Data-Driven Prediction of the High-Dimensional Thermal History in Directed Energy Deposition Processes via Recurrent Neural Networks

Mojtaba Mozaffar[*], Arindam Paul[**], Reda Al-Bahrani[**], Sarah Wolff[*], Alok Choudhary[**], Ankit Agrawal[**], Kornel Ehmann[*], and Jian Cao [*, 1]

[*] Department of Mechanical Engineering, Northwestern University, Evanston, IL 60208

[**] Department of Electrical Engineering & Computer Science, Northwestern University, Evanston, IL 60208

Abstract

Directed Energy Deposition (DED) is a growing additive manufacturing technology due to its superior properties such as build flexibility at multiple scales and limited waste. However, both experimental and physics-based models have limitations in providing accurate and computationally efficient predictions of process outcomes, which is essential for real-time process control and optimization. In this work, a recurrent neural network (RNN) structure with a Gated Recurrent Unit (GRU) formulation is proposed for predicting the high-dimensional thermal history in DED processes with variations in geometry, build dimensions, toolpath strategy, laser power and scan speed. Our results indicate that the model can accurately predict the thermal history of any given point of the DED build on a test-set database with limited training. The model's general applicability and ability to accurately predict thermal histories has been demonstrated through two overarching tests conducted for long time spans and non-trained geometries.

Keywords: Additive Manufacturing; Directed Energy Deposition; Artificial Neural Networks; Deep Learning; Recurrent Neural Network; Process Parameters.

[1] Corresponding author: jcao@northwestern.edu. Phone: 847-467-1032

1. Introduction

Additive manufacturing (AM) is a rapidly growing technology, which provides unique capabilities for building complex and one-of-a-kind parts especially in the aerospace and biomedical industries. Directed Energy Deposition (DED) is a metal powder-blown AM process that uses a laser beam to locally melt the deposited powder as it moves along a specific toolpath building a part layer-by-layer.

Despite of its many intriguing features, there are challenges associated with DED due to the inconsistency and sensitivity of the final properties on process conditions. Experimental studies have been conducted to analyze the influence of process parameters on microstructure and build properties in [1-5]. However, these models use a limited number of experiments and are incapable of taking into consideration the complex inter-connectivity of parameters in AM processes. Physics-based models were proposed to capture thermo-mechanical, thermo-fluidic, and/or microstructural evolution of the process in [6-9]. However, these models are not applicable in many cases because of their enormous computational cost that might take weeks or months of processing time, even on supercomputers [10], and their discrepancy with experimental results due to the simplifying assumptions made or incomplete physics.

Current trends in manufacturing, such as Industry 4.0 [11] and cyber-physical systems [12], increase the visibility and accessibility to information, which leads to extensive research in data-driven models in the manufacturing community [13-16]. The application of machine learning in a polymer powder bed fusion process was discussed in [17]. A data-driven model for characterizing geometrical dimensions of trace products using dense neural networks was proposed by [18]. In [19], the authors compared multiple regression models for developing a surrogate model for AM

simulation. The self-organizing map technique was used for quantifying the geometric accuracy of the Fused Filament Fabrication process in [20].

Due to the limited number of samples and the omission of crucial time-series features of the process such as the toolpath, which directly affects thermal history and hierarchical microstructure in AM, current data-driven models fail to provide a profound understanding of this process. To address this gap, we propose a recurrent neural network structure for predicting the thermal history of any given point in a DED build in a many-to-many configuration. The proposed approach best suits DED processes since it can accurately calculate the high-dimensional thermal history of the builds in a computationally efficient manner.

2. Methodology

A Recurrent Neural Network (RNN) is a special type of artificial neural network (ANN), designed with the idea that the outcome of each neuron is dependent on its input (like other types of ANN) and a history variable from past operations, which enables this structure to work with sequential data. In particular, the Long Short-Term Memory (LSTM) [21] and the Gated Recurrent Unit (GRU) [22] are two successful formulations of RNN structures for training long sequences of data. The underlying concept of RNN structures, i.e., the use of information from previous steps combined with the current state of the system, is in line with differential equations. Thus, physics that can be formulated with ordinary or partial differential equations, such as the finite element method (FEM), are good applications for RNN models. RNNs allow for the temporal dependency in the input data to be learned without the need to specify a fixed set of lagged observations. Traditional time series analysis such as auto-correlation requires the identification of seasonality and stationarity in a time series, which may change based on laser speed, size of the build, etc., and need to be explicitly adjusted for each simulation. Further, neural networks are robust to noise

in input data and the output variable and can learn in spite of missing values. These properties of RNNs led to the hypothesis that the RNN structure can predict the temperature history in DED regardless of its highly nonlinear nature.

2.1. RNN Model Architecture

A stacked RNN structure with GRU formulation is considered in this research as depicted in **Fig. 1**. Each GRU cell receives input features (x_t) for that time step and a hidden state from the previous time step (h_{t-1}) and outputs a new hidden state (h_t). The number of units in each GRU cell represents the dimension of the hidden state, which is connected to other cells using weights and biases. Using multiple layers of the GRU structure enables the model to comprehend deeply hidden correlations in the data. Fully connected layers are considered to combine the outputs of the GRU units into a single time-series temperature output.

The stacked GRU model is developed using the Keras deep learning library [23] with the Adam [24] optimization procedure and a mean-square-error (MSE) cost function between the model prediction and the database thermal history.

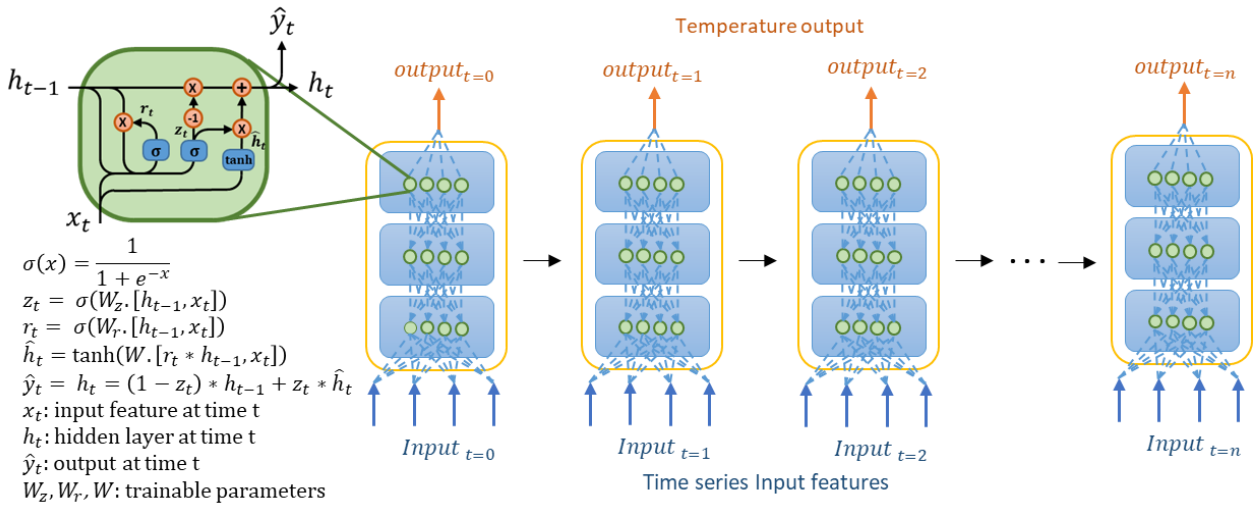

Fig. 1 Schematic of the many-to-many stacked RNN structure with GRU formulation; Green circles represent GRU units, blue rectangles represent GRU cells, yellow boundaries represent stacked GRU wrappers, and blue dashed lines within the GRU wrappers represent trainable parameters. The schematic and formulation of GRU units are provided on the left based on the formulation given in [22].

2.2. Database Development

The database considered for training the model is built using an in-house finite element code, GAMMA, for transient thermal analysis using an explicit solver [25]. This choice enables us to have access to high-volumes of input data required for training the model to learn about the phenomena occurring not only on the exterior of the AM build but also within its interior. A wide range of input parameters such as laser power (500-1000 W), scan speed (5-30 mm/s), toolpath strategy (e.g., zigzag, unidirectional or circular motion), geometric size (5-40 mm), and shapes (e.g., cubic, cylindrical, thin wall) are used to generate over 250,000 training points for DED simulation of stainless steel 316L objects.

The input features designed for each training point include toolpath feature (based on the distance between the coaxial powder nozzle outlet and the birthed element), the time of deposition, closest distance to the boundary of the build, layer height, laser intensity, and laser state (on or

off). Input features are extracted as time series data from FEM output files and stored in a three-dimensional tensor compatible with the RNN structure. To investigate the capability of the RNN structure to operate in noisy environments, an artificial noise with a standard deviation of 30 K is added to the thermal history in the database. All input features and thermal histories are normalized using a linear mapping from the minimum and maximum of each feature to the range of 0 to 1 in order to speed-up the optimization procedure. The model is trained over 80% of the database, while 20% of the database is left out for testing.

3. Results and Discussions

3.1. Test-Set Evaluation

The RNN model is trained for different configurations such as number of RNN layers (1-5 layers), GRU units (100-500 units), and fully connected layers (1-3 units). Even for configurations with a small number of layers and units the model reaches 1e-4 MSE after 100 epochs of training. An epoch is a complete pass of training through the database in batch mode. For a configuration of 3 stacked GRU layers with 500 units and 1 fully connected layer with 100 neurons, the model reaches 3.210e-5 MSE on training data and 3.84e-5 on testing data with 100 epochs of training. The training process took approximately 40 hours on a Nvidia Quadro P5000. The prediction for two random points of a test-set over 20 s of the process is demonstrated in **Fig. 2**. The smoothness of the predicted thermal histories and the similarity between the patterns of the applied noise and the prediction error (discrepancy between the noisy database and predicted thermal history) proves that the RNN structure can effectively avoid noise in the database and capture most of the critical features of the thermal history including sharp changes and flat regions that happen due to the phase transitions in the material.

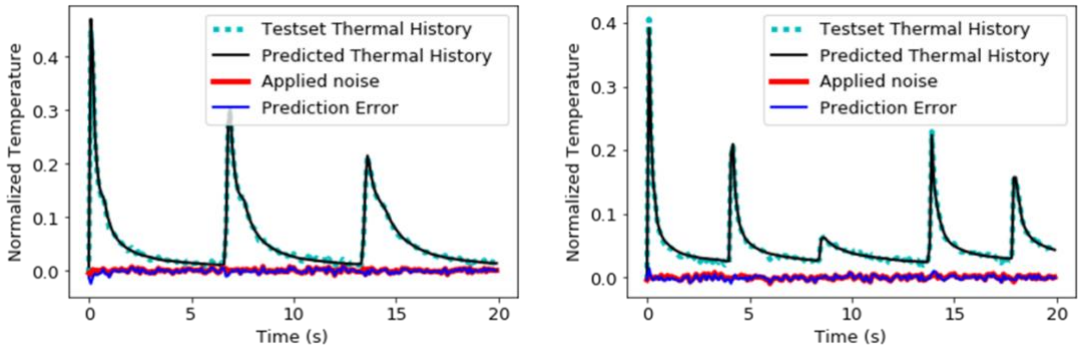

Fig. 2 Evaluation of the stacked RNN model on the test dataset for two random points over 20 s of the DED process; Comparison of the model prediction (black line) and the test-set value (cyan dashed line) for the thermal history of a point in a thin-wall build with uni-directional toolpath (left) and a cylindrical build circular toolpath (right).

3.2. Overarching Test I: Long Time Span Test

One of the key features of RNN structures for this application is that it can predict any arbitrary number of time increments (0.1 s in this case). To assess the generality of the model, a trained model is used for predicting a longer time span than it has been trained for. Particularly, two stacked RNN models trained for 20 s and 50 s of the process are used for predicting 100 s of the thermal history. The model trained on 20 s (**Fig. 3** top) performs well for the training period and continues to have a reasonable prediction for more than three times of the time span it has been trained on, to have an overall MSE of 7.05e-5 for 100 s of the process. The model trained on 50 s (**Fig. 3** bottom) decreases the MSE to 3.17e-5 for 100 s of the prediction. This significant drop in the long-term error of the model comes with the cost of almost two times the required training time and resources. It is noteworthy that the sharp gradients, melting and re-melting of the material, which are the main targets of this work, happen in a few layers after material deposition and after that the temperature changes smoothly. Therefore, the stacked RNN is an effective model for

predicting the most critical features of the thermal history, while the long-term predictions can be easily improved if necessary.

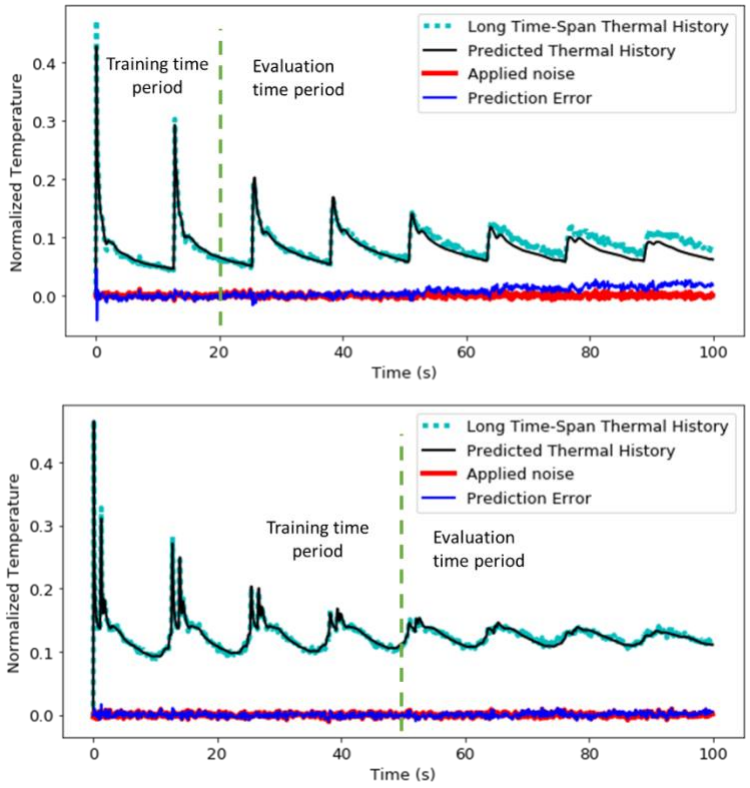

Fig. 3 Evaluation of the stacked RNN model on the test dataset for 100 s, while trained on 20 s (top) and 50 s (bottom) of the process; Comparison of the model prediction (black line) and the test-set value (cyan dashed line) for the long-time span thermal history of a point in a cubic build with zigzag toolpath.

3.3. Overarching Test II: Dissimilar Geometries

The capability of the trained model is further investigated for predicting the behavior of a dissimilar geometry type from the training database. A new database with different geometric features is developed, and the trained model is used to predict the thermal history of three points of the build. As depicted in **Fig. 4**, the model can accurately predict the behavior of points 1 and

3. However, there is significant error in the model prediction for point 2, which is because the geometric feature and the state of the boundaries close to this point is unprecedented to the model. The results indicate that even with two hand-picked geometric features introduced in 2.2 (i.e., closest distance to a boundary of the build and layer height features) the model can perform reasonably well at material points of untrained builds with similar geometric features as the training database. However, to represent complex geometries, a more flexible geometric feature extraction methodology will be needed.

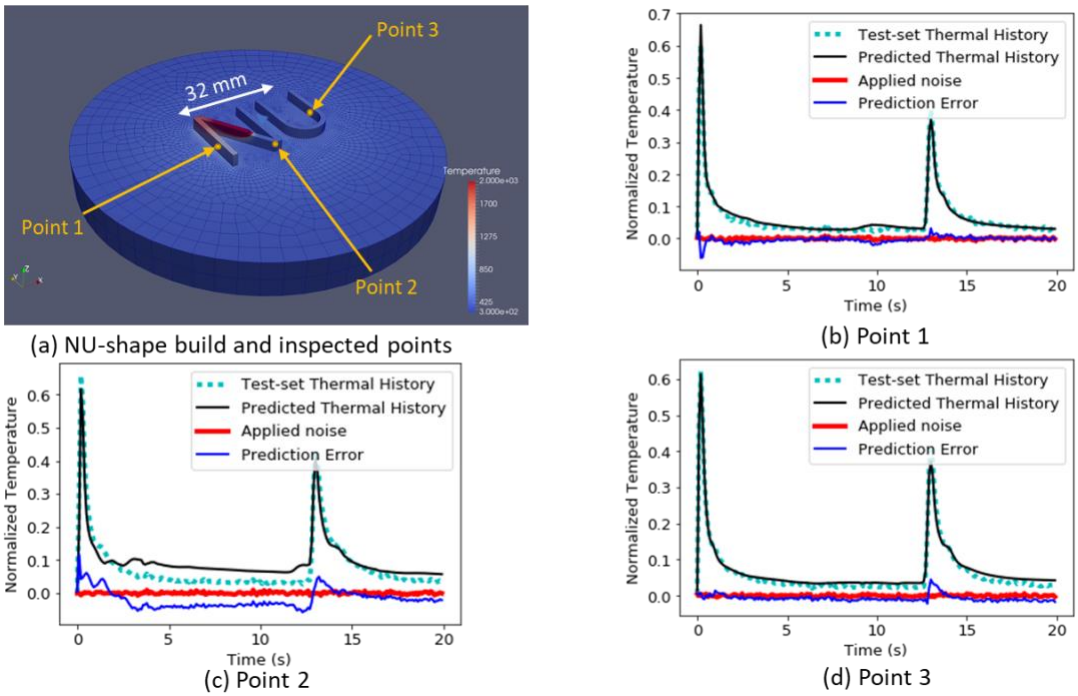

Fig. 4 Evaluating the trained model on a dissimilar geometry; The NU-shape build geometry and the inspected point locations (a), comparative figures for the points 1, 2 and 3 between model prediction and the test-set (b), (c), and (d), respectively. The toolpath of this build goes from the buttom left to the upper right side of the letter N and then moves from the upper left to the upper right side of the letter U.

4. Conclusions and future works

This paper presents a stacked RNN structure with a GRU formulation for predicting the thermal history of any given point in a DED build. The results show that the model reaches an MSE of 2.97e-5 on a test dataset after a 100 epoch training. Additionally, two overarching tests for predicting the thermal history over a longer time span and non-trained geometries are examined, showing the potential of RNN models to predict complex behaviors in AM processes, considering that the accuracy of the model can be further improved by increasing training epochs and geometry types included in the database.

Future work of this research includes directly training the model based on experimental data and developing an optimization framework for AM based on Reinforcement Learning.

5. Acknowledgement

The authors acknowledge the support by the National Institute of Standards and Technology (NIST) - Center for Hierarchical Materials Design (CHiMaD) under grant No. 70NANB14H012, the National Science Foundation (NSF) - Cyber-Physical Systems (CPS) under grant No. CPS/CMMI-1646592, and DMDII (Digital Manufacturing Design Innovation Institute).

6. References

[1] Shamsaei N, Yadollahi A, Bian L, and Thompson SM. An overview of Direct Laser Deposition for additive manufacturing; Part II: Mechanical behavior, process parameter optimization and control. Additive Manufacturing 2015; 8: 12-35.

[2] Stevens EL, Toman J, To AC, and Chmielus M. Variation of hardness, microstructure, and Laves phase distribution in direct laser deposited alloy 718 cuboids. Materials & Design 2017; 119: 188-198.

[3] Wang Z, Palmer TA, and Beese AM. Effect of processing parameters on microstructure and tensile properties of austenitic stainless steel 304L made by directed energy deposition additive manufacturing. Acta Materialia 2016; 110: 226-235.

[4] Fisher BA, Lane B, Yeung H, and Beuth J. Toward determining melt pool quality metrics via coaxial monitoring in laser powder bed fusion. Manufacturing Letters 2018.

[5] Li H, Ramezani M, Li M, Ma C, and Wang J. Effect of process parameters on tribological performance of 316L stainless steel parts fabricated by selective laser melting. Manufacturing Letters 2018.

[6] Chiumenti M, Lin X, Cervera M, Lei W, Zheng Y, and Huang W. Numerical simulation and experimental calibration of Additive Manufacturing by blown powder technology. Part I: thermal analysis. Rapid Prototyping Journal 2017; 23(2): 448-463.

[7] Wang Z and Beese AM. Effect of chemistry on martensitic phase transformation kinetics and resulting properties of additively manufactured stainless steel. Acta Materialia 2017; 131: 410-422.

[8] Bostanabad R, Zhang Y, Li X, Kearney T, Brinson LC, Apley DW, Liu WK, and Chen W. Computational microstructure characterization and reconstruction: Review of the state-of-the-art techniques. Progress in Materials Science 2018.

[9] Bostanabad R, Kearney T, Tao S, Apley DW, and Chen W. Leveraging the nugget parameter for efficient Gaussian process modeling. International Journal for Numerical Methods in Engineering 2018; 114(5): 501-516.

[10] Francois MM, Sun A, King WE, Henson NJ, Tourret D, Bronkhorst CA, Carlson NN, Newman CK, Haut TS, and Bakosi J. Modeling of additive manufacturing processes for metals: Challenges and opportunities. Current Opinion in Solid State and Materials Science 2017; 21(LA-UR-16-24513).

[11] Zühlke D. Industrie 4.0: From Vision to Reality. 2013, tech. report, SmartFactory KL.

[12] Lee J, Bagheri B, and Kao H-A. A cyber-physical systems architecture for industry 4.0-based manufacturing systems. Manufacturing Letters 2015; 3: 18-23.

[13] Lee J, Lapira E, Bagheri B, and Kao H-a. Recent advances and trends in predictive manufacturing systems in big data environment. Manufacturing Letters 2013; 1(1): 38-41.

[14] O'Donovan P, Gallagher C, Bruton K, and O'Sullivan DT. A fog computing industrial cyber-physical system for embedded low-latency machine learning Industry 4.0 applications. Manufacturing Letters 2018.

[15] Bessa M, Bostanabad R, Liu Z, Hu A, Apley DW, Brinson C, Chen W, and Liu WK. A framework for data-driven analysis of materials under uncertainty: Countering the curse of dimensionality. Computer Methods in Applied Mechanics and Engineering 2017; 320: 633-667.

[16] Bostanabad R, Bui AT, Xie W, Apley DW, and Chen W. Stochastic microstructure characterization and reconstruction via supervised learning. Acta Materialia 2016; 103: 89-102.

[17] Baturynska I, Semeniuta O, and Martinsen K. Optimization of Process Parameters for Powder Bed Fusion Additive Manufacturing by Combination of Machine Learning and Finite Element Method: A Conceptual Framework. Procedia CIRP 2018; 67(1): 227-232.

[18] Caiazzo F and Caggiano A. Laser Direct Metal Deposition of 2024 Al Alloy: Trace Geometry Prediction via Machine Learning. Materials 2018; 11(3): 444.

[19] Kamath C and Fan YJ. Regression with small data sets: a case study using code surrogates in additive manufacturing. Knowledge and Information Systems 2017: 1-19.

[20] Khanzadeh M, Rao P, Jafari-Marandi R, Smith BK, Tschopp MA, and Bian L. Quantifying Geometric Accuracy With Unsupervised Machine Learning: Using Self-Organizing Map on Fused Filament Fabrication Additive Manufacturing Parts. Journal of Manufacturing Science and Engineering 2018; 140(3): 031011.

[21] Hochreiter S and Schmidhuber J. Long short-term memory. Neural computation 1997; 9(8): 1735-1780.

[22] Cho K, Van Merriënboer B, Gulcehre C, Bahdanau D, Bougares F, Schwenk H, and Bengio Y. Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv preprint arXiv:1406.1078 2014.

[23] Chollet F. Keras. 2015.

[24] Kingma DP and Ba J. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 2014.

[25] Smith J, Xiong W, Cao J, and Liu WK. Thermodynamically consistent microstructure prediction of additively manufactured materials. Computational mechanics 2016; 57(3): 359-370.