The International Journal of Advanced Manufacturing Technology
https://doi.org/10.1007/s00170-026-17678-6

ORIGINAL ARTICLE

Bead geometry prediction of directed energy deposition based on spatio-temporal neural network model

Yuanzhi Chen[1,2] · Wenze Zhang[1,2] · Yichen Wang[2] · Pengcheng Hu[2] · Molong Duan[1,3] · Zichuan Fu[2] · Xiaoke Deng[1,2]

Received: 24 November 2025 / Accepted: 4 February 2026

Abstract

In metal additive manufacturing (AM), accurately predicting bead geometry in Directed Energy Deposition (DED) is crucial for ensuring the quality and precision of the final part. However, this task is often complicated by the dynamic and nonlinear nature of the DED process. This study introduces a novel spatio-temporal neural network (STNN) model designed to predict bead geometry using real-time melt pool images captured during the DED process. Multiple datasets are developed to capture dynamic process responses. The proposed STNN, which considers critical time intervals and overlap rates of the melt pools, demonstrates superior prediction accuracy compared to existing methods. The results also identify the optimal time interval for real-time control and highlight that future time instances have a greater impact than past instances, as evidenced by the asymmetrical melt pool observed in thermal images. Furthermore, the predictive bead geometry enables real-time feedback control in DED by dynamically compensating for deposition height variations. Experimental validation shows a 56.82% reduction in mean absolute error (MAE) and a 57.05% reduction in root mean square error (RMSE) in final component height accuracy compared to conventional open-loop control. These advancements offer a promising tool for real-time monitoring and optimization of the DED process.

Keywords Additive manufacturing · Directed energy deposition · Spatio-Temporal neural network · Bead geometry prediction · Real-Time monitoring

1 Introduction

In directed energy deposition (DED) additive manufacturing, a moving laser continuously generates melt pools that dynamically change in both space and time, as illustrated by the thermal maps in Fig. 1(a). These melt pools can be viewed as digital representations of the spatio-temporal (ST) domain, as they correspond to specific locations and times within the DED process. When melt pools are closely spaced, overlaps occur between adjacent pools, collectively influencing the geometry of the final bead. Understanding the ST dynamics of these melt pools is crucial for accurately modelling bead geometry. Existing methods, such as numerical derivation, response surface methodology, and data-driven modelling, often fall short due to the highly dynamic and nonlinear characteristics of the DED process. This study proposes a novel spatio-temporal neural network designed to predict bead geometry using real-time melt pool images captured during the DED process. This approach allows for the exploration and interpretation of the ST response of the melt pool through a deep learning-based framework.

Existing research on bead geometry prediction predominantly utilizes approaches such as numerical derivation and response surface methodology. Numerical derivation, as demonstrated by Mbodj et al. [1] and Xiong et al. [2], leverages critical processing information—including laser power, movement speed, and powder feed rate—along with material properties to predict bead geometry. A common technique within this framework is temperature-based

Yuanzhi Chen and Wenze Zhang joint first authors.

✉ Pengcheng Hu
pengchenghu@hkust-gz.edu.cn

1 Hong Kong University of Science and Technology, Hong Kong SAR, China

2 The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, China

3 HKUST Shenzhen-Hong Kong Collaborative Innovation Research Institute, Futian, Shenzhen, China

Published online: 10 February 2026

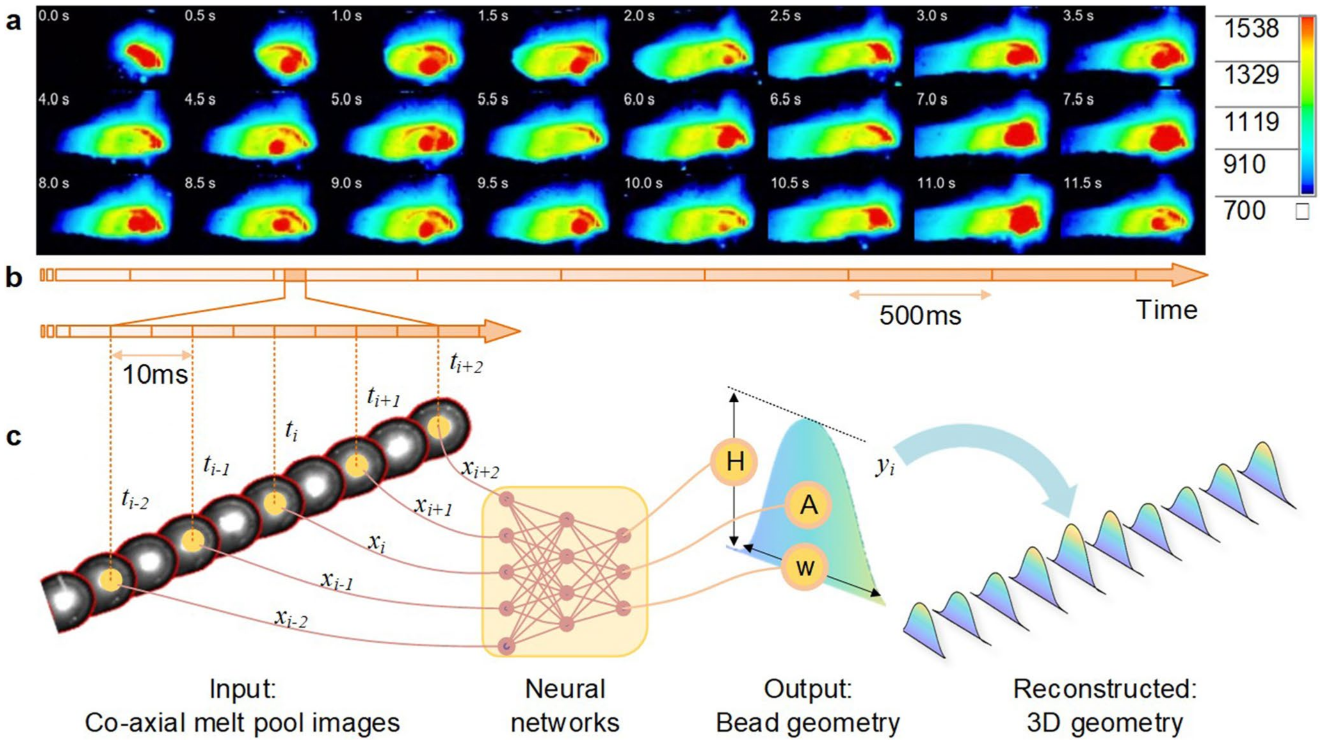

Fig. 1 Overview of our spatio-temporal neural network (STNN) for bead geometry prediction. (**a**) Thermal maps illustrating the dynamic development of melt pools; (**b**) Timestamps corresponding to thermal maps and co-axial images; (**c**) Schematics of melt pool geometry prediction process using neural networks

numerical derivation. Studies by Wei et al. [3, 4] and Huang et al. [5] focus on the evolution of temperature fields and the analysis of residual heat history to model the final clad geometry. Additionally, researchers like Lv et al. [6], Sun et al. [7], Wolff et al. [8], and Vundru et al. [9] investigate solid-liquid phase changes, melt pool formation, and subsequent geometry alterations to enhance the accuracy of deposition geometry predictions. Despite the comprehensive consideration of physical mechanisms, numerical derivation methods face two major limitations. First, simplifications of physical-chemical interactions in the solid-liquid mixture or the assumption of idealized profiles for laser or powder distribution can reduce the generalizability of these models. Second, finite-element-based methods are often computationally intensive and time-consuming. Response surface methodologies, as proposed by Khorram et al. [10] and Wolff et al. [8], aim to optimize bead geometry and product quality by establishing statistical relationships between printing parameters and process responses. However, these methods primarily focus on simplified static experimental conditions and gradual thermal gradient changes, neglecting the significant dynamic responses arising from continuously varying process parameters.

Recent advancements in machine learning (ML) technology have garnered significant attention due to their transformative impact [11–15]. ML has notably improved bead geometry prediction in both arc-based and laser-based metal additive manufacturing processes, enhancing both prediction accuracy and computational efficiency [16–18]. Data-driven regression models, such as those developed by Cai et al. [19] and Oh et al. [20], utilize algorithms like Random Forests (RF) and Support Vector Regression (SVR) to predict layer geometries from process parameters, providing an efficient foundation for process parameter optimization. Physics-informed models, as demonstrated by Cao et al. [21] and Yang et al. [22], integrate domain-specific knowledge into ML frameworks to enhance model robustness and interpretability. These models facilitate a seamless integration of data and mathematical representations, even in scenarios involving partially understood physical phenomena [23]. The advent of deep learning has further transformed DED additive manufacturing. Convolutional neural network (CNN)-based models, proposed by Jamnikar et al. [24], Asadi et al. [25], and Ribeiro et al. [26], leverage melt pool images for real-time geometry prediction and improved in-situ quality control, achieving high accuracy. Additionally, multi-sensor data fusion methods proposed by Chen et al. [27] and Kim et al. [28] integrate information from multiple sensors, yielding superior prediction results compared to image-only models. By extending ML-based approaches to more complex scenarios, Gihr et al. [29] developed a Multilayer Perceptron (MLP) model for corner structures in Wire Arc Additive Manufacturing (WAAM), while Oh et al. [20] and Wacker et al. [30] applied ML techniques to predict geometries in single and multi-layer structures, targeting applications closer to industrial settings. While these studies underscore ML's transformative potential in optimizing DED processes to enhance accuracy, efficiency, and robustness, few have addressed the spatio-temporal (ST) dependencies of the melt pool, which is critical for precisely predicting bead geometry.

ST-based deep learning, which captures both spatial and temporal dependencies in data, has the capability to predict bead geometry in metal additive manufacturing processes such as DED and arc welding. A natural approach is to leverage the strengths of deep CNNs for sequence analysis in visual pattern recognition [31, 32]. Recurrent neural network (RNN)-based models, such as Long Short-Term Memory (LSTM) networks [33] and Bidirectional LSTM (BiLSTM) [34], are particularly effective at capturing both long-term and short-term temporal dependencies. BiLSTM processes data in both forward and backward directions, providing a more comprehensive context for each time step and enhancing prediction accuracy compared to traditional LSTMs. Additionally, attention mechanisms, popularized by Vaswani et al. [35], dynamically weigh the importance of different time steps and spatial features, improving interpretability and accommodating varying input lengths. Graph Neural Networks (GNNs) [36, 37] can also capture ST relationships by representing data as graphs, where nodes encapsulate spatial information and edges reflect temporal relationships. For instance, Hu et al. [38] combine CNNs with BiLSTM to predict melt pool width in laser DED, while Yang et al. [39] utilize a BiLSTM-based model for process optimization and quality prediction. Zhou et al. [40] and Ren et al. [41] apply RNN-based models to predict melt pool thermal history based on trajectories, and Mozaffar et al. [42] employ GNNs for thermal modelling in DED processes. The results from these modelling approaches are consistent with finite element analysis. Furthermore, Hong et al. [43] implement a two-stream CNN for real-time quality monitoring, and Chen et al. [44] integrate multi-sensor data for defect prediction, both producing spatial feature maps that align with expert knowledge. Collectively, these studies [38–44] enhance the accuracy and generalizability of predictions by effectively capturing ST dependencies. However, there is a lack of clarity on how to determine the scale of ST that impacts the performance of DED, how these improvements are achieved, and what the relationships are between ST features and processing parameters.

Considering the aforementioned deficiencies, we propose a novel spatio-temporal neural network (STNN) that can precisely predict the bead geometry of DED manufacturing. The key contributions of this paper are outlined as follows:

1. **Accuracy Improvement with STNN**: The proposed STNN incorporates a core ST layer (STL) that effectively captures the dynamic interactions of the Directed Energy Deposition (DED) process, from which prediction accuracy is enhanced compared to models that rely on a single image for predictions.
2. **Optimal Time Interval for ST Response Prediction**: The overlap rate of the melt pool significantly impacts bead geometry prediction: excessive overlap obscures ST relationships, while insufficient overlap weakens them. We have determined the optimal time interval for predicting the ST response to achieve the best overlap rate.
3. **Asymmetrical Time Impact on Melt Pool Morphology**: Our findings, supported by thermal images, indicate that future time instances exert a greater influence on melt pool morphology than past time instances, highlighting the inherent asymmetry of the melt pool.

This paper is organized as follows: In Sect. 2, we introduce the STNN method, detailing the data generation and model design. Section 3 elaborates on the experimental approaches employed. In Sect. 4, we present the prediction performance of the STNN and interpret the results. Finally, Sect. 5 summarizes the conclusions, limitations, and directions for future work.

2 Methodology

In Sect. 2.1, we introduce the problem, outlining the objectives, inputs, and outputs of the proposed model. Section 2.2 describes the design of control parameters to stimulate the ST response. In Sect. 2.3, we cover the registration techniques used to align captured features to consistent temporal coordinates. Section 2.4 focuses on the dataset design for analyzing the model's response across various time intervals and sample sizes. Section 2.5 provides a detailed explanation of the neural network architecture. Finally, in Sect. 2.6, we present the metrics used to evaluate model performance.

2.1 Problem formation for the definition of the STNN

The objective of this research is to predict the deposit geometry during the DED process using real-time sequential image data (Fig. 1(c)). We employ a co-axial CCD camera to capture in-situ melt pool images, as it remains unaffected by factors such as part size, printing orientation, and powder flow, making it an ideal source of real-time data.

During DED, bead geometry is not determined solely by the instantaneous melt pool state, but is influenced by the cumulative thermal history and interactions with neighboring melt pools along the toolpath. As a result, both spatial and temporal dependencies play a critical role in geometry formation. Unlike prior approaches that rely on a single melt pool image for geometry prediction [24–26], this study formulates the problem as a spatio-temporal learning task and develops a STNN that exploits sequences of melt pool images to capture such dependencies.

Input to STNN Let X_0 denote the complete sequence of melt pool images captured at a constant sampling interval τ , containing a total N_0 images. For a given target image x_i, the input to the STNN is constructed as a local temporal window $X_i = \{x_{i-(N-1)/2}, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_{i+(N-1)/2},\}$, which consists of N images centered at x_i. Adjacent images in X_i are sampled at intervals of $\alpha\,\tau$, where α denotes the temporal subsampling factor relative to the original dataset X_0. This construction preserves the temporal structure of the melt pool evolution while enabling focused modelling of the local deposition context around x_i.

Output of STNN The output is the bead geometry $y_i = [W, H, A]$ corresponding to the central image x_i, where W, H, and A are the width, height, and area of the bead, respectively.

With the above-mentioned input and output defined, the STNN can be expressed as the functional mapping:

$$y_i = STNN\,(X_i) \tag{1}$$

Ground-truth geometric labels are obtained using a structured light camera, enabling supervised training of the STNN based on sequential melt pool images. Through this formulation, the network jointly leverages spatial information from individual images and temporal dependencies across image sequences to achieve accurate bead geometry prediction.

2.2 Control parameters design for generating the dataset of modelling

To create datasets suited for studying the ST response of the DED process, we employ a homebuilt AM-monitoring system to collect data from multiple sources, including real-time melt pool images from a co-axial CCD camera, laser head positions and velocities from the motion controller, and bead geometry profiles from a structured light camera. We select laser power and travel speed as control parameters because they can dynamically change with a fast response rate (on the scale of tens of milliseconds) during the printing process, while keeping all other parameters constant. The control parameters are designed to: (1) stimulate various dynamic responses in the melt pool using different frequencies and waveforms; (2) expand parameter ranges to enhance model generalizability; and (3) limit the number of experiments to improve efficiency in experimentation and model training. As shown in Fig. 2(a), we applied four periodic patterns—constant function, sine wave, square wave, and sawtooth wave—to both control parameters, resulting in a total of 16 combinations, each used to print a bead that is 30 mm in length. The power and speed patterns maintain a fixed phase difference of $\pi/2$ in each bead to produce dynamic changes in energy deposition rate (EDR), calculated by $\mathrm{EDR} = \text{laser power} / \text{travel speed}$. Furthermore, to ensure data diversity, four frequency options and three magnitude options are applied for each periodic pattern (Fig. 2(b)), yielding a combination of 12 groups, and a total of 192 beads. All data are stored in a database indexed by bead numbers. The parameters used are listed in Table 1.

2.3 Feature registration

Feature registration to the same ST coordinates is essential because the collected data have different sampling rates and timestamps, requiring resampling and temporal alignment to create a coherent dataset. During the DED process, the co-axial camera captures top-view images of the melt pool at a time interval of τ (Fig. 2(c)). The melt pool images recorded for the entire bead can be represented as $X_0 \in \mathbb{R}^{N_o\times H_o\times W_o}$, where N_0 is the number of images, and H_o and W_o are the height and width of the images in pixels. The timestamps corresponding to X_0 are described as $T_0 \in \mathbb{R}^{N_o}$. Concurrently, all positional data collected for the same bead are $S_c \in \mathbb{R}^{N_c\times 3}$, with the subscript c denoting the controller, 3 representing the position in the x, y, and z axes, and N_c being the number of signals collected from the motion controller. The associated timestamps are $T_c \in \mathbb{R}^{N_c}$. Given that T_0 and T_c share identical starting and ending points on the temporal axis, for any given melt pool image $x_i \in \mathbb{R}^{H_o\times W_o}$, we can ascertain the corresponding printing point location $s_i \in \mathbb{R}^3$ through spatial interpolation as depicted in Fig. 2(c). The resultant spatial coordinates are $S_0 \in \mathbb{R}^{N_o\times 3}$. After the DED process is completed, the structured light camera records the bead geometry in the format of point clouds. Similarly, the original point clouds are sliced according to the position S_0, and the geometry label $Y_0 \in \mathbb{R}^{N_0\times 3}$ contains cross-sectional width, height and area of the bead are calculated (Fig. 2(d)). To sum up, the collection of ST coordinates pertaining to the dataset can be defined as the concatenation of $[T_0, S_0, X_0, Y_0]$, signifying that each element encapsulates the timestamp, the location in the machine coordinate system, the co-axial melt pool image, and the cross-sectional geometrical labels.

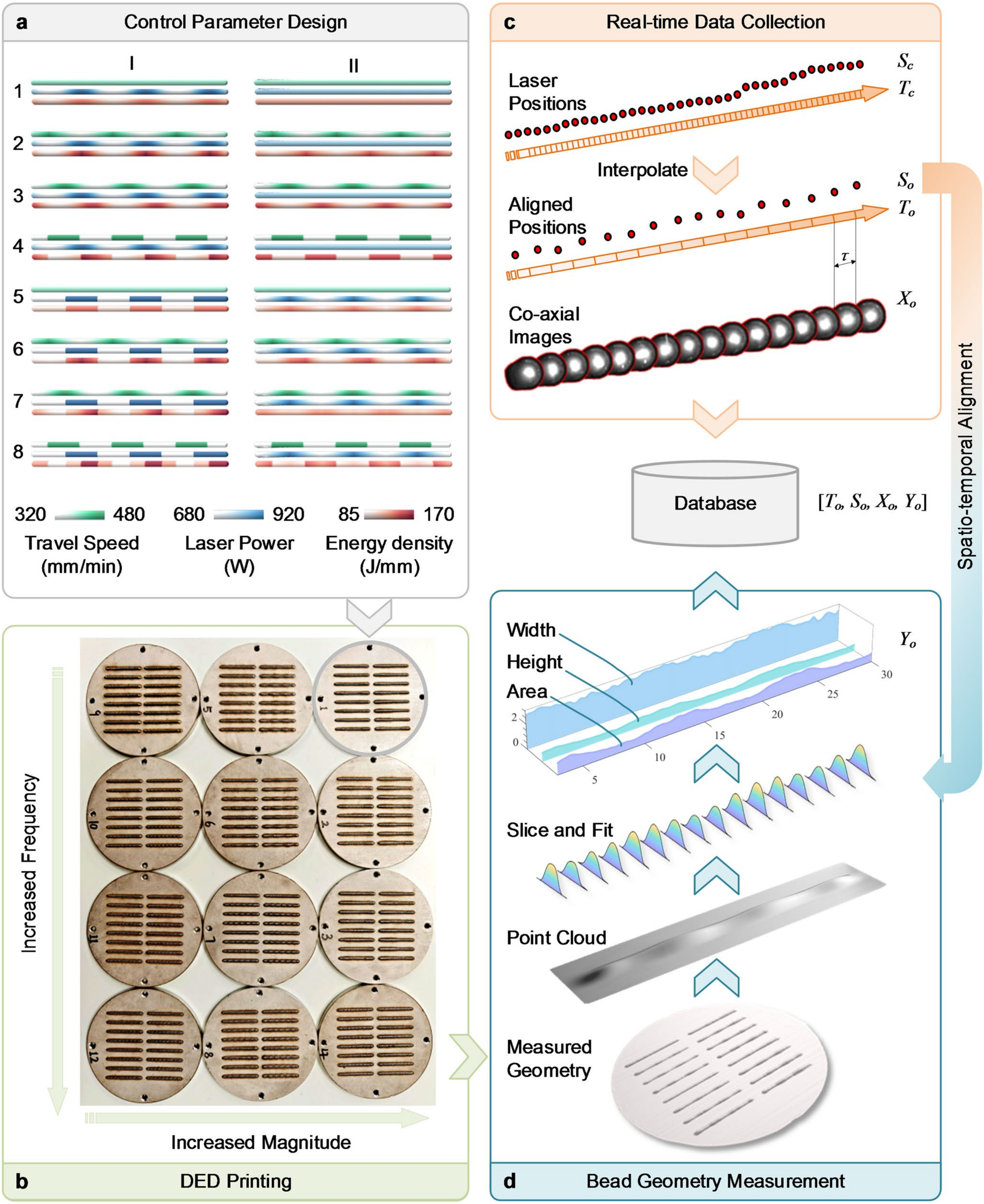

Fig. 2 Data collection processes. (**a**) Control parameters design for 16 combinations of travel speed and laser power; (**b**) DED printing of a total of 192 beads; (**c**) Real-time data collection and alignment; (**d**) Bead geometry measurement and data alignment

Table 1 Parameter options in creating datasets

Parameters	Options	Number of options
Laser power pattern	constant, sine, square, sawtooth	4
Travel speed pattern	constant, sine, square, sawtooth	4
Frequency*	very low: $\lambda = 30$; low: $\lambda = 10$; medium: $\lambda = 5$; high: $\lambda = 3.3$	4
Magnitude**	low: $P = 640 \pm 96$, $V = 320 \pm 64$; medium: $P = 800 \pm 120$, $V = 400 \pm 80$; high: $P = 960 \pm 144$, $V = 480 \pm 96$	3
Time interval (ms)	5, 10, 20, 40, 80, 120, 160, 200, 240, 280, 320, 360, 400	13
Sample size	100, 200, 300	3

* λ is the wavelength in mm; ** P is the laser power in W, and V is the travel speed in mm/min

2.4 Dataset design for building the STNN

To further examine the STNN models' ST response, sub-datasets were created from the database containing raw data (Fig. 3(a)-(b)), with all subsequent data preprocessing and model training procedures consistent across these datasets. Each item in the datasets includes N sequential melt pool images at equal time intervals of $\alpha\tau$, serving as input for the STNN models, with the bead geometry values as the output labels. The time intervals between melt pool images were varied to test the STNN models' response across different temporal domains (Table 1). A large interval (e.g. $\alpha > 48$, or $\alpha\tau > 240$ ms) means that two images are more separated in the ST domains, and an interval equal to 0 indicates that two images are 100% overlapping. To evaluate whether the response is dependent on dataset size, we varied the sample sizes to randomly fetch signals from each bead. Each dataset was randomly partitioned into training, validation, and test sets in an 8:1:1 ratio to facilitate cross-validation and prevent overfitting (Fig. 3(c)). For each training cycle, the training set is used to tune the model parameters, and the validation set is used to evaluate the models' performance. After the training-validation processes, the best model parameters as indicated by the validation set were saved. All results presented in this work are derived from the test set.

2.5 Architecture design of the STNN

Figure 3(d) illustrates the model pipeline for predicting geometrical features from chronological co-axial 2D melt pool images. The input $X_i \in \mathbb{R}^{N\times 512}$ of the model comprises features of size 512 from N images. For memory and computational efficiency, we selected N=5 in all our experiments. For instance, five images, sorted by temporal indices $\{i-2,\ i-1,\ i,\ i+1,\ i+2\}$, are used, where − 2 and − 1 indicate past times, 0 represents the current time, and 1 and 2 denote future times. ResNet-18 [45], an 18-layer CNN originally trained on the ImageNet dataset with 1.28 million images, functions as the spatial feature extractor. The images are first processed by the pre-trained ResNet-18, yielding $N \times 512$ spatial features, which are subsequently passed through a linear layer to reduce their dimensions to $h_i \in \mathbb{R}^{N\times H}$, where H is the hidden dimension. Following this, these features are processed by a spatio-temporal layer (STL) with three configurations:

1. **FC (fully connected)**: Consists of a linear layer that does not explicitly model sequential order or temporal dependencies, unlike recurrent neural networks. The output is computed as a weighted sum of the input features along the spatial dimension, serving as a non-sequential baseline.
2. **SA (self-attention)**: Employs self-attention mechanisms to focus on relevant parts of the input chronological images. At each time instance, the response is computed by attending to all time steps in the sequence, enabling the capture of long-range temporal dependencies and providing interpretability regarding the relative importance of different time instances.
3. **BL (BiLSTM)**: Employs a bi-directional long short-term memory (BiLSTM) network selected from the recurrent neural network (RNN) family to explicitly model sequential dependencies in the melt pool image sequence. By maintaining recurrent memory states and processing the sequence in both forward and backward directions, the BiLSTM captures temporal correlations arising from cumulative thermal history, making it a representative baseline for standard sequential modelling and a direct point of comparison with the simpler FC and SA configurations.

Each STL configuration preserves the dimensionality of the input feature sequence h_i. Following established multimodal representation learning practices [46, 47], raw melt pool images are transformed into low-dimensional embeddings that capture essential spatial features with high computational efficiency. A pre-trained ResNet-18 network is used as the image feature extractor, and its parameters are kept frozen during training. This common practice both stabilizes the spatial representation and enables a fair comparison among different STL configurations by isolating the effects of spatio-temporal modelling.

Furthermore, inspired by Zhang et al. [48] and Zilly et al. [49], a residual connection is incorporated by scaling the

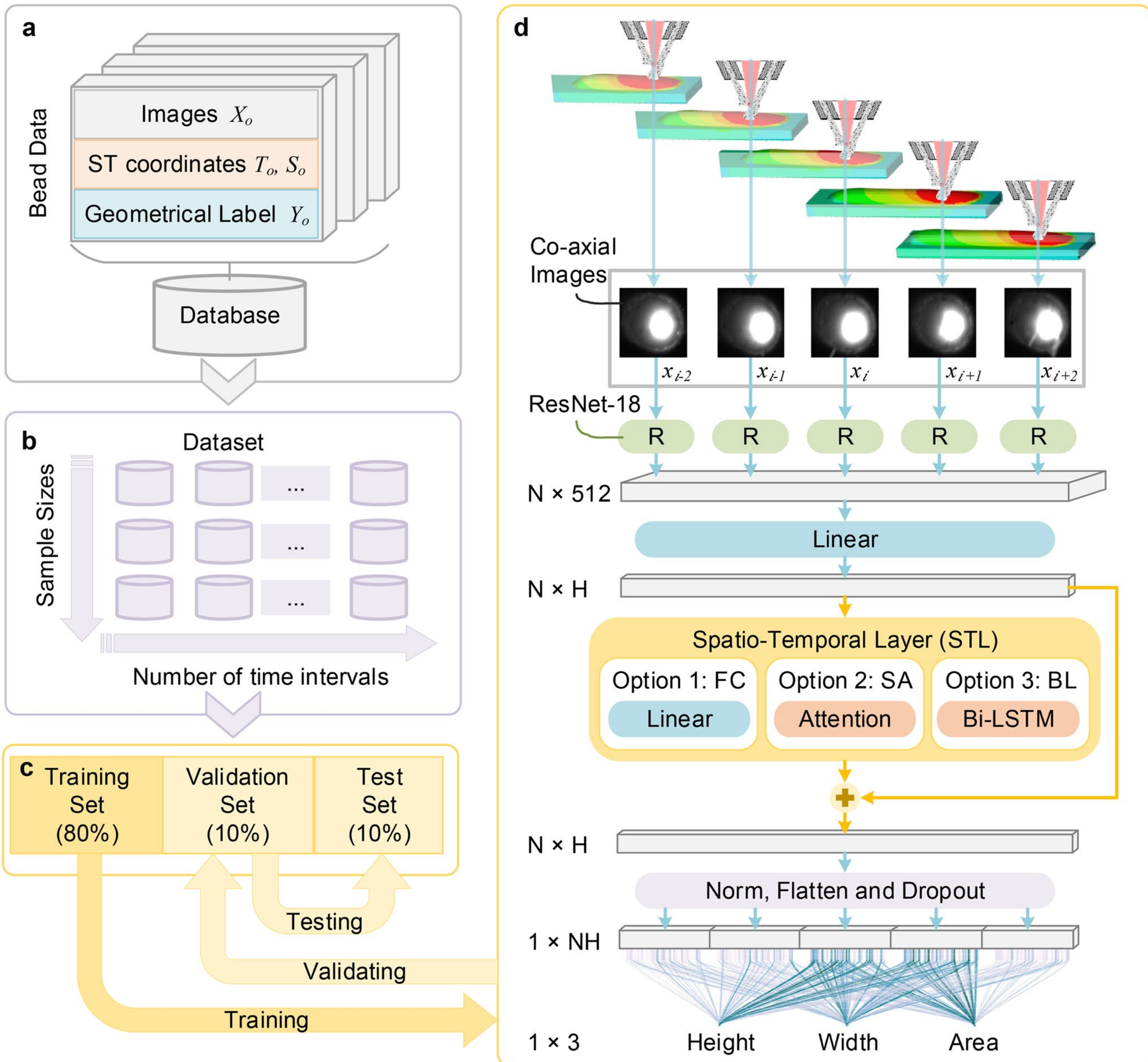

Fig. 3 Pipeline of the proposed STNN models. (**a**) Data collection to a database; (**b**) Creation of sub-dataset with different number of intervals and sample size; (**c**) Dataset splitting for training, validating, and testing; (**d**) Design of neural network structure, with three variants in the ST layer

output of the STL with a learnable parameter (γ) and adding it to the original STL input, as expressed by

$$h\prime_i = \gamma\, STL\,(h_i) + h_i. \tag{2}$$

.

Here, γ is a learnable scalar initialized to 0, h_i denotes the input feature vector, and $h\prime_i$ represents the final output feature vector after the residual connection. Introducing γ enables the network to initially focus on local cues; as training progresses, it gradually assigns more weight to the complex ST response, as shown in Fig. 11b. The STL output is normalized across the features using layer normalization [50]. The final output is a 1×3 vector representing the bead geometry (height, width, and area). This is achieved by flattening the output of the STL, followed by processing through dropout and a linear layer. For baseline comparison, a simpler linear model (F1) with $N = 1$ is used, which utilizes only one image as input, thereby lacking temporal interactions.

Metrics for evaluating the modelling performance.

The mean squared error (MSE) between the predicted height, width and area and the measured values is calculated as the loss function during the model training:

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y}_i)^2 \tag{3}$$

where y_i is the true value, $\widehat{y}_i$ is the predicted value, and n is the number of observations. To quantitatively compare the accuracy of different models, absolute percentage error (APE) is used to evaluate the error of each prediction, as given by

$$\mathrm{APE} = \left|\frac{y_i - \widehat{y}_i}{y_i}\right| \times 100\% \tag{4}$$

.

The mean absolute percentage error (MAPE) is used to evaluate the overall prediction error:

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \widehat{y}_i}{y_i}\right| \times 100\% \tag{5}$$

3 Experiments

In this section, we first briefly describe the setup used in our lab in Sect. 3.1. Next, we detail how the bead geometry is measured in Sect. 3.2. Section 3.3 covers the dataset preparation steps before model training. Finally, Sect. 3.4 provides details on the model setup and training process.

3.1 Experimental setup

For this study, we established a DED additive manufacturing and monitoring system consisting of a homebuilt DED printer, co-axial CCD camera, side-axial structured light camera, and off-axial infrared (IR) camera (Fig. 4). The DED printer deposits the bead while the motion controller records spatial and temporal coordinates at 1000 Hz. It uses Raycus RFL-C2000 laser (up to 6 kW) and a Precitec YC52 printing head with a 2 mm laser spot diameter. The powder is 316 L stainless steel with a size range of 75 to 150 µm, and the substrate is 304 stainless steel. The co-axial CCD camera (Daheng MER2-202-60GM/C, resolution of 640×480 pixels) captures top-view images of the melt pool at 200 Hz. The side-axial structured light camera (Deep Innovation M05G016-230) captures the bead geometry with an accuracy of 50 µm. After printing, the camera is positioned above the bead to acquire its point cloud. The Xiris VXIR-3000 shortwave infrared camera, mounted outside the printing chamber, observes the thermal map around the melt pool region, remaining fixed while the printing head moves. The camera measures temperatures from 350 to 3000 °C and has a resolution of 640×512 pixels (15 µm/pixel) at 100 Hz.

3.2 Measurement of bead geometry

Before printing, the structured light camera is calibrated so that the measured cloud data is aligned with the machine coordinate system, upon which a mesh is constructed based on the point cloud. By employing an on-machine Renishaw OMP40-2 ruby probe in conjunction with a structured light camera to measure standard spheres, the coordinate transformation between the camera coordinate system and the machine coordinate system can be accurately established. Subsequently, the channel mesh is clipped according to

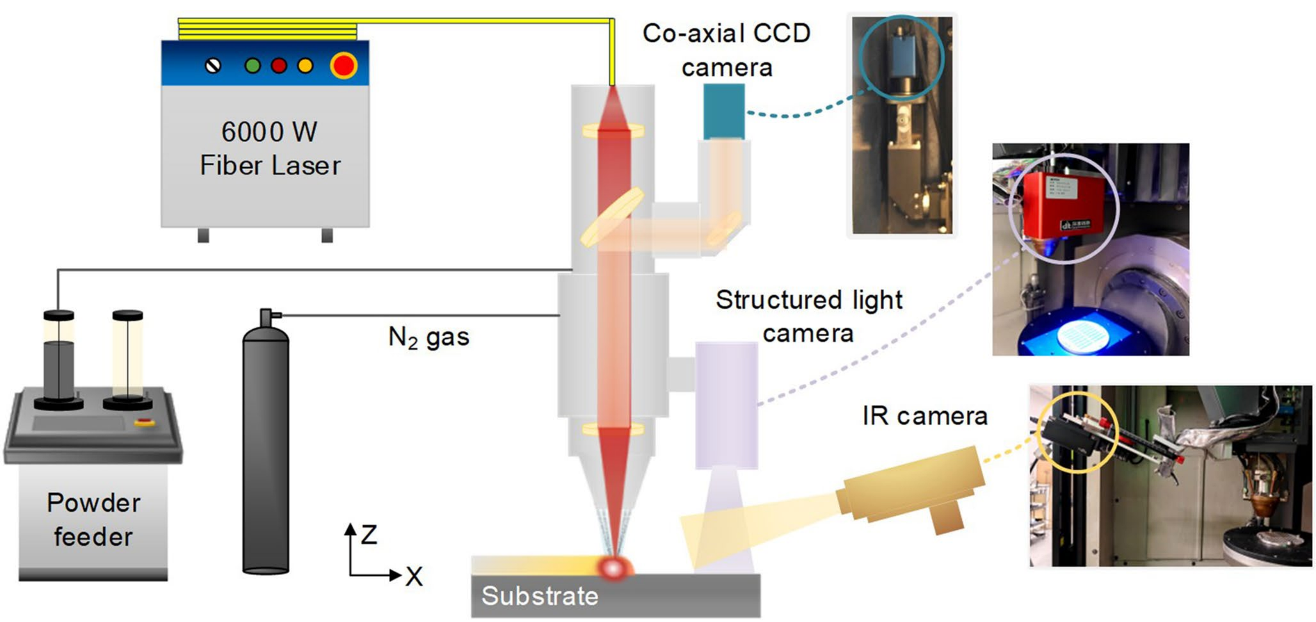

Fig. 4 The schematic diagram of the DED printing and monitoring system

the position corresponding to the melt pool image. For each cross-sectional shape, the width, height, and area are meticulously calculated, as illustrated in Fig. 5(a). Specifically, the width is defined as the distance across the cross-sectional shape at a height of 0.01 mm, the height is defined as the vertical distance from the substrate to the peak of the curve, and the area is defined as the enclosed area between the curve and the substrate. It is crucial to note that while the 0.01 mm threshold is situated extremely close to the substrate surface, this measurement is not directly applied to raw measurement points. Instead, the captured point cloud data is first utilized to construct a continuous mesh representing both the bead and the substrate. The width at the 0.01 mm height is then explicitly determined by applying a mathematical slicing plane to this reconstructed mesh surface. Consequently, this measurement serves as a robust and consistent baseline parameter derived from geometrical reconstruction, rather than being subject to raw data noise.

3.3 Preparation of dataset for building the STNN

Based on a database containing raw data from 192 beads, a total of 39 datasets were generated, composed of 3 sample size options and 13 interval options ranging from 50 to 400ms (Table 1). With each bead collecting 900 to 1100 melt pool images depending on the travel speed, the resulting dataset for the sample size of 100, 200, 300 contains approximately 20,000, 40,000, and 60,000 data items, respectively. Each dataset item contains five input images, each 640 × 480 pixels with a single-color channel, and corresponding labels comprising a 1 × 3 feature vector for bead geometry. Each grayscale image is converted to RGB, cropped to 320 × 320 pixels to focus on the region of interest, resized to 224 × 224 pixels to match ResNet-18's input size, and normalized using mean parameters [0.485, 0.456, 0.406] and standard deviation parameters [0.229, 0.224, 0.225]. The dataset is randomly partitioned into training, validation, and test sets in an 8:1:1 ratio with a random seed of 42 to ensure reproducibility. Note that data from three beads were excluded from the training and validation set, intended be used as a test of the best trained model (Fig. 8). Labels are standardized using the formula $label \leftarrow (label - mean) \ / \ std$, where mean and standard deviation (std) are calculated from the training set labels. A batch size of 16 is used to balance computational efficiency and training stability. For validation and testing, the batch size is set to 1.

3.4 Training of the STNN

The model architecture employs a ResNet-18 network as the feature extractor. The ResNet-18, pre-trained on the ImageNet dataset, is utilized with its final fully connected (FC) layer removed and its parameters frozen during the training phase to leverage the pre-learned features effectively. The ResNet-18 outputs 512 spatial features, which are reduced to a hidden dimension of H. These features are then processed through a STL, maintaining input and output sizes of $N \times H$. The STL has three variants: fully connected (FC) with a linear layer of the same input and output dimensions, self-attention (SA) with a single head, and bi-directional LSTM (BL) with a single layer. Following the STL, a linear layer with ReLU activation is applied, flattening and transforming the $N \times H$ input to a 1×3 output. This layer incorporates a dropout rate of 0.2 to mitigate overfitting. The activation functions and weight parameters of this final linear layer, sized $NH \times 3$, are saved for evaluation. Hidden dimensions of 10, 20, 40, 60, 80, 100, 120, 140, 180, 200 are applied for comparison.

Model training is conducted using PyTorch (version 2.3.0 + cu121) on a workstation equipped with a 13th Gen Intel Core i9-13900 K CPU and an NVIDIA GeForce RTX 4090 GPU with 24 GB of memory, spanning 80 epochs. The initial learning rate is set to 5×10^{-4}, with an adaptive learning rate strategy that employs exponential decay, reducing the learning rate by 5% per epoch for all STL configurations, as shown in Fig. 11a. The Adam optimizer [51] is utilized, configured without weight decay due to its negligible effect in this work. To prevent overfitting, the model weights from the epoch with the lowest MSE loss, as determined by validation performance, are saved. All results presented in this work are derived based on the test dataset using these saved

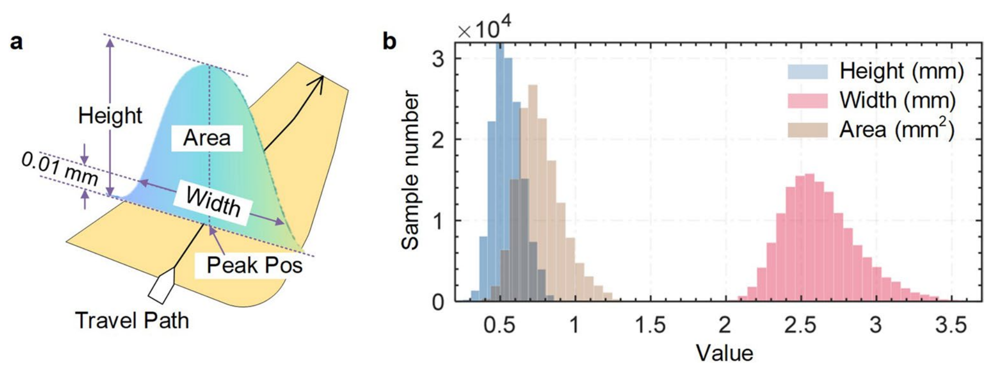

Fig. 5 Bead geometry definition and data distribution. (**a**) Definition of bead height, width (evaluated at 0.01 mm above the substrate), cross-sectional area, and peak position from a reconstructed bead cross-section; (**b**) Distribution of bead height, width, and area values over the full dataset

weights. Additionally, the predicted and true values for all data points are recorded to facilitate a comprehensive evaluation of the model's performance.

4 Results and discussion

In this results section, we first present the geometrical measurement results in Sect. 4.1. Next, we compare the prediction results of each model in Sect. 4.2. In Sect. 4.3, we analyse the ST response of each model and explain how time intervals affect the prediction results. Section 4.4 provides an ablation study to assess the importance of each layer in the STNN. Lastly, we evaluate the computational efficiency of each model in Sect. 4.4.

Results of geometrical measurement.

As illustrated in Sect. 2.2, 12 substrates, each containing 16 beads with control parameters specified in Fig. 2(a), were fabricated using varied frequency and magnitude options (Table 1), producing a total of 192 beads (Fig. 2(b)). The geometry (width, height, and area) labels were calculated from the corresponding point clouds by the method described in Sect. 3.2, yielding a total acquisition of 146,660 data points. The resultant data distribution of the geometry labels is plotted in Fig. 5(b), revealing an approximate normal distribution pattern for each label, highlighting the diversity and coherence in our dataset. More specifically, in Fig. 6, nine representative measurement results of the measured point clouds and calculated geometry are plotted. Overall, the correlation between bead geometry and EDR is most pronounced, particularly in the medium and low frequency situations. Conversely, at high frequencies, the spatial response to variations in input control parameters is minimal. Locally, significant oscillations are observed, underscoring the intricate nature of the process. Due to the complex multi-physics coupling involved in DED processes, determined by multiple input parameters and environmental factors, changes in bead geometry do not exhibit clear regularity despite consistent EDR variations.

4.1 Prediction performance of the STNN

The heatmap plot (Fig. 7(a)) displays the MSE loss of the prediction results across different STL configurations and time intervals. Larger sample numbers yield smaller MSE losses, with the F1 model exhibiting the highest loss, followed by the FC, SA, and BL models, the BL model having the lowest loss. The F1 model, serving as a baseline, shows higher losses across all intervals, emphasizing the benefits of incorporating temporal information. For the FC, SA, and BL models, an interval of around 120 ms consistently results in the lowest MSE values, indicating this interval's effectiveness in capturing the temporal dynamics of the melt pool. Figure 7(b) illustrates the MAPE of the prediction results across different intervals. The FC, SA and BL models consistently outperform the F1 baseline, particularly at the optimal interval of 120 ms. These findings highlight the significance of STL, particularly the BL configuration, in enhancing prediction accuracy, with temporal feature extraction substantially improving performance compared to the baseline.

Three beads from the test set were used to validate the model trained on a dataset with an interval of 120 ms and a sample size of 300, which yielded the best accuracy in previous analysis. An examination of the predicted width, height, and area of each data point on the channels by the model (Fig. 8(a)) reveals that the geometrical shapes predicted by the model closely follow the measured labels. However, the labels exhibit greater variations in amplitude, particularly in width prediction. Figure 8(b) shows the reconstructed geometry with colors indicating the APE of each model. For all models, the APE is better in regions with higher height, as seen in the reconstructed geometrical shapes. The model responds differently to frequency and magnitude, and performance degrades with high frequency and low magnitude. The prediction accuracy of STNN with the FC, SA, and BL configurations is superior to F1, which contains only one image input and lacks temporal information.

4.2 Investigation of spatio-temporal dynamics

Figure 9 illustrates the relationship between the absolute percentage error (APE) distribution and mean overlap rate for different models and time intervals. The overlap rate is the shared pixel number between two consecutive images divided by the total pixel number of the target image, as shown in Fig. 9(d). The mean overlap rate is the average overlap rate of the four surrounding images with respect to the current time instance.

Thermal images in Fig. 9(b) show the temperature distribution and characteristics of the melt pool, with boundaries defined by the melting temperature of 316 L stainless steel. These images reveal that the degree of spatial overlap between consecutive melt pool states varies systematically with the selected time interval. When the overlap rate is excessively high, consecutive images contain largely redundant thermal information, limiting the model's ability to learn meaningful temporal evolution. Conversely, as the time interval increases and the overlap rate decreases, adjacent melt pool states become increasingly separated in space, weakening the physical continuity between successive time steps. This reduced physical correlation diminishes the strength of spatio-temporal dependencies that the model can extract,

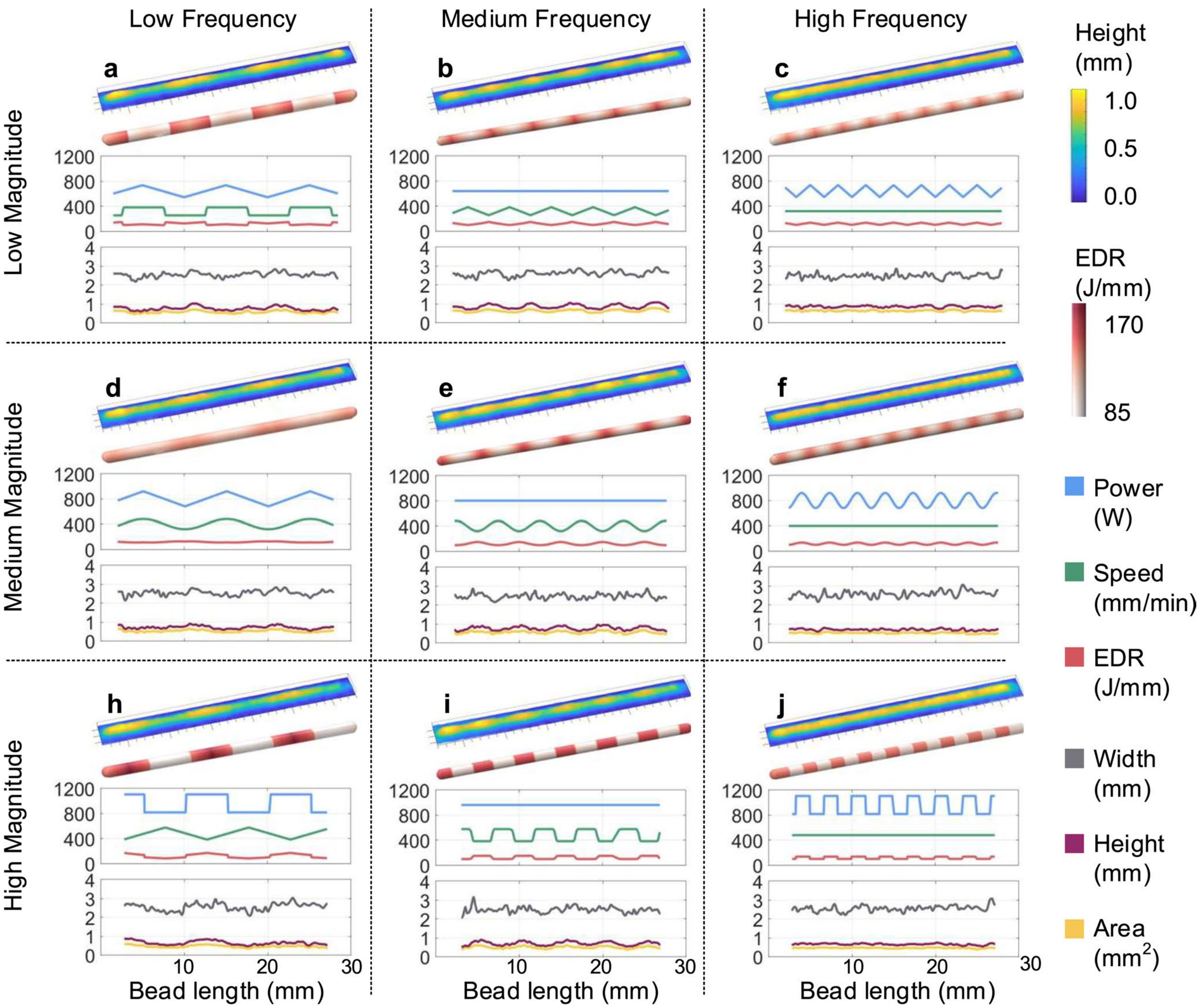

Fig. 6 Measured point clouds, input parameters, and calculated bead geometry (width, height, and area) for three different frequencies (low: $\lambda = 10$, medium: $\lambda = 5$, high: $\lambda = 3.3$) and three different magnitudes (low: $\bar{P} = 640$, $\bar{V} = 320$; medium: $\bar{P} = 800$, $\bar{V} = 400$; high: $\bar{P} = 960$, $\bar{V} = 480$). Here, λ is the wavelength in mm, $\bar{P}$ is the mean laser power in W, and $\bar{V}$ is the mean travel speed in mm/min

ultimately leading to degraded prediction accuracy. As shown in Fig. 9(a)-(b), when the interval is 5 ms, the five images span a size smaller than one melt pool length, resulting in strong redundancy. As the time intervals increase from 120 ms to 400 ms, consecutive images gradually separate and eventually exhibit little to no overlap. The quantitative results in Fig. 9(c) further show that, at a time interval of 120 ms, all STL configurations achieve their minimum prediction error at a mean overlap rate of approximately 0.6. These results suggest that neither overly large nor overly small overlaps are optimal. Instead, an intermediate overlap rate balances temporal continuity and informative variation, enabling effective capture of melt pool dynamics. Accordingly, an overlap rate of approximately 60%, corresponding to a time interval of 120 ms, yields the most accurate predictions.

Beyond finding that an intermediate time interval yields the best prediction accuracy, different time intervals have different impacts on the results. This is demonstrated by the activation status in the final layer of the STNN, as shown in Fig. 10(a). To access the average importance of each feature output of the STL layer, we use the mean activated weight (MAW):

$$\mathrm{MAW} = \frac{1}{n}\sum_{i=1}^{n} \left| W \cdot \mathrm{ReLU}\left(h\prime_i\right) \right|, \tag{6}$$

where $h\prime_i$ denotes the output feature vector of the STL, $W \in \mathbb{R}^{NH \times 3}$ is the weight matrix of the final linear

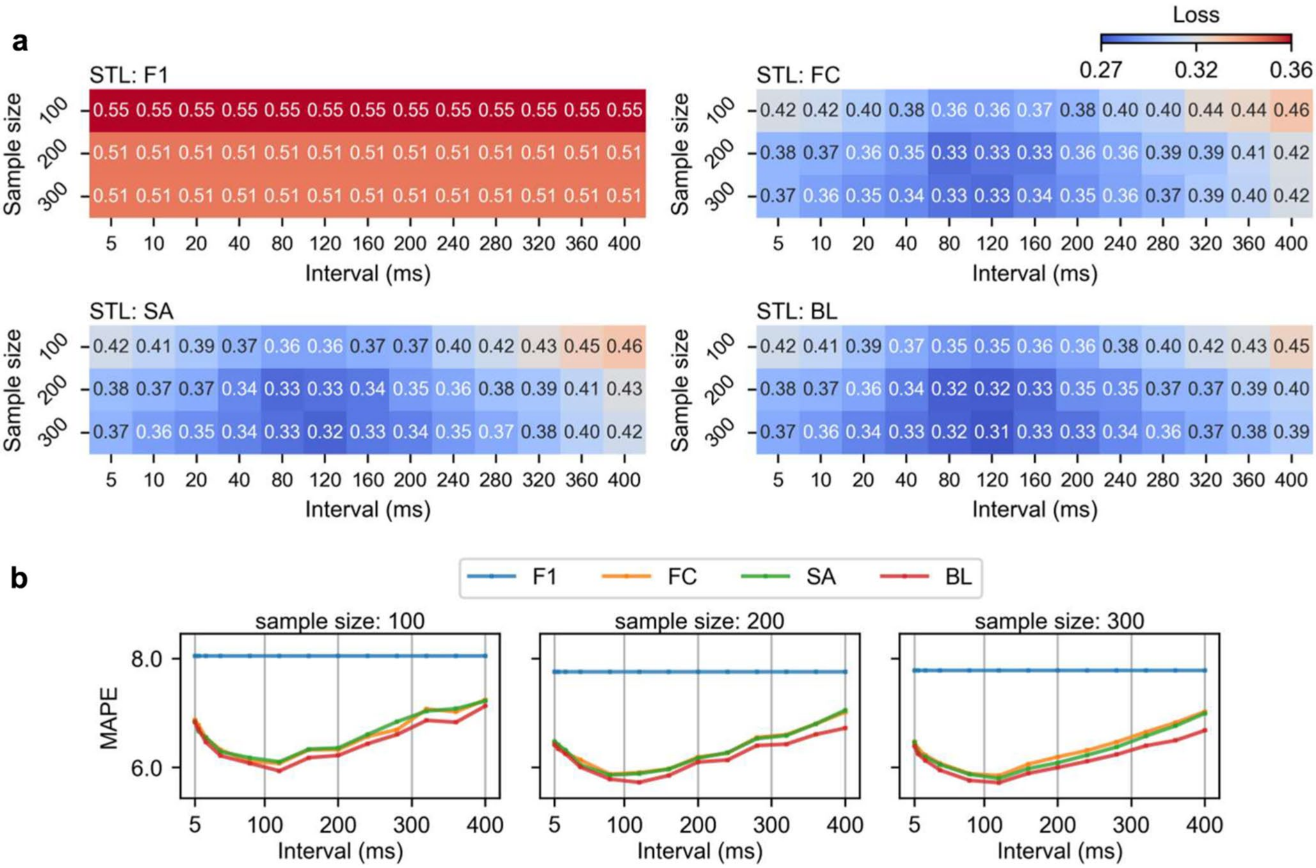

Fig. 7 Prediction accuracy of the proposed STNN with four STL configurations (F1, FC, SA, and BL), regarding different sample sizes and time intervals. (**a**) Heatmap of the MSE loss; (**b**) MAPE (lower values indicate better performance)

layer, ReLU is the rectified linear unit activation function defined as $\mathrm{ReLU(x)} = \max(0, \mathrm{x})$, and n is the number of time instances. In the neuron connection plot on the left of Fig. 10(b), colored edges represent the neuron connections, and scatter dots indicate the magnitude of the activated weights. Higher values indicate a greater contribution to the results. When the interval is small (e.g., 5 ms), all neurons are equally activated, indicating similar contributions from all-time instances. When the interval is very large (e.g., 400 ms), contributions primarily come from central neurons, indicating a dominant impact of the current time on results. When the interval is optimal (e.g., 120 ms), more neurons on the future side are activated, significantly contributing to the results. This highlights the greater impact of future instances on prediction accuracy. As shown in Fig. 10(b), the melt length is longer at the rear side than at the front side due to the traverse speed of the laser head. Additionally, below the melting temperature, larger regions exist above the phase shift temperature of the steel at the rear side of the melt pool. Therefore, future instances have a greater impact than the past because the thermal distribution around the melt pool area is asymmetric.

The thermal asymmetry contributes to data asymmetry, affecting how features from different time intervals are weighted. This is further illustrated in Fig. 10(c)-(e), where the data distribution of activated weights is plotted across five time instances $\{i-2,\ i-1,\ i,\ i+1,\ i+2\}$, with the x-axis representing various time intervals. The trends for height, area, and width are similar. For time instances $i-2$ and $i-1$, the overall values across time intervals are small and indistinguishable, indicating that the effects of past time instances are relatively minor. For the time instance i, the impact increases with the interval, suggesting that the effect of the current image becomes more significant as the time interval increases. For time instances $i+1$ and $i+2$, the impact first rises and then drops. However, the peak is more delayed for instance $i+1$ compared to $i+2$, indicating that closer images have a longer impact range because they are nearer to the center image. This aligns with previous heatmap findings, demonstrating that an interval of 120 ms effectively balances temporal dynamics, captures relevant ST features and ensures reliable feature extraction. This further supports the finding that an intermediate interval, e.g. 120 ms, provides the optimal balance for effectively capturing temporal information.

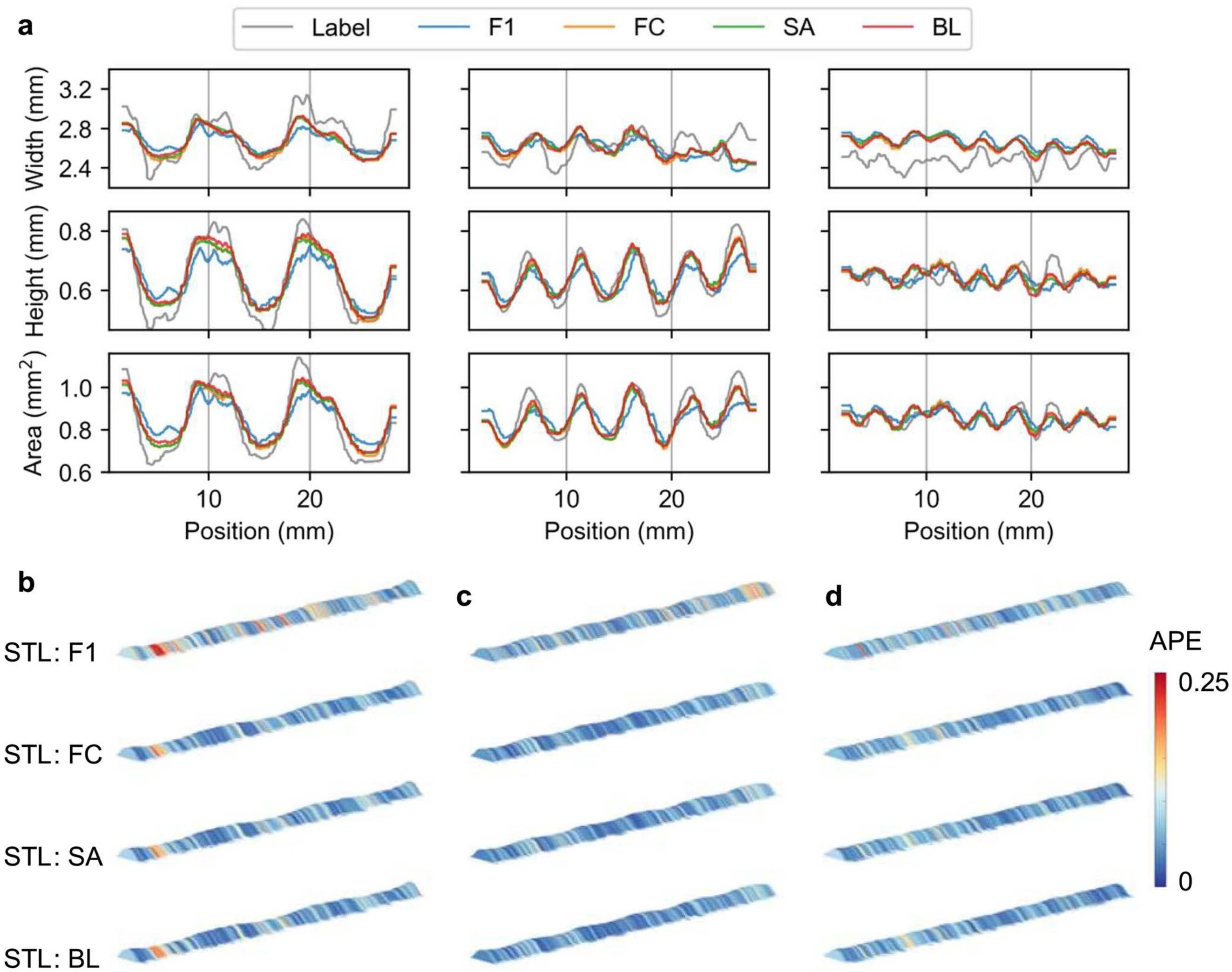

Fig. 8 The predicted geometry (width, height and area) for three beads using the best model trained from dataset with an interval of 120 ms and a sample size of 300. (a) The predicted values versus the true labels. The parameters from left to right are: frequency: low, medium, high; magnitude: low, low, low; (b-d) Reconstructed geometry with colors indicating the APE of each model (lower values indicate better performance)

4.3 Ablation study and computational efficiency of the proposed STNN

We conducted a combined analysis of the ablation study and computational efficiency to evaluate the impact of different model configurations on performance and resource utilization. Figure 11(b) indicates that as training progresses, the contribution from the STL increases and stabilizes, resulting in a consistent decrease in test loss. The trends of loss reduction during the training with respect to epoch are depicted in Fig. 11(c), with all training halted after 80 epochs. The FC configuration produces the best MAPE, whereas the F1 configuration with only one image input exhibits the worst performance. Figure 11(e) presents the loss from the best training results for hidden dimension H varying from 20 to 200. For H<150, the loss decreases with increasing hidden dimensions; however, for larger H, the loss increases, likely due to overfitting. Furthermore, the training time per epoch remains relatively constant across hidden dimensions, with the SA configuration having the highest time and the FC configuration the lowest. These results indicate that increasing H significantly is unnecessary, and the BL configuration is the optimal choice for the STL layer.

To determine which layers are critical for prediction performance, we performed an ablation study on each layer of the proposed STNN model, as shown in Table 2. This study was conducted on the dataset yielding the highest accuracy (with a time interval of 120 ms and sample size of 300). We removed the dropout layer, residual connection, layer normalization, and STL layer, respectively. The results demonstrate that the "Full" configuration consistently achieves the lowest MAPE across all hidden dimensions, underscoring

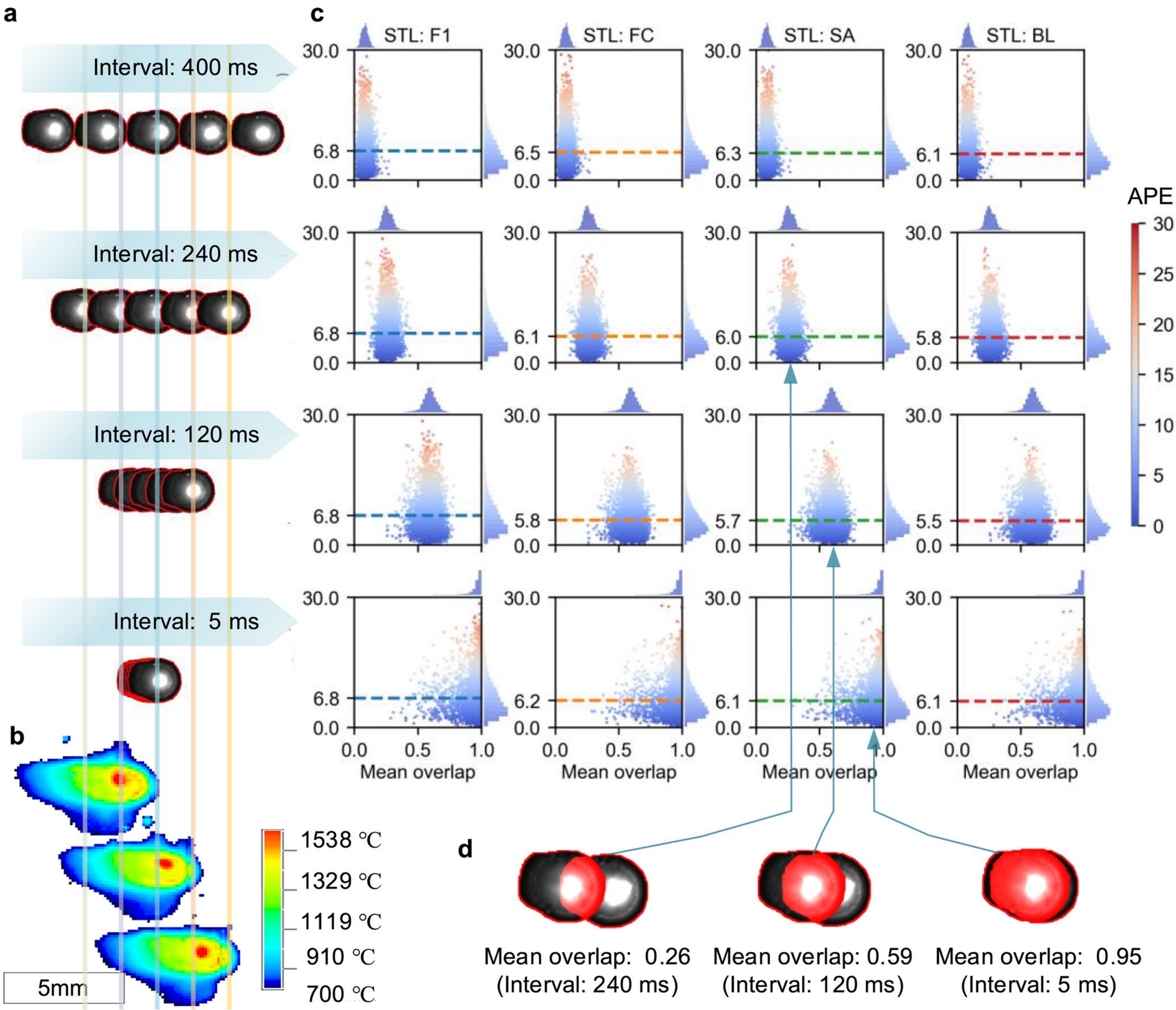

Fig. 9 Relationship between overlap, time interval, and prediction errors. (**a**) Co-axial melt pool images showing overlap conditions under different time intervals; (**b**) Thermal heatmaps of melt pool shown at the same spatial scale as the co-axial images; (**c**) Absolute percentage error (APE) as a function of mean overlap rate for different STL configurations; (**d**) Schematic illustration of the mean overlap rate corresponding to three representative time intervals

the necessity of including all layers (STL, Dropout, Norm, Residual) for optimal performance. For the BL configuration, which achieves the highest accuracy, the STL layer's impact is particularly significant; its removal leads to the largest increase in MAPE. Removing Dropout, Norm, or Residual layers results in a moderate performance drop. Notably, the residual connection substantially improves the accuracy of the SA configuration. The best performance for H = 20 is a MAPE of 5.68, for H = 50 it is 5.55, and for H = 80 it is 5.50. Interestingly, increasing the hidden dimension from 50 to 80 results in a performance drop, as evidenced by a higher MAPE and further supported by Fig. 11e. These findings emphasize the importance of retaining all components of the STNN model for optimal prediction accuracy while balancing model complexity and computational efficiency.

4.4 Intelligent layer height regulation via predictive feedback in DED printing

Layer height control represents a critical technological challenge in DED printing, where achieving geometric fidelity is paramount to ensuring component dimensional accuracy. However, in-situ layer height measurement presents significant challenges due to intense laser-induced light interference and elevated thermal conditions inherent to the process. Consequently, our proposed STNN model offers an

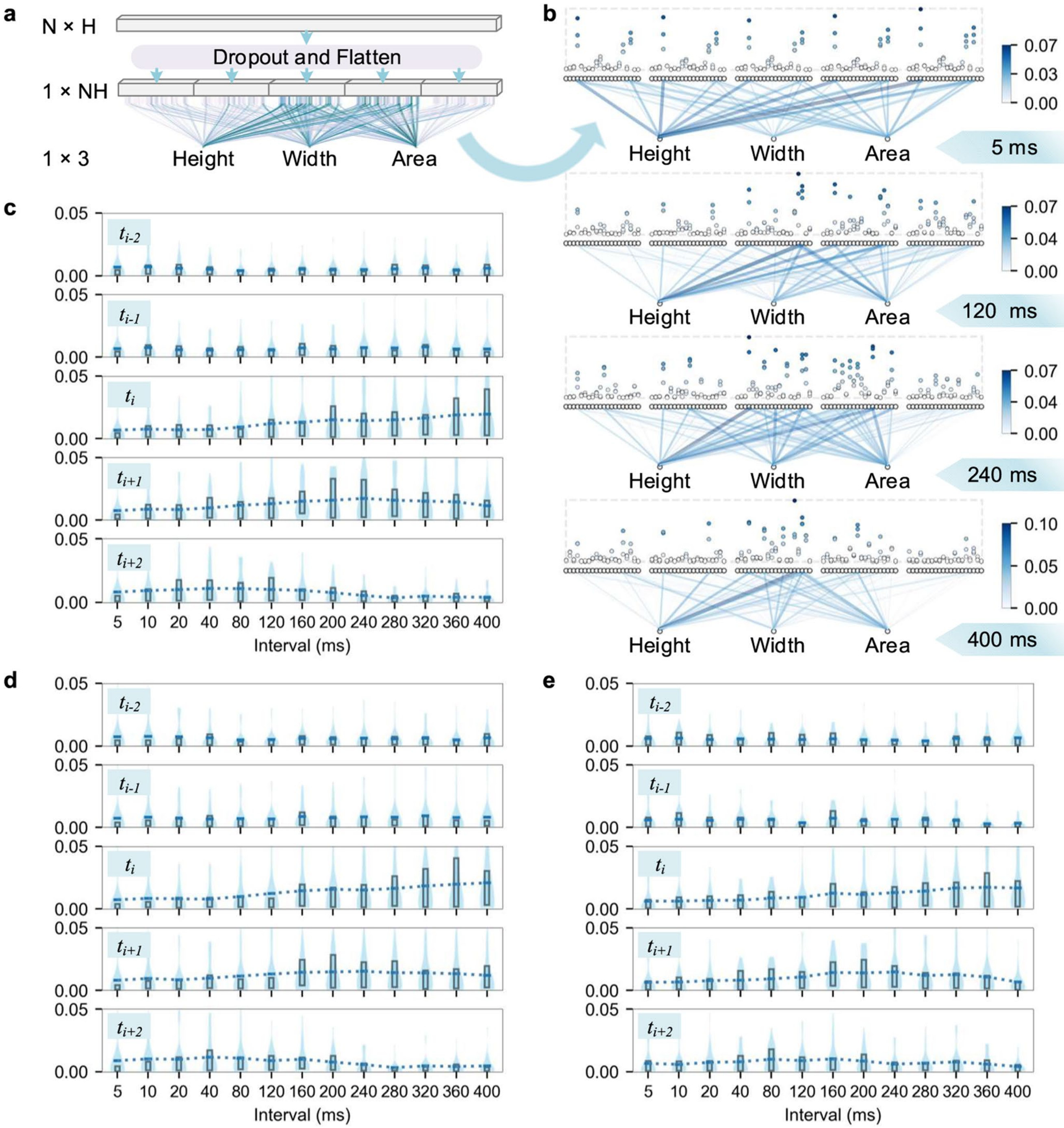

Fig. 10 Neuron activation status for different time intervals. (**a**) The final linear portion of the STNN model; (**b**) The neuron connection with color indicating activation statuses of the model with SA configuration and time interval equalling 5, 120, 240 and 400 ms; (**c**-**e**) The trends of activated weight for the model with SA configuration, with (**c**), (**d**) and (**e**) corresponding to the attribution to area, height, and width

accurate layer height prediction solution for such demanding environments, establishing an effective predictive feedback mechanism in DED printing.

A systematic comparative framework was developed to quantitatively assess performance differences between conventional open-loop control and our predictive feedback control, with vertical dimensional consistency serving as the primary quality metric. Experimental validation was conducted using standardized cube shell specimens (Fig. 12(a)) fabricated under identical printing paths (Fig. 12(b)). As evidenced in Fig. 12(c), the open-loop control group exhibited significant surface irregularities, manifesting as both

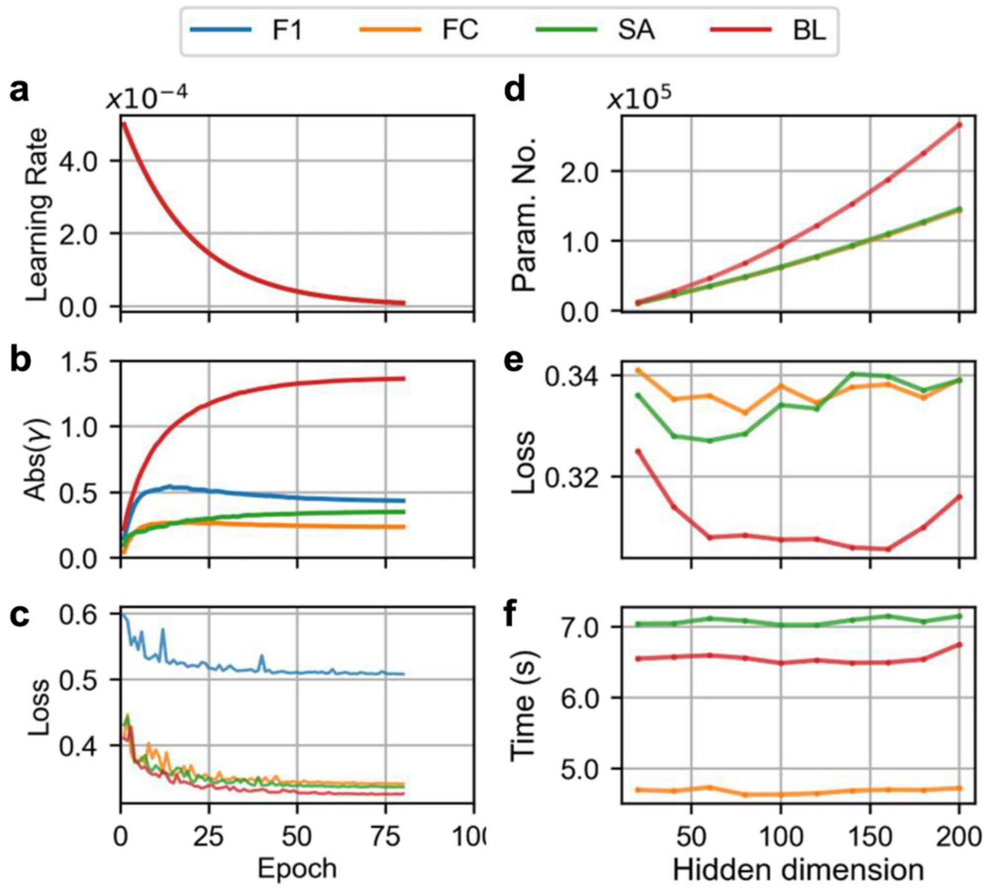

Fig. 11 Training and validation behaviors for epochs and hidden dimensions. (**a**-**c**) For epochs in training on a dataset with a time interval of 120 ms and a sample size of 300: (**a**) Adaptive learning rate; (**b**) Absolute value of trainable scalar in residual connection, gamma (γ); (**c**) Validation set loss. (**d**-**f**) For different hidden dimensions: (**d**) Total parameter number; (**e**) Test set loss; (**f**) Training time per epoch

convex and concave deformations at the top surface. In contrast, specimens produced with the predictive feedback control group maintained superior geometric fidelity, demonstrating a near-ideal planar surface profile.

Open-loop control maintains constant process parameters throughout printing, lacking real-time adaptation. The proposed predictive feedback system enables dynamic process control in DED through co-axial CCD–based melt pool monitoring. Captured melt pool images are processed by the STNN to estimate layer deposition height in real time. The predicted height is continuously compared with the target geometry, and the resulting deviation is provided to a PID controller for regulation. To compensate for height errors, the controller dynamically modulates the laser travel speed via real-time feed-rate override adjustments based on the PID output, rather than by generating or modifying static G-code. To preserve immediate responsiveness and maintain tight control loops, no post-processing smoothing filters are applied to the control signal. Through automated adjustment of laser travel speed, the system achieves precise dimensional control of deposited layer heights, as shown in Fig. 12(d).

Under predictive feedback control, the laser travel speed dynamically adjusts as shown in Fig. 13(a), while open-loop control maintains a constant speed. The resulting height error distributions of the top surface for both methods are compared in Fig. 13(b). The open-loop control yields a mean absolute error (MAE) of 0.4875 mm and a root mean square error (RMSE) of 0.5979 mm. In contrast, predictive feedback control significantly reduces these errors, achieving a 56.82% lower mean absolute error (MAE = 0.2105 mm) and 57.05% lower root mean square error (RMSE = 0.2568 mm) compared to open-loop operation. These results demonstrate the substantial improvement in dimensional accuracy enabled by predictive feedback control. Moreover, the accurate bead geometry prediction based on STNN plays a crucial role in enhancing process control in DED printing.

5 Conclusion

This research introduces the development and validation of a STNN model designed to predict bead geometry in DED additive manufacturing. By utilizing real-time melt pool images, our model adeptly captures the dynamic and nonlinear characteristics of the DED process, significantly enhancing prediction accuracy compared to methods that rely solely on single image inputs. By highlighting the potential of STNN in advancing the precision and reliability of bead geometry prediction in additive manufacturing, the

Table 2 Results of ablation study. The bold font indicates the results with the best MAPE

Hidden dimension	H = 20				H = 50				H = 80			
STL Configuration	F1	FC	SA	BL	F1	FC	SA	BL	F1	FC	SA	BL
STL Param. No.	421	421	1,681	2,561	2,551	2,551	10,201	15,401	6,481	6,481	25,921	39,041
Total Param. No.	10,784	11,024	12,284	13,164	28,454	29,054	36,704	41,904	47,924	48,884	68,324	81,444
STL Params %	3.9	3.8	13.7	19.5	8.97	8.78	27.8	36.8	13.5	13.3	37.9	47.9
Full STNN	**7.75**	**5.82**	**5.79**	**5.68**	**7.66**	**5.77**	**5.70**	**5.55**	**7.61**	**5.76**	**5.72**	**5.50**
No dropout	7.75	5.88	5.84	5.75	7.73	5.91	5.84	5.71	7.66	5.87	5.92	5.76
No norm	7.78	5.85	5.80	5.72	7.72	5.80	5.81	5.61	7.69	5.81	5.74	5.61
No residual	7.75	5.85	5.99	5.70	7.65	5.77	6.01	5.56	7.63	5.75	6.04	5.56
No STL	7.79	5.84	5.84	5.84	7.68	5.78	5.78	5.78	7.69	5.77	5.77	5.77

research paved the way for more efficient and effective real-time control of the DED process.

The key contributions of this study are as follows:

1. Accuracy Enhancement with STNN: The proposed STNN, particularly with the BiLSTM configuration, outperformed traditional models that depend on single image inputs in predicting bead geometry. By capturing both spatial and temporal dependencies, the STNN effectively models the dynamic behavior of the melt pool. Our datasets were carefully crafted to stimulate the ST response, ensuring a comprehensive representation of the DED process's temporal dynamics. This involved varying laser power, travel speed, and deposition patterns to generate a diverse set of melt pool images reflecting the process's inherent variability.
2. Optimal Time Interval for ST Response Prediction: Through extensive analysis, we determined that an intermediate time interval, approximately 120 ms, optimizes prediction accuracy. This interval balances the temporal dynamics, capturing relevant ST features and ensuring reliable feature extraction. A mean overlap rate of around 0.6 was identified as most effective, indicating that neither excessively high nor low overlap yield optimal results.
3. Asymmetrical Impact of Temporal Instances: Our study reveals the asymmetrical impact of past and future instances on melt pool morphology. Future time instances were found to exert a greater influence on prediction accuracy due to the asymmetric thermal distribution surrounding the melt pool. This thermal asymmetry leads to data asymmetry, affecting the weighting of features from different time intervals. This insight is crucial for refining real-time control strategies in DED processes. The observed temporal asymmetry may be qualitatively interpreted through the underlying thermophysical dynamics of the DED process. In the sequence of melt pool evolution, former time instances primarily reflect the initial interaction between the heat source and the relatively colder substrate or previous layer. A significant portion of the thermal input at this stage is consumed to overcome thermal inertia and establish a stable melt pool, rendering their thermal signatures less directly indicative of the final geometric variations. Conversely, future time instances, particularly those characterizing the tail region surrounding the melt pool, capture the interaction with previously deposited material that remains at elevated temperatures. This significant in-situ preheating effect alters local thermal boundary conditions and solidification dynamics, lowering the energy barrier required for subsequent metal powder to be captured and fully fused into the evolving track. Consequently, thermal signals

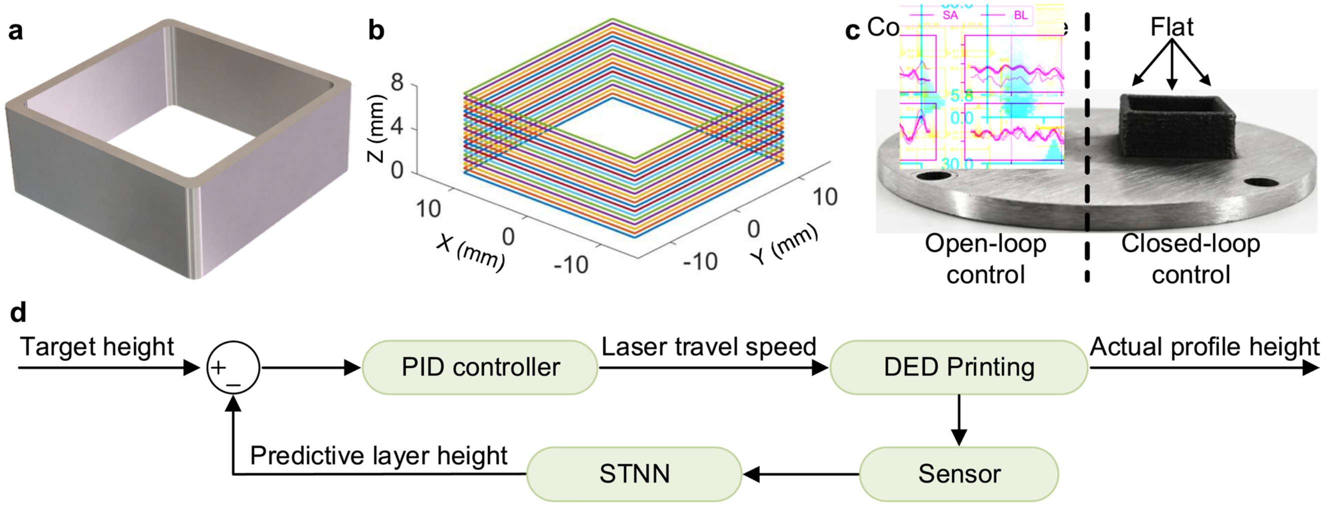

Fig. 12 The evaluation framework between open-loop control and predictive feedback systems: (**a**) cube shell model; (**b**) printing path; (**c**) surface morphology comparison of fabricated specimens; (**d**) schematic of the predictive feedback control system architecture

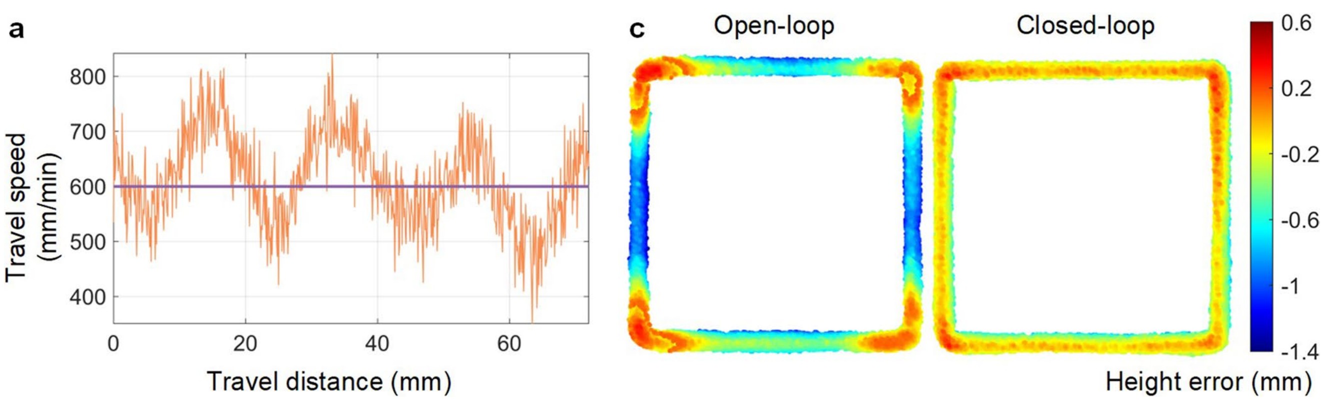

Fig. 13 Comparative analysis between the open-loop control and the predictive feedback control: (**a**) laser travel speed during final layer deposition; (**b**) dimensional accuracy of surface topography

reflecting this preheated state at the tail and the ongoing material capture process hold more direct and critical predictive power regarding the final formed geometry.

4. Improved printing accuracy control: The proposed STNN model generates predictive bead geometry, enabling dynamic closed-loop control of the printing process. In the instance of this study, deposition heights are estimated using the STNN, and deposition errors are derived by comparing these estimates with predetermined target values. Subsequently, a PID controller dynamically adjusts the laser travel speed to compensate for the deposition errors and precisely regulate the layer height. This approach achieves a 56.82% reduction in mean absolute error (MAE) and a 57.05% reduction in root mean square error (RMSE) compared to conventional open-loop control systems.

Despite these advancements, several limitations were identified. Feature extractor: The image feature extractor is kept frozen during training to enable fair comparison among spatio-temporal modelling strategies; fine-tuning on domain-specific melt pool images may further enhance spatial representation and prediction accuracy and will be explored in future work. Temporal Dynamics: The model currently uses only five images to maintain computational efficiency, which may restrict the capture of the complete temporal dynamics of the melt pool process. Temporal Encoding: The lack of explicit time stamps or positional information in the input data may limit the temporal modelling capabilities of configurations other than BiLSTM. Process Parameters: The effects of model parameters such as power, speed, and their rates of change were not explicitly discussed, with the assumption that this information is embedded in the melt pool images. Single-Track Deposition: The study focused on single-track deposition, while multi-track and multi-layer applications might present more complex dynamics.

Future research should address these limitations to enhance the model's accuracy and robustness. Investigating discrete features within melt pool images, exploring the ST effects of related physical parameters, and implementing the STNN in real-time industrial settings are promising directions for further study.

Acknowledgements This work is supported by supported by the Guangzhou-HKUST(GZ) Joint Funding Program (No. 2024A03J0680), Foshan HKUST Projects of China (No. FSUST20-SRI09E) and Guangdong S&T Program (No. 2025A0505000030).

Author contributions Yuanzhi Chen: Conceptualization, Methodology, Software, Formal analysis, Data curation, Writing – original draft. Wenze Zhang: Conceptualization, Methodology, Software, Investigation, Visualization, Writing – original draft. Yichen Wang: Conceptualization, Software, Validation, Visualization, Writing – original draft. Pengcheng Hu: Funding acquisition, Project administration, Writing – review & editing, Supervision. Molong Duan: Writing – review & editing, Supervision. Zichuan Fu: Conceptualization, Visualization. Xiaoke Deng: Conceptualization, Investigation, Methodology.

Data Availability Data will be made available on request.

Declarations

Competing interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Mbodj NG, Abuabiah M, Plapper P, El Kandaoui M, Yaacoubi S (2022) Modeling and control of layer height in laser wire additive manufacturing. Materials 15:4479
2. Xiong J, Zhang G, Gao H, Wu L (2013) Modeling of bead section profile and overlapping beads with experimental validation for robotic GMAW-based rapid manufacturing. Robot Comput Integr Manuf 29:417–423. https://doi.org/10.1016/j.rcim.2012.09.011
3. Wei HL, Cao Y, Liao WH, Liu TT (2020) Mechanisms on inter-track void formation and phase transformation during laser powder bed fusion of Ti-6Al-4V. Addit Manuf 34:101221. https://doi.org/10.1016/j.addma.2020.101221
4. Wei HL, Liu FQ, Liao WH, Liu TT (2020) Prediction of spatiotemporal variations of deposit profiles and inter-track voids during laser directed energy deposition. Addit Manuf 34:101219. https://doi.org/10.1016/j.addma.2020.101219
5. Huang Y, Khamesee MB, Toyserkani E (2019) A new physics-based model for laser directed energy deposition (powder-fed additive manufacturing): from single-track to multi-track and multi-layer. Optics & Laser Technology 109:584–99. https://doi.org/10.1016/j.optlastec.2018.08.015
6. Lv C, Wang J, Li H, Yin Q, Liu W, Shen S (2023) Dynamics of the laser–powder interaction in the Ti6Al4V powder feeding process of laser-directed energy deposition additive manufacturing. J Mater Res Technol 27:6376–6385. https://doi.org/10.1016/j.jmrt.2023.11.099
7. Sun Z, Guo W, Li L (2020) Numerical modelling of heat transfer, mass transport and microstructure formation in a high deposition rate laser directed energy deposition process. Addit Manuf 33:101175. https://doi.org/10.1016/j.addma.2020.101175
8. Wolff SJ, Gan Z, Lin S, Bennett JL, Yan W, Hyatt G et al (2019) Experimentally validated predictions of thermal history and microhardness in laser-deposited Inconel 718 on carbon steel. Addit Manuf 27:540–51. https://doi.org/10.1016/j.addma.2019.03.019
9. Vundru C, Singh R, Yan W, Karagadde S (2021) A comprehensive analytical-computational model of laser directed energy deposition to predict deposition geometry and integrity for sustainable repair. Int J Mech Sci 211:106790
10. Khorram A, Ghoreishi M, Yazdi MRS, Moradi M (2011) Optimization of bead geometry in CO2 laser welding of Ti 6Al 4V using response surface methodology. Engineering 3:708–712. https://doi.org/10.4236/eng.2011.37084
11. Han S-S, Kim H-K (2024) Artificial neural network-based sequential approximate optimization of metal sheet architecture and forming process. J Comput Des Eng 11:265–279. https://doi.org/10.1093/jcde/qwae049
12. Hilbig A, Vogt L, Holtzhausen S, Paetzold K (2023) Enhancing three-dimensional convolutional neural network-based geometric feature recognition for adaptive additive manufacturing: a signed distance field data approach. J Comput Des Eng 10:992–1009. https://doi.org/10.1093/jcde/qwad027
13. Shin S, Shin D, Kang N (2023) Topology optimization via machine learning and deep learning: a review. J Comput Des Eng 10:1736–1766. https://doi.org/10.1093/jcde/qwad072
14. Yang L, Li S, Li C, Zhu C (2024) Data-driven multi variate regression-based anomaly detection and recovery of unmanned aerial vehicle flight data. J Comput Des Eng 11:176–193. https://doi.org/10.1093/jcde/qwae023
15. Zhang Y, Moon SK (2021) Data-driven design strategy in fused filament fabrication: status and opportunities. J Comput Des Eng 8:489–509. https://doi.org/10.1093/jcde/qwaa094
16. Guo S, Agarwal M, Cooper C, Tian Q, Gao RX, Guo W et al (2022) Machine learning for metal additive manufacturing: towards a physics-informed data-driven paradigm. J Manuf Syst 62:145–63. https://doi.org/10.1016/j.jmsy.2021.11.003
17. He FY, Yuan L, Mu HC, Ros M, Ding DH, Pan ZX et al (2023) Research and application of artificial intelligence techniques for wire arc additive manufacturing: a state-of-the-art review. Robot Comput Integr Manuf 82:102525. https://doi.org/10.1016/j.rcim.2023.102525
18. Zhu K, Fuh JYH, Lin X (2022) Metal-based additive manufacturing condition monitoring: a review on machine learning based approaches. IEEE ASME Trans Mechatron 27:2495–2510. https://doi.org/10.1109/TMECH.2021.3110818
19. Cai Y, Wang Y, Chen H, Xiong J (2024) Searching optimal process parameters for desired layer geometry in wire-laser directed energy deposition based on machine learning. Virtual Phys Prototyp 19:e2352066. https://doi.org/10.1080/17452759.2024.2352066
20. Oh WJ, Lee CM, Kim DH (2022) Prediction of deposition bead geometry in wire arc additive manufacturing using machine learning. J Mater Res Technol 20:4283–4296. https://doi.org/10.1016/j.jmrt.2022.08.154
21. Cao X, Duan C, Luo X, Zheng S, Hao X, Shang D et al (2024) Physics-informed machine learning approach for molten pool morphology prediction and process evaluation in directed energy deposition of 12CrNi2 alloy steel. J Manuf Process 119:806–26. https://doi.org/10.1016/j.jmapro.2024.04.023
22. Yang Z, Verma AK, Smith L, Guzel A, Chen H, Pistorius PC et al (2022) Predicting melt pool dimensions for wire-feed directed energy deposition process. Integr Mater Manuf Innov 11:532–44. https://doi.org/10.1007/s40192-022-00278-z
23. Karniadakis GE, Kevrekidis IG, Lu L, Perdikaris P, Wang S, Yang L (2021) Physics-informed machine learning. Nat Rev Phys 3:422–440. https://doi.org/10.1038/s42254-021-00314-5
24. Jamnikar ND, Liu S, Brice C, Zhang X (2022) In-process comprehensive prediction of bead geometry for laser wire-feed DED system using molten pool sensing data and multi-modality CNN.

Int J Adv Manuf Technol 121:903–917. https://doi.org/10.1007/s00170-022-09248-3
25. Asadi R, Queguineur A, Wiikinkoski O, Mokhtarian H, Aihkisalo T, Revuelta A et al (2024) Process monitoring by deep neural networks in directed energy deposition: CNN-based detection, segmentation, and statistical analysis of melt pools. Robot Comput Integr Manuf 87:102710. https://doi.org/10.1016/j.rcim.2023.102710
26. Ribeiro KSB, Núñez HHL, Venter GS, Doude HR, Coelho RT (2023) A hybrid machine learning model for in-process estimation of printing distance in laser directed energy deposition. Int J Adv Manuf Technol 127:3183–3194. https://doi.org/10.1007/s00170-023-11582-z
27. Chen X, Yao M, Kong F, Fu Y, Wu J, Zhang H (2023) In-situ quality monitoring of laser hot wire cladding process based on multi-sensing diagnosis and machine learning model. J Manuf Process 87:183–198. https://doi.org/10.1016/j.jmapro.2023.01.031
28. Kim DB, Chong H, Mahdi MM, Shin S-J (2024) Data-fused and concatenated-ensemble learning for in-situ anomaly detection in wire and arc-based direct energy deposition. J Manuf Process 112:273–289. https://doi.org/10.1016/j.jmapro.2024.01.041
29. Gihr M, Rashid A, Melkote SN (2024) Bead geometry prediction and optimization for corner structures in directed energy deposition using machine learning. Addit Manuf 84:104080. https://doi.org/10.1016/j.addma.2024.104080
30. Wacker C, Köhler M, David M, Aschersleben F, Gabriel F, Hensel J et al (2021) Geometry and distortion prediction of multiple layers for wire arc additive manufacturing with artificial neural networks. Appl Sci 11:4694. https://doi.org/10.3390/app11104694
31. Donahue J, Hendricks LA, Rohrbach M, Venugopalan S, Guadarrama S, Saenko K et al (2017) Long-Term recurrent convolutional networks for visual recognition and description. IEEE Trans Pattern Anal Mach Intell 39:677–91. https://doi.org/10.1109/TPAMI.2016.2599174
32. Liu Q, Zhang Z, Guo P, Wang Y, Liang J (2023) Enhancing aircraft engine remaining useful life prediction via multiscale deep transfer learning with limited data. J Comput Des Eng 11:343–355. https://doi.org/10.1093/jcde/qwae018
33. Hochreiter S, Schmidhuber J (1997) Long Short-Term memory. Neural Comput 9:1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735
34. Graves A, Schmidhuber J (2005) Framewise phoneme classification with bidirectional LSTM and other neural network architectures. Neural Netw 18:602–610. https://doi.org/10.1016/j.neunet.2005.06.042
35. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN et al (2017) Attention is all you need. Proc. 31st Int. Conf. Neural Inf. Process. Syst., Red Hook, NY, USA: Curran Associates Inc.; pp. 6000–10
36. Scarselli F, Gori M, Ah Chung Tsoi, Hagenbuchner M, Monfardini G (2009) The graph neural network model. IEEE Trans Neural Netw 20:61–80. https://doi.org/10.1109/TNN.2008.2005605
37. Wei X, Cui W, Huang X, Yang L, Geng X, Tao Z et al (2024) Hierarchical RNNs with graph policy and attention for drone swarm. J Comput Des Eng 11:314–26. https://doi.org/10.1093/jcde/qwae031
38. Hu K, Wang Y, Li W, Wang L (2022) CNN-BiLSTM enabled prediction on molten pool width for thin-walled part fabrication using laser directed energy deposition. J Manuf Process 78:32–45. https://doi.org/10.1016/j.jmapro.2022.04.010
39. Yang X, Li Z, Cao L, Chen L, Huang Q, Bi G (2024) Process optimization and quality prediction of laser aided additive manufacturing SS 420 based on RSM and WOA-Bi-LSTM. Mater Today Commun 38:107882. https://doi.org/10.1016/j.mtcomm.2023.107882
40. Zhou Z, Shen H, Liu B, Du W, Jin J (2021) Thermal field prediction for welding paths in multi-layer gas metal arc welding-based additive manufacturing: a machine learning approach. J Manuf Process 64:960–971. https://doi.org/10.1016/j.jmapro.2021.02.033
41. Ren K, Chew Y, Zhang YF, Fuh JYH, Bi GJ (2020) Thermal field prediction for laser scanning paths in laser aided additive manufacturing by physics-based machine learning. Comput Methods Appl Mech Eng 362:112734. https://doi.org/10.1016/j.cma.2019.112734
42. Mozaffar M, Liao S, Lin H, Ehmann K, Cao J (2021) Geometry-agnostic data-driven thermal modeling of additive manufacturing processes using graph neural networks. Addit Manuf 48:102449. https://doi.org/10.1016/j.addma.2021.102449
43. Hong Y, He X, Xu J, Yuan R, Lin K, Chang B et al (2024) AF-FTTSnet: an end-to-end two-stream convolutional neural network for online quality monitoring of robotic welding. J Manuf Syst 74:422–34. https://doi.org/10.1016/j.jmsy.2024.04.006
44. Chen L, Bi G, Yao X, Tan C, Su J, Ng NPH et al (2023) Multisensor fusion-based digital twin for localized quality prediction in robotic laser-directed energy deposition. Robot Comput Integr Manuf 84:102581. https://doi.org/10.1016/j.rcim.2023.102581
45. He K, Zhang X, Ren S, Sun J (2015) Deep residual learning for image recognition. 2016 IEEE Conf Comput Vis Pattern Recognit CVPR :770–778
46. Guo W, Wang J, Wang S (2019) Deep multimodal representation learning: a survey. IEEE Access 7:63373–63394. https://doi.org/10.1109/ACCESS.2019.2916887
47. Vaezi Joze HR, Shaban A, Iuzzolino ML, Koishida K (2020) MMTM: multimodal transfer module for CNN fusion. 2020 IEEECVF Conf. Comput. Vis. Pattern recognit. CVPR, Seattle, WA, USA: IEEE;, pp 13286–13296. https://doi.org/10.1109/CVPR42600.2020.01330
48. Zhang H, Goodfellow I, Metaxas D, Odena A (2019) Self-attention generative adversarial networks. https://doi.org/10.48550/arXiv.1805.08318
49. Zilly JG, Srivastava RK, Koutník J, Schmidhuber J Recurrent highway networks 2017
50. Ba JL, Kiros JR, Hinton GE Layer normalization 2016
51. Kingma DP, Ba J, Adam (2017) A method for stochastic optimization