

ARTICLE OPEN



Superior printed parts using history and augmented machine learning

Meng Jiang^{1,2}[✉], Tuhin Mukherjee², Yang Du² and Tarasankar DebRoy²[✉]

Machine learning algorithms are a natural fit for printing fully dense superior metallic parts since 3D printing embodies digital technology like no other manufacturing process. Since traditional machine learning needs a large volume of reliable historical data to optimize many printing variables, the algorithm is augmented with human intelligence derived from the rich knowledge base of metallurgy and physics-based models. The augmentation improves the computational efficiency and makes the problem tractable by enabling the algorithm to use a small set of data. We provide a verifiable quantitative index for achieving fully dense superior parts, facilitate material selection, uncover the hierarchy of important variables that affect the density, and present easy-to-use visual process maps. These findings can improve the quality consistency of 3D printed parts that now limit their greater industrial adaptation. The approach used here can be applied to solve other problems of 3D printing and beyond.

npj Computational Materials (2022)8:184; <https://doi.org/10.1038/s41524-022-00866-9>

INTRODUCTION

3D printing or additive manufacturing (AM) enables one-step fabrication of intricate metallic parts that cannot be easily made by other manufacturing processes^{1–6}. Because of its advantages, metallic parts made by AM are of growing interest in aerospace, automotive, energy, and other industries^{7–9}. The printed metallic parts are the fastest-growing sector of AM¹. However, AM now represents only a small portion of the global manufacturing market because of persistent problems with the consistency and quality of the printed parts². Although progress is being made in overcoming the important problems of AM, it still underperforms traditional manufacturing processes in quality consistency, especially in the production of fully dense metallic parts that do not contain the lack of fusion and other voids^{1,2}.

Several approaches have been undertaken to improve the density of the printed metal parts by reducing the lack of fusion voids in powder bed fusion (PBF) as shown in the supplementary document^{10–21}. For example, dense parts, largely free of internal voids have been achieved by expensive hot isostatic post-processing^{22,23}, changing processing variables aimed at making denser parts by trial and error^{10–13}, and the use of machine learning using data obtained from previous runs and independent experiments. Post-processing imposes an extra cost and does not always eliminate the defects^{2,3}. The optimization of property by trial and error is tenuous because of the large number of variables. The evolution of the lack of fusion voids depends on multiple, simultaneously occurring, complex physical phenomena, and their mechanistic understanding is not yet fully developed^{3,4}. Therefore, mechanistic modeling of lack of fusion void formation for different alloys and process conditions is not always a viable path⁴. In contrast, machine learning can reveal the correlation between various process variables and lack of fusion using experimental data without the need to understand the mechanisms of their formation³.

The AM processes enjoy a tremendous advantage over the conventional manufacturing technologies in connectivity and communication with computers and the Internet^{3,24}. No other

processing technology has been designed to receive, use, and transmit digital data like AM. External optical sensors can collect micron-level detail using cameras and the data can be very useful for machine learning-based solutions²⁰. As a result, the implementation of machine learning solutions for AM has received considerable attention³ to overcome the current issues faced by AM. Three types of data, real-time data, historical data from the same AM unit, and literature data have been used to impact all aspects of the production of metallic parts³. The data and the machine learning can be used for closed-loop quality control, optimizing process variables, and expediting part qualification³ without depending solely on the inspection process.

One key requirement for building and deploying traditional machine learning solutions for defect remediation and other problems is the need for a large volume of high-quality data^{3,4}. Many variables affect the formation of defects such as the laser power, laser spot size, laser scanning speed, preheat temperature, hatch spacing, layer thickness, shape and size of the powder, specific heat, thermal conductivity, density, solidus and liquidus temperatures, and the latent heat of fusion^{25,26}. The number of data needed to consider the contributions of these 14 variables is at least 2^{14} (=16,384) using the 2-factor design of experiments²⁷. The sheer volume of the required data makes such an investigation challenging. Thus, the number of input variables needs to be reduced from 14 to a lower value to make the problem tractable.

An appropriate variable reduction technique that uses a combination of variables rather than the individual variables is needed. Augmentation of human intelligence based on the rich knowledge base of metallurgy and mechanics is needed for this purpose to construct non-dimensional numbers from raw process variables and materials properties to achieve this goal²⁸. Several such non-dimensional numbers may then be correlated with the part density or other attributes using an appropriate machine learning framework. The resulting reduction of the variables makes the problem tractable.

¹State Key Laboratory of Advanced Welding and Joining, Harbin Institute of Technology, Harbin 150001, China. ²Department of Materials Science and Engineering, The Pennsylvania State University, University Park, PA 16802, USA. ✉email: mengjiang@hit.edu.cn; debroy@psu.edu

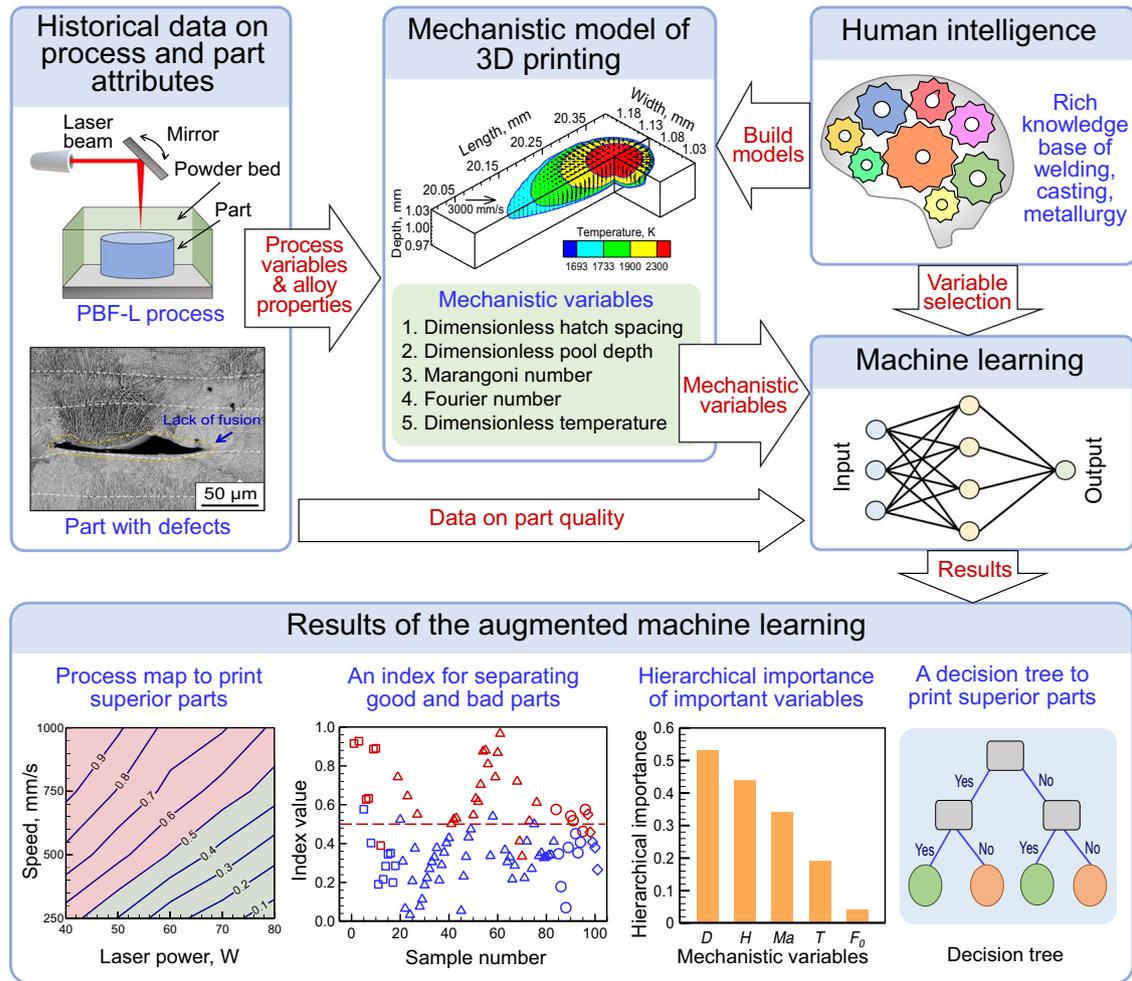


Fig. 1 Schematic representation of the methodology used. Process variables used in the experiments and the thermophysical properties of alloys are used in the mechanistic model to calculate the five mechanistic variables that influence the lack of fusion defect formation. These variables when used to train augmented machine learning algorithms can guide engineers to find out conditions for avoiding the lack of fusion defects in printed metallic parts. The experimentally observed lack of fusion defect⁶⁰ is for Inconel 718 made by powder bed fusion. The panel “Part with defects” is adapted with permission from reference⁶⁰, Elsevier. The panel “Human intelligence” is adapted with permission from reference⁶¹, Elsevier.

We implement the augmented machine learning^{29–31} strategy and synergistically combine a mechanistic model and historical experimental data to uncover the conditions necessary to reduce the lack of fusion void formation in laser powder bed fusion (PBF-L) (Fig. 1). We analyze one hundred and one independent experimental data for an aluminum alloy^{10,32–37}, AlSi10Mg, titanium alloy^{13,38–44}, Ti6Al4V, stainless steel^{45–47}, SS316, and a nickel alloy^{48,49}, Inconel 718. Based on the previous work^{4,12,15,38,49}, we identify (see the supplementary document) five important variables, dimensionless hatch spacing (hatch spacing/pool width), dimensionless pool depth (pool depth/layer thickness), dimensionless temperature (peak temperature/liquidus temperature), Marangoni number, and Fourier number that affect the lack of fusion defects. We use a well-tested model of heat transfer and fluid flow in the PBF-L to calculate the five mechanistic variables. These variables are analyzed using a decision tree and linear regression to forecast the lack of fusion. Furthermore, the hierarchical importance of these five mechanistic variables is determined using three feature selection indexes such as information gain, information gain ratio, and Gini index.

Here we show that using the augmented machine learning results, an easy-to-use, verifiable lack of fusion index can be constructed based on scientific principles. In addition, the

hierarchical influence of the important variables on the lack of fusion can be uncovered. The engineers can then know which variables to adjust to minimize the lack of fusion. Furthermore, process maps could be constructed to select the ranges of processing conditions to avoid the lack of fusion defects in parts. These maps were tested with independent experimental data and are useful for real-time use. In addition, the methodology helps in materials selection for obtaining fully dense parts and can support the discovery of new printable alloys.

RESULTS AND DISCUSSION

The mechanistic variables

The lack of fusion voids originate from the inadequate fusional bonding among the neighboring deposited tracks which are affected by heat and fluid flow and molten pool geometry. Here we select five mechanistic variables that capture the combined influence of process variables and alloy properties on the lack of fusion defects. The basis for the selection of mechanistic variables and their effects on the formation of lack of fusion voids^{4,12,15,38,49–52} are explained below.

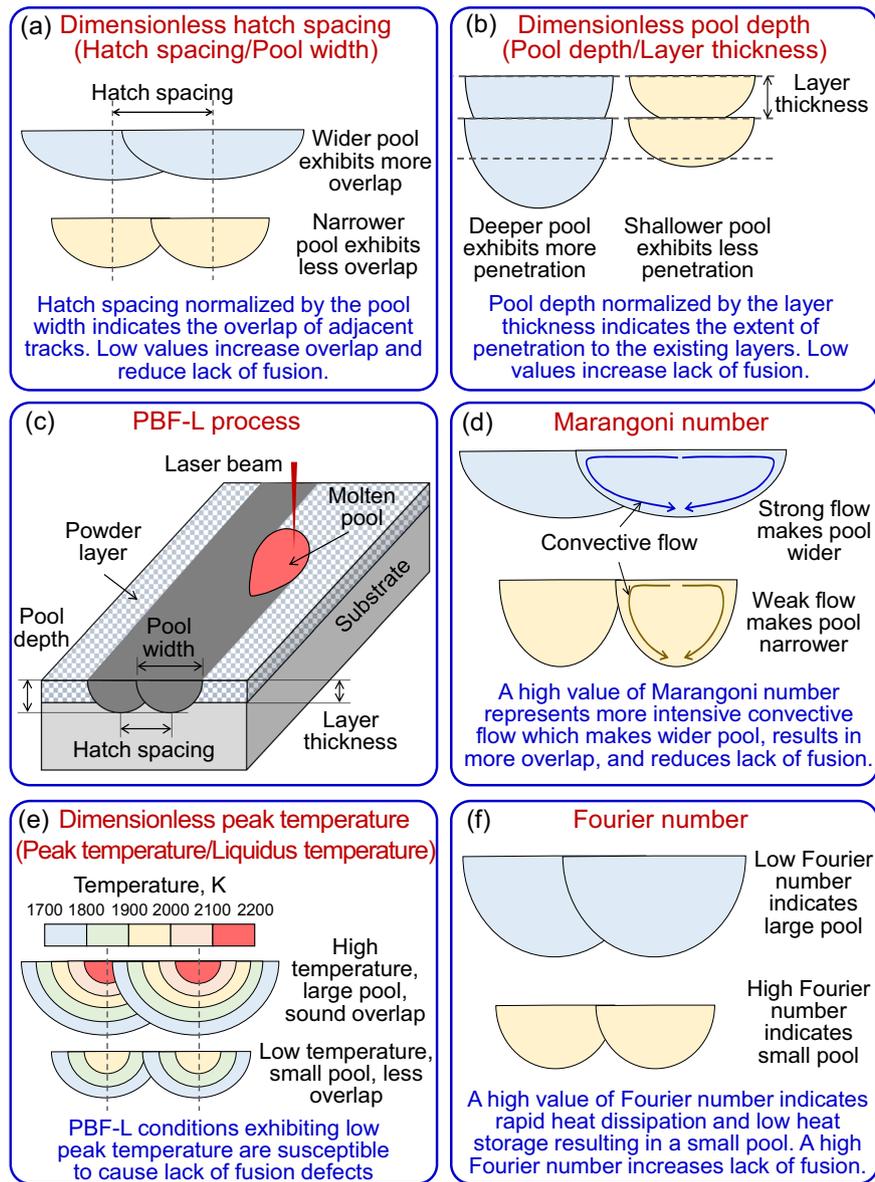


Fig. 2 The five mechanistic variables and their effects on the lack of fusion defect in PBF-L parts. **a** Dimensionless hatch spacing is represented by the ratio of hatch spacing to pool width. **b** Dimensionless pool depth is a ratio of pool depth to layer thickness. Figures **a** and **b** show that the lower dimensionless hatch spacing and higher dimensionless pool depth can reduce the lack of fusion. Figure **c** shows the schematic diagram of the PBF-L process. **d** Marangoni number indicates the intensity of the convective flow of liquid metal inside the molten pool. A high Marangoni number representing a vigorous flow of liquid metal can reduce the lack of fusion defects. **e** The dimensionless peak temperature is represented by the ratio of the peak temperature to the liquidus temperature of the alloy. **f** Fourier number indicates the ratio of the rate of heat dissipation to the rate of heat storage. Figures **d** and **e** show that a low Fourier number and a high dimensionless peak temperature can reduce the lack of fusion defect.

Dimensionless hatch spacing (H). Lack of fusion void formation is affected by the dimensions and geometry of the molten pool. For example, molten pool width determines the extent of fusional bonding between two neighboring hatches. If two neighboring hatches are separated by a large hatch spacing, the insufficient overlap results in gaps between the adjacent tracks. Sound fusional bonding between two hatches is affected by both the pool width and hatch spacing. Therefore, a dimensionless hatch spacing represented by the ratio of hatch spacing to pool width is considered a mechanistic variable affecting the lack of fusion (Fig. 2a). The pool width is estimated using a heat transfer and fluid flow model (Methods section). A low value of dimensionless hatch spacing due to a large molten pool width or a small hatch spacing

ensures good fusional bonding among the neighboring hatches and can reduce the lack of fusion (Fig. 2a).

Dimensionless pool depth (D). A deep molten pool increases the extent of remelting of the previously deposited layer and results in a sound fusional bonding among two successive layers. In contrast, thick layers reduce the extent of remelting. Therefore, both pool depth and layer thickness significantly affect the lack of fusion void formation. The depth of the pool is estimated using a heat and fluid flow model (Methods section). Therefore, we consider a dimensionless pool depth indicated by the ratio between the pool depth and the layer thickness as a mechanistic variable (Fig. 2b).

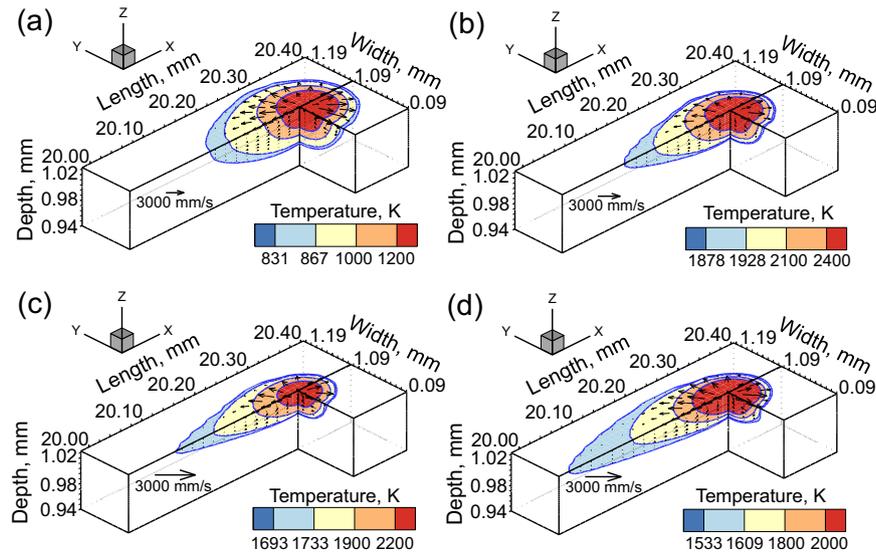


Fig. 3 Temperature and velocity fields calculated using the heat transfer and fluid flow model (see Methods section) during PBF-L of different alloys. 3D temperature and velocity fields during single track PBF-L of (a) AlSi10Mg, (b) Ti-6Al-4V, (c) SS 316, and (d) Inconel 718 using laser power of 60 W, scanning speed 1000 mm/s, hatch spacing of 110 μm and layer thickness of 30 μm . The color bands represent the ranges of temperatures corresponding to the figure legend. The solidus temperature (831 K for AlSi10Mg, 1878 K for Ti6Al4V, 1693 K for SS 316, and 1533 K for IN 718) isotherm indicates the molten pool boundary. The light blue region with the temperature range between solidus temperature and liquidus temperature represents the two-phase mushy region of the molten pool. The black arrows represent the velocity vectors inside the molten pool. The velocity vectors are radially outwards from the middle where the temperature is high to the rim of the molten pool where the temperature is the solidus temperature. The fluid flow is driven by the Marangoni stress that results from the spatial variation of surface tension due to the local temperature variation. The magnitude of the velocities can be obtained by comparing the length of the velocity vectors with the corresponding reference vectors. The molten pool is elongated in the negative X-axis direction (opposite to the scanning direction) and exhibits a teardrop shape due to the rapid scanning along the positive X-axis.

Marangoni number (Ma). The shape and dimensions of the molten pool that govern the fusional bonding among layers, are significantly affected by the convective flow of molten material inside the pool²⁶. Marangoni number (*Ma*) represents the strength of convective flow which determines the heat transfer inside the molten pool and the shape and size of the molten pool^{2,4}. A high value of *Ma* indicates a vigorous convective flow⁵⁰ that can increase the pool width and improve remelting and bonding with neighboring tracks to reduce the lack of fusion (Fig. 2d). Therefore, the Marangoni number is considered a mechanistic variable which is represented as⁵¹:

$$Ma = -\frac{d\gamma}{dT} \frac{\delta \Delta T}{\mu \alpha} \quad (1)$$

Where $d\gamma/dT$ is the derivative of surface tension with respect to the temperature, δ is a characteristic length which is considered as the pool width, ΔT indicates the difference between the solidus temperature of an alloy and the peak temperature inside the molten pool, μ and α represent the viscosity and thermal diffusivity of the alloy, respectively. Here, the width of the molten pool and the peak temperature are estimated using a heat and fluid flow model (Methods section).

Dimensionless peak temperature (T). A high peak temperature during PBF-L provides an indirect indicator of a large molten pool which is favorable for sound fusional bonding among neighboring tracks. Therefore, PBF-L conditions that provide a high peak temperature are useful to minimize the formation of lack of fusion (Fig. 2e). For each experimental condition, the peak temperature is estimated using a heat transfer and fluid flow model (Methods section). Therefore, a dimensionless peak temperature, represented by the ratio of the peak temperature inside the molten pool (T_p) to the liquidus temperature of alloy (T_l), is used as a mechanistic variable.

Fourier number (Fo). Fourier number indicates the ratio of the heat dissipation rate to the heat storage rate. A high value of the Fourier number represents a fast rate of heat dissipation and a low rate of heat accumulation both of which result in a small fusion zone⁵⁰ and increase the vulnerability to lack of fusion (Fig. 2f).

$$Fo = a/VL \quad (2)$$

where a is the thermal diffusivity, V and L are the scanning speed and pool length. The pool length is estimated using a well-tested heat transfer and fluid flow model (Methods section).

The five mechanistic variables are affected by the three-dimensional temperature and velocity fields and the shape and size of the molten pool. Figure 3 shows that the shape and dimensions of the molten pool are influenced by the alloy properties for a given set of process variables. For example, among the alloys considered, the largest molten pool is achieved for AlSi10Mg, which has the lowest density and liquidus temperature. The build of IN 718 alloy exhibits the largest two-phase mushy zone because it has a maximum temperature range between the solidus and liquidus temperatures. Therefore, the susceptibilities of different alloys to lack of fusion are different because of the difference in shape and dimensions of the molten pool.

The alloy properties, process parameters, and the five mechanistic variables computed corresponding to the one hundred and one experimental cases are supplied in the supplementary document. The computed data on the five mechanistic variables are used to find the hierarchical importance of these variables on lack of fusion as described below.

Hierarchical influence of mechanistic variables

The hierarchy of the effects of five mechanistic variables on the lack of fusion is useful to help the engineers identify the proper conditions to avoid this defect. Their importance is estimated

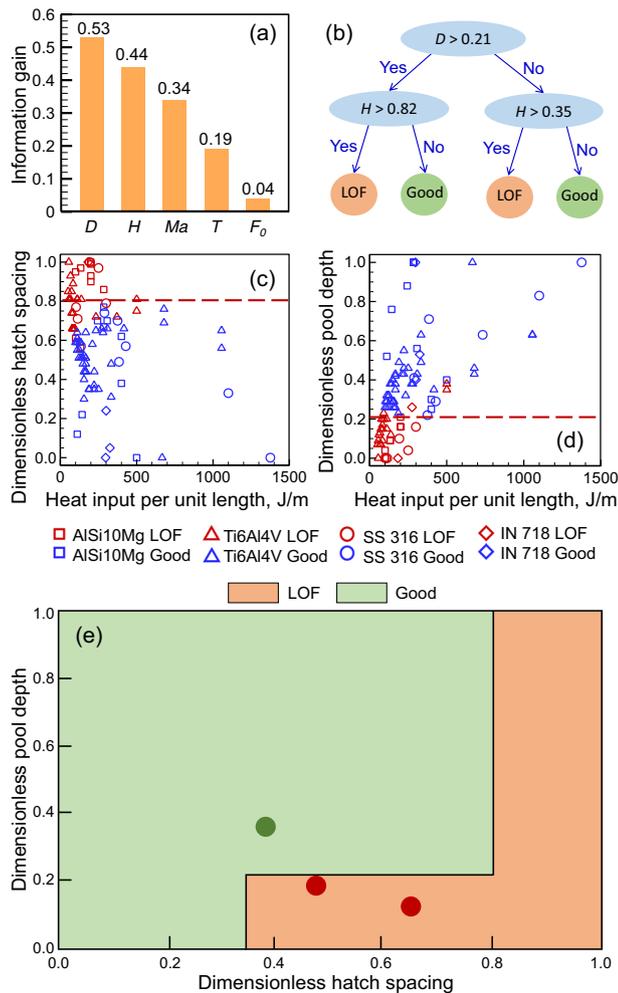


Fig. 4 Hierarchical importance of the mechanistic variables on lack of fusion and decision tree to predict the defects. **a** Hierarchical importance of the five mechanistic variables on lack of fusion defect computed using information gain. **b** A decision tree was constructed based on the 101 experimental cases to predict the formation of the lack of fusion defect. ‘LOF’ indicates the cases with lack of fusion defect and cases where lack of fusion was not observed experimentally are denoted by ‘Good’. Variations of **c** dimensionless hatch spacing and **d** dimensionless pool depth with heat input per unit length (laser power/scanning speed) for 101 experimental cases. The data for each alloy was made dimensionless by dividing by its maximum value for the alloy so that the data of all four alloys could be plotted on the same scale. **e** A process map of dimensionless hatch spacing and dimensionless pool depth indicating the regions of ‘LOF’ (with lack of fusion voids) and ‘Good’ (without lack of fusion voids) was constructed using the decision tree results. The map is tested using independent experimental data of Ti-6Al-4V parts⁵⁵ made using PBF-L. The green circle corresponds to a case without lack of fusion and the two red circles indicate two cases where voids were found.

by the information gain of the five mechanistic variables (Fig. 4a). A variable that has the highest value of information gain is the most important^{53,54}. The dimensionless hatch spacing and dimensionless pool depth are the two most influential variables. This is primarily because these two variables are the direct indicators of pool geometry which control the fusional bonding among neighboring tracks and the lack of fusion defect formation. The hierarchical importance computed by the information gain ratio and Gini index (see the supplementary document) is also the same as what is predicted by information

gain (Fig. 4a). This mechanistic variable can be used to construct a visual tool, a decision tree, to qualitatively predict the lack of fusion.

Decision tree

A decision tree is a machine learning algorithm that can categorize the five calculated mechanistic variables from the occurrence of lack of fusion. For a new set of processing parameters, the computed values of the five variables are categorized using the decision tree to forecast the lack of fusion. A decision tree is constructed to predict the lack of fusion defect (Fig. 4b). The procedure for constructing the decision tree is provided in the Methods section. The dimensionless pool depth and dimensionless hatch spacing are selected as the root node and classify node, respectively which are also the two most influential variables (Fig. 4a). When the calculated dimensionless hatch spacing and pool depth are available, a decision tree can forecast the lack of fusion with 93.3 percent accuracy as shown in the Methods section. Figure 4c and d show the distributions of the normalized values of numerically calculated dimensionless hatch spacing and dimensionless pool depth with the linear heat input (a ratio between the laser power and the laser scanning speed) based on 101 experimental cases for four alloys^{10,13,32–49}. Horizontal dashed lines in both figures approximately delineate the values of width and depth for cases with and without any lack of fusion.

Since a decision tree can qualitatively predict the formation of the defect for new processing conditions, an easy-to-use map can be constructed based on the decision tree results (Fig. 4e). The map can qualitatively predict the presence of lack of fusion which is also tested using independent experimental data of Ti-6Al-4V parts⁵⁵ made using PBF-L (Fig. 4e). We find that fully dense printed parts are obtained using the process parameter corresponding to the data point with the highest value of dimensionless pool depth and the lowest value of dimensionless hatch spacing in the region where the lack of fusion is not detected. For the printed part using the other two parameters corresponding to the two points in the region of defect formation, a lack of fusion defect is observed in both cases. Although a decision tree delivers a visual tool to forecast the lack of fusion, no quantitative relationship between the lack of fusion formation and the mechanistic variables can be obtained. A straightforward way of correlating multiple independent mechanistic variables with the formation of the lack of fusion voids is a linear regression as discussed below.

Lack of fusion index

Quantitative relations between the mechanistic variables and the lack of fusion defects can guide engineers to adjust the important variables to control the defect. Since linear regression is a simple but effective approach for this purpose, it is used to provide an equation to separate the cases with and without lack of fusion. One of the prerequisites of linear regression is the independence of the mechanistic variables. The independence of the five mechanical variables is evaluated using Pearson’s correlation coefficient (see Methods section). We found that these five variables are linearly independent of each other. The results of the linear regression are used to formulate the following lack of fusion index (LFI) containing the five calculated mechanical variables.

$$\text{LFI} = 0.473H - 0.15D - 2.61 \times 10^{-5}Ma - 0.122T + 0.306Fo + 0.569 \quad (3)$$

where H , D , Ma , T , and Fo are the calculated dimensionless hatch spacing, dimensionless pool depth, Marangoni number, dimensionless peak temperature, and Fourier number, respectively. The correlation is applicable within the ranges of values of each

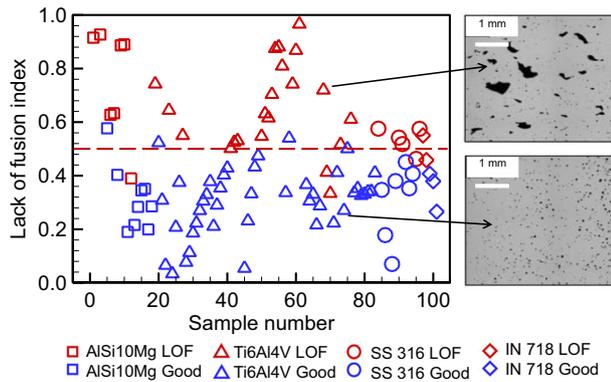


Fig. 5 Lack of fusion index from linear regression. Distribution of lack of fusion susceptibility index for 101 experimental cases. The dashed line at the index = 0.5 delineates the 'LOF' and 'Good' cases. Here, 'LOF' represents cases with defects and the cases where no defects were observed are indicated by 'Good'. The results are also tested using independent experimental data on the lack of fusion during PBF-L of AlSi10Mg alloy¹⁰. The two microstructures are adapted from ref. ¹⁰.

mechanistic variable indicated in the supplementary document. The sign of each coefficient in Eq. (3) represents how a particular mechanistic variable affects the lack of fusion defect. For example, H and Fo have positive coefficients which indicate an increase in the lack of fusion susceptibility for the higher values of these variables. In contrast, the negative coefficients of D , Ma , and T indicate a reduced lack of fusion susceptibility. The LFI can serve as a useful indicator for forecasting the formation of lack of fusion voids in printed metallic parts from the calculated values of the mechanistic variables. Figure 5 shows the calculated LFI values that correspond to the 101 experimental data sets^{10,13,32–49}. The data show that a threshold value of 0.5 for LFI can delineate the cases with and without the lack of fusion based on minimum classification error. This threshold value of 0.5 is valid for the four alloys studied here for the range of process conditions reported in the Supplementary document. When the calculated mechanical variables are available, LFI can serve as an index to predict the lack of fusion. Two sets of experimental data with or without the lack of fusion were tested using LFI and shown in Fig. 5. The validity of the procedure is observed from the two microstructures. The index LFI derived using linear regression can predict the lack of fusion defect with 90.0 percent accuracy as shown in the Methods section.

Relative susceptibility of alloys

The calculated values of LFI of various process conditions for different materials can be utilized to construct a lack of fusion susceptibility maps (Fig. 6a–d) that can be used to evaluate the relative susceptibility of alloys. These maps can visually represent the process windows for reducing the lack of fusion. The lack of fusion susceptible zone (for $LFI \geq 0.5$) is indicated by the red color, and the safe zone (for $LFI < 0.5$) is represented by the green color. These maps can be made available on the shop floor for indicating ranges of processing conditions suitable for reducing the voids. Furthermore, these maps can be used to compare the susceptibilities of different alloys to the formation of lack of fusion voids. Under the same PBF-L conditions, stainless steel 316 is the most susceptible to the formation of the lack of fusion voids among the four alloys. For the same scanning speed and laser power, the red zone for SS 316 occupies the largest area on the maps among the four alloys examined.

The aforementioned results show that wider and deeper pools obtained at a higher laser power and a slower scanning speed can

reduce the formation of lack of fusion voids. However, the conditions that can reduce the lack of fusion may cause other defects. For example, large pools at a high laser power and slow scanning speed shrink more during the solidification and may result in distortion. Apart from the differences in the molten pool size, the size of the mushy zones for different alloys also vary widely (Fig. 3). The solidification range that affects the shape and size of the mushy zone largely contributes to the susceptibility to solidification cracking. The vulnerability to lack of fusion is not largely affected by the mushy zone size but by the shape and size of the entire fusion zone. Therefore, the selection of processing conditions to reduce the lack of fusion voids should be carefully done so that it does not result in other defects such as solidification cracking, balling, and distortion.

In summary, we find that augmented machine learning with human intelligence can achieve superior printed parts by reducing the lack of fusion voids. Five mechanistic variables, dimensionless pool depth (pool depth/layer thickness), dimensionless hatch spacing (hatch spacing/pool width), dimensionless temperature (peak temperature/liquidus temperature), Marangoni number, and Fourier number are found to be influential in determining the susceptibility to lack of fusion in PBF-L parts. Two machine learning algorithms, decision tree and linear regression forecast the lack of fusion with 93% and 90% accuracy, respectively. The proposed lack of fusion index has a threshold value of 0.5 for delineating the parts with and without lack of fusion. The index is also used to generate easy-to-use lack of fusion susceptibility maps that showed that for the processing conditions investigated here, SS316 was the most susceptible alloy and AlSi10Mg alloy was the least susceptible alloy to lack of fusion. The same hierarchical importance of the mechanistic variables on the lack of fusion is obtained by using three feature selection indexes, information gain, information gain ratio, and Gini index. The dimensionless depth of the pool and dimensionless hatch spacing are the two most influential variables because these two variables are the direct indicators of pool geometry that control the fusional bonding among neighboring tracks and layers.

Considering that the knowledge base of manufacturing processes such as welding and casting evolved largely by empirical testing, this example shows how 3D printing can mature using the emerging digital tools. If the application of this procedure is verified for solving other problems of 3D printing, it will mature following a path that is rapid, scientifically based, and cost-effective. The traditional trial and error tests will be replaced by the advantages achievable from the application of mechanistic modeling and augmented machine learning. The methodology can reduce cost, improve the quality of the printed metallic parts, facilitate the printing of new alloys, and is equally attractive to solve important problems of other manufacturing processes.

METHODS

Data set for machine learning

The method used here is based on one hundred and one data for an aluminum alloy^{10,32–37}, AlSi10Mg, titanium alloy^{13,38–44}, Ti6Al4V, stainless steel^{45–47}, SS316, and a nickel alloy^{48,49}, Inconel 718. We indicated the presence or the absence of lack of fusion defects as "1" and "0", respectively. Among all one hundred and one data points, 38 cases were with the lack of fusion defects, and defects were not observed in the remaining 63 cases. Five mechanistic variables corresponding to all one hundred and one data points were calculated using a mechanistic model. For the two groups of cases ("0" and "1"), the data were randomly separated into three sub-sets. In machine learning analysis of decision tree and linear regression, 60% of the data were used for training, 10% for validation, and 30% for testing. The materials properties and process parameters were used to calculate the five mechanistic variables from a heat transfer and fluid flow model and the calculated results are given in the supplementary document.

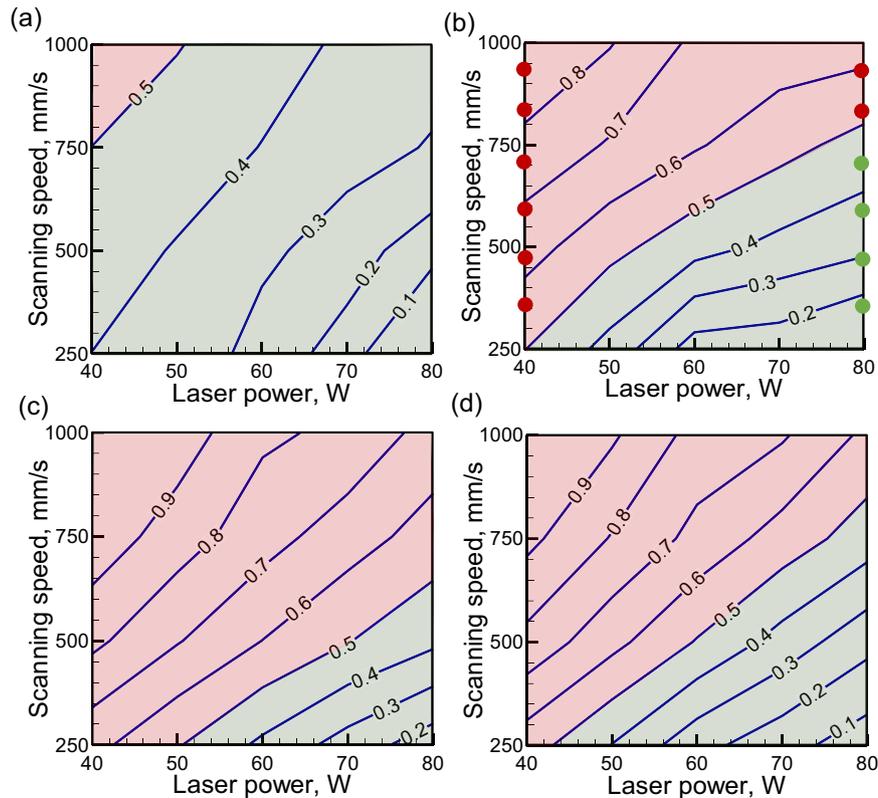


Fig. 6 Lack of fusion susceptibility maps. Lack of fusion susceptibility maps of different alloys, (a) AlSi10Mg, (b) Ti6Al4V, (c) SS316, and (d) IN 718. The red zones indicate a lack of fusion susceptible zones, and the green zones are safe regions. The values on the lines represent the values of LFI . The beam radius was $50\ \mu\text{m}$, the layer thickness was $30\ \mu\text{m}$, and the hatch spacing is $100\ \mu\text{m}$. The red dots and green squares are data points taken from the literature¹³ for validation. The green circles correspond to the cases without lack of fusion and the red circles indicate the cases where voids were found.

Heat transfer and fluid flow model of powder bed fusion

The transient, 3D heat transfer and fluid flow model of the PBF-L process considers process variables and alloy properties as input variables and supplies transient 3D temperature and velocity distribution and the shape and size of the molten pool as output. The details of this model are available in our previous works^{25,26,56} and are not repeated here. The calculation domain contains the substrate, powder bed, hatches and layers, and shielding gas. The calculations are performed by iteratively solving the equations of mass, momentum, and energy conservations. The thermophysical properties^{4,26} of Ti-6Al-4V, AlSi10Mg, SS 316, and Inconel 718 used for the calculations are presented in the supplementary document.

The model can simulate molten pool shape and dimensions for multi-layer multi-hatch deposits. For example, Fig. 7a–c show the computed shapes and dimensions of the molten pools at transverse sections of five hatch, three-layer, SS 316 builds at various processing conditions. Lack of fusion can originate in the unmelted region between the neighboring tracks (Fig. 7a). However, a higher laser power (Fig. 7b) or a lower scanning speed (Fig. 7c) results in a bigger pool and ensures complete fusional bonding with sufficient overlap of the neighboring tracks to decrease the lack of fusion. The calculated shape and dimensions of the molten pool are validated as discussed in the supplementary document.

Implementation of machine learning algorithms

Three commonly used feature selection algorithms⁵⁷, Iterative Dichotomiser (ID3), Classification and regression tree (CART), and a variant of ID3 commonly referred to as C4.5 were used to rank the importance of five mechanistic variables on the lack of fusion. Their importance was estimated by the ranking of the feature selection indexes using three algorithms. The feature selection indexes, information gain, and information gain ratio corresponding to ID3 and C4.5 were computed from the entropy. A variable with high values of information gain ratio and information gain indicated its high importance. The calculations of the Gini

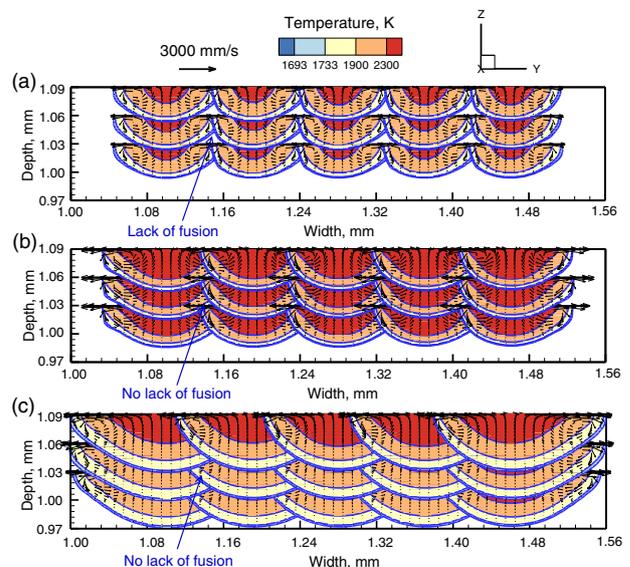


Fig. 7 Transverse view of fusion zones. The transverse views indicate lack of fusion during 3 layers 5 hatches PBF-L of (a) 60 W power 1000 mm/s speed, (b) 100 W power 1000 mm/s speed, and (c) 60 W power 250 mm/s speed.

index using the CART algorithm were based on the prediction of impurity^{54,57}. A relatively more important variable had a low Gini index.

A decision tree^{3,58} was constructed using the information gain values of the computed mechanistic variables. The mechanistic variable with the

highest information gain was selected as the root node of the tree. The tree contained multiple nodes and each node represented a mechanistic variable. A node classified the data by comparing the value of the data and a classifier value. The classifier value for a node was selected so that the node can classify the data with maximum accuracy^{54,57,58}. The calculation continues until all data can be classified into two classes.

A linear regression equation is used to connect the five mechanistic variables and lack of fusion. The coefficients of variables are optimized using a genetic algorithm⁵⁹ to achieve a least square error in the linear fitting between the lack of fusion results and the five mechanistic variables used in this work.

Evaluation of the mutual independence of mechanistic variables

The mutual independence of five mechanistic variables was examined using Pearson's correlation coefficient. The Pearson's coefficient (ρ) is estimated using the following equation:

$$\rho(x_1, x_2) = \frac{COV(x_1, x_2)}{\sigma_{x_1} \sigma_{x_2}} \quad (4)$$

where x_1 and x_2 are any two mechanistic variables, σ indicates the standard deviation of variables, and COV indicates a function to estimate covariance between two variables. Equation (4) can provide the mutual independence of each two mechanistic variables. The value of the Pearson's coefficient (ρ) varies between -1 and 1 . -1 and 1 indicate a strong negative and positive correlation, respectively. The five mechanical variables are dimensionless hatch spacing (H), dimensionless pool depth (D), Marangoni number (Ma), dimensionless peak temperature (T), and Fourier number (Fo). The calculated Pearson's coefficients can be used to construct a Pearson matrix, which is provided in Fig. 8. The absolute values of ρ less than 0.5 indicate no interdependence of the mechanistic variables.

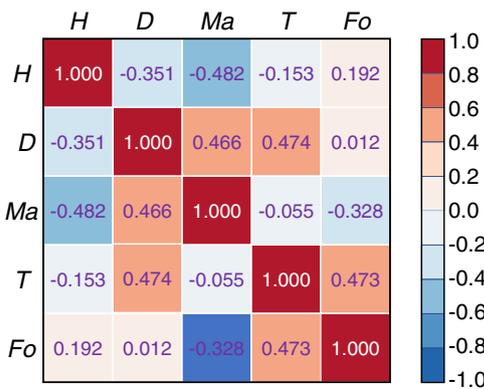


Fig. 8 The Pearson matrix. The matrix is used to show the correlation among the five mechanistic variables.

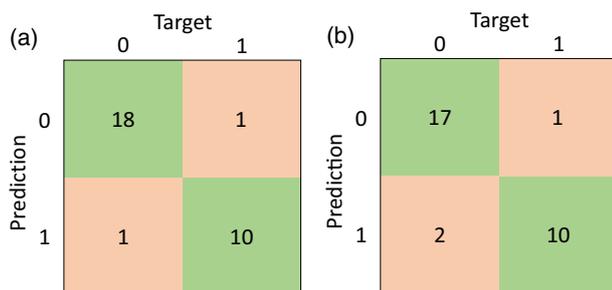


Fig. 9 Confusion matrices. Confusion matrices are shown for (a) decision tree and (b) linear regression. Here, '0' and '1' represent the cases without and with cracks, respectively.

Confusion matrices and calculations of accuracy

Confusion matrices were constructed to provide a visual description of the prediction ability of different machine learning algorithms used here to predict the lack of fusion⁵⁷. The method of constructing confusion matrices and their interpretations are described in the supplementary document. Figure 9a and b illustrate the confusion matrices for the decision tree and linear regression-based prediction of defects, respectively. The calculation methods of the accuracy of the above two machine learning methods from the data in confusion matrices are also discussed in detail in the supplementary document. The computed accuracies for testing in predicting the lack of fusion defects using a decision tree and linear regression are 93.3% and 90.0%, respectively.

DATA AVAILABILITY

Data used in this paper are available in the supplementary document.

CODE AVAILABILITY

The code used in this research cannot be shared due to the technical limitations.

Received: 2 March 2022; Accepted: 3 August 2022;

Published online: 23 August 2022

REFERENCES

1. DebRoy, T. et al. Scientific, technological and economic issues in metal printing and their solutions. *Nat. Mater.* **18**, 1026–1032 (2019).
2. DebRoy, T. et al. Additive manufacturing of metallic components – process, structure and properties. *Prog. Mater. Sci.* **92**, 112–224 (2018).
3. DebRoy, T. et al. Metallurgy, mechanistic models and machine learning in metal printing. *Nat. Rev. Mater.* **6**, 48–68 (2021).
4. Wei, H. L. et al. Mechanistic models for additive manufacturing of metallic components. *Prog. Mater. Sci.* **116**, 100703 (2021).
5. DebRoy, T., & Bhadeshia, H. K. D. H. *Innovations in everyday engineering materials* 1st edn (Springer, 2021).
6. Gu, D. D., Meiners, W., Wissenbach, K. & Poprawe, R. Laser additive manufacturing of metallic components: materials, processes and mechanisms. *Int. Mater. Rev.* **57**, 133–164 (2012).
7. Gu, D. D. et al. Material-structure-performance integrated laser-metal additive manufacturing. *Science* **372**, 6545 (2021).
8. Bose, S., Ke, D., Sahasrabudhe, H. & Bandyopadhyay, A. Additive manufacturing of biomaterials. *Prog. Mater. Sci.* **93**, 45–111 (2018).
9. Oliveira, J. P., Santos, T. G. & Miranda, R. M. Revisiting fundamental welding concepts to improve additive manufacturing: From theory to practice. *Prog. Mater. Sci.* **107**, 100590 (2020).
10. Aboulkhair, N. T., Everitt, N. M., Ashcroft, I. & Tuck, C. Reducing porosity in AlSi10Mg parts processed by selective laser melting. *Addit. Manuf.* **1**, 77–86 (2014).
11. Qiu, C. et al. On the role of melt flow into the surface structure and porosity development during selective laser melting. *Acta Mater.* **96**, 72–79 (2015).
12. Darvish, K., Chen, Z. W. & Pasang, T. Reducing lack of fusion during selective laser melting of CoCrMo alloy: Effect of laser power on geometrical features of tracks. *Mater. Des.* **112**, 357–366 (2016).
13. Gong, H., Rafi, K., Gu, H., Starr, T. & Stucker, B. Analysis of defect generation in Ti-6Al-4V parts made using powder bed fusion additive manufacturing processes. *Addit. Manuf.* **1**, 87–98 (2014).
14. Tang, M., Pistorius, P. C. & Beuth, J. L. Prediction of lack-of-fusion porosity for powder bed fusion. *Addit. Manuf.* **14**, 39–48 (2017).
15. Mukherjee, T., Zuback, J. S., De, A. & DebRoy, T. Printability of alloys for additive manufacturing. *Sci. Rep.* **6**, 1–8 (2016).
16. Bruna-Rosso, C., Demir, A. G. & Previtali, B. Selective laser melting finite element modeling: Validation with high-speed imaging and lack of fusion defects prediction. *Mater. Des.* **156**, 143–153 (2018).
17. Bayat, M., Mohanty, S. & Hattel, J. H. Multiphysics modelling of lack-of-fusion voids formation and evolution in IN718 made by multi-track/multi-layer L-PBF. *Int. J. Heat. Mass. Transf.* **139**, 95–114 (2019).
18. Kappes, B., Moorthy, S., Drake, D., Geerlings, H. & Stebner, A. Machine learning to optimize additive manufacturing parameters for laser powder bed fusion of inconel 718. In *Proceedings of the 9th International Symposium on Superalloy 718 & Derivatives: Energy, Aerospace, and Industrial Applications* 595–610 (Springer, Cham, 2018).
19. Zhang, B., Liu, S. & Shin, Y. C. In-Process monitoring of porosity during laser additive manufacturing process. *Addit. Manuf.* **28**, 497–505 (2019).

20. Khanzadeh, M., Chowdhury, S., Marufuzzaman, M., Tschopp, M. A. & Bian, L. Porosity prediction: Supervised-learning of thermal history for direct laser deposition. *J. Manuf. Syst.* **47**, 69–82 (2018).
21. Tapia, G., Elwany, A. H. & Sang, H. Prediction of porosity in metal-based additive manufacturing using spatial Gaussian process models. *Addit. Manuf.* **12**, 282–290 (2016).
22. Sizova, I. et al. A study on hot-working as alternative post-processing method for titanium aluminides built by laser powder bed fusion and electron beam melting. *J. Mater. Process. Technol.* **291**, 117024 (2021).
23. Plessis, A. D. & Macdonald, E. Hot isostatic pressing in metal additive manufacturing: X-ray tomography reveals details of pore closure. *Addit. Manuf.* **34**, 101191 (2020).
24. Gao, W. et al. The status, challenges, and future of additive manufacturing in engineering. *Comput. Aided Des.* **69**, 65–89 (2015).
25. Mukherjee, T. & DebRoy, T. Mitigation of lack of fusion defects in powder bed fusion additive manufacturing. *J. Manuf. Process.* **36**, 442–449 (2018).
26. Mukherjee, T., Wei, H. L., De, A. & DebRoy, T. Heat and fluid flow in additive manufacturing – Part II: Powder bed fusion of stainless steel, and titanium, nickel and aluminum base alloys. *Comput. Mater. Sci.* **150**, 369–380 (2018).
27. Antony, J. *Design of experiments for engineers and scientists* 2nd edn (Elsevier, 2014).
28. Bird, R., Stewart, W. & Lightfoot, E. *Transport phenomena* Revised 2nd edn (John Wiley & Sons, 2006).
29. Gennatas, E. D. et al. Expert-augmented machine learning. *Proc. Natl Acad. Sci.* **117**, 4571–4577 (2020).
30. Bankapalli, Y. S. & Wong, H. Y. TCAD augmented machine learning for semiconductor device failure troubleshooting and reverse engineering. In *2019 International Conference on Simulation of Semiconductor Processes and Devices*, 1–4, (IEEE, 2019).
31. Dhillon, H. et al. TCAD-augmented machine learning with and without domain expertise. *IEEE Trans. Electron Devices* **68**, 5498–5503 (2021).
32. Thijs, L., Kempen, K., Kruth, J. P. & van Humbeeck, J. Fine-structured aluminium products with controllable texture by selective laser melting of pre-alloyed AlSi10Mg powder. *Acta Mater.* **61**, 1809–1819 (2013).
33. Tang, M., Pistorius, P. C., Narra, S. & Beuth, J. L. Rapid solidification: selective laser melting of AlSi10Mg. *JOM* **68**, 960–966 (2016).
34. Manfredi, D. et al. Additive manufacturing of Al alloys and aluminium matrix composites (AMCs). In *Light metal alloys applications* 11, 3–34 (ed. Monteiro W. A.) (InTech, 2014).
35. Li, W. et al. Effect of heat treatment on AlSi10Mg alloy fabricated by selective laser melting: microstructure evolution, mechanical properties and fracture mechanism. *Mater. Sci. Eng. A Struct. Mater.* **663**, 116–125 (2016).
36. Xing, X. et al. Modification of residual stresses in laser additive manufactured AlSi10Mg specimens using an ultrasonic peening technique. *Materials* **12**, 455 (2019).
37. Brandl, E., Heckenberger, U., Holzinger, V. & Buchbinder, D. Additive manufactured AlSi10Mg samples using Selective Laser Melting (SLM): microstructure, high cycle fatigue, and fracture behavior. *Mater. Des.* **34**, 159–169 (2012).
38. Cunningham, R., Narra, S. P., Montgomery, C., Beuth, J. & Rollett, A. D. Synchrotron-based X-ray microtomography characterization of the effect of processing variables on porosity formation in laser powder-bed additive manufacturing of Ti-6Al-4V. *JOM* **69**, 479–484 (2017).
39. Promopattum, P. & Yao, S. C. Analytical evaluation of defect generation for selective laser melting of metals. *Int. J. Adv. Manuf. Technol.* **103**, 1185–1198 (2019).
40. Vandenbroucke, B. & Kruth, J. P. Selective laser melting of biocompatible metals for rapid manufacturing of medical parts. *Rapid Prototyp. J.* **13**, 196–203 (2007).
41. Thijs, L., Verhaeghe, F., Craeghs, T., Humbeeck, J. van & Kruth, J. P. A study of the microstructural evolution during selective laser melting of Ti-6Al-4V. *Acta Mater.* **58**, 3303–3312 (2010).
42. Parry, L., Ashcroft, I. A. & Wildman, R. D. Understanding the effect of laser scan strategy on residual stress in selective laser melting through thermo-mechanical simulation. *Addit. Manuf.* **12**, 1–15 (2016).
43. Gu, H. et al. Effects of powder variation on the microstructure and tensile strength of Ti6Al4V parts fabricated by selective laser melting. In *2014 International Solid Freeform Fabrication Symposium* (University of Texas at Austin, 2014).
44. Yadroitsev, I., Krakhmalev, P. & Yadroitsava, I. Selective laser melting of Ti6Al4V alloy for biomedical applications: temperature monitoring and microstructural evolution. *J. Alloy. Compd.* **583**, 404–409 (2014).
45. de Terris, T. et al. Optimization and comparison of porosity rate measurement methods of Selective Laser Melted metallic parts. *Addit. Manuf.* **28**, 802–813 (2019).
46. Ahmadi, A. et al. Effect of manufacturing parameters on mechanical properties of 316L stainless steel parts fabricated by selective laser melting: a computational framework. *Mater. Des.* **112**, 328–338 (2016).
47. Wang, D., Song, C., Yang, Y. & Bai, Y. Investigation of crystal growth mechanism during selective laser melting and mechanical property characterization of 316L stainless steel parts. *Mater. Des.* **100**, 291–299 (2016).
48. Jia, Q. & Gu, D. Selective laser melting additive manufacturing of Inconel 718 superalloy parts: Densification, microstructure and properties. *J. Alloy. Compd.* **585**, 713–721 (2014).
49. Kelley, P. F., Saigal, A., Vlahakis, J. K. & Carter, A. Tensile and fatigue behavior of direct metal laser sintered (DMLS) Inconel 718. In *ASME International Mechanical Engineering Congress and Exposition* 57359, p. V02AT02A001 (American Society of Mechanical Engineers, 2015).
50. Mukherjee, T., Manvatkar, V., De, A. & DebRoy, T. Dimensionless numbers in additive manufacturing. *J. Appl. Phys.* **121**, 064904 (2017).
51. Großmann, A. et al. Dimensionless process development for lattice structure design in laser powder bed fusion. *Mater. Des.* **194**, 108952 (2020).
52. Williams, R. J. et al. In situ thermography for laser powder bed fusion: effects of layer temperature on porosity, microstructure and mechanical properties. *Addit. Manuf.* **30**, 100880 (2019).
53. Du, Y., Mukherjee, T. & DebRoy, T. Conditions for void formation in friction stir welding from machine learning. *NPJ Comput. Mater.* **5**, 1–8 (2019).
54. Du, Y., Mukherjee, T., Mitra, P. & DebRoy, T. Machine learning based hierarchy of causative variables for tool failure in friction stir welding. *Acta Mater.* **192**, 67–77 (2020).
55. Dilip, J. J. S. et al. Influence of processing parameters on the evolution of melt pool, porosity, and microstructures in Ti-6Al-4V alloy parts fabricated by selective laser melting. *Prog. Addit. Manuf.* **2**, 157–167 (2017).
56. Mukherjee, T., Wei, H. L., De, A. & DebRoy, T. Heat and fluid flow in additive manufacturing—Part I: modeling of powder bed fusion. *Comput. Mater. Sci.* **150**, 304–313 (2018).
57. Jordan, M. I. & Mitchell, T. M. Machine learning: trends, perspectives, and prospects. *Science* **349**, 255–260 (2015).
58. Rokach, L. & Maimon, O. *Data mining with decision trees: theory and applications* 2nd edn (World scientific, 2007).
59. Rai, R. & DebRoy, T. Tailoring weld geometry during keyhole mode laser welding using a genetic algorithm and a heat transfer model. *J. Phys. D: Appl. Phys.* **39**, 1257 (2006).
60. Xia, M. et al. Porosity evolution and its thermodynamic mechanism of randomly packed powder-bed during selective laser melting of Inconel 718 alloy. *Int. J. Mach. Tools Manuf.* **116**, 96–106 (2017).
61. Mukherjee, T. & DebRoy, T. A digital twin for rapid qualification of 3D printed metallic components. *Appl. Mater. Today* **14**, 59–65 (2019).

ACKNOWLEDGEMENTS

M.J. thanks the support provided by the startup fund for new faculty from Harbin Institute of Technology. Funding for this research was also provided by the 2020 industryXchange Multidisciplinary research grant of the Pennsylvania State University, University Park, PA.

AUTHOR CONTRIBUTIONS

M.J. collected the data and performed the computations and data analysis. M.J., T.M. and Y.D. analyzed the data and prepared the display items. T.D. supervised the work. All authors contributed in writing the 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/s41524-022-00866-9>.

Correspondence and requests for materials should be addressed to Meng Jiang or Tarasankar DebRoy.

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) 2022