

## ARTICLE OPEN



# Deep learning approaches for instantaneous laser absorptance prediction in additive manufacturing

Runbo Jiang<sup>1</sup>✉, John Smith<sup>1</sup>, Yu-Tsen Yi<sup>1</sup>, Tao Sun<sup>2</sup>, Brian J. Simonds<sup>3</sup> and Anthony D. Rollett<sup>1,4</sup>

The quantification of absorbed light is essential for understanding laser-material interactions and melt pool dynamics in order to minimize defects in additively manufactured metal components. The geometry of a vapor depression formed during laser melting is closely related to laser energy absorption. This relationship has been observed by the state-of-the-art in situ high-speed synchrotron X-ray visualization and integrating sphere radiometry. These two techniques create a temporally resolved dataset consisting of vapor depression images and corresponding laser absorptance. In this work, we propose two different approaches to predict instantaneous laser absorptance. The end-to-end approach uses deep convolutional neural networks to learn implicit features of X-ray images automatically and predict the laser energy absorptance. The two-stage approach uses a semantic segmentation model to engineer geometric features and predict absorptance using classical regression models. While having distinct advantages, both approaches achieved a consistently low mean absolute error of less than 3.3%.

*npj Computational Materials* (2024)10:6; <https://doi.org/10.1038/s41524-023-01172-8>

## INTRODUCTION

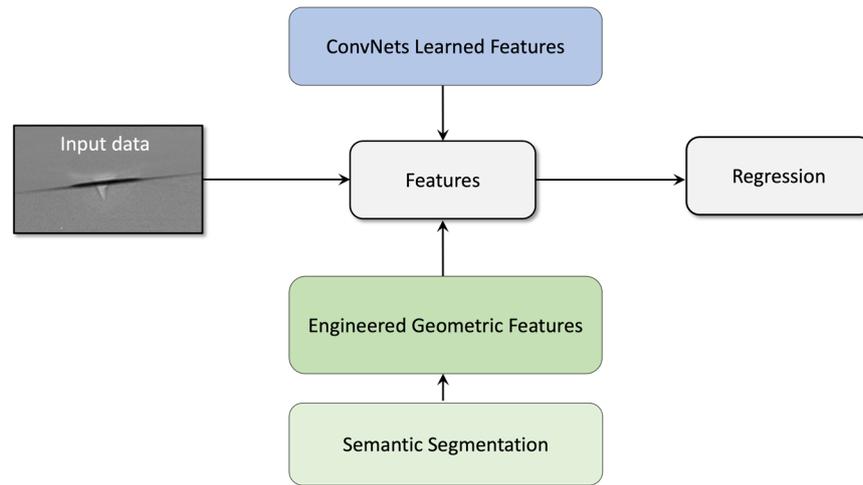
The phenomenon of keyholing is frequently observed in welding processes<sup>1</sup>, as well as in laser-based metal additive manufacturing (AM)<sup>2–4</sup>. When a sufficiently high laser power and/or slow scanning speed is used, vaporization of the metal occurs, creating pressure that pushes the melt pool surface downward. As a result, a vapor depression is formed, and as the local material temperature progressively exceeds the boiling point, the vapor depression could undergo a transition from conduction to keyhole mode<sup>5,6</sup>. This keyhole depression is characterized by multiple laser reflections and high laser absorptance by the metal substrate, which improves the overall energy efficiency and speed of the fabrication process. However, the uneven distribution of laser absorption on the keyhole wall can lead to the formation of local hotspots, resulting in an imbalance between recoil pressure, vapor dynamics, capillary forces, and Marangoni forces<sup>5,7</sup>. As a result, the gas-liquid interface within the keyhole often exhibits strong fluctuations, which are affected by laser parameters (such as power, scanning velocity, spot size, duty cycle, etc.) and the properties of the protective gas and substrate materials<sup>5,8</sup>. This significant variation in vapor depression geometry can affect the energy coupling mechanism between the laser and the material, leading to unusual melt pool dynamics and the formation of spatters and pores<sup>3,7,9,10</sup>. Therefore, it is essential to estimate or measure the absorptance of laser light, as it is the energy source causing the vapor depression formation.

Analytical and numerical modeling are powerful tools for estimating energy absorption, but they are not without costs. He et al. developed an analytical energy absorption model based on fitting the Rosenthal equation<sup>11</sup>, which provides an estimate for conduction-mode melting but has limited accuracy for keyhole mode owing to its simplifying assumptions. To address this limitation, others have used a high-fidelity multiphysics model to interpret energy absorption more accurately for deep keyholes<sup>10,12</sup>. This model can simulate 3D fields of temperature, fluid

flow velocity, pressure, and vapor chemical species with 2  $\mu\text{m}$  resolution and can predict absorptance and reflectance of laser light within a vapor depression using a laser ray-tracing method. Despite the high accuracy of this method, the accessibility of the software limits its usage. Gan et al. proposed an absorptance function that only involves a dimensionless keyhole number, a normalized diffusion length and a minimum absorptance<sup>10</sup>. The minimum absorptance is the absorptance of a material on a flat liquid surface but this exact absorptance value is difficult to obtain and therefore scarce for alloys<sup>13</sup>. Instead, Ye<sup>13</sup> used laser absorptance measured on flat sample surfaces at low laser power, which can be affected by laser power, local materials composition, surface condition, and actual surface temperature<sup>10,13</sup>. Therefore, minimum laser energy absorptance is not always available or accurate for all the materials of interest.

Direct experimental measurements of the laser light absorptance at high temperatures can be obtained by using calorimetry<sup>14,15</sup> and the integrating sphere method<sup>16,17</sup>. The former measures an average absorptance, whereas the latter measures absorptance in real time. Both methods show a significant increase in laser absorption when a keyhole is formed, with the magnitude of the absorptance increase being closely linked to the keyhole's shape<sup>15,17–19</sup>. To enhance the reliability and reproducibility of laser-based AM, it is necessary to have a real-time monitoring system that can detect defects as they occur and adjust the process on the fly. Recently, the National Institute of Standards and Technology (NIST) and Argonne National Laboratory (ANL) integrated in situ synchrotron X-ray imaging with integrating sphere radiometry (ISR), allowing real-time and simultaneous measurement of the projected vapor depression geometry and absolute laser energy absorptance<sup>17,18,20</sup>. This time-resolved measurement revealed a strong correlation between the instantaneous depression geometry (i.e., depth, width, and area) and laser absorptance that was not previously available<sup>20</sup>.

<sup>1</sup>Department of Materials Science and Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA. <sup>2</sup>Department of Materials Science and Engineering, University of Virginia, Charlottesville, VA 22904, USA. <sup>3</sup>Applied Physics Division, National Institute of Standards and Technology, Boulder, CO 80305, USA. <sup>4</sup>NextManufacturing Center, Carnegie Mellon University, Pittsburgh, PA 15213, USA. ✉email: runboj@alumni.cmu.edu



**Fig. 1** Schematic of the end-to-end approach and modular approach in predicting laser energy absorptance from a synchrotron X-ray keyhole image.

Although this experimental method provides cutting-edge performance, the setup is limited to large and expensive synchrotron facilities that are in high demand. To overcome these restrictions and extend the measurement of laser absorption, we built a temporally resolved vapor depression-absorption dataset using the in situ synchrotron X-ray imaging with ISR at ANL and applied machine learning (ML) models to interpret unprocessed X-ray images to predict absorptance. As the model only relies on the geometric features of the vapor depression, this ML approach can be generalized to unseen X-ray images that the model was not trained on, which are far more abundant throughout the years of development of the in situ monitoring techniques than the simultaneous X-ray images and absorption measurements<sup>5,9,21–23</sup>. Therefore, this pipeline greatly reduces the need for costly experiments and takes us one step closer to the goal of connecting industrial process parameters and predictive models that require knowledge of the amount of absorbed energy<sup>20,24</sup>.

In this study, our aim was to predict instantaneous laser absorptance for individual vapor depression X-ray images, irrespective of their position in the video sequence. To achieve this goal, we developed two methods: the end-to-end approach and the modular approach, as illustrated in Fig. 1. The end-to-end approach employs convolutional neural networks (ConvNets) to automatically learn feature kernels through training and directly produces an absorptance value for each X-ray image using the fully connected regression layer. The modular approach first generates a semantic segmentation mask for a vapor depression image, then extracts the geometric keyhole features, and finally applies classical regression models. Our motivation for the second approach is that many artificial intelligence (AI) tasks can be solved by designing the right set of features to extract for the task and then feeding these features to a simple machine learning algorithm<sup>25</sup>. This is especially applicable in this case, where we already have a clear understanding of which features are relevant and should be extracted. On the other hand, the main advantage of the end-to-end approach over the modular approach is that it automatically detects the important features and makes predictions without the explicit image segmentation and feature extraction process. However, the modular approach provides better model interpretability and understanding of the correlation between geometric keyhole features with laser energy absorption<sup>26</sup>.

Similar to traditional microstructure quantification<sup>27,28</sup>, manual measurements on X-ray vapor depression images using image processing programs such as ImageJ<sup>29</sup> are time-consuming and

laborious processes, subject to operator variability<sup>28</sup>. Classic computer vision techniques, such as image thresholding and morphology operations, can automate the segmentation process, which is much faster and reproducible<sup>30–32</sup>. The OpenCV library<sup>33</sup> is a viable option for this automation task. However, this method is not robust to slight changes in images or sample conditions because it requires multiple parameters to be tuned sequentially, which are highly dependent on the image characteristics<sup>29,32</sup>. Recently, ConvNets have produced superior segmentation results on biomedical images<sup>34</sup>, microstructures<sup>28,35</sup>, metal powders<sup>36</sup> and vapor depressions<sup>37</sup>. These networks are easier to implement and more robust against variations in images. Semantic segmentation is the process of classifying each pixel in an image into a specific category<sup>38,39</sup>. This task is often performed using ConvNets with encoder-decoder architectures<sup>39</sup>. The encoder uses learned filters to extract semantic content from the input image, converting it into a compact latent representation vector, which is subsequently mapped by the decoder onto each pixel to generate a pixel-wise classification of the objects in the original image, achieving semantic segmentation<sup>39</sup>. This approach is more accurate, robust, and less dependent on image characteristics compared to traditional techniques. To accomplish the semantic segmentation task, we created a vapor depression segmentation dataset using human-annotated ground-truth masks from four different metallic materials. This dataset covers both stationary and scanning laser conditions. In both absorptance prediction approaches, we experimented with different models and weights initialized from pretrained ImageNet weights and randomly initialized weights to evaluate their effectiveness. In addition, we compared the two approaches based on their accuracy, generalizability, interpretability, cost of training and inference, and suitable application scenarios for each method.

This paper contributes in the following ways:

- Establishing two benchmark datasets for X-ray vapor depression image segmentation and laser energy absorptance. A portion of the absorptance dataset is openly available, while the entire segmentation dataset is publicly released.
- Exploring two approaches to predict laser energy absorptance and discussing the trade-offs of each method. Additionally, we have publicly released the trained models.
- Establishing a relationship between the geometric characteristics of a vapor depression and the absorption of laser energy. Therefore, it provides an easy-to-use and experiment-validated laser absorption model for Ti-6Al-4V.

**Table 1.** Laser energy absorption dataset obtained without a powder layer.

Material	Laser condition	Laser power - scan speed	Special condition	Absorptance mean $\pm$ SD (%)	Image number
Ti64	Stationary	94 W - 0	-	40.5 $\pm$ 14.1	227
Ti64	Stationary	100 W - 0	-	48.6 $\pm$ 16.8	227
Ti64	Stationary	106 W - 0	-	53.1 $\pm$ 19.2	233
Ti64	Stationary	128 W - 0	-	67.4 $\pm$ 18.0	233
Ti64	Stationary	168 W - 0	-	75.6 $\pm$ 17.2	232
Ti64	Scanning	128 W - 0.7 m/s	-	31.1 $\pm$ 3.8	245
Ti64	Scanning	197 W - 0.7 m/s	-	50.7 $\pm$ 11.8	224
Ti64	Scanning	197 W - 0.7 m/s	in air	51.1 $\pm$ 9.8	250
Ti64	Scanning	197 W - 0.7 m/s	no skywriting	53.0 $\pm$ 13.3	250
Ti64	Scanning	254 W - 0.7 m/s	-	64.2 $\pm$ 11.5	225

**Table 2.** Laser energy absorption dataset obtained with a 100  $\mu$ m Ti64 powder layer for fine-tuning ConvNets models.

Material	Laser condition	Power - Scan speed	Special condition	Absorptance mean $\pm$ SD (%)	Image number
Ti64	Stationary	111 W - 0	-	51.1 $\pm$ 12.3	248
Ti64	Stationary	197 W - 0	-	83.1 $\pm$ 15.2	305
Ti64	Scanning	197 W - 0.7 m/s	-	70.4 $\pm$ 12.0	248
Ti64	Scanning	254 W - 0.7 m/s	-	78.6 $\pm$ 13.7	224

## RESULTS

### Laser energy absorptance datasets

Table 1 presents the absorptance dataset obtained from experiments conducted without a powder layer, and additional experimental information can be found in ref. 20. This dataset comprises 1152 X-ray images from spot welding (SW), i.e., stationary laser experiments, and 1194 images from scanning laser experiments on a 0.3-mm-thick Ti-6Al-4V (Ti64) substrate. To avoid the group leakage problem, we hold out the entire 100 W stationary experiment and 197 W with no skywriting scanning laser experiment as test dataset instead of using a random split strategy. In Table 1, the special condition *in air* indicates that the experiment chamber was filled with air instead of argon. The absence of the *skywriting* feature indicates that the laser scanning speed was gradually increased from zero to the desired value, while for all other runs, this feature was incorporated into the scanning optics to eliminate the initial ramp. The remaining 1870 images were randomly split into train/validation sets with an 80/20 ratio.

The absorptance dataset presented in Table 2 was acquired from experiments conducted on Ti64 substrates with a 100  $\mu$ m Ti64 powder layer. This dataset consists of 553 X-ray images obtained from stationary laser experiments and 472 images from scanning laser experiments. This dataset was held out for fine-tuning the top-performing models trained on the dataset without powder. During the fine-tuning process, 10% of the dataset (116 frames) was set aside for validation purposes. We incrementally varied the size of the fine-tuning dataset from 0% to 40% to investigate the influence of the training data size on the prediction outcomes. The remaining images were reserved for testing purposes.

Each X-ray image of a vapor depression is annotated with its corresponding laser energy absorptance. The time-resolved absorptance evolution for each experiment is depicted in Supplementary Fig. 1. In Tables 1 and 2, the laser absorptance column presents the mean and standard deviation for each experiment, starting from when the laser was turned on to when it was turned off. It is worth noting that as the melting can transition

between conduction and keyhole mode, a significant variance is usually observed.

Prior to training, the images were preprocessed, including denoising by dividing each image by a background image captured before the laser was turned on, normalization by stretching the histogram of each image to have a maximum pixel value of 255 and a minimum pixel value of 0, and center cropping to 350  $\times$  350 pixels while ensuring that the vapor depression remained approximately at the center of the image. Some parts of data from these experiments are openly available at the NIST Public Data Repository<sup>40</sup>.

### Segmentation dataset

To collect vapor depression images for segmentation, the experiment only required a less complicated X-ray imaging setup. This allowed us to cover a broader range of samples, including Ti64, Stainless Steel 316 (SS316), Inconel 718 (IN718), and aluminum alloy 6061 (AA6061) without powder layer, and diverse vapor depression geometries, including conduction, transition and keyhole mode (see Supplementary Table 1). Table 3 lists the processing parameter setup and the number of training images collected for each material. Creating a ground-truth mask for each X-ray image that accurately depicts the location and shape of the depression is important in training segmentation models. A variety of tools in ImageJ<sup>29</sup>, including the wand tool, polygon selection, and fit spline, were used to manually accomplish this goal. The limited segmentation dataset size is due to the manual and time-intensive process of generating accurate ground-truth masks.

X-ray images have varying brightness, contrast, and noise due to differences in sample thickness, atomic number, and surface roughness. Thicker and heavier metal plates result in blurred X-ray images due to increased X-ray scattering and higher attenuation, which diffuse and reduce the X-ray signal, respectively. Lighter materials such as Al6061 allow for a thickness of up to 1.2 mm while still achieving a strong contrast along the vapor depression boundary, while heavier materials such as IN718 must be as thin as 0.4 mm to achieve the same level of contrast. However, we

**Table 3.** Vapor depression segmentation dataset obtained from the melting process of alloys under various parameters using synchrotron X-ray imaging.

Materials	Laser condition	Sample thickness (mm)	Image number	Power (W) - Scan speed (m/s)
Al6061	Scanning	1.2	119	610 - 1.2, 630 - 1.2, 740 - 1.4, 850 - 0.9
IN718	Scanning	0.7	110	300 - 2.5, 400 - 3.0, 630 - 0.75, 630 - 1.25, 800 - 4.2
SS316	Scanning	0.4	153	99 - 0.4, 300 - 0.4, 300 - 0.6, 408 - 0.4, 408 - 0.6
Ti64	Scanning	0.4	153	111 - 0.4, 139 - 1.2, 154 - 0.4, 197 - 0.6, 311 - 0.7, 311 - 1.0, 520 - 1.0, 540 - 0.7
Ti64	Stationary	0.4	224	220 - 0, 270 - 0, 425 - 0

**Table 4.** End-to-end model performance on test absorption data without powder.

	Metric	Resnet-50	ConvNeXt-T
Trained from scratch	Train loss	0.74 ± 0.17	4.51 ± 0.27
	Val loss	1.42 ± 0.20	4.67 ± 0.47
	Test loss	4.70 ± 1.83	13.70 ± 3.04
	# epochs to converge	201 ± 34	204 ± 29
Pretrained on ImageNet	Train loss	0.79 ± 0.07	0.31 ± 0.03
	Val loss	1.5 ± 0.24	1.13 ± 0.16
	Test loss	4.19 ± 0.31	2.35 ± 0.35
	# epochs to converge	163 ± 49	63 ± 15

The loss values are evaluated using the smooth L1 loss with absorptance a unit of percentage (%).

intentionally used a 0.7 mm thick IN718 sample to naturally blur the keyhole boundary and increase the segmentation difficulty. For each set of experiments, images were denoised by dividing each image by a background image and normalized to the range of [0, 255], then center cropped with the depression at the center of the images. Owing to the differences in the size of the depression and the imaging field of view, the resulting images have various dimensions. To keep the input data of the same dimension, images were padded into 576 × 576 pixels using the mean pixel value of each image. The images were partitioned randomly into training, validation, and testing sets using a 75/15/15 ratio.

### End-to-end approach

Five different sets of training and validation splits were randomly generated to train the models. The performance of the end-to-end approach is presented in Table 4. The number of epochs required for the model to reach convergence is defined as the first occurrence of the validation loss being less than 105% of the final loss at 300 epochs. For ResNet-50, the ImageNet pretrained weights reduce the number of epochs required for the model to reach convergence by 19% and reduce the test loss by 11% compared with its counterpart trained from scratch. However, in ConvNeXt-T, the ImageNet weights reduce the test loss by 76% and the number of epochs to converge by 69%. Besides, pretrained ConvNeXt-T reached the lowest test loss and required the minimum number of training epochs to converge among all four models. Therefore, with the improved training configuration described in the method section, ImageNet pretrained weights were highly effective in ConvNeXt but had a less significant impact on ResNet. Both models achieved a test loss of less than

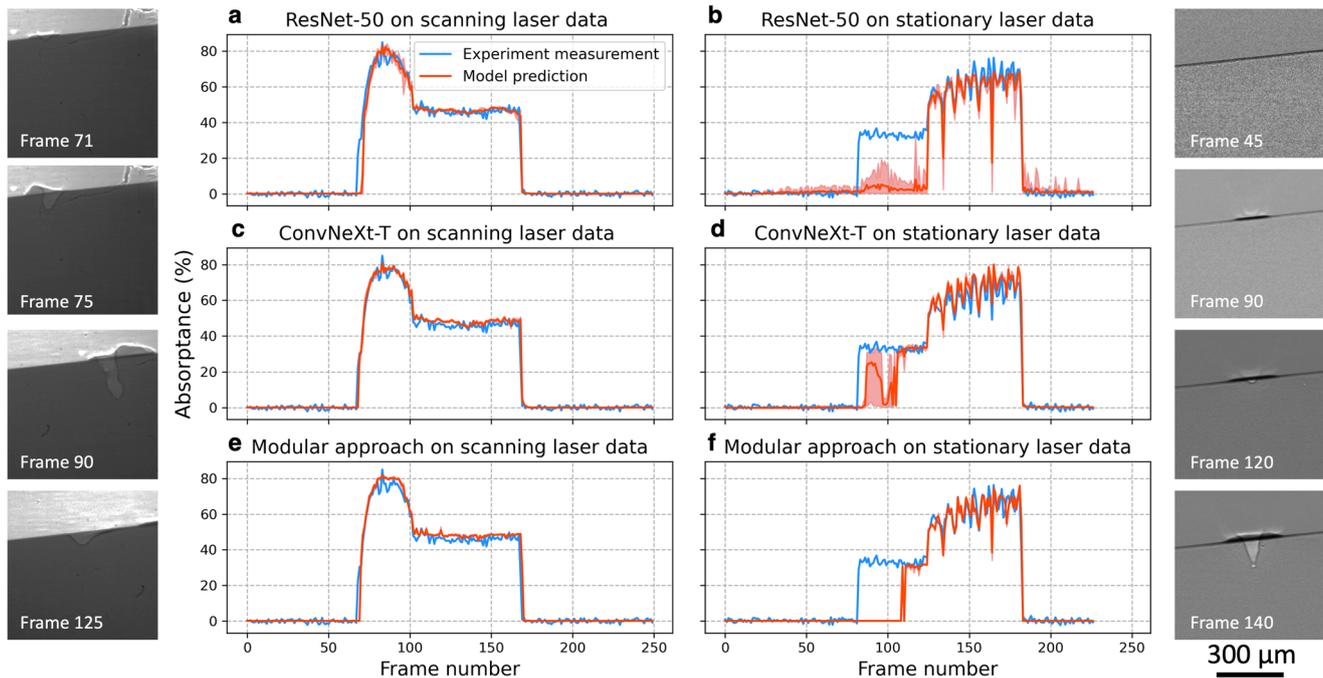
4.7, indicating excellent performance as the absorptance target values range from about 0 to 100.

Figure 2a–d displays the experimentally measured absorptance values and the predicted values using the ConvNets models. In the stationary laser test data, a lower energy of 100 W was used, which resulted in an initial stage between frames 83 to 105. During this period, heat accumulation and vapor plume development occurred before the laser vaporized the base material and a vapor depression formed. The 90th frame in the rightmost column of Fig. 2 is representative of this initial stage. This image shows a discernible semicircular melt pool and cone-shaped artifact at the material surface from X-ray scattering. In the second stage, from frames 106 to 125, a semicircular-shaped keyhole with a depth of approximately 10 μm formed. This stage is consistent with the stable formation and growth stage of the vapor depression and plume in ref. <sup>41</sup>. In the third stage, the depression depth expands abruptly to over 10 μm after the 125th frame, causing multiple reflections of the laser light within the depression before it escapes<sup>5</sup>. This results in a significant increase in laser absorptance, from approximately 33 to over 50%. As the laser is stationary, the same region continues to be heated, leading to a gradual increase in laser absorptance during the third stage. As shown in Fig. 2b, ResNet-50 is able to capture the third stages but fails to accurately predict the first and second stages. On the other hand, ConvNeXt-T in Fig. 2d succeeds in the second stage; however, it still exhibits significant variations in the first stage.

This stage transition was also observed in the moving laser case, as shown in the leftmost column of Fig. 2. However, as the laser energy used in this experiment was 197 W, which was strong enough to vaporize the base material temporarily within the current recording frame rate, the resulting transition could only be observed in two frames, 70 and 71. In the scanning laser experiment in Fig. 2, a high absorptance and a deep keyhole were initially observed. This observation occurred during the early stages of the experiment when vaporization was dominant before the steady-state convective fluid flow was established. As there was no trailing melt pool or associated convective flow to dissipate heat away from the laser-metal interaction region, a deeper keyhole formed, and the absorptance reached 80%. While the laser began to move faster across the material from frame 85, the absorptance quickly decreased to approximately 45% and stabilized at this value owing to the development of strong Marangoni convection near the rear keyhole wall. This convection moved the hotter molten metal toward the colder region, opposite to the direction of the laser scan<sup>5,20,42</sup>. Both models accurately predict the instantaneous changes in laser absorption during this scanning laser experiment.

### Modular approach

Similar to the end-to-end approach, five training and validation splits were used in segmentation model training and evaluation. Table 5 presents the performance of the five segmentation models. In segmentation evaluation, convergence is defined as the total number of epochs required during training when an early



**Fig. 2** The time-resolved laser absorption prediction results on test dataset without a powder layer. End-to-end approach using **a, b** ResNet-50 and **c, d** ConvNeXt-T. **e, f** Modular approach using engineered vapor depression geometric features and random forest regression model. The shaded region represents the range of maximum and minimum values predicted by the models trained on five different training data splits. The images on the leftmost column are representative scanning laser X-ray images, and the ones on the rightmost columns are stationary laser images, and all X-ray images share the same scale bar.

**Table 5.** Performance of semantic segmentation models on vapor depression X-ray images.

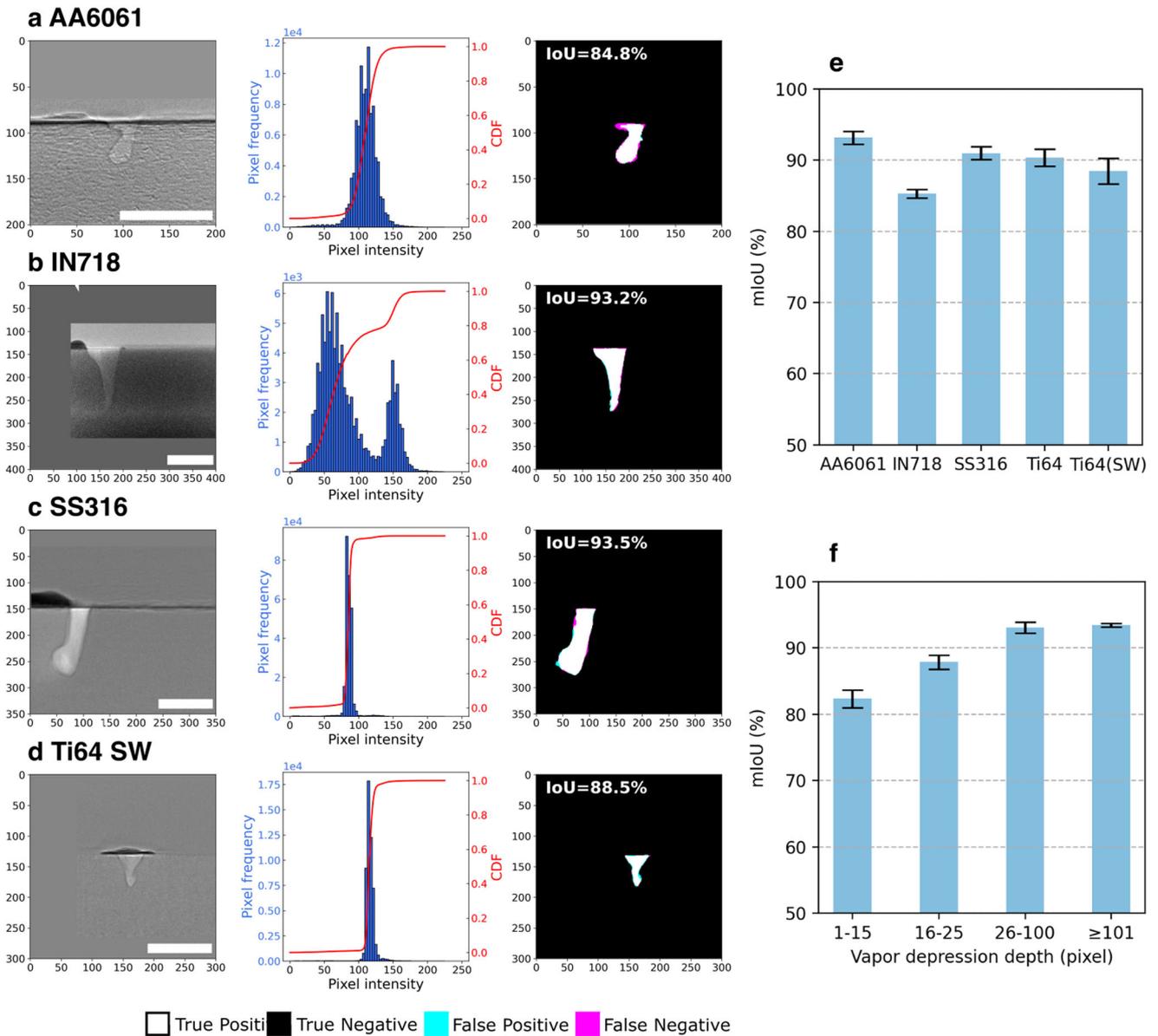
Segmentation model	# param.	Metric	Trained from scratch	Pretrained on ImageNet
UNet	17.2 M	Test mIoU (%)	90.4 ± 0.6	–
		Converge epoch	123 ± 11	–
UNet + ResNet-50	32.5 M	Test mIoU (%)	89.5 ± 1.3	89.5 ± 0.8
		Converge epoch	142 ± 23	103 ± 20
UNet++ + ResNet-50	49.0 M	Test mIoU (%)	90.1 ± 0.5	89.9 ± 0.6
		Converge epoch	137 ± 17	117 ± 9
DeepLabV3 + ResNet-50	39.6 M	Test mIoU (%)	87.6 ± 0.5	88.4 ± 0.9
		Converge epoch	146 ± 20	156 ± 32
DeepLabV3++ + ResNet-50	26.7 M	Test mIoU (%)	89.5 ± 0.9	89.8 ± 0.7
		Converge epoch	127 ± 11	134 ± 19

stop was triggered. We do not find a significant difference in mean intersection over union (mIoU) and converge epoch between the models with the encoder initialized using pretrained ImageNet weights and those with encoder initialized randomly.

Out of the five architectures explored, the simplest one, UNet, outperforms all the others. This could be due to the fact that the training data are grayscale X-ray images, providing relatively simple features that do not benefit from the more complex models. Hence, UNet is used for the downstream applications in this work. Figure 3 shows representative segmentation results of the UNet model on different materials. The pixel value used for padding the image to the same dimension was removed in the histogram to make the figures readable. As evident in Fig. 3f, UNet performs better as the size of the vapor depression increases and struggles with smaller depressions that have finer and less distinct boundaries. The performance of UNet on different materials is summarized in Fig. 3e. The lowest mIoU was observed for IN718,

whose pixel intensity follows a distinct bimodal distribution in Fig. 3b. The reason for this is that the IN718 sample thickness was 0.7 mm, which is considerably thicker than the standard 0.4 mm for heavy metals, resulting in more blurred and noisy images.

After extracting the geometric features from the segmented masks, classic regression models can be employed to predict laser absorbance. The hyperparameters used for training these models are listed in Supplementary Table 2. The only feature that has missing values is the front wall angle, which is not applicable for spot welding cases. To accommodate models that cannot handle missing values, we impute the missing values with 0, as it avoids introducing bias or artificially influencing the data analysis and modeling process. Such imputation was not performed for XGBoost as it can handle missing values directly<sup>43</sup>. Table 6 displays the performance of various models in the modular approach. Linear regression serves as a baseline due to its simplicity, computational efficiency, and ability to provide a

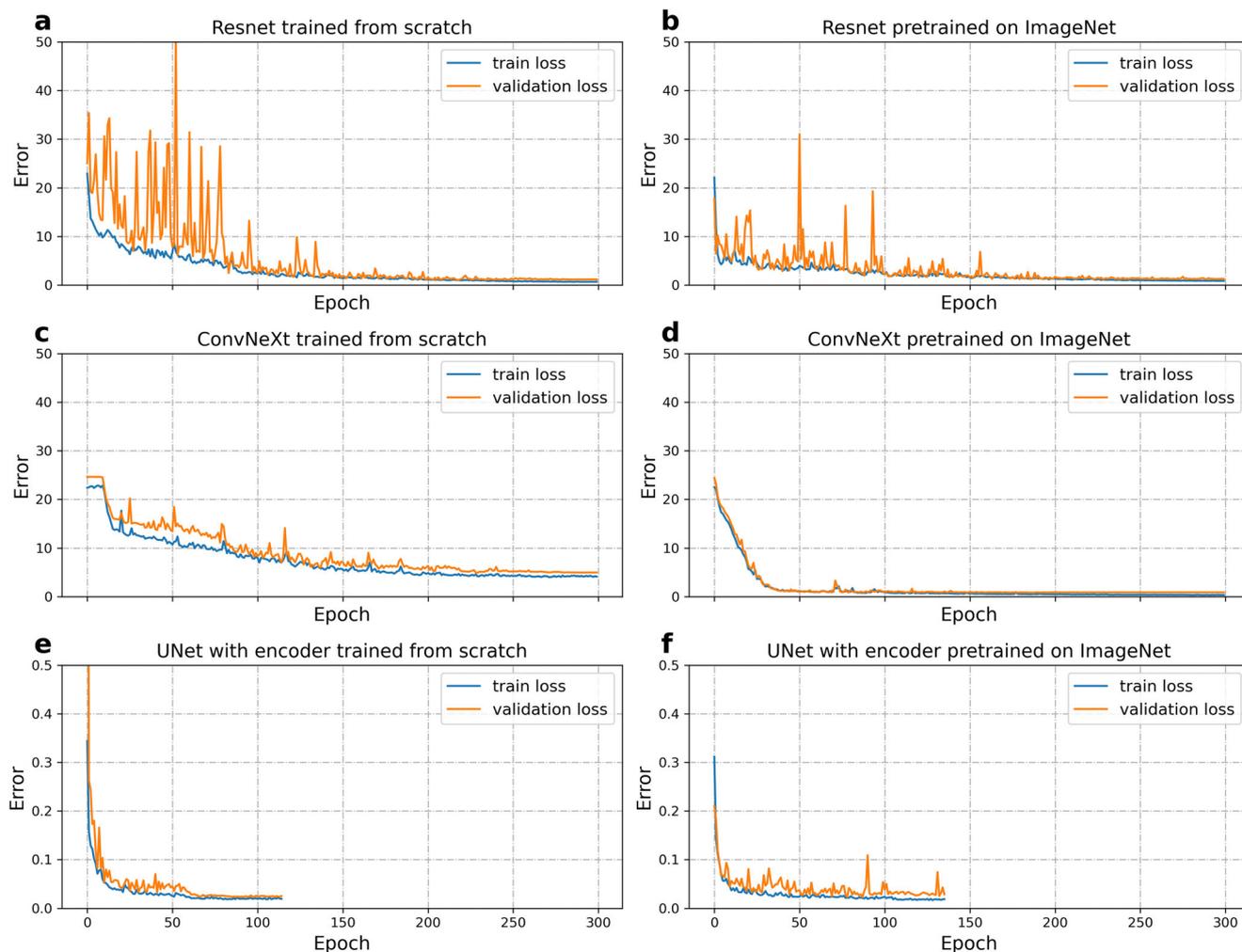


**Fig. 3 Segmentation results from test data on different materials and processing conditions using the UNet model.** Explementary scanning laser images for **a** AA6061, **b** IN718, **c** SS316, and stationary laser images for **d** Ti64. Each example in (a–d) includes three subfigures: the vapor depression X-ray image with the scale bar representing 200  $\mu\text{m}$ , the resulting pixel intensity histogram and cumulative distribution function (CDF) derived from the X-ray image, and the corresponding segmentation accuracy mask. In each mask, white pixels indicate true positive predictions, black represents true negatives, cyan represents false positives, and magenta represents false negatives. The performance of UNet for **e** different materials and **f** varying vapor depression depth, where 1 pixel is equivalent to 1.923  $\mu\text{m}$ . The statistics presented are averaged using the mIoU metric across 5 training splits.

Model	Train loss	Val loss	Test loss
Linear regression	6.34 $\pm$ 0.09	6.25 $\pm$ 0.27	6.59 $\pm$ 0.18
Decision tree	2.53 $\pm$ 0.06	2.75 $\pm$ 0.19	3.54 $\pm$ 0.07
Random forest	1.94 $\pm$ 0.03	2.44 $\pm$ 0.12	3.30 $\pm$ 0.02
XGBoost	1.78 $\pm$ 0.02	2.76 $\pm$ 0.2	3.55 $\pm$ 0.06

Loss values are calculated using smooth L1 loss.

benchmark for more advanced models. However, as the linear assumption is too simplistic to fully capture the intricacy of the relationship between the independent and dependent variables<sup>44</sup>, linear regression has the highest test loss and bias, as shown in Table 6. The decision tree model is able to capture nonlinear relationships between variables and achieves better performance than linear regression in this scenario. Bagging models like random forest (RF)<sup>45</sup> and boosting models like XGBoost<sup>43</sup> outperform decision tree since they combine multiple weak models into a stronger ensemble model. RF is the model that performs the best on the test dataset and its prediction is shown in Fig. 2e, f. Compared to the end-to-end approach, the performance of the



**Fig. 4** Training and validation error on the absorption dataset without a powder layer of both scanning laser experiments and stationary laser experiments. **a, b** ResNet-50, **c, d** ConvNeXt-T, and **e, f** on segmentation dataset with UNet whose encoder was replaced with ResNet-50. The error function in (**a–d**) is smooth L1 loss, and in (**e, f**) is the weighted binary cross entropy loss and dice loss.

modular approach is limited by the formation of a distinguishable keyhole and a successful image segmentation and feature extraction process. In the stationary laser case shown in Fig. 2f, during the initial stage, there was no vapor depression formed yet for the segmentation model to extract, which resulted in all feature values being zero and, in turn, resulted in zero (apparent) absorptance. However, the end-to-end approach is capable of learning such scenarios better than the modular approach. Once a vapor depression of significant size has formed (with a depth of at least 5 pixels or 10  $\mu\text{m}$ ), there is no discernible difference in the performance of the two approaches.

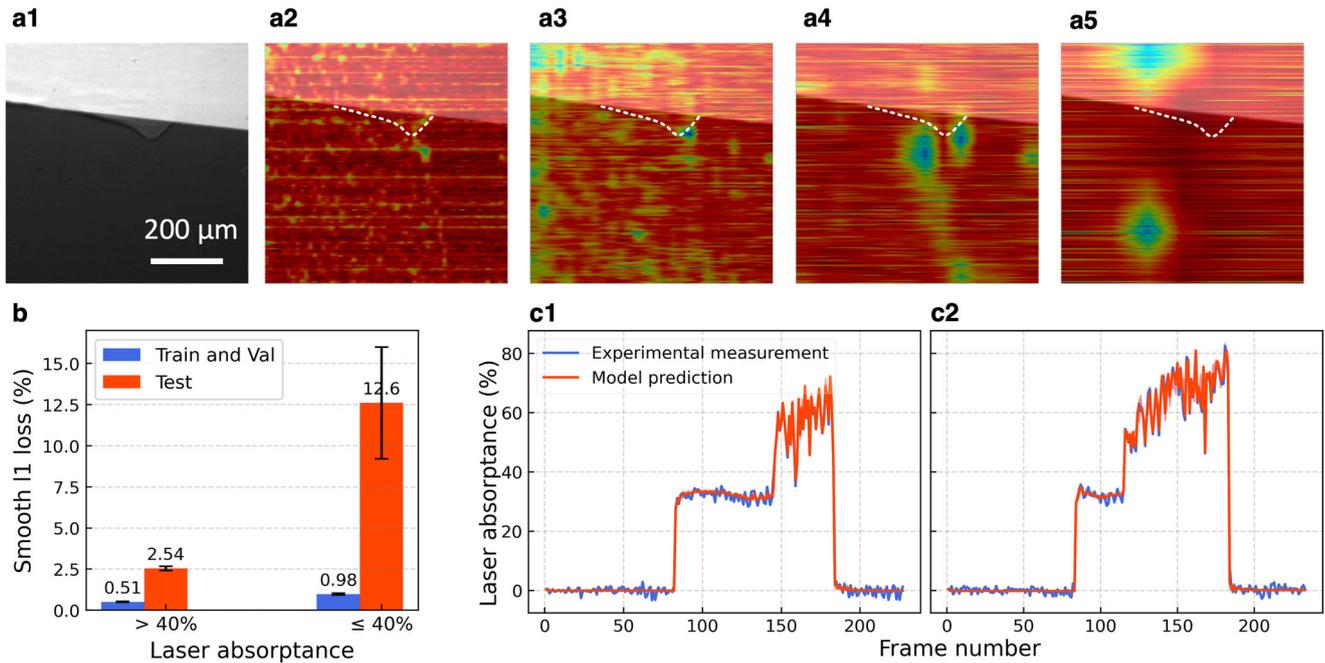
## DISCUSSION

We dissect the discussion into four critical facets. We start by examining the influence of pretrained ImageNet weights on model performance, followed by an exploration of model interpretation. We then dive into the efficacy of fine-tuning the powder dataset and conclude with practical recommendations for selecting the most appropriate approach. These elements collectively offer a more comprehensive view of our findings and their implications.

Transfer learning is an important research area in machine learning because it allows us to leverage knowledge learned from

one task or domain to improve the performance of another related task or domain. In this work, we investigated the impact of learned weights from ImageNet, which is a large dataset widely used in computer vision research communities for benchmarking state-of-the-art models<sup>46,47</sup>. ImageNet consists of over 1.4 million images and 1000 categories, including animals, vehicles, household items, natural phenomena, and more. In this study, the effectiveness of pretrained weights is particularly apparent in the end-to-end approach. Table 4 demonstrates that ResNet-50 and ConvNeXt-T with transfer learning converged in fewer iterations compared to their counterparts initialized from random, perhaps because of the more rapidly decreasing loss values and more stable loss curves observed in Fig. 4a–d. Besides, similar to this application to synchrotron X-ray images, transfer learning has been widely used in medical applications, such as the use of chest X-ray images to detect COVID-19 pneumonia<sup>48,49</sup>. These results indicate that transfer learning provides a superior initialization of ConvNets parameters compared to random initialization, even when using weights learned from vastly different natural images.

The effectiveness of transfer learning is influenced by various factors such as the similarity between the pretraining and target domains, the amount of target data, the complexity of the target task, and specific model architectures, as demonstrated by Huh and Stuckner<sup>28,47</sup>. According to Huh<sup>47</sup>, the high diversity and size



**Fig. 5 ConvNeXt-T model interpretation and error analysis.** **a1** original X-ray image depicts a vapor depression formed in a scanning laser experiment, with the laser scanning from left to right. The Grad-CAM visualization is generated from the last convolutional layer of different stages of the ConvNeXt-T model, specifically **a2** first stage, **a3** second stage, **a4** third stage, and **a5** last stage. The color bar represents the degree of importance. **b** Error analysis of the ConvNeXt-T model threshold on 40% laser absorptance. **c** Performance evaluation of ConvNeXt-T model on two SW experiments in training dataset with laser power: **c1** 94 W and **c2** 106 W.

of the ImageNet dataset, combined with the high level of visual similarity between different categories, enables the pretrained models to learn generalizable features that are relevant for many other visual recognition tasks. This could be the reason that pretrained weights significantly reduced the converge epoch and test loss in the end-to-end approach, especially in the ConvNeXt-T case. However, when using the pretrained weights in the ResNet-50 encoder of the segmentation models, the loss does not seem to be affected by the pretrained weights as much as shown in Fig. 4e, f and Table 5. The diminishing effect of ImageNet weights on the ResNet-50 encoder is likely because of the sequential combination of the encoder and decoder in segmentation models and the higher dissimilarity between the pretraining classification task and semantic segmentation task<sup>47</sup>.

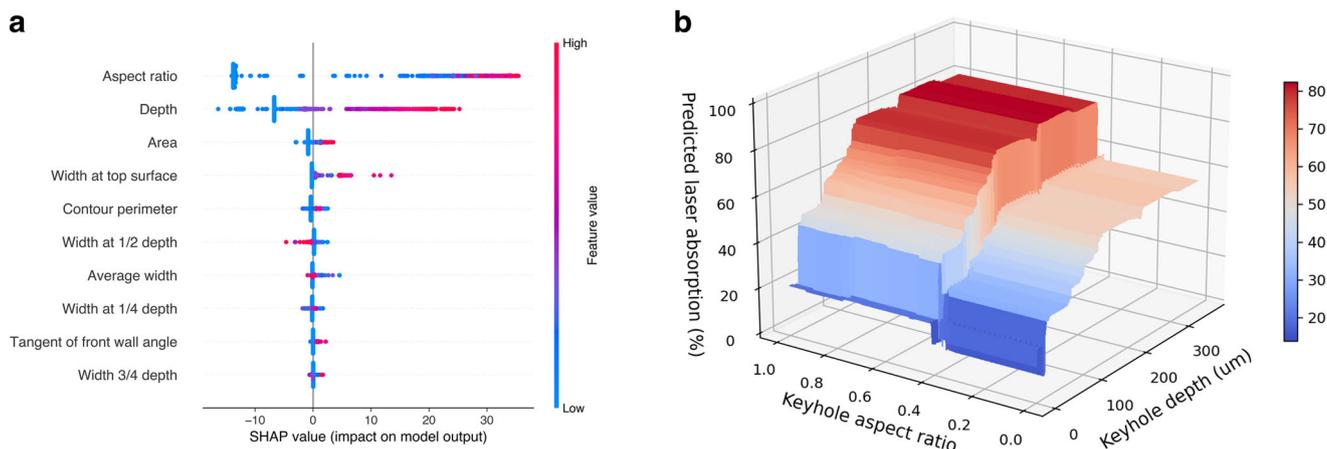
When discussing the trade-offs and considering which approach to deploy in production, it is important to not only evaluate the model's performance, as measured by standard metrics, but also its other properties, such as the amount of data, total compute time, and time required for training, as well as inference latency and model interpretability<sup>50</sup>. While deep learning models offer exceptional performance, their inability to decompose into individually understandable components makes it difficult for humans to comprehend the algorithms<sup>51,52</sup>. Researchers and practitioners have invested significant efforts in developing techniques to interpret deep learning models<sup>51</sup>. This helps to increase model transparency, which is essential for identifying the failure modes and building appropriate trust in the predictions made by the models<sup>52</sup>.

Selvaraju et al.<sup>52</sup> demonstrated the use of Grad-CAM (Gradient-weighted Class Activation Mapping) to generate "visual explanations" for ConvNet-based model decisions. This technique computes the gradients of the network's output with respect to the feature maps of the selected convolution layer. These gradients are then used to generate a heat map, which highlights the significance of regions in images for predicting the final outcome. Figure 5a shows the Grad-CAM at four stages of

ConvNeXt-T, superimposed on the 147th frame of the moving laser test data. The color bar indicates the degree of importance, with warm colors indicating higher importance. Clearly, at different stages of the ConvNeXt model, the image is being processed and the region of interest is gradually concentrated. Eventually, in Fig. 5a5, the model is able to focus more on the area where the depression appears rather than the entire image or the irrelevant background.

To investigate the cause of the failure in the conduction mode SW melting shown in Fig. 2, an error analysis was conducted for the ConvNeXt-T model. This analysis focused on the X-ray images where the laser was activated, excluding the background images. In Fig. 5c, the error on vapor depression images with absorptance greater than 40% is significantly smaller compared to images with absorptance less than 40%. A 40% threshold was used to separate the data into two groups, as this value conveniently separates conduction mode vapor depression from keyhole mode vapor depression in the absorptance dataset. The more severe test error in the conduction mode vapor depression can be explained by the model's behavior on the training data with similar absorptance. ConvNeXt-T's performance on the full absorption dataset is provided in Supplementary Fig. 2. In Fig. 5c, it is evident that the model effectively captures features in the keyhole mode, accurately predicting its absorptance and fluctuations. However, for the conduction mode, the model seems to learn an average value without significant fluctuations in its predictions. This could be attributed to the network's inability to learn impactful features in conduction mode, as the vapor depression remains stable and small in size. Additionally, the imbalanced amount of conduction mode images with keyhole mode images could make this problem even more.

There typically exists a trade-off between model performance and interpretability. Classical machine learning algorithms that involve using statistical and mathematical models to learn patterns in data and make predictions or decisions based on those patterns are generally more explainable than neural



**Fig. 6 Random forest model interpretation.** **a** The SHAP summary plot depicts the engineered geometric features of the vapor depression data and their impact on the random forest model. The color of the dots indicates the value of the feature. **b** The decision surface of the simplified random forest model assumes a constant vapor depression size of  $6000 \mu\text{m}^2$ . The color of the surface represents the predicted absorbance value.

networks. A distinct advantage of decision tree-based classifiers is the ability to measure the importance of each feature to the classification or regression being predicted<sup>53</sup>. In order to identify the features that have the greatest impact on the model, a summary plot of Shapley Additive exPlanation (SHAP) values<sup>54</sup> was created for the random forest model in Fig. 6a. This summary plot provides a way to visualize the average impact of each feature on the model's output across all predictions in a dataset. Therefore, it can be used to gain a global understanding of the feature importance. In Fig. 6a, the y-axis displays the variable name in order of global importance, with the most important feature at the top. The top three feature dependencies are vapor depression's aspect ratio, depth, and area. The SHAP values are represented on the x-axis. Each point in the plot corresponds to a row in the original dataset, with the gradient color indicating the original value of the respective variable. A positive SHAP value for a specific feature indicates a positive association between the presence of that feature and the model's prediction, and vice versa.

To evaluate the effectiveness of the top three features, as indicated in the summary plot in Fig. 6a, we trained random forest models exclusively on the top three features, resulting in train, validation, and test losses of  $2.19 \pm 0.04$ ,  $2.43 \pm 0.16$ , and  $3.36 \pm 0.02$ , respectively. The marginal increase in test loss when reducing features from 10 to 3 suggests that while the top three features were influential, the additional features in the initial RF model may have provided complementary information that led to improved prediction accuracy. The decision surface of the simplified model is depicted in Fig. 6b. To visualize the surface in a two-dimensional representation, we applied a constant vapor depression area of  $6000 \mu\text{m}^2$ , enabling us to observe the impact of the two most significant features on the prediction. Therefore, the modular approach offers exceptional interpretability.

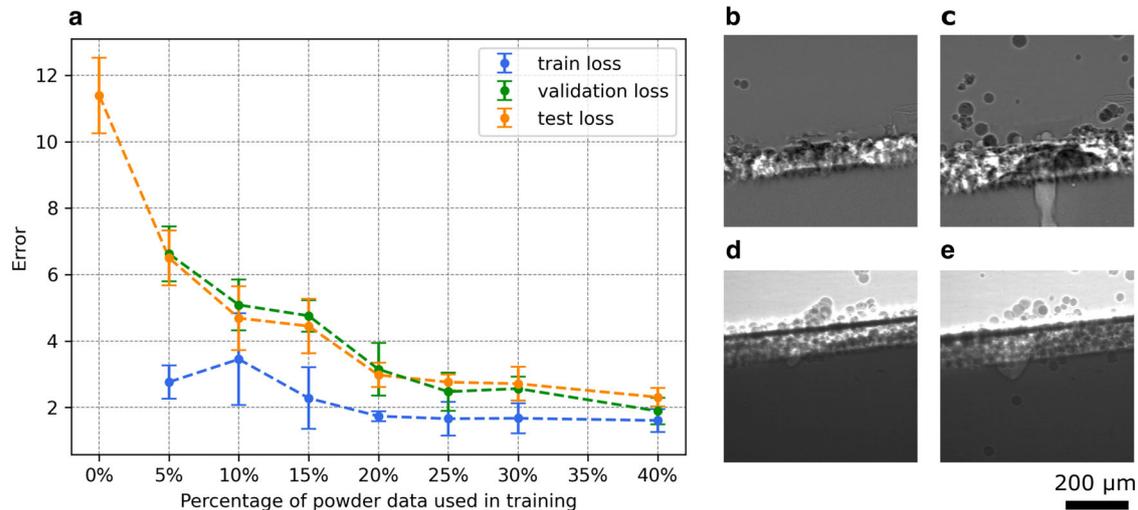
One of the major challenges in machine learning is generalization, which refers to the ability of AI systems to perform well on inputs that differ from their training data distribution<sup>55</sup>. As the absorption models in the end-to-end approach and modular approach were only trained on keyhole images without powder in Table 1, we do not expect these models to exhibit comparable performance on images acquired through powder experiments without fine-tuning. In fact, the test loss without fine-tuning measures  $11.39 \pm 1.14$  in the end-to-end approach, which is approximately five times higher than the test loss recorded on the dataset for which the model was originally trained.

To overcome this issue and improve the generalizability of the models to out-of-distribution data, we expanded the absorption dataset to incorporate data acquired with a layer of Ti64 powder, as listed in Table 2, and we fine-tuned the top-performing ConvNeXt-T models on this powder dataset, i.e., the ConvNeXt-T model pretrain on ImageNet and absorbance dataset without powder. By fine-tuning, we can leverage the knowledge that the model has already learned from the original absorbance dataset without powder and improve its performance on the new dataset with a powder layer.

Fine-tuning of the modular approach was not conducted. When a powder layer is added, vapor depressions form not only in the substrate but also within the powder layer. Identifying the depression boundary within the powder layer, as shown in Fig. 7b–e, becomes challenging. As the laser power decreases, the ratio of vapor depression within the powder layer becomes more significant. This poses considerable difficulties in generating accurate segmentation masks. Consequently, the modular approach is less applicable in the case of powder. Instead, an end-to-end approach that can directly process the raw image is preferred.

To evaluate how prediction performance is affected by the size of the training data, a training dataset sensitivity test is conducted, as depicted in Fig. 7. The performance of ConvNeXt-T was compared under 8 different conditions: without any fine-tuning and fine-tuning for 30 epochs using 5% (45 frames), 10% (90 frames), 15% (136 frames), 20% (182 frames), 25% (227 frames), 30% (273 frames), and 40% (364 frames) of the absorbance dataset with a powder layer as listed in Table 2. Upon fine-tuning the models using a mere 45 additional data points, the test error was reduced significantly from  $11.39 \pm 1.14$  to  $6.50 \pm 0.83$ . As the size of the training data increases to 364 frames, the test loss further decreases to  $2.31 \pm 0.26$ , which is in close proximity to the test loss of  $2.35 \pm 0.35$  on the absorption data without powder. Notably, a plateau in the loss was observed at around 20% of the training data, indicating that further increases in training data size may not result in significant improvements in model performance. This observation has important implications for future practice, especially when dealing with limited training data. It highlights the potential benefits of optimizing the use of available data and fine-tuning the model on a smaller dataset instead of collecting more data.

Though the modular approach gives better interpretability, the segmentation cost, including obtaining ground-truth masks,



**Fig. 7 Sensitivity analysis of ConvNeXt-T model on absorption data with a powder layer.** **a** Sensitivity analysis showing train, validation, and test errors evaluated by smooth L1 loss of the ConvNeXt models fine-tuned on the absorptance dataset with a powder layer in Table 2 using different amounts of training data. The right side displays selected images from the powder dataset, specifically the 177th frames for **b** stationary laser at 111 W, **c** stationary laser at 197 W, **d** scanning laser at 197 W, and **e** scanning laser at 254 W.

especially when powder is involved, is rather demanding compared to the ease of simply denoising and cropping images in the end-to-end approach. Therefore, for *operando* use of laser absorptance prediction, where accuracy, robustness and automation are valued over model interpretability, it is recommended to prioritize the end-to-end approach. However, when model interpretability is deemed important and affects the decision-making process for critical applications, the modular approach may be a preferable solution.

Furthermore, due to the limited interpretability of neural networks, the challenges in understanding the specific features learned from X-ray images, along with the constrained imaging variations in the training dataset, there is a concern regarding the generalizability of these features when applying the end-to-end model to new beamline data with different sample and image conditions. To effectively address this issue, fine-tuning can be utilized as a solution, as demonstrated in Fig. 7. In addition, using the weights trained in this study through transfer learning from an in-domain dataset, such as X-ray related images, can significantly reduce the costs associated with obtaining more training data and computational cost compared to using weights trained on dissimilar natural images in ImageNet. Therefore, we recommend fine-tuning and employing the end-to-end approach to beamline settings where new absorption experiments can be performed. This creates a positive feedback loop where the learned features become more robust and generalizable as new data is incorporated. In contrast to the implicitly learned features, the geometric features extracted from the segmentation masks exhibit high consistency and robustness regardless of sample and imaging variations. As a result, the modular approach, particularly the RF regression model, can be directly employed for off-line applications.

## METHODS

### Data collection

The training dataset includes all the data used in the development phase of ML models, including the different splits used for training, validation, and testing. We build both the absorption and segmentation datasets using data obtained from the synchrotron high-speed X-ray imaging system at the 32-ID-B beamline of the Advanced Photon Source (APS) at ANL. The imaging setup

integrates a scanning laser system and a sample chamber into a synchrotron X-ray imaging beamline<sup>3,56</sup>. The laser source is a ytterbium fiber operating with a wavelength of 1070 nm and a maximum power of 540 W. Enabled by a galvanometer scanner, the laser source can scan across the sample at a maximum speed of 2 m/s. Samples were enclosed in a stainless steel chamber with 1 atm Ar environment unless specifically noted. During the experiments, while the laser impinges and travels across the specimen, high-energy X-rays penetrate through the thickness, and a high-speed camera captures images at a frame rate of 50 kHz.

To measure laser energy absorptance, a calibrated integrating sphere was incorporated into the aforementioned X-ray imaging system to simultaneously collect the backscattered laser light. The experimental setup can be found in ref. 20. The amount of absorbed laser energy is calculated as the difference between input light and scattered light. Unlike the standard X-ray imaging setup in ref. 3, this setup was designed to capture the strong backscattered light from the initially mirror-like metal surface by placing the laser at a 7° angle relative to the sample surface. Though absorptivity is angularly dependent, the difference in absorption for unpolarized light below 30° in metal is very small, as evidenced in ref. 57. The backscattered light was detected through a photodiode on the sphere surface. The photodiode voltage was subsequently measured by a high-speed oscilloscope, which offered a 1% voltage uncertainty and a time resolution of 40 ns. To address the oversampling of the absorptance data, 62 data points from the integrating sphere were averaged to obtain a single absorptance value for each X-ray image. More detailed information about the experiment, including integrating sphere radiometry combined with high-speed X-ray imaging, can be found in ref. 20.

### Training end-to-end ConvNets models

Neural network models in both approaches were trained using Google Colab and Amazon Web Services (AWS) Deep Learning AMI with an EC2 G5 instance (Certain equipment, instruments, software, or materials, commercial or non-commercial, are identified in this paper in order to specify the experimental procedure adequately. Such identification is not intended to imply recommendation or endorsement of any product or service by NIST, nor is it intended to imply that the materials or equipment

**Table 7.** Laser absorptance prediction training settings.

Training config	ResNet-50 training	ConvNeXt-T training	ConvNeXt-T fine-tuning
optimizer	AdamW	AdamW	AdamW
base learning rate	4e-3	1e-4	1e-4
minimum learning rate	1e-6	1e-7	1e-7
weight decay	0.05	0.05	1e-8
optimizer momentum	$\beta_1, \beta_2 = 0.9, 0.999$	$\beta_1, \beta_2 = 0.9, 0.999$	$\beta_1, \beta_2 = 0.9, 0.999$
batch size	32	32	32
warmup epochs	20	20	0
warmup schedule	liner	linear	-
training epochs	300	300	30

identified are necessarily the best available for the purpose.). Two ConvNets models were trained, namely residual nets (ResNets)<sup>58</sup> and ConvNeXt<sup>59</sup>, on the laser energy absorptance dataset without a powder layer, as listed in Table 1. We selected the 50-layer ResNet (ResNet-50) and ConNeXt tiny (ConvNeXt-T) architectures because they have a similar number of parameters, which are 26 million and 29 million, respectively. To investigate the effect of weights learned from ImageNet, we trained models with randomly initialized weights and pretrained ImageNet weights. For both ResNet-50 and ConvNeXt-T training, we followed a similar procedure as described in the training configuration of ConvNeXt on ImageNet-1K classification<sup>59</sup>, as shown in the first two columns of Table 7. They demonstrated that a set of modern training techniques can significantly enhance the performance of the ResNet-50 model<sup>59–61</sup>. Note that as X-ray images are inherently different from those natural images in the ImageNet dataset, we used the same training configuration for models randomly initialized and models pretrained on ImageNet rather than taking the fine-tuning settings in ref. <sup>59</sup>. To extend the models to scenarios involving a powder layer, we selected the top-performing model trained on data without powder and fine-tuned using data containing a powder layer (see Table 2). The fine-tuning procedure adopted is detailed in the last column of Table 7. For data augmentations, we adopted common schemes, including random rotation by 7° and random horizontal flip. Images were normalized in the same way as ImageNet. The base learning rate was tuned to achieve optimal performance. The smooth mean absolute error loss (smooth L1 loss) function was used  $l(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n l_n$ , where

$$l_n = \begin{cases} 0.5 * (y_n - \hat{y}_n)^2 / \text{beta}, & \text{if } |y_n - \hat{y}_n| < \text{beta} \\ |y_n - \hat{y}_n| - 0.5 * \text{beta}, & \text{otherwise} \end{cases} \quad (1)$$

with ground truth  $y$ , predicted value  $\hat{y}$ , and a  $\text{beta}$  value 1.0. Smooth L1 loss is a modified version of L1 loss that is differentiable at every point, including at zero. It is less sensitive to outliers and exhibits less variability in gradients, which results in a more stable training process<sup>62</sup>.

### Training segmentation models

The decoder architectures that were tested in this study include UNet<sup>34</sup>, UNet++<sup>63</sup>, DeepLabV3<sup>64</sup>, and DeepLabV3+<sup>65</sup>, while the encoder architecture that was tested is ResNet-50<sup>58</sup>. To assess the impact of encoder training on the keyhole segmentation dataset from scratch, models for each architecture were initialized with random weights, and to examine the effect of the encoder's pretrained weights learned from ImageNet, models were also initialized with ImageNet weights<sup>66,67</sup>. To load the weights learned from the ImageNet dataset, grayscale images were converted to

RGB (red, green and blue) images by duplicating the gray channel twice.

The UNet training strategy relies on the use of data augmentation to effectively learn from a limited number of annotated images<sup>34,39</sup>. To achieve this, we implemented a comprehensive training data augmentation recipe<sup>68</sup>, which included random adjustment of contrast, brightness, and perspective; added blur or sharpening; added Gaussian noise and random horizontal flip; random changes to gamma or CLAHE (Contrast Limited Adaptive Histogram Equalization); applied random shift and rotation. These methods were used to improve the performance of the models by providing more diverse and representative training data. Additionally, image scaling with a maximum magnification/shrinkage factor of 0.2 was also used to ensure that the models were robust to changes in magnification and resolution.

During training, the RMSprop optimizer<sup>69</sup> was used with an initial learning rate of 1e-5. The learning rate was designed to decay by 10% every time after 15 epochs without improvement on the validation dataset. Early stop was triggered when the learning rate decayed for the third time. The loss function used was a combination of binary cross entropy (BCE) and dice loss, with a greater emphasis on BCE (70% weighting)<sup>28</sup>. BCE measures the error in segmentation predictions that take into account class imbalance by assigning more weight to the error of smaller area classes. Dice loss balances the error contribution of false negatives and false positives by taking the harmonic mean of precision and recall. This loss function was selected to balance the error contributions of different classes and optimize segmentation performance. Aiming to minimize overhead and make maximum use of GPU memory<sup>34</sup>, we utilized a small batch size of two. Accordingly, we employed a high momentum value of 0.99 to determine the update in the next step based on a large number of previously seen training samples. This helps accelerate gradients vectors in the correct direction and speed up the convergence.

### Geometric feature extraction

For the experiment without powder layer, we only consider the part of the depression that lies beneath the substrate surface. The geometric features extracted include vapor depression area, depth, width at the top surface, width at one-fourth of the depth, width at half of the depth, width at three-fourths of the depth, and mean width averaged over the entire depth, as well as aspect ratio and tangent of the front wall angle<sup>3,32</sup>. The size of each pixel is 1.923  $\mu\text{m}$ . Unlike the aspect ratio divided by the width at the top surface in ref. <sup>32</sup>, the aspect ratio in this study is defined as the depth of the depression divided by the width at the half depth. Due to a tilt angle of 7° applied during the experiment to collect backscattered light, the width at the top surface may differ from the measurement in untilted conditions, while the width at half depth is less affected by the incident angle. The front wall angle

for a laser moving from left to right is the angle between the rightmost pixel at the one-tenth of the depth and the rightmost pixel at the nine-tenths of the depth.

## DATA AVAILABILITY

A fragment of the laser energy absorption datasets is openly available at the NIST Public Data Repository<sup>40</sup> under the link <https://data.nist.gov/od/id/mds2-2525>. Full absorption datasets can be made available upon reasonable request to B.J.S. The X-ray keyhole image segmentation dataset is openly available at <https://rubyljiang18.github.io/keyholeofficial/>.

## CODE AVAILABILITY

The underlying code to train the end-to-end absorptance ConvNets models and the semantic segmentation models for this study is available and can be accessed via this link <https://rubyljiang18.github.io/keyholeofficial/>. In addition, we have made the trained models available on the same website for public access.

Received: 27 April 2023; Accepted: 21 November 2023;

Published online: 05 January 2024

## REFERENCES

- Matsunawa, A., Kim, J.-D., Seto, N., Mizutani, M. & Katayama, S. Dynamics of keyhole and molten pool in laser welding. *J. Laser Appl.* **10**, 247–254 (1998).
- King, W. E. et al. Observation of keyhole-mode laser melting in laser powder-bed fusion additive manufacturing. *J. Mater. Process. Technol.* **214**, 2915–2925 (2014).
- Cunningham, R. et al. Keyhole threshold and morphology in laser melting revealed by ultrahigh-speed x-ray imaging. *Science* **363**, 849–852 (2019).
- Huang, Y. et al. Keyhole fluctuation and pore formation mechanisms during laser powder bed fusion additive manufacturing. *Nat. Commun.* **13**, 1170 (2022).
- Kouraytem, N. et al. Effect of laser-matter interaction on molten pool flow and keyhole dynamics. *Phys. Rev. Appl.* **11**, 064054 (2019).
- Zhao, C. et al. Laser melting modes in metal powder bed fusion additive manufacturing. *Rev. Mod. Phys.* **94**, 045002 (2022).
- Ren, Z. et al. Machine learning-aided real-time detection of keyhole pore generation in laser powder bed fusion. *Science* **379**, 89–94 (2023).
- Khairallah, S. A., Sun, T. & Simonds, B. J. Onset of periodic oscillations as a precursor of a transition to pore-generating turbulence in laser melting. *Addit. Manuf. Lett.* **1**, 100002 (2021).
- Zhao, C. et al. Critical instability at moving keyhole tip generates porosity in laser melting. *Science* **370**, 1080–1086 (2020).
- Gan, Z. et al. Universal scaling laws of keyhole stability and porosity in 3D printing of metals. *Nat. Commun.* **12**, 2379 (2021).
- He, Y., Montgomery, C., Beuth, J. & Weibler, B. Melt pool geometry and microstructure of ti6al4v with b additions processed by selective laser melting additive manufacturing. *Mater. Des.* **183**, 108126 (2019).
- Cook, P. S. & Ritchie, D. J. Determining the laser absorptivity of ti-6al-4v during laser powder bed fusion by calibrated melt pool simulation. *Opt. Laser Technol.* **162**, 109247 (2023).
- Ye, J. et al. Energy coupling mechanisms and scaling behavior associated with laser powder bed fusion additive manufacturing. *Adv. Eng. Mater.* **21**, 1900185 (2019).
- Kawahito, Y., Matsumoto, N., Abe, Y. & Katayama, S. Relationship of laser absorption to keyhole behavior in high power fiber laser welding of stainless steel and aluminum alloy. *J. Mater. Process. Technol.* **211**, 1563–1568 (2011).
- Trapp, J., Rubenichik, A. M., Guss, G. & Matthews, M. J. In situ absorptivity measurements of metallic powders during laser powder-bed fusion additive manufacturing. *Appl. Mater. Today* **9**, 341–349 (2017).
- Tolochko, N. K. et al. Absorptance of powder materials suitable for laser sintering. *Rapid Prototyp. J.* **6**, 155–161 (2000).
- Simonds, B. J. et al. Time-resolved absorptance and melt pool dynamics during intense laser irradiation of a metal. *Phys. Rev. Appl.* **10**, 044061 (2018).
- Allen, T. R. et al. Energy-coupling mechanisms revealed through simultaneous keyhole depth and absorptance measurements during laser-metal processing. *Phys. Rev. Appl.* **13**, 064070 (2020).
- Wang, L., Zhang, Y., Chia, H. Y. & Yan, W. Mechanism of keyhole pore formation in metal additive manufacturing. *NPJ Comput. Mater.* **8**, 22 (2022).
- Simonds, B. J. et al. The causal relationship between melt pool geometry and energy absorption measured in real time during laser-based manufacturing. *Appl. Mater. Today* **23**, 101049 (2021).
- Miyagi, M. & Wang, J. Keyhole dynamics and morphology visualized by in-situ x-ray imaging in laser melting of austenitic stainless steel. *J. Mater. Process. Technol.* **282**, 116673 (2020).
- Schricker, K. et al. Characterization of keyhole dynamics in laser welding of copper by means of high-speed synchrotron x-ray imaging. *Procedia CIRP* **111**, 501–506 (2022).
- Martin, A. A. et al. Ultrafast dynamics of laser-metal interactions in additive manufacturing alloys captured by in situ x-ray imaging. *Mater. Today Adv.* **1**, 100002 (2019).
- Dilberoglu, U. M., Gharehpapagh, B., Yaman, U. & Dolen, M. The role of additive manufacturing in the era of industry 4.0. *Procedia Manuf.* **11**, 545–554 (2017).
- Goodfellow, I. J., Bengio, Y. & Courville, A. *Deep Learning* (MIT Press, 2016).
- Rudin, C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat. Mach. Intell.* **1**, 206–215 (2019).
- ASTM. *Standard Test Methods for Determining Average Grain Size E112-10*. ASTM E112 (ASTM, 2010).
- Stuckner, J., Harder, B. & Smith, T. M. Microstructure segmentation with deep learning encoders pre-trained on a large microscopy dataset. *NPJ Comput. Mater.* **8**, 200 (2022).
- Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675 (2012).
- Smith, T. et al. Characterization of nanoscale precipitates in superalloy 718 using high resolution SEM imaging. *Mater. Charact.* **148**, 178–187 (2019).
- Simonds, B. J. et al. Simultaneous high-speed x-ray transmission imaging and absolute dynamic absorptance measurements during high-power laser-metal processing. *Procedia CIRP* **94**, 775–779 (2020).
- Pyeon, J. et al. Time-resolved geometric feature tracking elucidates laser-induced keyhole dynamics. *Integr. Mater. Manuf. Innov.* **10**, 677–688 (2021).
- Bradski, G. The OpenCV Library. *Dr. Dobb's J. Softw. Tools* **120**, 122–125 (2000).
- Ronneberger, O., Fischer, P. & Brox, T. U-Net: convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI*, 234–241 (Springer International Publishing, 2015).
- DeCost, B. L., Lei, B., Francis, T. & Holm, E. A. High throughput quantitative metallography for complex microstructures using deep learning: a case study in ultrahigh carbon steel. *Microsc. Microanal.* **25**, 21–29 (2019).
- Cohn, R. et al. Instance segmentation for direct measurements of satellites in metal powders and automated microstructural characterization from image data. *JOM* **73**, 2159–2172 (2021).
- Zhang, J. et al. Image segmentation for defect analysis in laser powder bed fusion: deep data mining of X-ray photography from recent literature. *Integr. Mater. Manuf. Innov.* **11**, 418–432 (2022).
- García-García, A., Orts-Escobedo, S., Oprea, S., Villena-Martínez, V. & García-Rodríguez, J. A review on deep learning techniques applied to semantic segmentation. Preprint at <https://arxiv.org/abs/1704.06857> (2017).
- Minaee, S. et al. Image segmentation using deep learning: a survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **44**, 3523–3542 (2022).
- Simonds, B. J. et al. Asynchronous AM Bench 2022 challenge data: real-time, simultaneous absorptance and high-speed x-ray imaging <https://data.nist.gov/od/id/mds2-2525> (2022).
- Bitharas, I. et al. The interplay between vapour, liquid, and solid phases in laser powder bed fusion. *Nat. Commun.* **13**, 2959 (2022).
- Fabbro, R. Melt pool and keyhole behaviour analysis for deep penetration laser welding. *J. Phys. D Appl. Phys.* **43**, 445501 (2010).
- Chen, T. & Guestrin, C. Xgboost: a scalable tree boosting system. In *Proc. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16*, 785–794 (Association for Computing Machinery, 2016).
- James, G., Witten, D., Hastie, T. & Tibshirani, R. *An Introduction to Statistical Learning: With Applications in R* (Springer, 2013).
- Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32 (2001).
- Deng, J. et al. Imagenet: a large-scale hierarchical image database. In *Proc. 2009 IEEE Conference on Computer Vision and Pattern Recognition*, 248–255 (2009).
- Huh, M., Agrawal, P. & Efron, A. A. What makes ImageNet good for transfer learning? Preprint at <https://arxiv.org/abs/1608.08614> (2016).
- Minaee, S., Kafieh, R., Sonka, M., Yazdani, S. & Jamalipour Soufi, G. Deep-COVID: predicting COVID-19 from chest X-ray images using deep transfer learning. *Med. Image Anal.* **65**, 101794 (2020).
- Chouhan, V. et al. A novel transfer learning based approach for pneumonia detection in chest x-ray images. *Appl. Sci.* **10**, 559 (2020).
- Huyen, C. *Designing Machine Learning Systems* (O'Reilly Media, 2022).
- Radford, A., Metz, L. & Chintala, S. Unsupervised representation learning with deep convolutional generative adversarial networks. Preprint at <https://arxiv.org/abs/1511.06434> (2015).
- Selvaraju, R. R. et al. Grad-CAM: visual explanations from deep networks via gradient-based localization. *Int. J. Comput. Vis.* **128**, 336–359 (2019).

53. Parsa, A. B., Movahedi, A., Taghipour, H., Derrible, S. & Mohammadian, A. K. Toward safer highways, application of XGBoost and SHAP for real-time accident detection and feature analysis. *Accid. Anal. Prev.* **136**, 105405 (2020).
54. Lundberg, S. M. & Lee, S.-I. A unified approach to interpreting model predictions. *Adv. Neural Inf. Process. Syst.* **30**, 4765–4774 (2017).
55. Kingma, D. P. & Dhariwal, P. Glow: generative flow with invertible 1x1 convolutions. In *Pred. Advances in Neural Information Processing Systems*, 10215–10224 (2018).
56. Kouraytem, N. et al. Solidification crack propagation and morphology dependence on processing parameters in aa6061 from ultra-high-speed x-ray visualization. *Addit. Manuf.* **42**, 101959 (2021).
57. Hipp, D., Mahrle, A. & Beyer, E. Beyond Fresnel: absorption of fibre laser radiation on rough stainless steel surfaces. *J. Phys. D Appl. Phys.* **52**, 355302 (2019).
58. He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2016).
59. Liu, Z. et al. A ConvNet for the 2020s. In *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 11976–11986 (2022).
60. Bello, I. et al. Revisiting ResNets: improved training and scaling strategies. In *Proc. Advances in Neural Information Processing Systems (NeurIPS)*, 34 (2021).
61. Wightman, R., Touvron, H. & Jégou, H. ResNet strikes back: an improved training procedure in timm. Preprint at <https://arxiv.org/abs/2110.00476> (2021).
62. Girshick, R. Fast R-CNN. In *Proc. IEEE International Conference on Computer Vision (ICCV)*, 1440–1448 (2015).
63. Zhou, Z., Siddiquee, M. M. R., Tajbakhsh, N. & Liang, J. UNet++: a nested U-Net architecture for medical image segmentation. In *Proc. Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, 3–11 (2018).
64. Chen, L.-C., Papandreou, G., Schroff, F. & Adam, H. Rethinking atrous convolution for semantic image segmentation. Preprint at <https://arxiv.org/abs/1706.05587> (2017).
65. Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F. & Adam, H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In *Proc. European Conference on Computer Vision (ECCV)*, 801–818 (2018).
66. He, K., Zhang, X., Ren, S. & Sun, J. Delving deep into rectifiers: surpassing human-level performance on ImageNet classification. In *Proc. IEEE International Conference on Computer Vision (ICCV)*, 1026–1034 (2015).
67. Iakubovskii, P. Segmentation models pytorch. GitHub [https://github.com/qubvel/segmentation\\_models.pytorch](https://github.com/qubvel/segmentation_models.pytorch) (2019).
68. Buslaev, A. et al. Albumentations: fast and flexible image augmentations. *Information* **11**, 125 (2020).
69. Ruder, S. An overview of gradient descent optimization algorithms. Preprint at <https://arxiv.org/abs/1609.04747> (2017).

## ACKNOWLEDGEMENTS

We gratefully acknowledge funding for this work through the support of the Department of Defense, Office of Economic Adjustment, under grant number ST1605-17-02, Michael Gilroy, CTO & Program Director. R.J. acknowledges the contributions of Joseph Aroh for the initial idea inception and Jiahui Zhang for extensively utilizing the absorbance models and providing constructive feedback. A.D.R. acknowledges support from the NASA ULI program under grant number

80NSSC19M0123. B.J.S. acknowledges partial support of his work by NIST. The acquisition of the X-ray videography used resources from the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357.

## AUTHOR CONTRIBUTIONS

R.J. conceived and designed the study, performed experiments and collected data, developed the software, evaluated results, implemented formal analysis, and writing of the original draft. J.S. and Y.Y. performed experiments and labeled the segmentation dataset. T.S. advised on experiments, data collection and analysis. B.J.S. conducted experiments and analysis and provided the absorption dataset. A.D.R. provided guidance and advice for the project. All authors reviewed the final manuscript.

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41524-023-01172-8>.

**Correspondence** and requests for materials should be addressed to Runbo Jiang.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024