Contents lists available at ScienceDirect

Chemical Engineering Research and Design

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

Finite element modeling of direct metal laser solidification process: Sensor data replication and use in defect detection and data reduction via machine learning

Yi Ming Ren[a], Yichi Zhang[a], Yangyao Ding[a], Tao Liu[b], Cody S. Lough[b], Ming C. Leu[b], Edward C. Kinzel[c], Panagiotis D. Christofides[a,d,*]

[a] Department of Chemical and Biomolecular Engineering, University of California, Los Angeles, CA 90095-1592, USA
[b] Department of Mechanical and Aerospace Engineering, Missouri University of Science and Technology, Rolla, MO 65409, USA
[c] Department of Aerospace and Mechanical Engineering, University of Notre Dame, Notre Dame, IN 46556, USA
[d] Department of Electrical and Computer Engineering, University of California, Los Angeles, CA 90095-1592, USA

ARTICLE INFO

Article history:
Received 23 April 2021
Received in revised form 14 May 2021
Accepted 17 May 2021
Available online 24 May 2021

Keywords:
Additive manufacturing
Direct metal laser solidification
Computational fluid dynamics
Sensor analytics
Machine learning
Data reduction

ABSTRACT

In the additive manufacturing (AM) industry, the direct metal laser solidification (DMLS) process has received significant research interest because of its ultra-high precision and geometry variability. However, many possible process errors can happen during the printing process and in-situ quality control is quite challenging because of the lack of appropriate monitoring and analysis capabilities. Thus, in this work, a part-scale finite element method (FEM) model is developed to investigate the heat transfer behavior of the DMLS process and to extract experimentally relevant thermal features. Specifically, a microscopic and a mesoscale sub-model are initially developed to describe powder properties and the laser behaviors, respectively, and their outputs are directly incorporated in the part-scale FEM model. The FEM model-generated data are then processed to replicate the long-wave infrared (LWIR) camera sensor outputs. Finally, both the experimental and FEM model-generated images are used to train machine learning algorithms for in-situ defect detection applications. In addition, we are endeavoring to look into a convolutional neural network utilizing such realistic thermal features from LWIR images with and without a variety of simulated data augmentation to examine its contribution to the CNN training. The thermal feature extraction and the machine learning algorithm are then utilized for the data reduction purpose. Also, a transmission strategy is demonstrated to filter out a significant amount of redundant data while maintaining high model prediction quality.

1. Introduction

Additive manufacturing (AM) has been developed for various materials like polylactic acid (PLA) plastic, polymers, and metals (Gibson et al., 2014). In particular, metal 3D printing has been growing significantly each year expanding at a compound annual growth rate of 27.8% and has projected to reach USD 5.51 billion by 2027. Two metal 3D printing methods

* *Corresponding author at*: Department of Chemical and Biomolecular Engineering, University of California, Los Angeles, CA 90095-1592, USA.
E-mail address: pdc@seas.ucla.edu (P.D. Christofides).
https://doi.org/10.1016/j.cherd.2021.05.018

are popular in the industry: direct metal laser solidification (DMLS) (Frazier, 2014) and directed energy deposition (DED). Among the two, DMLS is the more suitable option for applications in the aerospace (Liu et al., 2017) and medical (Singh and Ramakrishna, 2017) industry, due to DMLS's ability to achieve greater dimensional accuracy and produce net shape parts. The DMLS process utilizes a laser source to melt powders locally and allows them to resolidify to the desired shape. First, the desired build geometry along with the necessary support structure needs to be constructed with computer-aided design (CAD) software. Then, a printing recipe will be developed according to the geometry, which specifies the laser parameters and scanning strategy. In the actual printing process, a layer of powder is spread on the printing platform by the recoater, and the printing is executed according to the operating recipe.

Because of the thermal and mechanical complexity, process failures are often encountered in DMLS, which may lead to part defects or even detrimental damage to the printing platform. For example, unexpected defocusing of the laser beam will form cold and hot zones on the powder bed. These heating abnormalities may induce thermal and mechanical stresses on the build part and eventually cause physical problems, such as keyholing or lack of fusion (Grünberger and Domröse, 2014). The generated thermal stresses may cause part distortion, which interacts with the recoater to change the layer thickness and ultimately interferes with the recoater on a layer-to-layer timescale (Gong et al., 2014). Thus, it is important to identify these kinds of problems and control the process parameters to suppress them when they are emerging during layer-to-layer printing (Yoon et al., 2016; Yang et al., 2016; McLouth et al., 2020; Li et al., 2020; Khalil et al., 2020). To monitor the printing process and to facilitate engineers to make crucial operating recipe updates, in-situ sensors are developed to investigate the printing process information, (Craeghs et al., 2010; Khanzadeh et al., 2019), including the Optical Tomography (OT) sensor from the EOSTATE suite, (Scott, 2017), the IR sensors from FLIR Systems (Lough et al., 2020), and the physical powder bed images from P. Basler visible-light cameras (Scime et al., 2020).

Nevertheless, due to the limitation of the platform setup, not all types of monitoring can be included in the printing platform. Thus, a significant part of the monitoring information might be lost or hidden beyond what can be captured visually. If everything needs to be experimentally determined, the cost and time of experiments may be quite high. Also, it is difficult to characterize the observed relationship between process error and observed heat map for different build geometries and materials (Delgado et al., 2012). Thus, first-principles-based modeling and analysis methods are of great importance for DMLS processes to offer more physical insights into the process and to enhance the existing process monitoring (Wang et al., 2020). Among other recent works, Arısoy et al. (2019) used a multi-physics model to understand the relationship of the melt pool and the microstructure mechanical property. Dong et al. (2019) looked into the impact of hatch spacing, and Ramos et al. (2019) delved into the influence of scanning strategy on melt pool dimension. Nevertheless, part-scale information is important for defect detection, as a lot of local problems are associated with regional geometry and printing history. Despite the existing modeling efforts, very few efforts have been able to successfully explore a reasonable parametric analysis of an industrial geometry because of the computational cost. Thus, such a part-scale model is deemed to be valuable for actual process-relevant analysis and to provide critical insights and information to couple with the current monitoring technology. Motivated by these considerations, in Section 2, we endeavor to construct an accurate yet computationally efficient part-scale model with the aid of small micro and mesoscale sub-models tests. In addition, we successfully extract experimentally relevant thermal features that are deemed useful for determining problematic heating. It is noted that we calibrate and compare our model and results with respect to the experimental results.

In addition, the proliferation of sensor data for AM processes requires a smart way of integrating to help process engineers analyze sensor images and efficiently filter monitoring data for transport and storage. Modern machine learning and data processing algorithms are often implemented. These algorithms integrate the functionality of automated image processing, transferring, and analytics. In particular, image analytics for printing error detection has been a really important aspect. Zhang et al. (2018) utilized the support vector machine (SVM) technique to extract plume and spatter information out of powder bed images and evaluate their impact. Scime and Beuth (2019) looked into error feature extraction using scale-invariant feature transform (SIFT) and then related abnormality in melt pool to printing defects using histogram of oriented gradients (HOG) method. In their later work, a convolutional neural network (CNN) was applied to further analyze the classification of physical disturbances in powder bed images including recoater jams and super-elevation. Additionally, Yuan et al. (2019) investigated a semi-supervised learning algorithm based on the temporal ensemble method, in order to bootstrap the CNN training with limited sensor images. In our previous study, we have investigated a unique error detection direction in the IR outputs with purely simulated images (Ren et al., 2020), and successfully demonstrated the possibility of automated detection of errors from the thermal features. In this work, we are endeavoring to extend the training to realistic IR images from experiments and examine the benefit of CNN training with and without simulated data augmentation, as explained in Section 3. In addition, one of the current challenges in large-scale DMLS manufacturing is the massive amount of data generated during the process monitoring. Thanks to the advancement of optical and sensor technology, high-quality cameras are installed on AM machines. While this advancement helps the detection of defects through the generation of more detailed images, the increased image quality has also made data storage a challenge in monitoring DMLS processes. Therefore, the thermal feature extraction and the machine learning algorithm developed are utilized for data reduction purposes. Also, the proposed transmission strategy is demonstrated in Section 4 to filter a significant amount of redundant data while maintaining model prediction accuracy.

2. Finite-element method (FEM) modeling for DMLS

As introduced in the previous section, there has not been an existing part-scale simulation model for the DMLS process. The major blocking factors include the high computational demand and the long computational time. Thus, we first review the validity of a previously developed mesoscale heat transfer model in Ren et al. (2020). Then, to reduce the computational demand of the large-scale model, complexity reduction strategies are investigated with microscopic

models, including a Gaussian heat source reduction to an equivalent uniform heat source, and a radiation simplification using a comprehensive 1-D transport model. Next, we associate these two model outputs to our heat transport modeling framework for experimental build part investigation and thermal feature analysis. In addition, parametric analyses for various operating parameters, geometric variations, and disturbances are carried out using cluster computation. The heat transfer module from COMSOL Multiphysics is used to capture the major thermal characteristics of the DMLS process.

2.1. Mesoscale FEM model setup

To investigate the heat transport on the mesoscale, a detailed model has been constructed with a small-sized simple build geometry in Ren et al. (2020), as shown in Fig. 1(a). Thermal boundary condition assumptions are considered for the mesoscale FEM modeling. In particular, the top surface consists of a powder layer, where the laser is applied. A Gaussian laser source is considered, as well as convection of the ambient gas and the radiation. In contrast, no heat loss is assumed for the remaining surfaces, the substrate, and the peripheral sides, shown in Fig. 1(b). The time-dependent heat transport is dictated by the following equation with the assumption of no convection in the melt pool:

$$\frac{\partial(\rho(T)C_p(T)T)}{\partial t} - \nabla(k(T)T) = 0 \tag{1}$$

where ρ is the density of the material, where k is the thermal conductivity, C_p is the specific heat capacity, and T is the temperature. The top boundary condition consists of convection, radiation, and laser heating:

$$-\frac{\partial(k(T)T)}{\partial z}(x, y, z = 0) = -\frac{2AP}{\pi\omega^2}e^{-2r^2/\omega^2} + h(T - T_b) + \varepsilon\sigma_0(T^4 - T_b^4) \tag{2}$$

where P is the laser power, A is the absorptivity, ω is the laser beam radius, and r is the radial distance, h is the heat transfer coefficient, T_b is the ambient gas temperature, ε is the emissivity, and σ_0 is the Stefan's constant. For the other surfaces, no heat loss is assumed, as shown in the following equation:

$$\frac{\partial(k(T)T)}{\partial x}(x = (0, l), y, z) = 0, \quad \frac{\partial(k(T)T)}{\partial y}(x, y = (0, w), z) = 0, \quad \frac{\partial(k(T)T)}{\partial z}(x, y, z = h) = 0 \tag{3}$$

where w, l, h are the width, length, and height of the build platform.

Three different phases of alloy are considered which include powder, liquid, and re-solidified metal. The powder properties change irreversibly once the powder has been melted. A transitional zone of solid or liquid exists between solidus temperature, T_s and liquidus temperature, T_l. The thermal conductivity k, specific heat capacity C_p, and the density ρ of the material, are governed by the latent heat of fusion L_f and the ratio of each phase θ:

$$C_p(T) = \frac{L_f}{T_l - T_s} + \theta_s C_{p,s}(T_s) + \theta_l C_{p,l}(T_l) \quad \text{if} \quad T_s < T < T_l \tag{4}$$

where subscripts s and l refer to the solid and liquid phases, respectively. In addition, the powder phase thermal properties k_{powder} is defined in terms of the porosity, ϕ:

$$k_{powder} = k_{solid}(1 - \phi(T))^4 \tag{5}$$

which is dependent on temperature:

$$\phi(T) = \begin{cases} \phi_0 & \text{if} \quad T \le T_s \\ 0 & \text{if} \quad T \ge T_l \\ \dfrac{\phi_0}{T_s - T_l}(T - T_l) & \text{if} \quad T_s < T < T_l \end{cases} \tag{6}$$

The key thermodynamic parameters and their functional forms are shown in Ren et al. (2020) and in Fig. 2 (Romano et al., 2016). The resulting melt pool dynamics are compared with that of the experimental works from Sadowski et al. (2016). It is demonstrated in Ren et al. (2020) that the two results deviate less than 10%. Also, the model confirms the analytical estimate from the Rosenthal equation (Rosenthal, 1941), which is investigated in detail in Ren et al. (2020).

2.2. Microscale investigation for computational demand reduction

Despite the accuracy provided by the mesoscale simulation, it is computationally heavy to directly incorporate the developed model into a larger-scale FEM model. The two aspects that especially require large quantity and high-quality meshes are the Gaussian laser source and the radiation. Thus, we are looking for valid simplification to efficiently characterize these two phenomena without losing model fidelity.

When looking at a model that incorporates the entire build part, the focus range is around 0.25–0.65 mm in radius, as prescribed by the pixel dimension of the LWIR camera, which will be explained in further detail in Section 2.3.1. Thus, the exact heat map around the laser focus does not need to be precisely reproduced as in the Gaussian type of laser simulation. Instead, a uniform heat source is considered, which assumes a constant heat flux inside the beam radius, and zero heat flux outside, as shown in the following equation:

$$q(x, y, t) = \begin{cases} \dfrac{2P}{\pi R^2} & \text{if} \quad r \le R \\ 0 & \text{if} \quad r > R \end{cases} \tag{7}$$

where q is the heat flux simulating the uniform laser distribution, P is the laser power, R is the laser diameter, r is the location of the laser in relation to the x and y coordinates, the velocity of the laser movement, v, and time, t. A temperature comparison is made between the two kinds of laser sources under the same power. The maximum temperature difference is less than 3% for a moving laser, and more importantly, a pixel-wise average temperature difference is less than 4%. The reciprocal of step size is around 5×10^4 with and 8×10^4 without this simplification.

Thus, it can be concluded that there is no significant change in temperature and little change when producing sensor images by using the assumption of a uniform laser distribution. Under such uniform laser assumption, the mesh size demand is significantly reduced and the number of mesh points can be cut by about 80%.

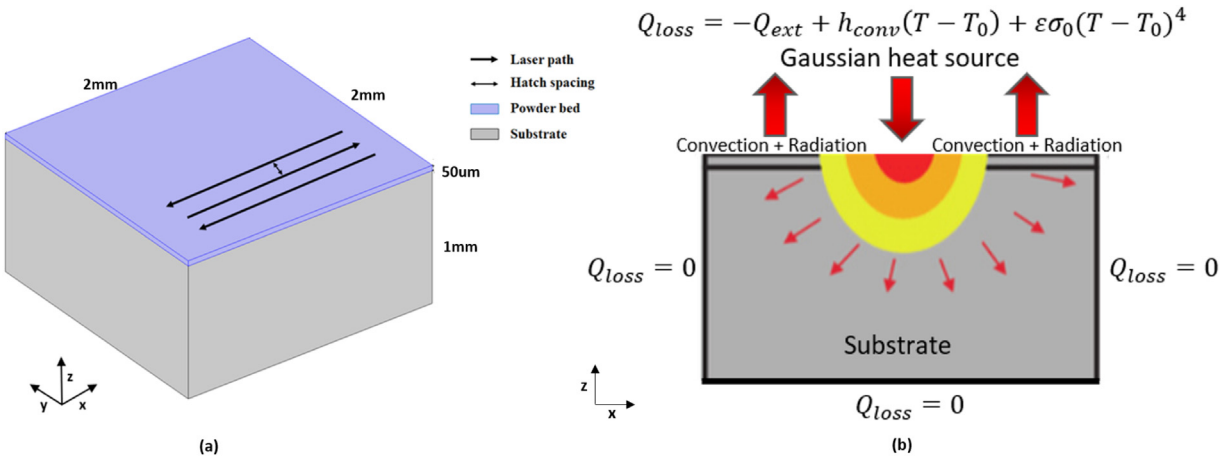


Fig. 1 – Mesoscale build part geometry: (a) 3D view and scan pattern, (b) cross-sectional view with boundary conditions.

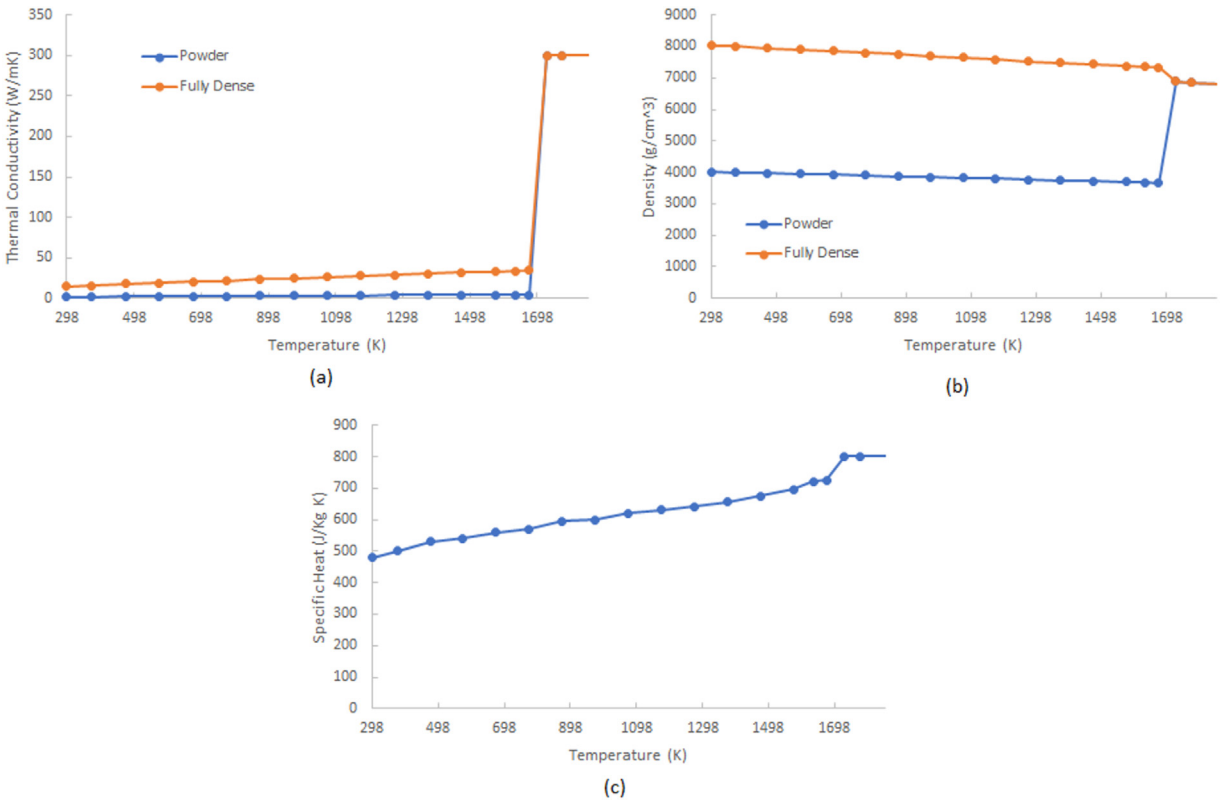


Fig. 2 – Temperature dependent thermal properties of Stainless Steel 304L: (a) thermal conductivity, (b) density, and (c) specific heat capacity.

In addition, to evaluate the effect of radiation and to seek a solution to efficiently describe its contribution, we apply a 1-D heat transfer model for the Stainless Steel 304L powder/air systems as shown in Fig. 3. A closest packed structure is considered for the powders with uniform size according to experimental and industrial relevant recipes. Constant temperature boundary conditions are applied at the two sides, which are around the laser temperature and the ambient temperature, respectively. Conduction is applied in both powder phases and air, whereas natural convection is considered in the air. Radiation is considered at the front and on every powder beads surface. It can be seen from Fig. 3 that the temperature gradient decreases sharply before the first contacting surfaces of powder due to radiation, and the air is the primary source of conductive heat transfer after the first powder surface. Thus, an effective thermal conductivity of powder k_{eff} can be considered for the porosity, conductivity of air and powder, as well as for the effect of radiation. It has been demonstrated that the calculated value lies within the prediction from Moser et al. (2016).

2.3. *Experimental based part-scale thermal feature analysis*

After the thermal performances are validated, we apply the developed FEM scheme with respect to experimental data from Missouri University of Science and Technology (MST) to reproduce results from the long-wave infrared (LWIR) thermal sensor, which monitors the overall temperature map of the whole layer. As shown in the previous section, the mesoscale FEM model is capable of producing realistic thermal results at any time interval. In addition, aided by the microscale and mesoscale model outputs, a significant level of computational demand can be reduced so that part-scale

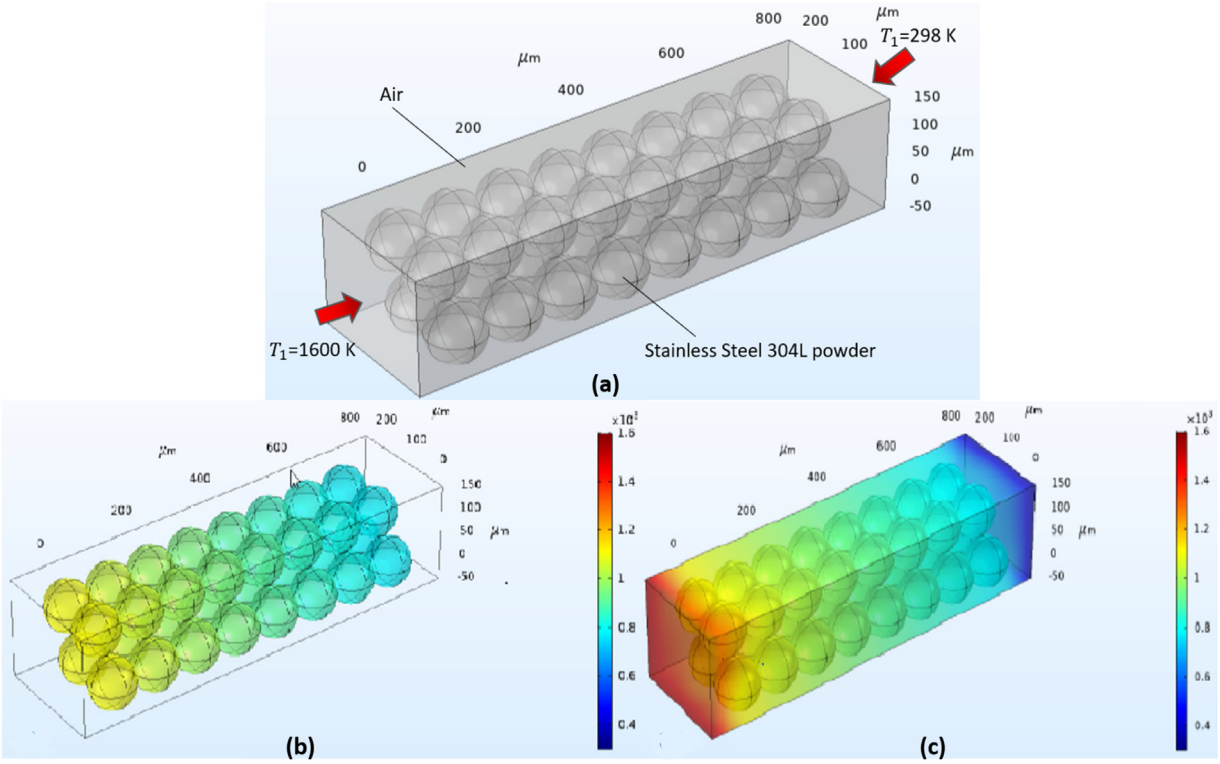


Fig. 3 – (a) The air and closely packed powder system. Temperature distribution: (b) only in the powder, and (c) both in powder and in air.

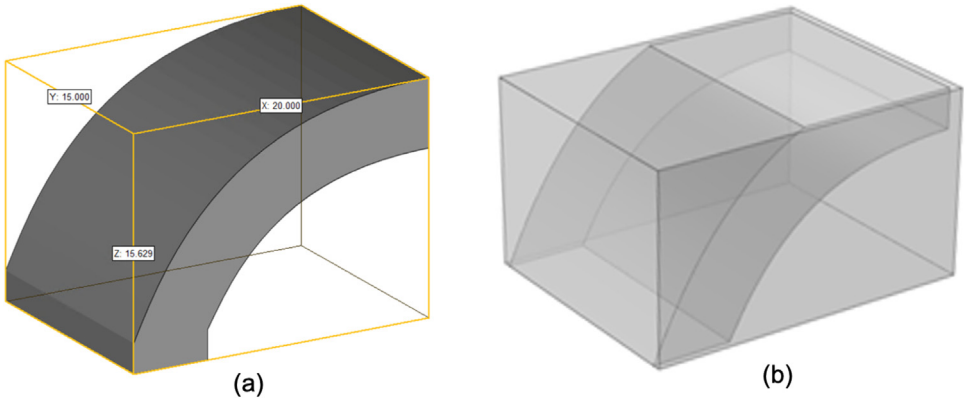


Fig. 4 – (a) Experimental semi-arch geometry, and (b) FEM simulation model geometry.

Table 1 – FEM model operating process parameters.

Process parameters	Value
Laser power	200 W
Exposure time	75 μs
Point distance	65 μm
Hatch spacing	0.1 mm
Powder thickness	40 μm
Laser spot diameter	70 μm
Absorbance	0.7

geometry can be simulated. In addition, in order to ensure model convergence, a dynamic mesh is implemented, which simultaneously allows for mesh Independence and an efficient yet accurate solution. These modifications demonstrate a great potential to reproduce the realistic thermal feature map and sensor output. Thus, we adapt our FEM model to that of the experimental plan of a desired 15 mm by 20 mm by 10 mm semi-arch geometry, as shown in Fig. 4. As presented in Table 1, a pulsed laser is simulated according to the experimental setup, which is governed by the point distance, the movement between each laser pulse, and the exposure time, the duration between each laser pulse. Other numerical specifications and details are similar to Ren et al. (2020).

2.3.1. *Experimental temperature feature analysis*

As shown in Fig. 4, the target geometry is a semi-arch with varying degrees of overhang. The experimental plans and trials were conducted at MST and we aim to create FEM simulations that produce similar results as the experimental trials. We are particularly interested in recreating the defects relating to the thermal aspects of the build process. This can be a disturbance in the laser causing a region of over-melting or under-melting, or a geometry-based defect caused by its large overhanging angle. The heat transfer rate between materials is mainly governed by the thermal conductivity, k, of the material. During the DMLS process, the thermal conductivity of the powder k_{powder} and the solid, k_{solid}, differ around one or two orders of magnitude resulting in the solid being much

better at conducting heat than its powder counterpart. At the bulk region, the heat from the laser quickly dissipates through the solid metal beneath it. However, at overhanging regions, the heat from the laser cannot be dissipated in time due to being surrounded by mainly powder material, causing a heat saturation problem. As a result of this heat oversaturation, over-melting may occur at this region, when the laser repetitively scans near this region before heat is fully dissipated. This in turn can cause a bad surface finish or even collapse of the part. In order to detect sample quality relating to the thermal aspects of the DMLS process, an LWIR camera (FLIR T420) is installed above Renishaw AM 250 vacuum-tight build chamber with a ZnSe window to monitor the defects caused by the overhang.

The LWIR camera attempts to generate a layer-wise 2D thermal history of each build layer. It picks up the radiation in a fixed field of view (0.65 mm/pixel) and calibrates these radioactive signals into temperature values. Each temperature value is assigned to a region/pixel of the final image depending on the sensor specification. Under normal process operating conditions, the highest temperature values within the melt pool would contribute predominantly to the signal produced (Mohr et al., 2020). This is due to the nonlinearity of the temperature-intensity dependence relationship despite the rapid cooling of the material after the laser passes. The LWIR camera has a frame rate of 30 Hz which in turn generates thousands of images for one layer of the build process. The maximum temperature of each pixel can be extracted from thousands of images and remapped into a single layer-wise image. This image can then be easily stored and analyzed in real-time by machine learning methods to detect defects. Also, this maximum temperature feature allows us to spot any non-uniform temperature distribution along the surface of the build part. A heat saturation region caused by large overhanging angles can be directly reflected as shown in Fig. 5.

2.3.2. *FEM data processing and validation*

To reproduce the LWIR camera results produced by the experiments, we need to further process our FEM model-generated data. After extracting the time-series temperature data from the FEM model, we calibrated the timestep of the FEM model to that of the actual LWIR camera capture rate. The FEM model generates temperature values at a rate of 300 Hz which is ten times faster than the capture rate of the experimental LWIR of 30 Hz. Therefore, the FEM raw data is temporally averaged by taking the mean of every ten data points to reflect this difference in capture rate. After the FEM data is in the same temporal frequency as the experimental data, we also need to map the FEM data to the same spatial domain as the pixel dimension. The last step is similar to the maximum temperature feature extraction step of the experimental data processing step. While the spatial averaging step is performed for every time step, in this work, we are focusing on the maximum temperature achieved by each pixel space. Therefore, for the final surface pixel temperature mapping, only the maximum temperature of the entire printing period is saved. The data calibration and preprocessing step is done within MATLAB similar to the experimental work.

We evaluate the FEM model-generated sensor image both qualitatively and quantitatively with respect to the experimental sensor images. For qualitative comparison, we look for general trends in the temperature distribution displayed in sensor images. For quantitative comparison, we examine the difference in temperature values for each pixel. It should be noted that qualitative comparison is more important due to its machine learning applications. This is because we do not want duplicate images to supplement our experimental training set, instead, we want images that share the same general trends but with some noise. In the experimental pictures shown in Fig. 5(a), the heat saturation effect causes a zone of high temperature on the right edge corresponding to the overhanging region of the build part. A similar trend of high temperature on the right edge of the FEM generated image is reproduced on the FEM simulation, as shown in Fig. 5(b).

To compare the sensor images quantitatively, we take one set of images from layer 212 of the build part and examine the temperatures by each pixel. The maximum temperature region represents the temperature of the overhanging zones while the minimum temperature region represents the bulk part above printed metal. It can is demonstrated that the maximum temperature difference is around 9% while the minimum temperature difference is around 8.2%. This difference is within a reasonable range as the experimental images' temperature can also vary due to different camera calibrations and non-uniform correction methods. We are more interested in the qualitative comparison between the images and if the temperature trends make physical sense.

2.4. *Parametric analysis using cluster computation*

The aforementioned part-scale FEM model serves as a general-purpose guideline for DMLS modeling, and it is readily customizable with respect to model accuracy and computational demands. In this work, we are also trying to look beyond experimental analysis with the aid of simulation. In particular, defects and disturbances are costly to be investigated in experiments, as those errors lead not only to higher material and operating costs but also detrimental damage to the printing platform. The purpose of the CNN classifier is to detect defects in real-time and having more types of defects in our training dataset will improve its performance. Therefore, we want to leverage the part-scale FEM model to intentionally create disturbances in order to augment the existing training dataset. Even though the FEM model can generate data at a lower cost than experimental builds, it is also not optimal to arbitrarily generate a massive dataset solely for data augmentation. The FEM model should aim to generate rare or machine damaging defect data that AM machines cannot produce in a large capacity. The first type of defect implemented is the excessive overhang which is caused by the build geometry. This type of defect can cause oversaturation of heat near the overhang due to the powder's low conductivity of heat. This can lead to a bad surface finish and even collapse of the part. The second type of defect we implemented is line streaking which is caused by scanning pattern and laser power surge. We created this defect in our FEM model by applying a laser power increase for three to five hatch spacings. This can lead to surface protrusion which can potentially cause recoater arm jams and damage. The final type of defect we implemented is the corner saturation defect. This type of defect is generally caused by suboptimal scanning patterns and laser movement errors. We created this defect in the FEM model through the slow turning of the laser at corners along with a suboptimal scanning pattern. Similar to line streaking, corner saturation can cause surface protrusion which can lead to recoater jams. Fig. 6 shows a summary of the defects introduced in the FEM model along with visualizations.

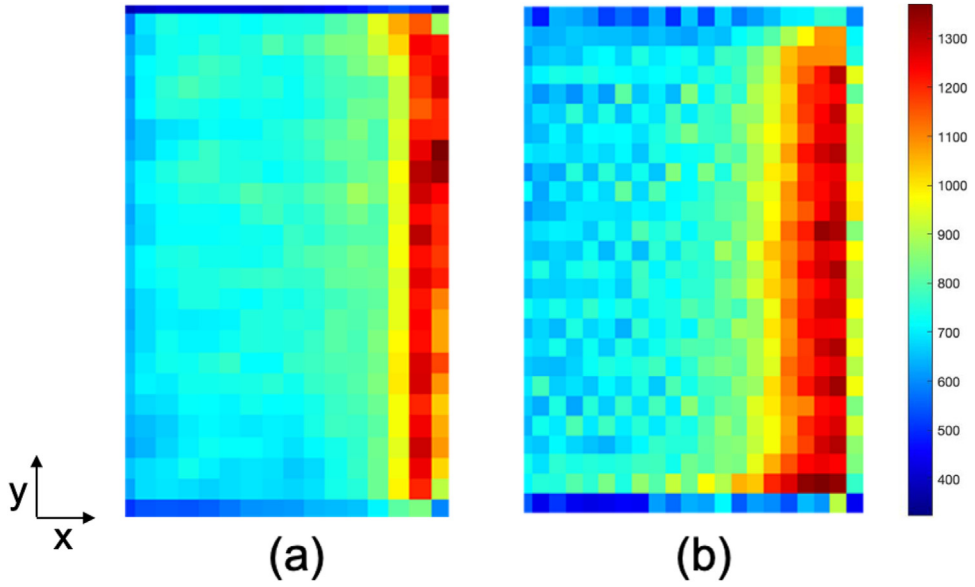


Fig. 5 – Top view of (a) experimental sensor image, and (b) FEM generated sensor image.

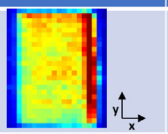
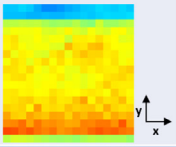
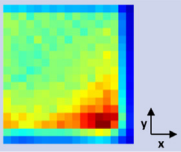
Type of Defect	Augmentation Methodology	Cause	Top-view result	Effect
Excessive overhang	Created by using the same geometry as experimental. No external disturbances added.	Caused by too much overhang in geometry. Most common type of defects included in simulation.		Bad surface finish. Extreme cases may lead to part collapse.
Line streaking	Created by applying 3-5 hatch spacing of laser power increase. Causes a line of over-melting at certain locations.	Caused by outside disturbance and potentially scanning pattern.		Bad surface finish. Extreme cases can lead to extrusion which lead to recoater jams.
Corner saturation	Created by slow turning (pause of 5-8 exposure time) of laser at corners of part. Causes heat saturation at corners.	Caused by scanning pattern.		Bad surface finish. Extreme cases can lead to extrusion which lead to recoater jams

Fig. 6 – Table showing the three types of defects introduced in our FEM simulation.

Due to a large amount of potential operating conditions, geometries, and disturbances scenarios to explore, a variety of simulations need to be carried out. To realize such mass simulation production, parametric analysis from COMSOL is utilized. Furthermore, to enhance the simulation computational efficiency, job distribution is realized with distributed cloud computing, where more than 100 of 8 GB RAMs are allocated on the UCLA hoffman2 cluster. The number of jobs and adopted cores is based on the mesh memory demand, the available core on the cluster, and the parallel computing overhead.

3. Machine learning using CNN

Pretrained CNN backbones are used for transfer learning to investigate the DMLS process adopting the simulated and experimental thermal features. In order to show the importance of using FEM model simulated images, we will be training two different types of CNNs, one with simulation augmentation and one without, to compare their respective performance in classifying the testing datasets.

3.1. *Dataset preparation*

Our total dataset contains images from four different sources. One is the FEM model simulated semi-arch images, and the rest are the experimental images of cantilever beam and semi-arch images from two different builds. All images first undergo data processing before they are separated into the training and testing datasets. For our training dataset, we use one set of experimental semi-arch images, and the FEM model simulated semi-arch images as augmentation. For our testing dataset, we use the other set of experimental semi-arch images and the experimental cantilever beam images. In our testing dataset, we wish to evaluate how well our CNN classifier performs on classifying known geometry from a different build and an unknown geometry. As a result, three different training datasets are created, composed of different combinations of experimental and simulated images. The three different sets of data are all used to train CNN classifiers in order to display the effectiveness of the simulated data in helping the detection of defects. Set 1 contains only experimental data and is used as a baseline to compare with if there were not any

simulated data to use as augmentation. Set 2 and set 3 both contain half and half of each type but differ in their simulated data composition. Set 2 only contains geometry-based defects such as those caused by overhangs while set 3 also contains intentionally created defects such as lines and edges caused by laser power disturbances. It is important to note that the FEM model generated dataset should not be used as a replacement but as a supplement to the experimental dataset. In this work, the ratio between the simulated and experimental data is kept below 0.5, so that the training dataset distribution is more realistic and closer to the experimental builds than the FEM simulations.

In any AM build, without any major machine and build design problem, there is little room for error for the majority of the layers. As a result, defects only start appearing toward the latter layers, often causing the build to stop. This trend is reflected in our experimental dataset, which causes there to be much more non-defective data than defective data. When training any machine learning model, we want to avoid the problem of class imbalance as much as possible as it can lead to the classifier having an inherent balance toward one class. Thus, a sampling strategy is included when selecting our training dataset in which we oversample the less saturated class (defective images) and undersample the more saturated class (non-defective images). Other ways to combat the class imbalance would be to adjust the weights when learning one class to the other while training. However, in our work, we found that using a sampling method was enough to solve the class imbalance problem and achieving good performance.

3.2. CNN construction and training

A residual network (ResNet) backbone is used as the basis for the transfer learning of the CNN classifier. Residual networks contain many residual blocks which create skip connections between layers, which allow the CNN to classify both simple and complex features without overfitting. In a typical residual network, there are two different types of skip connections: the identity block and the convolutional block, shown in Fig. 7(a) and (b) respectively. In the identity block, there exists a shortcut path from the input that allows the skipping of three convolutional blocks which consists of a 2-D convolutional layer, a batch normalization layer, and sometimes a Rectified Linear Unit (ReLU). Similar to the identity block, the convolutional block allows the input to skip over three convolutional blocks but the shortcut consists of a single convolutional block. Both of these skip connection types allow the input to skip many computations depending on the classification problem. Due to the implementation of these skip connections, residual networks are able to be constructed much deeper than other neural networks without adding extra computation power and worry about parameter overfitting. Among the popular CNN structures, we found the ResNet-18, a variation of the residual network, architecture to work best due to its ability to train deeper networks while also maintaining good performance.

After selecting the backbone architecture for our CNN classifier, we applied transfer learning on the last three classification layers to make it specialize in classifying LWIR sensor images. Next, we conducted hyperparameter tuning (learning rate, maximum training epochs, and the type of optimizer) on a subset of the training images in order to save computational cost. The convolutional and pooling layer hyperparameters are not changed, such as dimensions, strides, and channels, as during our hyperparameter tuning, we found the original combinations performed better. Instead of the traditional grid search algorithm, we utilized a random "coarse-to-fine" search algorithm as it allows for more efficiency by not training CNN with suboptimal hyperparameter values multiple times.

3.3. CNN model evaluation

After our CNN classifiers are trained from the three different training sets, we compare the difference in their performance on the pre-constructed testing dataset. All three CNN classifiers achieve similar high training accuracy, around 96%, which indicates that they can accurately distinguish between the defective and non-defective classes with the same distribution as the training set. Therefore, the main purpose of the testing dataset is to evaluate how well our CNN classifiers handle less known geometries and defects.

The detailed performance of each CNN classifier is shown in Fig. 8, which demonstrates a clear difference in the testing accuracy achieved by the three different training sets. The CNN trained with only experimental data (Set 1) has the lowest testing accuracy which likely indicates overfitting. Specifically, it achieves accuracy below 50% in classifying different builds and new geometries which indicates that it fails to generalize beyond one specific build. Set 2 contains FEM simulated images to provide a bit more noise to reduce overfitting. The CNN trained from set 2 achieves around the same training accuracy as set 1 while having a higher testing accuracy of 75.4% from 46.4%. Upon a close inspection of the error set 2 has made, we realized that set 2 has trouble with classifying defects coming from different sources than the training. For example, the CNN trained from set 2 has difficulties in detecting the corner and edge defects from the cantilever beam images. Within set 3, the common intentionally created defects are added to the training dataset. As a result, the CNN trained from set 3 achieves the highest testing accuracy of 94.5% while also maintaining a good training accuracy.

A common method in comparing different classifiers is through the plotting of the receiving operating characteristic (ROC) curve. The ROC curve plots the true positive rate against the false positive rate at different classification thresholds. The final ROC curve for the augmented CNN classifier is shown in Fig. 9. One quantitative metric to reflect the performance of the classifier is to look at the area under the ROC curve (AUC), which provides an aggregate measure of performance across all possible classification thresholds. One interpretation of AUC is the probability that the model ranks a random positive example higher than a random negative example. A perfect classifier would have an AUC of 1.0 as it can perfectly predict all positive and negative cases correctly. On the other hand, a no-skill classifier would have an AUC of 0.5 as it randomly predicts between positive and negative classes. With these numbers as a reference, we can evaluate how well each classifier does depend on its proximity to a perfect or a no-skill classifier. As shown in Fig. 9, the CNN trained from set 1, set 2, and set 3 has an AUC of 0.6998, 0.8774, and 0.9845 respectively.

Finally, we will discuss the applicability limits of the proposed CNN-based defect detection approach. The CNN classifier does not take into consideration of the physical properties of the DMLS process. The CNN classifier only sees the numerical values associated with the processed images and classifies them based on the gradient and intensity differences. Therefore, the pre-processing of the data into a standardized and suitable format is crucial to the success of the algorithm. For example, if we applied a different transfor-

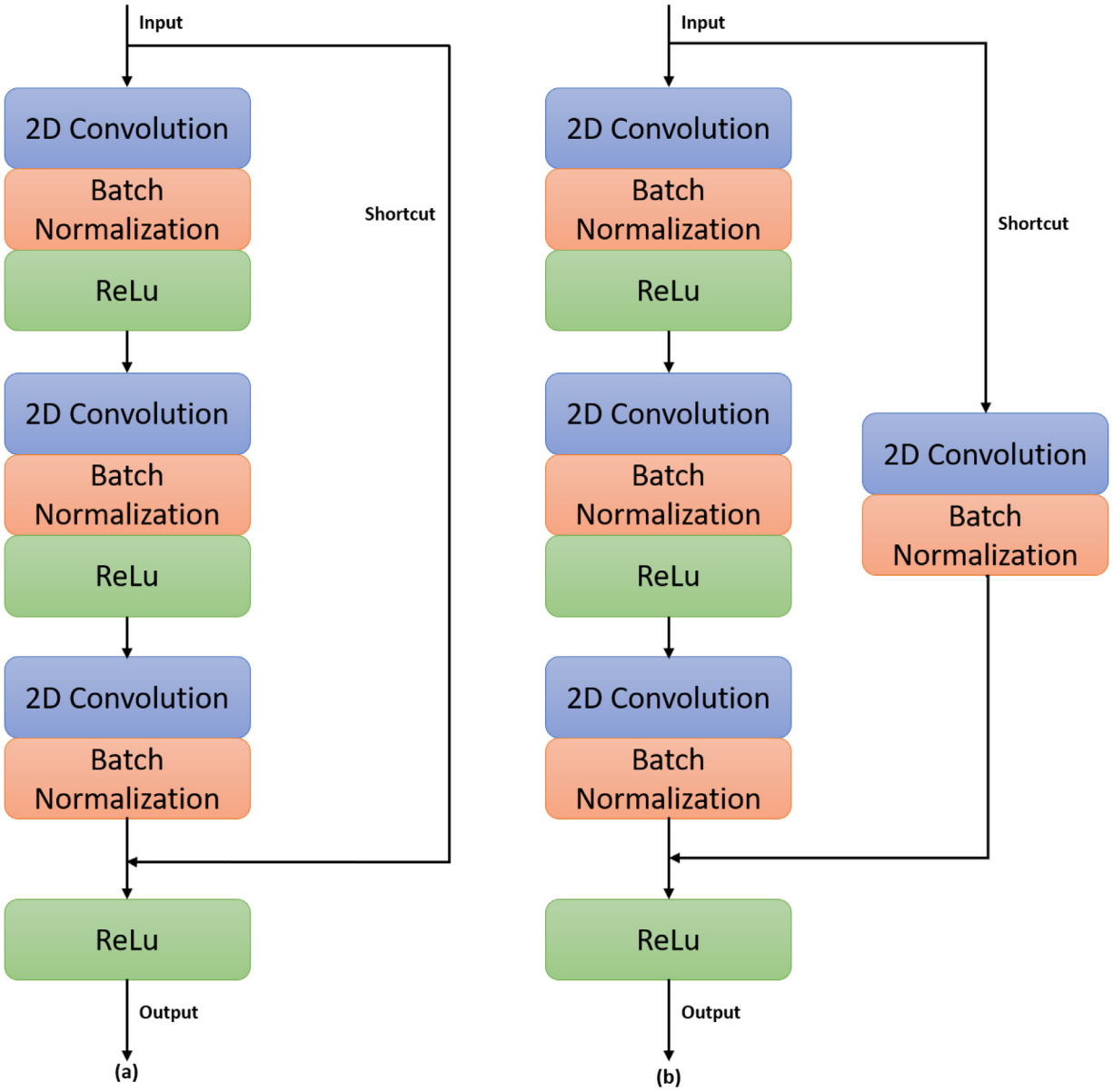


Fig. 7 – Two types of skip connections in residual networks. (a) The identity block skip connection, and (b) the convolutional block skip connection.

mation for the temperature to pixel value or used a different material, such as Inconel 718, the proposed CNN will have to undergo minor adjustment and retraining to achieve similar performance. Overall, the proposed ML defect detection framework would still be applicable to material or thermal feature transformation changes.

4. Data reduction

As mentioned in Section 1, one of the current challenges in scaling up DMLS manufacturing is the gigantic dataset size produced during the monitoring process. In the following section, we will connect our proposed workflow with data reduction implications and demonstrate several methods to reduce data.

4.1. *Data reduction in DMLS process*

The first strategy is thermal feature reduction, which refers to the transformation of the raw LWIR sensor data to the thermal feature images. The LWIR sensor first picks up light intensity signals from the DMLS build process and stores them in a time-series format. In a typical experiment example, there are seven different build parts on the build plate and each is built up to layer 240. The raw LWIR sensor data for the build plate for one layer is around 1.485 Gb which translates to around 350 Gb of data in total. Thus, for large and lengthy production processes, direct storage of the raw data is not feasible. To deal with this problem, thermal feature algorithms can be applied to avoid storing unnecessary information such as the light intensity of a location that is not currently along the laser path. For example, the maximum temperature thermal feature extracts the highest light intensity for each pixel in each layer. In the same experiment above, we can reduce 7429 frames of raw data images to a single T_{max} image for each layer, which results in around a 99.94% reduction from raw data.

Even though a large portion of the data reduction is handled by the thermal feature reduction, the CNN detection reduction is crucial when further scaling up the DMLS manufacturing process. The main goal of the CNN detection reduction is to determine which thermal feature images to transmit for storage and which to discard. The CNN detection reduction can

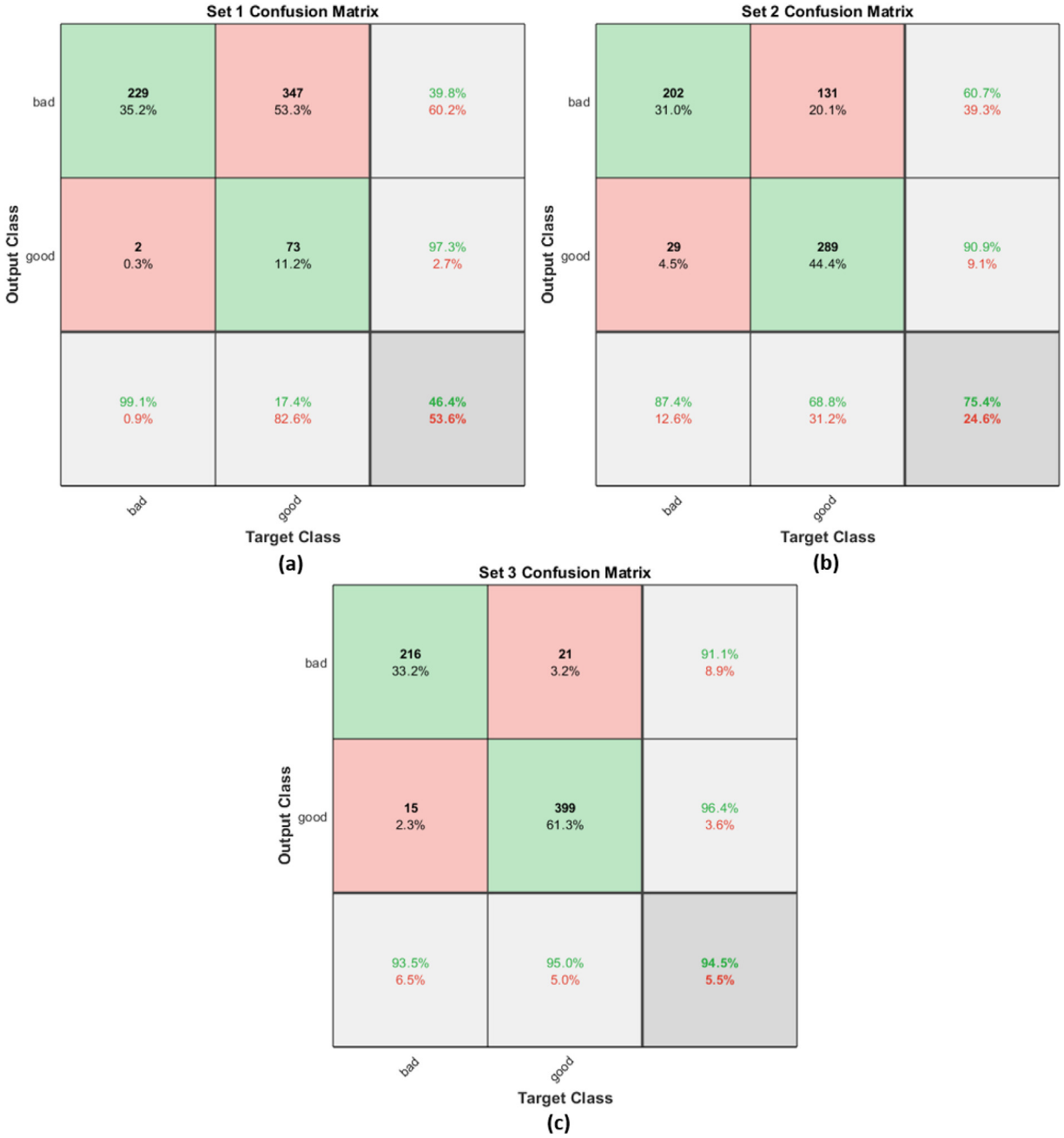


Fig. 8 – Confusion matrices for the CNN classifiers trained by (a) set 1, (b) set 2, and (c) set 3.

be separated into two parts: the CNN classifier and the transmission strategy. Adjusting the CNN classifier will affect the classification decisions made by the CNN. This is generally done by setting a classification threshold and varying this threshold to find the optimal operating point based on the importance of false positive and false negative errors. The trade-off for adjusting the CNN classifier is the performance of the classifier and the amount of data reduced. The optimal operating point based on performance is unlikely the optimal point for reducing the most data. On the other hand, adjusting the transmission strategy will affect what to do with the already classified results. The main trade-off of adjusting the transmission strategy is the amount of data reduced and the effectiveness of the data transmitted. In certain scenarios, while a large portion of data can be discarded and thus saving storage space, some key information may be lost.

4.2. *CNN classifier*

One objective of our CNN model is to prevent unnecessary data from being transmitted and stored and this can be controlled by adjusting the CNN classifier threshold, which will vary the classifications of each image. The specific amount of data reduced is based on the transmission strategy which considers what percentage of defective and non-defective images will be transmitted to storage. Different transmission strategies will be discussed in detail in the following section. To demonstrate the effect of varying the classification threshold, we assume a naive transmission strategy of transmitting all defective images while discarding all non-defective images in this section. While other approaches may vary the amount of defective and non-defective images kept, the general trend displayed by the model performance and data reduction trade-off is the same.

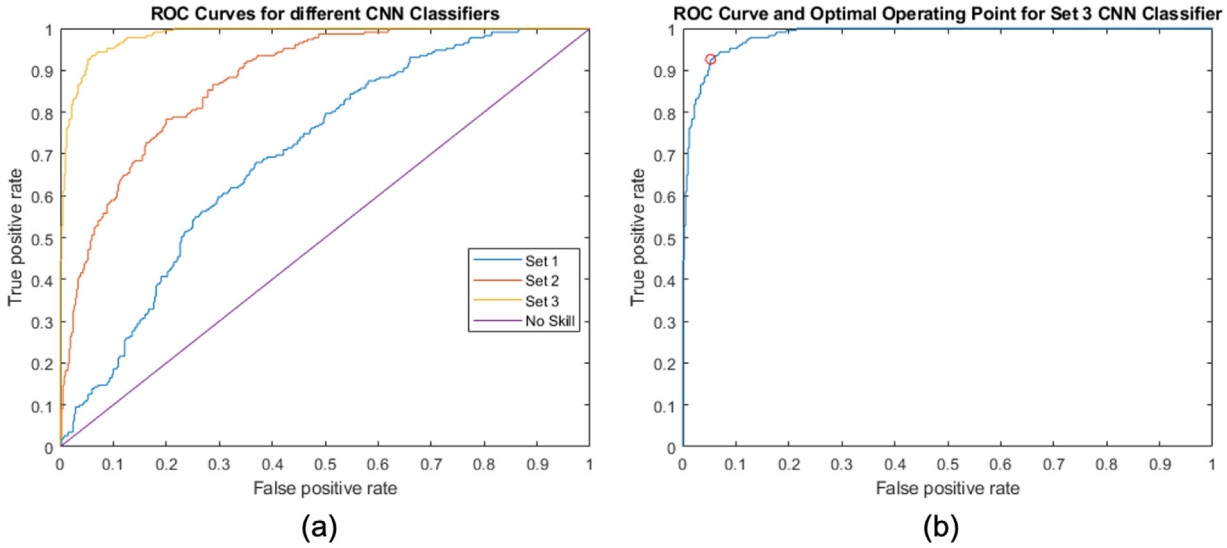


Fig. 9 – (a) Four ROC curves are drawn to compare the classifiers trained by three different sets and the no skill classifier. (b) ROC curve for the CNN classifier trained from set 3 with the red circle shows the optimal classification threshold. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

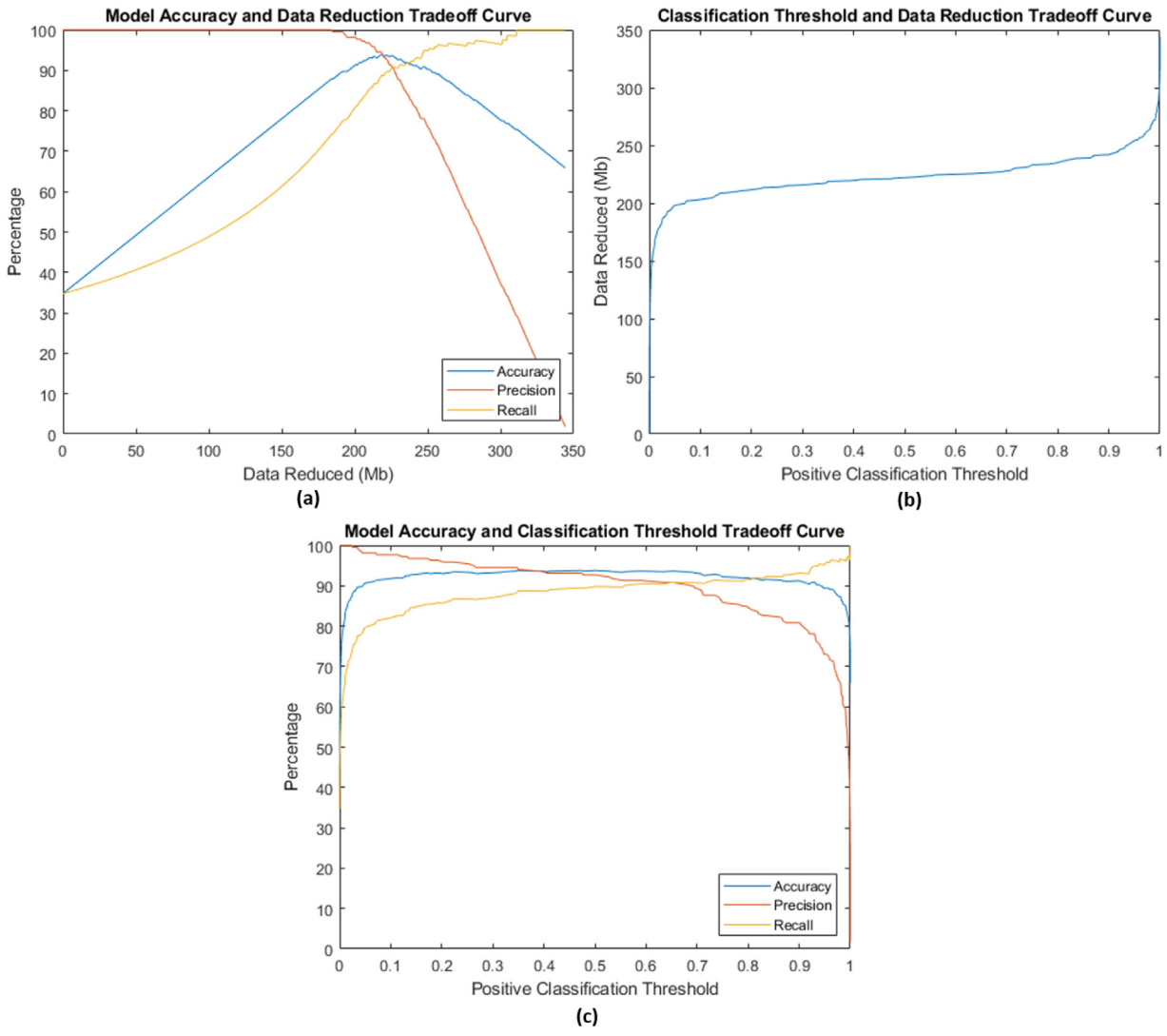


Fig. 10 – Different trade-offs between: (a) performance and data reduced, (b) data reduced and classification threshold, and (c) performance and classification threshold. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2 – Performance summary of the different transmission strategies both on a regular CNN and CNN with FEM augmentation.

	CNN	CNN with FEM augmentation
Original (no retraining)	Accuracy: 45.3%, data reduction: n/a	Accuracy: 50%, Data reduction: n/a
Complete transmission	Accuracy: 50.3%, data reduction: 0%	Accuracy: 95.1%, data reduction: 0%
Random sampling transmission	Accuracy: 51.4%, data reduction: 58.3%	Accuracy: 83.1%, data reduction: 57.4%
Score-based transmission	Accuracy: 60.3%, data reduction: 36.6%	Accuracy: 95.2%, data reduction: 75.0%

Our CNN classifies the sensor images into two classes: defective (positive cases) and non-defective (negative cases). However, simply classifying the image based on the higher confidence scores of the two classes may not result in the best industrial application. For example, if an image is given a score of 55% defective and 45% non-defective, classifying this image as defective may not be the best in practice as we want to be more certain when we are categorizing an image as defective. Therefore, an absolute classification threshold should be determined based on the importance of false positive and false negative classifications. In the DMLS process, a false negative report can range from a missed hot spot on the part leading to a bad surface finish or a recoater jam that can damage the whole AM machine. A false positive report is generally considered as a false alarm and may stop the AM machine build prematurely until a process engineer is dispatched. By setting a higher classification threshold for detecting defective cases, there will be a decrease in false positive reports as it would be less inclined to classify some uncertain errors as defective. Similarly, setting a lower classification threshold will increase the amount of false positive reports while decreasing negative reports. Fig. 10 shows the general trend of varying different thresholds and its effect on model performance and data reduction.

In order to determine the optimal operating point of the CNN classifier, we refer back to the previously drawn ROC curves in Fig. 9. A guideline with slope S, calculated by Eq. (8):

$$S = \frac{Cost(P|N) - Cost(N|N)}{Cost(N|P) - Cost(P|P)} * \frac{N}{P} \tag{8}$$

where $Cost(x|y)$ is the cost of misclassifying a y class as an x class where x and y can be any combination of P and N, N represents a negative class, and P represents a positive class. In our work, we assumed equal value of the positive and negative classes. After determining the slope of the line, we can find the optimal operating point by moving the guideline from the top left corner of the ROC plot in the direction of the bottom right corner, until it intersects the ROC curve. The optimal operating point is indicated by the red circle in Fig. 9. To avoid confusion, only the classifier trained from set 3 is drawn as it had the highest accuracy and AUC.

4.3. *Transmission strategy*

In the previous section, we have varied the classification threshold of the CNN classifier to examine the trade-off between model performance and data reduction. Another key aspect of the data reduction trade-off is the transmission strategy, which considers which images should be transmitted for further analysis and training purposes. Within the total pool of sensor images, not all images are helpful for error analysis or CNN improvement. Therefore, we want to transmit this type of image at a reduced rate compared to images containing rare defects. Another goal of transmitting sensor images is to allow the further tuning of the CNN classifier and to improve its accuracy during manufacturing cycles. Later in this section, we will compare the differences between using a complete selection strategy versus a random selection versus a score-based image transmission strategy.

To describe the trade-offs between data reduction and data effectiveness, we use the retrained CNN accuracy as the metric. The higher accuracy of the retrained CNN corresponds to a higher quality of data being transmitted. Regardless of the selected transmission strategy, we take a portion, 60% in our case, of the original CNN training set and combine it with the newly transmitted images as our new training set. We still use a portion of the original CNN training dataset because we want to use it as an anchor dataset in case of dramatic disturbance shift or error. However, not the entire original training dataset is used because we want to allow the CNN to improve upon each iteration by emphasizing the importance of the incoming images, which generally corresponds to more relevant and up-to-date data. Therefore, more weight should be given to the dataset when conducting retraining.

Table 2 provides a summary of the different transmission strategies and their respective performance and data reduction level. With the complete selection strategy, we transmit all of the classified images based on the current cycle. This method guarantees none of the crucial information being lost but is deficient from a data reduction perspective. With the random selection strategy, we transmit a portion of classified images selected at random. Since defective images are generally more valued than non-defective images, we select a higher number of defective images than non-defective images. In our example, we selected 80% of the total defective images and only 20% of the non-defective images. This selection strategy achieves a reasonable amount of data reduction around 57%, but cannot achieve the same level of accuracy as the other strategies. The final proposed strategy is the score-based selection strategy. In our CNN predicted result, a significant portion of the classified pictures scores above the maximum score threshold, which indicates that the previous CNN iteration can already detect these types of errors. Thus, transmitting all of the images may result in inefficient use of the data storage, we can eliminate the uncertain classifications through the implementation of a minimum score threshold and also achieve data reduction through the implementation of a maximum score threshold. With the score-based transmission strategy, we are able to reduce the incoming data by 75%. As shown in Table 2, the proposed CNN with FEM augmentation paired with score-based transmission strategy achieves the highest accuracy as well as the most data reduction. Performance can be further improved if more detailed tuning is conducted concerning specific processes. Example tuning parameters include the previously mentioned classification threshold and the maximum and minimum score-based thresholds.

5. Conclusion

Several levels of FEM modeling and simulation were investigated in this work. A mesoscale model was first constructed which demonstrates that the heat transfer and melt pool behavior is successfully reconstructed with respect to experimental works and analytical relationships. Microscale tests were then investigated. In particular, the laser source simplification and heat conductivity rectification are proven to achieve significant computational demand reduction while preserving model fidelity. Based on these modeling works, a part-scale model was developed to reproduce experimentally relevant thermal feature analysis. This FEM model simulated an experimental semi-arch part geometry. In addition, the model will be useful in understanding experimental defects, and also in the CNN training data augmentation. Transfer learning was performed on a modified ResNet CNN backbone to reduce computational load. Three different CNN classifiers trained from different datasets are compared to each other in performance through examining their respective confusion matrices and ROC curves. It can be concluded that the CNN classifier augmented with FEM thermal images has shown much improved performance than its non-augmented counterpart by increasing detection accuracy from 45.2% to 92.3%. Finally, data reduction implications were discussed in relation to the CNN classifier to overcome the problem of data transmission and storage.

Conflict of interest

None declared.

Declaration of competing interest

The authors report no declarations of interest.

Acknowledgements

Financial support from the Department of Energyis gratefully acknowledged.

References

Arısoy, Y.M., Criales, L.E., Özel, T., 2019. Modeling and simulation of thermal field and solidification in laser powder bed fusion of nickel alloy IN-625. Opt. Laser Technol. 109, 278–292.

Craeghs, T., Bechmann, F., Berumen, S., Kruth, J.P., 2010. Feedback control of layerwise laser melting using optical sensors. Phys. Proc. 5, 505–514.

Delgado, J., Ciurana, J., Rodríguez, C.A., 2012. Influence of process parameters on part quality and mechanical properties for DMLS and SLM with iron-based materials. Int. J. Adv. Manuf. Technol. 60, 601–610.

Dong, Z., Liu, Y., Wen, W., Ge, J., Liang, J., 2019. Effect of hatch spacing on melt pool and as-built quality during selective laser melting of stainless steel: modeling and experimental approaches. Materials 12, 50.

Frazier, W.E., 2014. Metal additive manufacturing: a review. J. Mater. Eng. Perform. 23, 1917–1928.

Gibson, I., Rosen, D.W., Stucker, B., et al., 2014. Additive Manufacturing Technologies, Vol. 17. Springer.

Gong, H., Rafi, K., Gu, H., Starr, T., Stucker, B., 2014. Analysis of defect generation in Ti-6Al-4V parts made using powder bed fusion additive manufacturing processes. Addit. Manuf. 1, 87–98.

Grünberger, T., Domröse, R., 2014. Optical in-process monitoring of direct metal laser sintering (DMLS): A revolutionary technology meets automated quality inspection. Laser Techn. J. 11, 40–42.

Khalil, A., Loginova, I., Solonin, A., Mosleh, A., 2020. Controlling liquation behavior and solidification cracks by continuous laser melting process of AA-7075 aluminum alloy. Mater. Lett. 277, 128364.

Khanzadeh, M., Chowdhury, S., Tschopp, M.A., Doude, H.R., Marufuzzaman, M., Bian, L., 2019. In-situ monitoring of melt pool images for porosity prediction in directed energy deposition processes. IISE Trans. 51, 437–455.

Li, J., Hu, J., Zhu, Y., Yu, X., Yu, M., Yang, H., 2020. Surface roughness control of root analogue dental implants fabricated using selective laser melting. Addit. Manuf. 34, 101283.

Liu, R., Wang, Z., Sparks, T., Liou, F., Newkirk, J., 2017. Aerospace applications of laser additive manufacturing. Laser Additive Manufacturing, 351–371.

Lough, C.S., Wang, X., Smith, C.C., Landers, R.G., Bristow, D.A., Drallmeier, J.A., Brown, B., Kinzel, E.C., 2020. Correlation of SWIR imaging with LPBF 304L stainless steel part properties. Addit. Manuf. 35, 101359.

McLouth, T.D., Witkin, D.B., Bean, G.E., Sitzman, S.D., Adams, P.M., Lohser, J.R., Yang, J.M., Zaldivar, R.J., 2020. Variations in ambient and elevated temperature mechanical behavior of in718 manufactured by selective laser melting via process parameter control. Mater. Sci. Eng.: A 780, 139184.

Mohr, G., Altenburg, S.J., Ulbricht, A., Heinrich, P., Baum, D., Maierhofer, C., Hilgenberg, K., 2020. In-situ defect detection in laser powder bed fusion by using thermography and optical tomography-comparison to computed tomography. Metals 10, 103.

Moser, D., Pannala, S., Murthy, J., 2016. Computation of effective thermal conductivity of powders for selective laser sintering simulations. J. Heat Transfer 138, 8.

Ramos, D., Belblidia, F., Sienz, J., 2019. New scanning strategy to reduce warpage in additive manufacturing. Addit. Manuf. 28, 554–564.

Ren, Y.M., Zhang, Y., Ding, Y., Wang, Y., Christofides, P.D., 2020. Computational fluid dynamics-based in-situ sensor analytics of direct metal laser solidification process using machine learning. Comput. Chem. Eng. 143, 107069.

Romano, J., Ladani, L., Sadowski, M., 2016. Laser additive melting and solidification of Inconel 718: finite element simulation and experiment. JOM 68, 967–977.

Rosenthal, D., 1941. Mathematical theory of heat distribution during welding and cutting. Weld. J. 20, 220–234.

Sadowski, M., Ladani, L., Brindley, W., Romano, J., 2016. Optimizing quality of additively manufactured Inconel 718 using powder bed laser melting process. Addit. Manuf. 11, 60–70.

Scime, L., Beuth, J., 2019. Using machine learning to identify in-situ melt pool signatures indicative of flaw formation in a laser powder bed fusion additive manufacturing process. Addit. Manuf. 25, 151–165.

Scime, L., Siddel, D., Baird, S., Paquit, V., 2020. Layer-wise anomaly detection and classification for powder bed additive manufacturing processes: a machine-agnostic algorithm for real-time pixel-wise semantic segmentation. Addit. Manuf. 36, 101453.

Scott, C., 2017. EOS introduces EOSTATE Exposure OT, First Commercial Optical Tomography System for Additive Manufacturing, Accessed from https://3dprint.com/178624/eos-eostate-exposure-ot.

Singh, S., Ramakrishna, S., 2017. Biomedical applications of additive manufacturing: present and future. Curr. Opin. Biomed. Eng. 2, 105–115.

Wang, X., Lough, C.S., Bristow, D.A., Landers, R.G., Kinzel, E.C., 2020. A layer-to-layer control-oriented model for selective laser melting. Proceedings of American Control Conference, 481–486.

Yang, J., Yu, H., Yin, J., Gao, M., Wang, Z., Zeng, X., 2016. Formation and control of martensite in ti-6al-4v alloy produced by selective laser melting. Mater. Des. 108, 308–318.

Yoon, Y., Yim, J., Choi, E., Kim, J., Oh, K., Joo, Y., 2016. by local laser melting and directional solidification. Mater. Lett. 185, 43–46.

Yuan, B., Giera, B., Guss, G., Matthews, I., Mcmains, S., 2019. Semi-supervised convolutional neural networks for in-situ video monitoring of selective laser melting. In: Proceedings of Winter Conference on Applications of Computer Vision (WACV), Waikoloa Village, Hawaii, pp. 744–753.

Zhang, Y., Hong, G.S., Ye, D., Zhu, K., Fuh, J.Y., 2018. Extraction and evaluation of melt pool, plume and spatter information for powder-bed fusion AM process monitoring. Mater. Des. 156, 458–469.