Tailoring Laser Powder Bed Fusion Process Parameters for Standard and Off-Size Ti6Al4V Metal Powders: A Machine Learning Approach Enhanced by Photodiode-Based Melt Pool Monitoring

Farima Liravi [1], Sebastian Soo [1], Sahar Toorandaz [1], Katayoon Taherkhani [1], Mahdi Habibnejad-Korayem [2] and Ehsan Toyserkani [1,*]

[1] Multi-Scale Additive Manufacturing (MSAM) Laboratory, Department of Mechanical and Mechatronics Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada; f2liravi@uwaterloo.ca (F.L.); sjxsoo@uwaterloo.ca (S.S.); sahar.toorandaz@uwaterloo.ca (S.T.); ktaherkhani@uwaterloo.ca (K.T.)

[2] AP&C, Colibrium Additive, A GE Aerospace Company, Montreal, QC J7H 1R8, Canada

* Correspondence: ehsan.toyserkani@uwaterloo.ca

Abstract: An integral part of laser powder bed fusion (LPBF) quality control is identifying optimal process parameters tailored to each application, often achieved through time-consuming and costly experiments. Melt pool dynamics further complicate LPBF quality control due to their influence on product quality. Using machine learning and melt pool monitoring data collected with photodiode sensors, the goal of this research was to efficiently customize LPBF process parameters. A novel aspect of this study is the application of standard and off-size powder feedstocks. Ti6Al4V (Ti64) powder was used in three size ranges of 15–53 μm, 15–106 μm, and 45–106 μm to print the samples. This facilitated the development of a process parameters tailoring system capable of handling variations in powder size ranges. Ultimately, per each part, the associated set of light intensity statistical signatures along with the powder size range and the parts' density, surface roughness, and hardness were used as inputs for three regressors of Feed-Forward Neural Network (FFN), Random Forest (RF), and Extreme Gradient Boosting (XGBoost). The laser power, laser velocity, hatch distance, and energy density of the parts were predicted by the regressors. According to the results obtained on unseen samples, RF demonstrated the best performance in the prediction of process parameters.

Keywords: laser powder bed fusion; machine learning; melt pool monitoring; process parameters prediction

1. Introduction

Additive manufacturing (AM) has developed into one of the world's most advanced manufacturing technologies over recent years due to its ability to manufacture parts with highly complex geometries with lightweight features [1]. Laser powder bed fusion (LPBF) is one of the most prominent techniques in metal AM. LPBF involves spreading a thin layer of powder on the build plate, which is then selectively melted with a laser beam [2] according to an operator-defined digital model of the desired part and optimum process parameters such as laser power, laser speed, hatch distance, etc. The subsequent layers of parts are printed on top of each other similarly.

Although LPBF has achieved notable success in industrial applications in the aerospace, health, and automobile fields, it still lacks a comprehensive quality assurance system to satisfy the repeatability and reproductivity criteria of large-scale commercialization. The development of a real-time comprehensive quality assurance system for LPBF is challenging due to the enormous number of quality-affecting factors involved in the process. For instance, feedstock powder properties, process parameters, melt pool dynamics, powder

bed smoothness feature, and environmental parameters are among these influential factors that have a significant impact on part physical and mechanical properties such as density, hardness, and surface roughness.

Current studies propose partially effective quality control methodologies that mainly rely on experimental assessment and data-driven techniques such as machine learning (ML) algorithms. The design of experiments (DoE) is usually developed in such studies to print metal parts with various process parameter sets to obtain parts with different levels of quality in terms of density and surface roughness. In data-driven analysis, the obtained data from these designs are usually used to develop ML models for either prediction/classification of part's properties (output) based on given process parameters (input) or vice-versa. This method has been adopted in many recent scientific studies for the optimization of part properties such as hardness [3,4], density and surface roughness [5–7], or dimensional accuracy control [3,8,9].

ML algorithms have also been practiced for quantitative materials analysis and optimization as well. In the studies carried out by Zhiwei Zheng et al., self-supervised algorithms have been developed for segmentation of the microscopic images across various materials enabling quantitative analysis of the materials [10]. The modified algorithm was then applied to detect and analyze of helium bubbles in 304 L stainless steel printed parts [11]. Estimation of the mechanical properties such as intrinsic strengths of the parts has also been conducted using ML algorithms [12].

In more recent investigations, however, ML-based quality control techniques in LPBF have seen a major shift toward the analysis of datasets collected from the melt pool by different devices such as photodiodes or cameras. This method is referred to as melt pool monitoring (MPM), recognized as one of the leading advanced techniques extensively employed in LPBF [13]. The capability of MPM in capturing melt pool variations which intensely alter the outcome quality of the parts and could signify the flaw formation during the print is an appealing factor for its usage. Accordingly, utilizing the MPM data significantly aids in improving quality-controlling systems in defect detection and process optimizations. Among the current sensor types widely employed in LPBF for data collection (e.g., digital camera, thermal camera, photodiode, and acoustic), photodiodes provide valuable time-series datasets regarding the melt pool condition but in a lower volume of data than imaging sensors. This is beneficial in terms of data processing computation costs and the decision-making pace of ML algorithms [14]. A photodiode records the light intensity irradiated from the melt pool, which is strongly correlated with its temperature [14]. Thus, it can be used to monitor the melt pool variations that could imply important information about the process such as hot spots and cold spots (indicators of defects) and laser–powder interactions. The process's MPM data can then be used to control/adjust the process parameters to improve the quality of printed components. In addition, the light intensity data recorded by photodiode sensors throughout the printing could be adopted as a strong tool for effective and precise LPBF process optimization and quality estimation. Many studies focused on the quality control of LPBF in recent years have used photodiode-based MPM and ML. The main goal of the reviewed studies is mostly defect detection based on MPM data alterations on which the ML model extracts patterns to detect different types of defects. Taherkhani et al. used intensity signals and a Self-Organizing Maps (SOMs) algorithm to detect and localize intentional/seeded and random lack of fusion defects [15]. Zhuang Mao et al. also developed a Long Short-Term Memory Network (LSTM) algorithm that surpassed other tested ML models in the classification of quality (defective or non-defective) based on photodiode signals [16]. Yadav et al. applied photodiode signals and a Support Vector Machine (SVM) to detect the drift anomaly in the powder bed [17].

Quality level estimation based on physical and mechanical properties is also another region of exploration in MPM and ML-based LPBF quality control analysis. For instance, Jayasinghe et al. adopted k-means and Gaussian Mixture models to cluster the photodiode signals and developed a Gaussian Process Regression (GPR) to predict the density of

parts [18]. Cao et al. classified the tracks' melting classes using a Convolutional Neural Network (CNN), based on process parameters and light intensity signals (inputs) [19]. Many similar investigations have also been published in which data fusion and the use of other sensor types in addition to photodiodes are adopted [20–23]. However, in the majority of these studies, either the classification/prediction of parts' properties or the detection and localization of defects were the subject.

Prediction of required process parameters given the user-determined part properties (density, surface roughness, hardness, etc.) could provide an advanced alternative for conventional experimental LPBF process parameter optimization methods. This gains advantages in terms of time and cost of design of the printing process. In a recent investigation, Lapointe et al. used photodiode's intensity signals during the print of tracks to develop a Feed-Forward Neural Network (FFN) to predict the process parameters of tracks based on intensity signal and geometrical features [24]. The results were validated by the improvement of the dimensional accuracy of complex printed parts using the algorithm-predicted process parameters.

In the context of the LPBF process optimization as reviewed in the relevant literature, it is evident that most studies focusing on ML-based optimization systems typically integrate part properties and other physical characteristics, ranging from geometric to mechanical properties, alongside various process parameters. However, this study distinguishes itself in two key aspects: (1) It incorporates a more comprehensive dataset, supported by quantitative and qualitative studies, which have shown significant influence on the LPBF outcome quality, including melt pool behavior and the size distribution of powder feedstock. (2) Unlike similar studies that primarily focus on single or multi-track design of experiments, this research extends the scope of process optimization to encompass a part-scale design of experiments using MPM and ML methodologies.

AM and LPBF, even in their most recent advanced state, continue to encounter major challenges, such as being highly cost-intensive in production, primarily due to the extensive requirements for experimental design development and post-process quality inspection. On a production scale, in these methods, major investments (e.g., USD 100,000+ per material and product) would need to be invested in the process parameters development and experimental post-process quality assessment. The integration of MPM and ML into process optimization is still in its early stages; however, a comprehensive system based on these techniques holds the potential to substantially reduce the costs and time associated with optimization processes. By automating optimization through self-supervised algorithms and real-time data integration in future advancements, such systems could streamline the optimization process significantly. Therefore, although printers equipped with built-in sensors for MPM may entail initial costs, the practicality and effectiveness of ML-based systems are poised to surpass conventional methods.

Moreover, traditional process optimization techniques rely heavily on fully experimental approaches, involving multiple cycles of design of experiments, printing, and destructive quality assessments to address the complex and often nonlinearly interconnected factors influencing quality outcomes. This iterative process demands a substantial budget and time investment to achieve reliable optimization results.

These systems have also demonstrated limitations in incorporating comprehensive factors essential into the process optimization. For example, they often overlook considerations such as melt pool stability, despite its significant impact on quality and the optimization of process parameters. Moreover, traditional experimental and statistical-based techniques struggle to effectively identify and integrate complex, nonlinear relationships between factors influencing quality outcomes. When attempts are made to scale these methods, they frequently result in prohibitively expensive large-scale prints.

These drawbacks stem from the inherent constraints of statistical optimization methods typically used for both optimization and experimental design. It is important to note that ML and MPM-based techniques, including the approach proposed in this study, can potentially be adapted into in situ, non-destructive process optimization systems. In

contrast, traditional experimental methods are typically conducted off situ and involve destructive testing.

Hence, the goal of this study is to develop an ML-based model that can identify optimum process parameters (power, velocity, energy density, and hatch distance) using MPM datasets collected during printing coupons and artifacts to expand on the previously reviewed ML-based LPBF process parameter optimizations. Due to the fast-paced data processing nature of ML algorithms and the availability of powerful yet low-cost platforms for the development and execution of these algorithms, it is possible to offer solutions to the time/cost-intensive process development tailored to regular coupons and complex shape components. Furthermore, the post-process quality inspection of the printed products could effectively be replaced by the effective recording and processing of in situ monitoring datasets, considering the recent rise of LPBF printers with built-in monitoring setups.

Figure 1 shows the interconnected aims of this study where the trained algorithms/ regressors are employed to provide precise estimation of process parameters (power, velocity, energy density, and hatch distance) as the output, based on the input variables of the powder size range, part properties (density, hardness, and surface roughness), and melt pool characteristics (light intensity signature variations). A comparative analysis is also presented between various ML regressors to select the most accurate algorithm for the task.
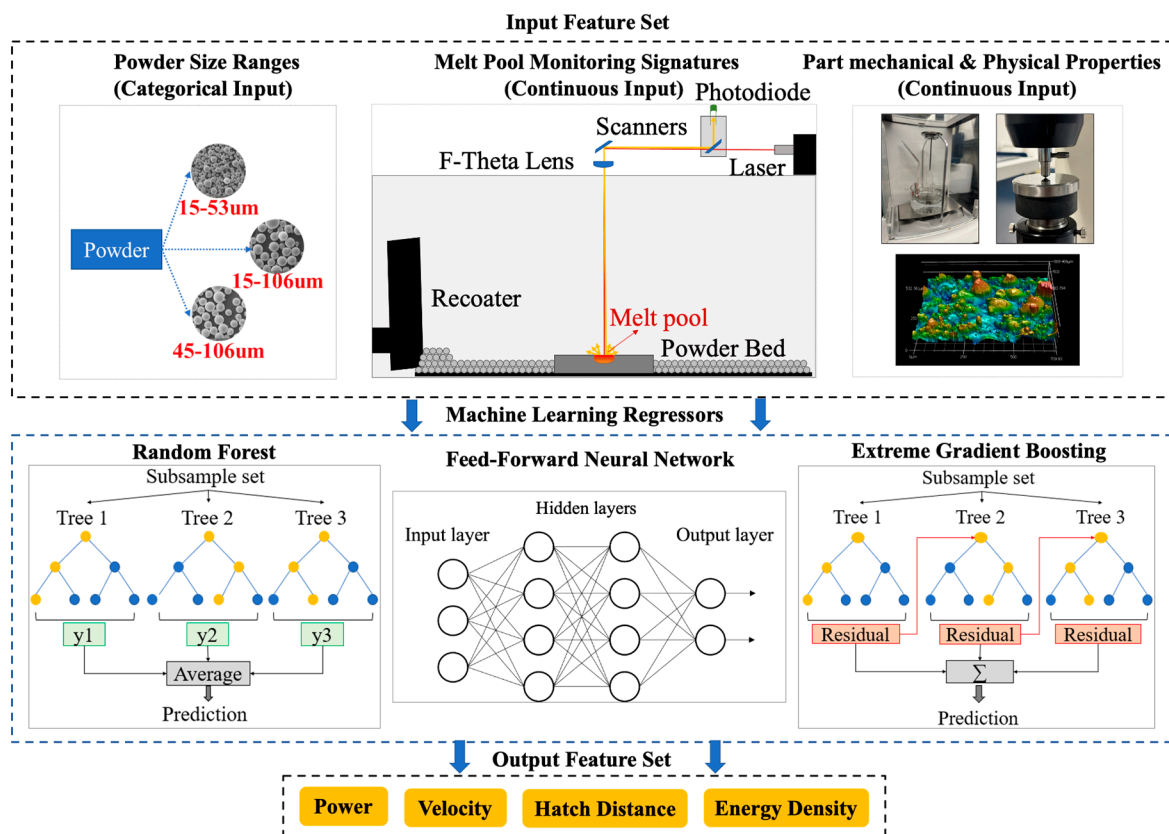

Figure 1. Graphical abstract highlighting the study workflow including application of ML on MPM data, part properties, and different powder particle size distributions for process parameters estimation.

2. Materials and Methods

2.1. Design of Experiment

A commercial LPBF system (EOS M290, EOS GmbH, Krailling, Germany) was utilized for printing Ti64 parts. The powder bed recoating is handled by an in-chamber automatic system. The machine has an Ytterbium (Yb) fiber laser with 400 W nominal power and a laser wavelength of 1060 nm. Argon gas is employed in the chamber for Ti64 parts printing.

Ti64 metal powder (AP&C, Boisbriand, QC, Canada) in three size ranges of 15–53 µm (fine batch/batch A), 15–106 µm (broad batch/batch B), and 45–106 µm (coarse batch/off size/batch C) is used as the feedstock material to print 10 mm × 10 mm × 10 mm cubical and cylindrical samples with a height and diameter of 10 mm × 10 mm (a subset of printed parts is shown in Figure 2) using the Full Factorial Design (FFD) technique.

Figure 2. Printed cubical and cylindrical parts.

Three separate sets of prints were carried out for the printing of cuboid and cylindrical parts of 15–53 µm, 15–106 µm, and 45–106 µm powder batches, i.e., 54 cuboid and 54 cylindrical parts of each powder size range were printed in a single print.

Laser power (W), laser energy density (J/mm^2), hatch distance (mm), and scanning strategy are varied in the ranges listed in Table 1. The layer thickness of 40 µm is kept fixed in all prints. Laser velocity (mm/s) was determined according to specified energy density and hatch distance using Equation (1).

$$Energy\ density = \frac{Laser\ Power}{Laser\ Velocity \times Hatch\ Distance} \tag{1}$$

Parts geometry (cubical/cylindrical) and powder size range are the other changing variables in the design of the experiment.

Table 1. Design of the experiment.

Tagging Class	Power (W)	Energy Density (J/mm^2)	Hatch Distance (mm)	Scan Strategy	
Low	150	1.5	0.08		
Mid	250	2.75	0.09	Stripe	Chess
High	350	4	0.1		

In total, 324 parts were printed which consists of 108 parts (54 cubical and 54 cylindrical samples) printed with each of the three powder size ranges. The same variation of the process parameters is repeated for each set of 108 parts, allowing for inspection of the powder size range changes' influence on the final parts' quality and MPM data alterations.

The density, surface roughness, and hardness of the cuboid and cylindrical parts are measured by Archimedes, Keyence laser profilometer, and Phase II Rockwell hardness tester. The hardness and surface roughness are measured on the vertical sides. Nine points across the side surface of the cuboid and cylindrical parts were measured, controlled for potential noisy dataset elimination, and averaged to obtain the side surface roughness measures per part. Regarding the hardness measurements, the hardness in five points on the side surface of the parts were tested by the Phase II Rockwell equipment. The averaged hardness over the five measured points per part was then used as the hardness input feature of the algorithms.

2.2. Melt Pool Monitoring

Light intensity signals irradiated from the melt pool area were collected by an on-axis photodiode sensor during the printing of 108 parts of each powder batch (each single run of print), as shown in Figure 3. The on-axis photodiode sensor records the visible to near-infrared (NIR) range of light intensity reflected parallel to the laser radiation direction. The on-axially recorded light intensity signals are gathered in the form of raw data containing the laser x-y location on the build plate, the modulation of the laser (on-off logs of the lasers), and the light intensity at each corresponding x-y location. Geometry and light intensity data corrections were applied to the raw MPM data to compensate for the chromatic aberration and position-related errors that appear in the light-intensity data recording due to the placement of the sensor and the difference between the optical elements of monitoring and the light wavelength (for more information about the melt pool monitoring technique and the data corrections, refer to [25]).
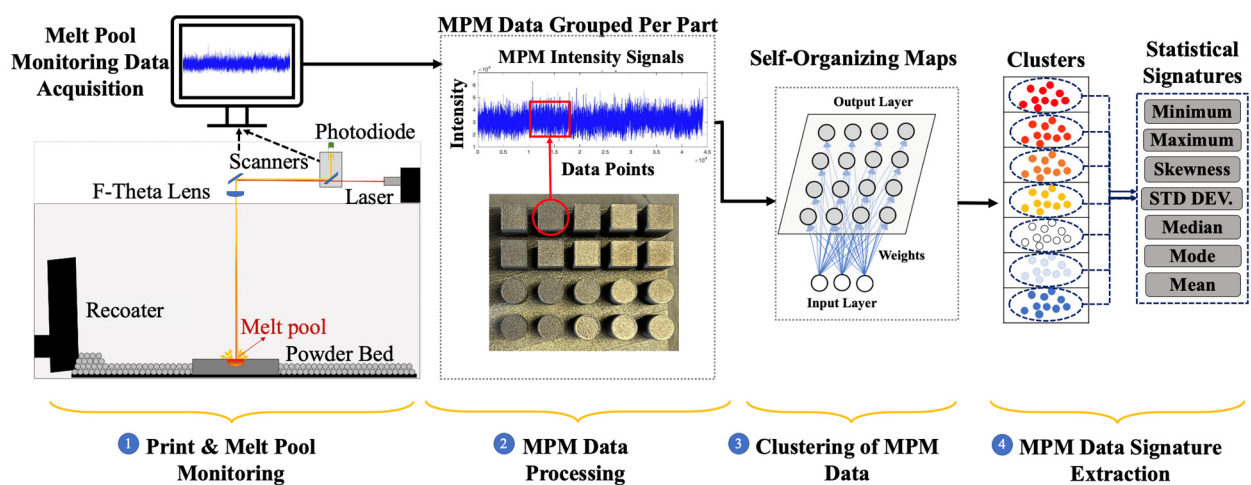

Figure 3. Light intensity collection using an on-axis photodiode, processing, and clustering using self-organizing maps.

2.3. Machine Learning Algorithms

2.3.1. Self-Organizing Maps

Each part's MPM data were then clustered into seven groups (other cluster numbers were tested and results are shown in the following sections) based on the intensity level using the unsupervised Self-Organizing Maps (SOMs) algorithm. As a result of changes in heat conduction, heat accumulation, and emission rate across different regions of a part, this technique was adopted to extract process signatures from multiple classes/clusters of light intensity data of the parts rather than extracting one general measure (e.g., one average intensity signature for the entire part) per each part. It also benefits the ML algorithms' performance due to the larger feature set.

The SOMs algorithm offers several advantages over other clustering methods, such as k-means clustering. The primary reason for its application in this study is its robustness to noise. MPM datasets, i.e., the light intensity signals, are highly susceptible to noise. The SOMs algorithm can effectively discern the true underlying patterns within noisy data. This capability arises from the neural network nature of the algorithm, wherein learning occurs through multiple iterations and continuous improvement. Moreover, a previous study conducted by our team [15] has demonstrated that the SOMs algorithm can effectively cluster in situ light intensity signals obtained using the same printer and sensor employed in this study.

As illustrated in Figure 3, seven statistical metrics were chosen to be extracted from each cluster of MPM signals (per part) as the printed part's signatures.

As described in a study published by Taherkhani et al., SOMs training is based on competition while other neural network algorithms' optimization is based on back-

propagating. The SOMs nodes with the same dimension as the input data have weight vectors that are randomly initialized. In general, the SOMs training process could be summarized in three steps: competition, cooperation, and adaptation. In the competition step, the winning node for each of the input data points is identified based on the lowest distance between the node weight values and the input data points. In the cooperation step, the neighboring nodes to the winner node which have a high similarity with it (low distance) are identified using the neighborhood kernel function. The last step of adaptation consists of the modification of node weights regulated with a learning rate decay parameter. For a detailed description of the method, see [15]. The weight optimization continues until no further changes are seen in the winning nodes and then the final map of the clustered data set can be generated using the final weight values of the nodes.

The processed MPM datasets, including the clustering, signature extraction, and scaling, in addition to categorical input features such as powder size ranges and continuous input features such as parts' density, hardness, and surface roughness, are then fed to three various supervised regression algorithms, FFN, RF, and XGBoost, to predict the process parameters (algorithm outputs/targets), as shown in Figure 1.

Ultimately, the design of the study was planned on providing information regarding the powder size range, the mechanical and roughness properties of the parts, and the statistical representation of the light intensity signals (light intensity signatures) to feed the regressor algorithms. The algorithms were intended to estimate the proper set of process parameters associated to the given powder, part, and melt pool properties.

The MPM light intensity signals, recorded per layer, were initially clustered using SOMs. This clustering step aimed to group signals obtained from the entire part into multiple clusters, considering various underlying patterns influenced by factors such as part geometry (e.g., cuboid/cylindrical), process parameters (e.g., power, hatch distance), gas flow, and recoater direction. These factors affect light intensity differently based on the location of the melt pool on the build surface and the height of the printed layer within the part (layer number), making clustering essential for accurate representation of light intensity signals. SOMs incorporates topological characteristics of the signals, generating clusters with similar behavioral patterns in the data.

Subsequently, signals classified into each cluster were characterized using statistical signatures. For each cluster, the average, standard deviation, mode, and median of light intensity were computed. Additionally, the minimum and maximum intensity levels were calculated by averaging the intensity over 10 melt pool points with the lowest and highest intensities, respectively, to mitigate the impact of outlier points potentially affected by noise, particularly in extreme intensity levels. This approach yielded seven statistical signatures from each cluster of photodiode signals recorded in each part, effectively reducing the dimensionality of the MPM data.

These seven light intensity signatures per cluster were combined with categorical indicators of powder size and part properties (density, hardness, and side surface roughness) to serve as input features for the algorithms. In total, 55 input features were utilized in this study, comprising seven signatures from each of the seven clusters of MPM data, one-hot-encoded indicators for three batches of powder sizes, and three features representing part physical properties. The algorithms aimed to predict the four major quality-impacting process parameters: laser power, scan velocity, hatch distance, and energy density, serving as the outputs of the models.

2.3.2. Feed-Forward Neural Network

FFN is a supervised artificial neural network (ANN) that is comprised of the input layer, hidden layers, and the output layer. Each layer is constructed from multiple information processing units denoted as neurons/nodes. Information transfers between the neurons of the subsequent layers by weight parameter. Weights connect the neurons, and their value determines the importance of the information conveyed between each pair of neurons to the accurate calculation of the known target of the network. FFN algorithm

training operates using the back-propagation technique which optimizes the user-defined error/loss functions (difference between network calculations and actual outputs) by moving along the steepest slope of error decrease in the loss function space at each training iteration [26]. Training and loss function updating continue until no further improvement is made in the calculations. In this study, the FFN regressor is developed and trained using the tensor flow library and keras API on the Visual Studio Code interface. Hyper-parameters such as the number of layers, neurons, drop-out rate, and initial learning rate (used in the Adam model optimizer) are tuned in this study by the application of the keras-tuner tool in keras API. Adam optimizer, tanh activation function, and the inverse time decay model optimization technique were chosen for the model based on error and trial attempts operated on the train (0.8 fraction of initially selected train set) and validation data (0.2 fraction of initially selected train set). The model's final structure was derived from the tuner outcome demonstrated in Table 2. The activation function of tanh was used in the input layer and hidden layers. Linear activation was employed in the output layer for the regression task.

Table 2. FFN-tuned hyper-parameters and the final structure of the algorithm. L1 and L2 denote the first hidden layer and the second hidden layer, respectively.

Algorithm	Layers	Nodes		Drop Out		Initial Learning Rate
FFN	2	L1	311	L1	0.2	0.01
		L2	119	L2	-	

2.3.3. Random Forest

The RF regressor simultaneously generates several decision trees (DT) on the various subsets of the dataset. Each tree starts with a parent node that splits into child nodes, which contain subgroups of data points, based on a partitioning condition which is determined at each node of the tree. The dataset splits according to whether each data point satisfies the parent node's partitioning condition or not. The impurity reduction between parent and child nodes of the dataset at each tree is then measured based on the defined metrics (e.g., variation and error). The partitioning of data continues (along the tree length) until no further improvement is made on the impurity reduction in the last partitions of the dataset. An ensemble (average) of random DT learners is then used to implement the final prediction [27].

2.3.4. Extreme Gradient Boosting

XGBoost regressor is also a modified RF algorithm that generates DT learners on random subsets of samples and features but in successive order. Each DT (weak learner) is developed to predict the residual. The prediction of each weak learner is improved compared to the previous weak learner by modification of features' coefficients in the error equation based on the previous learner's results [28]. XGBoost enables a more powerful model regularization by employing L1 and L2 methods and can avoid overfitting more efficiently with the tree-pruning tool. The trained RF and XGBoost algorithms' hyper-parameters in this study are tuned by implementing the CV and grid search on the train set. The main fine-tuned parameters are listed in Table 3.

Table 3. RF and XGBoost tuned hyper-parameters.

Algorithm	Estimators	Learning Rate	Max Depth	Max Leaf Nodes
RF	50	_	9	50
XGBoost	200	0.1	4	_

To summarize, according to the graphical abstract shown in Figure 3, three powder size ranges (15–53 μm, 15–106 μm, and 45–106 μm), the melt pool signatures (mean, mode, median, minimum, maximum, skewness, and standard deviation), and the printed part physical properties (density, surface roughness, and hardness) comprise the input features set of the algorithms. The categorical variables of powder size are transformed into numerical variables through one-hot encoding. Ultimately, three algorithms of RF, FNN, and XGBoost are then trained to predict the outputs which are the process parameters of power, velocity, hatch distance, and energy density of the laser.

3. Results and Discussion

3.1. Experimental Data Analysis

3.1.1. Correlation between Powder Size Distribution and Part Properties

The alteration in part properties (surface roughness, density, and hardness) against the powder batch shift from the finest size range (15–53 μm) to the widest (15–106 μm) and the coarsest range (45–106 μm) are illustrated in Figure 4. According to Figure 4a, the surface roughness of cylindrical parts (samples 54–108) exhibits a relative rise in comparison to cubical parts (samples 0–53), particularly accentuated in the coarser batch ranging from 45 to 106 μm. For example, for the cubical sample 2 in this batch, the surface roughness is ~23.01 μm whereas for cylindrical sample 2 in the same batch, which is printed with the same set of process parameters as the mentioned cubical sample, the surface roughness increased to ~29.02 μm. Considering the same set of process parameters are used for the cylindrical and cubical parts, this therefore implies the impact of geometry on the surface roughness of parts. Meanwhile, this trend of variation was not observed in density and hardness characteristics meaning the change of geometry in this design mainly affected the surface roughness. For all three studied part properties (Figure 4a–c), it may be concluded that parts in all three powder batches mostly behave the same in response to process parameters variations and have similar fluctuations. As expected, Figure 4a shows that the parts printed with the coarsest powder batch resulted in the highest roughness while parts printed with the fine powder batch had the lowest roughness. According to Figure 4b,c, the best quality of density and hardness are respectively obtained with parts printed with the fine powder batch as well but much less difference is seen between the powder batches for these characteristics (the least difference in density). Although the set of all influential factors should be addressed in the inspection of LPBF parts quality, it could generally be concluded that the fine powder tends to promote a higher density packing, powder bed uniformity, and energy absorption, thus, less perturbation in the melt pool dynamics. This could lead to improved part properties.

The mentioned features of the fine powder have led to the application of the finer powders for higher quality production in the industry (15–53 μm is a standard size range for LPBF applications). However, the optimization of LPBF for wider and coarser ranges of powder is a vital requirement in LPBF since it is more cost-effective due to the notably lower production cost of wider-range powder. The coarse/wider powder normally inherits better flowability which can significantly alter the powder bed characteristics positively. The process parameter optimization for unconventional/off-size powder sizes is a major gap in the recent state of LPBF research which inspired this study to integrate the standard and off-size powders into the ML-based design of process parameters. The results of this study enable data-driven process parameters arrangement for wider adaptation of off-size Ti64 powder in the LPBF applications.
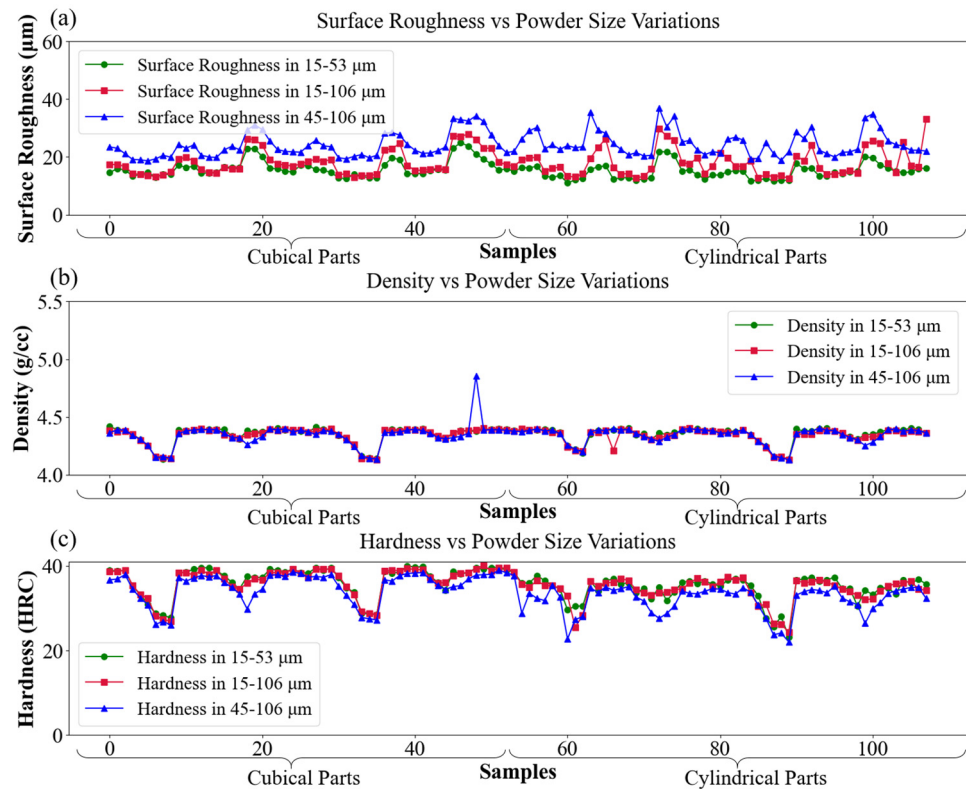
Figure 4. Variation of parts physical characteristics in cubical ([0–53] on *x*-axis) and cylindrical ([54–108] on *x*-axis) parts with change of powder size. The process parameters per each sample in different powder batches are identical (to observe the impact of powder size change). The part characteristics variations are shown in (**a**) surface roughness; (**b**) density; (**c**) hardness.

3.1.2. Correlation between Melt Pool Light Intensity, Process Parameters, and Part Properties

Aside from the inspection of relationships between part properties and powder size ranges, the correlation between the MPM data and these factors is also highly influential in the LPBF process optimization. This inspection was also carried out in this study, and it was observed that the MPM data were highly impacted by power and energy density levels rather than powder size range. The correlations between part properties (density and roughness), powder size, and MPM signatures (the average intensity per layer) are demonstrated in 3D plots in Figure 5, for parts with the highest and the lowest surface roughness as well as density in all three batches of powder. For clarification on the mapped information in this figure refer to Figure A1 in Appendix A section. As seen in Figure 5a,d, regardless of the powder batch used for the prints, higher power causes a higher level of intensity in parts. In cases with equivalent power levels, energy density should be inspected to interpret the intensity level. This means that either hatch distance or velocity would influence the level of light intensity irradiated from the melt pool. In Figure 5b, the highest and lowest density for cubical samples in batch A (pink square and diamond markers) are both attributed to the laser power of 150 W; but the higher energy density in the latter caused the higher level of intensity. Both laser velocity and hatch distance also become relevant even when both power and energy density are equivalent. As such condition could be seen in Figure 5b between the lowest density cubical samples in A and B (pink and gray diamonds), the smaller hatch distance of the latter caused higher intensity due to heat accumulation within a certain volume of process zone. The same applies to the sample with the lowest surface roughness in batch A (pink circles shown in Figure 5c) which has the highest light intensity compared to the two samples with the lowest surface roughness in batch B and C (gray and blue circles) due to the smaller hatch distance. In the same plot, the sample with the highest surface roughness in batch B (gray triangle) shows

a lower intensity regardless of having an equivalent power but a higher energy density due to a lower velocity. The significantly lower velocity and lower hatch distance in this sample could generate a keyhole in the melted material causing light scattering within the process zone, which in turn could decrease the intensity which also enables the detection of potential flaws (keyhole in this case) based on the MPM data.
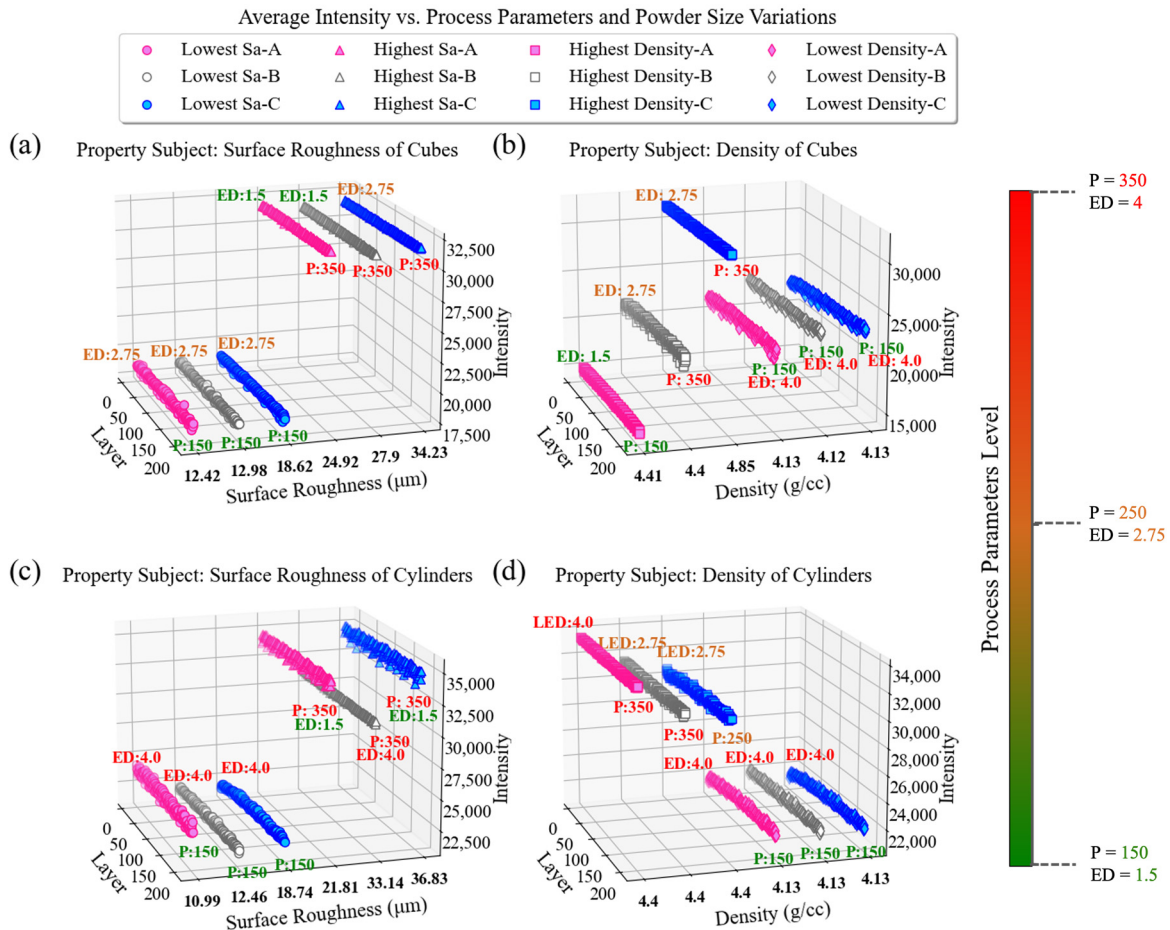

Figure 5. The average intensity level per layer is plotted for cubical and cylindrical parts with the highest and lowest surface roughness/density in three powder batches of 15–53 μm (A), 15–106 μm (B), and 45–106 μm (C). The average intensity per layer in parts is plotted for (**a**) cubical parts with highest and lowest surface roughness; (**b**) cubical parts with highest and lowest density; (**c**) cylindrical parts with highest and lowest surface roughness; (**d**) cylindrical parts with highest and lowest density.

The light intensity data recorded in each layer, particularly at the edges of the parts, directly links melt pool stability to the resulting roughness, including the side surfaces where roughness measurements were taken. These light intensity signals are closely related to the temperature and thermal distribution on the build surface. Instabilities in the melt pool, caused by thermal variations, are reflected in the light intensity signals. Anomalies such as spatter and balling manifest as perturbations in these signals. Since phenomena like balling and spatter affect the roughness of the parts, the light intensity signal signatures can approximate the resulting roughness in layers and, ultimately, the entire part. Similarly, the melt pool condition in each layer is recorded by the light intensity, with the overall extracted light intensity signatures (per part) statistically representing the light intensity across all layers, correlating to the density and hardness of the final fabricated parts. Light intensity can also indicate potential lack of fusion or keyhole defects based on signal perturbations caused by thermal disturbances during these defects. These defects determine the quality of fusion and ultimately define the part density. Lastly, the cooling rate and thermal history

of the melt pool, which significantly impact the microstructure of the parts, can also be captured by the light intensity data. This is due to the light intensity's ability to reflect temperature and thermal conditions within the melt pool, as previously discussed.

These analyses collectively underscore the importance of powder PSD, process parameters, and melt pool characteristics in shaping the physical properties of LPBF-made parts. It should be also noted that these factors are intricately interconnected through complex and nonlinear relationships, which often defy unraveling through simple statistical modeling or experimental studies alone. The volume of informative data from these factors further urges for leveraging advanced data processing tools capable of complex data learning and modeling, such as ML and deep learning algorithms.

3.2. Machine Learning Analysis

3.2.1. Evaluation Metrics and Techniques

In the first approach of model assessments, the algorithms were trained and evaluated using the repeated CV method. Twelve samples with optimum levels of surface roughness and density were selected from the original dataset to be used as unseen samples. Therefore, the train set on which CV was performed is comprised of the remaining 312 samples. It should be noted that the targets of the analysis contain low variance. Hence, the regression results analysis could not be solely based on the conventional error measuring metrics, such as Mean Squared Error (MSE) or Mean Absolute Error (MAE). These metrics mainly normally imply the difference between the actual and predicted values, where they are highly sensitive to scale and outliers in data. Due to the low variance in the targets, the study assessed regression performance using the MAPE metric. This metric, being scale-independent, was employed for analyzing accuracy results in CV and predictions on unseen samples. Furthermore, due to the low variance in the targets, the deviation span of the errors in the CV analysis was also inspected to attain a reliable insight into the performance of the models. By this approach, an estimation of the uncertainty in predictions could be obtained. The coefficient of determination (R^2 score) is reported in both CV and unseen samples regression results to assess the extent to which the models explain variation in the targets or how effectively a regressor can model the overall target data population (goodness of fit). Additional statistical examinations, such as the Kolmogorov–Smirnov (KS) test and comparison of probability density distributions of prediction and actual targets, are conducted in the unseen samples regressions to provide a comprehensive assessment of regressors' generalization on the unseen samples. Integrating all these assessments helps address the low variance in the targets and provides unbiased insight accordingly.

3.2.2. Self-Organizing Maps Clusters Optimization

The number of SOMs clusters on which the dataset was generated was also optimized using the CV technique. Initially, seven clusters were chosen to be applied in SOMs clustering and generation of the baseline dataset, according to the results of a previous similar study involving a smaller dataset with similar DoE. Therefore, the initial dataset was generated with seven SOMs clusters and was subsequently used for regressor algorithms hyper-parameters tuning. For additional assessment, the tuned algorithms were then re-examined with datasets derived from various SOMs cluster numbers, using cross-validation (CV) with five folds. In CV, the data are divided into multiple folds to make predictions once on each left-out fold while being trained on the rest of the folds. While there are no strict rules regarding the number of folds and it is often dependent on the characteristics of datasets and modeling requirements, five folds of CV were opted for in this study, as this aligns with a common practice where 5 or 10 are typically chosen for the CV implementation [29]. The overall mean absolute error percentage (MAPE) of prediction obtained by the regressors in the CV results was averaged to choose an appropriate number of clusters among 5, 7, and 10 clusters for this application.

The results of this test are illustrated in Figure 6, indicating that the dataset generated using seven clusters of the SOMs led to a marginally better average performance in the

regression models for all four targets, achieving an average MAPE of approximately 3.7%. Although the difference in performance between using five and seven clusters was slight, both configurations were deemed valid choices. However, based on previous preliminary investigations with different datasets (cuboids and cylindrical parts), the seven-cluster SOMs consistently yielded the best performance. This suggests better generalization of algorithms with this clustering approach across various datasets, leading to the decision to proceed with the seven-cluster SOM dataset.
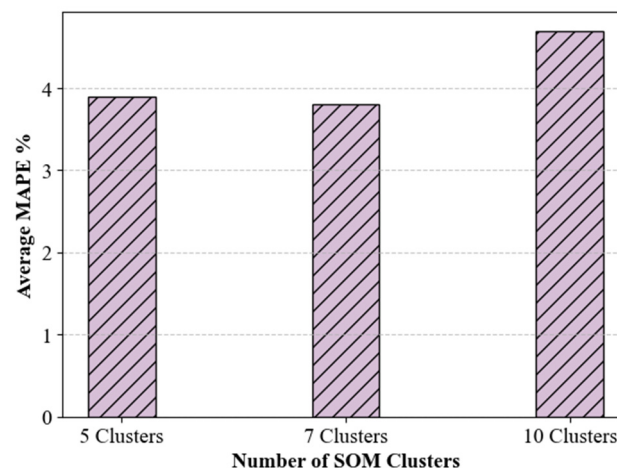

Figure 6. Impact of SOMs cluster numbers on overall averaged performance of three regressors showed by average MAPE metric obtained in five-fold CV test implemented on train set samples.

3.2.3. Cross-Validation on Train Set Analysis

Regression Accuracy Evaluation

Five repeats of the five-fold CV were implemented on the train set meaning that the iterative process of CV is repeated five times to randomly change the set of samples in each fold, resulting in a total of 25 various validation sets. The FFN algorithm was trained here with 50 epochs in each round of CV.

As shown in Figure 7a, the R^2 score obtained by all three models in the regression of validation folds indicates that the models can fit the dataset very well. The obtained R^2 scores show that above 90% of the target variance is explained by the models in all 25 configurations for power, velocity, and energy density targets. However, the results vary for hatch distance, with XGBoost showing the best modeling with a ~0.57 R^2 score, while RF and FFN could not achieve meaningful modeling of the data. This shows that although the XGBoost model exhibits relatively better capability in learning the hatch distance variance, it could still lack a reliable level of fitness due to a low R^2 score. It should also be considered that low target variability of hatch distance (~9% coefficient of variability in the train set while other targets have above 30%) impacted the R^2 score, regardless of the level of error in prediction. By inspecting the actual values and prediction values obtained in each CV, it was observed that a low level of variability in hatch distance relative to cumulative error significantly decreased the R^2 score in each CV split compared to other targets with high variability. R^2 score mainly reflects how well a model fits the data and how effectively it estimates the variation of the target based on inputs/predictors and low variability of targets could cause bias in its result. On the other hand, error metrics, such as MAPE, indicate the proximity of model predictions to actual data values on a point-by-point basis. For a comprehensive inspection of regression performance, both aspects should then be analyzed. The average MAPE (average of five iterations of random five-fold CV) of models in the estimation of targets is shown in Figure 7b. The MAPE level obtained by the FFN for all four targets exceeds those of the XGBoost and RF regressors. Comparing RF and XGBoost, RF demonstrates marginally more accurate predictions in energy density and

power. However, the difference in performance between the two regressors for hatch distance and velocity predictions is minimal. It could also be interpreted from the hatch distance results that, as it was expected by R^2 scores, XGBoost resulted in a lower prediction error. Additionally, although the R^2 score was low, the model regressed with good accuracy (MAPE of ~0.047). This again implies that low target variability could have influenced the low R^2 score.
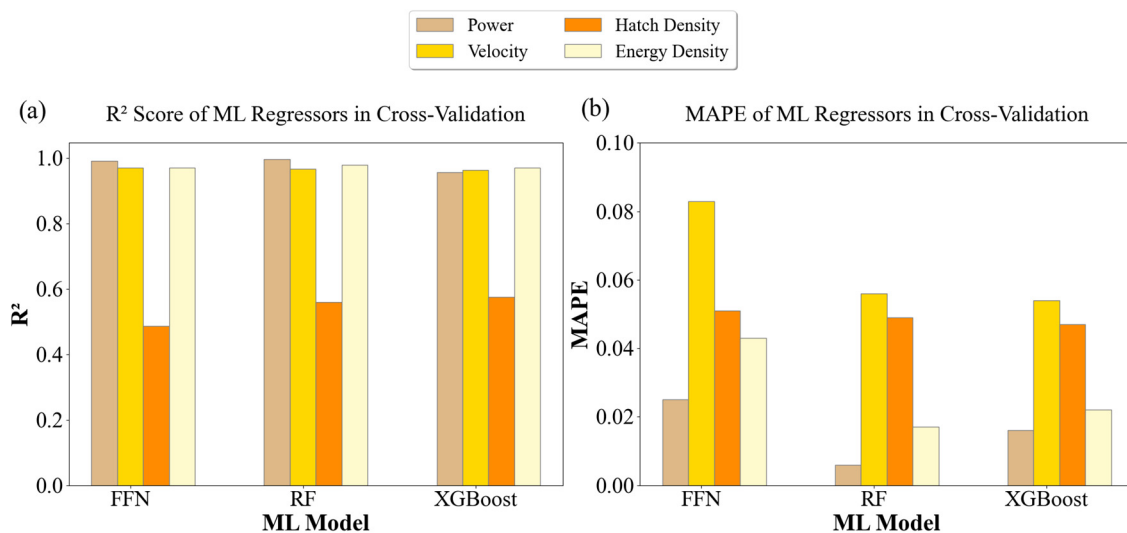

Figure 7. Average of: (**a**) R^2 score; (**b**) MAPE of the regressor algorithms in CV analysis implemented on the train set of data.

Statistical Evaluation

The accuracy of the models' prediction is further assessed from a statistical perspective, to identify which model exhibits more reliable regression accuracy. The notched box plot of the yielded average MAPE of models in all 25 configurations of five-fold CV is illustrated in Figure 8. Both XGBoost and RF regressors show favorable process parameters regression results in the CV analysis for all four targets. The box plots in Figure 8a–d show that RF, with small superiority to XGBoost, achieves the highest accuracy for all targets. This marginal difference, compared to the RF model, is also seen in the prediction accuracy results shown in Figure 7a. Additionally, considering the interquartile ranges and whiskers caps, it can also be interpreted that RF exhibits a smaller range of variation in error levels throughout the 25 CV combinations for all four targets, this underscores the reliability of this regressor, suggesting its higher robustness to data perturbations and potential for improved generalization. In contrast, the FFN regressor shows a wider range of errors in the CV analysis for all four targets, indicating its less robustness to noise or any potential data variation. Hence, it is anticipated to result in poorer generalization to new unseen samples. The overlapping notches in the RF and XGBoost boxes illustrated in Figure 8b–d show the median in these two populations is not significantly different, showing a similar level of prediction accuracy, in contrast to FFN. In summary, it can be concluded that XGBoost and RF are the better regressors, in the CV analysis focused on the training set, in terms of goodness of fit and prediction accuracy.

However, it should be noted that RF and XGBoost achieved a low R^2 score in hatch distance modeling despite exhibiting good point-wise prediction accuracy. This may suggest a potential issue of overfitting to the training samples. Accordingly, further inspection into model generalization was implemented relying on algorithms evaluation on unseen samples.
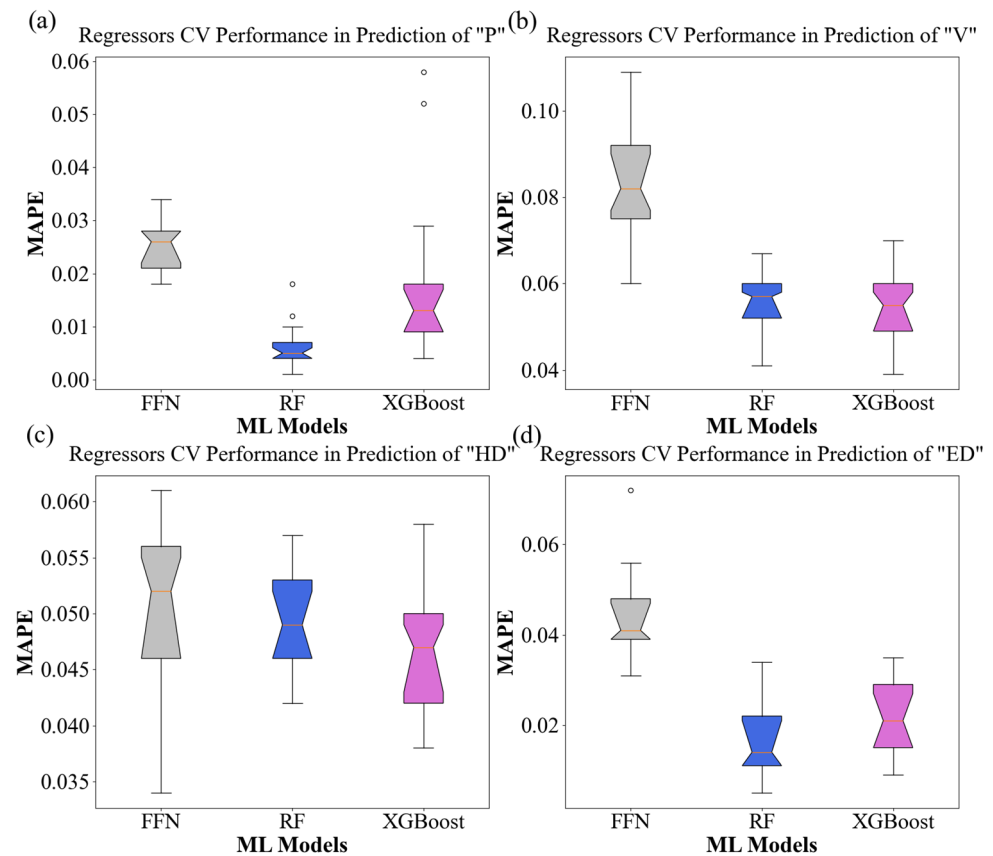
Figure 8. Notched box plot of regression errors of three regressor algorithms obtained in CV analysis implemented on the train set. The results are shown for (**a**) power (P); (**b**) velocity (V); (**c**) hatch distance (HD); (**d**) energy density (ED).

3.2.4. Train Set Size Impact

While CV allows for assessing the model performance based on variations in sample set combination, the number of training samples remains fixed in each fold. However, the size of the training set is one of the most influential factors in the performance of regressors, necessitating an examination of the performance of the regressors across different training set sizes. Limited training data sizes can result in overfitting and limited generalization performance. Increasing the size of the training set typically enhances the accuracy of regression models, especially those with strong generalization abilities.

The regressors' overall prediction error (average error of all targets) was assessed in various fractions of a train set in the [60%, 70%, 80%, 90%] range taken from the set of 312 samples. Therefore, in each train size fraction, models predicted the samples in the remained fraction of 312 samples. As shown in Figure 9, it is observed that an increase in train set samples mostly led to higher accuracy (except for a negligible increase from 70% to 80% in RF and XGBoost models). This implies that there is no significant overfitting/sensitivity to data which confirms the admissible level of generalization of the models.

Overall, FNN consistently resulted in higher errors in all target estimations, which agrees with the previous analyses. However, it should be considered that none of these error levels are considered major in LPBF applications and do not disqualify any of the models. In conclusion, in a comparative study among the three regressors, FFN seems to have lower accuracy and generalization. RF and XGBoost, in most cases, demonstrated similar levels of performance and better generalization compared to FFN.

Figure 9. Impact of train set size on the overall performance of three regressors in the estimation of process parameters. The overall error is shown by the $\log_{10}$(MAPE) metric.

3.2.5. Regression of Unseen Samples Analysis

Regression Accuracy Evaluation

The final evaluation of the models in terms of generalization performance involves regression on unseen samples that were not included in previous analyses. The models were subsequently reset and retrained on the entire training set (i.e., 312 samples) to make predictions on a pre-selected set of 12 unseen samples. It should be noted that since these 12 samples were the optimum samples of the DoE in terms of surface roughness and density, they count as extreme points compared to the utilized training set. Consequently, it enables the analysis to evaluate the robustness of the models to outliers/extreme points too.

In addition, an epoch number examination was carried out separately to identify the optimal epoch number that balances the overfitting and bias of the FFN model. In this test, 20% (from 312 samples) of the training set was used to evaluate the FFN performance in various epoch numbers. Figure 10 illustrates a significant decrease in error within the initial 20 epochs, followed by a highly gradual decline in error from 20 to 100 epochs. The epoch number of 50 was chosen since it promotes a good level of accuracy in estimations. The model's training had also reached stability in the learning process at this epoch number.
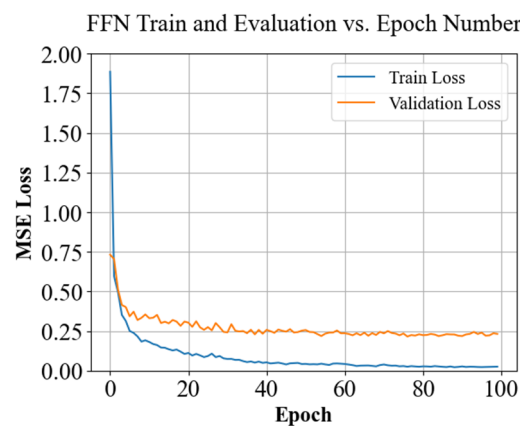

Figure 10. The impact of epoch numbers on FFN performance was tested and shown by the MSE metric. The algorithm was tested on 80% of the initial train set as the new train set and 20% of the initial train set as the validation set.

The RF algorithms and XGBoost regressors used in the prediction of the unseen samples are the same tuned models that were used in the previous analysis. In this step, all three algorithms demonstrated good prediction accuracy for the majority of targets. Among four targets, the lowest prediction accuracy was seen in the hatch distance estimations for all three candidate models. This was expected according to the performance of the models in the previously reported analysis. The total R^2 score achieved by FFN in predicting

unseen samples was ~0.84. The performance of FFN on unseen samples for each four targets (R^2 and MAPE) is shown in Figure 11a. FFN shows very accurate generalization for power, velocity, and energy density estimations, yet achieved a lower R^2 score in the hatch distance prediction. However, the observed performance of FFN in the hatch distance estimation for unseen samples is consistent with the results obtained in the CV analysis (training samples).
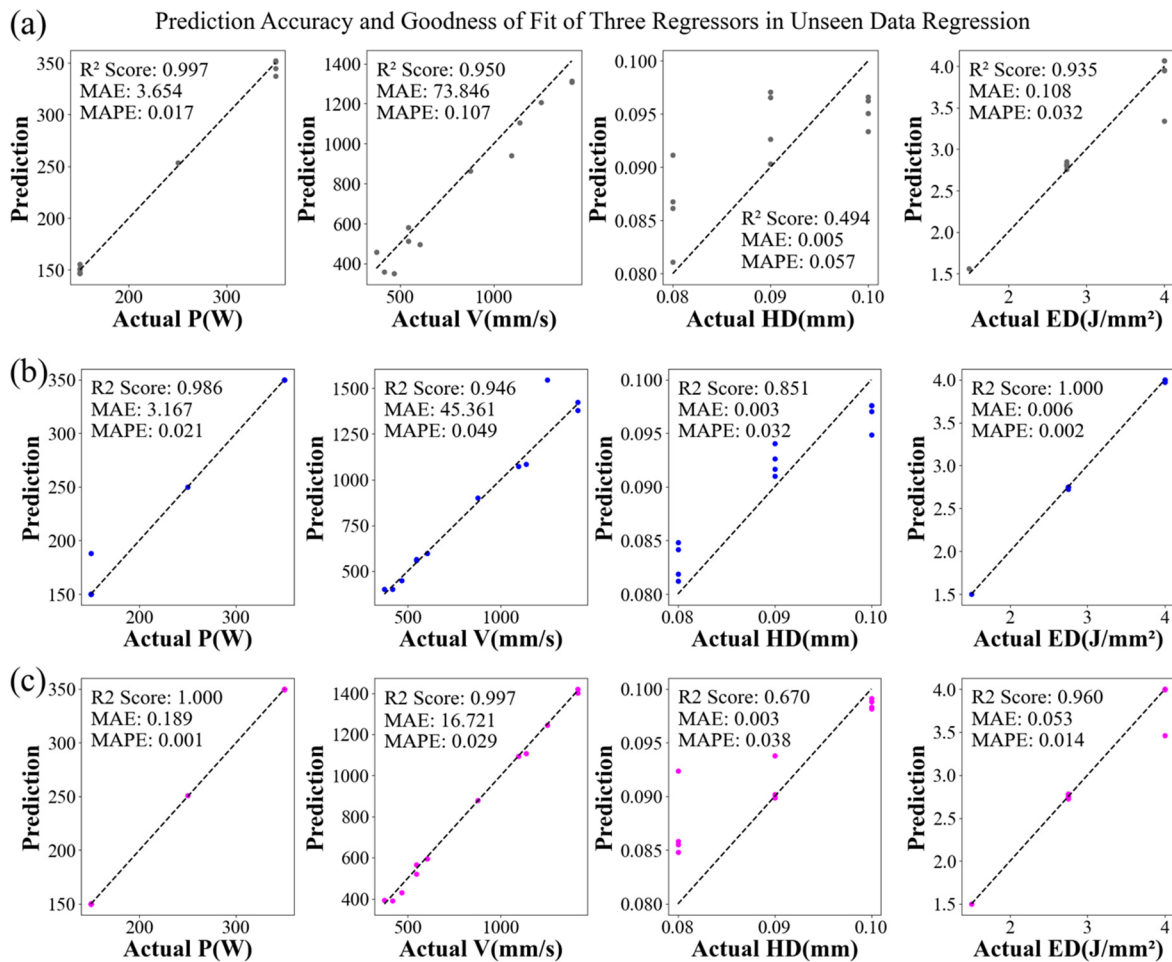

Figure 11. Prediction vs. actual target plots obtained from the unseen samples' regression in the four targets. Prediction accuracy of regressions is also shown by MAE and MAPE for (**a**) FFN regressor (**b**) RF regressor, and (**c**) XGBoost regressor. Power, velocity, hatch distance, and energy density are denoted by (P), (V), (HD), and (ED), respectively.

A great total R^2 score of 0.94 was achieved by the RF regressor in predicting unseen samples. As demonstrated in Figure 11b, RF achieved above 90% accuracy in power, velocity, and energy density estimations, as well as FFN. However, RF showed more reliable results than FFN in hatch distance estimations, considering both R^2 and MAPE measures. This suggests an overall good generalization and robustness in the RF regressor.

XGBoost regressor showed results slightly less accurate than RF and marginally better than FFN on the unseen samples set, with a total R^2 of 0.9. The R^2 and MAPE obtained by XGBoost regressor in unseen sample estimation are shown in Figure 11c. XGBoost yielded results almost as accurate as RF across all targets, except in hatch distance estimations where an R^2 score of ~0.67 was obtained which explains that both FFN and XGBoost could not provide accurate data fitting for this target. Meanwhile, RF resulted in robust hatch distance modeling (R^2 score ~0.8) and accurate predictions (lowest MAPE).

Statistical Evaluation

As previously mentioned, due to the low variance of targets in the dataset, regression performance needs to be assessed in various aspects using different techniques. The R^2 score, error variability, and prediction accuracy were discussed and compared for the three trained models. However, to further ensure the generalization of the models, the results obtained by the three regressors in estimating the unseen samples were also examined from a statistical perspective. This was implemented to examine the predictions in terms of distribution. The Kolmogorov–Smirnov (KS) test, a non-parametric technique that does not assume any distribution for the data, was applied to test the regression results. The test results indicated that the predicted values by all three models are from the same population distribution as the targets, further confirming the credibility of the models' performance.

In addition, the Probability Density Function (PDF) was used to visualize the probability density distribution of the predicted population and the target population in the unseen samples tested by the KS tests. Based on the PDF in Figure 12a, it can be concluded that FFN yields less overlap between predictions and actual target distribution in velocity, hatch distance, and energy density targets. By comparing the PDF plots in Figure 12b,c, obtained for each of the targets, it can be concluded that RF and XGBoost generally exhibit better data learning, but RF achieves slightly more accurate data learning overall. This aligns with the previous analysis and the fact that RF generates a density distribution closer to the actual targets.
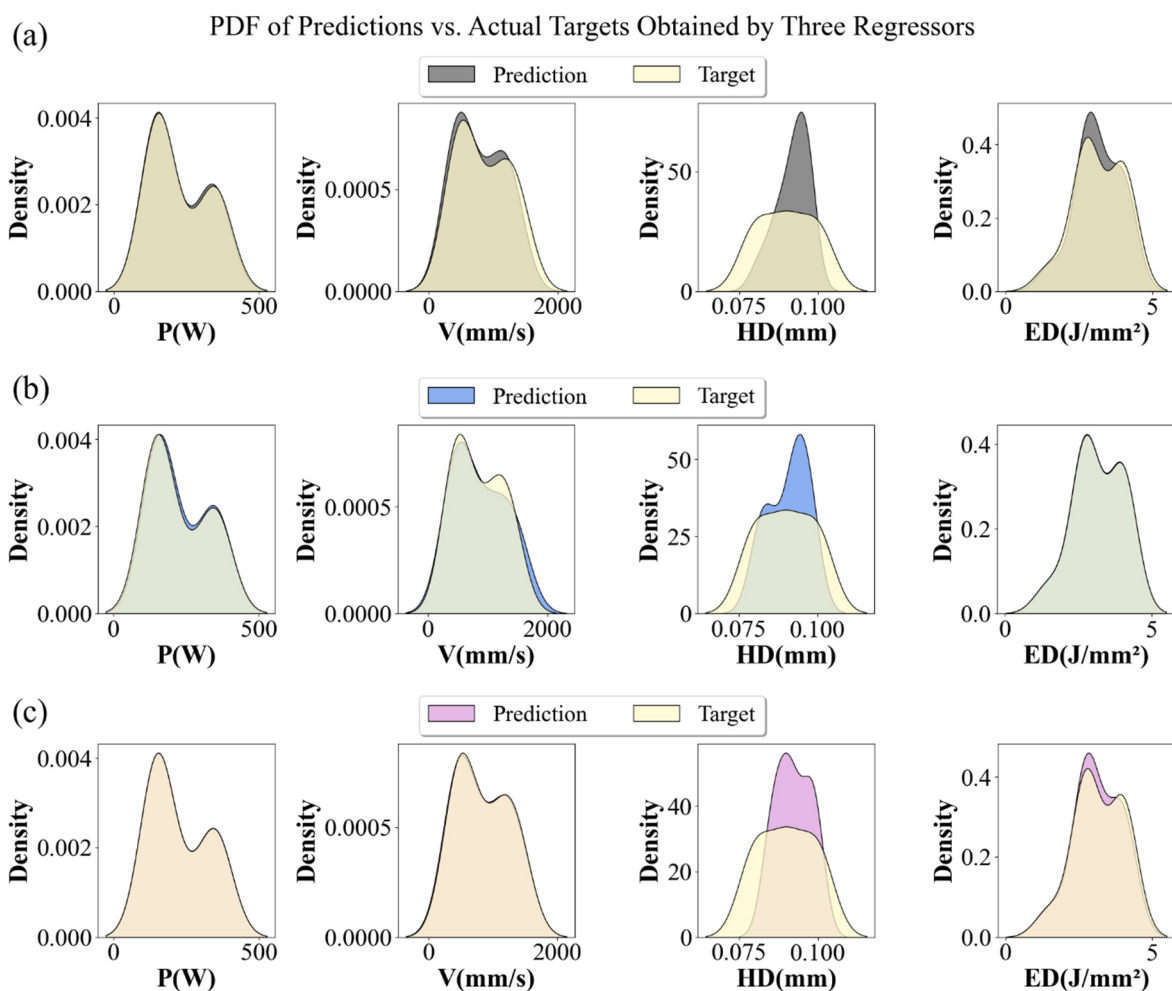

Figure 12. Probability density distribution of predicted values and actual target values in unseen samples regression resulted by (**a**) FFN regressor (**b**) RF regressor, and (**c**) XGBoost regressor. Power, velocity, hatch distance, and energy density are denoted by (P), (V), (HD), and (ED), respectively.

The plot of hatch distance density distributions indicates that all three models frequently estimated a hatch distance of 0.09 (instead of 0.08 and 0.1), suggesting that they lacked accurate data variation learning concerning hatch distance changes at different extents. This deviation, however, is very marginal when it comes to the optimization of LPBF. Notably, the hatch distance variations in the DoE, ranging from 0.08 to 0.1, do not have a major impact on the quality of the parts, unlike the other three target changes. In addition, the prediction MAE of models in hatch distance fell within an acceptable small range [0.003–0.005]. Therefore, the hatch distance estimation accuracy obtained by the models could still be considered an admissible result for such a DoE in terms of the quality and functionality of the LPBF-made product.

To select the final regressor as the best candidate for future studies, the overall MAPE of FFN, RF, and XGBoost in the estimation of all four process parameters were reviewed, yielding values of 0.05, 0.025, and 0.02, respectively. This, along with the R^2 score levels shown in Figure 11, state that both RF and XGBoost can achieve high prediction accuracy in terms of overall performance (considering all targets simultaneously); however, RF demonstrates higher goodness of fit and consequently better generalization.

4. Conclusions

In this paper, three distinct regression models, namely Feedforward Neural Network (FFN), Random Forest (RF), and Extreme Gradient Boosting (XGBoost), were trained and optimized to forecast process parameters, including power, velocity, hatch distance, and energy density, utilizing light intensity signatures and physical properties of cubical and cylindrical shape samples, made by laser powder bed fusion (LPBF), across standard and off-size Ti64 powder size distributions. The efficacy of these regression models was evaluated through metrics such as the coefficient of determination (R^2 score), mean absolute percentage error (MAPE), and a comprehensive analysis of error distribution within cross-validation (CV) assessments. Additionally, the performance of these algorithms was assessed through the prediction accuracy observed in evaluations of unseen samples.

According to the reported assessments of the regression results, the following conclusion can be drawn:

- Overall, all three algorithms demonstrated satisfactory performance in estimating process parameters, as evidenced by the magnitude of prediction errors. These errors, within acceptable bounds, indicate the potential for maintaining high-quality parts despite the inherent uncertainties associated with prediction in this application.
- In CV, although RF and XGBoost performed similarly closely, the lower deviation observed in prediction errors underscores the greater robustness of RF compared to XGBoost.
- Among four targets/process parameters, three regressors showed the lowest accuracy and goodness of fit, in both the CV analysis and the unseen samples evaluation, in the hatch distance estimation.
- In the CV analysis, while all three regressors achieved practically acceptable levels of prediction errors for hatch distance estimation (ranging from 4% to 6% MAPE), the highest goodness of fit (R^2 score) was attained by XGBoost, approximately 0.57, surpassing that of RF. This finding suggests that while acceptable results were achieved in the unseen samples tested here, the regressors may not exhibit strong generalization ability for other unseen samples. However, the predicted results fall within the range of outcomes considered satisfactory for optimizing a typical LPBF process.
- Among all regressors, RF showed the most reliable results, showing promising generalization in unseen sample evaluation with an overall MAPE error of 2.5% and R^2 score of 0.94 in the prediction of all four process parameters.
- Following RF, XGBoost emerged as the second-best performer in estimating process parameters for unseen samples, albeit with marginally lower goodness of fit. XGBoost achieved an overall MAPE of 2% and an R^2 score of 0.9.

- Despite exhibiting less precise model fitting and generalization, the FFN algorithm still produced acceptable results, achieving an R^2 score of 0.84 and a MAPE of 5% in evaluations of unseen samples.

The discussed results of this study confirm the applicability of such ML-based process parameters determination techniques for similar applications involving cubical and cylindrical parts fabricated by LPBF using Ti64 powder in standard and off-size ranges. Further evaluation of the proposed technique could be examined on more complex geometries of Ti64 parts, given initial light intensity data of printed parts. The practicality of this technique could be extended by developing additional algorithms providing a rough estimation of light intensity signatures in response to user-defined part physical properties. The rough estimation could be fed to the developed algorithm in this study to provide the proper process parameters costumed to the defined powder size range and part properties, eliminating the need for available light intensity data. This methodology could serve as a valuable tool for fast process optimization, reducing the number of labor-intensive and costly experimental tasks, thanks to the integration of MPM into the ML platform.

Author Contributions: Conceptualization, F.L. and E.T. methodology, F.L. and E.T.; software, F.L. and S.T.; validation, F.L. and S.S.; formal analysis, F.L.; investigation, F.L.; resources, M.H.-K.; data curation, F.L., S.T. and K.T.; writing—original draft preparation, F.L.; writing—review and editing, F.L., K.T., S.T., M.H.-K. and E.T.; visualization, F.L.; supervision, E.T.; project administration, E.T.; funding acquisition, E.T. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: Author Mahdi Habibnejad-Korayem was employed by the company Colibrium Additive. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Appendix A

Two information graphs for 3D plots (Figure 5a–d) are provided which map the information of 3D plots in (1) cubical parts and (2) cylindrical parts. The average intensity level per layer is plotted for cubical and cylindrical parts with the highest and lowest surface roughness/density in three powder batches of 15–53 μm (A), 15–106 μm (B), and 45–106 μm (C).
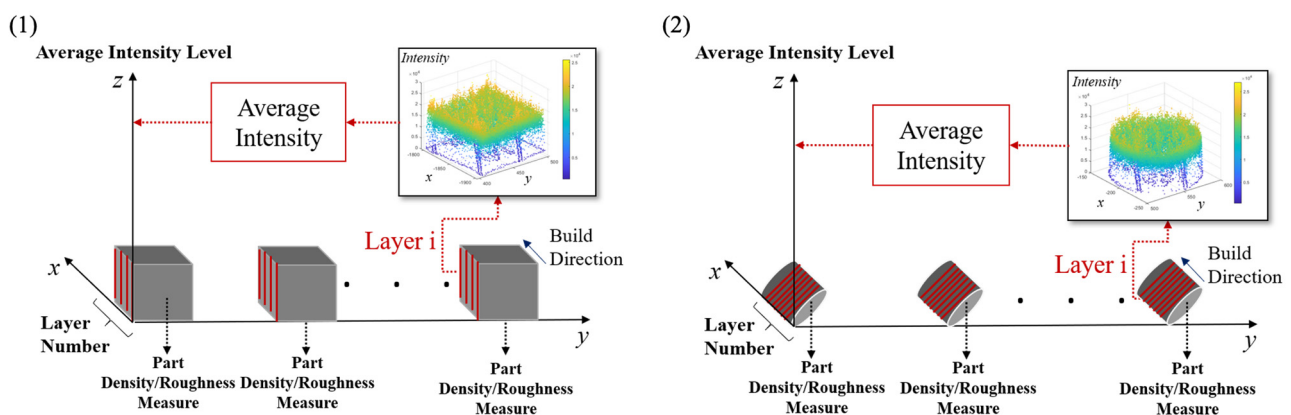

Figure A1. Information Map of Figure 5.

References

1. Toyserkani, E.; Sarker, D.; Ibhadode, O.O.; Liravi, F.; Russo, P.; Taherkhani, K. *Metal Additive Manufacturing*; Wiley: Hoboken, NJ, USA, 2022; pp. 1–90.
2. Singh, R.; Gupta, A.; Tripathi, O.; Srivastava, S.; Singh, B.; Awasthi, A.; Rajput, S.K.; Sonia, P.; Singhal, P.; Saxena, K.K. Powder Bed Fusion Process in Additive Manufacturing: An Overview. *Mater. Today Proc.* **2020**, *26*, 3058–3070. [CrossRef]
3. Chung Baek, A.M.; Park, E.; Seong, M.; Koo, J.; Jung, I.D.; Kim, N. Multi-Objective Robust Parameter Optimization Using the Extended and Weighted k-Means (EWK-Means) Clustering in Laser Powder Bed Fusion (LPBF). *Expert Syst. Appl.* **2024**, *236*, 121349. [CrossRef]
4. Gogulamudi, B.; Bandlamudi, R.K.; Bhanavathu, B.; Guttula, V.S.K. A Prediction Model for Additive Manufacturing of AlSi10Mg Alloy. *Trans. Indian Inst. Met.* **2023**, *76*, 571–579. [CrossRef]
5. Gu, Z.; Sharma, S.; Riley, D.A.; Pantawane, M.V.; Joshi, S.S.; Fu, S.; Dahotre, N.B. A Universal Predictor-Based Machine Learning Model for Optimal Process Maps in Laser Powder Bed Fusion Process. *J. Intell. Manuf.* **2023**, *34*, 3341–3363. [CrossRef]
6. Shen, T.; Zhang, W.; Li, B. Machine Learning-Enabled Predictions of as-Built Relative Density and High-Cycle Fatigue Life of Ti6Al4V Alloy Additively Manufactured by Laser Powder Bed Fusion. *Mater. Today Commun.* **2023**, *37*, 107286. [CrossRef]
7. Muhammad, W.; Kang, J.; Ibragimova, O.; Inal, K. Experimental Investigation and Development of a Deep Learning Framework to Predict Process-Induced Surface Roughness in Additively Manufactured Aluminum Alloys. *Weld. World* **2023**, *67*, 897–921. [CrossRef] [PubMed]
8. Cao, L.; Li, J.; Hu, J.; Liu, H.; Wu, Y.; Zhou, Q. Optimization of Surface Roughness and Dimensional Accuracy in LPBF Additive Manufacturing. *Opt. Laser Technol.* **2021**, *142*, 107246. [CrossRef]
9. Theeda, S.; Jagdale, S.H.; Ravichander, B.B.; Kumar, G. Optimization of Process Parameters in Laser Powder Bed Fusion of SS 316L Parts Using Artificial Neural Networks. *Metals* **2023**, *13*, 842. [CrossRef]
10. Zheng, Z.; Yue, X.; Wang, J.; Hou, J. A Framework for General-Purpose Microscopic Image Analysis via Self-Supervised Learning. *Mater. Charact.* **2024**, *213*, 114003. [CrossRef]
11. Zheng, Z.; Qiu, S.; Yue, X.; Wang, J.; Hou, J. Detecting Irradiation Defects in Materials: A Machine Learning Approach to Analyze Helium Bubble Images. *J. Nucl. Mater.* **2024**, *596*, 155117. [CrossRef]
12. Yan, Y.-G.; Wang, K. The Intrinsic Strength Prediction by Machine Learning for Refractory High Entropy Alloys. *Tungsten* **2023**, *5*, 531–538. [CrossRef]
13. Fidan, I.; Huseynov, O.; Ali, M.A.; Alkunte, S.; Rajeshirke, M.; Gupta, A.; Hasanov, S.; Tantawi, K.; Yasa, E.; Yilmaz, O.; et al. Recent Inventions in Additive Manufacturing: Holistic Review. *Inventions* **2023**, *8*, 103. [CrossRef]
14. Taherkhani, K.; Ero, O.; Liravi, F.; Toorandaz, S.; Toyserkani, E. 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.* **2023**, *99*, 848–897. [CrossRef]
15. Taherkhani, K.; Eischer, C.; Toyserkani, E. An Unsupervised Machine Learning Algorithm for In-Situ Defect-Detection in Laser Powder-Bed Fusion. *J. Manuf. Process.* **2022**, *81*, 476–489. [CrossRef]
16. Mao, Z.; Feng, W.; Ma, H.; Yang, Y.; Zhou, J.; Liu, S.; Liu, Y.; Hu, P.; Zhao, K.; Xie, H.; et al. Continuous Online Flaws Detection with Photodiode Signal and Melt Pool Temperature Based on Deep Learning in Laser Powder Bed Fusion. *Opt. Laser Technol.* **2023**, *158*, 108877. [CrossRef]
17. Yadav, P.; Rigo, O.; Arvieu, C.; Le Guen, E.; Lacoste, E. Data Treatment of In Situ Monitoring Systems in Selective Laser Melting Machines. *Adv. Eng. Mater.* **2021**, *23*, 2001327. [CrossRef]
18. Jayasinghe, S.; Paoletti, P.; Sutcliffe, C.; Dardis, J.; Jones, N.; Green, P.L. Automatic Quality Assessments of Laser Powder Bed Fusion Builds from Photodiode Sensor Measurements. *Prog. Addit. Manuf.* **2022**, *7*, 143–160. [CrossRef]
19. Cao, L.; Hu, W.; Zhou, T.; Yu, L.; Huang, X. Monitoring of Single-Track Melting States Based on Photodiode Signal during Laser Powder Bed Fusion. *Sensors* **2023**, *23*, 9793. [CrossRef] [PubMed]
20. Raj, A.; Huang, D.; Stegman, B.; Abdel-Khalik, H.; Zhang, X.; Sutherland, J.W. Modeling Spatial Variations in Co-Axial Melt Pool Monitoring Signals in Laser Powder Bed Fusion. *J. Manuf. Process.* **2023**, *89*, 24–38. [CrossRef]
21. Li, J.; Zhang, X.; Zhou, Q.; Chan, F.T.S.; Hu, Z. A Feature-Level Multi-Sensor Fusion Approach for in-Situ Quality Monitoring of Selective Laser Melting. *J. Manuf. Process.* **2022**, *84*, 913–926. [CrossRef]
22. Raj, A.; Owen, C.; Stegman, B.; Abdel-Khalik, H.; Zhang, X.; Sutherland, J.W. Predicting Mechanical Properties from Co-Axial Melt Pool Monitoring Signals in Laser Powder Bed Fusion. *J. Manuf. Process.* **2023**, *101*, 181–194. [CrossRef]
23. Li, J.; Cao, L.; Zhou, Q.; Liu, H.; Zhang, X. Imbalanced Quality Monitoring of Selective Laser Melting Using Acoustic and Photodiode Signals. *J. Manuf. Process.* **2023**, *105*, 14–26. [CrossRef]
24. Lapointe, S.; Guss, G.; Reese, Z.; Strantza, M.; Matthews, M.J.; Druzgalski, C.L. Photodiode-Based Machine Learning for Optimization of Laser Powder Bed Fusion Parameters in Complex Geometries. *Addit. Manuf.* **2022**, *53*, 102687. [CrossRef]
25. Taherkhani, K.; Sheydaeian, E.; Eischer, C.; Otto, M.; Toyserkani, E. Development of a Defect-Detection Platform Using Photodiode Signals Collected from the Melt Pool of Laser Powder-Bed Fusion. *Addit. Manuf.* **2021**, *46*, 102152. [CrossRef]
26. Ketkar, N.; Moolayil, J. Feed-Forward Neural Networks. In *Deep Learning with Python: Learn Best Practices of Deep Learning Models with PyTorch*; Ketkar, N., Moolayil, J., Eds.; Apress: Berkeley, CA, USA, 2021; pp. 93–131. [CrossRef]
27. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

28. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016. [CrossRef]
29. Hastie, T.; Friedman, J.; Tibshirani, R. Model Assessment and Selection. In *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*; Hastie, T., Friedman, J., Tibshirani, R., Eds.; Springer: New York, NY, USA, 2001; pp. 193–224. [CrossRef]