The International Journal of Advanced Manufacturing Technology
https://doi.org/10.1007/s00170-025-15591-y

ORIGINAL ARTICLE

Laser powder bed fusion parameter estimation with k-NN

Patrick Jung[1] · **Nathaniel DeVol**[1] · **Christopher Saldaña**[1] · **Katherine Fu**[2]

Received: 7 November 2024 / Accepted: 17 April 2025

Abstract
Laser powder bed fusion (L-PBF) is a technique within additive manufacturing that uses a high power density laser to build parts from fused powdered metal alloy. This technology is well equipped to produce complex parts with otherwise impossible features, such as hidden voids or lattice structures. Alongside capability, reliability and quality are key characteristics considered when choosing a manufacturing method, and these are gaining attention as this method becomes more prevalent in industry. One main indicator of a stable L-PBF process is consistent melt pool geometry, and the properties of which are likely to determine the quality of the part produced. As computing power and sensing technologies become more advanced, this melt pool geometry could be studied in real time. This work addresses the challenge by leveraging a k-nearest neighbor (k-NN) model to identify key features within melt pool imagery and predict the energy density. The k-NN model was trained on data provided by the National Institute of Standards and Technology (NIST). Data preprocessing was performed on the images to extract features that were used in the k-NN model. This approach was used to accurately infer the energy density of unseen layers within the same part. The algorithm was subsequently tested with unique scan strategies and found to reasonably estimate the energy density of different parts. A fivefold cross validation found the algorithm to be consistently predicting the class of 91.4% of the in situ melt pool images.

Keywords Additive manufacturing · Laser powder bed fusion · Melt pool monitoring · Machine learning · Smart manufacturing

Nomenclatures

σ	Standard deviation
AMMT	Additive Manufacturing Metrology Testbed
AM	Additive manufacturing
COV	Coefficient of variance
EBM	Electron-beam melting
ED	Energy density
I	Mean
k-NN	K-nearest neighbor
L-PBF	Laser powder bed fusion
ML	Machine learning
NIST	National Institute of Standards and Technology
P	Power
STL	Standard tessellation language
V	Velocity
XYPT	X position, Y position, power, camera triggered

1 Introduction

Research in automating the process level of manufacturing operations is motivated by a strong belief that it will provide increased productivity, improved part quality, reduced costs, and relaxed machine design constraints. The basis for this belief is two-fold. First, manufacturing process automation can be applied to both large batch production environments and small batch jobs. Second, process automation can autonomously tune machine parameters (velocity, power, scan strategy, etc.) on-line and off-line to substantially increase the machine's performance in terms of part tolerances and surface finish, operation cycle time, etc.

✉ Katherine Fu
kate.fu@wisc.edu

Patrick Jung
pjung@gatech.edu

Nathaniel DeVol
ndevol3@gatech.edu

Christopher Saldaña
christopher.saldana@me.gatech.edu

[1] George W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA

[2] Department of Mechanical Engineering, University of Wisconsin - Madison, Madison, WI 53706, USA

Published online: 28 April 2025

Process automation holds the promise of bridging the gap between product design and process planning, while reaching beyond the capability of a human operator. The success of manufacturing process automation hinges primarily on the effectiveness of the process monitoring and control systems [35]. Particularly for metal additive manufacturing (AM), successful process control with precision measurement science will not only extend the capabilities for this technology, but also introduce reliability to advanced manufacturing methods that have not yet been widely incorporated into production environments. The most widely used metal AM technique in the commercial industries is laser powder bed fusion (L-PBF), a method for melting thin layers of metal powder with a high-powered laser, recoating the build area with powder, and repeating this process until a dense part is completed.

A promising future is in store for L-PBF AM. However, widespread adoption of L-PBF with metallic parts hinges on solving a main challenge: the requirement that the final product should meet engineering quality standards. This includes increasing the predictability of melt pool geometry, since the melt pool has a strong impact on the final mechanical properties [39]. Modeling advances on this front typically rely on physics-based simulations, which require a deep understanding of the process involved, as well as full access to the input parameters and machine specifications [17]. An attractive alternative to answering this challenge is through machine learning and predictive simulation where a system is understood from data collected on it instead of a physics-based simulation. In this work, an inexpensive machine learning approach, which utilizes open-source algorithms, is explored to understand its novel application to predicting machine parameters using in situ monitoring.

The shape of a melt pool is significantly affected by laser power and scan velocity and melt pool formation has been shown to determine material properties [34]. An understanding of the fundamental science behind these relationships is currently being developed in the literature [22]. However, an intermediate process, such as melt pool formation, is challenging to model because it includes several complex physical phenomena, including thermal conduction, fluid dynamics, and phase changes of materials [17, 18]. Machine learning (ML) provides a different path to correlate the process parameters and results without the need for advanced modeling, and it has been shown to be effective in additive manufacturing applications [18].

The objective of this work is to identify key features in monitoring the melt pool of L-PBF AM systems and to use those features in a supervised machine learning algorithm to predict energy density. An understanding of how melt pool images can predict AM machine commands or scan strategy for closed loop control or anomaly detection is not well presented in the literature. To complement the current work being done to predict melt pool geometry from velocity and/or power, k-nearest neighbor (k-NN) models were trained and evaluated for interlayer prediction and transferability to other parts. This model type was chosen based on its ability to work well with smaller datasets, as this work explicitly used data and images from the National Institute of Standards and Technology (NIST) and their Additive Manufacturing Metrology Testbed (AMMT) [21]. In addition to its advantages there, k-NN models generally require smaller computational resources and have not been explored as thoroughly in the literature, making them an interesting addition to the field of closed loop control of L-PBF AM systems.

The following sections are structured as follows. Section 2 will begin with a background of additive manufacturing while specifically focusing on laser powder bed fusion. A brief introduction to machine learning will also be presented, and the intersection of these two subjects will be discussed as it currently is addressed in the literature. Section 3 will include the methodology used in developing the training and testing matrices from the experimental dataset presented in [21]. The pre-processing of the in situ monitoring images will be reviewed, as well as the flow of information into the k-NN classification algorithm. In Section 4, the results of interlayer and transferability analysis will be presented. The k-fold validation will also be included with some inferences about the results. A further discussion is held in Section 5 regarding the challenges faced, along with the assumptions made. Section 6 delivers the conclusions made and postulates future possibilities for improvement in this area.

2 Background

2.1 Laser powder bed fusion

Powder bed fusion, one of the seven additive manufacturing (AM) techniques defined by ASTM F2792, uses a thermal energy source, such as a high-powered laser, to melt metal powder to the base plate or previous layers [8]. This technique is particularly well suited for applications that require better precision and accuracy but can spare more time per build. Powder bed fusion is the parent technology of electron-beam melting (EBM) and laser powder bed fusion (L-PBF). The main difference between the two technologies is the thermal energy source used to heat the metal powder to melting temperature. This work focuses on L-PBF, which is typically more cost-effective than EBM, and is the most widely used metal additive process in the field [15].

The process of L-PBF is executed as follows: first, the build area is filled with a layer of metal powder and the material is evenly distributed with a recoating system. A galvanometer scanner orients the energy source, such as a laser, and transfers thermal energy to the top layer of material

through a predetermined scan strategy to add a layer typically 30–50 μm in thickness to the part. The build platform is then lowered by this thickness without removing the excess metal powder, and more powder is dispensed to recoat the build area for the next layer. The laser is oriented with great precision in the XY plane and has a focus diameter of less than five thousandths, and with a layer thickness less than half of a sheet of paper, features can be built with high fidelity [11]. This process typically takes place in an environment filled with nitrogen, argon, or another inert gas to keep the solidified layers from oxidizing after they have been melted. Once the final layer has been fused to the part, the part will need to be removed from the build area and from its encapsulation of packed metal powder.

Numerous layers are fused together to create a part through a three-step process: thermal energy absorption, melt pool formation, and solidification [19]. This work focuses on the second step of this process, melt pool formation, which is significantly influenced by process parameters such as energy density and laser scan strategy [28, 37]. Several physical phenomena are at play as well during this process, including fluid mechanics, heat transfer, and phase transformations.

2.2 Influence of Melt Pool Geometry

Within the process of laser powder bed fusion, the characteristics of the melt pool have been shown to influence the part's microstructure and the resultant material properties [27]. The melt pool geometry, as is directly influenced by energy density, can be controlled to improve part quality as well [11, 30]. This could be partly due to the fact that energy density has been shown to be correlated to part density [41]. The effects of energy density parameters, scan speed and laser power, on material properties have been studied by many in the literature [7, 13, 16, 33, 41].

The underlying effects from reduced energy density manifest primarily as an increased level of porosity in the part, which in turn weakens the mechanical properties of the affected area by introducing stress concentrators during manufacturing in the form of external pores, voids, and micro-cracks [9, 20]. Parts with 100% relative density, on the other hand, have been experimentally proven to mimic the material properties of wrought materials, if not surpass them. It has been shown that poor energy density may cause this because of incomplete fusion or trapped gases within the part [17]. At higher energy densities, the pores identified were typically smaller in size and were found to be more regular and even in their spherical shape. Porosity increases are essentially being linked to improperly formed melt pools, which emphasizes the importance of this aspect for PBF [6]. With energy density controlling melt pool geometry and melt pool geometry heavily influencing porosity and subsequently the strength of the material, the melt pool acts as an intermediate step in the efforts to predetermine the properties of the resultant part.

In order to intelligently control melt pool geometry, a well-defined relationship between process parameters and melt pool dimensions will need to be established. Figure 1 illustrates the typical geometric boundaries that a melt pool shape is defined by, including length, width, and hatch space h [23, 24].

Guo et al. conducted experiments where energy densities were kept constant for three different sets of experiments focused on the relationship between melt pool dimensions and the individual parameters that make up energy density [12]. Correlations were found between energy density and melt pool volume overall, but within each energy density tier, there is another linear relationship between melt pool dimension and the process parameters. These two correlations compound upon one another, forcing the models of this problem into the nonlinear optimization space.

Despite increases among energy density and melt pool volume being correlated, it is difficult to linearly predict when the process is open to influence from uncontrollable or unexpected factors [5]. Even if the surrounding environment is controlled, there is a possibility of variation in melt pool size as a build progresses, and defects can grow to cause build failure if the signs of a defect are not caught early. Open loop control would likely be inadequate when quality and/or complexity is highly valued. As computing power becomes cheaper, including closed loop control in additive

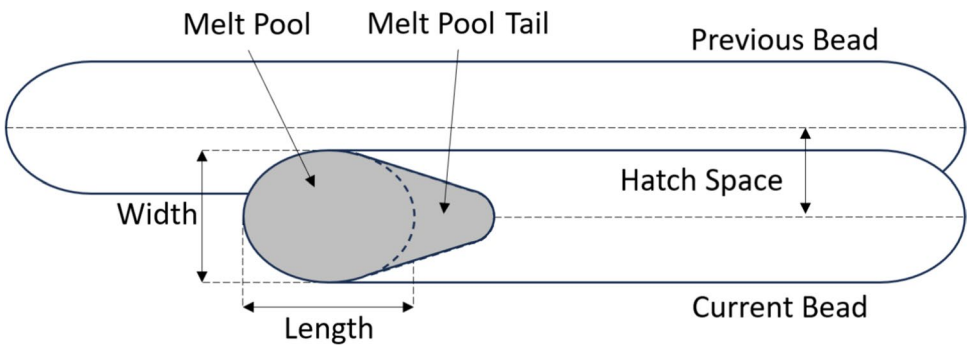

Fig. 1 Melt pool geometry in L-PBF

processes is becoming more accessible, and algorithms need to be developed to address this issue.

2.3 In-situ process monitoring

The necessity of high feature resolution has encouraged more interest in the development of sensor technologies to monitor laser powder bed fusion, and defect avoidance has become critically important. NIST has outlined measurement science needs for metal additive manufacturing, and these include qualification and certification of the parts and processes that make up powder bed fusion, the part's mechanical properties, and the observation of machine performance [32]. Closed-loop control in metal AM has subsequently been identified as an important focus for the scope of measurement science work, as robust process control in situ can reduce outcome variability and increase the quality of parts produced by these means [40].

Closed-loop control has been investigated before in a similar context for the process of directed energy deposition by monitoring the infrared radiation emitted by the melt-pool [3]. The results of this study were encouraging, as the microstructure became more homogeneous with the addition of a PID controller connected to the laser power responding to changes in temperature.

2.4 Machine learning

Machine learning algorithms have been introduced as a revolutionary addition to additive manufacturing, as they increase the probability of success in high-value and complex parts [19, 36]. ML algorithms are oriented around prediction and interpolation within a large dataset where it is advantageous to discover patterns. Some algorithms, called unsupervised learning methods, do not even require labels for their data points, as they simply search for natural subsets of data. If the expected data classification or response vector is available, then a supervised learning algorithm is preferred.

ML algorithms can be used in a monitoring and control context due to the potential for lower processing power needed to predict new values. As an additional benefit, ML models can improve over time as the pattern boundaries automatically shift to optimally incorporate new data. Therefore, these algorithms are suitable for interpreting parameters in processes where the environment and part geometry may vary over time. With respect to AM, machine learning is already used to recognize issues and predict the quality of parts [2, 10, 38].

The k-nearest neighbor algorithm is a supervised classification algorithm that seeks to make predictions on a sample data point by determining the k nearest points by Euclidean distance. Euclidean distance is used because it has been shown to be applicable across different distributions and is not sensitive to class imbalances [4]. The most frequent class label among the k nearest neighbors is applied to the sample data point.

3 Methodology

The goal of this work is to develop a prediction method for the energy density of laser powder bed fusion systems in additive manufacturing. The significant process parameters that affect melt pool geometry, scan velocity and laser power, are more accurately represented in their effects by looking at energy density. Energy density (ED) can be approximated with the following:

$$\mathrm{ED} \sim \frac{P}{V} \tag{1}$$

where

ED energy density, J/mm

P power, W

V velocity, mm/s

Instead of predicting multiple dependent variables, this method rather delivers an estimation of the largest influence on melt pool size and requires less computing power. The faster method is desirable in this application for low-cost additive manufacturing prediction.

3.1 Data interpretation

The data studied consists of input command files, in situ process monitoring data, and metadata provided by the NIST, who performed a build with the Additive Manufacturing Metrology Testbed (AMMT) [21]. Twelve parts, identical in shape, were built out of IN625 on the same build platform, which consisted of hot rolled and annealed wrought nickel alloy. The standard tessellation language (STL) file, as illustrated in Fig. 2, represents the near rectangular prism that was built for this dataset.

None of the 12 parts varied in geometry, but rather, they each had a unique scan strategy, and some of the differences can be observed in Fig. 3. The first category of information within the dataset provided by NIST, the build command data, can be used to highlight these differences, and there were two types of such data provided: AM G-Code and XYPT commands, where XY indicates the Cartesian position (in mm) in the current layer, P represents power (in Watts) of the laser, and T represents the binary indication of

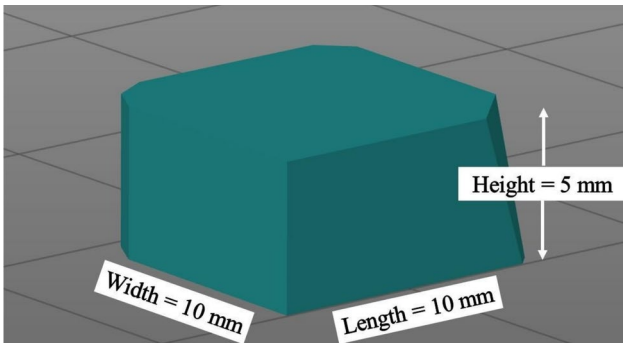

Fig. 2 Graphical representation of studied part in STL format

capturing melt pool images. In this work, the XYPT commands were used because of the interest in directly comparing programmed position and power to the second category of information, the in situ measurement data. This second category of information has two subcategories: co-axial melt pool images and grayscale layer images. From these, the melt pool images were more important to this work because their capture could be directly placed in the context of the location and power of the scanning laser.

The melt pool camera captured thousands of images per layer, and altogether, this build required 250 layers each 20 µm thick, with a laser spot diameter of 85 µm. The images in this subcategory, all 120 × 120 pixels, focused on a single part out of the 12 built during any layer. For example, during layer 2, the in situ melt pool camera was only triggered when building part 2, and similarly, images from layers 1, 13, 25, 37, etc. (~ 20 equally distributed layers in total) were only attributed to part 1. Collecting a training set of images for studying a certain part was then approached by filtering out the other layers.

3.2 Data preprocessing

The XYPT commands and melt pool images were studied layer by layer and, within each layer, frames 1-n were individually considered for viability in the training set. First, the gray image was binarized based on the ISO 50 method, where any pixel with greater intensity than 50% of the maximum intensity would be counted as a true value and all else were considered false. Once binarized, there were several conditions that had to be met in order for an image to be included: the shape made up of true values must have one and only one centroid, and it must also be larger than 40 pixels squared in area. That said, if a secondary shape in the image is recognized and it is smaller than the required area, that shape was removed and the primary shape was not considered valid. These cases are represented in Fig. 4.

Once an image was classified as viable, features of the image were extracted to increase the speed and efficiency of a k-NN algorithm, including descriptive statistics of the melt pool area, dimensions of the melt pool shape, and centroid locations. Two centroid locations were found in each

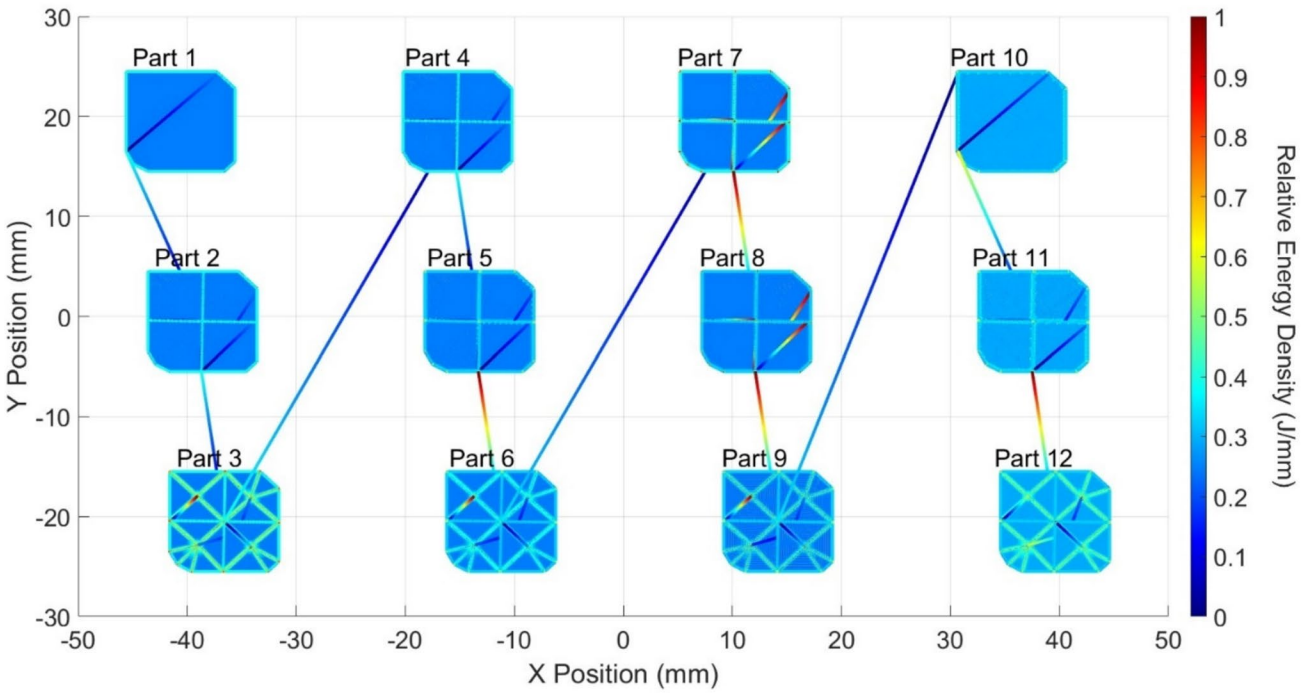

Fig. 3 Unique scan strategies of the 12 parts

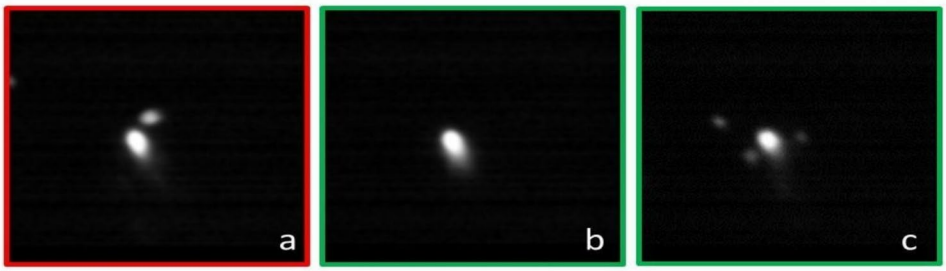

Fig. 4 Image viability: **a** images with two objects not considered, **b** typical melt pool image studied, **c** small, low intensity melt pools were removed

typical melt pool image, that of the shape greater than the ISO 50 threshold, and the centroid of the melt pool *tail*. The tail is defined in this work as the collection of pixels that lie between the ISO 50 and a similarly obtained 35% threshold. In order to extract this information, the original image was again binarized with these new limits, but no size requirements were imposed on the tail section. The process of binarization and isolation of the necessary features is illustrated in Fig. 5.

The original gray image was characterized by Matlab's *regionprops* function in the Image Processing Toolbox, because most of the melt pool images represented an elliptical shape. The centroid of the binarized shape and the *regionprops* estimation of the elliptical axes are superimposed on the original gray image in Fig. 5a. The same major axis estimation is carried over to the binary image in Fig. 5b. In Fig. 5c, the binary image of the isolated tail is shown, and a separate centroid is calculated from the new shape. In Fig. 5d, the culmination of these efforts is displayed, as the line fit through both centroids is taken to be the correct major axis orientation, while the original major and minor axis lengths are kept. Extra input features such as these can be used as additional model inputs to improve model output fidelity such as in (J. [26]). These modified orientations give the k-NN algorithm better context and could allow for path planning prediction in future work. However, the in situ monitoring camera was not aligned with the coordinate system of the substrate, and thus, the orientations found using this method must be manipulated with a transformation matrix to be reconciled with the build command data.

Some of these independent variables can be combined or manipulated to give the model more relevant information, defined here as *derived variables*, such as coefficient of variance (COV), which is calculated with the equation:

$$\mathrm{COV} = \sigma / \underline{I} \tag{2}$$

where

COV coefficient of variation (unitless).

σ standard deviation (unit of original data).

I mean or average (unit of original data).

Figure 6 illustrates the flow of information from the datasets to the k-NN classification algorithm to predict the process parameters velocity, power, or energy density.

The ground truth in relation to these images lies in the XYPT commands, as each image correlates to a trigger indicator in the "T" column of this build command data. These process parameters provided by this data give the algorithm its class names, represented by velocity, power, or energy density around the time each image was captured. The range of energy density class labels was categorized into three main groups: low (energy density of 0.115 $\pm$0.01 J/mm), medium (energy density of 0.245 $\pm$0.01 J/mm), and high (energy density of 0.39 $\pm$0.01 J/mm). The maximum number of viable images per energy density category was found and the size of the smallest category was used as the ceiling for the other categories. Matlab's randperm function chose the images to delete from the set in order to normalize the distribution of energy density categories represented in the set. Even though this work studied energy density prediction, the same algorithm and training matrix could be used to predict velocity or power as well. The information provided by the melt pool image is the culmination of many input features, but could also be linked to output features as well using this algorithm. Porosity would be particularly interesting as it directly relates to material strength, but such information was not provided in this dataset.

3.3 Training/testing split

This algorithm was validated through the evaluation of interlayer prediction within a certain part, as well as through the use of k-fold cross validation. The prediction of process parameters within a certain part required a testing set and a training set split that gives the model enough data to train with, while adequately representing the error of the result with a large enough testing set. As there were typically 50,000 melt pool images available for each part, a testing set of 20% was used, leaving 80% of the data

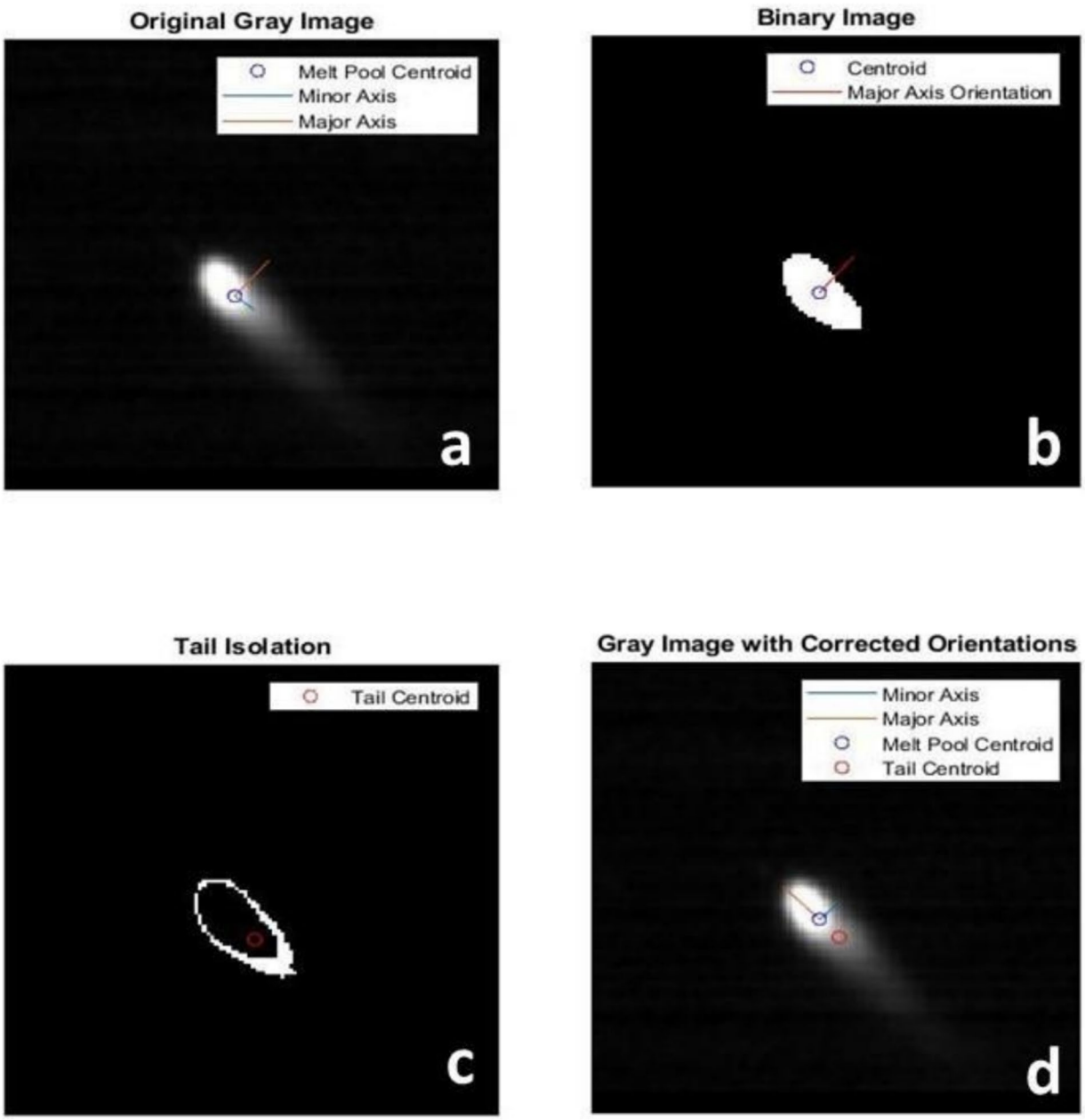

Fig. 5 Process of image feature extraction: **a** original gray image with Matlab estimated orientation, **b** binarized image of pixels > ISO 50, **c** binarized image to isolate melt pool *tail*, **d** new estimated orientation superimposed upon original gray image

to train the k-NN. Once the model has been tested and evaluated, the performance is recorded and the model is replaced with the next. Bin *i* out of parameter *k* is chosen as the new testing set and this process is repeated until *k* models have been trained, tested, and evaluated. *k*-values of 5 and 10 are typical numbers of folds, and a value of 5 folds was chosen as computational time remained a priority. With variations in scan strategies near the shell compared to the infill, biases of higher energy density are likely to occur near the beginning and end. The k-fold cross validation was used in this case to improve reliability across variations in scan strategy.

This process was also used to initially establish the *k-value* of the k-NN model, or number of neighbors considered, due to its higher reliability. During this initial testing, it was found that when *k-value* = 9, the model produced results with the lowest margin of error. Methodical analysis to optimize the *k-value* could be investigated for this application in future work, but as this is a proof-of-concept, it was not within the scope of this work. Matlab's

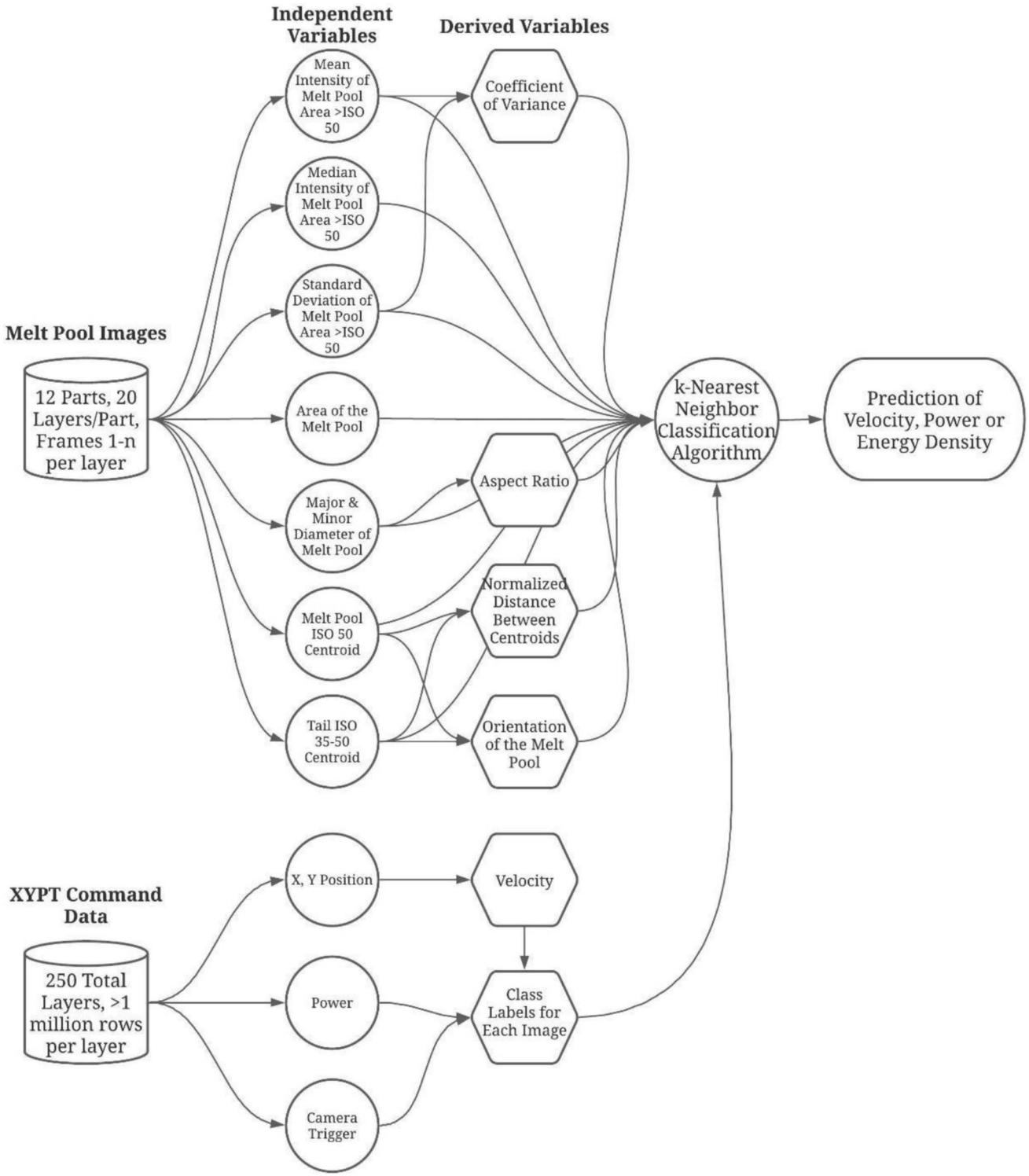

Fig. 6 Flow chart of information throughout the prediction process

fitcknn function was used to establish the model using the *seuclidean* distance metric and a *k-value* of 9. This distance metric uses Euclidean distance, but scales the distance from neighbors and the query point by the standard deviation of the array, removing the need for a separate normalization function.

The model was also tested for transferability across the different parts, as the future work in this area would ideally

use transfer learning to predict process parameters of any part printed using L-PBF. A different path was taken for each part, and their general descriptions can be seen in Table 1.

4 Results and discussion

The results of the trained k-NN on the process command and monitoring data are presented in this section. The dataset was divided up based on part number (1–12) and then further divided for the training/test split. With a usable dataset per part of ~ 30,000 images, the relevant features were derived and saved into the training matrix with 15 columns of independent variables and 1 column of class labels. The matrix was then normalized so that each class is equally represented, as without such post-processing, the algorithm may be more likely to predict lower energy density as that class label would be most prevalent.

After normalization, a typical training matrix had a length of 4000 rows, which was a significant reduction (87%) from the identified training dataset. This training matrix was input into the testing program, and each new image found in the applicable bounds was assigned an estimate of energy density. A testing matrix could be then saved from those estimates, and error was generated by finding the percent difference between the predicted class label and actual energy density value.

4.1 Feature extraction results

The first step in this process, extracting features from the images, played a crucial role in building this low-bandwidth training matrix. Some of these features represent commonly used ones, such as the features that describe a fit ellipse [31], and others were introduced in this work to include additional information about the motion of the melt pool as captured by the tail features. The histograms of the training matrix variables, including the features of the image data and their respective class names, are displayed in Fig. 7. The distribution of independent variables, including normal, bimodal, or skewed distributions, provided context for the relationship between these variables and their respective class names. Using extracted features allows for smaller models to be used and does not require high-resolution images as in some deep learning methods [25].

4.2 Inflection points and transitions

The transition points between scan strategies in the various parts were studied to understand the potential sources of error in the results. Results from parts #2 and #9 are presented here as examples of the patterns seen throughout the entire dataset. This includes both expected behavior and confounding outliers. In Fig. 8, the change in the melt pool area over a 0.1-s period in the build when the laser power was increased is plotted. This transition from 100 to 195 Watts instigated an immediate directly proportional response from the resultant melt pool area increased from 112 to 238 pixels squared. It is also important to note that scan velocity did not significantly vary over these periods during the build, and therefore, the energy density scaled with laser power.

Part #2 however employs a simple serpentine island scan strategy, which is more predictable than others, such as the alternating power concentric island strategy used by part #9. Shown in Fig. 9, the standard laser power transition for this part is more difficult to observe from studying the melt pool area. The difference in bias between the two sections is only 50 pixels squared, a 60% decrease from the typical shift in part #2. This increases the difficulty a prediction algorithm would have if only fed the melt pool area but melt pool area is not the only indication of a change in energy density. This emphasizes the advantage of using a machine learning algorithm, as the relationship

Table 1 Scan strategy descriptions

Part #	Scan strategy
1	Standard rectilinear
2	Island
3	Concentric island
4	Island with alternate starting quadrant
5	Island with higher vertical shell laser power
6	Concentric island with higher vertical shell laser power
7	Island with fewer vertical shells
8	Island with fewer vertical shells and alternate starting quadrant
9	Concentric island with alternating power
10	Concentric island with alternating island and concentric island interleaved
11	Interleaved island
12	Interleaved concentric island

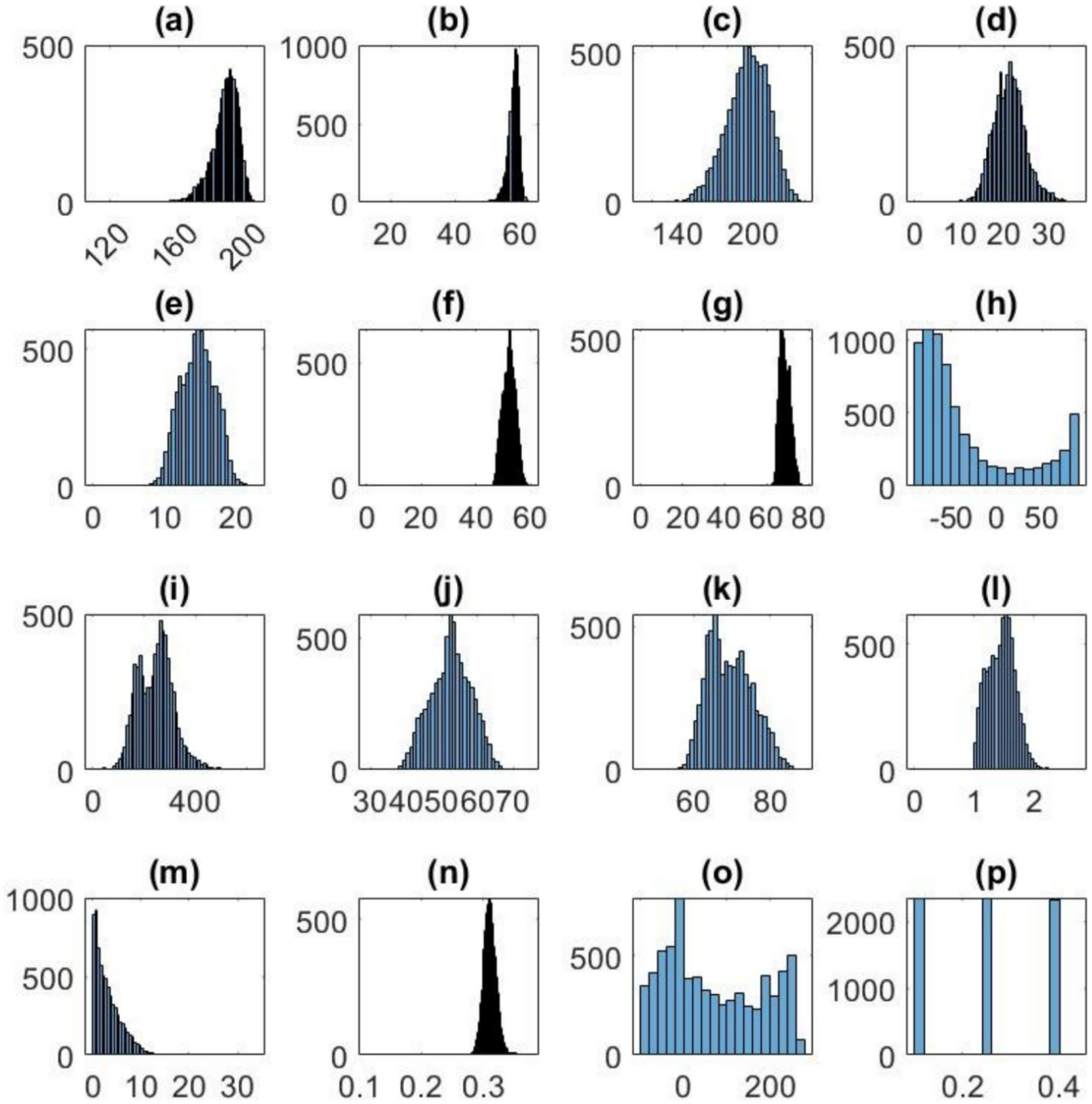

Fig. 7 Elements of the training matrix: **a** average image intensity, **b** standard deviation of image intensity, **c** median image intensity, **d** major axis length, **e** minor axis length, **f** centroid X position, **g** centroid Y position, **h** image processing toolbox's regionprops estimate of orientation, **i** melt pool area, **j** tail centroid X position, **k** tail centroid Y position, **l** aspect ratio, **m** distance between centroid and tail, **n** coefficient of variance, **o** estimated orientation, **p** energy density class labels

between the process parameters and aspects of the images is complicated by additional physical factors. For example, the physics of thermal energy retention could be influencing the lack of melt pool area response, as the powder bed absorbs the extra thermal energy when the laser does not cycle on and off, such as in part #2.

Patterns, such as changes in melt pool size with shifts in power, can be noted for future prediction, and they will be the features that influence the weighting scheme in ML algorithms like the k-NN approach. Since this work is focused on completing the control loop by estimating process parameters, the overlap in the spans in melt pool area

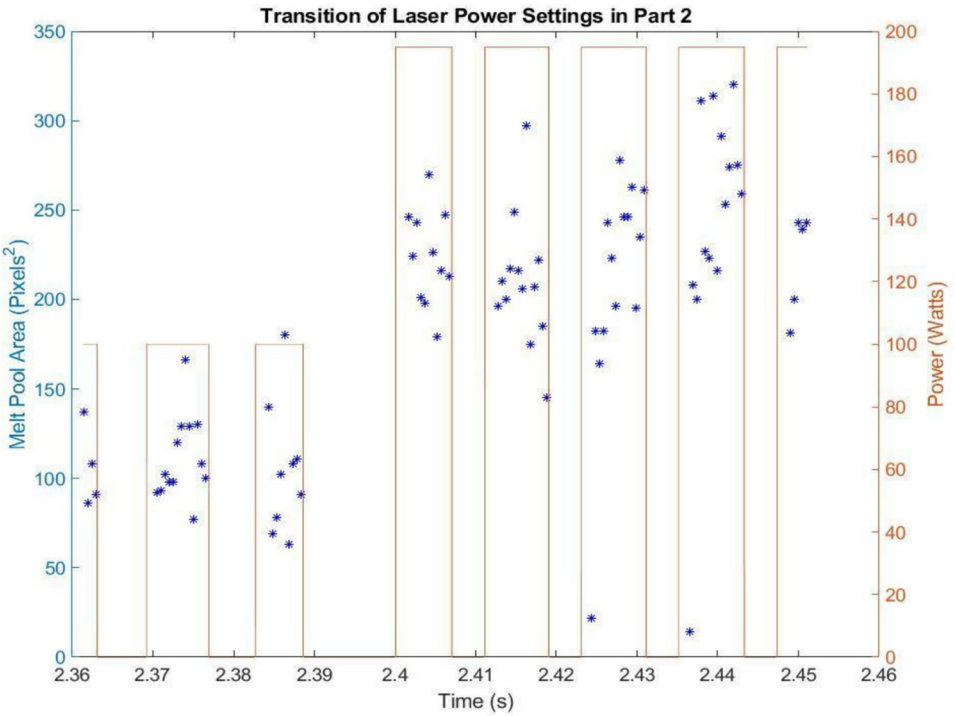

Fig. 8 Typical delineation between power settings and melt pool areas over time for part #2

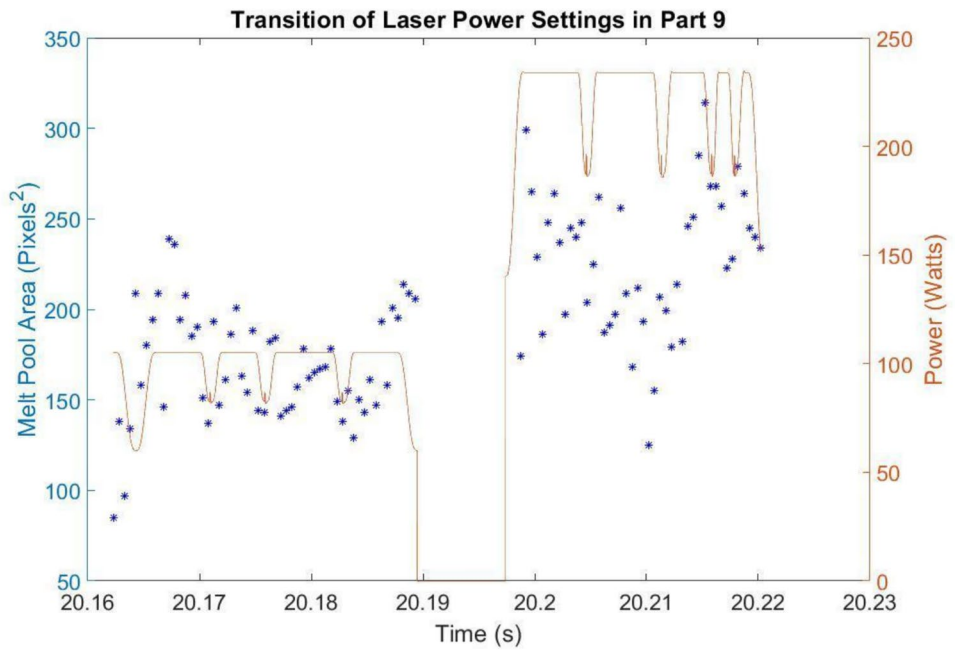

Fig. 9 Typical delineation between power settings and melt pool areas over time for part #9

make it tough for the algorithm to be perfectly accurate, but it does indicate a general trend that should match up with the programmed laser power and scan velocity. This lack of precision is also exacerbated by phenomena occurring at the edge cases where the parameters do not match up with the in situ monitoring data.

In Fig. 10, the unique pattern of drastic increases in melt pool area can be seen over the course of 0.07 s during the build of part #9. This pattern occurs regularly for parts with concentric island scan strategies as the laser moves in a spiral motion towards the center of each island. The cause for this spike in melt pool area without changing energy density could fall on the increase in thermal energy being retained when the Euclidean distance between the current melt pools and recently melted powder continues to be small. This concentration of thermal energy does indicate an increase in

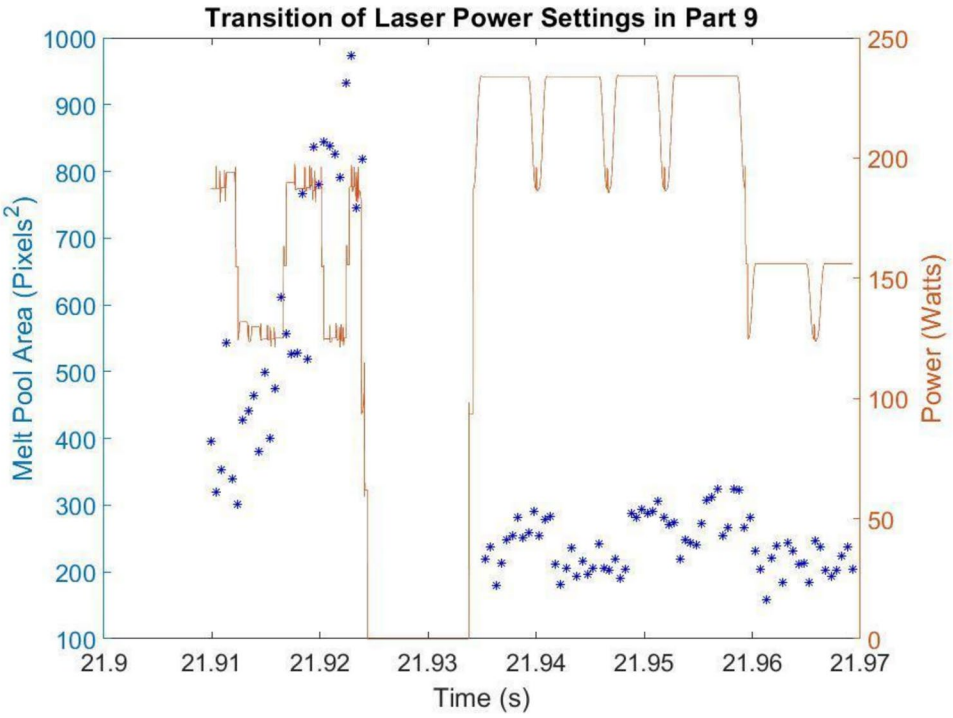

Fig. 10 Atypical delineation between power settings and melt pool areas over time for part #9

Table 2 Layer comparison

Training part: layer number	Testing part: layer number	Mean accuracy
1 – Standard: 2–188	1 – Standard: 189–250	82.3%
2 – Island: 2–188	*2 – Island: 189–250*	*66.2%*

energy density within the part at this point, but it is no longer monotonically related based on process parameters. Therefore, this type of behavior has been classified in this work as atypical, and the algorithm does not attempt to predict such occurrences in the future.

4.3 k-NN layer comparison results

The simplest scan strategies represented by part #1 and #2 were the first studied for validation of the model. Table 2 displays the results of a k-NN ($k = 5$) employed on parts with a split between training and test sets at layer 188. From the table, it is clear that training on a subset of data from a given part provides an effective model for classifying other data from that same part. This initial check evaluates the viability of predicting other layers' energy density values within the same part.

4.4 k-NN part comparison results

Examples of the various scan strategies can be seen in Fig. 11. The standard scan strategy is a vertical serpentine pattern with a constant hatch spacing, and the laser scans from one side of the part across the entire depth to the other side. In contrast, the island technique divides the part into smaller identical "islands," typically four, and there, the laser would only scan halfway across the depth and then move orthogonal to the nominal scanning direction before scanning back to the edge. The serpentine strategy is replaced in part #3 by the "concentric" style, which begins near an edge of the part and moves inward with a centering spiral pattern. Different variations of these basic strategies and their permutations make up the remaining parts.

Using parts #1–3 as the pool for training sets, the variations from similar parts were studied for transferability of the k-NN model. First, the interchangeability between the training set parts was tested, and the experiment was organized with enough permutations so that each subsequent part tested could be related back to any of the training parts. The scan strategies at the core of each pair of parts compared were built upon the same foundation, as parts #2 and #4 both use serpentine island techniques, but part #4 starts with a different island or quadrant.

The training/testing pairs and their respective errors are presented in Table 3. For each row in the table, the k-NN model was trained on the training part number indicated in column 1, and the trained model was evaluated by quantifying its performance on the part number indicated in the second column. The strategy used for that part's manufacture is copied from Table 1 alongside part number for reference. The reason for doing this is to evaluate the transferability of the model between parts. The model's predictions were

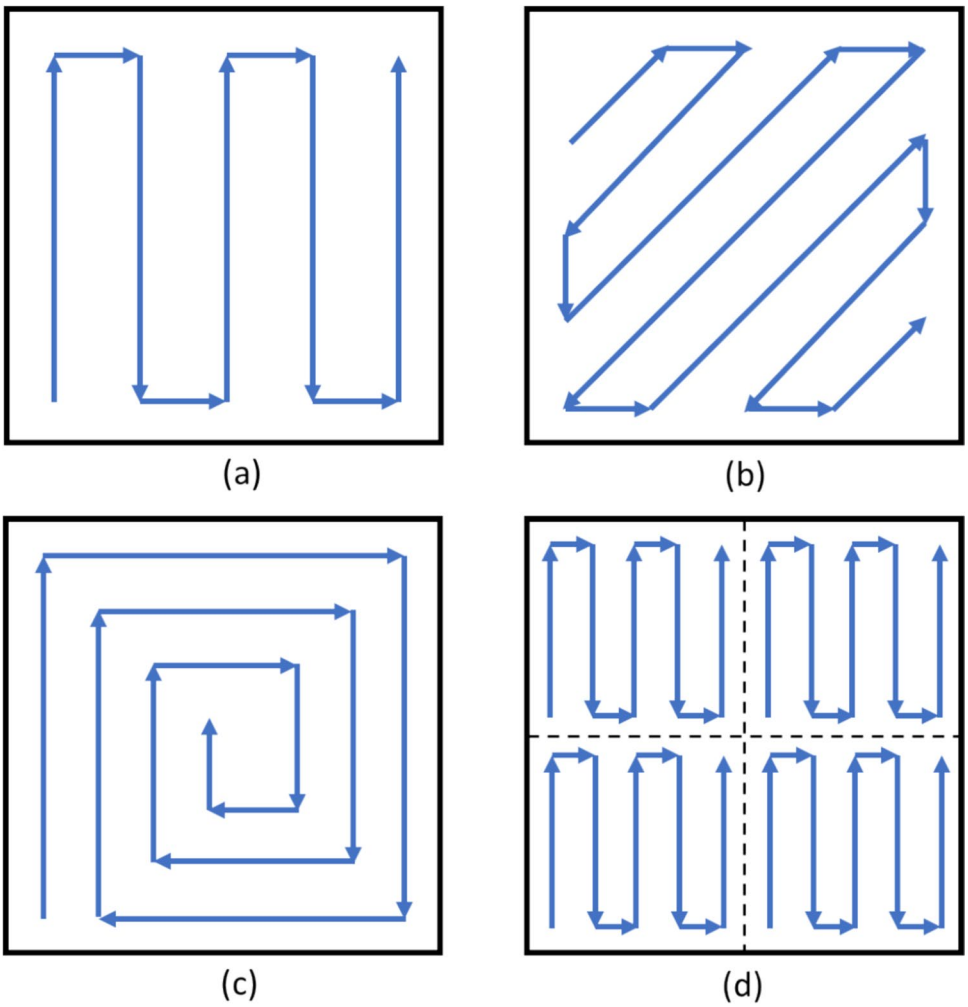

Fig. 11 Examples of different scan strategies: **a** standard/serpentine, **b** off-axis serpentine, **c** concentric, **d** serpentine island

Table 3 Part comparison

Training part – strategy Layers 1–250	*Testing part – strategy* Layers 1–250	Mean accuracy	*F*-score
1 – Standard	2 – Island	65.5%	0.79
2 – Island	3 – Concentric Island	65.5%	0.79
1 – Standard	3 – Concentric Island	65.3%	0.79
2 – Island	4 – Alt. Start Island	64.0%	0.78
2 – Island	5 – High Power Island	62.2%	0.77
3 – Concentric Island	6 – High Power Con. Island	59.8%	0.75
3 – Concentric Island	*9 – Alt. Power Con. Island*	*91.4%*	*0.95*

evaluated against the XYPT command data, and the mean accuracy and *F*-score are presented in column 3.

The transferability of the model between various scan strategy differences was studied to produce the above data, as each part (#1–12) has consistent geometry. The training part pool, parts #1–3, was initially investigated to establish a baseline set of results for the three simplest scan strategies that were used. For three unique paths to compare, three permutations of the train/test split were established and executed to achieve connections between each baseline part from another. An average misclassification rate of 34.5% was reached with the combinations in the first three rows of Table 3, and this reduced error was only matched in the future comparisons with parts #2 and #4. As more dissimilar parts or more complex scan strategies were compared, the error was observed to significantly increase in probability. As mentioned, training part #4 on a test matrix generated by part #2 produced similar error to that of the baseline part

Table 4 Results of training on part #3 and testing on part #9

		Predicted category		
		Low	Medium	High
Actual category	Low	2520	103	27
	Medium	69	1219	219
	High	16	110	1381

pool due to the only difference originating from the alternating starting quadrants. Otherwise, the features that the model could learn from were discovered in the testing part data as well, which resulted in fewer misclassifications. Inversely, the proportion of features possibly mapping from parts #3 to #6 decreases due to the exponential increase in thermal absorption and subsequently thermal energy density with a concentrically focused scan strategy. The inflection points mentioned in Sect. 4.2 are critical here in the part comparison, as the introduction of more variability in the strategy, as in part #9 for example, will reduce the possibility for the model to project trends from the simpler scan strategy. The overall higher error with pairs (2,5), (3,6), and (3,9) seems to be consistent with more of these fluctuations in laser power.

Table 4 illustrates the context behind the error presented in Table 3, as each predicted energy density is recorded in relation to the actual energy density. Rigid class labels were defined to simplify the modeling process so that the algorithm only estimates the best class for a sample as was done in [1]. The range of energy density class labels was categorized into three main groups: low (energy density of 0.115 $\pm$0.01 J/mm), medium (energy density of 0.245 $\pm$0.01 J/mm), and high (energy density of 0.39 $\pm$0.01 J/mm).

When broken out into categories, such as in Table 4, it was observed that the algorithm performed consistently well when the XYPT commands indicated a low energy density, as the algorithm predominately estimated the low category with 91% accuracy. In contrast, actual densities in the medium category resulted in an average ratio of correct predictions near 68%. The majority of melt pools fell into this category, and it is reasonable to assume this category would have the largest error with the edge cases introduced and large variation in melt pools, which resulted in similar average features distributed normally. This issue may be addressed in future work if more data can be collected to describe the behavior caused by changes in energy density. Different models could also be generated for different sections of the part, such as separating perimeters from the infill.

Finally, the high category for energy density was estimated correctly 75% of the time. This variation in accuracy is interesting as it highlights a challenge with this method: melt pools may present as higher or lower density based on their surroundings and position within the scan strategy. Quick, low power melt pools characterized by low energy density could be easier to spot with elongated, ovular shapes. On the other hand, when the tool speed decreases and energy density proportionally grows in magnitude, the surrounding heat becomes more of a factor as contact time increases. Melt pools may appear hotter than was planned in the scan strategy. However, for the purposes of establishing a proof of concept algorithm, these accuracies are promising enough to move forward with integration into a closed loop AM system for further study. More time-consuming analysis can be done after the build is completed if necessary, and future work could even reconcile ML predictions with physics-based simulations. This comparison demonstrates a capability for the k-NN algorithm to accurately transfer between similar parts with a mean accuracy of 68%.

4.5 Time usage analysis

The average time needed to complete each image prediction and layer analysis is shown in Tables 5 and 6. Mean time per image or layer represents the average time across just the testing matrices, including the time required to parse through the data and construct those matrices. Sources of error were explored, and the relationship between error and average time required was studied.

From these results, it was observed that testing the algorithm that was trained on serpentine scan strategies generally required less time, an average of 64 s per layer and 12.5 ms per image. In contrast, the concentric island scan strategies encourage higher computation times, as training on part 3 resulted in an average increase of 6.3 ms per image, which translated to a jump in processing time of 54 to 118 s needed to evaluate each layer.

These trends held true in the layer wise comparison, as training on a serpentine island (part #1 or #2) and testing on part #3 resulted in a mean increase of 16 s from the standard training results. The effect concentric scan strategies had on the computation time could be attributed to the irregularity that these techniques introduce with respect to the energy

Table 5 Image processing time comparison

Training part – strategy Layers 1–250	Testing part – strategy Layers 1–250	Mean time/image (ms) $\pm\sigma$
1 – Standard	2 – Island	11.9 $\pm$3.0
2 – Island	3 – Concentric Island	13.4 $\pm$4.0
1 – Standard	3 – Concentric Island	12.1 $\pm$2.5
2 – Island	4 – Alt. Start Island	12.6 $\pm$3.2
2 – Island	5 – High Power Island	13.2 $\pm$3.0
3 – Concentric Island	6 – High Power Con. Island	19.6 $\pm$3.0
3 – Concentric Island	*9 – Alt. Power Con. Island*	*19.5 $\pm$3.0*

Table 6 Layer time comparison

Training part – strategy Layers 1–250	Testing part – strategy Layers 1–250	Mean time/layer (s) $\pm\sigma$
1 – Standard	2 – Island	39.8 ± 1.80
2 – Island	3 – Concentric Island	84.3 ± 13.7
1 – Standard	3 – Concentric Island	76.7 ± 11.9
2 – Island	4 – Alt. Start Island	41.0 ± 2.09
2 – Island	5 – High Power Island	58.0 ± 13.6
3 – Concentric Island	6 – High Power Con. Island	121.7 ± 18.5
3 – Concentric Island	*9 – Alt. Power Con. Island*	*116.1 ± 15.0*

density build up surrounding the center of the inward spiral path. If the algorithm failed to correctly predict the energy density classification, no change in the analysis time per image could be observed on average. The computing time was also observed to be higher during trials where scan strategy greatly differed, which is expected if the model is having difficulty placing new values that are far from the established class labels.

This work was performed on a desktop computer with an octo-core CPU with a base clock of 3.60 GHz and a GPU with a boost clock of 1.785 GHz and 8 GB of GDDR6 memory. No parallel processing techniques were employed and Matlab was the only program running at the time of the study. Currently, Matlab is not meant to be run directly on a Raspberry Pi or similar microcomputer, but this algorithm could be performed remotely and receive the live data over the internet from such a device to keep a low footprint in the machine. This layout could provide an opportunity to increase computing power, thus lowering time to process the image and opening a time window to decide an improvement on the scan strategy in real-time.

4.6 Correlation of model accuracy and melt pool location

To understand how one might improve this method in future work, the relationship between melt pool location and prediction accuracy was studied. This investigation may provide context to the prior results, as energy density tends to follow patterns relative to geometry. For example, scan strategies typically reduce speeds at the part perimeters to increase the homogeneity of the surface, which increases the energy density if power is held constant.

This can be seen in Fig. 12, where the scan strategy of layer 98 in part 2 is shown and higher energy density is represented by brighter colors. Higher energy densities can also be seen where the edges of each quadrant or island within the part are outlined. The speed is increased elsewhere for infill, which pushes the energy density downward, hence the darker colors. In Fig. 13, the predictions of the k-NN algorithm can be seen plotted in their respective physical

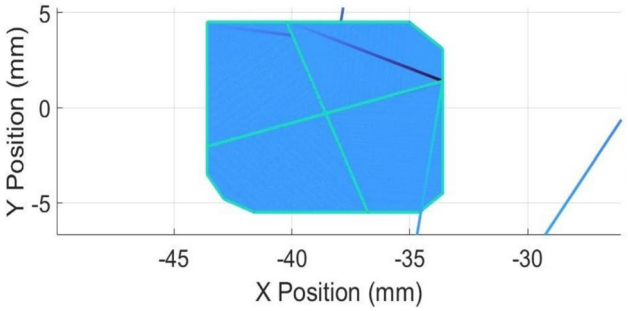

Fig. 12 Scan strategy for part #2, layer 98

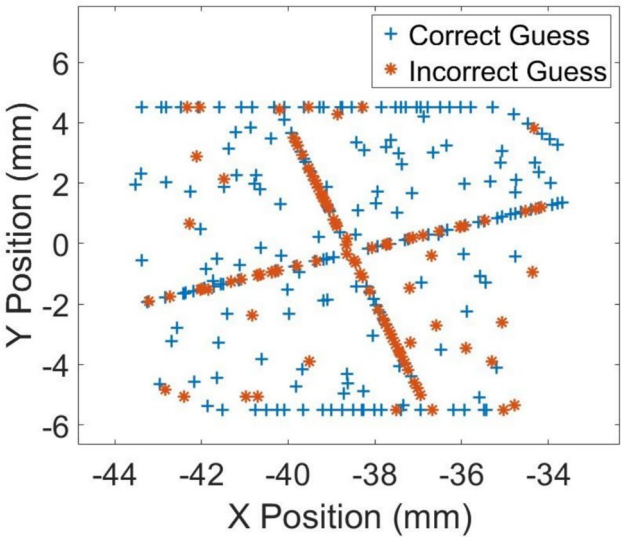

Fig. 13 Error scatter plot for layer 98 of prediction of part #2 with k-fold validation

locations. For reference, the overall accuracy of the ML prediction for this layer was calculated to be 82.6%. The main source of error appears to follow the contours of the inner vertical shells, which aligns with the results presented in the confusion matrices, as category 1 predictions were more accurate than categories 2 and 3, where energy density is higher.

4.7 Assumptions

For clarity, the assumptions made in this work are detailed in this section. A critical assumption was made to estimate velocity around the point of the image capture. It was assumed that the XYPT data accurately represented the discrete position of the laser, and traveled distance could be interpreted from the surrounding cluster of points by plotting a line of best fit through those XY positions and assuming the length of the best fit line segment to be the average distance through the path of the laser. That average distance was divided by the average time, represented by the number of passed time intervals multiplied by the time step of 10 μs. This provided the energy density equation with the average velocity component, and the power of the laser was evaluated instantaneously by using the index of the image capture.

5 Conclusion

In this work, a novel application of k-NN modeling to predict process parameters in laser powder bed fusion has been presented. The k-NN was trained on in situ monitoring images and build command data from the Additive Manufacturing Metrology Testbed at NIST, which contains 12 parts of identical geometry but unique scan strategies. Before input into the model, each image was characterized by 15 parameters, which not only reduced computing time but also enabled the inclusion of a greater number of images in the training/testing matrix split.

The model was initially validated by studying interlayer prediction within the same part, with average accuracy ranging from 66.2 to 82.3%. This performance is an improvement upon other image classification models using k-NN algorithms [14, 29].

This approach was then used to investigate transferability between parts of differing scan strategies, as training matrices were used to predict the energy density of images taken of a different part. A lower mean accuracy was expected with this experiment, and as such, this ranged from 59.8 to 91.4%. The deviation increased as the parts became more dissimilar, and especially as the process parameters of the test parts began to fall outside of the bounds of the identified classes.

Finally, a k-fold cross validation was performed to standardize the evaluation of this modeling application. Then, a scatter plot of error relative to the physical position was generated to understand the main source of misclassifications made by the algorithm. From this scatter plot, it was observed that higher error followed the vertical shells of the scan strategy, and conversely predictions of the infill sections were more accurate. This could be correlated in future work to the increase in error related to higher energy density classes.

This work demonstrated the capability of k-NN in predicting the process parameters within additive manufacturing applications, and that these approaches are especially robust when changes to scan strategy are minimal. It also plays an important role in demonstrating the possibility of using ML to predict these relationships instead of physics-based modeling. Opening up the possibilities for closed-loop control in the metal additive manufacturing field allows for more choice and competition, which encourages more widespread adoption of laser powder bed fusion, satisfying an overarching goal of this work. Without more basic research needed, the ML approach can improve greatly over time as more data becomes available from the application of this advanced manufacturing technique, and that combined with the open-source nature of these algorithms implies a promising future in store for low-cost quality assurance of powder bed fusion.

Author contribution All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by Patrick Jung and Nathaniel DeVol. The first draft of the manuscript was written by Patrick Jung toward the requirements for his master's thesis degree. All authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding This work was supported in part by the U.S. Department of Energy DE-EE0008303.

Declarations

Competing interests The authors declare no competing interests.

References

1. Akhavan J, Lyu J, Manoochehri S (2023) A deep learning solution for real-time quality assessment and control in additive manufacturing using point cloud data. J Intell Manuf. https://doi.org/10.1007/s10845-023-02121-4
2. Aminzadeh M, Kurfess TR (2019) Online quality inspection using Bayesian classification in powder-bed additive manufacturing from high-resolution visual camera images. J Intell Manuf 30(6):2505–2523. https://doi.org/10.1007/s10845-018-1412-0

3. Bi G, Gasser A, Wissenbach K, Drenker A, Poprawe R (2006) Characterization of the process control for the direct laser metallic powder deposition. Surf Coatings Technol 201(6):2676–2683. https://doi.org/10.1016/j.surfcoat.2006.05.006
4. Chomboon K, Chujai P, Teerarassammee P, Kerdprasop K, Kerdprasop N (2015) An empirical study of distance metrics for k-nearest neighbor algorithm. Proceed 2nd Int Conf Indust Applic Eng 2015:280–285
5. Criales LE, Arısoy YM, Lane B, Moylan S, Donmez A, Özel T (2017) Laser powder bed fusion of nickel alloy 625: experimental investigations of effects of process parameters on melt pool size and shape with spatter analysis. Int J Mach Tools Manuf 121:22–36. https://doi.org/10.1016/j.ijmachtools.2017.03.004
6. Cunningham R, Narra SP, Montgomery C, Beuth J, Rollett AD (2017) Synchrotron-based X-ray microtomography characterization of the effect of processing variables on porosity formation in laser power-bed additive manufacturing of Ti-6Al-4V. JOM 69:479–484. https://doi.org/10.1007/s11837-016-2234-1
7. Dong L, Makradi A, Ahzi S, Remond Y (2009) Three-dimensional transient finite element analysis of the selective laser sintering process. J Mater Process Technol 209(2):700–706. https://doi.org/10.1016/j.jmatprotec.2008.02.040
8. F42.91, A. C (2012) Standard terminology for additive manufacturing technologies. ASTM Int. https://doi.org/10.1520/F2792-12
9. Frazier WE (2014) Metal additive manufacturing: a review. J Mater Eng Perform 23:1917–1928. https://doi.org/10.1007/s11665-014-0958-z
10. Gobert C, Reutzel EW, Petrich J, Nassar AR, Phoha S (2018) Application of supervised machine learning for defect detection during metallic powder bed fusion additive manufacturing using high resolution imaging. Addit Manuf 21:517–528. https://doi.org/10.1016/j.addma.2018.04.005
11. Grasso M, Colosimo BM (2017) Process defects and in situ monitoring methods in metal powder bed fusion: a review. Meas Sci Technol 28(4):44005. https://doi.org/10.1088/1361-6501/aa5c4f
12. Guo Q, Zhao C, Qu M, Xiong L, Escano LI, Hojjatzadeh SMH, Parab ND, Fezzaa K, Everhart W, Sun T, Chen L (2019) In-situ characterization and quantification of melt pool variation under constant input energy density in laser powder bed fusion additive manufacturing process. Additive Manuf 28:600–609. https://doi.org/10.1016/j.addma.2019.04.021
13. Gusarov AV, Yadroitsev I, Bertrand Ph, Smurov I (2007) Heat transfer modelling and stability analysis of selective laser melting. Applied Surface Science 254(4):975–979. https://doi.org/10.1016/j.apsusc.2007.08.074
14. Hasanlou M, Samadzadegan F (2012) Comparative study of intrinsic dimensionality estimation and dimension reduction techniques on hyperspectral images using K-NN classifier. IEEE Geosci Remote Sens Lett 9(6):1046–1050. https://doi.org/10.1109/LGRS.2012.2189547
15 Herzog D, Seyda V, Wycisk E, Emmelmann C (2016) Additive manufacturing of metals. Acta Materialia 117:371–392. https://doi.org/10.1016/j.actamat.2016.07.019
16. Hussein A, Hao L, Yan C, Everson R (2013) Finite element simulation of the temperature and stress fields in single layers built without-support in selective laser melting. Mater Design (1980-2015) 52:638–647. https://doi.org/10.1016/j.matdes.2013.05.070
17. Khairallah SA, Anderson AT, Rubenchik A, King WE (2016) Laser powder-bed fusion additive manufacturing: physics of complex melt flow and formation mechanisms of pores, spatter, and denudation zones. Acta Mater 108:36–45
18. King WE, Anderson AT, Ferencz RM, Hodge NE, Kamath C, Khairallah SA, Rubenchik AM (2015) Laser powder bed fusion additive manufacturing of metals; physics, computational, and materials challenges. Appl Phys Rev 2(4):41304. https://doi.org/10.1063/1.4937809
19. Ko H, Witherell P, Ndiaye NY, & Lu Y (2019) Machine learning based continuous knowledge engineering for additive manufacturing. *2019 IEEE 15th Int Conf Autom Sci Eng (CASE)*, 648–654. https://doi.org/10.1109/COASE.2019.8843316
20. Kruth JP, Badrossamay M, Yasa E, Deckers J, Thijs L, & Van Humbeeck J (2010) Part and material properties in selective laser melting of metals. *Proceedings of the 16th International Symposium on Electromachining (ISEM XVI)*, 3–14
21. Lane B, Yeung H (2019) Process monitoring dataset from the additive manufacturing metrology testbed (AMMT):"three-dimensional scan strategies." J Res Nat Inst Stand Technol 124:1–14. https://doi.org/10.6028/jres.124.033
22. Leitz K-H, Grohs C, Singer P, Tabernig B, Plankensteiner A, Kestler H, Sigl LS (2018) Fundamental analysis of the influence of powder characteristics in Selective Laser Melting of molybdenum based on a multi-physical simulation model. Int J Refractory Metal Hard Mater 72:1–8. https://doi.org/10.1016/j.ijrmhm.2017.11.034
23. Letenneur M, Brailovski V, Kreitcberg A, Paserin V, Bailon-Poujol I (2017) Laser powder bed fusion of water-atomized iron-based powders: process optimization. J Manuf Mater Process 1(2):23. https://doi.org/10.3390/jmmp1020023
24. Letenneur M, Kreitcberg A, Brailovski V (2019) Optimization of laser powder bed fusion processing using a combination of melt pool modeling and design of experiment approaches: density control. J Manuf Mater Process 3(1):21. https://doi.org/10.3390/jmmp3010021
25. Li X, Jia X, Yang Q, Lee J (2020) Quality analysis in metal additive manufacturing with deep learning. J Intell Manuf 31(8):2003–2017. https://doi.org/10.1007/s10845-020-01549-2
26. Liu J, Ye J, Silva Izquierdo D, Vinel A, Shamsaei N, Shao S (2023) A review of machine learning techniques for process and performance optimization in laser beam powder bed fusion additive manufacturing. J Intell Manuf 34(8):3249–3275. https://doi.org/10.1007/s10845-022-02012-0
27. Liu YJ, Li SJ, Wang HL, Hou WT, Hao YL, Yang R, Sercombe TB, Zhang LC (2016) Microstructure, defects and mechanical behavior of beta-type titanium porous structures manufactured by electron beam melting and selective laser melting. Acta Materialia 113:56–67. https://doi.org/10.1016/j.actamat.2016.04.029
28. Lu, Y., Choi, S., & Witherell, P. (2015). *Towards an integrated data schema design for additive manufacturing: conceptual modeling. 1A:*, V01AT02A032. https://doi.org/10.1115/DETC2015-47802
29. Ma L, Crawford MM, Tian J (2010) Local manifold learning-based k-nearest-neighbor for hyperspectral image classification. IEEE Trans Geosci Remote Sens 48(11):4099–4109. https://doi.org/10.1109/TGRS.2010.2055876
30. Mani, M., Feng, S., Brandon, L., Donmez, A., Moylan, S., & Fesperman, R. (2017). Measurement science needs for real-time control of additive manufacturing powder-bed fusion processes. In *Additive manufacturing handbook* (629–652). CRC Press. https://doi.org/10.1201/9781315119106-34
31. Mitchell JA, Ivanoff TA, Dagel D, Madison JD, Jared B (2020) Linking pyrometry to porosity in additively manufactured metals. Addit Manuf 31:100946. https://doi.org/10.1016/j.addma.2019.100946
32 NIST (2013) Measurement science roadmap for metal-based additive manufacturing. Department of Commerce, US
33. Shiomi M, Yoshidome A, Abe F, Osakada K (1999) Finite element analysis of melting and solidifying processes in laser rapid prototyping of metallic powders. Int J Mach Tools Manuf 39(2):237–252. https://doi.org/10.1016/S0890-6955(98)00036-4

34. Suzuki A, Miyasaka T, Takata N, Kobashi M, Kato M (2021) Control of microstructural characteristics and mechanical properties of AlSi12 alloy by processing conditions of laser powder bed fusion. Addit Manuf 48:102383. https://doi.org/10.1016/j.addma.2021.102383
35. Taherkhani K, Ero O, Liravi F, Toorandaz S, Toyserkani E (2023) On the application of in-situ monitoring systems and machine learning algorithms for developing quality assurance platforms in laser powder bed fusion: a review. J Manuf Process 99(May):848–897. https://doi.org/10.1016/j.jmapro.2023.05.048
36. Tapia G, Elwany A (2014) A review on process monitoring and control in metal-based additive manufacturing. J Manuf Sci Eng 136(6):60801. https://doi.org/10.1115/1.4028540
37. Witherell P, Feng S, Simpson TW, Saint John DB, Michaleris P, Liu Z-K, Chen L-Q, Martukanitz R (2014) Toward metamodels for composable and reusable additive manufacturing process models. J Manuf Sci Eng 136(6):61025. https://doi.org/10.1115/1.4028533
38. Xia C, Pan Z, Polden J, Li H, Xu Y, Chen S (2022) Modelling and prediction of surface roughness in wire arc additive manufacturing using machine learning. J Intell Manuf 33(5):1467–1482. https://doi.org/10.1007/s10845-020-01725-4
39. Yang Z, Lu Y, Yeung H, Krishnamurty S (2020) From scan strategy to melt pool prediction: a neighboring-effect modeling method. J Comput Inf Sci Eng 20(5):51001. https://doi.org/10.1115/1.4046335
40. Yeung H, Yang Z, Yan L (2020) A meltpool prediction based scan strategy for powder bed fusion additive manufacturing. Addit Manuf 35:101383. https://doi.org/10.1016/j.addma.2020.101383
41 Yin J, Zhu H, Ke L, Lei W, Dai C, Zuo D (2012) Simulation of temperature distribution in single metallic powder layer for laser micro-sintering. Comput Mater Sci 53(1):333–339. https://doi.org/10.1016/j.commatsci.2011.09.012