE (↓)	MAPE (↓)	R^2 (↑)	RMSE (↓)	MAE (↓)	MAPE (↓)	R^2 (↑)	RMSE (↓)	MAE (↓)	MAPE (↓)	R^2 (↑)	CP (↑)	IoU (↑)
ANN	1	31.1060	27.8307	11.9908	0.5834	20.5495	16.4502	6.8790	0.8504	45.3976	42.9460	19.5070	−0.1028	49.1036	0.1704
Proposed		**22.2597**	**19.0321**	**8.3405**	**0.7859**	**14.9371**	**11.9434**	**5.2642**	**0.9207**	**32.9241**	**30.5041**	**13.7957**	**0.4176**	**76.1525**	**0.2983**
ANN	2	20.4231	16.4158	6.8388	0.8770	28.6574	23.5214	8.3420	0.7915	39.4762	34.0143	13.1334	0.4693	**93.3536**	0.4031
Proposed		**18.1180**	**14.4883**	**5.5030**	**0.9029**	**22.1694**	**18.3837**	**6.2410**	**0.8748**	**28.0929**	**22.9561**	**8.6372**	**0.7302**	86.5142	**0.4589**
ANN	3	20.8810	17.0986	9.2034	0.7867	**22.7944**	**19.1502**	**9.3845**	**0.7660**	29.3836	24.6555	13.3292	0.5412	76.6000	0.3414
Proposed		**17.8058**	**14.6137**	**8.2112**	**0.8454**	25.8243	21.9447	10.6194	0.7006	**23.5505**	**19.7667**	**10.8896**	**0.7063**	**79.7292**	**0.359**
ANN	4	29.6442	23.7974	10.6361	0.7793	41.2884	34.3715	13.5051	0.6638	32.0054	26.3039	11.8059	0.6485	**85.6414**	0.4188
Proposed		**25.2568**	**20.8923**	**8.5151**	**0.8405**	**31.7907**	**26.4649**	**9.7486**	**0.8014**	**21.2675**	**16.4662**	**7.7210**	**0.8456**	77.5174	**0.4755**
ANN	5	19.1310	14.9700	6.8780	0.8645	16.5298	**12.4992**	**5.5506**	0.9145	22.4538	18.5731	8.9592	0.7730	66.4583	0.3763
Proposed		**13.2336**	**10.2544**	**5.1422**	**0.9343**	**15.5790**	12.5508	5.8371	**0.9229**	**12.1669**	**9.7347**	**5.0920**	**0.9325**	**77.9350**	**0.4891**

These trends are consistent across all examined cases. For instance, Figure 7 illustrates the temperature histories at TC3 for Case 1. Here, the predictions of the proposed method during the early printing time (0–300 s) closely match the measured temperatures, outperforming the ANN in accuracy. In addition, Figures 8 and 9 show that the temperature histories predicted by the proposed method exhibit a greater resemblance to the ground truth temperature. Upon the detailed examination of Case 2, depicted in Figure 8, the proposed method better captures local maxima as well as local minima corresponding to the dwell times. Complete predictions of the proposed method for all test cases are presented in Figures S3 to S7.

Figure 9 provides a visual representation of the confidence intervals for Case 1 generated by both models. The confidence intervals of the proposed method more accurately enclose the actual temperature measurements from the three replicates, compared to those generated by ANN. This suggests that the introduction of a deposition strategy image provides valuable insights that enhance the model's ability to capture the variability and uncertainty in the temperature. However, during the initial printing stages, the confidence intervals are narrower, reflecting minimal variations in the actual temperature measurements. As the printing process progresses, these intervals widen, reflecting the increasing uncertainties in the DED process, such as changes in material properties, heat accumulation, or variations in the cooling rate during component build-up [91].

A significant observation from comparing the ground truth and predicted results in Figures 7–9 is the pronounced fluctuations in the temperature predictions generated by the proposed method. These fluctuations stem from the image data, which contain detailed spatial aspects of the deposition state at each time step. As the model is architected to effectively process the image, it demonstrates heightened sensitivity in capturing subtle cross-sectional geometric variations that influence heat distribution and transient thermal behaviour. Nevertheless, the method maintains robust performance metrics, indicating that the fluctuations do not substantially compromise prediction accuracy. Moreover, enhanced sensitivity can offer advantages for real-time process monitoring and optimisation.

The Pearson correlation coefficients were computed to quantify the strength of the direction of potential linear relationships between the numerical input variables. The correlation matrix depicted in Figure 10 reveals valuable insights into these interrelationships. Most variables exhibited weak correlations; however, several notable exceptions were observed. For instance,

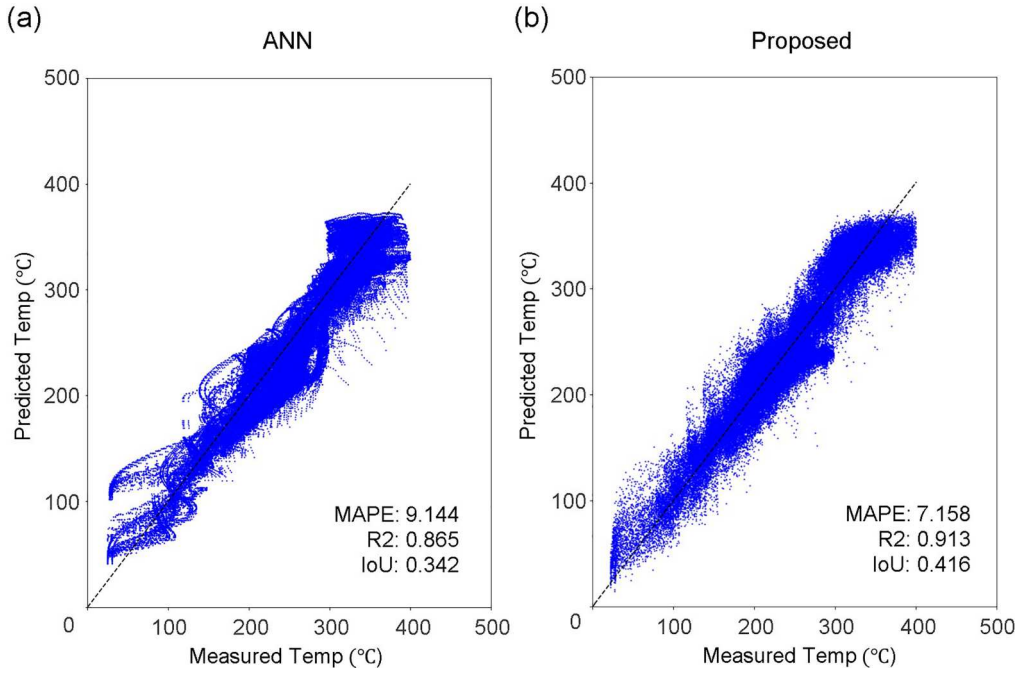

Figure 6. Measured temperature vs. predicted temperature of (a) ANN trained with numerical input data only and (b) proposed method.

the Euclidean distance between the heat source and cooling surfaces d_s^1, d_s^2 and d_s^3, d_s^4 had perfect negative correlations, which is expected since these distance represent the opposing measures relative to the substrate surfaces as illustrated in Figure S1. The dwell sequence *D* demonstrated a strong negative correlation (−0.90) with laser state LP_s, reflecting the fact that dwell periods occur when the laser is inactive. The layer index *L* showed strong positive correlations with both dwell index *I* (0.68) and the total printing time *t* (0.71), indicating that these variables increase concurrently during the printing process. For similar reasons, dwell index *I* had a positive correlation with intralayer dwell time t_l^{intra} (0.58). Figure 10 also provides a comprehensive view of the relationships between the numerical input variables and the process temperature. In particular, temperature was found to have moderate positive correlations with laser power *LP* (0.57), *t*(0.44), interlayer dwell time t_l^{inter}(0.41), and *L*(0.32).

The influences of all input, including the image, were also identified through sensitivity analysis using the variance-based Sobol decomposition method [91,108], which measures the individual contribution of each input variable to the variance in the temperature and confidence bounds. As observed in Figure 11, time *t* and laser power *LP* emerged as the primary contributors to the prediction variance. This insight is consistent not only with the previous correlation results, but also with theoretical expectations, as these variables directly affect the temperature dynamics. In contrast, the deposition strategy played a minor role in inducing variance, corroborating the earlier analysis in the previous section: the deposition strategy alone is insufficient for accurate temperature predictions, also evidenced by the poor performance of the convolutional models trained with deposition strategy image.

However, this does not imply that the deposition strategy image is ineffective. The gradient-weighted class activation map (Grad-CAM) was applied to the VAE encoder to visualise the contribution of the image data. Grad-CAM is a technique used to interpret CNN models by generating a localisation map that highlights the regions of an image most influential for the prediction. Regions of high importance are marked in red, while less relevant areas are shaded in blue [109,110]. For the deposition strategy images, the Grad-CAM visualisation reveals that the critical regions are concentrated around the most recent deposition paths, indicating that the proposed method actively utilises the temporal and spatial information from the image for temperature prediction, as illustrated in Figure 12.

As such, the improvements in temperature prediction and uncertainty estimation achieved arise from the enriched spatial and temporal information provided into the thermal modelling through incorporating deposition sequence details. The reproduced images offer a dynamic perspective of the printing process, delivering insights beyond what numerical input data alone can

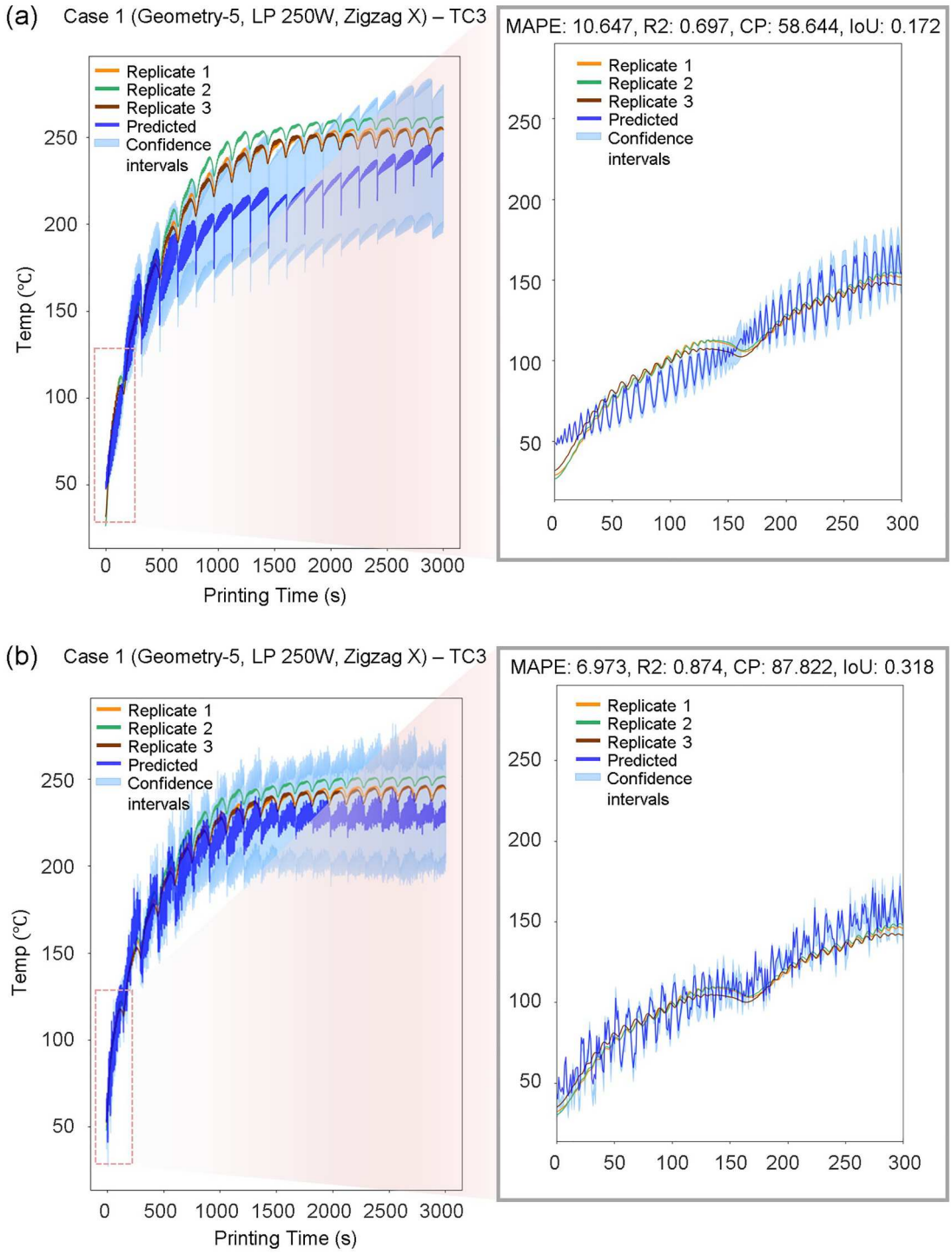

Figure 7. A detailed analysis of the early printing period (0–300 s) to compare the predictions of (a) ANN trained with numerical input data only and (b) the proposed method for Case 1 at TC3.

provide. Spatially, these images represent the distribution of heat source throughout the printing process and their interaction with the deposited material, which are critical for understanding thermal distribution. Temporally, they contain the heat evolution including the duration and frequency of heat application, which influences heat accumulation and the subsequent cooling rates, helping the understanding of temperature changes and resulting thermal gradients.

4.3. Ablation study

An ablation study was conducted to systematically evaluate the contributions of the two key components

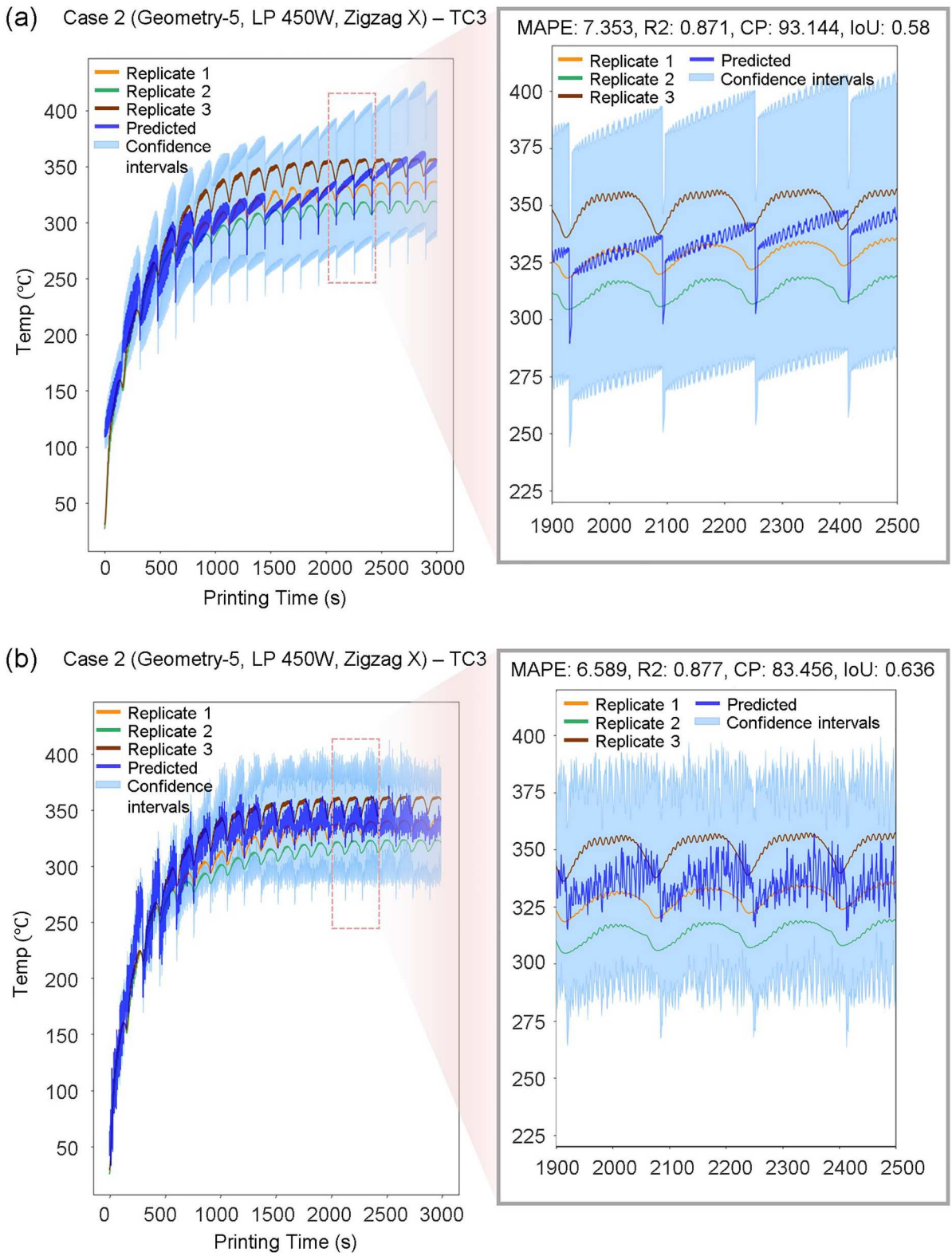

Figure 8. A detailed analysis between 1900–2500 s to compare the predictions of (a) ANN trained with numerical input data only and (b) the proposed method for Case 2 at TC3.

in the proposed method – specifically, the VAE encoder for image feature extraction and the custom multi-task loss function – to both temperature prediction and uncertainty quantification, as summarised in Tables 9 and 10. The baseline model is an ANN with architecture similar to that described in Section 4.1, differing in the

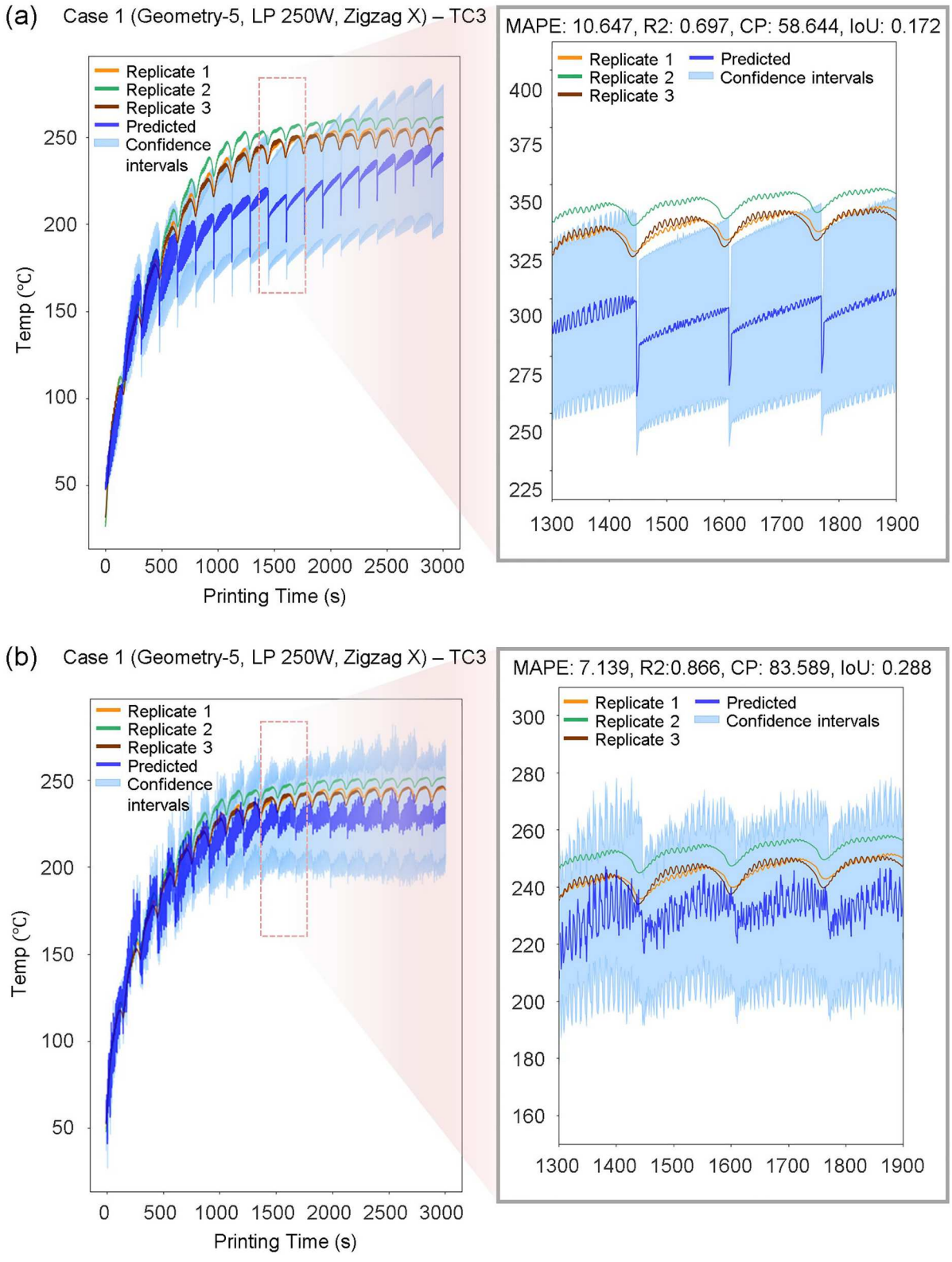

Figure 9. A detailed analysis between 1300–1900 s to compare the predictions of (a) ANN trained with numerical input data only and (b) the proposed method for Case 1 at TC3.

last two layers, which were split to handle temperature prediction and uncertainty estimation tasks separately. Three variants were examined: Variant 1 incorporates the proposed VAE encoder with the standard MSE loss function, Variant 2 excludes the VAE encoder while employing the proposed loss function, and Variant 3

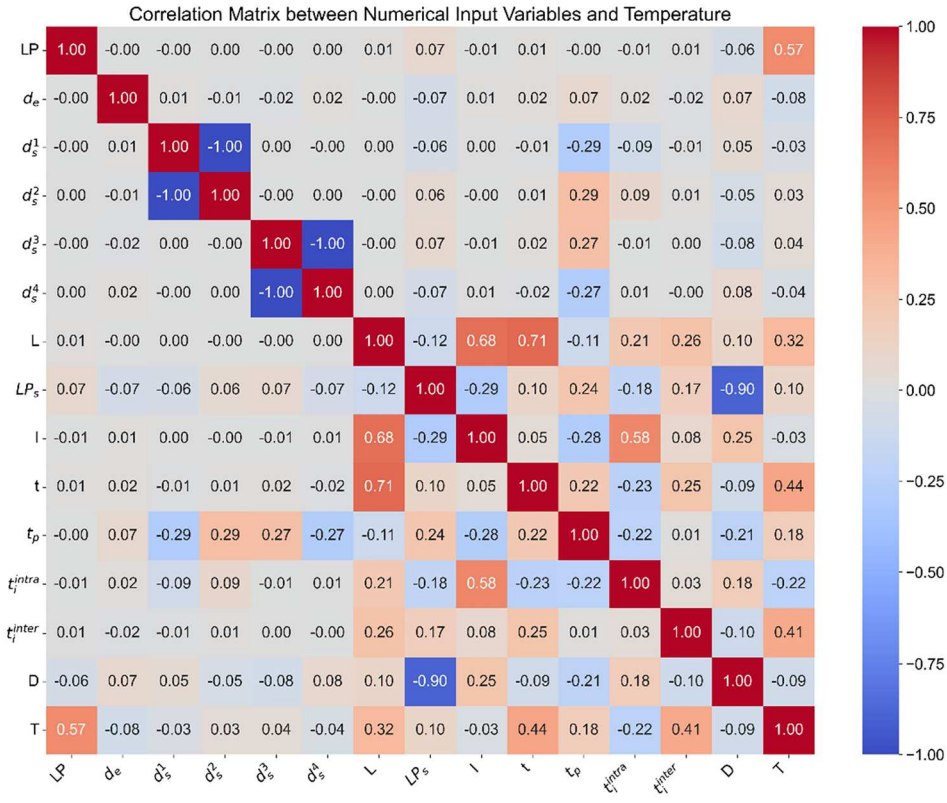

Figure 10. Correlation matrix showing Pearson correlation coefficients between numerical input variables and temperature in the DED process. Red and blue colours indicate positive and negative correlations, respectively, with colour intensity proportional to the correlation strength.

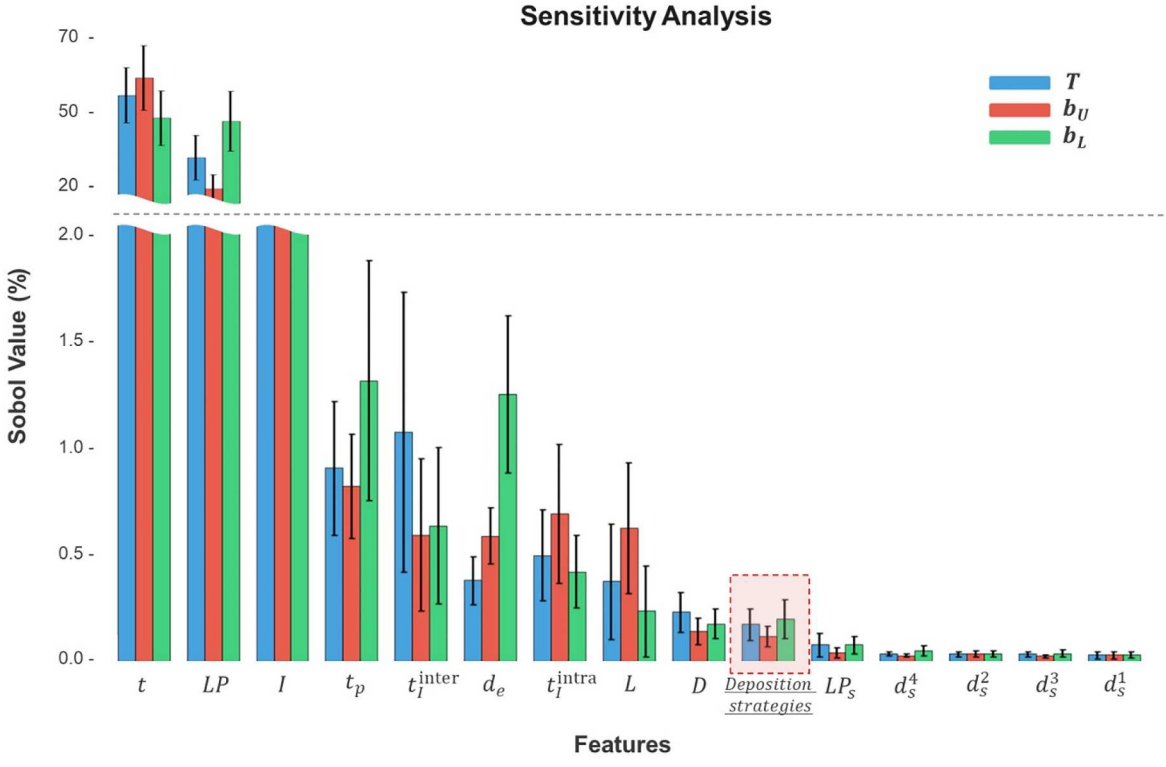

Figure 11. Sobol indices for each input variable inducing variance on the temperature and upper and lower bound. T is the temperature, b_U is the upper bound, and b_L is the lower bound.

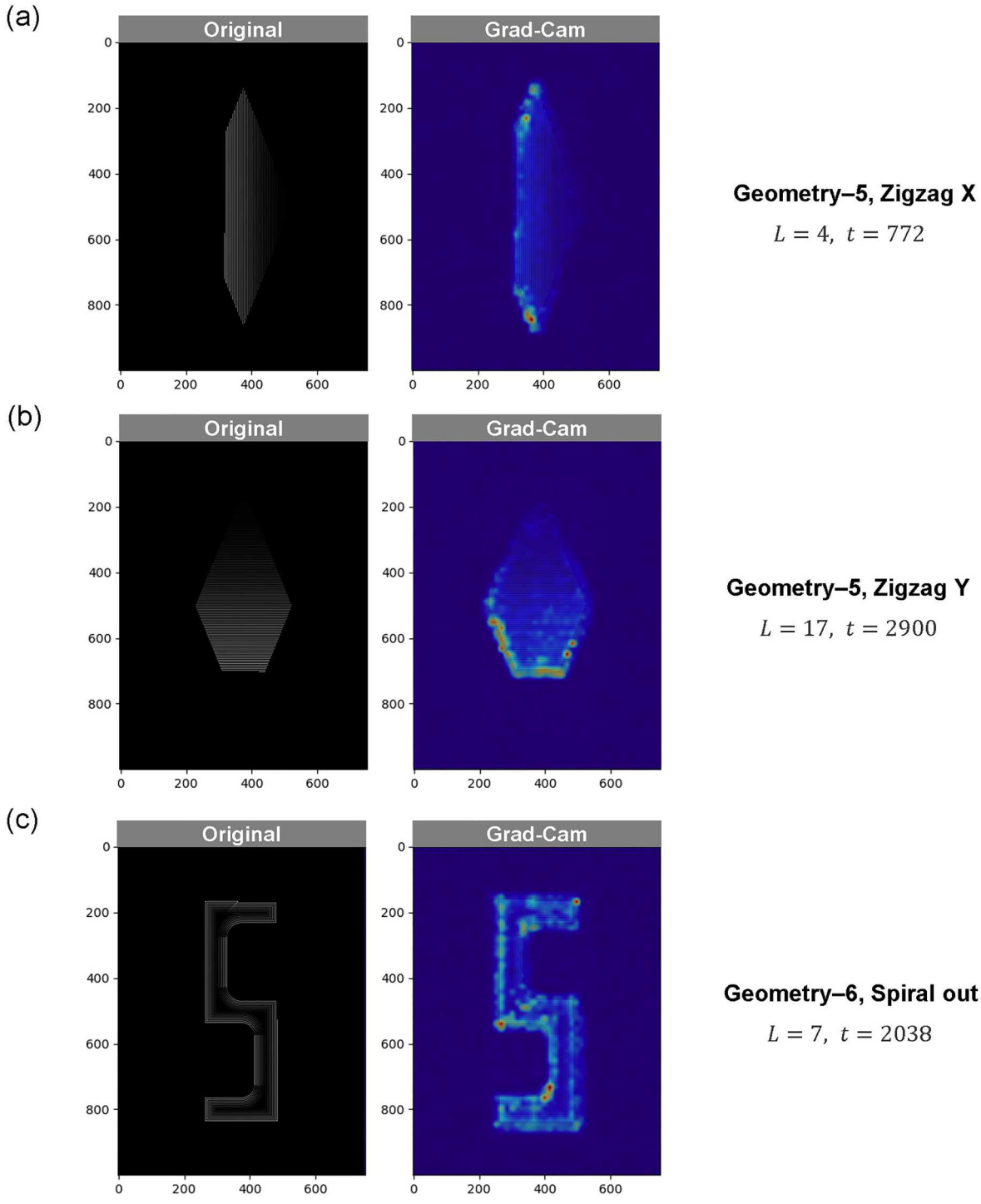

Figure 12. Grad-CAM activation maps on deposition strategy image for example cases: (a) geometry-5 with a zigzag X, (b) geometry-5 with a zigzag Y, and (c) geometry-6 with a spiral out. These maps highlight the regions most influential for the model's predictions. Red and yellow denote high-importance areas and blue marks less relevant regions.

replaces the proposed VAE encoder with a simplified ResNet described in Section 4.1. The same evaluation metrics were adopted to compare all variants.

The introduction of the VAE encoder in Variant 1 resulted in improvements in both temperature prediction and uncertainty quantification compared to the baseline, underscoring the encoder's ability to capture the spatial details and latent features critical for modelling the thermal behaviours of the printed part. However, relying solely on the MSE loss function led to higher upper bound predictions and, in some cases, a marginally diminished lower bound, implying that a

Table 9. Temperature prediction performance in the ablation study evaluating the modules of the proposed method. The best result is highlighted in bold.

Model	Component		Metric				Computational efficiency	
	Image encoder	Loss function	RMSE(↓)	MAE(↓)	MAPE(↓)	R^2(↑)	Number of parameters	Inference time (ms)
Baseline	–	–	24.5273	19.7832	8.6140	0.8681	771,171,331	3.2963
Variant1	√	–	20.8998	16.8718	7.2416	0.9042	4,744,195	1.7029
Variant2	–	√	23.6184	18.8919	8.3067	0.8777	771,171,331	3.2973
Variant3	√	√	24.0937	19.2641	9.1840	0.8726	10,084,145	**1.1581**
Proposed	√	√	**19.9069**	**16.0029**	**7.1527**	**0.9131**	**4,744,195**	1.6998

Table 10. Uncertainty quantification performance in the ablation study evaluating the modules of the proposed method. The best and second highest results are highlighted in bold and underlined, respectively.

			Confidence interval									
	Component		Upper bound				Lower bound					
Model	Image encoder	Loss function	RMSE (↓)	MAE (↓)	MAPE(↓)	R^2 (↑)	RMSE (↓)	MAE (↓)	MAPE(↓)	R^2 (↑)	CP (↑)	IoU (↑)
Baseline	–	–	24.0646	18.4678	7.6769	0.8953	30.1767	25.1611	11.2977	0.7525	71.2578	0.3665
Variant1	√	–	**20.8613**	**16.2040**	**6.5991**	**0.9213**	30.3294	25.1384	10.9311	0.7500	77.5769	0.3933
Variant2	–	√	23.6792	<u>18.3138</u>	7.5696	0.8986	30.5846	25.5880	11.3987	0.7457	76.7457	0.3778
Variant3	√	√	23.6558	18.3999	8.3853	0.8988	29.5318	24.2638	11.6261	0.7628	74.5950	0.3875
Proposed	√	√	<u>23.1049</u>	18.4099	<u>7.5480</u>	<u>0.9035</u>	**24.9158**	**20.1947**	**9.2953**	**0.8312**	**79.7429**	**0.4156**

single-objective loss inadequately balances both regression tasks. Variant 2, which excluded the VAE encoder, yielded only modest improvements over the baseline, as direct feedforward inputs cannot efficiently deliver the same level of geometric details available through learned image representation. Variant 3 with a simplified ResNet in place of the VAE encoder, exhibited moderate performance gains relative to the baseline but did not match the performance levels achieved by Variant 1. While the ResNet architecture effectively extracts general image features, it appears less capable of capturing complex spatial dependencies related to temperature compared with the latent space learned by the VAE.

Notably, the use of VAE encoder also reduced computational overhead. As shown in Table 9, the proposed method has fewer parameters and achieves faster inference than most other variants. This efficiency is attributed to the architecture of VAE, which balances robust feature extraction with a compact design.

Overall, the ablation study confirms that while alternative feature extraction methods can offer certain benefits, the specialised latent representation by the VAE model proves most effective for the temperature prediction task. Furthermore, the proposed loss function plays a crucial role in refining uncertainty quantification. The combination of the VAE encoder and an appropriate multi-task loss function ensures precise temperature prediction, reliable uncertainty estimation, and computational efficiency, thereby validating the efficacy of the proposed methodology.

4.4. Thermal defect inference from predicted temperature analysis

To gain a deeper insight into the thermal behaviour in the DED process, the temperature distribution on the bottom surface of the substrate was inferred using the proposed multimodal method, as depicted in Figure 13. Examining the temperature distribution at the 8th layer ensures that the temperature has adequately converged, providing a reliable representation of the thermal behaviour. The figure reveals that a

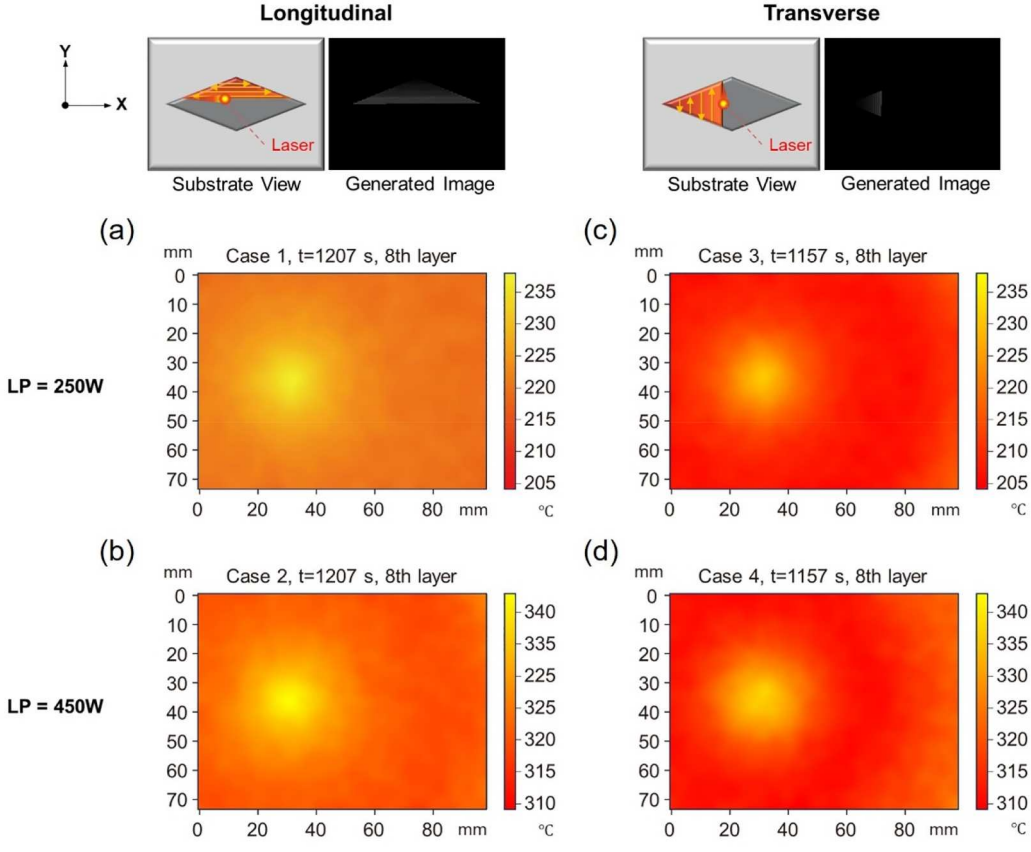

Figure 13. Predicted temperature distributions for different deposition strategies and laser power conditions. (a) Case 1: zigzag X deposition strategy with LP 250 W, (b) Case 2: zigzag X deposition strategy with LP 450 W, (c) Case 3: zigzag Y deposition strategy with LP 250 W, and (d) Case 4: zigzag Y deposition strategy with LP 450 W. The colour maps illustrate temperature gradients, where higher temperatures are concentrated around the laser.

Table 11. Measured distortion for test cases 1–4.

	Case 1 (250W)		Case 2 (450W)		Case 3 (250W)		Case 4 (450W)	
	Distortion direction							
Replicate	X-axis	Y-axis	X-axis	Y-axis	X-axis	Y-axis	X-axis	Y-axis
1	0.356	0.183	0.578	0.232	0.276	0.502	0.419	0.933
2	0.419	0.136	0.640	0.230	0.312	0.401	0.427	0.879
3	0.395	0.141	0.636	0.262	0.332	0.462	0.432	0.779
Average	0.390	0.154	0.618	0.241	0.307	0.455	0.426	0.864

zigzag deposition strategy in the Y-direction (Cases 3 and 4) produced a more pronounced temperature gradient compared to the zigzag in the X-direction (Cases 1 and 2). A similar trend in temperature distribution was observed for higher power settings, that is, 450W. This is obvious as zigzag X involves more frequent back-and-forth movements, which distribute heat more evenly across the substrate.

The intensified temperature gradients observed in Cases 3 and 4 indicate an increased susceptibility to thermal defects, particularly geometric distortion. Distortion in metal AM primarily results from steep temperature gradients between the molten pool and surrounding material, where the heated region experiences restricted expansion and subsequent shrinkage during cooling, generating internal stresses. The magnitude and direction of these thermal gradients directly influence the severity of part distortion, typically manifesting as upward curling of the top layer [111].

To validate the prediction results, the distortion of the substrates from the four cases was measured using a Mitutoyo ® SV-C3100 formtracer equipped with a SPH-51 Stylus, operating at a speed of 1 mm/s with a sampling rate of 50 Hz [93]. Figure S8 illustrates the form tracing line used to measure the distortion, specifically measuring the maximum z-deflection relative to the reference plane.

The measurements, recorded across all replicates and listed in Table 11, confirmed that Cases 3 and 4 exhibited greater average distortion compared to Cases 1 and 2 respectively, thereby validating the predicted thermal distribution. In addition, the measurement results revealed a correlation between deposition path direction and distortion characteristics, with maximum distortion aligning along the deposition axis [112]. This directional behaviour can be attributed to the cyclical thermal loading during the deposition process, where alternating patterns produce repetitive cycles of thermal expansion and contraction, leading to concentrated residual stresses that align with the deposition path direction.

To conclude, the findings presented in this section highlight both the predictive capability and practical applicability of the proposed method.

5. Conclusions

In this work, we presented a novel multimodal DL-based approach for temperature prediction and uncertainty estimation in the DED process. The proposed method consists of two main components: a multimodal framework that combines deposition strategy image with numerical process features to improve temperature predictions, and an uncertainty quantification method mimicking the concepts in the object detection models to achieve reliable confidence interval estimation. The framework integrates a VAE encoder to extract salient features from the 2D representations of deposition strategies. For efficient model training and evaluation, a customised multi-task loss function was utilised paired with a modified IoU metric for robust assessment of confidence intervals.

Through extensive experimental validation, including benchmarking against deterministic, probabilistic DL models and SOTA multimodal methods, the proposed method had proven effective in predicting process temperatures while offering uncertainty estimates that better align with actual deviations, thereby mitigating both over- or under-confidences. An ablation study further confirmed the efficacy of the architectural design. Sensitivity analysis and Grad-CAM visualisations offered deeper insights into the relative contributions of input variables to the prediction outcomes. The practical value of the proposed method was underscored through an in-depth analysis of predicted temperature distributions.

The focus on capturing aleatoric uncertainty leads to some limitations in addressing other sources of uncertainty in AM processes. In addition, the accuracy of the confidence intervals depends heavily on the quality and representativeness of the experimental data, while efforts to increase data collection incur substantial costs and material waste. To address these challenges, future work should explore the integration of experimental data with simulation results. The current study also omits certain process parameters, such as powder flow rate and nozzle speed, which may have a significant impact on the process outcomes.

Nevertheless, the proposed method shows promise for optimising deposition strategies to minimise

thermal distortion. Furthermore, although applied specifically to the DED process, it serves as a versatile framework that could be extended to other manufacturing processes and broader manufacturing applications.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by Electronics and Telecommunications Research Institute (ETRI) grant funded by the Korean government [Grant No: 24ZB1210, Development of human-mobile-space autonomous collaboration intelligence technology development in industrial sites], by the Korea Institute of Energy Technology Evaluation and Planning (KETEP) and the Ministry of Trade, Industry & Energy(MOTIE) of the Republic of Korea [Grant No. RS-2023-00240918], and by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) [Grant No. RS-2024-00416891].

Data availability

Data will be made available upon request from the corresponding author.

ORCID

Adrian Matias Chung Baek http://orcid.org/0000-0002-4372-1866
Taehwan Kim http://orcid.org/0009-0000-9670-4667
Minkyu Seong http://orcid.org/0009-0001-5389-4939
Seungjae Lee http://orcid.org/0009-0008-4234-8546
Ho Gyeong Kang http://orcid.org/0009-0005-2878-6126
Eun Ju Park http://orcid.org/0009-0002-6339-3844
Im Doo Jung http://orcid.org/0000-0003-0883-1848
Namhun Kim http://orcid.org/0000-0003-4429-2191

References

[1] Wang C, Sun CN, Zhang L, et al. Improving component dimensional accuracy in electron beam powder bed fusion by addressing nonlinear deformations with 3D compensation strategies. Virtual Phys Prototyp. 2024;19(1):e2430319. doi:10.1080/17452759.2024.2430319
[2] Li X, Wang P, Zhao M, et al. Customizable anisotropic microlattices for additive manufacturing: machine learning accelerated design, mechanical properties and structural-property relationships. Addit Manuf. 2024;89:104248. doi:10.1016/j.addma.2024.104248
[3] Schmidt M, Merklein M, Bourell D, et al. Laser based additive manufacturing in industry and academia. CIRP Ann. 2017;66 (2):561–583.
[4] Cagirici M, Guo S, Ding J, et al. Additive manufacturing of high-entropy alloys: current status and challenges. Smart Mater Manuf. 2024;2:100058. doi:10.1016/j.smmf.2024.100058
[5] Ren K, Chew Y, Zhang YF, et al. Thermal field prediction for laser scanning paths in laser aided additive manufacturing by physics-based machine learning. Comput Methods Appl Mech Eng. 2020;362:112734. doi:10.1016/j.cma.2019.112734
[6] Gu D, Shi X, Poprawe R, et al. Material-structure-performance integrated laser-metal additive manufacturing. Science. 2021;372(6545):eabg1487. doi:10.1126/science.abg1487
[7] Cook PS, Murphy AB. Simulation of melt pool behaviour during additive manufacturing: underlying physics and progress. Addit Manuf. 2020;31:100909. doi:10.1016/j.addma.2019.100909
[8] Li C, Liu ZY, Fang XY, et al. Residual stress in metal additive manufacturing. Procedia CIRP. 2018;71:348–353.
[9] Bontha S, Klingbeil NW, Kobryn PA, et al. Thermal process maps for predicting solidification microstructure in laser fabrication of thin-wall structures. J Mater Process Technol. 2006;178 (1–3):135–142.
[10] Guo S, Zamiela C, Bian L. Knowledge-transfer-enabled porosity prediction for new part geometry in Laser Metal Deposition. J Manuf Process. 2023;103:64–77.
[11] Yan H, Shen L, Wang X, et al. Stress and deformation evaluation of the subarea scanning effect in direct laser-deposited Ti-6Al-4 V. Int J Adv Manuf Technol. 2018;97 (1–4):915–926.
[12] Yang Q, Zhang P, Cheng L, et al. Finite element modeling and validation of thermomechanical behavior of Ti-6Al-4 V in directed energy deposition additive manufacturing. Addit Manuf. 2016;12:169–177.
[13] Oliveira JP, LaLonde AD, Ma J. Processing parameters in laser powder bed fusion metal additive manufacturing. Mater Des. 2020;193:108762. doi:10.1016/j.matdes.2020.108762
[14] Kiss AM, Fong AY, Calta NP, et al. Laser-induced keyhole defect dynamics during metal additive manufacturing. Adv Eng Mater. 2019;21(10):1900455. doi:10.1002/adem.201900455
[15] Chen H, Liu X, Liu X, et al. MeltpoolGAN: melt pool prediction from path-level thermal history. Addit Manuf. 2024;84:104095. doi:10.1016/j.addma.2024.104095
[16] Mozaffar M, Liao S, Lin H, et al. Geometry-agnostic data-driven thermal modeling of additive manufacturing processes using graph neural networks. Addit Manuf. 2021;48:102449. doi:10.1016/j.addma.2021.102449
[17] Mozaffar M, Paul A, Al-Bahrani R, et al. Data-driven prediction of the high-dimensional thermal history in directed energy deposition processes via recurrent neural networks. Manuf Lett. 2018;18:35–39.
[18] Perumal V, Abueidda D, Koric S, et al. Temporal convolutional networks for data-driven thermal modeling of directed energy deposition. J Manuf Process. 2023;85:405–416.
[19] Xie J, Chai Z, Xu L, et al. 3D temperature field prediction in direct energy deposition of metals using physics informed neural network. Int J Adv Manuf Technol. 2022;119 (5–6):3449–3468.
[20] Ness KL, Paul A, Sun L, et al. Towards a generic physics-based machine learning model for geometry invariant

thermal history prediction in additive manufacturing. J Mater Process Technol. 2022;302:117472. doi:10.1016/j.jmatprotec.2021.117472

[21] Sideris I, Crivelli F, Bambach M. GPyro: uncertainty-aware temperature predictions for additive manufacturing. J Intell Manuf. 2023;34 (1):243–259.

[22] Zhou Z, Shen H, Liu B, et al. Thermal field prediction for welding paths in multi-layer gas metal arc welding-based additive manufacturing: a machine learning approach. J Manuf Process. 2021;64:960–971.

[23] Sajadi P, Dehaghani MR, Tang Y, et al. Real-Time 2D temperature field prediction in metal additive manufacturing using physics-informed neural networks. 2024 Jan 4; Available from: https://arxiv.org/abs/2401.02403

[24] Liao S, Xue T, Jeong J, et al. Hybrid thermal modeling of additive manufacturing processes using physics-informed neural networks for temperature prediction and parameter identification. Comput Mech. 2023;72 (3):499–512.

[25] Nijhuis B, Geijselaers HJM, van den Boogaard AH. Efficient thermal simulation of large-scale metal additive manufacturing using hot element addition. Comput Struct. 2021;245:106463. doi:10.1016/j.compstruc.2020.106463

[26] Bandyopadhyay A, Traxel KD. Invited review article: metal-additive manufacturing – modeling strategies for application-optimized designs. Addit Manuf. 2018;22:758–774.

[27] Perumal VI, Najafi AR, Kontsos A. A novel digital design approach for metal additive manufacturing to address local thermal effects. Designs (Basel). 2020;4 (4):1–16.

[28] Gan Z, Liu H, Li S, et al. Modeling of thermal behavior and mass transport in multi-layer laser additive manufacturing of Ni-based alloy on cast iron. Int J Heat Mass Transf. 2017;111:709–722.

[29] Mukherjee T, Wei HL, De A, et al. Heat and fluid flow in additive manufacturing – part I: modeling of powder bed fusion. Comput Mater Sci. 2018;150:304–313.

[30] Roberts IA, Wang CJ, Esterlein R, et al. A three-dimensional finite element analysis of the temperature field during laser melting of metal powders in additive layer manufacturing. Int J Mach Tools Manuf. 2009 Oct;49 (12–13):916–923.

[31] Francois MM, Sun A, King WE, et al. Modeling of additive manufacturing processes for metals: challenges and opportunities. Curr Opin Solid State Mater Sci. 2017;21:198–206.

[32] Li J, Wang Q, Michaleris P. An analytical computation of temperature field evolved in directed energy deposition. J Manuf Sci Eng Trans ASME. 2018;140(10):101004. doi:10.1115/1.4040621

[33] Plotkowski A, Kirka MM, Babu SS. Verification and validation of a rapid heat transfer calculation methodology for transient melt pool solidification conditions in powder bed metal additive manufacturing. Addit Manuf. 2017;18:256–268.

[34] Forslund R, Snis A, Larsson S. Analytical solution for heat conduction due to a moving Gaussian heat flux with piecewise constant parameters. Appl Math Model. 2019;66:227–240.

[35] Yang Y, Knol MF, van Keulen F, et al. A semi-analytical thermal modelling approach for selective laser melting. Addit Manuf. 2018;21:284–297.

[36] Reza Yavari M, Cole KD, Rao P. Thermal modeling in metal additive manufacturing using graph theory. J Manuf Sci Eng Trans ASME. 2019;141(7):071007. doi:10.1115/1.4043648

[37] Roy M, Wodo O. Data-driven modeling of thermal history in additive manufacturing. Addit Manuf. 2020;32:101017. doi:10.1016/j.addma.2019.101017

[38] Paul A, Mozaffar M, Yang Z, et al. A real-time iterative machine learning approach for temperature profile prediction in additive manufacturing processes. Proceedings – 2019 IEEE International Conference on Data Science and Advanced Analytics, DSAA 2019; 2019; Institute of Electrical and Electronics Engineers Inc. p. 541–550.

[39] Lo YL, Liu BY, Tran HC. Optimized hatch space selection in double-scanning track selective laser melting process. Int J Adv Manuf Technol. 2019;105 (7–8):2989–3006.

[40] Stathatos E, Vosniakos GC. Real-time simulation for long paths in laser-based additive manufacturing: a machine learning approach. Int J Adv Manuf Technol. 2019;104 (5–8):1967–1984.

[41] Cao X, Duan C, Luo X, et al. Deep learning-based rapid prediction of temperature field and intelligent control of molten pool during directed energy deposition process. Addit Manuf. 2024;94:104501. doi:10.1016/j.addma.2024.104501

[42] Anandan Kumar H, Kumaraguru S, Paul CP, et al. Faster temperature prediction in the powder bed fusion process through the development of a surrogate model. Opt Laser Technol. 2021;141:107122. doi:10.1016/j.optlastec.2021.107122

[43] Le VT, Nguyen HD, Bui MC, et al. Rapid and accurate prediction of temperature evolution in wire plus arc additive manufacturing using feedforward neural network. Manuf Lett. 2022;32:28–31.

[44] Pham TQD, Hoang TV, Van Tran X, et al. Fast and accurate prediction of temperature evolutions in additive manufacturing process using deep learning. J Intell Manuf. 2023;34 (4):1701–1719.

[45] Chen J, Pierce J, Williams G, et al. Accelerating thermal simulations in additive manufacturing by training physics-informed neural networks with randomly synthesized data. J Comput Inf Sci Eng. 2024;24(1). doi:10.1115/1.4062852

[46] Kim SI, Hart AJ. A spiral laser scanning routine for powder bed fusion inspired by natural predator-prey behaviour. Virtual Phys Prototyp. 2022;17 (2):239–255.

[47] Ren K, Chew Y, Liu N, et al. Integrated numerical modelling and deep learning for multi-layer cube deposition planning in laser aided additive manufacturing. Virtual Phys Prototyp. 2021;16 (3):318–332.

[48] Vandone A, Baraldo S, Valente A. Multisensor data fusion for additive manufacturing process control. IEEE Robot Autom Lett. 2018;3 (4):3279–3284.

[49] Chen L, Bi G, Yao X, et al. In-situ process monitoring and adaptive quality enhancement in laser additive manufacturing: A critical review. J Manuf Syst. 2024;74:527–574.

[50] Kozjek D, Porter C, Carter FM, et al. Iterative closest point-based data fusion of non-synchronized in-situ and ex-situ data in laser powder bed fusion. J Manuf Syst. 2023;66:179–199.

[51] Zhang H, Caputo AN, Vallabh CKP, et al. Evaluating and correlating multimodal process dynamics, microstructure features, and mechanical properties in laser powder bed fusion. J Manuf Process. 2024;127:511–530.
[52] 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. doi:10.1016/j.addma.2022.103075
[53] Snow Z, Scime L, Ziabari A, et al. Scalable in situ nondestructive evaluation of additively manufactured components using process monitoring, sensor fusion, and machine learning. Addit Manuf. 2023;78:103817. doi:10.1016/j.addma.2023.103817
[54] Akhavan J, Manoochehri S. Sensory data fusion using machine learning methods for in-situ defect registration in additive manufacturing: a review. 2022 IEEE International IOT, Electronics and Mechatronics Conference, IEMTRONICS 2022; 2022; Institute of Electrical and Electronics Engineers Inc.
[55] Pandiyan V, Baganis A, Axel Richter R, et al. Qualify-as-you-go: sensor fusion of optical and acoustic signatures with contrastive deep learning for multi-material composition monitoring in laser powder bed fusion process. Virtual Phys Prototyp. 2024;19(1):e2356080. doi:10.1080/17452759.2024.2356080
[56] Yu R, Tan X, He S, et al. Monitoring of robot trajectory deviation based on multimodal fusion perception in WAAM process. Measurement (Lond). 2024;224:113933. doi:10.1016/j.measurement.2023.113933
[57] Li J, Cao L, Zhou Q, et al. Imbalanced quality monitoring of selective laser melting using acoustic and photodiode signals. J Manuf Process. 2023;105:14–26.
[58] Zhang J, Wang P, Gao RX. Deep learning-based tensile strength prediction in fused deposition modeling. Comput Ind. 2019;107:11–21.
[59] Wu Q, Yang F, Lv C, et al. In-Situ quality intelligent classification of additively manufactured parts using a multi-sensor fusion based melt pool monitoring system. Addit Manuf Front. 2024;3 (3):200153.
[60] Karthikeyan A, Balhara H, Hanchate A, et al. In-situ surface porosity prediction in hybrid-directed energy deposition process using explainable multimodal sensor fusion. ArXiv. 2023; Available from: https://arxiv.org/pdf/2304.08658
[61] Petrich J, Snow Z, Corbin D, et al. Multi-modal sensor fusion with machine learning for data-driven process monitoring for additive manufacturing. Addit Manuf. 2021;48:102364. doi:10.1016/j.addma.2021.102364
[62] Khusheef AS, Shahbazi M, Hashemi R. Predicting 3D printed plastic part properties: a deep learning approach with thermographic and vibration data fusion. Expert Syst Appl. 2024;255:124605. doi:10.1016/j.eswa.2024.124605
[63] Biehler M, Li J, Shi J. FUSION3D: multimodal data fusion for 3D shape reconstruction: a soft-sensing approach. IISE Trans. 2024;1–15. doi:10.1080/24725854.2024.2376650
[64] Perani M, Baraldo S, Decker M, et al. Track geometry prediction for Laser Metal Deposition based on on-line artificial vision and deep neural networks. Robot Comput Integr Manuf. 2023;79:102445. doi:10.1016/j.rcim.2022.102445
[65] Chen L, Bi G, Yao X, et al. Multisensor fusion-based digital twin for localized quality prediction in robotic laser-directed energy deposition. Robot Comput Integr Manuf. 2023;84:102581. doi:10.1016/j.rcim.2023.102581
[66] Chen L, Moon SK. In-situ defect detection in laser-directed energy deposition with machine learning and multi-sensor fusion. J Mech Sci Technol. 2024;38 (9):4477–4484.
[67] Zou Z, Zhang K, Liu T, et al. In situ monitoring with melt pool data based on multi-signal fusion method in laser powder bed fusion. Measurement (Lond). 2024;234:114877. doi:10.1016/j.measurement.2024.114877
[68] Li J, Zhang X, Zhou Q, et al. A feature-level multi-sensor fusion approach for in-situ quality monitoring of selective laser melting. J Manuf Process. 2022;84:913–926.
[69] Li J, Zhou Q, Cao L, et al. A convolutional neural network-based multi-sensor fusion approach for in-situ quality monitoring of selective laser melting. J Manuf Syst. 2022;64:429–442.
[70] Kim DB, Chong H, Mahdi MM, et al. Data-fused and concatenated-ensemble learning for in-situ anomaly detection in wire and arc-based direct energy deposition. J Manuf Process. 2024;112:273–289.
[71] Bevans B, Barrett C, Spears T, et al. Heterogeneous sensor data fusion for multiscale, shape agnostic flaw detection in laser powder bed fusion additive manufacturing. Virtual Phys Prototyp. 2023;18(1):e2196266. doi:10.1080/17452759.2023.2196266
[72] Saunders RN, Teferra K, Elwany A, et al. Metal AM process-structure-property relational linkages using Gaussian process surrogates. Addit Manuf. 2023;62:103398. doi:10.1016/j.addma.2023.103398
[73] Hu Z, Mahadevan S. Uncertainty quantification and management in additive manufacturing: current status, needs, and opportunities. Int J Adv Manuf Technol. 2017;93 (5–8):2855–2874.
[74] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks [Internet]. Available from: http://code.google.com/p/cuda-convnet/.
[75] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition [Internet]. Available from: http://image-net.org/challenges/LSVRC/2015/.
[76] Wang J, Lan C, Liu C, et al. Generalizing to unseen domains: a survey on domain generalization. 2021 Mar 2; Available from: http://arxiv.org/abs/2103.03097.
[77] Kurach K, Lucic M, Zhai X, et al. A large-scale study on regularization and normalization in GANs [Internet]. Available from: www.tensorflow.org/hub.
[78] Kaae Sønderby C, Raiko T, Maaløe L, et al. Ladder variational autoencoders. Adv Neural Inf Process Syst. 2016;29:3738–3746.
[79] He YL, Li XY, Ma JH, et al. Attribute-relevant distributed variational autoencoder integrated with LSTM for dynamic industrial soft sensing. Eng Appl Artif Intell. 2023;119:105737. doi:10.1016/j.engappai.2022.105737
[80] Tsutsumi M, Saito N, Koyabu D, et al. A deep learning approach for morphological feature extraction based on variational auto-encoder: an application to mandible shape. NPJ Syst Biol Appl. 2023;9(1):30. doi:10.1038/s41540-023-00293-6

[81] Kingma DP, Welling M. Auto-encoding variational bayes. 2013 Dec 20; Available from: https://arxiv.org/abs/1312.6114

[82] Fei R, Wan Y, Hu B, et al. A novel network core structure extraction algorithm utilized variational autoencoder for community detection. Expert Syst Appl. 2023;222:119775. doi:10.1016/j.eswa.2023.119775

[83] Liu D, Hao P, Xu T, et al. Intelligent optimization of stiffener unit cell via variational autoencoder-based feature extraction. Struct Multidiscipl Optim. 2023;66(1):8. doi:10.1007/s00158-022-03463-y

[84] Abdar M, Pourpanah F, Hussain S, et al. A review of uncertainty quantification in deep learning: techniques, applications and challenges. Inf Fusion. 2021;76:243–297.

[85] Hüllermeier E, Waegeman W. Aleatoric and epistemic uncertainty in machine learning: an introduction to concepts and methods. Mach Learn. 2021;110 (3):457–506.

[86] Kaur R, Singh S. A comprehensive review of object detection with deep learning. Digital Signal Process: A Rev J. 2023;132:103812. doi:10.1016/j.dsp.2022.103812

[87] Cannon AJ. Quantile regression neural networks: Implementation in R and application to precipitation downscaling. Comput Geosci. 2011;37 (9):1277–1284.

[88] Ren S, He K, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks. 2015 Jun 4; Available from: http://arxiv.org/abs/1506.01497.

[89] Wu S, Yang J, Wang X, et al. IoU-Balanced loss functions for single-stage object detection. Pattern Recognit Lett. 2022;156:96–103.

[90] Park JM, Choi M, Um J. Convolutional LSTM based melt-pool prediction from images of laser tool path strategy in laser powder bed fusion for additive manufacturing. Int J Adv Manuf Technol. 2024;130 (3–4):1871–1886.

[91] Pham TQD, Hoang TV, Tran XV, et al. Characterization, propagation, and sensitivity analysis of uncertainties in the directed energy deposition process using a deep learning-based surrogate model. Probab Eng Mech. 2022;69:103297. doi:10.1016/j.probengmech.2022.103297

[92] Everton SK, Hirsch M, Stavroulakis PI, et al. Review of in-situ process monitoring and in-situ metrology for metal additive manufacturing. Mater Des. 2016;95:431–445.

[93] Lee J, Chung Baek AM, Kim N, et al. In situ monitoring-based feature extraction for metal additive manufacturing products warpage prediction. Smart Struct Syst. 2022;29 (6):767–775.

[94] Kim H, Seo J, Chung Baek AM, et al. Direct energy deposition for smart micro reactor. Virtual Phys Prototyp. 2024;19(1):e2411024. doi:10.1080/17452759.2024.2411024

[95] Mirazimzadeh SE, Pazireh S, Urbanic J, et al. Investigation of effects of different moving heat source scanning patterns on thermo-mechanical behavior in direct energy deposition manufacturing. Int J Adv Manuf Technol. 2022;120 (7–8):4737–4753.

[96] Mazzarisi M, Angelastro A, Latte M, et al. Thermal monitoring of Laser Metal Deposition strategies using infrared thermography. J Manuf Process. 2023;85:594–611.

[97] Kim T, Kim JG, Park S, et al. Virtual surface morphology generation of Ti-6Al-4 V directed energy deposition via conditional generative adversarial network. Virtual Phys Prototyp. 2023;18(1):e2124921. doi:10.1080/17452759.2022.2124921

[98] Takahashi K, Yamamoto K, Kuchiba A, et al. Confidence interval for micro-averaged F 1 and macro-averaged F 1 scores. Applied Intelligence. 2022;52 (5):4961–4972.

[99] Li L, Chang J, Vakanski A, et al. Uncertainty quantification in multivariable regression for material property prediction with Bayesian neural networks. Sci Rep. 2024;14(1):10543. doi:10.1038/s41598-024-61189-x

[100] Rezatofighi H, Tsoi N, Gwak J, et al. Generalized intersection over union: a metric and a loss for bounding box regression. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition.

[101] Kim G, Yang SM, Kim DM, et al. Developing a deep learning-based uncertainty-aware tool wear prediction method using smartphone sensors for the turning process of Ti-6Al-4 V. J Manuf Syst. 2024;76:133–157.

[102] Tan M, Le QV. Efficientnet: rethinking model scaling for convolutional neural networks. International Conference on Machine Learning.

[103] Gao J, Li P, Chen Z, et al. A survey on deep learning for multimodal data fusion. Neural Comput. 2020;32:829–864.

[104] Huang W. Multimodal contrastive learning and tabular attention for automated alzheimer's disease prediction. IEEE/CVF International Conference on Computer Vision (ICCV); 2023; Paris. p. 2473–2482.

[105] Ryu S, Kim H, Kim SG, et al. Probabilistic deep learning model as a tool for supporting the fast simulation of a thermal–hydraulic code. Expert Syst Appl. 2022;200:116966. doi:10.1016/j.eswa.2022.116966

[106] Mamun O, Taufique MFN, Wenzlick M, et al. Uncertainty quantification for Bayesian active learning in rupture life prediction of ferritic steels. Sci Rep. 2022;12(1): 2083. doi:10.1038/s41598-022-06051-8

[107] Ahmad R, Alsmadi I, Al-Ramahi M. Optimization of deep learning models: benchmark and analysis. Adv Comput Intell. 2023;2023;3(2):7. doi:10.1007/s43674-023-00055-1

[108] Saltelli A, Annoni P, Azzini I, et al. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. Comput Phys Commun. 2010;181 (2):259–270.

[109] Wang K, Xu J, Zhang S, et al. Towards low-budget energy efficiency design in additive manufacturing based on variational scale-aware transformer. J Clean Prod. 2023;393:136168. doi:10.1016/j.jclepro.2023.136168

[110] Westphal E, Seitz H. A machine learning method for defect detection and visualization in selective laser sintering based on convolutional neural networks. Addit Manuf. 2021;41:101965. doi:10.1016/j.addma.2021.101965

[111] Raut R, Kumar Ball A, Basak A. Temperature distribution prediction in laser powder bed fusion using transferable and scalable graph neural networks. 2024 Jul 18; Available from: https://arxiv.org/abs/2407.13838

[112] Monu MCC, Afkham Y, Chekotu JC, et al. Bi-directional scan pattern effects on residual stresses and distortion in as-built nitinol parts: a trend analysis simulation study. Integr Mater Manuf Innov. 2023;12 (1):52–69.