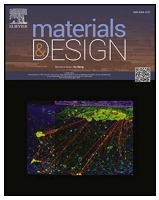
Materials & Design 258 (2025) 114546

Contents lists available at ScienceDirect

Materials & Design

journal homepage: www.elsevier.com/locate/matdes



# AI-based real-time prediction of cross-sectional phase distribution during laser heat treatment via sequential thermal imaging and image-to-image synthesis

Myeonggyun Son, Hyungson Ki *

*Department of Mechanical Engineering, Ulsan National Institute of Science and Technology (UNIST), 50 UNIST-gil, Ulsan 44919, South Korea*

ARTICLE INFO

*Keywords:*
Deep learning
Laser heat treatment
Infrared thermal imaging
Phase distribution
Process monitoring

ABSTRACT

Real-time monitoring of laser heat treatment is essential for ensuring microstructural consistency and process stability, but remains challenging due to complex thermal behavior and phase transformations. While deep learning has been applied to various laser-based processes, no prior study has addressed the real-time prediction of cross-sectional phase distributions in laser heat treatment. This study proposes the first AI-driven framework for real-time phase mapping during laser heat treatment of S45C steel, using sequential infrared surface temperature images. A convolutional gated recurrent unit architecture was developed to extract both spatial and temporal features, enabling pixel-wise image-to-image translation from thermal images to phase maps. Experimental data were collected from 16 experimental conditions by varying laser power and scanning speed. The model was trained using cross-sectional phase maps obtained from metallographic analysis, achieving high prediction accuracy across heat-affected, heat-treated, and melted regions. The model successfully captured key phenomena such as heat accumulation, geometric growth, and asymmetric phase evolution due to directional thermal gradients. This approach demonstrates a non-invasive, data-driven method for high-resolution monitoring of phase evolution in laser heat treatment. It offers new insights into thermally driven phase evolution and holds promise for integration into intelligent thermal processing and materials design strategies.

## 1. Introduction

Laser heat treatment is important because it enables precise, localized hardening with minimal thermal distortion, making it ideal for enhancing the surface properties of components with complex geometries. The effectiveness of this process is closely related to the hardenability of carbon steel, which refers to its ability to form martensite during cooling. A faster cooling rate promotes greater martensitic transformation, resulting in higher hardness and deeper hardened layers [1,2]. Since both heating and cooling rates are governed by process parameters such as laser power and scanning speed [3], precise control of these parameters is essential for achieving the desired microstructure. Consequently, researchers have extensively studied thermal monitoring techniques in laser heat treatment to improve process predictability and reliability.

Santhanakrishnan et al. [4] developed a process model based on temperature data obtained from a thermal monitoring camera during multi-pass laser heat treatment. Their study demonstrated that uniform case depth hardness is highly dependent on tempering temperature and cooling rate variations. Xu et al. [5] explored an infrared (IR) monitoring technique for laser heat treatment of ferrous alloys and identified a linear correlation between IR signal voltage and surface roughness, as well as hardening distribution. This finding confirmed the feasibility of real-time monitoring for process optimization. Similarly, Martínez et al. [6] proposed a closed-loop temperature control system for laser transformation hardening, integrating PID control and numerical simulations. Their results showed strong agreement between real-time monitoring data and simulated temperature trends. Furthermore, Martínez et al. [7] investigated scanner-based laser hardening of AISI 1045 steel, emphasizing the scanning speed as a key parameter affecting hardening layer depth. They also introduced a control-loop strategy for temperature monitoring based on scanning speed variations. Farshidianfar et al. [8] proposed a real-time thermal monitoring system for laser heat treatment of AISI 1020 low-carbon steel, demonstrating that microstructural characteristics and phase transformations correlate with real-time measurements of cooling rate, heating rate, and peak

* Corresponding author.
*E-mail address:* hski@unist.ac.kr (H. Ki).

https://doi.org/10.1016/j.matdes.2025.114546
Received 4 April 2025; Received in revised form 7 August 2025; Accepted 8 August 2025
Available online 8 August 2025

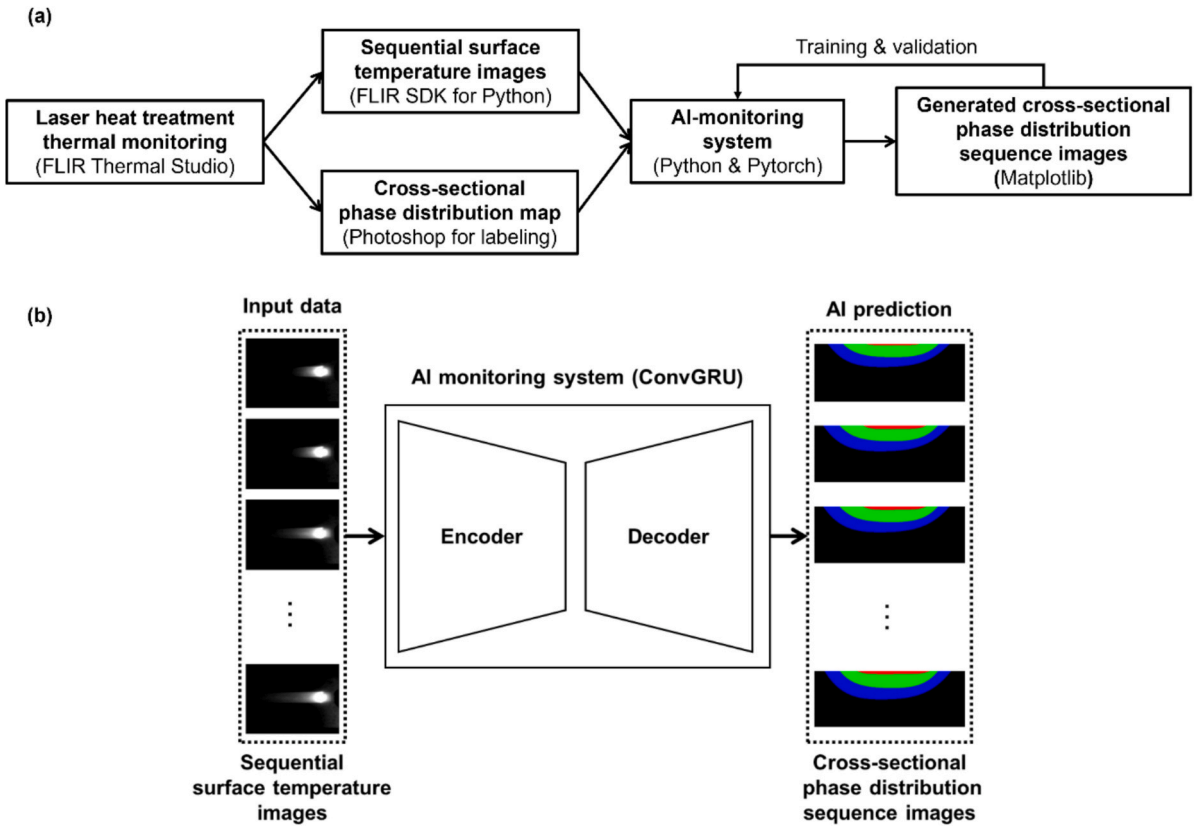



**Fig. 1.** Overall framework of the proposed AI-based monitoring system. (a) Data acquisition and processing pipeline, including surface temperature measurement (FLIR), ground truth labeling, model training using Python and PyTorch, and prediction output. (b) Structure of the ConvGRU-based AI model, which translates sequential surface temperature images into cross-sectional phase distribution maps.

temperature. Their findings further established that hardening depth is directly related to peak temperature, suggesting the potential for feedback control during laser heat treatment. Van Blitterswijk et al. [9] developed an adaptive thermal model capable of accurately predicting peak temperature and cooling rate during laser heat treatment. The model demonstrated robust performance under varying process conditions and was suitable for integration with closed-loop control systems. Furthermore, they introduced a model predictive control algorithm for real-time peak temperature regulation using in-situ measurements [10]. This controller effectively optimized scanning speed to achieve the target peak temperature, resulting in consistent hardness and hardening depth. While these physics-based models provide reliable surface-level thermal control, they are inherently limited in their ability to infer internal microstructural transformations such as phase evolution. As a result, conventional approaches often fall short in capturing the full complexity of thermally driven phase changes, which motivates the need for AI-based monitoring systems capable of learning spatiotemporal patterns and predicting internal microstructural states directly from thermal image sequences.

In recent years, AI-driven thermal monitoring methods have attracted increasing attention for their ability to enhance predictive accuracy in laser-based processes, including laser direct metal deposition [11] and selective laser sintering [12]. Generative adversarial network (GAN)-based models [13] have also been employed to predict hardness and phase distributions during the laser heat treatment process [14,15], demonstrating promising predictive performance. However, no thermal monitoring model has yet been reported specifically for laser heat treatment, and most previous AI applications in thermal monitoring have focused on classification tasks [11] rather than continuous process prediction. Furthermore, prior studies have primarily addressed surface defects or variations [12] that are observable on the surface. To the best of our knowledge, no previous work has demonstrated a real-time, frame-by-frame deep learning model capable of predicting cross-sectional phase distributions that are not directly observable during laser heat treatment.

In this study, we propose an AI-based monitoring system that predicts sequential cross-sectional phase formation from monitored surface temperature data during the laser heat treatment of S45C carbon steel. We applied real-time image-to-image synthesis, wherein sequences of surface temperature images were translated into cross-sectional phase distribution maps. The ground truth phase distribution maps were obtained from optical microscope (OM) images, while the sequential surface temperature images recorded during laser heat treatment served as the model's input.

Our AI model is based on a convolutional gated recurrent unit (ConvGRU) architecture, which combines convolutional neural networks (CNNs) with recurrent neural networks (RNNs) to effectively capture spatiotemporal dependencies. The proposed model successfully tracks phase transformations from surface temperature variations and accurately generates cross-sectional phase distribution maps.

Fig. 1 provides an overview of the complete framework and system architecture. As shown in Fig. 1(a), thermal images were acquired using FLIR Thermal Studio and processed via the FLIR SDK for Python, while the ground truth phase distribution maps were manually labeled from OM images using Adobe Photoshop. These paired data were used to train the AI model, which was implemented in Python using the PyTorch™ library. The resulting predictions were visualized using Matplotlib.

Fig. 1(b) illustrates the core architecture of the monitoring model. The ConvGRU-based encoder-decoder system takes 20 sequential thermal images as input and outputs a sequence of predicted cross-sectional phase maps. The encoder processes spatial and temporal dependencies from the temperature images, and the decoder reconstructs the

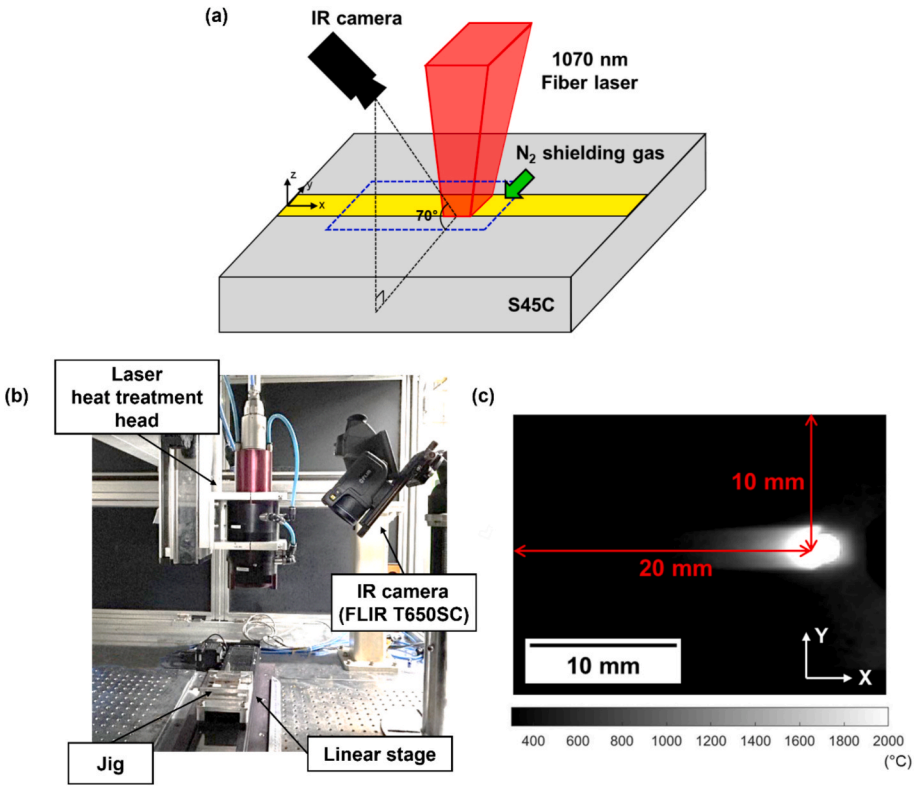



**Fig. 2.** (a) Experimental setup for laser heat treatment of S45C carbon steel using a 2 kW multi-mode fiber laser (1070 nm) and an IR camera (FLIR T650SC) for surface temperature monitoring. Nitrogen gas (20 L/min) was used as a shielding gas. (b) Photograph of the experimental setup. (c) IR thermal image showing the measured temperature distribution at x = 50 mm under 1800 W and 10.4 mm/s.

predicted internal phase structures. This end-to-end framework allows real-time, non-invasive monitoring of microstructural evolution in laser heat treatment processes and lays the foundation for intelligent control and optimization of thermal manufacturing systems.

## 2. Dataset preparation: heat treatment experiment

The experimental setup used for the laser heat treatment of S45C carbon steel is shown in Fig. 2. Fig. 2(a) illustrates the schematic diagram of the setup, and Fig. 2(b) presents a photograph of the actual experimental setup. A 2 kW multi-mode fiber laser (IPG YLS-2000) with a wavelength of 1070 nm was employed for the laser heat treatment of S45C carbon steel. The laser beam was delivered through a 200 μm process fiber, collimated using a 160 mm focal length collimation lens, and focused onto the specimen surface using a 250 mm focusing lens. A beam homogenizer reshaped the laser beam from a top-hat circular profile to a 4.4 mm × 3.7 mm rectangular top-hat beam. The specimen had dimensions of 120 mm in length (the length of the specimen in the laser scanning direction), 80 mm in width, and 10 mm in thickness.

In Fig. 2 (a), the x-, y-, and z-axes represent the laser scanning direction (length), the transverse direction (width), and the through-thickness direction (depth), respectively. The laser heat treatment was applied over a 100 mm section along the length of the specimen. To prevent oxidation, nitrogen gas was supplied as a shielding gas at a flow rate of 20 L/min. The surface temperature during the laser heat treatment process was monitored using an IR camera (FLIR T650SC), as shown in Fig. 2(c).

Both the laser head and the IR camera were fixed during the process. The specimen was mounted on a linear stage that moved along the x-direction beneath the stationary laser and IR camera. The camera was configured with an emissivity setting of 0.25 and recorded thermal images at 30 frames per second (fps). The IR camera had a temperature detection range of 300 °C to 2000 °C and a noise equivalent temperature difference (NETD) of 0.03 °C. It was mounted at a 70° angle with respect to the specimen surface, as illustrated in Fig. 2(a). Due to this angled setup, geometric calibration was performed along the y-axis using a reference grid to correct for perspective distortion in the captured thermal images. Fig. 2(c) presents the measured surface temperature distribution at the center of the scanning path (x = 50 mm) under a laser power of 1800 W and a scanning speed of 10.4 mm/s (marked by the blue dashed box in Fig. 2(b)). The chemical composition of the S45C steel specimen is provided in Table 1. To ensure surface consistency, the specimen's top surface was polished to minimize roughness and cleaned with acetone to remove dust and impurities before the laser heat treatment process.

A total of 16 process conditions were considered to construct the dataset by selecting four laser power levels (600, 865, 1248, and 1800 W) and four scanning speeds (5, 7.2, 10.4, and 15 mm/s). To obtain cross-sectional heat-treated region images, the specimens were cut in the transverse direction using wire electrical discharge machining (EDM) to minimize thermal effects. Cross-sectional images were obtained at x = 5 mm, 50 mm and 95 mm for all specimens, perpendicular to the laser scanning direction. Each specimen was polished to achieve a surface roughness of approximately 1 μm and then etched using a Nital solution.

**Table 1**
Chemical composition of S45C carbon steel (wt.%).

| C | Si | Mn | P | S | Cr | Ni | Cu | Fe |
|---|---|---|---|---|---|---|---|---|
| 0.451 | 0.253 | 0.731 | 0.013 | 0.003 | 0.039 | 0.022 | 0.041 | balanced |

**Table 2**
Summary of key experimental parameters and instrumentation used in the laser heat treatment process, including laser optics, specimen characteristics, shielding gas, and IR thermal imaging system.

| | |
|---|---|
| Laser | |
| Wavelength | 1070 nm |
| Focal length | 250 mm |
| Collimation length | 160 mm |
| Beam size | 4.4 mm × 3.7 mm (top-hat profile) |
| Specimen | |
| Material | S45C |
| Dimension | 80 mm × 120 mm × 10 mm |
| Chemical composition | Table 1 |
| Shielding gas | |
| Gas type | $N_2$ |
| Flow rate | 20 L/min |
| IR camera | |
| Resolution | 480 × 640 pixel (41.6 μm/pixel) |
| Frame rate | 30 fps |
| Temperature range | 300 ~ 2000 °C |
| NETD | 0.03 °C |

The key process parameters and instrumentation details used in the laser heat treatment experiment are summarized in Table 2.

Fig. 3 illustrates an OM image of the specimen cross-section at x = 50 mm obtained under a laser power of 1800 W and a scanning speed of 10.4 mm/s, where four distinct regions are identified as (a), (b), (c), and (d). Magnified OM images for each region are also presented. As shown in Fig. 3(a), the base metal consists of ferrite and pearlite phases. In region (b), the transformation of ferrite and pearlite phases leads to a rapid increase in hardness. In region (c), the microstructure is fully converted from austenite to martensite, where the hardness remains stable. Region (d) represents the transition between the melted and heat-treated zones, characterized by a dark-etched area. Using this classification method, all 16 process conditions were mapped into phase distribution maps, as shown in Fig. 4, where regions (a), (b), (c), and (d) are color-coded in black, blue, green, and red, respectively.

The top image shows the cross-sectional OM image, with colored boxes indicating the locations corresponding to the microstructural images shown in (a)–(d). (a) Base metal consisting of ferrite and pearlite. (b) Transition region with partial phase transformation. (c) Hardened region composed of martensite. (d) Melted, transition, and heat-treated regions. A scale bar of 150 μm is provided.

As shown in Fig. 4, the cross-sectional phase distribution varies notably with changes in laser power and scanning speed. Higher laser power increases energy input, resulting in deeper and wider transformation zones, while higher scanning speed shortens laser-material interaction time, leading to shallower and narrower regions.

Specifically, the heat-affected zone (blue) increases gradually in width and depth with energy input. The heat-treated region (green), where martensitic transformation occurs, becomes dominant at higher power and lower speed, forming a thicker hardened layer. The melted region (red) appears only under the highest energy conditions, where excessive heat input causes localized melting near the surface. This progression reflects the transition from partial phase transformation to full hardening and eventual melting, depending on the thermal energy delivered to the material.

The ground truth cross-sectional phase distribution images had a resolution of 240 × 640 pixels and comprised four channels (i.e., [H, W, C] = [240, 640, 4]), with each channel representing a distinct phase

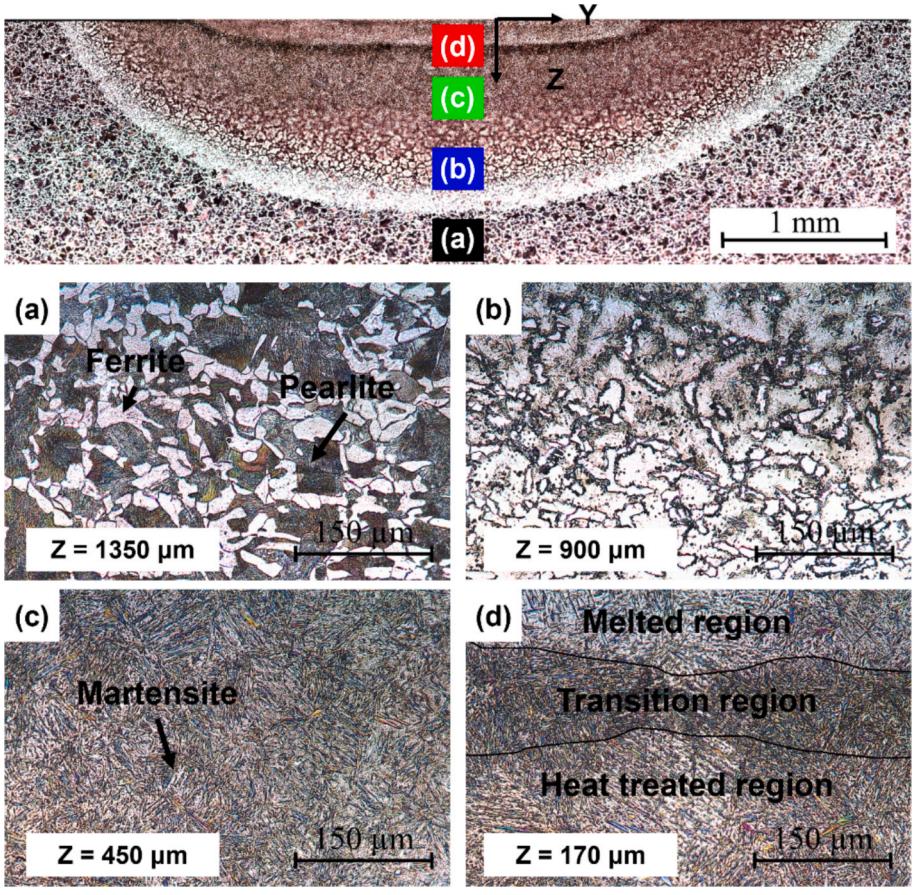



**Fig. 3.** Microstructural analysis of laser heat-treated S45C steel (1800 W, 10.4 mm/s) at x = 50 mm.

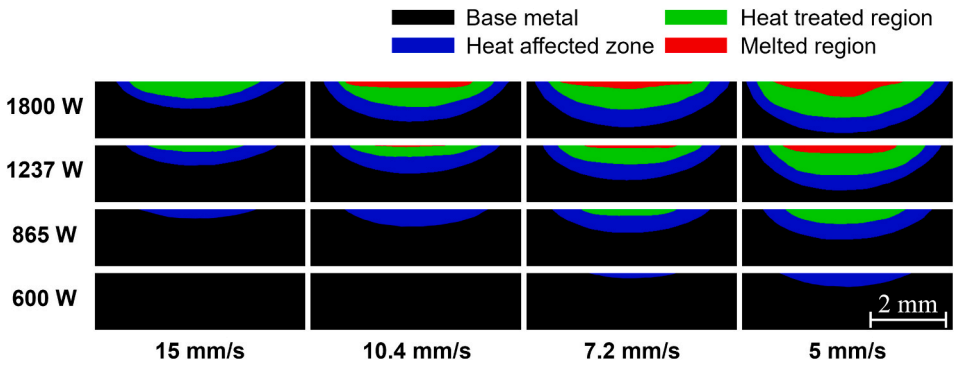



**Fig. 4.** Cross-sectional phase distribution maps of laser heat-treated S45C steel under different laser power and scanning speed conditions. The regions are color-coded as follows: base metal (black), heat-affected zone (blue), heat-treated region (green), and melted region (red). A scale bar of 2 mm is provided. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

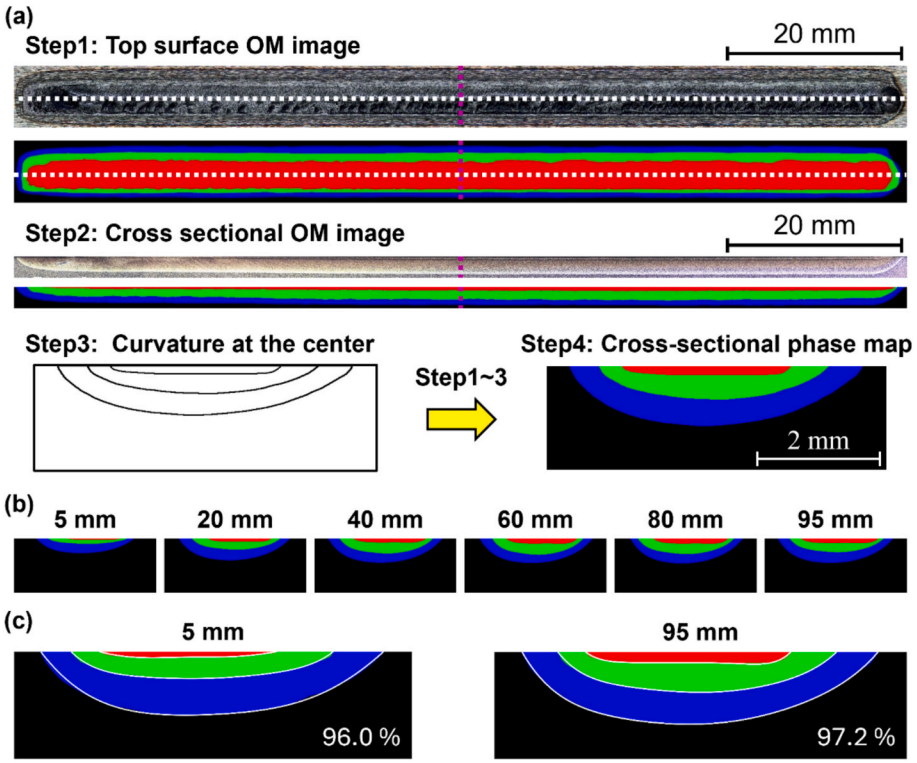



**Fig. 5.** (a) Procedure for synthesizing ground truth cross-sectional phase distribution maps: Step 1—Top surface OM image analysis to extract phase boundaries along the x-direction (x–y plane); Step 2—Cross-sectional OM imaging along the centerline (x–z plane); Step 3—Curve fitting of phase boundaries using sixth-order polynomials; Step 4—Generation of cross-sectional phase maps at arbitrary x-locations. (b) Synthesized phase distribution images at six x-positions (5, 20, 40, 60, 80, and 95 mm). (c) Validation of synthesized boundaries at x = 5 mm and x = 95 mm for 1800 W and 10.4 mm/s. White curves indicate the fitted phase boundaries, and $R^2$ values represent the goodness of fit between synthesized and experimental data.

region.

Generating ground-truth phase distributions at various cross-sections along the entire heat-treated length is essential for training the deep learning model. Fig. 5(a) illustrates the three-step process used to construct the ground truth data.

In Step 1, the top surface of the heat-treated specimen was analyzed using OM to obtain information on the different phase regions along the x-direction (x–y plane). In Step 2, the specimen was sectioned along the x-direction to reveal the full heat-treated region along the centerline (x–z plane). From this view, the different phase regions were identified. In Step 3, the phase boundary data from the x–y and x–z planes were used to fit sixth-order polynomial curves to each phase boundary. These fitted curves were then used to synthesize cross-sectional phase distribution maps at arbitrary locations along the scanning direction.

Fig. 5(b) shows examples of the synthesized ground truth phase distribution images at six x-positions, ranging from 5 mm to 95 mm. The results indicate that the hardening depth gradually increases due to cumulative heat input. Additionally, minor asymmetries in the shape of the heat-treated region were observed.

To validate the accuracy of the synthesized phase boundary curves, we prepared additional cross-sections at two critical locations, x = 5 mm and x = 95 mm—where acceleration and deceleration occur and phase transitions are most pronounced. Therefore, in total, 48 OM images (16 processing conditions × 3 x-positions) were analyzed. Fig. 5(c) presents the comparison results for 1800 W and 10.4 mm/s. The white curves indicate the fitted phase boundaries obtained using the method described in Fig. 5(a). The $R^2$ value shown in the corner of each image quantifies the goodness of fit. As shown, the synthesized phase

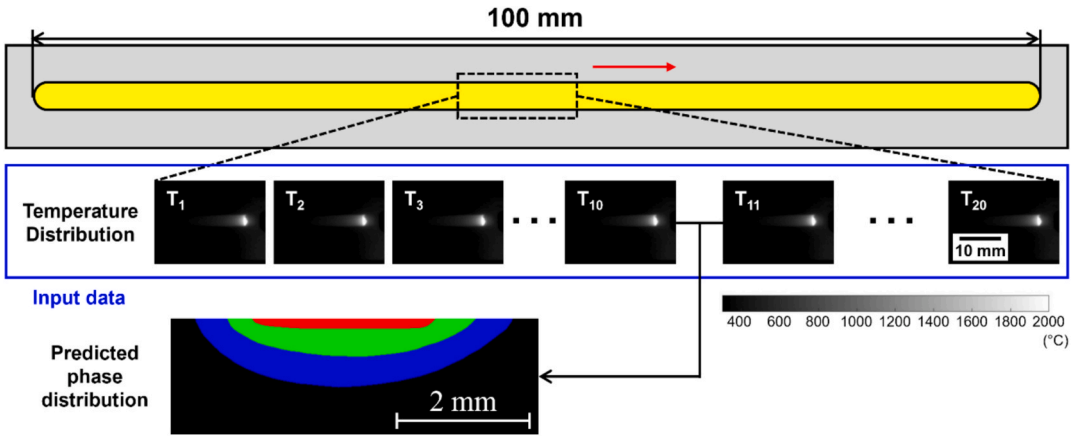



**Fig. 6.** Schematic of the laser scanning process and dataset configuration. A sequence of 20 IR temperature images, captured before and after the target location, was used as input to predict the cross-sectional phase distribution at that location.

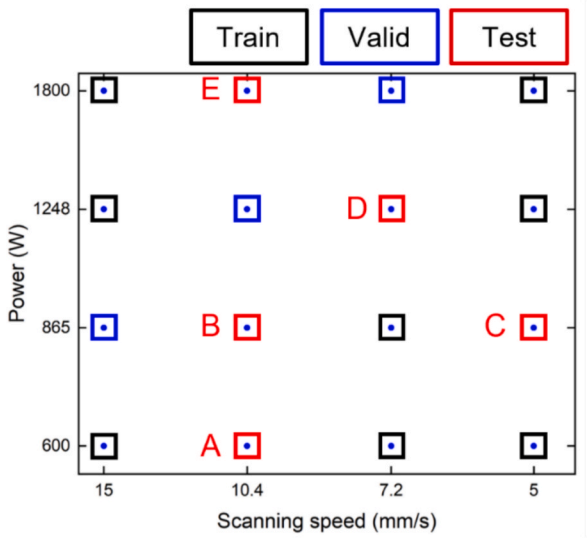



**Fig. 7.** Distribution of the 16 experimental conditions on the laser power–scanning speed plane. The dataset was divided into training (black), validation (blue), and test (red, labeled A–E) sets. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

boundaries closely match the actual experimental data. The average $R^2$ accuracy across all process conditions (excluding the non-heat-treated case) was 97.6 % at x = 5 mm and 97.9 % at x = 95 mm.

To accurately capture the heating and cooling rates during the laser heat treatment process, a sequence of temperature data was required for training the AI model. As schematically illustrated in Fig. 6, the model takes as input 20 sequential top-surface temperature images acquired by an IR camera to predict the cross-sectional phase distribution at a specific location. These input images include 10 frames captured before and 10 frames after the target prediction location.

We empirically evaluated the optimal number of input temperature frames for achieving the best prediction performance. It was found that increasing the number of input frames beyond 20 did not significantly improve the prediction accuracy, while it substantially increased the computational cost. Therefore, using 20 temperature images was determined to be sufficient for accurate prediction of the cross-sectional phase distribution. Each temperature image had a resolution of 480 × 640 pixels.

Although the time interval between successive frames was constant at 1/30 s, additional information about the spatial position of each frame was necessary because the linear stage did not always move at a constant speed. For instance, during the start or end of the heat treatment, the linear stage may accelerate or decelerate, resulting in temporal frames that are unevenly spaced in space. To compensate for this, the beam position corresponding to each temperature image was provided to the model as an auxiliary input.

Fig. 7 illustrates the 16 experimental process conditions mapped on the laser power–scanning speed plane, covering a broad range of energy inputs relevant to laser heat treatment. These conditions were selected to ensure diversity in thermal histories and phase transformation behaviors. The dataset was divided into training, validation, and test subsets in a ratio of 8:3:5, allowing the model to learn generalizable patterns while avoiding overfitting. The test dataset comprises the experimental conditions labeled A to E in Fig. 7.

From the eight test conditions, a total of 3,200 infrared (IR) images were collected. For the training set, left–right flipping augmentation was applied to double the number of images, resulting in 6,400 IR images used for model training. This simple yet effective method doubled the training data, improving the model's robustness to spatial variance and directional asymmetry.

## 3. Deep learning model

The architecture of the AI-based monitoring model is schematically illustrated in Fig. 8. The model is based on an RNN encoder-decoder framework [16–19] combined with a CNN [20–22]. This model translates a sequence of 2D surface temperature distributions $x = (x_1, x_2, x_3, \ldots, x_{20})$ into a segmented cross-sectional phase distribution map y. The primary advantage of this model is its ability to simultaneously capture both local heating and cooling patterns within a single surface temperature image, as well as global temporal dependencies across multiple surface temperature distributions.

As depicted in Fig. 8(a), each layer of the monitoring model processes an individual surface temperature distribution concatenated with a position vector ($P_t$), which represents the center position of the laser beam corresponding to each input frame. The inclusion of the position vector is particularly important due to certain limitations of the thermal imaging system. Since the IR camera captures only surface temperature, it cannot directly detect changes in scanning speed, especially during acceleration and deceleration at the start and end of each scan path. Such variations in speed significantly affect the heating and cooling rates, which in turn influence phase transformations and microstructural evolution. Moreover, the limited field of view of the IR camera prevents full coverage of the processing region. To overcome these limitations, the spatial position of the laser beam corresponding to each thermal frame was incorporated as an input feature. This enables the model to

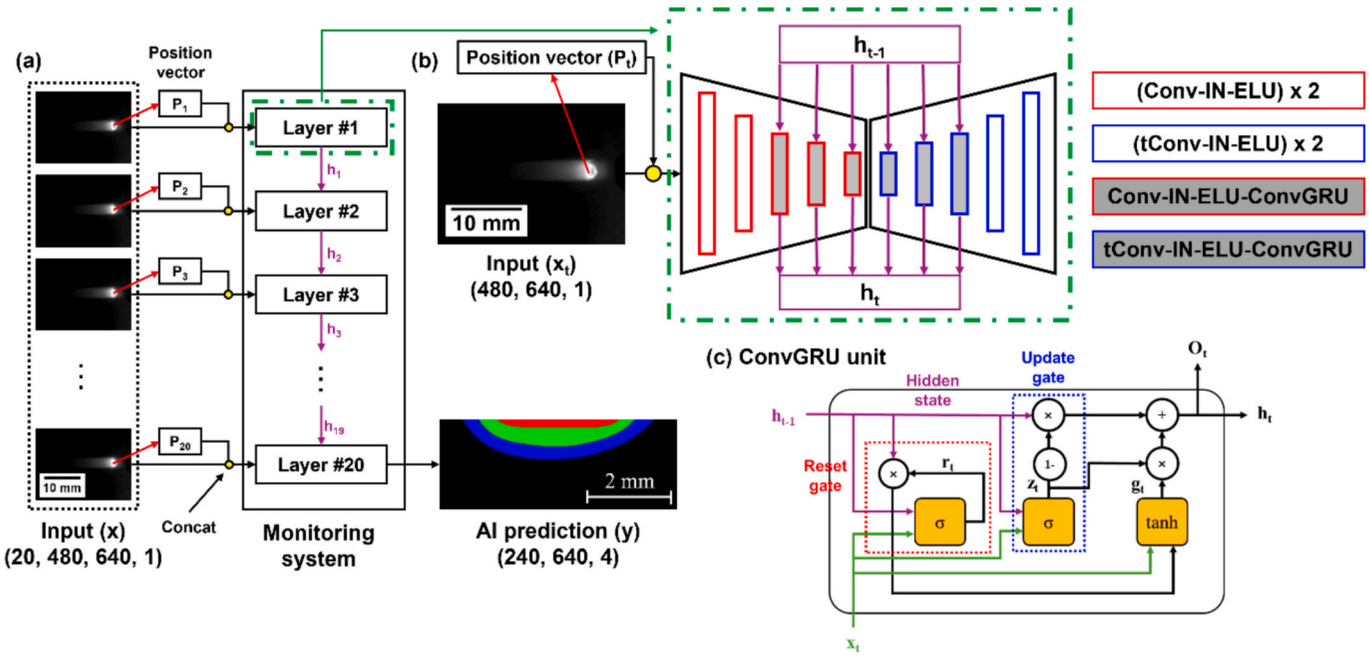



**Fig. 8.** Architecture of the AI-based monitoring model. (a) Sequential surface temperature images and position vectors are input into the model, where each layer processes a frame and passes the extracted features through hidden states. (b) Encoder-decoder structure incorporating ConvGRU for spatiotemporal feature extraction. (c) Detailed structure of the ConvGRU unit, illustrating the reset and update gates used for sequential learning.

infer relative motion between the laser and the workpiece and improves its ability to capture thermal history under variable scanning conditions and restricted observation windows. The number of layers corresponds to the number of input frames, and the hidden state ($h_t$) carries extracted feature maps to the subsequent layer. At the final layer, the model outputs the predicted cross-sectional phase distribution map.

The encoder extracts spatial and temporal features using a sequence of convolutional layers (Conv) [20], instance normalization (IN) [23], and exponential linear unit (ELU) activation functions [24], as shown in Fig. 8(b). The decoder reconstructs the phase distribution map using transposed convolutional layers (tConv) [25], IN, and ELU activations. To ensure the correct output height, the stride of the last transposed convolution layer is set to one. Three out of the five layers in both the encoder and decoder incorporate ConvGRU.

ConvGRU [21] was applied only to the three bottleneck layers of the encoder–decoder structure. The ConvGRU layers were excluded from the first two encoder layers and the last two decoder layers, as these stages primarily focus on low-level spatial encoding or high-resolution reconstruction. When ConvGRU was applied to all layers, the prediction accuracy did not improve significantly, while GPU memory usage and computational time increased considerably.

Fig. 8(c) illustrates the structure of the ConvGRU unit, which consists of a reset gate, an update gate, and a hidden state. The reset gate ($r_t$, marked by a red dotted line) determines the extent to which past information is forgotten and is defined as

$$r_t = \sigma(w_{xr} \cdot x_t + w_{hr} \cdot h_{t-1} + b_r) \tag{1}$$

where σ represents the logistic sigmoid function, and t, $x_t$, $w_x$, $w_h$, $h_{t-1}$ and b denote the number of layers, the input, the weight calculated with the input, the weight calculated with the previous hidden state, and the bias, respectively. The update gate ($z_t$, marked by a blue dotted line) controls how much past information is retained versus how much new information is incorporated, and is similarly defined as

$$z_t = \sigma(w_{xz} \cdot x_t + w_{hz} \cdot h_{t-1} + b_z) \tag{2}$$

Finally, the hidden state ($h_t$) is updated based on the reset and update gates using the following equation:

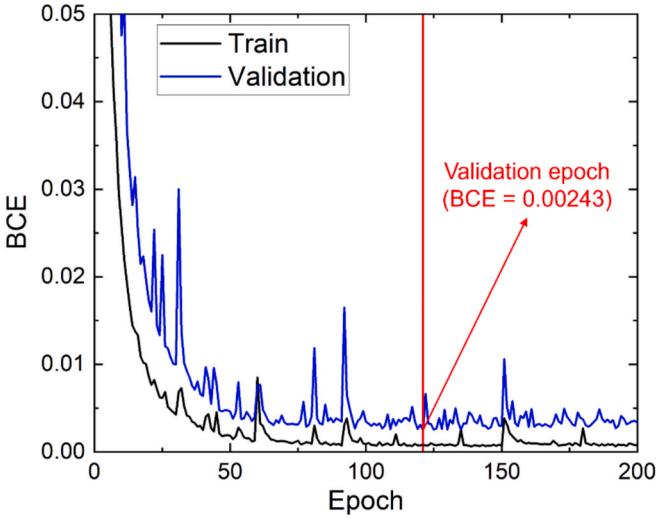



**Fig. 9.** Training and validation loss curves of the AI-based monitoring model. The black line represents the training loss, while the blue line represents the validation loss. The red vertical line indicates the validation epoch where the BCE loss reaches its minimum value of 0.00243. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

$$\begin{aligned} g_t &= \tanh(w_{hg} \cdot (r_t \odot h_{t-1}) + w_{xg} \cdot x_t + b_g) \\ h_t &= z_t \odot h_{t-1} + (1 - z_t) \odot g_t \end{aligned} \tag{3}$$

where $g_t$, and ⊙ represent the activation process before updating the hidden state and the pointwise operation, respectively. The update gate determines how much temporal information to send to the next layer.

The training and validation performance of the AI monitoring model was evaluated using binary cross-entropy (BCE) loss [26], as shown in Fig. 9. The loss function is defined as

**Table 3**
Training hyperparameters for the AI-based monitoring model.

| | |
|---|---|
| Batch size | 2 |
| Learning rate | $10^{-4}$ |
| Weight decay | $10^{-5}$ |
| Validation epoch | 121 (BCE = 0.00243) |
| Optimizer | Adam |
| Loss function | Binary cross-entropy loss (BCE) |

$$L(y,t) = -\frac{1}{N}\sum_{i=1}^{N}[t_i \cdot \log y_i + (1-t_i)\cdot \log(1-y_i)] \quad (4)$$

where y is the predicted value, t is the ground truth, and N is the total number of elements. The training and validation losses are represented by black and blue lines, respectively, in Fig. 9.

The model was trained and optimized to minimize the BCE loss on the validation set, and the minimum validation loss (0.00243) was reached at epoch 121, as indicated by the red vertical line. At this point, early stopping was applied to avoid overfitting, and the model weights corresponding to the validation epoch were used for evaluation on the test dataset.

The AI monitoring model was implemented using Python™ and PyTorch™ and trained on a TITAN RTX GPU. The Adam optimizer [27] was used for optimization. The training hyperparameters are summarized in Table 3.

## 4. Results and discussion

Fig. 10 presents the monitoring results obtained using the developed AI model under a laser power of 1248 W and a scanning speed of 7.2 mm/s. Fig. 10(a) displays the top surface image of the laser-processed specimen over a length of 100 mm. As shown in the image, the heat-affected zone and heat-treated region remain relatively stable along the length, whereas the melted region exhibits an irregular shape. Fig. 10(b) compares the ground truth and AI-predicted cross-sectional phase distribution images at several positions from 10 mm to 90 mm. Overall, the predicted results closely match the ground truth, particularly for the heat-affected and heat-treated regions. Fig. 10(c) quantitatively compares the width and depth of the phase regions along the scanning direction. As shown in the figure in the lower left corner, width was defined as the horizontal distance along the y-direction between the leftmost and rightmost segmented boundaries of each phase region in a given cross-section. This definition reflects the overall span of each region, regardless of asymmetry. The width and depth of the phase regions (melted, heat-treated, and heat-affected zones) were determined from the segmentation images using pixel-based measurements. These values were then converted to micrometers (μm) using a fixed resolution of 8.7 μm/pixel. The red, green, and blue lines represent the melted region,

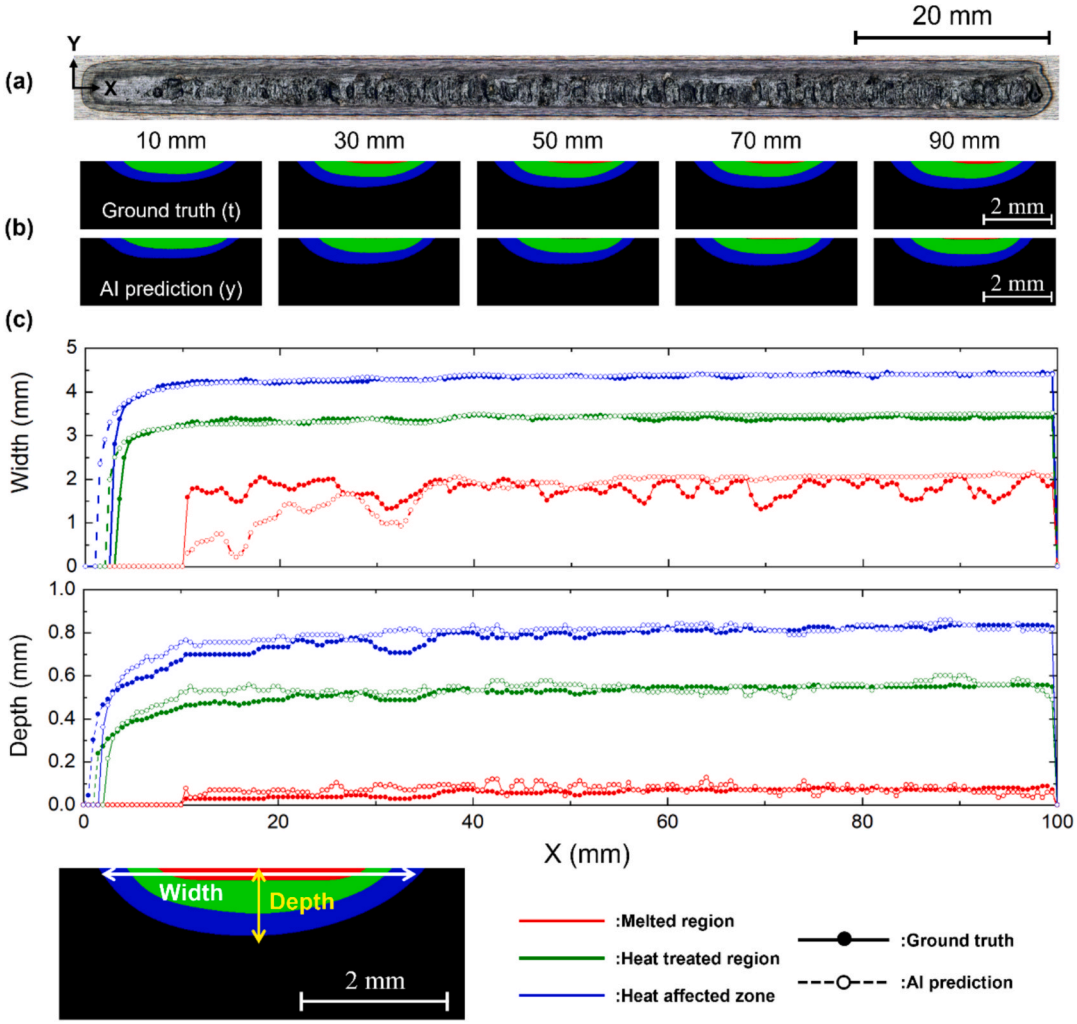



**Fig. 10.** Monitoring results of the AI-based monitoring model under 1248 W and 7.2 mm/s. (a) Top surface image of the heat-treated specimen. (b) Comparison of ground truth and AI-predicted cross-sectional phase distributions at selected positions. (c) Width and depth profiles of each phase region along the scanning direction. Definition of heat-treated region width and depth in the cross-sectional phase distribution was plotted at left bottom side of figure.

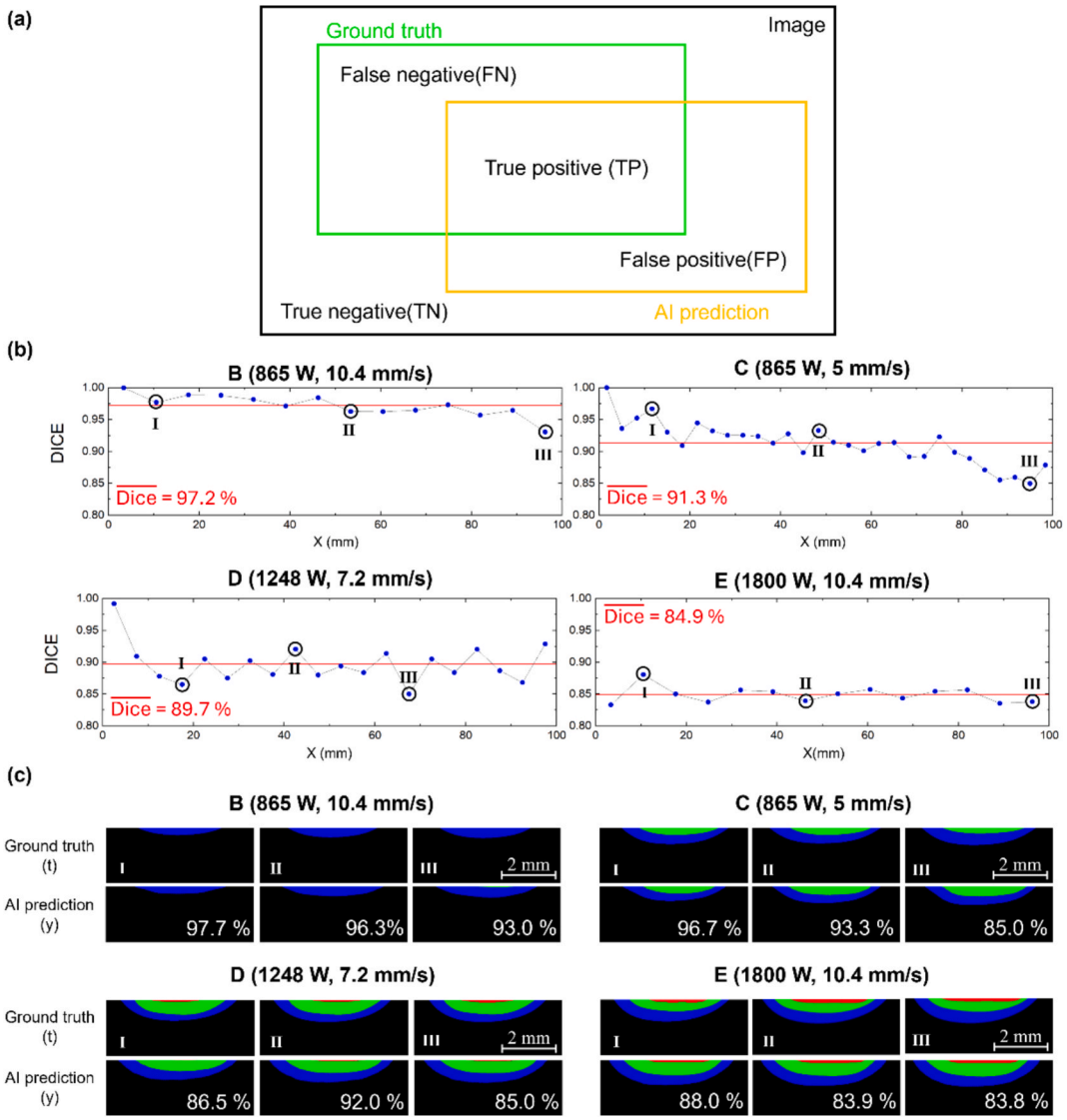



**Fig. 11.** (a) Schematic illustration of the evaluation method for AI prediction performance using the Dice score. The green and yellow rectangles represent the ground truth and AI prediction regions, respectively, with classifications of true positive (TP), false positive (FP), false negative (FN), and true negative (TN). (b) Dice score variation along the x-direction (laser scanning direction) for four test conditions: B (865 W, 10.4 mm/s), C (865 W, 5 mm/s), D (1248 W, 7.2 mm/s), and E (1800 W, 10.4 mm/s). The red horizontal lines represent the average Dice scores, and the selected positions I, II, and III are marked for comparison in (c). (c) Comparison of ground truth and AI-predicted cross-sectional phase distributions at the selected positions I, II, and III for the four test cases. Local Dice scores are shown at the bottom of each prediction. A scale bar of 2 mm is included in each image. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

heat-treated region, and heat-affected zone, respectively. Solid lines indicate the ground truth, while dashed lines represent the AI predictions. As shown in the plots, the predicted widths and depths are in good agreement with the ground truth across most of the scanned length.

In particular, the model effectively captures the gradual increase in the width and depth of each region in the early stage of scanning, which is attributed to heat accumulation. However, the prediction accuracy for the melted region is lower than that for the other regions. This can be attributed to three main factors. First, the melted region tends to have an irregular and unstable geometry, making it inherently more difficult for the model to learn consistent spatial features. Second, the melted zone is relatively small in size, which limits the amount of usable information within each sample. Third, melted regions appeared in only a few specimens, resulting in a relative lack of training data. Furthermore, since laser heat treatment is not primarily intended to induce melting, the melted zone is not a dominant feature of the process. These factors combined make accurate prediction of the melted region particularly challenging for the AI model.

To evaluate the performance of the AI-generated phase distribution images, the Dice score [28] was calculated as a metric to assess segmentation accuracy. The Dice score, which represents the harmonic mean of precision and recall, was used because the ground truth dataset consists of binary values, where 0 indicates the absence of a phase and 1 indicates its presence. Fig. 11(a) illustrates the evaluation process, where the green and yellow rectangular regions represent the ground truth and AI prediction areas, respectively. The predicted region is further categorized into true positive (TP), false positive (FP), false negative (FN), and true negative areas (TN). The Dice score is calculated as

$$\text{Dice} = \frac{2 \times TP}{2 \times TP + FP + FN} \tag{5}$$

Fig. 11(b) shows the Dice score variations along the x-direction (i.e.,

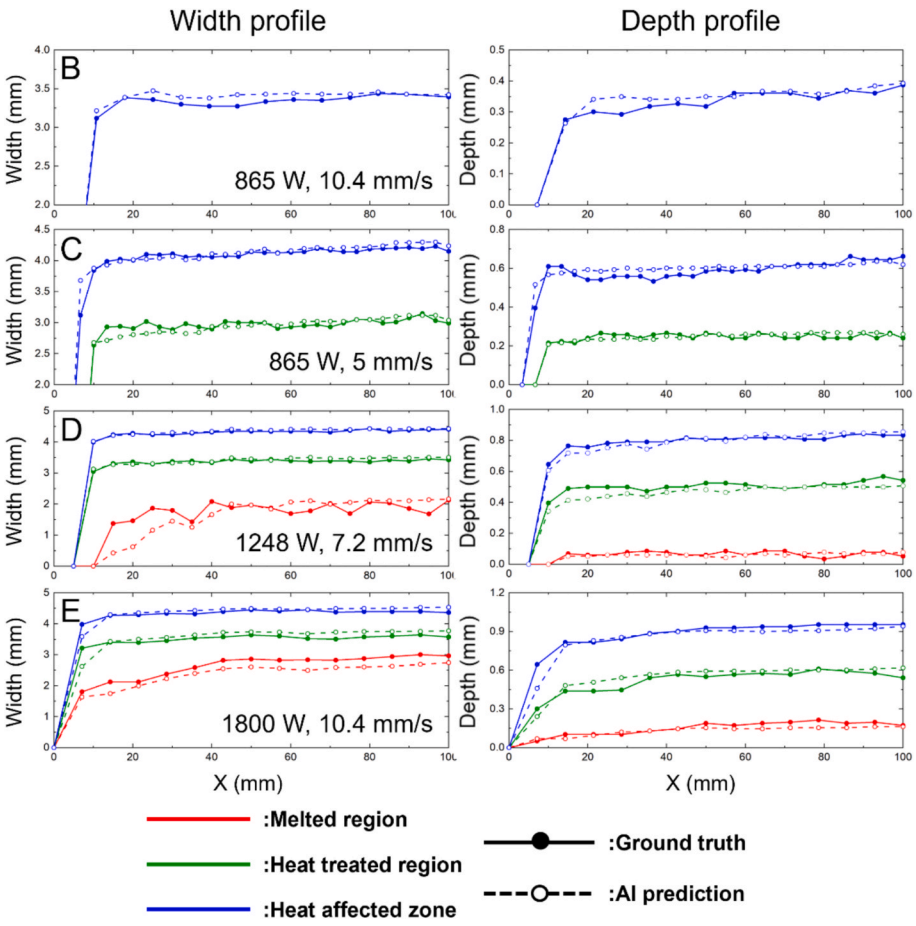



**Fig. 12.** Comparison of heat-treated region width (left) and depth (right) between AI predictions (dashed lines) and ground truth (solid lines) for test conditions B (865 W, 10.4 mm/s), C (865 W, 5 mm/s), D (1248 W, 7.2 mm/s), and E (1800 W, 10.4 mm/s). The red, green, and blue lines represent the melted region, heat-treated region, and heat-affected zone, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the laser scanning direction) for four test conditions B (865 W, 10.4 mm/s), C (865 W, 5 mm/s), D (1248 W, 7.2 mm/s), and E (1800 W, 10.4 mm/s), as defined in Fig. 7. The red horizontal line in each plot represents the average Dice score.

Fig. 11(c) compares the ground truth and AI-predicted cross-sectional phase distributions at selected positions (I, II, III) for each test case, with the corresponding local Dice scores provided below each prediction. The AI-generated results closely resemble the ground truth across all conditions. However, as the input energy density increases from Case B to Case E, the overall prediction accuracy tends to decrease due to the increased complexity of the phase distributions. In particular, the accuracy for the melted region, especially in Cases D and E, is lower than that for the heat-treated and heat-affected zones. This discrepancy is attributed to the relatively small number of training samples containing melted regions, as well as the irregular geometry of these zones. Despite these challenges, the model successfully captures the overall shape and boundaries of the phase regions. Additionally, it is worth noting that Case A (600 W, 10.4 mm/s), representing a non-hardening condition, was accurately predicted as having no hardening effects, which further demonstrates the reliability of the model in identifying boundary conditions.

To further assess the accuracy of the AI predictions, the phase distribution map was measured along the specimen length for all predicted images. Fig. 12 presents the measured heat-treated region width and depth for the four test datasets. The dimensions of the melted region, heat-treated region, and heat-affected zone are represented in red, green, and blue, respectively. The ground truth and AI-predicted results are indicated by solid and dashed lines, respectively. The geometric trends of the heat-treated region cross-section show strong agreement between AI predictions and the ground truth. Additionally, the heat-treated region depth and width generally increase along the laser scanning direction due to heat accumulation effects, which the AI model successfully captures. However, for Cases D and E, the AI model struggles to predict the melted region, likely due to the limited number of training samples and the irregular shape of this phase.

Table 4 summarizes the accuracy of the predicted heat-treated region width and depth for different regions across the four test cases. The

**Table 4**
Accuracy of AI-predicted width and depth for different phase regions in test conditions B (865 W, 10.4 mm/s), C (865 W, 5 mm/s), D (1248 W, 7.2 mm/s), and E (1800 W, 10.4 mm/s). The accuracy is calculated for the melted region, heat-treated region, and heat-affected zone, with the average accuracy presented for each case.

| Width | | | | | |
|---|---|---|---|---|---|
| Accuracy (%) | **B** | **C** | **D** | **E** | **Average** |
| Melted region | – | – | 83.2 | 90.4 | 86.8 |
| Heat treat region | – | 95.5 | 98.1 | 95.2 | 96.3 |
| Heat affected zone | 96.7 | 98.3 | 99.1 | 97.5 | 97.9 |
| Depth | | | | | |
| Accuracy (%) | **B** | **C** | **D** | **E** | **Average** |
| Melted region | – | – | 77.9 | 83.0 | 80.5 |
| Heat treat region | – | 91.8 | 93.1 | 91.5 | 92.1 |
| Heat affected zone | 90.0 | 94.8 | 97.5 | 95.6 | 94.5 |

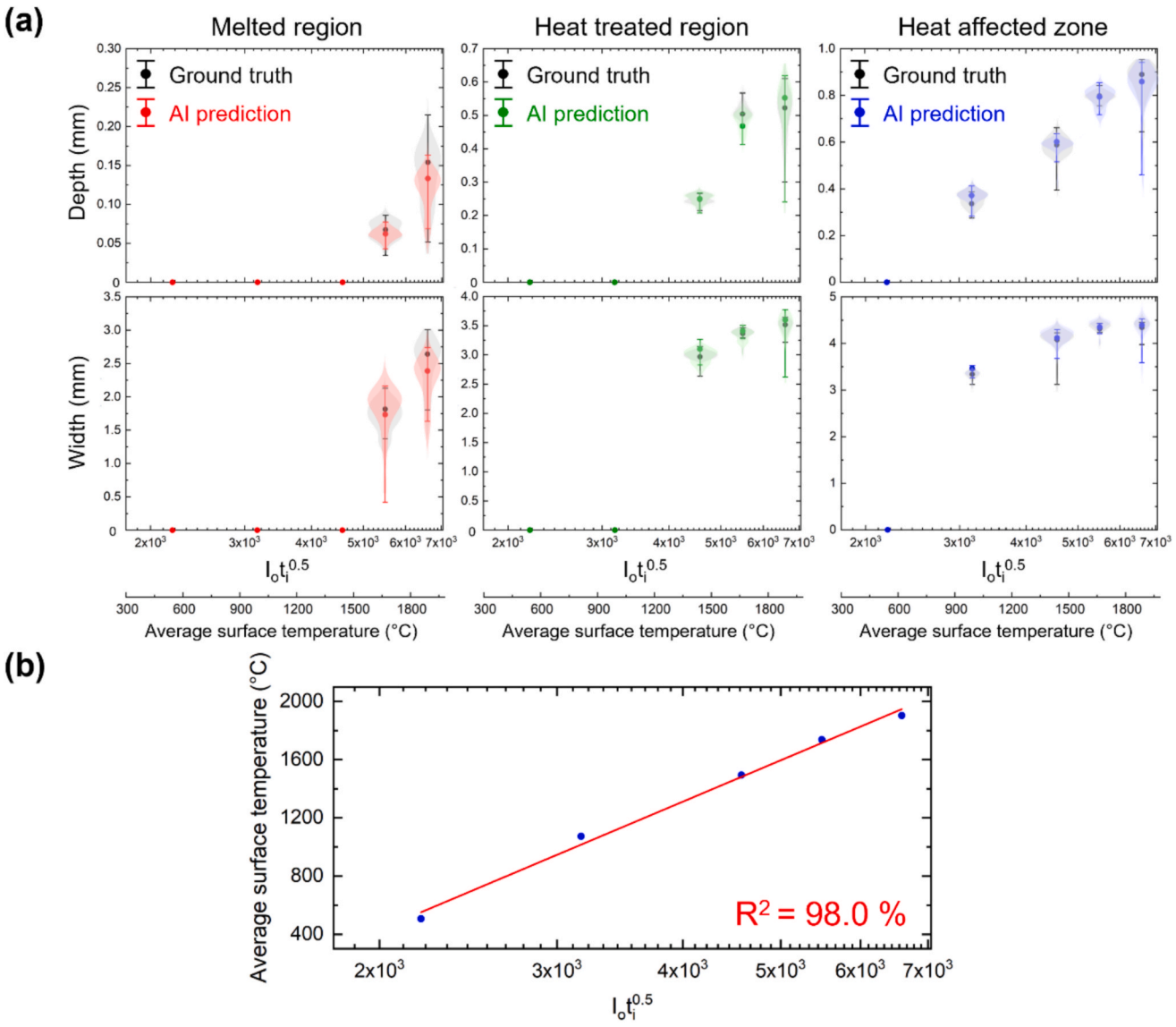



**Fig. 13.** (a) Comparison of AI-predicted and ground truth widths and depths of the melted region, heat-treated region, and heat-affected zone as functions of $I_o t_i^{0.5}$. The x-axis also shows the corresponding average surface temperature within the laser-irradiated region. Ground truth is shown in black, while AI predictions are in red (melted region), green (heat-treated region), and blue (heat-affected zone). Solid circles indicate the mean values, and the violin plots represent the data distribution at each point. (b) Correlation between the average surface temperature and $I_o t_i^{0.5}$, demonstrating a strong linear relationship. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

average width accuracy was 86.8 % for the melted region, 96.3 % for the heat-treated region, and 97.9 % for the heat-affected zone. The corresponding values for depth were 80.5 %, 92.1 %, and 94.5 %, respectively. These results indicate that the model exhibits high geometric prediction accuracy, particularly in the heat-treated and heat-affected zones. However, as previously noted, the accuracy in the melted region remains lower due to the challenges associated with limited training data and geometric complexity.

Fig. 13(a) presents the average width and depth of the heat-treated region as functions of the laser processing parameter $I_o t_i^{0.5}$ [29], where $I_o$ is the laser intensity, defined as the laser power divided by the beam area of the top-hat beam, and $t_i$ is the interaction time, defined as the beam length along the scanning direction divided by the scanning speed. This parameter characterizes the degree of thermal exposure, with higher values indicating greater temperature rise [30].

In the figure, the ground truth, melted region, heat-treated region, and heat-affected zone are represented by black, red, green, and blue markers, respectively. As expected, both the width and depth increase with increasing thermal exposure, consistent with previous experimental observations. The AI model effectively captures this temperature-dependent trend, demonstrating its ability to predict laser-induced phase transformations.

The bright-colored violin shapes illustrate the data distribution at each point. A wider horizontal spread indicates a larger number of data samples and greater variation. As shown in the figure, the data distributions of the AI predictions closely match the ground truth for the heat-treated and heat-affected zones. However, the prediction for the melted region is less accurate, as also observed in Fig. 11 and Fig. 12.

The maximum prediction errors in average width and depth were 120 μm and 36 μm, respectively. Given that the ground truth resolution was 8.7 μm/pixel, these correspond to deviations of approximately 13 and 4 pixels, respectively, indicating reasonably high prediction accuracy.

To provide a clearer link between thermal behavior and process outcomes, an additional x-axis representing the average surface temperature within the laser-irradiated region was added to Fig. 13(a).

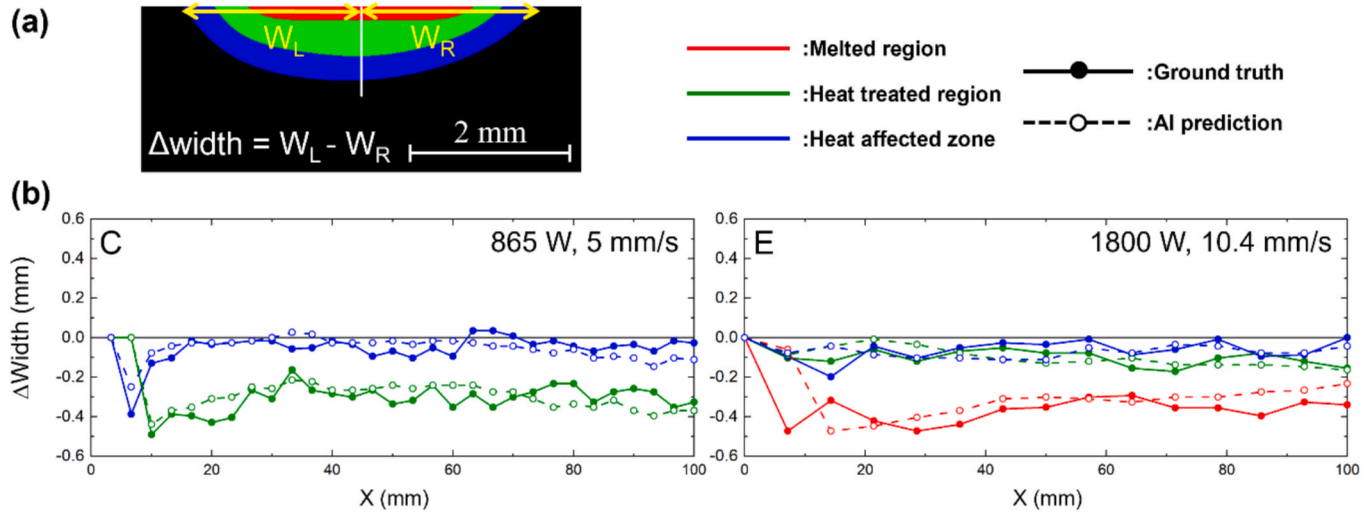



**Fig. 14.** (a) Definition of heat-treated region width asymmetry ($\Delta W = W_L - W_R$) as the difference between the left and right widths. (b) Comparison of ΔW profiles between AI predictions (dashed lines) and ground truth (solid lines) for test conditions C (865 W, 5 mm/s) and E (1800 W, 10.4 mm/s). The red, green, and blue lines represent the melted region, heat-treated region, and heat-affected zone, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

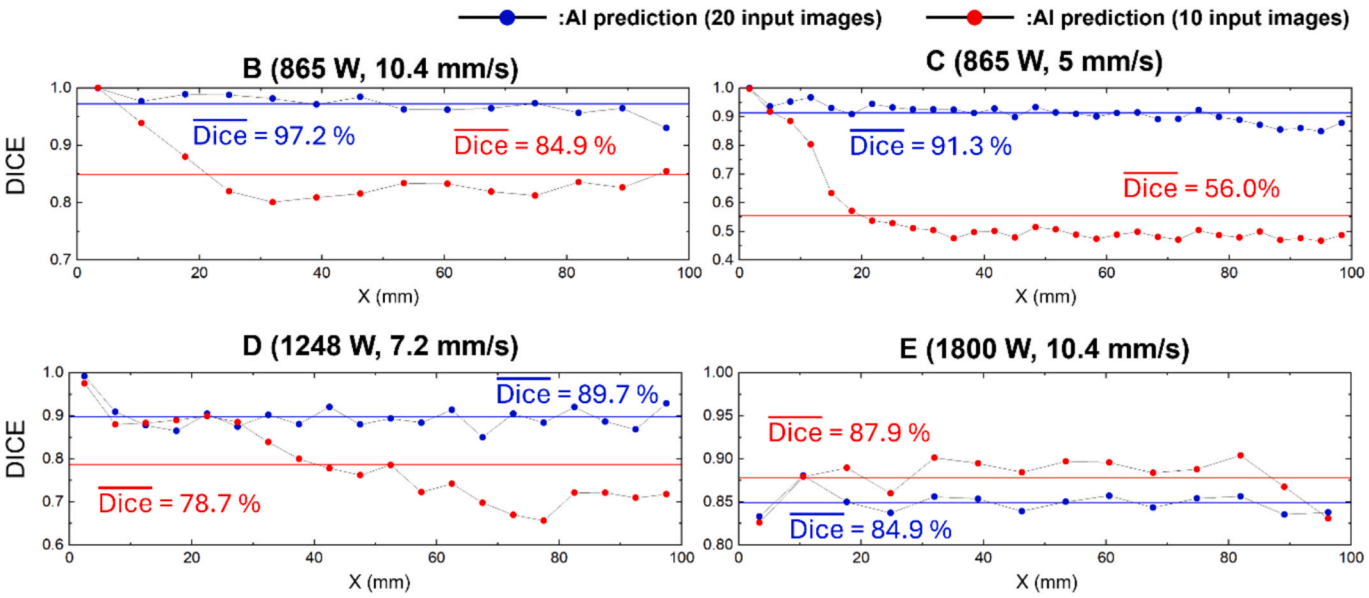



**Fig. 15.** Comparison of Dice scores along the x-direction for test conditions B (865 W, 10.4 mm/s), C (865 W, 5 mm/s), D (1248 W, 7.2 mm/s), and E (1800 W, 10.4 mm/s) using 10 and 20 input frames. Red and blue markers indicate results with 10 and 20 frames, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

These temperature values were experimentally measured using thermographic data for the same five process conditions used in Fig. 13(a). As shown in Fig. 13(b), the average surface temperature is nearly proportional to $I_o t_i^{0.5}$ ($R^2 = 98.0$ %).

This dual-axis approach allows for a more intuitive interpretation of how thermal input influences the geometry of different phase zones. It also demonstrates that the AI model successfully captures the temperature-dependent nature of laser-induced transformations, even without explicit temperature input.

During the experiments, slight asymmetry in the cross-sectional profiles was observed, primarily due to experimental variability, such as minor fluctuations in beam quality. To evaluate whether the AI model could capture such asymmetry, an asymmetry analysis was conducted. As illustrated in Fig. 14(a), width asymmetry was defined as the difference between the left ($W_L$) and right ($W_R$) widths relative to the centerline of the specimen. This definition was applied separately to each of the three phase regions.

Fig. 14(b) presents the measured asymmetry profiles along the scanning direction for test conditions C and E. The red, green, and blue lines correspond to the melted region, heat-treated region, and heat-affected zone, respectively. Solid lines indicate the ground truth, while dashed lines represent the AI predictions. As shown, the model successfully captures the overall trend and relative direction of asymmetry across the entire length. This implies that the model has learned not only the general thermal behavior, but also the subtle directional tendencies arising from asymmetric heating and cooling during the process.

Importantly, this result demonstrates the model's capability to reflect real-world imperfections in thermal fields, which is particularly valuable for practical applications where ideal symmetry cannot always be guaranteed.

In this study, we also conducted an ablation study to validate the use of 10 frames before and after the target location as input to the model. In particular, we investigated whether using only the thermal history after the target location (i.e., 10 future frames) would be sufficient for accurate phase distribution prediction. Since final microstructure and grain size are primarily governed by the cooling rate, post-target thermal data are expected to have a stronger influence than pre-target data.

Fig. 15 presents the Dice score variation along the x-direction for four representative test conditions. Red and blue markers indicate results from models trained using 10 and 20 input frames, respectively. In the early scanning region ($x < 10$ mm), both models produced comparable accuracy. However, beyond this region, the 10-frame model showed a significant accuracy drop due to its inability to fully capture heat accumulation effects.

For condition E, the 10-frame model achieved a slightly higher Dice score; however, it still failed to correctly predict the melt region. These findings confirm that using 10 frames both before and after the target location is generally more effective for accurately predicting phase distributions.

## 5. Conclusions

- This study proposed the first AI-based monitoring framework for predicting cross-sectional phase distributions during laser heat treatment of S45C steel using sequential surface temperature images.
- A deep learning model based on a convolutional gated recurrent unit architecture was developed to extract both spatial and temporal features, enabling real-time image-to-image synthesis.
- The model translated thermal image sequences into cross-sectional phase distribution maps with reasonably high accuracy.
- The model achieved an average prediction accuracy of 97.9 %, 96.3 %, and 86.8 % for width and 94.5 %, 92.1 %, and 80.5 % for depth in the heat-affected, heat-treated, and melted regions, respectively.
- It effectively captured heat accumulation effects, showing a gradual increase in the width and depth of the heat-treated region along the scanning direction.
- The model also identified cross-sectional asymmetry caused by directional thermal gradients, reflecting real-world imperfections in the thermal field.
- Overall, the proposed approach offers a non-invasive, data-driven solution for real-time monitoring of laser heat treatment, with strong potential for integration into intelligent manufacturing systems.

## CRediT authorship contribution statement

**Myeonggyun Son:** Writing – original draft, Validation, Software, Investigation, Formal analysis, Data curation. **Hyungson Ki:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition, Conceptualization.

## Declaration of 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.

## Acknowledgement

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (RS-2024-00334308).

## Data availability

Data will be made available upon reasonable request.

## References

[1] G.E. Totten, Steel Heat Treatment: Metallurgy And Technologies, CRC Press, 2006.
[2] M. Yonemura, et al., Fine microstructure formation in steel under ultrafast heating and cooling, Sci. Rep. 12 (1) (2022) 2237.
[3] F.I. Jamhari, et al., Influence of heat treatment parameters on microstructure and mechanical performance of titanium alloy in LPBF: a brief review, J. Mater. Res. Technol. 24 (2023) 4091–4110.
[4] S. Santhanakrishnan, R. Kovacevic, Hardness prediction in multi-pass direct diode laser heat treatment by on-line surface temperature monitoring, J. Mater. Process. Technol. 212 (11) (2012) 2261–2271.
[5] Z. Xu, K.H. Leong, C.B. Reed, Nondestructive evaluation and real-time monitoring of laser surface hardening, J. Mater. Process. Technol. 206 (1–3) (2008) 120–125.
[6] S. Martínez, et al., Control loop tuning by thermal simulation applied to the laser transformation hardening with scanning optics process, Appl. Therm. Eng. 98 (2016) 49–60.
[7] S. Martínez, et al., Analysis of the regimes in the scanner-based laser hardening process, Opt. Lasers Eng. 90 (2017) 72–80.
[8] M.H. Farshidianfar, A. Khajepouhor, A. Gerlich, Real-time monitoring and prediction of martensite formation and hardening depth during laser heat treatment, Surf. Coat. Technol. 315 (2017) 326–334.
[9] R.H. van Blitterswijk, et al., Adaptive thermal model for real-time peak temperature and cooling rate prediction in laser material processing, J. Manuf. Process. 101 (2023) 1301–1317.
[10] R.H. van Blitterswijk, M.H. Farshidianfar, A. Khajepour, Model predictive control for real-time microstructure control in laser heat treatment, J. Manuf. Process. 118 (2024) 116–127.
[11] X. Li, et al., Deep learning-based intelligent process monitoring of directed energy deposition in additive manufacturing with thermal images, Procedia Manuf. 48 (2020) 643–649.
[12] V. Klamert, M. Schmid-Kietreiber, M. Bublin, A deep learning approach for real time process monitoring and curling defect detection in Selective Laser Sintering by infrared thermography and convolutional neural networks, Procedia CIRP 111 (2022) 317–320.
[13] I. Goodfellow, et al., Generative adversarial networks, Commun. ACM 63 (11) (2020) 139–144.
[14] S. Oh, H. Ki, Deep learning model for predicting hardness distribution in laser heat treatment of AISI H13 tool steel, Appl. Therm. Eng. 153 (2019) 583–595.
[15] M. Son, H. Ki, Deep-learning model for predicting hardness and phase distributions from two cross-sectional temperature distribution images in laser heat treatment of AH36 steel, J. Mater. Res. Technol. 27 (2023) 718–732.
[16] S. Hochreiter, Long Short-Term Memory, Neural Computation MIT-Press, 1997.
[17] D.E. Rumelhart, G.E. Hinton, R.J. Williams, Learning representations by back-propagating errors, Nature 323 (6088) (1986) 533–536.
[18] Sutskever, I., Sequence to Sequence Learning with Neural Networks. arXiv preprint arXiv:1409.3215, 2014.
[19] Cho, K., Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv preprint arXiv:1406.1078, 2014.
[20] Y. LeCun, et al., Gradient-based learning applied to document recognition, Proc. IEEE 86 (11) (1998) 2278–2324.
[21] Ballas, N., et al., Delving deeper into convolutional networks for learning video representations. arXiv preprint arXiv:1511.06432, 2015.
[22] X. Shi, et al., Convolutional LSTM network: a machine learning approach for precipitation nowcasting, Adv. Neural Inf. Proces. Syst. 28 (2015).
[23] Ulyanov, D., Instance normalization: The missing ingredient for fast stylization. arXiv preprint arXiv:1607.08022, 2016.
[24] Clevert, D.-A., *Fast and accurate deep network learning by exponential linear units (elus)*. arXiv preprint arXiv:1511.07289, 2015.
[25] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015.
[26] R.Y. Rubinstein, D.P. Kroese, The Cross-Entropy Method: A Unified Approach to Combinatorial Optimization, Monte-Carlo simulation, and Machine Learning, Springer, 2004.
[27] Kingma, D.P., Adam: A method for stochastic optimization. arXiv preprint arXiv: 1412.6980, 2014.
[28] L.R. Dice, Measures of the amount of ecologic association between species, Ecology 26 (3) (1945) 297–302.
[29] H. Ki, S. So, Process map for laser heat treatment of carbon steels, Opt. Laser Technol. 44 (7) (2012) 2106–2114.
[30] J. Kim, H. Ki, Scaling law for penetration depth in laser welding, J. Mater. Process. Technol. 214 (12) (2014) 2908–2914.