DEVCOM
ARMY RESEARCH
LABORATORY

Physics-Guided Neural Network for Regularization and Learning Unbalanced Data Sets: A Priori Prediction of Melt Pool Width Variation in Directed Energy Deposition

by Brett Diehl, Clara Mock, Lester Hitch, Christopher Rinderspacher, and Brandon McWilliams

NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.

Physics-Guided Neural Network for Regularization and Learning Unbalanced Data Sets: A Priori Prediction of Melt Pool Width Variation in Directed Energy Deposition

Clara Mock, Christopher Rinderspacher, and Brandon McWilliams
DEVCOM Army Research Laboratory

Brett Diehl
Oak Ridge Associated Universities

Lester Hitch
SURVICE Engineering

REPORT DOCUMENTATION PAGE

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED	
		START DATE	END DATE
March 2023	Technical Report	2 November 2020	2 November 2022

4. TITLE AND SUBTITLE

Physics-Guided Neural Network for Regularization and Learning Unbalanced Data Sets: A Priori Prediction of Melt Pool Width Variation in Directed Energy Deposition

5a. CONTRACT NUMBER W15P7T-19-D-0126	5b. GRANT NUMBER	5c. PROGRAM ELEMENT NUMBER
5d. PROJECT NUMBER	**5e. TASK NUMBER**	**5f. WORK UNIT NUMBER**

6. AUTHOR(S)

Brett Diehl, Clara Mock, Lester Hitch, Christopher Rinderspacher, and Brandon McWilliams

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER ARL-TR-9655
DEVCOM Army Research Laboratory ATTN: FCDD-RLA-MD Aberdeen Proving Ground, MD 21005	

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release: distribution unlimited.

13. SUPPLEMENTARY NOTES

ORCID IDs: Clara Mock, 0000-0001-8055-2371; Brandon McWilliams, 0000-0002-0494-3140; Berend Christopher Rinderspacher, 0000-0003-1333-1653

14. ABSTRACT

Directed energy deposition is of interest to the aerospace and defense industries for the production of novel and complex geometries, as well as repair applications. However, variability during the build process can result in deviations in final component geometry, structure, and mechanical properties, which adds to the complexity of process planning and slows down adoption of this technology. To address geometric tolerances, neural networks were trained to predict melt pool width, given input drivers such as build height, laser power, laser speed, and thin wall length. Physical constraints on the relationships between the 1st and 2nd derivatives of input drivers and melt pool width were enforced using custom loss functions, yielding physics-guided neural networks (PGNNs). PGNNs predicted the melt pool width with a higher performance (R-square = 0.991) than traditional neural networks (R-square = 0.884). Physics-based loss functions performed superior to traditional methods of regularization and were a superior method of training on unbalanced data sets versus sample/class weighting. This work demonstrates the benefits of enforcing physical constraints on machine learning predictions of additive manufacturing processes using finite estimates of mathematical expressions of physical laws.

15. SUBJECT TERMS

physics-informed machine learning, additive manufacturing, in-situ sensor prediction, laser-directed energy deposition, build planning, melt pool, Sciences of Extreme Materials

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES
a. REPORT UNCLASSIFIED	**b. ABSTRACT** UNCLASSIFIED	**C. THIS PAGE** UNCLASSIFIED	UU	36

19a. NAME OF RESPONSIBLE PERSON Clara Mock	19b. PHONE NUMBER (Include area code) (410) 306-0821

STANDARD FORM 298 (REV. 5/2020)
Prescribed by ANSI Std. Z39.18

Contents

List of Figures

List of Tables

Acknowledgments

1. Introduction

Directed energy deposition (DED) is a method of metal additive manufacturing (AM) by which parts are built layer by layer from 3-D computer-aided design models. In laser DED, powder is blown in front of a moving laser, which melts the powder, before it solidifies into part of the component being manufactured (Fig. 1). AM provides a more direct pathway between the digital model and the manufactured part, which allows for more rapid design iterations and testing of parts. This further enables more advanced design methods and integrated features that would be operationally difficult by way of traditional subtractive manufacturing. In addition to being used to produce entire components, DED offers advantages through the ability to deposit additional material to existing components for the use of repair or cladding. Due to these advantages, DED is being adopted for specialized applications in defense and aerospace (DebRoy et al. 2019).
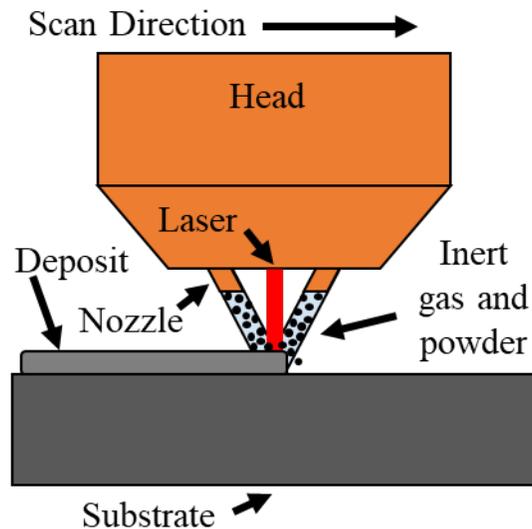
Fig. 1   Schematic of DED process

While there is great enthusiasm to use AM, process deviations are common, which can result in undesired final geometry, poor surface finish, or insufficient mechanical performance (Colosimo et al. 2018). Some of these drifts in the structure and property of metals fabricated by DED are not caused by stochastic process deviations, but by predictable trends. For example, Feenstra et al. (2020) observed changes in grain structure of 316L stainless steel with build height. Fine grains were observed in the AM structure near the build plate due to high cooling rates during building on a cold substrate. As the structure continued to build, heat buildup in the part created larger, columnar grains. At some point the grain structure was very predictable as the part reached a thermally, steady-state condition. Thus,

1

the deviation in microstructure and mechanical properties can be attributed to the local thermal history of the part. Observation of the melt pool during the DED process provides a means to measure an instantaneous point in the thermal history of a part. These observations can then be mapped in 3-D to the built geometry such as what was done in Kriczky et al. (2015). As a result, there is great interest in using sensors for in-situ certification of components, as well as responding to process deviations in real time, enhancing the performance of AM components, and increasing the yield of the AM process (Everton et al. 2016).

Researchers have sought to predict the value of sensor data using machine learning, given process parameters and build plan information. The sensor data is reflective of all the nuance and interdependences in the inputs. Using machine learning to predict sensor data may elucidate relationships that may not be discoverable with simpler statistics (DebRoy et al. 2018). For example, Diehl and Nassar (2020) demonstrate that the time between hatches is an important factor in predicting near-surface void occurrence. While they explain that the time between hatches is a function of toolpath and geometry, they cannot fully describe the relationship between these input drivers and void formation, due to the simplicity of their model. In contrast, Gobert et al. (2019) trained a machine learning algorithm to produce synthetic thermal images, given processing data, toolpath information, and geometric data as input drivers. While they do not make any conclusions on the relationship between sensor data and voids, their work clearly demonstrates the relationship between toolpath/geometry and thermal buildup, yielding a better understanding of the process. Supervised machine learning is advantageous to manage the large amount of data in these problems, and the complexity of the input drivers' interactions (Tapia and Elwany 2014). Supervised machine learning is a collection of algorithms that generalize a known relationship from examples, such as mapping inputs to outputs, as opposed to unsupervised machine learning, which seeks to find unknown patterns from examples (Bishop 2006). A limitation of neural nets is that they are general purpose approximators that do not incorporate prior knowledge of, for example, AM physics. It has been demonstrated that physical limitations can be enforced on neural networks, by detecting physical violations in the model's predictions, and adding a penalty term to their loss function (Daw et al. 2021). Daw et al. (2021) demonstrated that using these methods, neural nets can achieve results superior to state-of-the-art physics-based models and, unlike typical neural nets, do not predict physical inconsistencies. Physics-guided neural networks (PGNNs) have been shown to outperform traditional neural nets while predicting no physical inconsistencies, but Daw et al. only applied 1st-order differential limitations to their models.

In this study, DED melt pool width was predicted as a function of input drivers. Physical trends in the DED melt pool width were identified in literature, predicted via modeling, and experimentally verified. These physical trends are then enforced on neural nets as constraints. Huang et al. (2019) developed models that predicted a melt pool width increase with each subsequent layer, and that as the wait time between layers increased, melt pool width decreased. Akbari and Kovacevic (2019) confirm experimentally that melt pool width only increases with increasing layer height. Liu et al. (2019) experimentally show that melt pool temperature increases with increasing laser power and—below the speed at which balling occurs—decreases with increasing laser speed. It could be argued that a hotter melt pool would be wider than a cooler one.

The previously described physical trends can be enforced via simple finite difference estimates of derivatives; for example, a model can be penalized when it predicts a decrease in melt pool width in response to an arbitrarily small increase in laser power—everything else held constant. While this work predicts DED melt pool width as a function of input drivers, the aim of this study is not to develop a framework to predict sensor readings during DED, but to dynamically predict melt pool width, such that power and speed can be adjusted, to improve the ability of DED to meet geometric tolerances. The ultimate goal is to develop a generalized framework for adding physics-based loss functions to advanced manufacturing models and demonstrate the benefits thereof.

In this work, neural networks were trained, given process parameters and build plan information, to predict melt pool width during a DED build, given input drivers of speed, power, length, and height. Length and height are geometry dependent and may vary over the course of a build, and cause drift in melt pool width. In contrast, power and speed are process-specific and can be changed dynamically to adjust the melt pool width in response to drift caused by changes in local geometry.

Traditional neural nets were trained alongside their physics-guided counterparts on an unbalanced data set. Advantages of PGNN models over traditional neural nets were demonstrated, in terms of model performance, physical consistency, and training on unbalanced data. PGNNs were compared against neural nets with weighted loss functions, a traditional method of training on unbalanced data. As well, it was hypothesized that PGNNs act as physics-regularized neural networks, and their performance was compared against neural networks with traditional kernel, bias, and activation regularizers. Afterward, the best-performing PGNN was validated by inverting it to predict processing parameters that maintain constant melt pool width over build height.

2.   Methods

2.1 Build Procedure

A low alloy steel powder with composition provided in Table 1 was procured from Carpenter Technology Corp (Reading, Pennsylvannia). Particle size analysis was conducted using a Retsch Camsizer X2 (Newtown, Pennsylvannia). The particle size distribution is provided in Fig. 2a and shows that the majority of the powder particles have a diameter between 40–130 µm with a median particle diameter (D50) of 71.5 µm. A Hitachi S4700 (Schaumburg, Illinois) scanning electron microscope was used to further examine the particle size and shape of the powder (see Fig. 2b). The particles were all fairly spherical with some satellites present on the surface of the steel powder.

Table 1   Elemental composition of steel powder, all units are weight percent

Al	C	Cr	Cu	H	N	Mn	Mo	Ni	O	P	S	Si	Ti	V	Fe
0.004	0.278	2.41	0.004	<0.0005	0.003	0.65	0.84	1.04	0.008	0.007	0.002	1.04	0.003	0.11	93.6

a)

b)
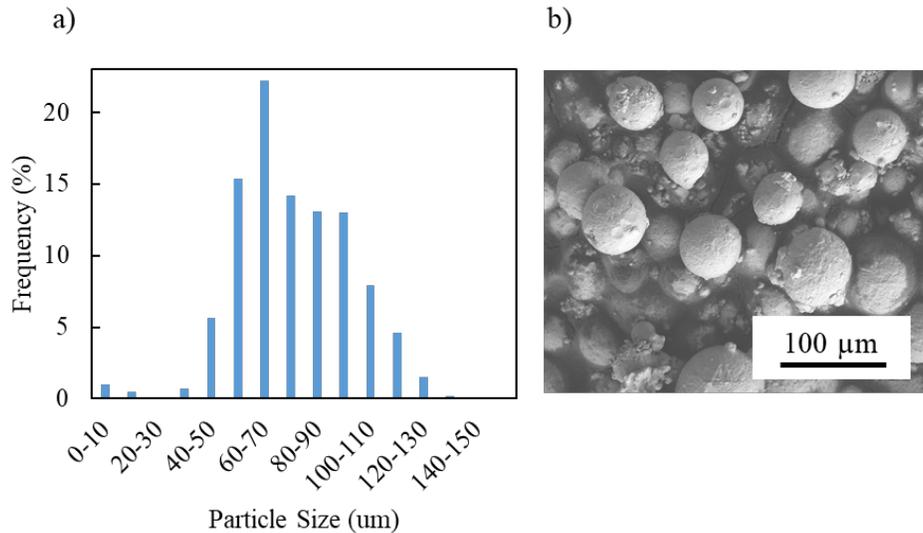
Fig. 2   Steel powder a) particle size distribution and b) imaged via scanning electron microscopy

Fifteen steel thin walls were fabricated in an RPM 222 powder-based laser DED system (RPM Innovations, Inc, Rapid City, SD), with a 2-kW ytterbium laser of 1070 nm wavelength. Each component was fabricated with a layer height of 102 µm, and a powder feed rate of 15 g/min. This work used the "open contour," in which each layer is processed a single time, with each layer consisting of a single bead of material and deposited in a bidirectional (serpentine) raster pattern.

Two sets of thin walls were produced: one in which the power and speed were varied (parameter set) and a set in which the thin wall lengths were varied (length set). The geometry and processing parameters of all of the thin walls are shown in Table 2. Samples for the parameter set were named *X-Y* where *X* is the power in watts (W) and *Y* is the travel speed (mm/s) for the deposition head. The nomenclature for the length set is as follows: *Z*L, where *Z* is the approximate length of the wall in millimeters.

Table 2   Process parameters and lengths associated with each thin wall

Sample name	Sample set	Power (W)	Speed (mm/s)	Wall length (mm)
12L	Length	600	16.9	12.7
25L	Length	600	16.9	25.4
50L	Length	600	16.9	50.8
100L	Length	600	16.9	101.6
150L	Length	600	16.9	152.4
25LB	Length	600	16.9	25.4
25L-12H	Length	600	16.9	25.4
650-16	Parameter	650	16.9	25.4
550-16	Parameter	550	16.9	25.4
450-16	Parameter	450	16.9	25.4
350-16	Parameter	350	16.9	25.4
650-12	Parameter	650	12.7	25.4
550-12	Parameter	550	12.7	25.4
450-12	Parameter	450	12.7	25.4
350-12	Parameter	350	12.7	25.4

The sample 25LB was fabricated to a height of 25.4 mm, with an additional layer deposited after a 30-min wait. The modification to 25LB was to study an effect separate from this study. Each of the thin walls was built to a height of 25.4 mm, with the exception of 25L-12H, which was only built to 12.7 mm. The modifications to 25L-12H were done to investigate an effect not addressed in this study; however, this additional data was useful to demonstrate the repeatability of the results. The data from 25LB and 25H-12H were held out from training due to being extremely similar to the data from 25L. The neural nets were trained on 10 of the data sets shown in Table 2 (the others were held out for validation and testing). There were 2510 points in total—one for each of the 251 layers of the 10 thin walls used for training.

The best-performing neural net was then used to predict the required combination of laser power and contour speeds to produce thin walls 2 ±0.02 mm thick and 25.4 mm tall. These multiple parameter set (MPS) walls were 25.4 mm long (25L MPS) and 63.5 mm long (63L MPS), respectively. The tolerance in the wall

thickness reduced the number of unique laser power and contour speed combinations to allow easier tool path programming.

2.2 Sensing Setup and Analysis

A coaxial Xiris XVC-1000 grayscale welding camera (Burlington, Ontario, Canada) was used to record the fabrication process, with a framerate of 55 frames/second and an approximate resolution of 24 microns/pixel. The XVC-1000 was equipped with a high dynamic range (140+ dB) complementary metal oxide semiconductor sensor and UV + IR cut filter. A representative image of the melt pool during the DED process is provided in Fig. 3. A Gaussian mixture model with two mixtures was used to determine the grayscale threshold between the melt pool and the surrounding region and the grayscale threshold for the laser spot region. Both thresholds were manually verified. That is, all pixels with intensities below the laser spot size threshold were removed from randomly sampled images and it was confirmed they only contained the laser spot. Likewise, all pixels with intensities below the melt pool threshold were removed from randomly sampled images and it was confirmed they only contained the melt pool and the laser spot. The presence of the laser spot size was used to manually register the data.
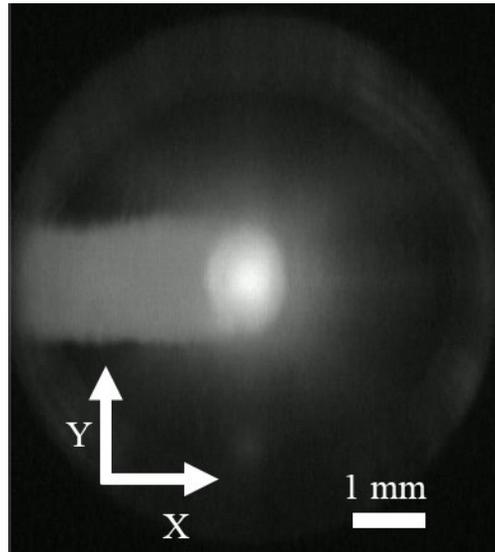
Fig. 3   Characteristic image of melt pool during processing; melt pool (gray) and laser spot region (gray-white) shown. Image shown is of the native resolution of the weld camera.

Each image was transformed from an 8-bit grayscale image of the melt pool to a Boolean image, where 1 corresponds to regions that are part of the melt pool, and 0 corresponds to those that are not. For every column of the melt pool, the width of the melt pool was recorded. Of these recorded values, the 90th percentile was recorded as the width associated with the frame. This was to remove outliers, such

as regions of the melt pool that were ejecting material at the instant the image was recorded, an example of which is shown in Fig. 4.
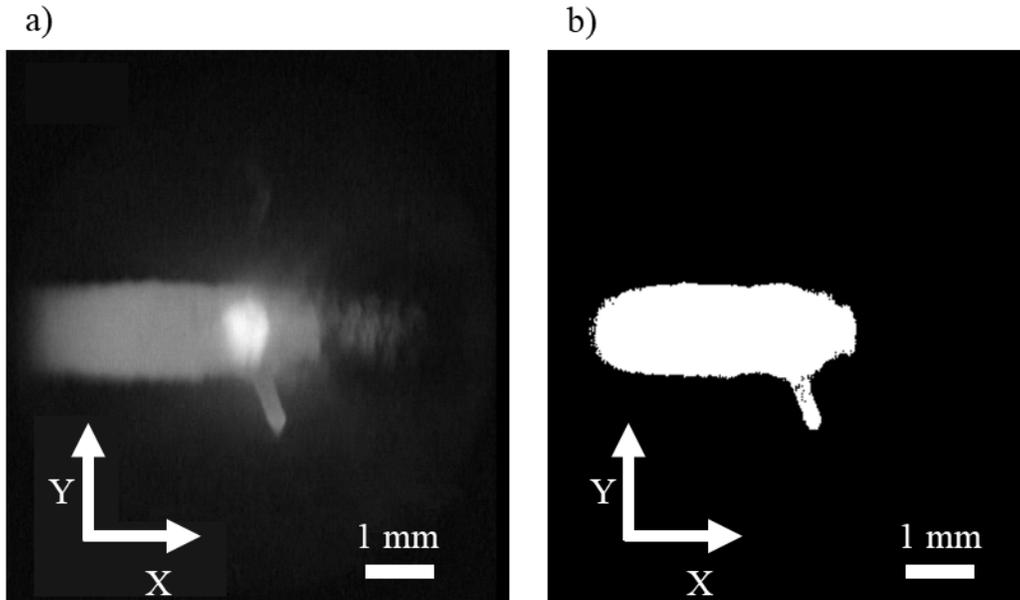
Fig. 4 **(a) Ejecta from the lower side of the melt pool makes the melt pool appear artificially wider, when a simple grayscale threshold is used. (b) Boolean transform of the melt pool image.**

The width measurement process is shown in Fig. 5. Figure 4a shows the width of the melt pool at each image column. Note the peak associated with the ejecta. Figure 5b shows the positive values in Fig. 5a sorted, with the 90th percentile highlighted by a dotted line. The measured width of this melt pool is 1.54 mm. Visually, this value is closer to the width of the melt pool demonstrating the robustness of the 90th percentile method to remove spatter-induced outliers.
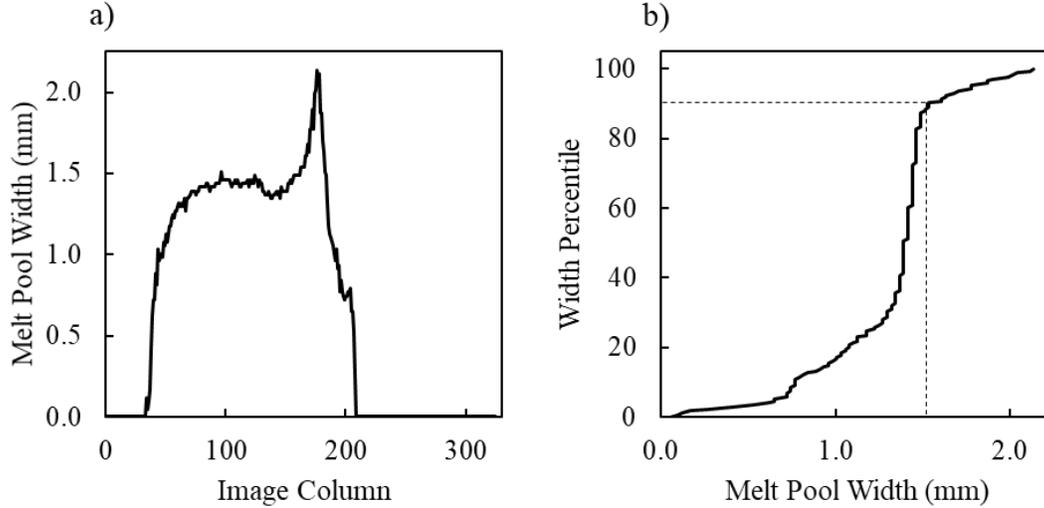
Fig. 5   **(a) Width of melt pool on a column-wise basis, for the melt pool image shown in Fig. 4b, every positive number in (a), sorted**

The presence or lack of the laser spot could be determined using the gray value threshold. A script was used to determine the position of the camera during each frame as follows: frames in which the laser spot was not present correspond to when the laser was off, before or after layers. Between these "laser-off" frames, frames were equally spaced between the start and stop points on the thin wall. The first set of consecutive frames in which the laser was on correspond to the first layer; the second set of consecutive frames in which the laser was on correspond to the second layer, and so forth. After frames were associated to layers, the median width of all frames associated with a layer was considered the characteristic width associated with that layer. While the serpentine build path did affect melt pool width near turnaround points, that is, the melt pool was qualitatively confirmed to bulge at the beginning of each layer, this affects a relatively small portion of each layer. Because this bulging occurs in far less than half of the layer, it does not affect the median width of any layers and therefore does not affect our analysis.

2.3 Machine Learning

Single-layer fully connected neural networks were used to regress from input drivers: power (P), speed (S), wall length (L), and build height (H) to an output of melt pool width (W). The structure of this network is shown in Fig. 6. In this work, the structure of the network is held constant. Various physics-based and non-physics-based loss functions are used to train the network.

8
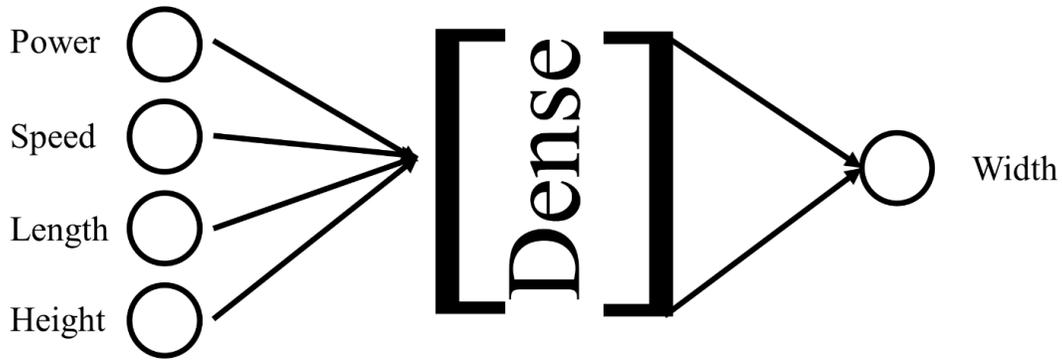
Fig. 6 **Structure of the neural network that predicted melt width from input drivers**

All neural networks were implemented using the Keras package (Chollet 2015) and Tensorflow backend. The activation function for the hidden layer was the hyperbolic tangent function (tanh), and activation for the output layer was linear. The learning rate was set to 0.0001, batch size was set to 50, and the Adam algorithm (Diederik and Kingma 2014) was used for training. No maximum number of epochs was chosen; instead, testing and validation data were held out. The network was trained until the network performance on the validation data did not improve for 100 epochs. Unless otherwise specified, the loss function used was the mean squared error (MSE). All other training values were left to Keras default. To determine the ideal number of hidden neurons in the hidden layer, neural nets with hidden layer sizes 10, 25, 50, 100, and 200 were trained on the training data, and the performance of the validation data was used to estimate optimal network architecture.

There were 251 data points along the build direction (251 layers/sample) and a significantly smaller number of powers (5), speeds (2), and wall lengths (5). To prevent the neural networks from learning to interpolate along the H axis (along the build height), each thin wall's data was either included in only the train, validation, or testing data set. As each thin wall's data must be only found in the training, validation, or testing data, and never more than one of them, the entirety of the data from sample 450-12 was held out for testing and the data from samples 50L and 450-16 were held out for validation. The division of data used for training, validation, and testing is shown in Table 3.

Table 3   Identification of experiments used for training, validation, and testing

Training	Validation	Testing
12L	50L	450-12
25L	450-16	. . .
100L	. . .	. . .
150L	. . .	. . .
650-16	. . .	. . .
550-16	. . .	. . .
350-16	. . .	. . .
650-12	. . .	. . .
550-12	. . .	. . .
350-12	. . .	. . .

Because there are eight experiments in the parameter set and five in the length set, two sets of data were held out from the parameter set for validation and testing, while only one set of data was held out for validation. Still, there are twice as many variable parameter set experiments as there are variable length set experiments used for training. As a result of this imbalance, the parameter set will comprise a majority of the training data's contribution to the loss function. The training data was batch-normalized, and the same transformation was applied to the validation and test data sets; that is, input drivers had the following transformation applied before being inputted into the network:

$$P_{norm} = \frac{P - \mu_{P_{train}}}{\sigma_{P_{train}}}$$

$$S_{norm} = \frac{S - \mu_{S_{train}}}{\sigma_{S_{train}}}$$

$$L_{norm} = \frac{L - \mu_{L_{train}}}{\sigma_{L_{train}}}$$

$$H_{norm} = \frac{H - \mu_{H_{train}}}{\sigma_{H_{train}}}$$

where $\mu_{P_{train}}$, $\mu_{S_{train}}$, $\mu_{L_{train}}$, and $\mu_{H_{train}}$ are the corresponding averages for the *Power*, *Speed, Length*, and *Height* of the training data, respectively, and $\sigma_{P_{train}}$, $\sigma_{S_{train}}$, $\sigma_{L_{train}}$, $\sigma_{H_{train}}$ are their respective standard deviations.

2.4 PGNNs

A similar approach to Daw et al. (2021) was followed, adding an additional component to the loss function that quantifies physical inconsistencies in a model's prediction by defining various metrics, which should be nonpositive on a physical basis. Daw et al. defined 1st-order derivatives of their input drivers and regressor

target, which must be positive or negative due to the laws of physics. Here, this work was expanded: 1st-order differential constraints were defined and enforced on a neural network. As well, 2nd-order constraints were developed and were estimated using a finite difference method.

The physical contribution to the loss function is denoted λ_{PHY}*Loss.PHY($\hat{Y}$), where λ_{PHY} is the hyper-parameter that determines the relative importance of minimizing physical inconsistences, versus minimizing prediction error. For this work, λ_{PHY} was set equal to 1, to demonstrate that no tuning was performed to any particular data set. The physical component of the loss function does not rely on accurate predicted values but requires that predicted values obey physical limitations. Therefore, neural network predictions were sampled to evaluate the physical contribution to the loss function. Specifically, 160,000 samples of input drivers were used, spanning every combination of: 1) 20 equally spaced values of power from 350 to 650 W, 2) 20 equally spaced values of speed from 12.7 to 16.9 mm/s, 3) 20 equally spaced values of wall length from 0.5 to 6, and 4) 20 equally spaced values of build height from 0 to 25.4 mm. These input driver sets do not correspond to any real manufactured component and are only added to enforce the physics-based constraints.

It can be predicted and experimentally verified that within the regime of typical processing parameters, some changes in melt pool width should not be expected with particular changes in input drivers. For example, increasing the laser power should not result in a decrease in melt pool width, all other input drivers held constant. That is, for any two values of power, P_i and P_{i+1}, where $P_i < P_{i+1}$, the following value should be nonpositive:

$$\boldsymbol{\Delta_P}[\text{i, S, L, H}] = \widehat{W}[P_i, \text{S, L, H}] - \widehat{W}[P_{i+1}, \text{S, L, H}] \leq 0$$

Similar relationships can be presented for the other three input drivers, within the regime of typical processing parameters and with all other input drivers held constant. When laser speed is increased, melt pool width should not increase. For any two values of speed, S_i and S_{i+1}, where $S_i < S_{i+1}$, the following value should be nonpositive: (note that the subscripts are reversed):

$$\boldsymbol{\Delta_S}[\text{P, j, L, H}] = \widehat{W}[\text{P}, S_{j+1}, \text{L, H}] - \widehat{W}[\text{P}, S_j, \text{L, H}] \leq 0$$

As well, it could be hypothesized that some melt pool widening is due to thermal buildup. After heat is applied in Layer i, it has a finite amount of time to dissipate before Layer i+1 is processed. As a result, if that amount of time is higher, when the wall length is increased, the melt pool width should not increase:

$$\boldsymbol{\Delta_L}[\text{P, S, k, H}] = \widehat{W}[\text{P, S}, L_{k+1}, \text{H}] - \widehat{W}[\text{P, S}, L_k, \text{H}] \leq 0$$

Similarly, when build height is increased, the melt pool width should not decrease.

$$\boldsymbol{\Delta_H}[\text{P, S , L, l}] = \widehat{W}[\text{P, S, L, } H_l] - \widehat{W}[\text{P, S, L, } H_{l+1}] \leq 0$$

Every possible consecutive pair of width predictions in the unlabeled data set can be evaluated. For each of them, all positive occurrences are added up. This can be expressed as the loss contribution:

$$\begin{aligned}
\text{Loss. PHY}_1(\hat{Y}) = \ &\sum_{l=1}^{n_H} \sum_{k=1}^{n_L} \sum_{j=1}^{n_S} \sum_{i=1}^{n_P-1} \boldsymbol{ReLU}(\boldsymbol{\Delta_P}[\text{I, } S_j \text{ , } L_k, H_l]) + \\
&\sum_{l=1}^{n_H} \sum_{k=1}^{n_L} \sum_{j=1}^{n_S-1} \sum_{i=1}^{n_P} \boldsymbol{ReLU}(\boldsymbol{\Delta_S}[P_i, \text{j, } L_k, H_l]) + \\
&\sum_{l=1}^{n_H} \sum_{k=1}^{n_L-1} \sum_{j=1}^{n_S} \sum_{i=1}^{n_P} \boldsymbol{ReLU}(\boldsymbol{\Delta_L}[P_i, S_j, \text{k, } H_l]) + \\
&\sum_{l=1}^{n_H-1} \sum_{k=1}^{n_L} \sum_{j=1}^{n_S} \sum_{i=1}^{n_P} \boldsymbol{ReLU}(\boldsymbol{\Delta_H}[P_i, S_j, L_k, \text{k}])
\end{aligned}$$

where ReLU is the rectified linear unit function: ReLU(x) = max(0,x).

Additionally, there are constraints on the 2nd derivatives of the input drivers with respect to melt pool width. The thermal gradients in DED are larger near the beginning of the build and decrease as the build progresses, and changes in melt pool width are a function of changes in temperature (Feenstra et al. 2020). Therefore, the rate at which the melt pool width increases per change in build height should be nonpositive. This was experimentally verified by the results presented here and that of Akbari and Kovacevic (2019). The finite difference approximations of this value can be calculated as a function of $\boldsymbol{\Delta_H}$ (the negative change in melt pool width with respect to height) in a similar construction as prior:

$$\boldsymbol{\Delta^2_H}[\text{P, S, L, l}] = \boldsymbol{\Delta_H}[P, \text{S, L, } H_l] - \boldsymbol{\Delta_H}[P, \text{S, L, } H_{l+1}] \leq 0$$

As well, it can be argued that if melt pool widening is due to thermal buildup, this effect will be most pronounced in smaller samples and will decrease as the wall length increases due to a longer time between laser exposures for the longer wall; such results were experimentally verified in this work. As the melt pool width is expected to decrease with increasing wall length, the second derivative of melt pool width with respect to wall length could be argued to be positive. That is, the function of melt pool width with respect to wall length is a monotonically decreasing function that approaches some lower bound. In this case, the finite different approximation of the opposite of the second derivative of melt pool width with respect to wall length is expected to be nonpositive:

$$\boldsymbol{\Delta^2_L}[\text{P, S, k, H}] = \boldsymbol{\Delta_L}[P, \text{S, } L_k, H] - \boldsymbol{\Delta_L}[P, \text{S, } L_{k+1}, H] \leq 0$$

These contributions can be added to the previous loss contribution to yield a physics-based loss contribution that enforces all of the constraints discussed.

$$\text{Loss.PHY}_2(\hat{Y}) = \text{Loss.PHY}_1(\hat{Y}) + \sum_{l=1}^{n_{H-2}} \sum_{k=1}^{n_L} \sum_{j=1}^{n_S} \sum_{i=1}^{n_P} ReLU(\Delta^2{}_H[P_i, S_j, L_k, l])$$

$$+ \sum_{l=1}^{n_H} \sum_{k=1}^{n_{L-2}} \sum_{j=1}^{n_S} \sum_{i=1}^{n_P} ReLU(\Delta^2{}_L[P_i, S_j, k, H_l])$$

Otherwise, identical neural nets were trained using 1) MSE, 2) MSE + Loss.PHY$_1$, and 3) MSE + Loss.PHY$_2$. The results of the training were analyzed to include test performance and the presence of physical inconsistences in the network predictions. The neural nets with loss functions MSE + Loss.PHY$_1$ (PGNN1) and MSE + Loss.PHY$_2$ (PGNN2) trained for more epochs before training terminated. For another direct comparison of PGNNs against traditional neural nets, a neural net with loss function MSE (NN2) was trained for the same number of epochs as the best-performing PGNN. As well, PGNNs are compared against traditional methods for regularization and training on unbalanced data. The best-performing PGNN was then inverted to produce parameter sets (power and speed) with the aim of maintaining a constant melt pool width over a build.

3. Results and Discussion

3.1 Width Measurements

Melt pool width was measured as a function of height for the thin walls with variable length and constant power and speed (see Fig. 7). In general, the melt pool width increased with increasing height. This was reported elsewhere in the literature (Akbari and Kovacevic 2019) and is due to thermal buildup resulting in an increase in the amount of material captured by the melt pool and/or a decrease of the viscosity of the melt, which encourages spreading. Other effects in DED also contribute to the variable melt pool width, independent of the thermal environment, such as the spacing between the nozzle and deposition being smaller than ideal, resulting in the initial layers being narrower due to the high thermal gradient induced by a cold substrate than at steady state where deposition is deeper into previous layers slowing down the cooling rate and resulting in heat buildup (thus wider melt pool widths). Over the course of the build, the melt pool width approaches a steady state. Identifying specific athermal effects, which increase the melt pool width with increasing build height, is beyond the scope of this report.
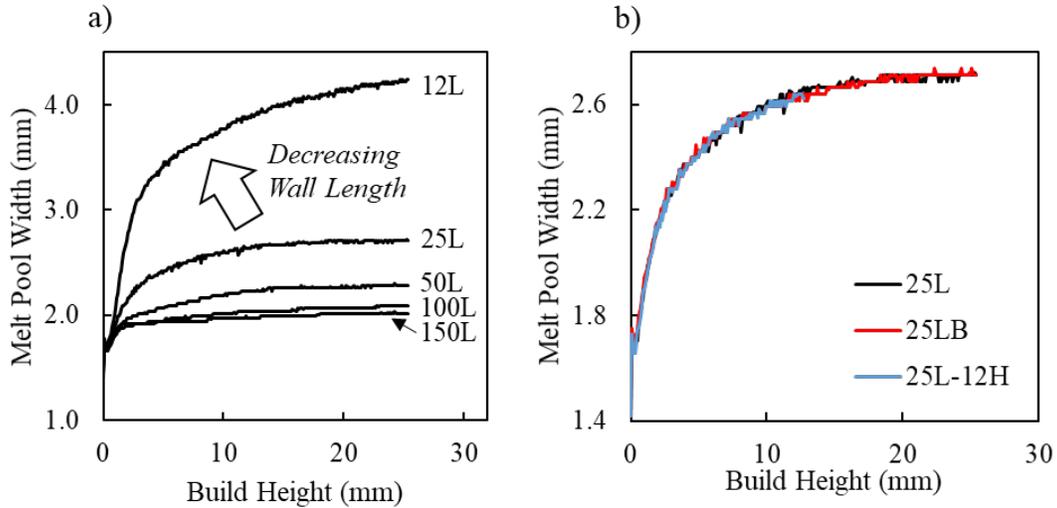
Fig. 7 **Melt pool width vs. height, (a) decreasing the wall length increased the melt pool width. For each label, the number before the L is the approximate sample length in millimeters. (b) The three 25.4-mm-long samples overlaid (black – 25L, red – 25LB, and blue – 25L-12H). A large degree of repeatability is observed.**

For walls with smaller lengths, the rate of the change in melt pool width increased faster with increasing layer height, and the melt pool width increased to an overall larger size, confirming trends found in literature.

Samples 25L, 25LB, and 25L-12H had identical processing conditions except for the variation in height or the addition of a layer after a delay. The width and height relationships are plotted in Fig. 7b. Note, the final layer (layer added after a wait) for sample 25LB is not pictured in Fig. 7b.

Melt pool width versus height of 25.4-mm length thin walls of variable power and speed can be found in Fig. 8. The trends found in literature are reproduced here: at any given height, samples fabricated at a higher power have wider melt pools, and samples fabricated at slower speeds have wider melt pools.

14
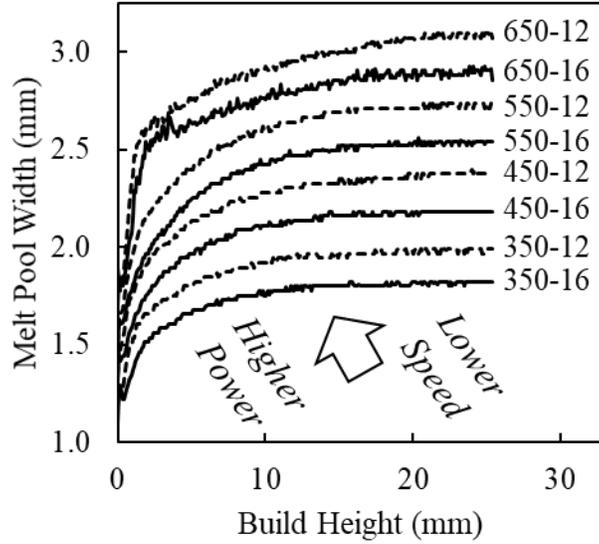
Fig. 8   Melt pool width vs. height, variable power/speed data set. Increasing power and reducing speed are observed to increase melt pool width. For each sample label, the first number is the power in watts, and the second number is the approximate speed in millimeters/second.

3.2 Machine Learning: PGNNs as Regularizers

Before neural networks were trained to predict melt pool width, optimal hidden layer size should be determined. The hidden layer size was varied, and the network was trained until validation loss did not decrease with a patience of 100. The results are in Table 4.

Table 4   Validation loss as a function of network size

Hidden layer size	Validation loss
10	0.1806
25	0.0254
50	0.0176
100	0.0221
200	0.0273

Because the network with a single hidden layer size of 50 performed the best on the validation data, this architecture was used to explore each loss function. As this network has a single hidden layer of size 50, 4 input neurons, and 1 output neuron, it has 250 weights and 51 biases, for 301 learnable parameters in total, therefore, this model is underparameterized, as it has 2510 training data points. New networks were initialized before each loss function was used. It should be noted that while hidden layer size has been tuned here for a traditional neural network (NN1), it cannot be assumed that this is the optimal network size for PGNNs. As the network

15

size is the same for each of them, and the physics-based constraints must be encoded into the network, there is less network capacity available to learn the relationship between input drivers and melt pool width. Therefore, performance of PGNNs here are lower bounds. The result of each network's test data can be found in Table 5.

Table 5   Results and number of training epochs for each model

Model	Loss	Test R-square	Epochs	Training method
NN1	MSE	0.866	967	Patience
PGNN1	MSE + Loss.PHY_1	0.951	2571	Patience
PGNN2	MSE + Loss.PHY_2	0.986	3486	Patience
NN2	MSE	0.901	3486	Same epochs as PGNN2
NN-R	MSE + Regularizers	0.912	1462	Patience

It is clear that the physics-guided neural nets (PGNN1 and PGNN2) outperformed the non-physics-guided neural net (NN1), and that the PGNNs trained for more epochs before the loss of the validation data increased. Adding the additional physics-based constraint to the loss in PGNN2 furthered these trends, training longer than and outperforming PGNN1. The performances of each network are further discussed in detail. The test predictions for NN1, from sample 450-12, are shown in Fig. 9.
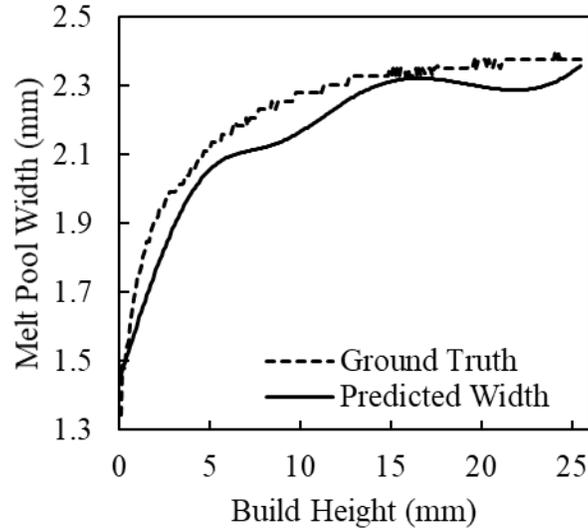

Fig. 9   Test target vs. prediction of NN1

It is immediately clear that NN1 predicted trends that are not physically expected. The melt pool width decreased temporarily at greater build height, which is not consistent with experimental data or expected from processing physics. In addition, the predicted rate at which the melt pool widened increased above 15-mm build

height where experimentally the melt pool width has leveled off. NN1 predicts the target melt pool width in the test data with a root mean square error of 0.0793 mm.

Because NN1 was outperformed by models that trained for more epochs, another comparison is to train a traditional neural net for 3486 epochs, as much as model PGNN2. The resulting model, NN2, fits the test data with an R^2 of 0.901, shown in Fig. 10, and outperforms NN1, but underperforms compared to both of the PGNNs.
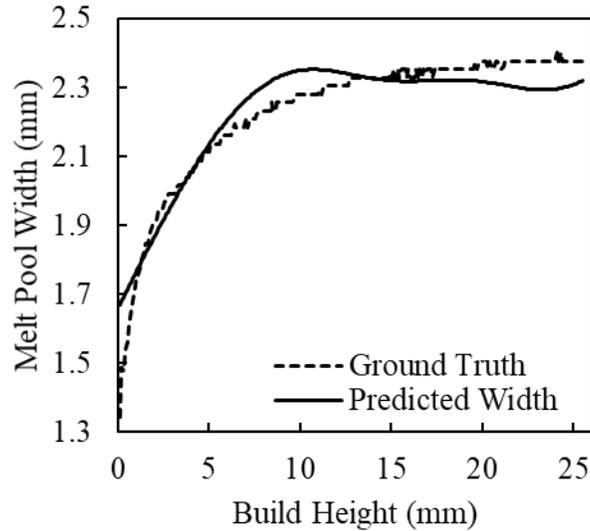

Fig. 10   Test target vs. prediction of NN2

While NN2 trained longer and outperformed NN1, it still produced predictions that are not consistent with the physical limitations that were previously described: it predicts decreases in melt pool width with increasing height and it predicts that the rate of melt pool widening increases with increasing height. However, NN2, trained for 3486 epochs, outperforms NN1, trained for 967 epochs. This suggests that the training may have halted too early.

More of the network updates pertained to patterns that extend to the validation data (and likely all data), as opposed to pertaining to patterns that only applied to the training data because the PGNNs trained for more epochs before the validation loss increased. This is entirely expected, because part of the loss function in the PGNNs corresponded to the physics-based trends applied to all the data. Since the network was required to obey these physics-based trends when updating its model weights, it trained for longer before it produced poor predictions for the validation data. Training a normal neural network for longer did not re-create these improvements, as the later model updates did not necessarily correspond to patterns that exist in nontraining data. Results from PGNN1 can be found in Fig. 11.
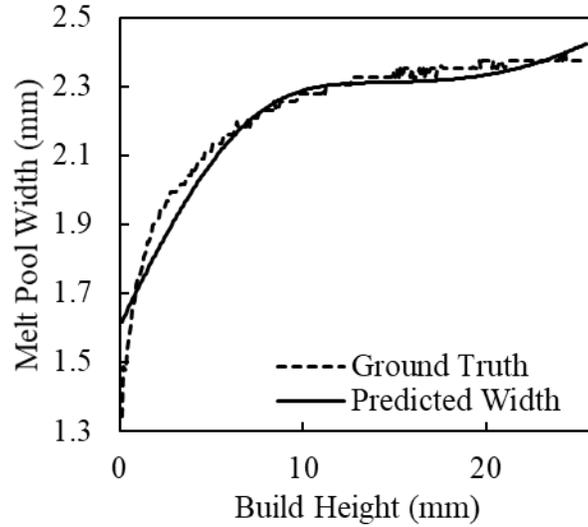
Fig. 11   Test target vs. prediction of PGNN1

Since the network was penalized for predicting a decrease in melt pool width as a function of height, it did not make any such predictions. However, as there was no penalty for predicting increases in the rate at which melt pool width increased with increasing height, this model predicted such an increase, above 20 mm in build height. Given the processing physics, there is no known mechanism that would cause rapid heat buildup or otherwise cause melt pool width increases at larger heights. PGNN1 outperformed non-physics-guided models.

Results from PGNN2, which was trained with a loss function that penalizes all physical inconsistencies discussed, can be found in Fig. 12. PGNN2 produced predictions that perform with an R^2 of 0.986 on the test data and no physical inconsistencies were predicted by this model. This model demonstrated the ability for physics-guided models to be rapidly trained to produce results both high in accuracy and respecting physical limitations.
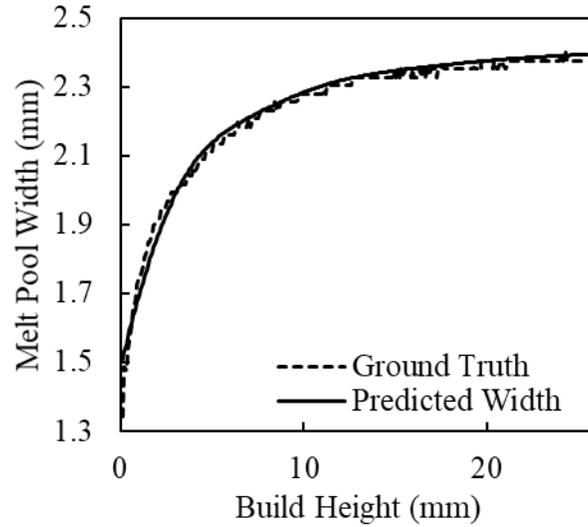
Fig. 12   Test target vs. prediction of PGNN2

It can be argued that physics-guided loss functions act as regularizers: they bias the model to solutions obeying physical rules. While additional regularizers were not employed in this work, Daw et al. (2021) show that PGNNs with traditional regularizers outperform standard neural networks with traditional regularizers. In this work, PGNNs without other regularizers are compared against a neural network with traditional regularizers. Another neural network, NN-R, with kernel regularizers ($L_1 = 10^{-5}$, $L_2 = 10^{-4}$), bias regularizers ($L_2 = 10^{-4}$), and activity regularizers ($L_1 = 10^{-5}$), was trained. These regularizer values were taken from the Keras page on layer regularizers to demonstrate they were not tuned to any data set (Chollet 2015). The testing results of NN-R are shown in Fig. 13.
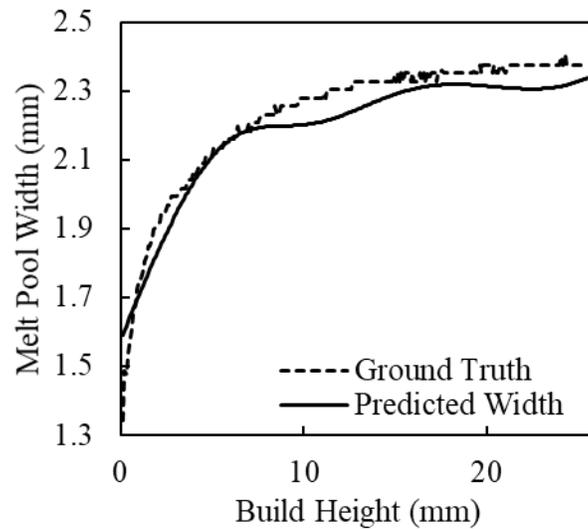

Fig. 13   Test target of NN-R

19

While NN-R performed better (R-square = 0.912) than the other neural networks, it was outperformed by both PGNNs and predicted both types of physical inconsistency. Conceptually, it can be argued that the physics-guided loss functions offer regularization that cannot be replaced by kernel/bias/activation regularization. Increasing the latter runs the risk of underfitting, biasing the model away from solutions that have large parameters or many nonzero parameters, which may actually be characteristic of the underlying function the model is trying to learn. However, there is no such risk of underfitting with a physics-based loss function as physics-based loss functions bias the model against learning physically impossible functions.

3.3 Machine Learning: PGNNs for Unbalanced Data

All of the test results were of sample 450-12, which is within the variable parameter set of samples. There were five samples in the variable length set, three of which were used to train the model. There were eight samples in the variable parameter, six of which were used to train the model. As there were twice as many variable power/speed samples than variable length samples, it could be argued that the model learned to better approximate the variable parameter set than the variable length set because the variable parameter set's data will comprise a majority of the training data's contribution to the loss function. The experiment was repeated, but with only 450-16 used for validation, and 50L and 450-12 used for testing, so that the testing performance can display model accuracy on both variable length and variable parameter data. The new division of data used for training, validation, and testing is shown in Table 6.

Table 6   Split of training, validation, and testing data in experiments

Training	Validation	Testing
12L	450-16	450-12
25L	. . .	50L
100L	. . .	. . .
150L	. . .	. . .
650-16	. . .	. . .
550-16	. . .	. . .
350-16	. . .	. . .
650-12	. . .	. . .
550-12	. . .	. . .
350-12	. . .	. . .

The neural nets trained were NN1 (standard method), NN2 (trained for as many epochs as PGNN2), PGNN1 (trained with 1st derivative physics loss contributions), and PGNN2 (trained with both 1st and 2nd derivative physics loss contributions).

An additional neural net, NN-W, in which the variable length samples were weighted as twice as important as the variable power/speed samples, was trained. This is a common strategy for training on unbalanced data. Testing data from samples 50L and 450-12 are both presented in Table 7.

Table 7   Results and number of training epochs for each model

Model	Loss	R-square 50L	R-square 450-12	Epochs	Training method
NN1	MSE	0.439	0.884	1134	Patience
PGNN1	MSE + Loss.PHY$_1$	0.950	0.962	3177	Patience
PGNN2	MSE + Loss.PHY$_2$	0.972	0.991	3948	Patience
NN2	MSE	–21.139	0.904	3948	Same epochs as PGNN2
NN-W	Weighted MSE	0.493	0.846	843	Patience

Similar trends were recorded for sample 450-12, as with the previous batch of models. However, traditional neural networks did not fit well to the data from sample 50L. Training NN2 for longer than NN1 did not result in increased performance on 50L, likely due to the model fitting more closely to variable power/speed samples than the variable length samples. Both PGNN models performed better than traditional neural nets on both samples. However, the PGNNs performed particularly better on 50L, compared to their traditional counterparts, as the physics-based contribution to their loss function forced them to fit to an underlying distribution of widths. NN1 and NN2 are more prone to predicting more effectively where there is more training data provided because they do not have access to the physics-based prior knowledge that the PGNNs do. PGNNs performed superior to weighted loss functions (NN-W), when working with unbalanced data, which is useful when sparse data sets are added to existing large ones. This may be because there are many more sets of model parameters that have low MSE errors for a given small data set than there are sets of model parameters that meet the physical constraints of the distribution that the data set came from.

3.4 Model Verification

To validate the model further, an experiment was conducted. Two thin walls were built: a 25.4-mm-long sample (25L MPS) and a 63.5-mm-long sample (63L MPS). The power and speed were varied by layer such that the predicted wall width would be 2 ±0.02 mm. The same width measurement algorithm used on the previous set of samples was applied to the in-situ monitoring videos from these builds. The width versus build height of sample 25L MPS is displayed in Fig. 14. The melt pool width for sample 25L MPS underpredicted the target width initially but then overshoots. The steady state width was 2.05 mm, 2.5% larger than the target value.
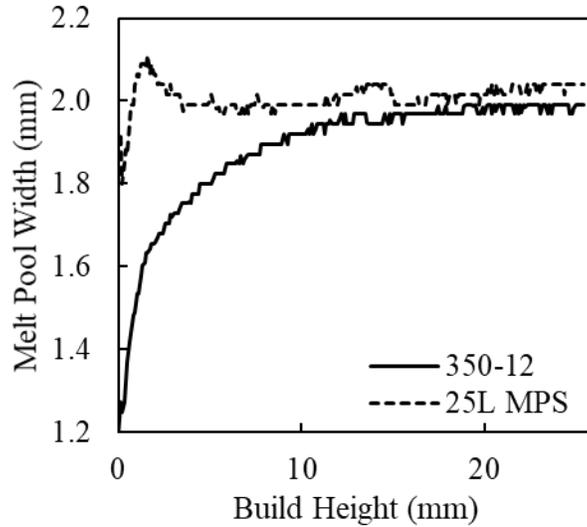
Fig. 14   Melt pool width vs. build height for sample 25L MPS. Compared to sample 350-12, sample 25L MPS had significantly less height deviation and quickly reached a steady state of 2.05-mm melt pool width.

Similar trends were seen in the melt pool width for sample 63L MPS compared to sample 50L (seen in Fig. 15): a slight undershoot in the beginning, followed by an overshoot, and then a plateau that was above the target width. The melt pool width stabilized at 2.1 mm, 5% higher than the target.
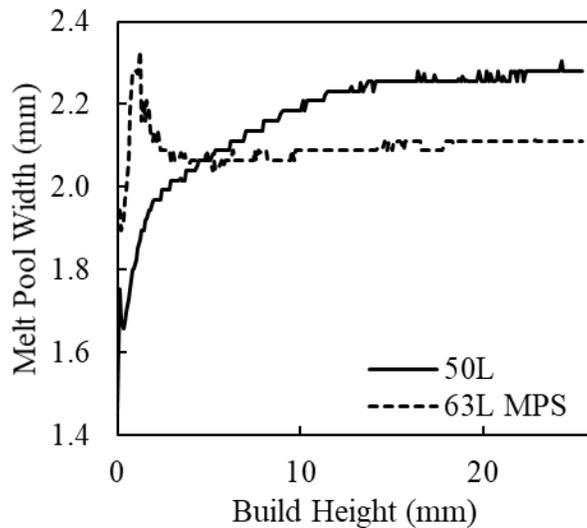

Fig. 15   Height vs. width curves for samples 50L and 63L MPS

The melt pool width is not just a function of the power, speed, and geometry of that layer, but also affected by thermal buildup from the residual heat retained in the layers processed before the current layer. Here, the power was increased during the first portion of the build, increasing the amount of heat built up for subsequent

layers. As a result, the melt pool may have been wider than the model predicts. Future models could be improved by taking inputs from all previous layers.

It should be stressed that while this work uses finite difference approximations of inequalities of 1st- and 2nd-order derivatives as constraints on neural networks, any equation can be enforced as a constraint, including but not limited to, equalities or inequalities comprised of multivariate differential equations and multivariate integrals. As long as a finite difference can be calculated, such an equation can be transformed into a constraint on a neural network.

4.   Conclusions and Future Work

This work demonstrated the ability of machine learning to predict DED melt pool width given relevant input drivers based on specific process parameters and geometry of a printed part. PGNNs were shown to outperform traditional neural networks, with a limitation that the mathematical descriptions of physical laws must be identified. This work identified trends in DED melt pool width, justified by modeling, experimental results, and physical reasoning, and enforces those trends as constraints on neural networks. The following conclusions were drawn:

- PGNNs are able to model the interdependences between process input drivers (power and speed) and geometric input drivers (built height and wall length) to DED melt pool width.

- This has implications for driving down cost and process time, postprocessing cost and time, as well as opening up the use of DED for broader applications.

- Neural networks often produce predictions not consistent with known physics, such as melt pool width decreasing as a function of build height.

- PGNNs are constructed to predict things that are physically consistent.

- Adding additional physics-based terms not only improved compliance with known physics but also improved performance.

- PGNNs act as regularized NNs, at no increased prediction cost and at minimal extra training cost as they bias the network away from physically unsound solutions, as opposed to traditional regularization that bias the model away from solutions that have large weights or many non-zero weights.

- PGNNs are superior to train on unbalanced data than traditional neural networks with weighted loss functions.

Because physics-guided loss functions restrict the family of functions that the network can represent to functions that do not predict physical inconsistencies, the network is provided incentive to make weight/bias updates that are extensible to the validation data. As a result, PGNNs can train for more epochs before patience-based training terminates. It is demonstrated that the increased performance is not merely a result of longer training, as a traditional neural network trained for the same number of epochs as a PGNN will still underperform compared to the PGNN. It should be noted that physics-guided loss functions can be employed on any loss-supervised model, not just neural networks.

In this work, length and height of a thin wall are the geometric input drivers. In practice, more complex representations of local geometry may be employed, with similar physics-based constraints applied. PGNNs may have applications in any materials science or AM problem where physics-based insights exist. For example, it could be hypothesized that cooling rates may be partially predicted from process parameters and geometric information, as evidenced by the relationship between wall width and process parameter and wall length. Understanding the relationships between cooling rates, process parameters, and geometry may allow the prediction of final microstructure: phase and grain size. As a result, constraints may be placed on microstructure trends as a function of process parameters and geometric information. These constraints may be enforced on machine learning models, using the methods described, to predict these properties with higher accuracy and without predicting physical inconsistencies. Ongoing work seeks to apply these methods to predict properties such as microstructure, as well as to apply this work to arbitrary geometries.

5. References

Akbari M, Kovacevic R. Closed loop control of melt pool width in robotized laser powder–directed energy deposition process. Int J Adv Manuf Technol. 2019;104(5–8):2887–2898.

Bishop CM. Pattern recognition and machine learning. Springer; 2006.

Chollet F. Keras. GitHub, Inc.; 2015. https://github.com/fchollet.

Colosimo BM, Huang Q, Dasgupta T, Tsung F. Opportunities and challenges of quality engineering for additive manufacturing. J Qual Technol. 2018;50(3):233–252.

Daw A, Karpatne A, Watkins W, Read J, Kumar V. Physics-guided neural networks (PGNN): an application in lake temperature modeling. ArXiv:1710.11431. 2021;1–16. doi:10.48550/arXiv.1710.11431.

DebRoy T, Mukherjee T, Milewski JO, Elmer JW, Ribic B, Blecher JJ, Zhang W. Scientific, technological and economic issues in metal printing and their solutions. Nat Mater. 2019;18(10):1026–1032.

DebRoy T, Wei HL, Zuback JS, Mukherjee T, Elmer JW, Milewski JO, Beese AM, Wilson-Heid A, De A, Zhang W. Additive manufacturing of metallic components – process, structure and properties. Prog Mater Sci. 2018;92:112–224.

Diederik P, Kingma JLB. Adam: a method for stochastic optimization. arXiv preprint. 2014. https://arxiv.org/abs/1412.6980.

Diehl B, Nassar A. Reducing near-surface voids in metal (Ti-6Al-4V) powder bed fusion additive manufacturing: the effect of inter-hatch travel time. Addit Manuf. 2020;36.

Everton SK, Hirsch M, Stravroulakis P, Leach RK, Clare AT. Review of in-situ process monitoring and in-situ metrology for metal additive manufacturing. Mater Des. 2016;95:431–445.

Feenstra DR, Cruz V, Gao X, Molotnikov A, Birbilis N. Effect of build height on the properties of large format stainless steel 316L fabricated via directed energy deposition. Addit Manuf. 2020;34:101205.

Gobert C, Arrieta E, Belmontes A, Wicker RB, Medina F, McWilliams B. Conditional generative adversarial networks for in-situ layerwise additive manufacturing data. Proceedings of the 30th Annual International Solid Freeform Fabrication Symposium; 2019 Aug 12–14; Austin, TX. p. 192–201.

Huang Y, Khamesee MB, Toyserkani E. A new physics-based model for laser directed energy deposition (powder-fed additive manufacturing): from single-track to multi-track and multi-layer. Opt Laser Technol. 2019;109:584–599.

Kriczky DA, Irwin J, Reutzel EW, Michaleris P, Nassar AR, Craig J. 3D spatial reconstruction of thermal characteristics in directed energy deposition through optical thermal imaging. J Mater ProcessTechnol. 2015;221:172–186.

Liu Z, Li T, Ning F, Cong W, Kim H, Jiang Q, Zhang H. Effects of deposition variables on molten pool temperature during laser engineered net shaping of Inconel 718 superalloy. Int J Adv Manuf Technol. 2019;102(1–4):969–976.

Tapia G, Elwany A. A review on process monitoring and control in metal-based additive manufacturing. J Manuf Sci Eng. 2014;136(6):060801–060810.

List of Symbols, Abbreviations, and Acronyms

3-D	three-dimensional
AM	additive manufacturing
ARL	Army Research Laboratory
DED	directed energy deposition
DEVCOM	US Army Combat Capabilities Development Command
H	build height
IR	infrared
L	wall length
MPS	multiple parameter set
MSE	mean squared error
NN	neural network
P	power
PGNN	physics-guided neural network
ReLU	rectified linear unit
S	speed
UV	ultraviolet
W	melt pool width