Additive Manufacturing 111 (2025) 104982

Contents lists available at ScienceDirect

Additive Manufacturing

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

Research paper

Digital 3D defect maps: Detecting localised porosity with high-speed melt pool imaging data in LPBF

Patrick L. Taylor [a,1], Richard J. Williams [a,1], Henry C. de Winton [a,1], Vincent Fernandez [b], Sebastian Larsen [a], Paul A. Hooper [a,*]

[a] Department of Mechanical Engineering, Imperial College London, Exhibition Road, London, SW7 2AZ, UK
[b] The Natural History Museum, Cromwell Road, London, SW7 5BD, UK

ARTICLE INFO

Keywords:
Porosity detection
Laser powder bed fusion
In-process monitoring
Machine learning
Additive manufacturing

ABSTRACT

Adoption of metal additive manufacturing for critical applications is hindered by the costs of post-build quality inspection. In-process monitoring offers a promising alternative by enabling parallel construction of digital 3D defect maps for every component manufactured. In this work, we present a system to detect local regions of porosity, containing both keyhole and lack-of-fusion defects, in laser powder bed fusion parts. A coaxial high-speed melt pool imaging setup operating at 20 kHz acquires feature-rich data, capturing images approximately every 37.5 µm along scan tracks and records over 30 million melt pool images per hour of build time. Using these data, a gradient-boosted decision tree model is trained to classify porosity levels in localised 2 mm voxel bins. The system achieves a state-of-the-art detection threshold of 0.11% porosity, defined by the standard non-destructive evaluation criterion of 90% probability of detection at 95% confidence. By training on datasets containing realistic, organically generated porosity and demonstrating the most accurate localised porosity detection yet reported, this work represents a significant advance towards practical, industrially relevant in-process defect detection for additive manufacturing.

1. Introduction

In-process monitoring systems for metal additive manufacturing (AM) processes have long been promised to provide online assurances of part quality and rapid component qualification. However, the reality is that widespread adoption in industrial settings has been slow and limited. Data from monitoring systems are large, complex, and difficult to manage, and it is often not clear how to turn this data into robust decision making tools. Other factors that hamper adoption include a lack of rigour in defect detection performance characterisation and disregard to the generalisability of detection methods between both builds and machines [1]. Frequently, monitoring systems are trained on datasets in which defects have been introduced in an artificial manner that is not reflective of natural defect formation encountered in practical use cases.

Increased porosity is known to be detrimental to many mechanical properties and it is therefore a primary indicator for part quality [2]. Common industrial practice is to include at least one specimen for porosity measurement. These specimens can provide assurance that no gross process failure modes have occurred but give limited assurance to the porosity content in the part of interest since local differences in porosity can occur due to both build geometry and process variations [3]. Consequently, structurally significant porosity may go undetected without detailed post-manufacture inspection of every part.

In this work, we aim to address some of these challenges by presenting a new in-process monitoring and classification system for laser powder bed fusion (LPBF). Our approach meets industrially relevant non-destructive evaluation (NDE) criteria for localised porosity detection and is further tested for generalisability across repeated builds on the same machine.

1.1. Monitoring approaches

In-process monitoring systems for LPBF have been developed extensively in recent years [4–8]. They present the possibility of constructing real-time digital models of the defects contained within a part as it is being built, enabling assessment of process health, supporting rapid part qualification, and informing digital twin frameworks for managing the component throughout its operational life [9].

* Corresponding author.
E-mail address: paul.hooper@imperial.ac.uk (P.A. Hooper).
[1] These authors contributed equally to this work as co-first authors.

https://doi.org/10.1016/j.addma.2025.104982
Received 15 April 2025; Received in revised form 29 September 2025; Accepted 1 October 2025
Available online 9 October 2025
2214-8604/© 2025 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Table 1
A summary of defect generation approaches in porosity detection literature.

Type	Subtype	Studies
Naturally induced porosity		[11,16,21,25–28]
Sub-optimal laser parameters	Power	[16,24,25,28–36]
	Speed	[16,25,28–33,35–40]
	Hatch spacing	[16,23,25,28–30,39,40]
	Focus	[1]
Organic porosity		[3,21,35,41–46]

The most sophisticated systems employ high-fidelity sensors and complex post-processing to measure localised build quality [10,11]. However, these studies have not recorded entire builds on industrial length and time scales. This limitation is compounded by the fact that production builds usually achieve high density ($\geq$ 99.9%), meaning that defects are rare and require very large data volumes to reliably capture. For example, Goosens et al. developed a system capable of long recording durations [12], but without explicit xy position tracking. The data were instead used to extract melt pool width and predict depth rather than to identify defects directly.

Commercially, the most production-ready systems are those integrated into LPBF machines by equipment manufacturers [7,8,13]. Building on early work by Kruth et al. (2007), Concept Laser introduced the QM melt pool system in 2010, which uses a coaxial camera alongside a photodiode for melt pool intensity measurement. These signals can be combined to generate build quality estimates [14–20]. EOS machines, in contrast, incorporate fixed, high-resolution NIR cameras [8], while other commercial monitoring systems rely primarily on photodiode-based measurements [5,7,8].

Beyond production readiness, in-process monitoring approaches can also be distinguished by their spatiotemporal resolution [8,13]. Four categories can be defined in a hierarchy of length scale: melt pool, track-wise, layer-wise (i.e. part-scale), and powder-bed (i.e. process-scale) monitoring. At the lowest scale, the melt pool can be directly imaged using cameras [21–23]. The resulting data volumes are large and typically reduced to a limited set measurements [16,19,20,22,23], which may serve either as direct quality indicators [20] or as inputs to models of build quality [16]. Alternative strategies apply non-physical models such as convolutional neural networks (CNNs), classifying melt pool images either into processing-parameter regimes with anticipated links to quality [24], or into categories directly related to part mechanical properties [18].

1.2. Porosity generation methods

Studies on porosity detection in LPBF differ in how porosity is generated. These approaches can be divided into three categories: designed pores, sub-optimal laser parameters, and naturally induced porosity. Table 1 summarises the distribution of these methods across studies in the literature.

Designed pores are introduced by creating voids in the part geometry, so the laser does not scan small regions within the part [16,21,25–27]. In-process detection of such pores is trivial while the laser is off, but models can be trained to identify these regions in subsequent layers scanned with standard parameters. This approach is commonly used to benchmark defect detection methods due its repeatability, but it does not reflect natural defect formation.

Sub-optimal laser parameters are used to induce defective melting modes and generate pores. Variations in laser power, speed, hatch spacing, or focus create localised defects that can be observed through changes in melt pool behaviour rather than direct measurement of printer outputs. While direct monitoring of laser parameters is straightforward, melt pool monitoring demonstrates the system's ability to infer process outcomes from physical signatures. Some studies combine sub-optimal parameters with designed pores to induce defects in small regions of the build [28]. In practice, industrial LPBF machines maintain fixed parameters during a build, so large deviations typically only occur under gross malfunction. Carter et al. demonstrated that models trained on such artificially induced porosity distributions can still predict defects under normal conditions, achieving measurable improvements in build quality through parameter feedback control [47,48].

Naturally induced porosity (NIP) refers to defects that form while the process operates under optimal parameters and no machine faults are present. This method provides the most realistic representation of defect formation for in-process melt pool monitoring systems. Although waiting for stochastic defects is infeasible, the frequency and type of porosity can be influenced by build layout, part geometry, layer time, and part placement on the build plate [3,13,35,46,49–52]. By deliberately designing builds to exploit these factors, it is possible to induce porosity in a controlled yet natural manner. Snow et al. used a spatter generator to correlate in-process spatter observations with ex-situ porosity measurements, demonstrating that NIP builds can train predictive models [46,51]. Bevans et al. varied part geometry and laser parameters to study heat dissipation effects, finding that lack-of-fusion porosity occurred primarily under non-optimal laser conditions [35].

By focusing on naturally induced porosity, monitoring systems can be developed and validated on defects that resemble those occurring in real production conditions, ensuring that predictive models are both realistic and industrially relevant.

1.3. Detection resolution

Identifying the spatial location of defective material is a key aspect of any defect detection system. In the literature, reported detection resolutions vary widely. Many studies quantify porosity on a part-wide basis [1,24,30,39,40], or over large sub-regions of a part [3,37,38,44]. Detection has also been reported at the layer level [40] or within sub-regions of a layer [44]. An alternative approach is to estimate mean or maximum pore size within a region, for example using major axes measured in μm [32]. Such metrics have the advantage of direct relevance to fracture mechanics and failure prediction.

However, recent analyses suggest a more nuanced picture. Joshi et al. found that overall porosity correlates more strongly with yield strength and failure strain in LPBF Ti–6Al–4V than maximum pore size [53]. They further showed that ligament parameters, the ratio of pore size to pore spacing, were particularly predictive of failure strain. These findings indicate that both the level and geometric characteristics of localised porosity influence local and part-wide mechanical properties, as well as failure modes in AM components.

In terms of spatial resolution, the finest scales of in-process data collection are typically limited to tens of microns, constrained by optical resolution or acquisition frequency [21,25,41,42]. Ren et al. [54] demonstrated very fine spatial and temporal resolution in synchrotron X-ray experiments, capturing keyhole instabilities in real time with a machine learning model achieving 100% prediction accuracy. However, defects in built parts are often delocalised from the originating melt pool, true for both keyhole and lack-of-fusion defects. Detection methods based on data binning accommodate this delocalisation, even though they have lower nominal resolution.

Voxel-based approaches have been used to represent parts at high resolution. Data are binned into voxels, typically sized so that each voxel bin is either defective (belongs to a pore) or not (belongs to solid material). This produces outputs analogous to binarised μCT data, allowing precise measurement of pore size and position. A major challenge is the abundance of dense material, which can make up over 99% of even relatively poor parts, creating a high risk of false positives [25]. For high-resolution voxel maps, reported Probability of False Alarm (PFA) values range from 0.27 to 0.5 depending on the defect generation method [25,40], while some studies do not report PFA at all [21,33].

Lack-of-fusion pores are typically flat, aligning with build layers, rather than spherical like keyhole porosity [50,55]. While their lateral extent can be large, their thin geometry means they occupy a relatively small volume. At very fine voxel bins scales, such defects may be fragmented across multiple voxel groups, complicating reliable detection, increasing the risk of false positives, and underestimating defect size. Using larger voxel bins mitigates this issue, allowing these structurally significant pores to be captured accurately while balancing sensitivity and precision [25]. Voxel bin size can therefore be tuned to optimise detection performance for industrial monitoring.

1.4. Machine learning for detection

Machine learning (ML) has become a key tool in AM in-process monitoring. A recent review by Herzog et al. estimated that 77% of studies employ supervised learning, though there is a wide variety of model architectures and input data sources across both powder bed fusion and directed energy deposition research [56].

Among deep learning approaches, convolutional neural networks (CNNs) and convolutional autoencoders (CAEs) have proven particularly popular due to their ability to extract latent features from high-dimensional imaging data without manual feature engineering [51,57–59]. Previous studies using similar monitoring hardware to that employed here have leveraged semi-supervised deep learning to detect non-optimal processing conditions, embedding melt pool images into latent spaces using CNNs [11]. Other neural network models, including back-propagation neural networks (BPNNs), multi-layer perceptrons (MLPs), deep belief networks (DBNs), and recurrent architectures such as LSTMs, have been explored, often as benchmarks for comparison against more complex models [27,36].

A limitation of neural networks is their limited interpretability and training complexity. Some studies have addressed this by extracting physically interpretable features from melt pool data: for example, using geometric considerations to estimate lack-of-fusion defect formation [58]. These features can then be used as inputs to simpler models such as k-nearest neighbours (k-NN) [35], random forests (RF) [28], or support vector machines (SVMs) [36], which offer greater interpretability and reduced training complexity.

Evaluation metrics are critical for meaningful comparison. De Winton et al. introduced the 90% probability of detection with 95% confidence (90/95 POD) metric to AM in-process monitoring, adapting a standard from the wider field of non-destructive evaluation (NDE) [1,55,60]. Compared with common ML metrics such as ROC-AUC, F1 score, or confusion matrices, the 90/95 POD is more directly interpretable in a part qualification context, representing the likelihood of detecting defects of a given size with high confidence. ROC-AUC retains value in illustrating the trade-off between probability of detection (POD) and probability of false alarm (PFA), which is important for balancing defect detection with resource efficiency in AM.

1.5. This work

We present a system for in-process detection of localised porosity in LPBF parts using a coaxial high-speed camera. Porosity arises naturally from part geometry and build layout, without manipulating laser parameters, creating industrial-relevant processing conditions. Multiple machine learning models were trained to classify porosity at a 0.05% threshold, roughly an order of magnitude lower than previously reported. The best performing model meets the 90/95 POD requirement at 0.11% porosity, representing one of the most sensitive in-process detections of localised porosity reported to date.

The paper is structured as follows. First, we describe the in-process monitoring hardware and data processing pipeline. Next, we detail the LPBF build layout and processing parameters. Ground truth data collection using μCT is then presented, followed by a description of the machine learning algorithms. In the results section, we present and analyse the ground truth data, followed by quantitative evaluation and performance metrics for the detection model, including performance across repeated builds. Finally, we discuss the relevance of these results to industrial process qualification and compare our findings to previous studies.

2. Materials and methods

Our detection system is built around a high-speed coaxial camera system to continuously monitor the melt pool throughout a build. The system includes a pipeline for efficiently offloading and preprocessing the large volumes of imaging data collected, as well as segmenting and extracting features to create a more manageable dataset. The system is hereafter referred to as the Assurance Through Layer-wise Anomaly Sensing (ATLAS) system (Section 2.1).

Two repeated LPBF builds were monitored using ATLAS. To ensure industrial relevance, the parts and their layout on the build plate were designed to induce realistic, naturally occurring porosity arising from part geometry and build layout (Section 2.2). Selected parts from each build were characterised ex-situ using μCT, providing the ground truth dataset for ML defect classification (Section 2.3).

For ML analysis, the data were grouped into voxel bins to calculate statistical measures, forming the input dataset. Corresponding μCT data were binarised and porosity values computed for each voxel bin. After thresholding into defective and non-defective voxel bins, these formed the target dataset. Data preprocessing is described in Section 2.5, with training and validation of the ML model detailed in Section 2.6. An overview of the complete workflow is shown in Fig. 1.

2.1. ATLAS in-process melt pool monitoring system

2.1.1. Hardware

A high-speed melt pool imaging system was used to monitor the LPBF process. Fig. 2 shows the optical train for the ATLAS system, adapted from previous work [1,61]. A coaxially mounted high-speed streaming camera captures the melt pool, with light filtered using a 950 nm narrow bandpass filter to mitigate chromatic aberrations.

The camera has a resolution of 128 × 128 pixels with 8 bit depth. A projected pixel size of 31.5 μm on the build plate gives a 4 × 4 mm field of view. ATLAS records images at 20 kHz with a 20 μs exposure time, along with laser power, focus, and (x, y, z) coordinates for each layer. At an approximate scan speed of 750 mm/s, this corresponds to a frame spacing of 37.5 μm along the scan track.

The camera permits continuous recording at this frame rate, with no theoretical duration limit. Consequently, entire builds can be captured at full frame rate, generating over 1 TB of data (≈72 million melt pool images) per hour of laser on-time. This enables full-resolution tracking of melt pool behaviour across realistic industrial builds, capturing rare defect events.

2.1.2. Feature engineering

Fig. 3 shows an example coaxial image, with raw intensity in Fig. 3(a) and the segmented result in Fig. 3(b). Each image is segmented into three regions based on intensity and connectivity, with intensity serving as a proxy for melt pool features as follows:

- High-intensity spot (red in Fig. 3(b)): estimated to include the vapour depression at the laser contact point.
- Low-intensity surround (blue): representing cooler features such as the melt pool tail and vapour cloud. The balance between these depends on the prevailing melting mode.
- Spatter particles (green): identified as disconnected regions not belonging to the main high- or low-intensity areas.

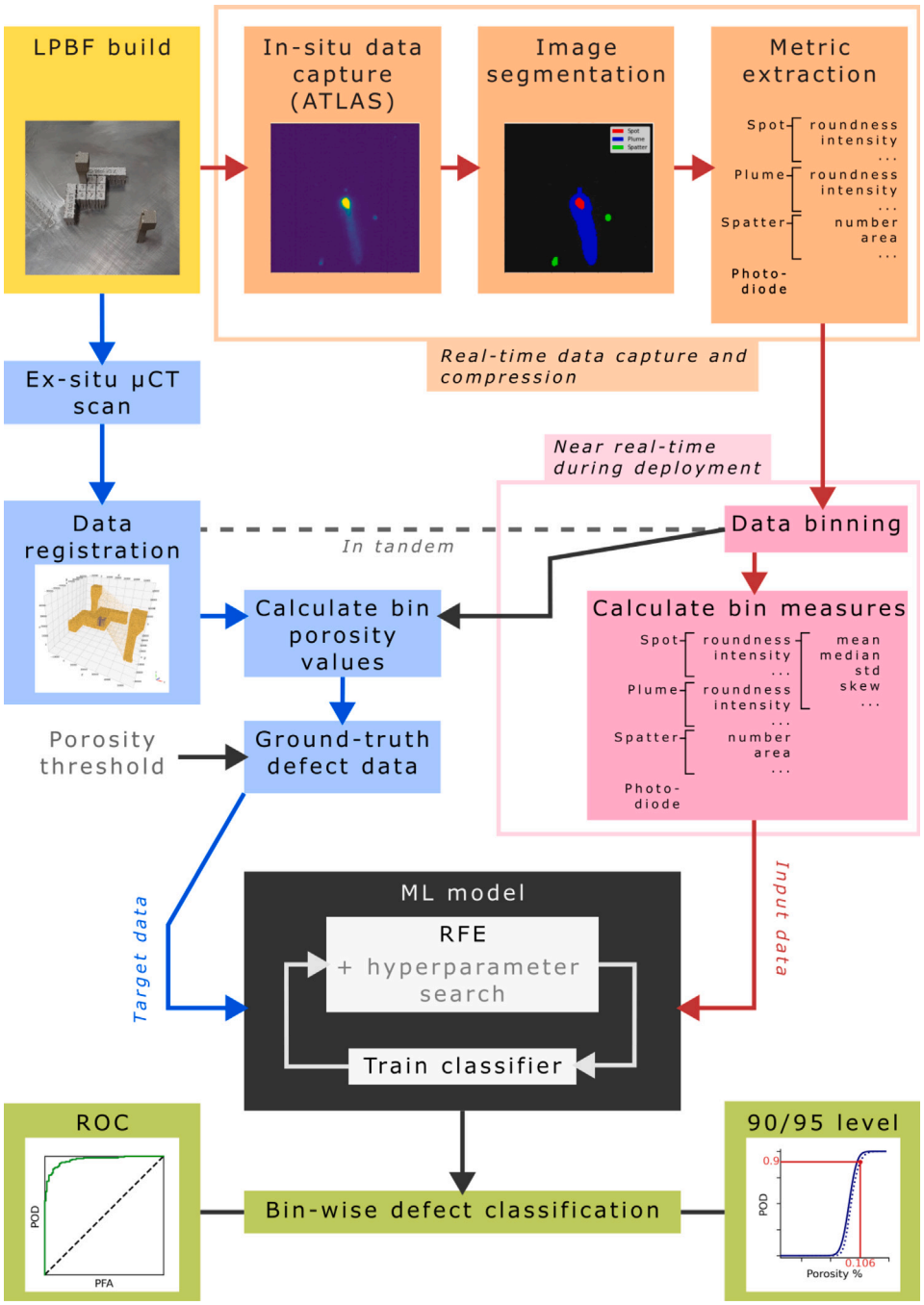


Fig. 1. Overview of the in-process defect detection method in this paper.

These segments are not direct physical boundaries such as the true liquidus–solidus boundary, but are likely to be distinct for different melt pool behaviours and vary when defects form. From each image, 21 geometric and intensity-based metrics were extracted across the three regions (listed in Table 2), including one metric to approximate the response of a photodiode by summing all the pixel values in an image. Spatter was quantified by averaging across disjointed particles so the number of metric did not vary from image-to-image. This feature extraction reduced the raw imaging data by ≈ 10×, yielding ≈ 10 GB of engineered feature data per hour of laser on-time.

2.2. Experiment design

A central aim of this study was to evaluate defect detection under conditions representative of industrial practice. To this end, porosity was induced naturally by exploiting part geometry and build layout, rather than by altering laser processing parameters. The build was therefore designed to promote a range of realistic defect mechanisms while maintaining manufacturer-recommended parameters for stainless steel 316L.

Fig. 4 shows the layout of the parts on the build plate. Three simple geometries were used: cubes, horizontal bars, and overhang columns (Fig. 5). Each part incorporates small registration features to enable precise alignment of in-process data with µCT scans, and all parts were sized to fit within the µCT scanner without sectioning.

Four defect inducing build features, which are also likely to be present in industrial builds, were targeted through part design and placement:

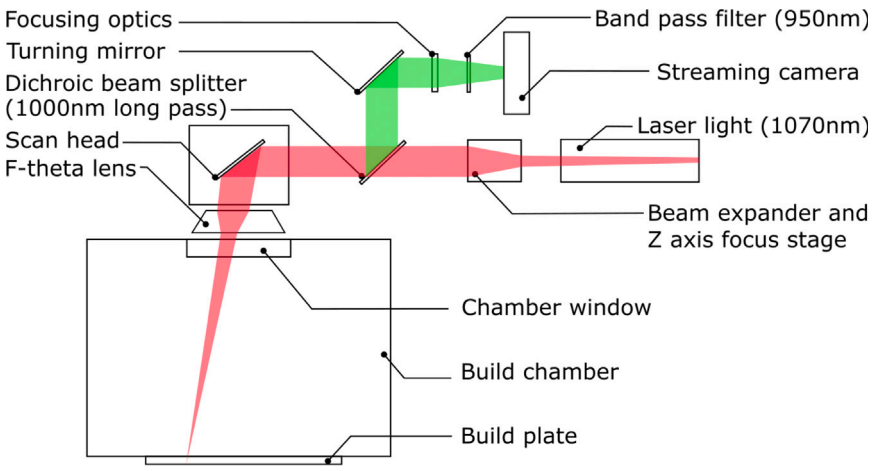
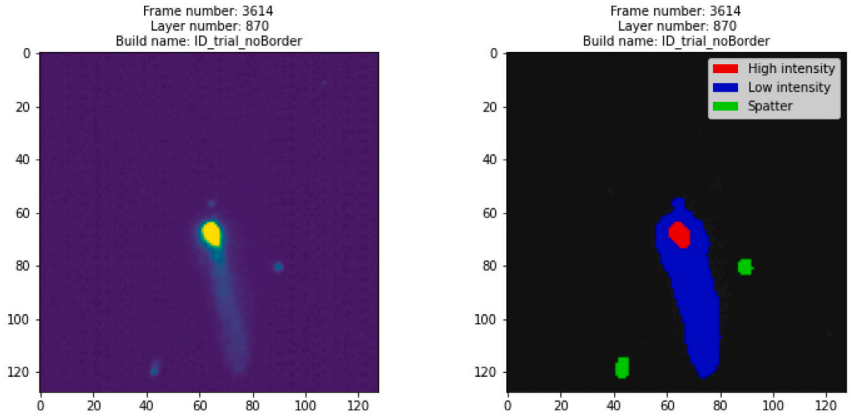


Fig. 2. Coaxial imaging optics showing the original optical path in red and the imaging optical path in green. Reproduced from [1] with permission.



(a) Raw coaxial camera image (units are pixels). (b) Segmented coaxial camera image (units are pixels).

Fig. 3. Raw coaxial camera image and segmented equivalent highlighting the three regions from which metrics were calculated. The field of view on the powder bed is 4 mm × 4 mm.

Table 2
The metrics calculated for each captured image.

Segmentation region	Metric
High-intensity spot	Roundness
	Orientation (angle)
	Max. intensity
	Mean intensity
	Area
	Minor axis length
	Major axis length
Low-intensity surround (melt pool tail and plume)	Roundness
	Min. intensity
	Max. intensity
	Mean intensity
	Area
	Minor axis length
	Major axis length
Spatter particles	Mean of mean intensities
	Mean area
	Median of mean intensities
	Median area
	Total area
	Number
Photodiode equivalent	Sum of all pixel values in image

- Inter-layer cooling promotes lack-of-fusion type defects by causing low workpiece temperatures, reducing melt pool size. Increasing the scanned area in the lower 10 mm of the build gives the tall overhang columns more time to cool between layers.
- Intra-part spatter is where spatter particles land back onto the part's surface, blocking the laser and causing spatter induced lack-of-fusion porosity. This is promoted by the horizontal bar geometry with a large surface area in the gas flow direction.
- Inter-part spatter is a similar mechanism but the spatter comes from neighbouring parts. Nesting the cube parts up-stream and in close proximity encourages spatter to land onto the overhang column part, resulting in further spatter induced lack-of-fusion porosity.
- Thermal bottlenecks promote keyhole porosity by locally increasing workpiece temperature, increasing the depth of the melt pool depression due to increased vaporisation. In the top overhang section of the column part the scan area is larger than the cross-sectional area in the lower section, constraining cooling and increases the build surface temperature as heat accumulates.

Laser parameters were kept fixed at manufacturer-recommended values for 316L Stainless Steel (see Table 3), with border scanning disabled but standard top- and down-skin settings retained to ensure robust build initiation and clean registration features. Border scans

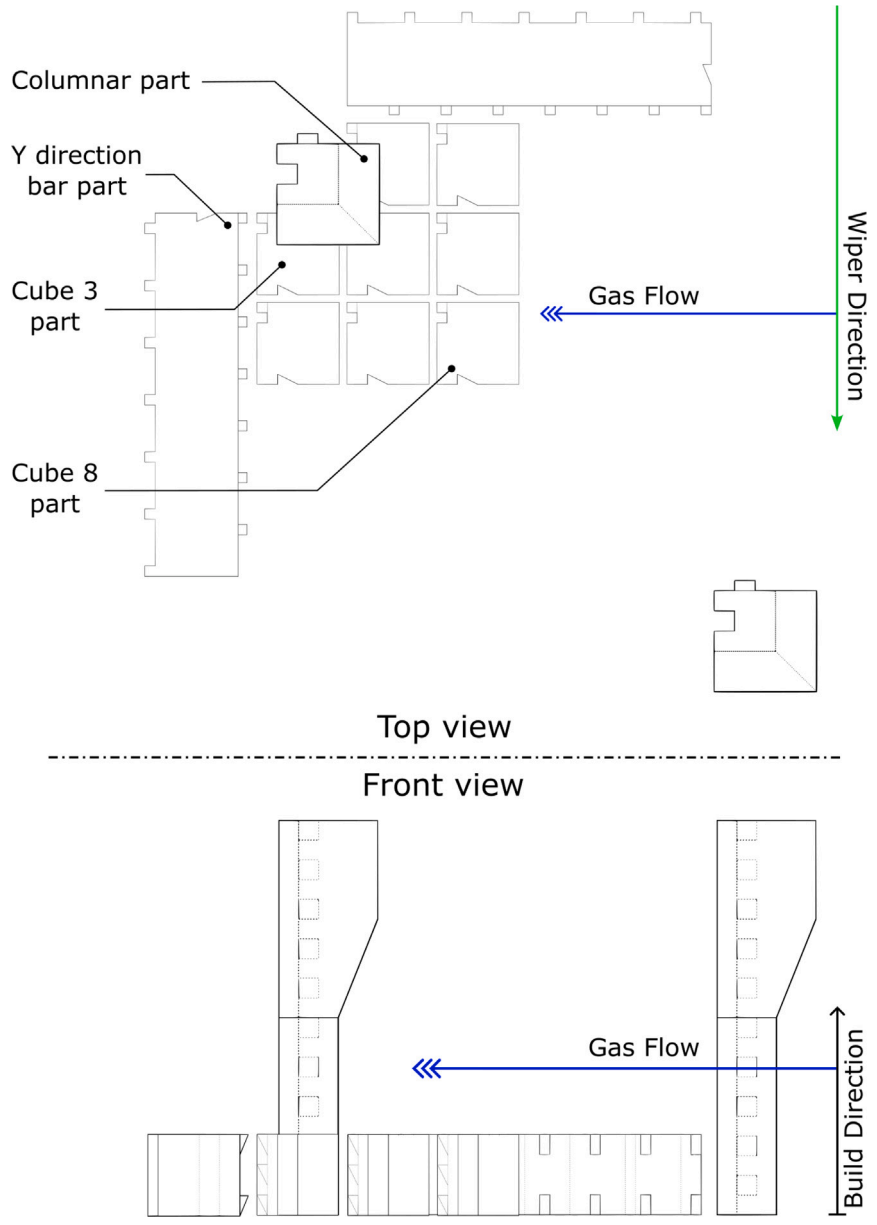


Fig. 4. Layout of the overhang columns, cubes, and horizontal bars in the build. The small regular features on the outside of each part are to assist with registration with μCT, breaking symmetry and allowing both matching of layer number and confirmation of orientation.

were omitted to avoid introducing distinct melt conditions that would complicate model training. The Renishaw stripe hatching strategy was used for all parts [62]. Any porosity was induced into the parts solely in a natural manner, and not by adversely manipulating the processing parameters.

Each build consisted of 900 layers and took approximately three hours. The experiment was repeated twice (build A and build B), producing two independent datasets to allow transferability of models to be evaluated. In total, more than 200 million melt pool images were collected (over 4 TB of raw image data), reduced to ≈100 GB after feature extraction.

Table 3
Processing parameters used to manufacture the samples. Note that border scans were omitted.

Parameter	Value	Unit
Laser power	200	W
Point distance	60	μm
Exposure time	80	μs
Hatch spacing	110	μm
Spot size	65	μm
Layer thickness	50	μm

2.3. Ground truth data and alignment

Ex-situ μCT was used to obtain ground truth porosity for a subset of the printed parts. The resulting volumetric data were aligned with in-process monitoring data to enable voxel-wise comparison of porosity and image-derived features.

A key reason for adopting a voxel-wise approach is to mitigate the challenges associated with registering individual pores. Natural variation in layer height and the remelting of previous layers (which can both repair pores and create new ones below the monitored layer) introduce uncertainty into the ground-truth data and its alignment with the in-process data. This uncertainty weakens the training datasets, as

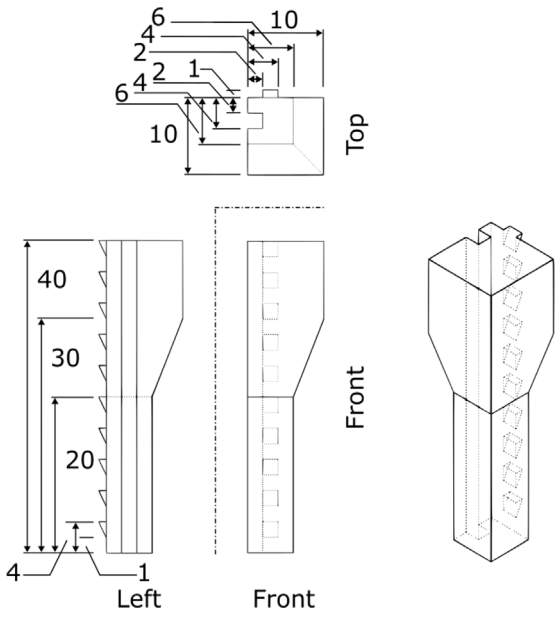


Fig. 5. Design of the overhang column part, where the change in cross-section acts as a thermal bottleneck. Again, the small regular features on the outside of each part are to assist with registration with μCT. All dimensions are in mm.

valid detections in the in-process monitoring data may not appear in the ground-truth data, and conversely, defects present in the ground-truth data may show no corresponding signal variation on that layer. By avoiding the registration of individual pores and instead calculating voxel-wise statistics, we avoid these inconsistencies.

2.3.1. μCT scanning

Scans were performed on a Nikon Metrology XT H 225 ST using a rotating reflection tungsten target at 215 kV and 47 μA (10 W). A 1 mm tin filter was applied to minimise beam hardening, and the source-object and object-detector distances were set to 31 mm and 748 mm, respectively. This yielded a 6 μm pixel size, sufficient to resolve lack-of-fusion and keyhole porosity. The exposure time was 708 ms.

From each build, four parts were scanned: one overhang column, two cubes, and the y-direction bar (Fig. 4). Scans were reconstructed and binarised, producing solid and porosity volumes for each part. Optical microscopy was used to confirm the fidelity of the scans.

2.3.2. Registration and voxelisation

In-process data and μCT volumes were registered prior to voxelisation using the designed registration features. A Python GUI was developed to aid manual alignment (Fig. 6). The alignment precision was limited by the resolution of the two data sources, giving uncertainties of ±50 μm in z and ±6 μm in x–y. However, the largest uncertainty was caused by surface roughness in the manual alignment step, and added an estimated 100 μm in z and 25 μm in x–y. Overall this corresponds to a maximum possible error of 5.3% in volume overlap at a 2 mm voxel bin scale.

Both datasets were transformed into equivalent voxelised form. For μ CT data, average porosity values were calculated per voxel bin. For in-process monitoring, image metrics were binned and summarised (mean, median, mode, mode count, standard deviation, skew, kurtosis), producing a 189-feature dataset. The voxelisation reduced the dataset size to ≈50 MB per build, depending on voxel bin size.

Voxel bin alignment error is not expected to have a significant effect. Assuming a worst case scenario, where a voxel bin with a level of porosity that is deemed to be unacceptable (taken to be 0.05% porosity as discussed in Section 2.4) neighbours a voxel bin with no porosity, the truth data may underestimate the true value by 0.003% porosity at most. In contrast, if the same alignment uncertainty was applied to register individual defects of 200 μm diameter, there would be at least a 50% uncertainty in volume overlap. This is one of the key advantages of the voxelised approach, particularly at small defect sizes.

2.4. Defect equivalencies

To support comparison across different NDE techniques, it is useful to relate voxel bin porosity levels to equivalent defect sizes. The distribution of porosity within a voxel bin can vary considerably. In the most benign case, porosity is dispersed as many small, isolated pores. In the most critical case, the same porosity is concentrated within a thin planar feature spanning the voxel bin, essentially forming a crack. While the latter could in principle be estimated from porosity fraction, the resulting "crack thickness" is not directly meaningful for fracture mechanics, where crack length is the governing parameter.

As a practical middle ground, we define an equivalent spherical defect (ESD): the diameter d of a single sphere that would account for the total porosity ϕ in a voxel bin of volume l^3:

$$d = l\left(\frac{6\phi}{\pi}\right)^{\frac{1}{3}} \qquad \phi = \frac{\pi}{6}\left(\frac{d}{l}\right)^3 \tag{1}$$

For instance, 0.05% porosity in a 2 mm voxel bin corresponds to a 200 μm diameter spherical pore. This assumption is most representative of keyhole pores, which are approximately spherical. Lack-of-fusion pores, however, are typically flatter and more irregular, and may exceed the equivalent spherical dimension in one direction. This definition is therefore a convenient approximation for cross-technique comparison, but caution is warranted if it is used in a fracture mechanics context.

2.5. Defective voxel bin definition

Ground truth voxel bin porosity is measured as a continuous fraction between 0 and 1. To enable supervised learning, these values were discretised into two classes by applying a porosity threshold: voxel bins with porosity above the threshold were labelled defective, and those below it non-defective. Multiple threshold values were investigated.

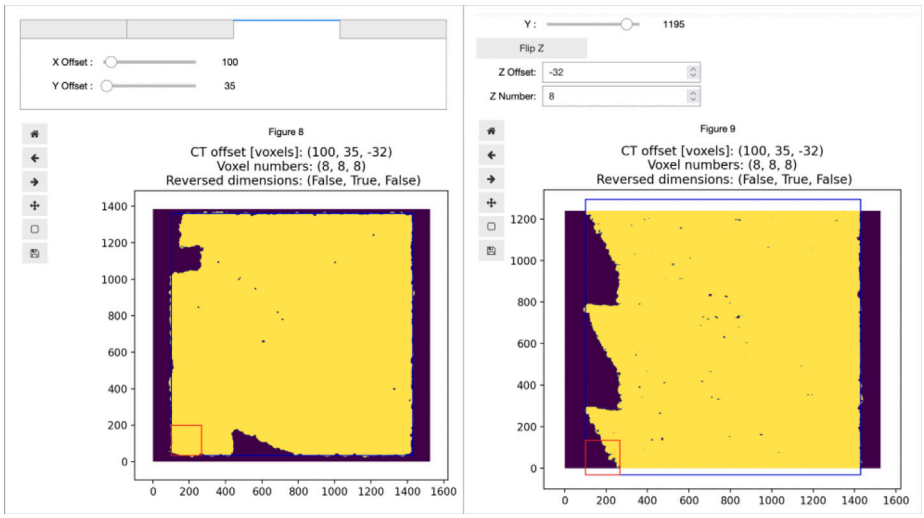
Defining an industrially relevant porosity threshold is challenging, since the acceptable defect level depends strongly on the application. To provide context, we compare our thresholds to the capabilities of established NDE techniques. Reported minimum detectable pore sizes in LPBF include 200 μm for radiography [63] and a 230 μm 90/95 value for μCT (validated against synchrotron tomography) [64]. However, what is achievable in practice could be significantly different depending on many factors such as source intensity, material, part geometry and size

We have selected a 0.05% porosity threshold in a 2 mm voxel bin, corresponding to an equivalent spherical defect of approximately 200 μm in diameter (see Section 2.4), placing our in-process monitoring approach in line with these traditional NDE benchmarks.

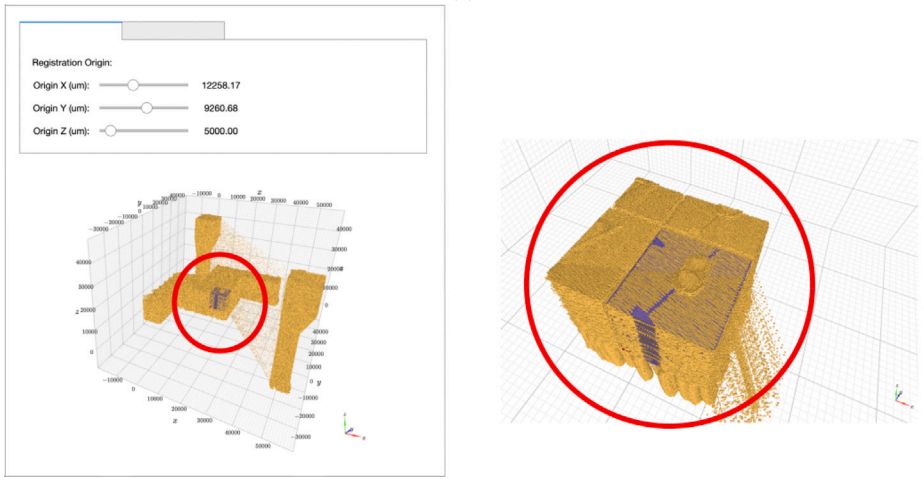
This pore size also approaches the scale of the melt track width, implying that most defects at this level are likely to be lack-of-fusion pores. Such a definition of defective voxel bins therefore provides both practical relevance and comparability with NDE standards such as ASTM E1742 (radiography) and ASTM E1695 (μCT) [65,66].

2.6. Machine learning model

We frame ground truth porosity y as the outcome of a stochastic process governed by design factors D, machine operation parameters, and semi-random variables ε such as local powder distribution. Our ATLAS system records melt pool metrics X, which are causally linked to defect formation (Fig. 7). In realistic use cases, D and machine



(a)



(b)

Fig. 6. The GUI used for manual registration of the CT and in-process monitoring datasets. 6(a) CT registration GUI. 6(b) Metrics point cloud registration GUI (left), alongside a close-up image of part of the point cloud corresponding to a CT scanned region.

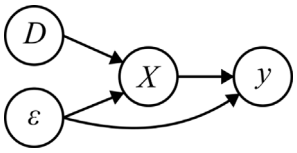


Fig. 7. The causal relationship between observed and controlled variables.

settings are controlled, leaving X as the observable proxy for the unmeasured factors driving porosity. Our objective is therefore to model the relationship $X \rightarrow y$.

While some individual melt pool metrics correlate with porosity, none do so strongly enough to be reliable indicators [1]. We therefore apply ML to map the 189 statistical features extracted from voxelised bins (Section 2.3) to the binary porosity labels defined in Section 2.5.

A range of classifiers were screened: logistic regression as a simple baseline, a multi-layer perceptron (MLP) [67], a Gaussian process classifier (GPC) [68,69], and a gradient boosted decision tree (GBDT) [70]. For each model a thorough grid-search over relevant hyperparameters was carried out, with hyperparameters selected using cross-validation and optimisation of receiver operating characteristic area under the curve (ROC-AUC) score. Model selection was carried out using the combined dataset of build A and build B, which we call dataset C. A standard procedure of withholding 20% of the data for testing was used (datasets C-train/C-test). To assess generalisability, inter-build validation was also performed: models trained on one build were tested on the other (A-train/B-test and B-train/A-test).

Since the datasets contained a large number of features, great care was taken to address the possibility of model overfitting. Model pipelines included a stage that filtered out features with zero variance, and also features that were highly correlated to another feature in the dataset (after applying a standard scaling). For the 2 mm cubic bin dataset obtained in Section 2.3, this approximately halved the number of input features. Next, a recursive feature elimination (RFE) strategy was used for the logistic regression, GBDT, and GPC models. This entails training a model, selecting the most important features, then retraining the model using just those. The number of selected input features was included as a hyperparameter to be searched over. For the MLP, where feature importance is more difficult to obtain, input features pre-selected by the GBDT model were used. Full details of the respective model pipelines and RFE strategies are provided in a supplement.

In all cases ROC curves were used to quantify classification performance and the trade-off between detection and false alarms. Building on previous work by de Winton et al. [1] and Snow et al. [51], a POD curve was constructed to assess detection capability at different levels

Table 4
Overview of datasets and binning into 2 mm voxel bins. Porosity values are a summary of ground truth from μCT.

Dataset	Raw images	No. 2 mm voxel bins			Mean porosity (%)
		Total	Porosity		
			> 0.05%	> 0.2%	
A	28,140,000	670	346	10	0.059
B	28,392,000	676	441	189	0.138
C (A+B)	56,532,000	1,346	787	199	0.099

of porosity. Following the method of Harding et al. only the lower confidence interval on the POD curve is calculated [71]. The 90/95 metric derived from the POD [55,60] was used as the benchmark for assessing the industrial readiness of the proposed in-process monitoring system.

3. Results

3.1. Datasets and ground truth analysis

Table 4 shows an overview of the resulting datasets after the μCT scans and in-process data were registered and voxelised into overlapping voxel bins. The datasets were derived from approximately 28 million melt pool images per build (less that the total data acquired as not all parts were scanned by μCT). The 2 mm voxel bin size covers an average of 42,000 individual melt pool images per voxel bin, resulting in approximately 670 total voxel bins per build. Each voxel bin is described by 189 statistical features (derived from the 21 feature metrics extracted from the image data), giving over 250,000 data points for ML training and testing across both builds. Voxel bin sizes above 2 mm were not examined since this left too few data points to reliably train and test models. With average porosity levels below 0.2%, both builds represent dense, production grade samples. A histogram of the comparative datasets is shown in Fig. 8

A comparison between voxel bin porosity values and pore sizes is shown in Fig. 9. In this figure, pore sizes have been calculated using the length of the major axis of the ellipse that has the same normalised second central moments as the pore region as calculated from raw μCT voxels [72]. The figure shows a clear correlation between the maximum pore size in each voxel bin and the porosity value. Only a single defect above 200 μm was observed, with the largest defect being 300 μm. This defect was contained in a voxel bin that had a mean porosity of 0.6%. A 2 mm voxel bin with 0.6% corresponds to an equivalent spherical diameter of 450 μm. These results suggest that the proposed ESD measure is conservative in estimating defect size from porosity.

Meanwhile, the plots of porosity against mean pore size and max pore size against mean pore size show two separate trends, present in both build A and build B. Here the low porosity trend region, showing a strong linear relationship between mean and maximum pore size with maximum values capped around ~ 50 μm, is indicative of a 'background' of keyhole porosity formation, with these particular data points corresponding to voxel bins with only keyhole porosity present. Build B also has both more voxel bins with greater levels of porosity and larger defect sizes when compared to build A, despite no known differences in build setup and procedure. This makes the datasets ideal for testing robustness and transferability.

Fig. 10 visualises porosity variation in both builds (1 mm voxel bins). The overhang column exhibits increased porosity in the chamfer region; this feature was designed to induce higher workpiece temperatures and hence lead to keyhole porosity. Likewise, the lower 10 mm of the column exhibits higher porosity, although it is difficult to disentangle the effects of inter-layer cooling and inter-part spattering on this phenomena. The *y*-direction bar exhibits weak correlation between *y*-direction porosity and adjacent cube placement, suggesting limited pore formation due to inter-part spattering, but does show a significant trend of increased porosity in the *x*-direction, away from the gas flow; likely due to intra-part spattering. Cube 8 was found to be more dense than cube 3 in both builds. Due to these cubes' placements with respect to other cubes and the gas flow direction, this also suggests that inter-part spattering was limited, or smaller than the effect of inter-layer cooling.

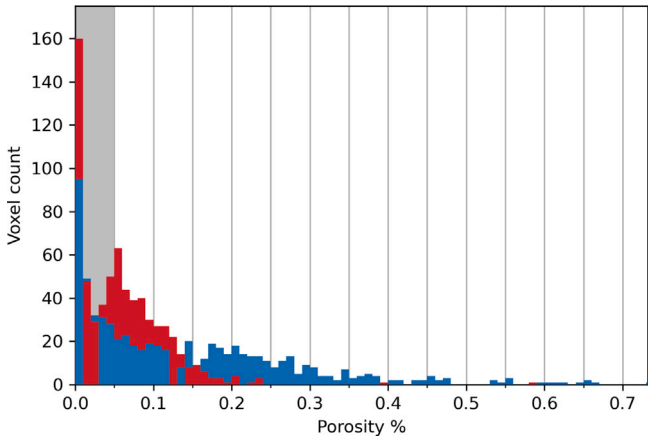


Fig. 8. Porosity distribution for build A (red) and build B (blue) from μCT. The grey region indicated the 0.05% porosity threshold that was used in the defect classification model.

The goal of employing the naturally induced porosity approach was to generate a realistic distribution of organic porosity, and *not* to induce each identified mode of pore formation in isolation. The preceding analysis satisfied us that we had successfully induced keyhole porosity and porosity due to inter-layer cooling, as well as some porosity due to intra-part spattering. Porosity varied widely across each part, sections of the parts, and also varied between builds. For these reasons, we were satisfied that we had met our goal and generated a suitable dataset for our modelling aims.

3.2. Machine learning model comparison

All of the models trained had comparable quantitative performance when trained and tested on the combined dataset C. They also had comparable qualitative performance when trained then tested on the inter-build datasets—every model was better able to transfer from training on dataset B to testing on dataset A, than the other way round (inter-build transfer is discussed in detail in Section 3.4). The gradient boosted decision tree (GBDT) model had a significant improvement in AUC score on C-test, attaining 0.969 vs 0.934 for logistic regression (2 mm voxel bin dataset with a 0.05% porosity threshold), as well as achieving a marginal improvement over the Gaussian Process Classifier (GPC) model with a value of 0.963.

Table 5 provides a comparison between the four ML models that were tested. The table is for the 2 mm voxel bin dataset with a 0.05% porosity threshold for defect classification. On ROC-AUC score, the three complex models (GBDT, multi-layer perceptron (MLP) and GPC) outperformed simple logistic regression. The MLP model outperformed GBDT and GPC on ROC-AUC score when trained and tested on the same dataset (both A and B), but scored worse on transfer learning. This suggests this highly flexible model was prone to overfitting. Additionally, measured on the crucial 90/95 metric, the MLP (0.142%) was worse than any other model, including logistic regression (0.107%). From this we can infer that the MLP model lacked sensitivity close to the decision threshold, likely a further contributing factor to its comparatively poor performance on the inter-build transfer learning tasks.

The GBDT and GPC were the two best models, with a trade-off in performance from score metric to score metric. The GPC model

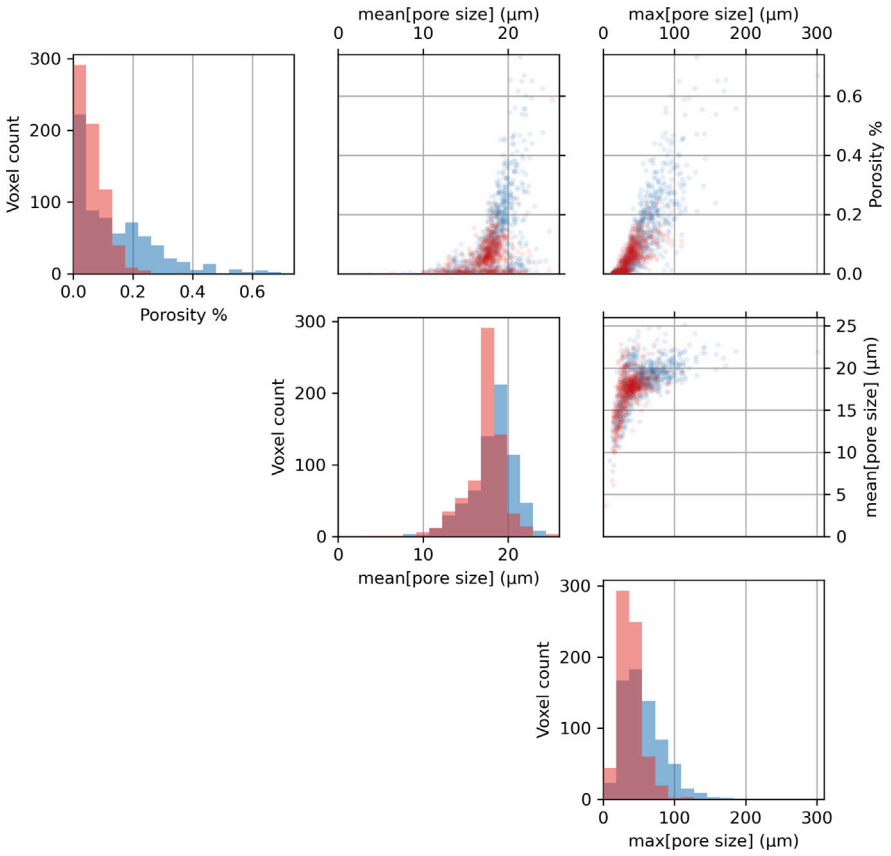


Fig. 9. Distributions and correlations between pore size and porosity of voxel bins in the 2 mm voxel bin datasets. Histograms are shown on the grid axis, with correlation plots off-axis. Build A data corresponds to the red bars and points, build B data to the blue. Note that the histogram from Fig. 8 is replotted here with larger porosity bins (top-left histogram).

outperformed GBDT on the 90/95 metric, which is the most relevant to practical application. It also performed better on transfer learning from the build B dataset to testing on build A; the test on which all models grossly underperformed (see Section 3.4). This might suggest that the GPC model is better able to generalise between dissimilar builds.

However, the GPC model comes with considerable limitations. Training and inference scale as $\mathcal{O}(N^3)$ with the number of data points, which would pose a serious limitation on this detection system's ability to scale-up to industrially relevant application. CPU training on a modest notebook-laptop with a 10-core intel i5 processor and 16 GB memory, the GBDT model was trained on dataset C, including full recursive feature elimination (RFE) with single feature increments, in around 15 min. Training the GPC took around 90 min, even after limiting RFE to ten feature increments. For these reasons, the GBDT model was selected for further analysis.

3.3. ML model performance by voxel bin size and thresholds

In Fig. 11(a) to 11(c) detection performance at a range of porosity thresholds and voxel bin sizes is shown by ROC curves produced for dataset C-test using GBDT models trained and validated on C-train. Larger porosity thresholds were not tested due to the low number of voxel bins with high porosity. AUC values are similar for both the 2 mm voxel bin models, and higher than for the 1 mm voxel bin model. However, the 2 mm voxel bin model with a 0.05% threshold has favourable behaviour in the low PFA region of the ROC curve, which is the desired region of operation for a practical monitoring and detection system. It achieves PFA = 0.088 for POD = 0.9, and a POD = 0.841 for PFA = 0.05 (superior to any other model that was tested). The confusion matrix for this model is shown in Fig. 12. From here on out we will refer to the GBDT model trained on 2 mm voxel bins with a 0.05% threshold as the optimal model.

Table 5
Comparison between four ML models: logistic regression (LR), gradient boosted decision tree (GBDT), multi-layer perceptron (MLP) and Gaussian process classifier (GPC). Each trained on the 2 mm voxel bin dataset with a 0.05% porosity threshold for defect classification. Each model was trained on 80% of the respective dataset with 20% withheld for testing, except when testing on across datasets, where the full dataset was used for testing. C refers to the combined dataset and A and B refer to the two respective builds. ESD stands for equivalent spherical defect.

ML model	ROC-AUC score					90/95 level	
						Porosity (%)	ESD (μm)
	Train/Test datasets						
	C/C	A/A	B/B	A/B	B/A	C/C	C/C
LR	0.934	0.935	0.960	0.933	0.794	0.107	253.4
GBDT	0.969	0.940	0.969	0.961	0.891	0.106	252.8
MLP	0.958	0.949	0.976	0.939	0.846	0.142	279.0
GPC	0.963	0.945	0.970	0.942	0.911	0.102	250.0

We find that the optimal model uses just 19 features, as selected by recursive feature elimination (RFE). The models for the other two voxel bin size/porosity thresholds use ~50 features. An explanation of their comparative performance is that the higher threshold models are 'grasping' for more obscure features in order to separate out the more porous defective voxel bins from the non-defective voxel bins. This leads to overfitting and consequently poorer model performance in the low-PFA regime. Meanwhile, the model trained to classify at a lower porosity threshold is able to construct a simpler and more robust model of defectiveness.

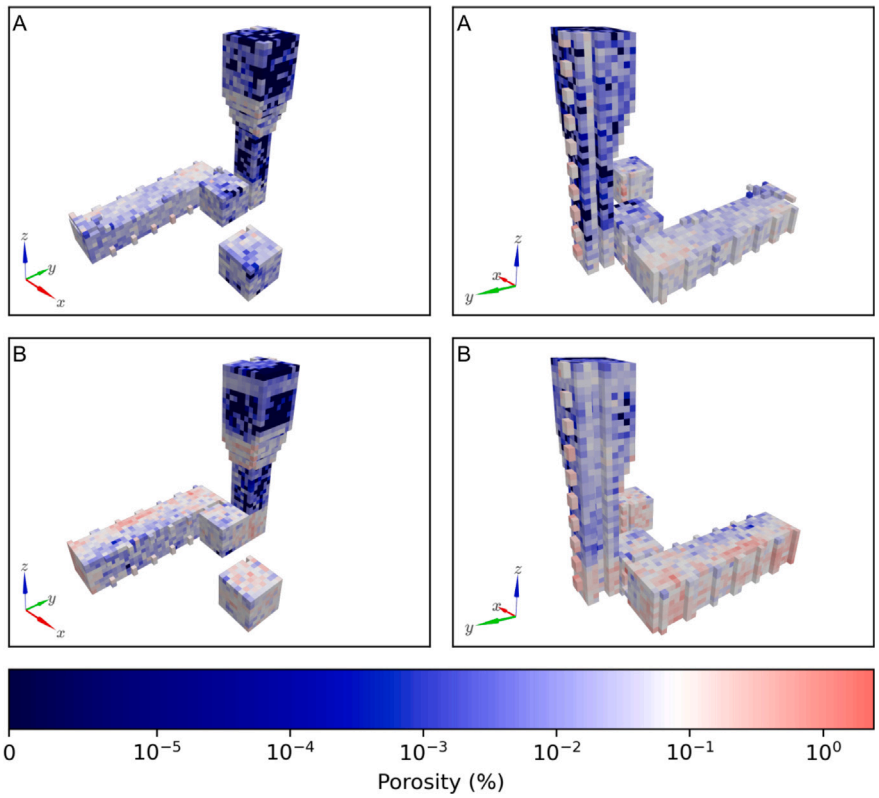


Fig. 10. Volume renders of the 1 mm voxel bin results for the four parts scanned from both builds. Top row: build A; bottom row: build B. Note that the white colour on the scale coincides with the 0.05% porosity level used as a classification threshold in Section 3.2.

This effect is particularly significant near the decision threshold. The combined dataset has a concentration of data around the 0.05% porosity threshold, allowing for higher sensitivity in this crucial region, but also making too-complex models susceptible to overfitting (this is likely responsible for the poor performance of the MLP model, see Section 3.2). This feature of the dataset, in combination with the RFE strategy employed to prevent overfitting on the non-linear models, is responsible for the low PFA and high POD, along with the high 90/95 level, that is attained for the optimal model. This data feature is even more pronounced in the build A dataset, and the results for that model (discussed in Section 3.4) provide further support for the explanation above.

The 19 features selected by RFE show a surprising degree of interpretability. These features and their importances in the optimal model are shown in Fig. 13, whilst Figs. 14 and 15 visualise the spatial distribution of the four most important features for builds A and B respectively. The gain of features in the GBDT model are used for feature importance (see the supplementary material). We have used 'spot', 'spatter', and 'plume' to describe the three regions of the segmented images (image segmentation is shown in Fig. 3(b)); as discussed in Section 2.1.2 the segmented image regions should be understood as merely *proxies* for these actual physical regions. We find the three features that are most important to classifying a voxel bin as high or low porosity are the average spot roundness throughout the vixel bin (a larger roundness value equates to a shape closer to circular), the standard deviation of spot intensities (averaged over the spot region) throughout the voxel bin, and the average count of spatter regions throughout the voxel bin. Meanwhile, the 'photodiode' proxy feature does not appear amongst the 19 RFE-selected features.

Fig. 16 plots the POD curve for the optimal model. This shows the trade-off between the size of the defect detected, in this case the percentage of porosity in the voxel bin, and rate of detection. The ATLAS system achieves 90% POD with 95% confidence (90/95) for approximately 0.11% porosity; in the 2 mm voxel bin, this is equivalent to a 253 µm spherical defect. The POD has shifted leftward considerably in comparison work detailed in de Winton et al. [1]. This represents a marked improvement in detection performance as there is a far higher POD for a given defect size.

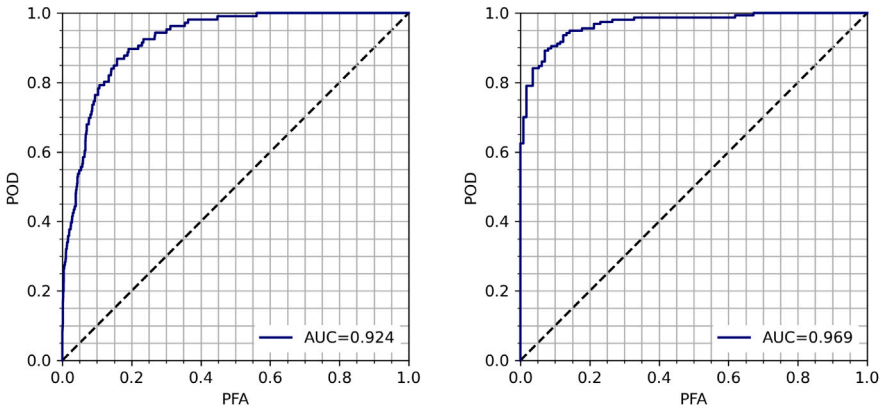
3.4. ML transferability between builds

Parts from each build were found to have different porosity distributions, suggesting there may be an unknown factor influencing porosity, providing an ideal test of model performance when trained on data from a given build and tested on data from a separate build. Resulting ROC curves are shown in Fig. 17. Hyperparameters were further fine-tuned for each such model, including feature selection via RFE.

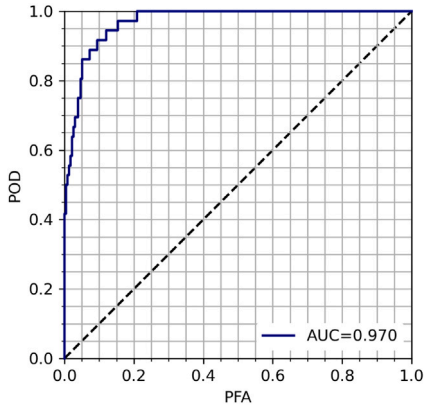
We found that the model trained on the more porous build B did not transfer well to the other build A. In particular, this model only achieved a high POD for a high PFA, suggesting this model is less sensitive to porosity values closer to the threshold for defectiveness. This is unsurprising considering the porosity distribution for each part (Fig. 8), in which build A had a concentration of voxel bins with porosity values near to the 0.05% threshold.

However, for the reverse case of transfer this was not true. As seen in Fig. 17, the model trained on dataset A actually performed better on transfer to build B than it did on its own dataset for the optimal model. This is likely due to the same concentration of voxel bins with porosity values near to the 0.05% threshold. This data distribution allowed build A to better learn information about the decision threshold, in turn leading to better performance on build B in which the data distribution has a longer and fatter tail above the threshold.

The other ML models' transfer performance are summarised in Table 5. The other models show similar transfer performance, but did not outperform training and testing on the same build. The transfer case



(a) 1 mm voxel bins / 0.4% porosity threshold. (b) 2 mm voxel bins / 0.05% porosity threshold.



(c) 2 mm voxel bins / 0.2% porosity threshold.

Fig. 11. ROC curves from the GBDT models trained on the combined builds dataset and calculated on a withheld 20% test dataset. For these voxel bin sizes/porosity thresholds, corresponding equivalent spherical defect diameters are (a) 200 µm, (b) 200 µm, (c) 310 µm.

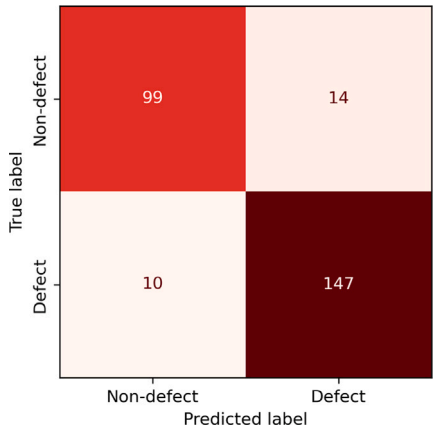


Fig. 12. Confusion matrix for the optimal model.

used the full dataset B, whereas the test case used the 20% reserved as test data, a fifth the number of total data points, so these small differences could result from a statistical artefact.

Fixing the final number of features selected via RFE to 19 (as found for the C-train model) allowed for comparison of feature importance between each of the three models. Since the models performed well on inter-build transfer (when the underlying distribution of porosity was favourable to this task), this suggests that different features can have strong covariances, rather than resulting from variation in the underlying mechanisms of defect formation between datasets A and B. These results highlight the importance of training defect detection models on datasets spanning a number of different builds.

4. Discussion

4.1. Industrial applicability

The United States Federal Aviation Administration (FAA) advises that Non-Destructive Evaluation (NDE) techniques applied to critical parts should meet the 90/95 requirement for the defect size of interest. Our best model attains this 90/95 requirement at a porosity level of 0.11%, which can be approximated to a single 253 µm diameter spherical pore within a 2 mm voxel bin.

Use cases for LPBF are diverse, as are defects sizes of interest, but this detection performance is objectively comparable to other established NDE methods. For example, Sun et al. found a 230 µm 90/95 pore size by µCT in 12 mm diameter Ti6Al4V bars [64]. Rometsch et al. found a minimum detectable pore size of 200 µm by radiography in 2 mm thick Hastelloy samples, rising to 500 µm in 2 mm thick samples [63]. These comparisons suggest that our system could potentially replace traditional NDE methods in a range of practical applications.

In cases where inspection requirements are not fully met, the system could alternatively serve as a pre-screening tool, reducing the number of parts that undergo costly NDE stages or enabling more targeted inspection, thereby shortening scan times and improving overall throughput. This dual role makes the system valuable both in safety-critical contexts, where strict qualification standards apply, and in

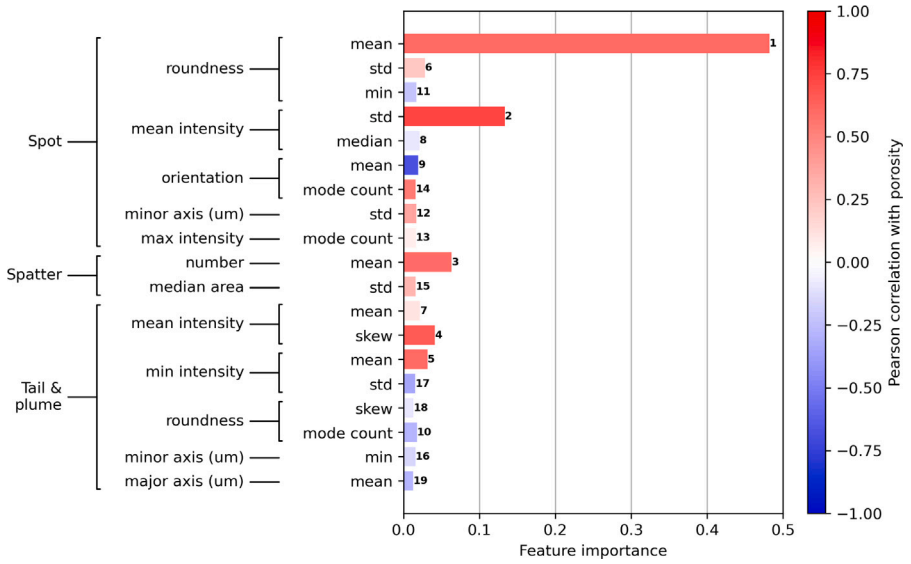


Fig. 13. Feature importance for the optimal machine learning model. The numbers to the rights of the bars indicate their importance rankings. Colours indicate the Pearson correlation coefficient between a feature and porosity % in the raw dataset.

Table 6
Quantitative comparison of performance for closest studies in literature that report 90/95 performance metric.

Generation	90/95 porosity level %	Porosity classification threshold %	Mean sample porosity %	Lengthscale	Source
Laser defocusing	1.7	0.5	<0.1–1.8	part-wide	[1]
Organic porosity	1.9[a]	0.1[a]	<0.1	position in layer	[51]
Organic porosity	0.11	0.05	0.099	2 mm voxel bins	This work

[a] Estimated from the reported 90/95 defect size of 661 μm and a defect classification threshold of 250 μm using our equivalent spherical diameter conversion approach (assuming 2 mm voxel bins).

less demanding applications, where detection of only larger defects is required.

As noted in Section 1.3, recent work has suggested that porosity percentage can be more informative for yield stress and failure strain compared with pore size [53]. However, they also found a ligament parameter (related to pore size and spacing) to be a better predictor of failure strain alone. Further work is required to link in-process monitoring results to mechanical failure prediction.

4.2. Comparison to other works

The closest points of comparison are the studies by de Winton et al. and Snow et al. which also reported the 90/95 requirement [1,51]. De Winton et al. achieved 90/95 at 1.7% porosity using a 0.5% part-wide porosity threshold, while Snow et al. reported 90/95 at a defect size of 661 μm based on a 250 μm flaw threshold. These Snow et al. values can be converted into a porosity values for comparison using our equivalent spherical defect concept, which gives a 90/95 porosity level of 1.9% for a 0.1% porosity classification threshold. By contrast, we achieve 90/95 at 0.11% porosity for classification threshold of 0.05%, corresponding to an equivalent spherical defect of 253 μm for a 200 μm classification threshold in 2 mm voxel bins (Table 6). This represents an order-of-magnitude improvement in porosity detection sensitivity and significantly reduces the size of the largest defect that could be missed.

Some key methodological differences underpin these results:

- Fidelity and data volume: Our datasets were derived from over 1 TB of raw melt pool imaging data, three orders of magnitude larger than the closest comparable work. Snow et al. labelled 18 layers of in-process imaging data (approximately 298 million pixels, or ≈2 GB). By comparison, we used ≈0.5 GB of raw image data *per voxel bin statistic set*, with 56 million melt pool images overall (≈1 trillion pixels). The fidelity of this dataset reveals more information about the process and enables more robust feature extraction and model training.
- Closer proximity to defect formation: Melt pool imaging directly observes the physics of defect initiation, unlike layer-wise imaging methods (e.g. EOS OT), which superimpose temporal events at the same location and therefore dilute causal signals.
- Voxel-based approach: By voxelising both CT and in-process data after aligning, we mitigated alignment errors that otherwise obscure information about smaller pores. This improved sensitivity at smaller defect sizes.
- Naturally induced porosity: Using the build layout and part geometry, rather than deliberately altering laser parameters, generated organic porosity across the target classification threshold, ensuring there was enough training data for the ML models to learn realistic defect signatures rather than artificially induced anomalies.

These results define a new state-of-the-art for in-process porosity detection: naturally occurring defects are detected at significantly lower porosity levels, localised, and quantified with an industrially interpretable metric (90/95), rather than generic classification scores. Future work should now extend to more complex geometries, systematic induction of other failure modes, materials and different machines to further test the limits of detection capability.

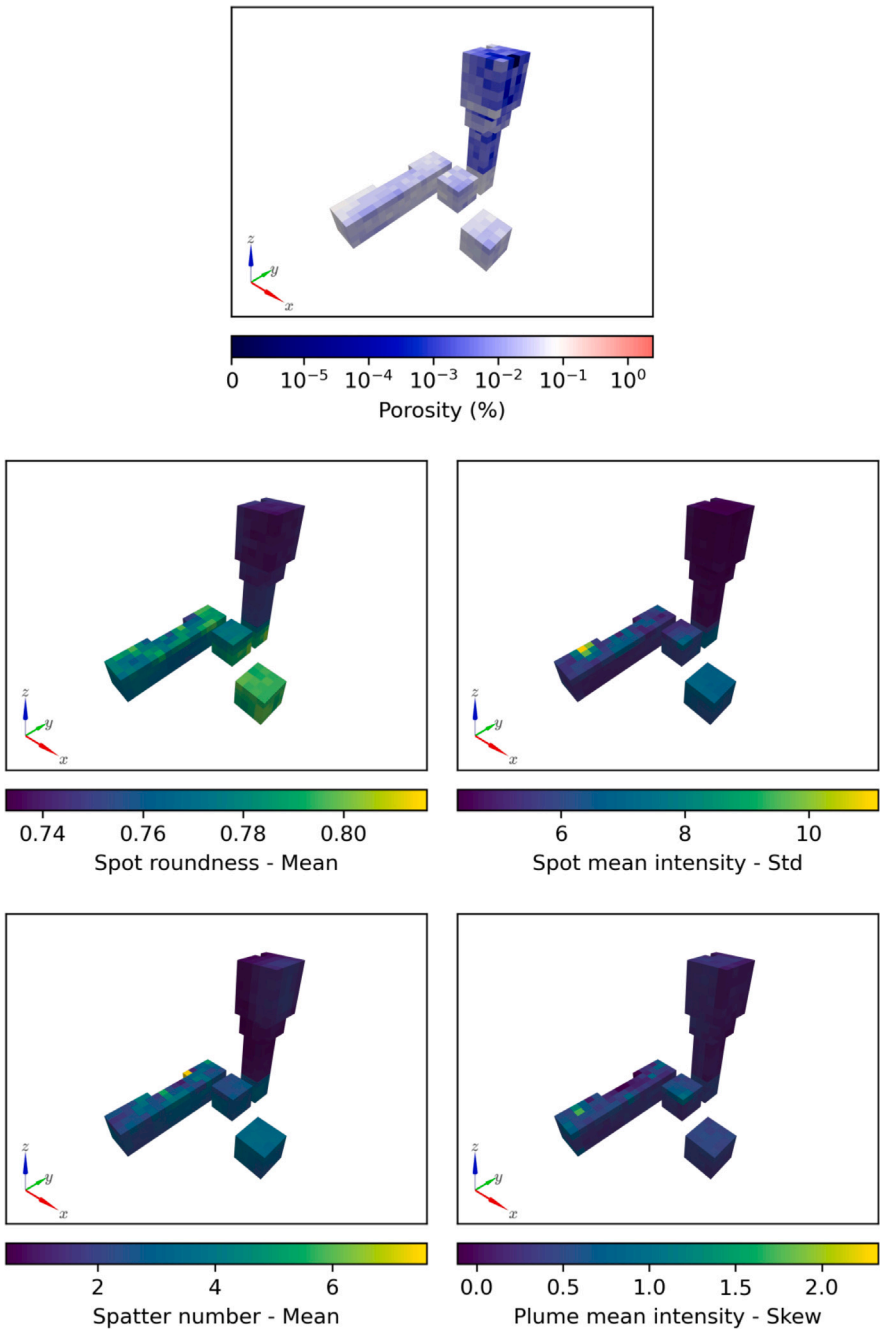


Fig. 14. Volume renders of the 2 mm voxel bins for build A, coloured by voxel bin porosity values (top) and values of the four most important features as selected by the optimal model. Note that colour scales align with Fig. 15.

4.3. Localisation and voxel binning length scale

Doubling the voxel bin length scale from 1 mm to 2 mm improved detection performance, as shown in Fig. 11. Averaging over a larger region of both ground-truth and sensor data increases homogeneity and reduces the impact of anomalous or outlying voxel bins. At the 2 mm scale, the system effectively performs melting regime classification, since porosity distribution and melt behaviour are assumed uniform at this length scale. Feature averages across the voxel bin therefore become reliable indicators of whether melting in that region is defective or not.

At smaller voxel bin sizes, this assumption of homogeneity breaks down. When the voxel bin scale approaches the pore size itself (below ≈0.5 mm), it becomes increasingly likely that some voxel bins are fully porous, producing extreme values in both porosity and monitoring signals. This explains the drop in detection performance at finer scales and highlights that the current approach is not designed for individual pore detection.

The voxel bin size also sets the localisation resolution of predicted defects. A 2 mm voxel bin is relatively coarse compared to other efforts that aim to resolve individual pores within a part [21,25,41,42,51]. While this limits the precision of localisation, the larger voxel bins deliver improved detection performance. Predicting porosity fractions robustly at this resolution could provide significant value in production environments, even without pinpointing each pore. One could also take a conservative approach and assume the equivalent spherical defect is located at the most high-stressed point in a voxel bin.

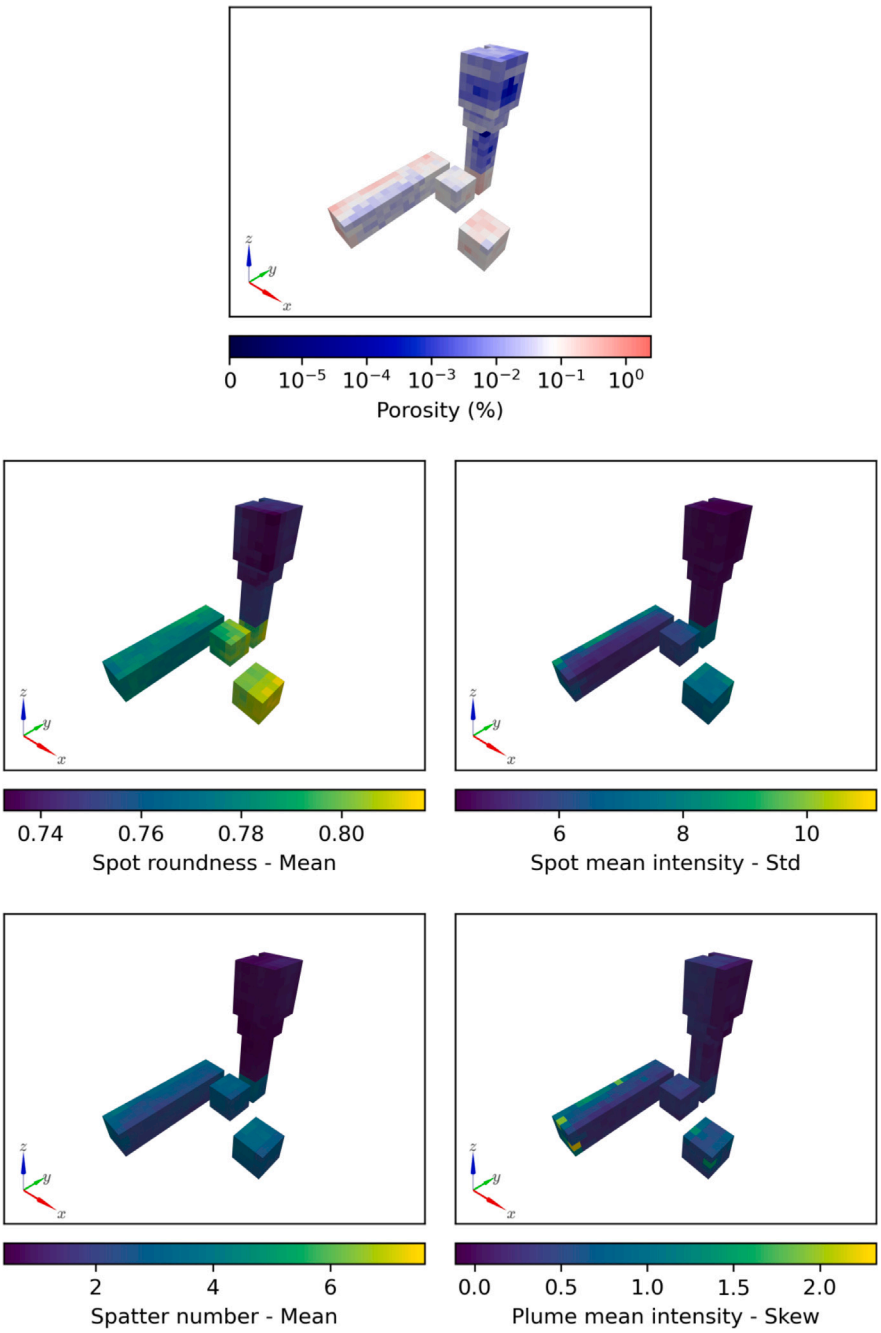


Fig. 15. Volume renders of the 2 mm voxel bins for build B, coloured by voxel bin porosity values (top) and values of the four most important features as selected by the optimal model. Note that colour scales align with Fig. 14.

A potential limitation of the voxel binning method lies in regions where different processing parameters are used within a single voxel bin, such as where contour or boarder scan paths are used at the edge of parts. These scan strategies produce melt pool feature values that differ significantly from those of bulk hatching, and averaging them within the same voxel bin may not yield meaningful predictions. This is a vital avenue for future work given the importance of these locations to fatigue and fracture behaviour, with possible solutions outlined in the future work section.

4.4. Features and machine learning

As shown in Fig. 13, features serving as proxies for melt pool shape, variations in melt pool emissions, and amount of spatter are the most important for defect prediction. This aligns well with previous studies that have studied the use of melt pool shape [12], melt pool intensity [23], and spatter [46,73] as predictors for defect formation or related physical phenomena.

RFE identified roundness of the bright spot region as the most significant feature for porosity classification. It is important to clarify that this measure does not directly quantify the roundness of the melt pool itself, but rather the roundness of the most intense region in the coaxial image, located at or near the laser–melt pool interaction zone. Direct physical interpretation is therefore non-trivial.

One plausible explanation is that roundness provides a proxy for delivered energy density and its influence on melt pool dynamics. At low energy input (e.g., insufficient overlap or shallow melting), the bright region reflects the shape of the incident laser beam and is therefore

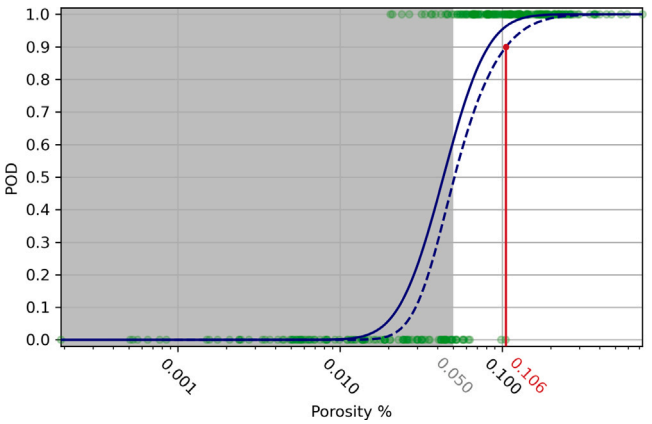


Fig. 16. POD curve plotted against porosity % for the optimal model. The grey region shows the 0.05% porosity threshold used for classification. The dashed lined gives the 95% confidence interval for the POD curve. The green data points are the signals from the GBDT model, and the red line gives the 90/95 porosity level.

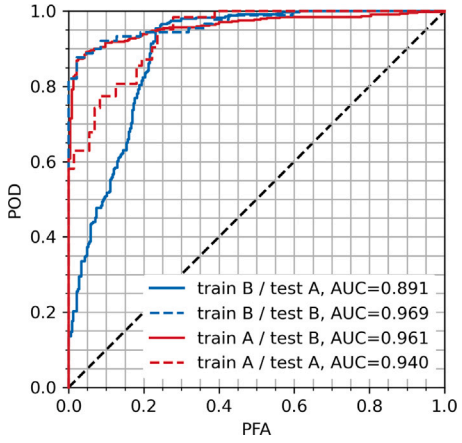


Fig. 17. ROC curves for GBDT models trained and tested on different samples of the dataset for each build. The 2 mm voxel bin dataset with 0.05% porosity threshold is used.

highly round. With increasing energy input, melting and vaporisation distort the surface and vapour plume, leading to less circular intensity distributions. At even higher input energies, keyhole formation and associated vapour fluctuations further reduce the roundness of the bright region, even if the keyhole geometry itself is circular.

Spatter may also play a role. Large spherical particles that intercept the beam can produce localised bright spots with high roundness, introducing further correlation between roundness and porosity formation. Taken together, we hypothesise that high roundness values act as an indicator of insufficient energy delivery, conditions that are strongly associated with lack-of-fusion porosity.

Mean spot intensity ranked lower than expected, while the standard deviation of spot intensity within a voxel bin was the second most important feature. This may be due to fundamental limits in surface temperature increases once vaporisation commences: the melt pool surface reaches the boiling point of the material, capping the mean intensity under normal melting conditions. Variations in intensity, however, may arise from anomalous melt pool behaviour, explaining the strong association between increased standard deviation and higher porosity levels.

Features analogous to photodiode signals did not appear among the RFE-selected features of the best-performing model. Likewise, with the exception of the spot minor axis length, no features directly representing melt pool size were retained, despite extensive prior work highlighting size as a predictor of defects. One likely explanation is that shape- and spatter-based features are inherently more robust to variability in image capture and post-processing. Their scale independence may preserve more physical information than absolute size measurements, which are sensitive to imaging artefacts. Conversely, mean intensity features serve as proxies for effects that would act to obscure more geometric features, such as the vapour cloud. This may indirectly encode information about unstable melting regimes.

These results show that machine learning models can be highly interpretable. Feature relevance analysis does not simply echo prior assumptions, but highlights which process signatures carry the most consistent predictive power. This, in turn, points to promising directions for future research into the underlying physics of melt pool dynamics.

4.5. Dataset standardisation and route to industrial adoption

This work demonstrates significant improvements in defect detection performance and shows that models can transfer between builds on the same machine. Such reproducibility is encouraging for de-risking deployment in serial production scenarios, where a given part may be manufactured thousands of times on the same machine under controlled conditions. However, broader industrial adoption is significantly more challenging. The diversity of machine types, materials, geometries, build layouts, parameter sets, and inspection requirements across industry is vast. A fully generalisable model, applicable to any machine, material or part, is an appealing goal, but in practice models tailored to specific serial production scenarios may deliver higher reliability and more immediate industrial value.

Nevertheless, improved access to large datasets containing realistic, organically generated defects across multiple failure modes will be critical. The naturally induced porosity approach and the defect detection performance analysis outlined here provide a pathway to generating standardised datasets with industrially relevant defect populations. When combined with well-aligned ground truth data (ideally captured and processed in automated, repeatable ways) such datasets can underpin robust model training and build confidence in deployment across a wide range of use cases.

Achieving this is likely beyond the capacity of any single laboratory. Instead, it calls for coordinated community efforts. Round-robin style testing programmes, where standardised builds are independently reproduced and shared across institutions, represent one promising approach. Establishing such shared, high-quality datasets will be essential for translating in-process defect detection into reliable, certifiable tools for industrial adoption.

5. Conclusions

In this work, we developed an in-process monitoring system for LPBF with the capability to indefinitely record high-speed coaxial melt pool images at 20 kHz. Over 56 million melt pool images were analysed over repeated builds. A voxel-based binning framework was introduced, enabling statistics of melt pool, spatter, and plume features to be quantified for voxelised regions within the part. These image-derived features were then used to classify porosity against a 0.05% threshold, enabling construction of a digital 3D defect map of the part during fabrication.

The following conclusions can be drawn:

1. The system achieved the NDE-standard 90/95 requirement at a porosity level of 0.11% in 2 mm voxel bins, indicating potential for industrial assurance at this level.

2. The concept of an equivalent spherical defect was introduced to convert between voxel bin porosity and single defect diameter, aiding comparisons between different approaches. This was 253 μm for this system. Analysis of μCT data showed this to provide a conservative estimate of true defect size.
3. Increasing the voxel bin length scales improved detection performance by averaging out noise and anomalous signals.
4. Transferability was demonstrated between repeat builds on the same machine, but the most robust results were achieved by including training data from multiple builds.
5. Features related to melt pool shape, spatter, and plume behaviour correlated strongly with porosity. By contrast, features related to melt pool size were not selected by recursive feature elimination, suggesting that the former either capture more relevant defect information or are more robust in coaxial imaging.

The system achieves a step-change in the performance of in-process monitoring systems for LPBF, enabling localised porosity with industrially relevant defect sizes, outperforming the closest comparable work by a factor of 10. By combining high-fidelity melt pool imaging with robust voxel-based ML analysis, this work provides a new methodology for real-time defect detection, with direct applicability to industrial quality assurance and part qualification.

5.1. Future research

Future research should focus on improving detection performance at industrially meaningful thresholds, particularly the 90/95 metric. Refinement of the naturally induced porosity strategy will allow datasets to be engineered with distributions of defects concentrated near decision thresholds, enhancing ML model training.

Predictive capability near the part surface, in thin-wall and fine-feature regions are an important avenue for future work. Potential strategies include incorporating processing parameters directly as input features (so the model explicitly accounts for contour or boarder scans) and refining voxel bin size to keep parameter mixing in a voxel bin to a minimum. This may also require more sophisticated ML architectures capable of capturing correlations across multiple length scales, and will need to be supported by larger and more diverse training datasets. The extension of naturally induced porosity to a broader range of defect generation mechanisms could enable the creation of standardised, high-fidelity datasets for in-process monitoring research.

Finally, standardising builds, round-robin testing, and automation of data registration and alignment pipelines are critical to improving reproducibility and scalability. This will support the development of large high-quality databases, underpinning future industrial adoption and providing a robust foundation for the verification and certification of LPBF components by in-process monitoring.

CRediT authorship contribution statement

Patrick L. Taylor: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Formal analysis. **Richard J. Williams:** Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation. **Henry C. de Winton:** Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation. **Vincent Fernandez:** Writing – original draft, Data curation. **Sebastian Larsen:** Writing – review & editing. **Paul A. Hooper:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Paul A Hooper reports financial support was provided by Engineering and Physical Sciences Research Council. Paul A Hooper reports financial support was provided by AWE plc. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors gratefully acknowledge the financial support from EPSRC for funding Patrick Taylor and Sebastian Larsen (EP/X024288/1) and Harry de Winton (EP/R513052/1). We also acknowledge the financial support and technical input from AWE plc for supporting Richard Williams and Sebastian Larsen (30458871 and 40041621).

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.addma.2025.104982.

Data availability

Data will be made available on request.

References

[1] H.C. de Winton, F. Cegla, P.A. Hooper, A method for objectively evaluating the defect detection performance of in-situ monitoring systems, Addit. Manuf. 48 (2021) 102431, http://dx.doi.org/10.1016/J.ADDMA.2021.102431.
[2] W.H. Kan, L.N.S. Chiu, C.V.S. Lim, Y. Zhu, Y. Tian, D. Jiang, A. Huang, A critical review on the effects of process-induced porosity on the mechanical properties of alloys fabricated by laser powder bed fusion, J. Mater. Sci. 57 (2022) 9818—9865, http://dx.doi.org/10.1007/s10853-022-06990-7.
[3] R.J. Williams, A. Piglione, T. Rønneberg, C. Jones, M.-S. Pham, C.M. Davies, P.A. Hooper, In situ thermography for laser powder bed fusion: Effects of layer temperature on porosity, microstructure and mechanical properties, Addit. Manuf. 30 (2019) 100880, http://dx.doi.org/10.1016/j.addma.2019.100880.
[4] T. Purtonen, A. Kalliosaari, A. Salminen, Monitoring and adaptive control of laser processes, Phys. Procedia 56 (2014) 1218–1231.
[5] M. Grasso, V. Laguzza, Q. Semeraro, B.M. Colosimo, In-process monitoring of selective laser melting: spatial detection of defects via image data analysis, J. Manuf. Sci. Eng. 139 (5) (2017).
[6] R. McCann, M.A. Obeidi, C. Hughes, E. McCarthy, D.S. Egan, R.K. Vijayaraghavan, A.M. Joshi, V.A. Garzon, D.P. Dowling, P.J. McNally, In-situ sensing, process monitoring and machine control in laser powder bed fusion: A review, Addit. Manuf. (2021) 102058.
[7] B. Wu, X.-y. Ji, J.-x. Zhou, H.-q. Yang, D.-j. Peng, Z.-m. Wang, Y.-j. Wu, Y.-j. Yin, In situ monitoring methods for selective laser melting additive manufacturing process based on images—A review, China Foundry 18 (4) (2021) 265–285.
[8] C. Chua, Y. Liu, R.J. Williams, C.K. Chua, S.L. Sing, In-process and post-process strategies for part quality assessment in metal powder bed fusion: A review, J. Manuf. Syst. 73 (2024) 75–105, http://dx.doi.org/10.1016/J.JMSY.2024.01.004.
[9] T. Shen, B. Li, Digital twins in additive manufacturing: a state-of-the-art review, Int. J. Adv. Manuf. Technol. 131 (1) (2024) 63–92, http://dx.doi.org/10.1007/S00170-024-13092-Y/METRICS.
[10] L. Scime, J. Beuth, Using machine learning to identify in-situ melt pool signatures indicative of flaw formation in a laser powder bed fusion additive manufacturing process, Addit. Manuf. 25 (2019) 151–165, http://dx.doi.org/10.1016/j.addma.2018.11.010.
[11] S. Larsen, P.A. Hooper, Deep semi-supervised learning of dynamics for anomaly detection in laser powder bed fusion, J. Intell. Manuf. 33 (2) (2022) 457–471, http://dx.doi.org/10.1007/S10845-021-01842-8/FIGURES/10.
[12] L.R. Goossens, B. Van Hooreweder, A virtual sensing approach for monitoring melt-pool dimensions using high speed coaxial imaging during laser powder bed fusion of metals, Addit. Manuf. (2021) 101923.
[13] M. Grasso, B.M. Colosimo, Process defects and in situ monitoring methods in metal powder bed fusion: A review, Meas. Sci. Technol. 28 (4) (2017) http://dx.doi.org/10.1088/1361-6501/aa5c4f.

[14] J. Kruth, P. Mercelis, J.V. Vaerenbergh, T. Craeghs, Feedback control of selective laser melting, in: Proceedings of the 3rd International Conference on Advanced Research in Virtual and Rapid Prototyping, 2007, pp. 1–7.
[15] S. Berumen, F. Bechmann, S. Lindner, J.P. Kruth, T. Craeghs, Quality control of laser- and powder bed-based additive manufacturing (AM) technologies, in: Physics Procedia, vol. 5, (PART 2) Elsevier B.V., 2010, pp. 617–622, http://dx.doi.org/10.1016/j.phpro.2010.08.089.
[16] T. Toeppel, P. Schumann, M.-C. Ebert, T. Bokkes, K. Funke, M. Werner, F. Zeulner, F. Bechmann, F. Herzog, 3D analysis in laser beam melting based on real-time process monitoring, in: Mater Sci Technol Conf, 2016, pp. 123–132.
[17] S. Bagg, Z. Jones, J. Rowe, Investigation into the use of the Concept Laser QM system as an in-situ research and evaluation tool, Tech. Rep., NASA, 2014.
[18] M.H. Kunkel, A. Gebhardt, K. Mpofu, S. Kallweit, Quality assurance in metal powder bed fusion via deep-learning-based image classification, Rapid Prototyp. J. (2019).
[19] E. Boos, M. Schwarzenberger, M. Jaretzki, H. Wiemer, S. Ihlenfeldt, Melt pool monitoring using fuzzy based anomaly detection in laser beam melting, in: Metal Additive Manufacturing Conference, 2019.
[20] T. Kolb, R. Elahi, J. Seeger, M. Soris, C. Scheitler, O. Hentschel, J. Tremel, M. Schmidt, Camera signal dependencies within coaxial melt pool monitoring in laser powder bed fusion, Rapid Prototyp. J. (2020).
[21] J.A. Mitchell, T.A. Ivanoff, D. Dagel, J.D. Madison, B. Jared, Linking pyrometry to porosity in additively manufactured metals, Addit. Manuf. 31 (2020) 100946, http://dx.doi.org/10.1016/j.addma.2019.100946.
[22] H. Yeung, Z. Yang, L. Yan, A meltpool prediction based scan strategy for powder bed fusion additive manufacturing, Addit. Manuf. 35 (2020) 101383.
[23] F.H. Kim, H. Yeung, E.J. Garboczi, Characterizing the effects of laser control in laser powder bed fusion on near-surface pore formation via combined analysis of in-situ melt pool monitoring and X-ray computed tomography, Addit. Manuf. (2021) 102372.
[24] O. Kwon, H.G. Kim, M.J. Ham, W. Kim, G.-H. Kim, J.-H. Cho, N.I. Kim, K. Kim, A deep neural network for classification of melt-pool images in metal additive manufacturing, J. Intell. Manuf. 31 (2) (2018) 375–386.
[25] K. Taherkhani, E. Sheydaeian, C. Eischer, M. Otto, E. Toyserkani, Development of a defect-detection platform using photodiode signals collected from the melt pool of laser powder-bed fusion, Addit. Manuf. 46 (2021) 102152.
[26] D.S. Egan, C.M. Ryan, A.C. Parnell, D.P. Dowling, Using in-situ process monitoring data to identify defective layers in Ti-6Al-4V additively manufactured porous biomaterials, J. Manuf. Process. 64 (2021) 1248–1254.
[27] Z. Mao, W. Feng, H. Ma, Y. Yang, J. Zhou, S. Liu, Y. Liu, P. Hu, K. Zhao, H. Xie, G. Guo, Z. Liu, Continuous online flaws detection with photodiode signal and melt pool temperature based on deep learning in laser powder bed fusion, Opt. Laser Technol. 158 (2023) 108877, http://dx.doi.org/10.1016/J.OPTLASTEC.2022.108877.
[28] S. Feng, Z. Chen, B. Bircher, Z. Ji, L. Nyborg, S. Bigot, Predicting laser powder bed fusion defects through in-process monitoring data and machine learning, Mater. Des. 222 (2022) 111115, http://dx.doi.org/10.1016/J.MATDES.2022.111115.
[29] F. Imani, A. Gaikwad, M. Montazeri, P. Rao, H. Yang, E. Reutzel, Layerwise in-process quality monitoring in laser powder bed fusion, in: ASME 2018 13th International Manufacturing Science and Engineering Conference, American Society of Mechanical Engineers Digital Collection, 2018.
[30] M. Montazeri, A.R. Nassar, A.J. Dunbar, P. Rao, In-process monitoring of porosity in additive manufacturing using optical emission spectroscopy, IISE Trans. (2019) 1–16, http://dx.doi.org/10.1080/24725854.2019.1659525.
[31] J.-B. Forien, N.P. Calta, P.J. DePond, G.M. Guss, T.T. Roehling, M.J. Matthews, Detecting keyhole pore defects and monitoring process signatures during laser powder bed fusion: A correlation between in situ pyrometry and ex situ X-ray radiography, Addit. Manuf. 35 (2020) 101336.
[32] N.H. Paulson, B. Gould, S.J. Wolff, M. Stan, A. Greco, Correlations between thermal history and keyhole porosity in laser powder bed fusion, Addit. Manuf. (2020) 101213.
[33] G. Mohr, S.J. Altenburg, A. Ulbricht, P. Heinrich, D. Baum, C. Maierhofer, K. Hilgenberg, In-situ defect detection in laser powder bed fusion by using thermography and optical tomography—Comparison to computed tomography, Metals 10 (1) (2020) http://dx.doi.org/10.3390/met10010103.
[34] M. Grasso, A.G. Demir, B. Previtali, B.M. Colosimo, In situ monitoring of selective laser melting of zinc powder via infrared imaging of the process plume, Robot. Comput.-Integr. Manuf. 49 (2018) 229–239.
[35] B.D. Bevans, A. Carrington, A. Riensche, A. Tenequer, C. Barrett, H.S. Halliday, R. Srinivasan, K.D. Cole, P. Rao, Digital twins for rapid in-situ qualification of part quality in laser powder bed fusion additive manufacturing, Addit. Manuf. 93 (2024) 104415, http://dx.doi.org/10.1016/J.ADDMA.2024.104415.
[36] J. Li, L. Cao, J. Xu, S. Wang, Q. Zhou, In situ porosity intelligent classification of selective laser melting based on coaxial monitoring and image processing, Measurement 187 (2022) 110232, http://dx.doi.org/10.1016/J.MEASUREMENT.2021.110232.
[37] S.A. Shevchik, C. Kenel, C. Leinenbach, K. Wasmer, Acoustic emission for in situ quality monitoring in additive manufacturing using spectral convolutional neural networks, Addit. Manuf. 21 (2018) 598–604.
[38] S.A. Shevchik, G. Masinelli, C. Kenel, C. Leinenbach, K. Wasmer, Deep learning for in situ and real-time quality monitoring in additive manufacturing using acoustic emission, IEEE Trans. Ind. Inform. 15 (9) (2019) 5194–5203, http://dx.doi.org/10.1109/TII.2019.2910524.
[39] M. Grasso, B.M. Colosimo, A statistical learning method for image-based monitoring of the plume signature in laser powder bed fusion, Robot. Comput.-Integr. Manuf. 57 (2019) 103–115, http://dx.doi.org/10.1016/j.rcim.2018.11.007.
[40] H. Merschroth, J. Harbig, M. Weigold, Defect detection based on sensor data fusion of optical monitoring systems in laser based powder bed fusion, in: Laser Applications Conference, Optical Society of America, 2020, pp. LTh1B–3.
[41] J. Petrich, C. Gobert, S. Phoha, A.R. Nassar, E.W. Reutzel, Machine learning for defect detection for PBFAM using high resolution layerwise imaging coupled with post-build CT scans, in: Proceedings of the 27th International Solid Freeform Fabrication Symposium, 2017.
[42] C. Gobert, E.W. Reutzel, J. Petrich, A.R. Nassar, S. Phoha, Application of supervised machine learning for defect detection during metallic powder bed fusion additive manufacturing using high resolution imaging, Addit. Manuf. 21 (2018) 517–528.
[43] M. Khanzadeh, S. Chowdhury, M.A. Tschopp, H.R. Doude, M. Marufuzzaman, L. Bian, In-situ monitoring of melt pool images for porosity prediction in directed energy deposition processes, IISE Trans. 51 (5) (2019) 437–455.
[44] S. Keaveney, A. Shmeliov, V. Nicolosi, D.P. Dowling, Investigation of process by-products during the selective laser melting of Ti6AL4V powder, Addit. Manuf. 36 (2020) 101514.
[45] P. Yadav, O. Rigo, C. Arvieu, E. Le Guen, E. Lacoste, In situ monitoring systems of the SLM process: On the need to develop machine learning models for data processing, Crystals 10 (6) (2020) 524.
[46] Z. Snow, L. Scime, A. Ziabari, B. Fisher, V. Paquit, Observation of spatter-induced stochastic lack-of-fusion in laser powder bed fusion using in situ process monitoring, Addit. Manuf. 61 (2023) 103298, http://dx.doi.org/10.1016/J.ADDMA.2022.103298.
[47] F.M. Carter, D. Kozjek, C. Porter, S.J. Clark, K. Fezzaa, M. Fujishima, N. Irino, J. Cao, Melt pool instability detection using coaxial photodiode system validated by in-situ X-ray imaging, CIRP Ann 72 (1) (2023) 205–208, http://dx.doi.org/10.1016/J.CIRP.2023.03.031.
[48] F.M. Carter, C. Porter, D. Kozjek, K. Shimoyoshi, M. Fujishima, N. Irino, J. Cao, Machine learning guided adaptive laser power control in selective laser melting for pore reduction, CIRP Ann 73 (1) (2024) 149–152, http://dx.doi.org/10.1016/J.CIRP.2024.04.043.
[49] S. Yoder, P. Nandwana, V. Paquit, M. Kirka, A. Scopel, R.R. Dehoff, S.S. Babu, Approach to qualification using E-PBF in-situ process monitoring in Ti-6Al-4V, Addit. Manuf. 28 (2019) 98–106, http://dx.doi.org/10.1016/J.ADDMA.2019.03.021.
[50] T. Ronneberg, C.M. Davies, P.A. Hooper, Revealing relationships between porosity, microstructure and mechanical properties of laser powder bed fusion 316L stainless steel through heat treatment, Mater. Des. 189 (2020) 108481.
[51] Z. Snow, L. Scime, A. Ziabari, B. Fisher, V. Paquit, Scalable in situ non-destructive evaluation of additively manufactured components using process monitoring, sensor fusion, and machine learning, Addit. Manuf. 78 (2023) 103817, http://dx.doi.org/10.1016/J.ADDMA.2023.103817.
[52] B.M. Sharratt, Non-destructive Techniques and Technologies for Qualification of Additive Manufactured Parts and Processes, Technical Report No. DRDC-RDDC-2015-C035, Sharratt Research and Consulting Inc., Victoria, BC, 2015.
[53] K. Joshi, P. Promoppatum, S.S. Quek, S. Raghavan, N.S. Johan, S. Shukla, S. Samudrala, S. van der Veen, M.H. Jhon, Effect of porosity distribution on the strength and strain-to-failure of laser-powder bed fusion printed Ti–6Al–4V, Addit. Manuf. 75 (2023) 103738, http://dx.doi.org/10.1016/J.ADDMA.2023.103738.
[54] Z. Ren, L. Gao, S.J. Clark, K. Fezzaa, P. Shevchenko, A. Choi, W. Everhart, A.D. Rollett, L. Chen, T. Sun, Machine learning-aided real-time detection of keyhole pore generation in laser powder bed fusion, Science 379 (6627) (2023) 89–94, http://dx.doi.org/10.1126/SCIENCE.ADD4667.
[55] NASA, Standard for additively manufactured spaceflight hardware by laser powder bed fusion in metals, 2017.
[56] T. Herzog, M. Brandt, A. Trinchi, A. Sola, A. Molotnikov, Process monitoring and machine learning for defect detection in laser-based metal additive manufacturing, J. Intell. Manuf. 35 (4) (2023) 1407–1437, http://dx.doi.org/10.1007/S10845-023-02119-Y.
[57] Q. Wu, F. Yang, C. Lv, C. Liu, W. Tang, J. Yang, In-situ quality intelligent classification of additively manufactured parts using a multi-sensor fusion based melt pool monitoring system, Addit. Manuf. Front. 3 (3) (2024) 200153, http://dx.doi.org/10.1016/J.AMF.2024.200153.
[58] H. Liu, C. Gobert, K. Ferguson, B. Abranovic, H. Chen, J.L. Beuth, A.D. Rollett, L.B. Kara, Inference of highly time-resolved melt pool visual characteristics and spatially-dependent lack-of-fusion defects in laser powder bed fusion using acoustic and thermal emission data, Addit. Manuf. 83 (2024) 104057, http://dx.doi.org/10.1016/J.ADDMA.2024.104057.

[59] J. Kim, Z. Yang, H. Ko, H. Cho, Y. Lu, Deep learning-based data registration of melt-pool-monitoring images for laser powder bed fusion additive manufacturing, J. Manuf. Syst. 68 (2023) 117–129, http://dx.doi.org/10.1016/J.JMSY.2023.03.006.

[60] J. Mochache, R.M. Taylor, A review of fatigue and damage tolerance life prediction methodologies toward certification of additively manufactured metallic principal structural elements, in: AIAA Scitech 2021 Forum, 2021, p. 1509.

[61] P.A. Hooper, Melt pool temperature and cooling rates in laser powder bed fusion, Addit. Manuf. 22 (May) (2018) 548–559, http://dx.doi.org/10.1016/j.addma.2018.05.032.

[62] Renishaw, Design for metal AM - a beginner's guide, in: Renishaw Newsletter, vol. 44, Renishaw, 2017, pp. 1–16, URL https://www.renishaw.com/en/design-for-metal-am-a-beginners-guide–42652.

[63] P.A. Rometsch, D. Pelliccia, D. Tomus, X. Wu, Evaluation of polychromatic X-ray radiography defect detection limits in a sample fabricated from hastelloy X by selective laser melting, NDT E Int. 62 (2014) 184–192.

[64] S. Sun, J.D. Miller, Y.R. Choi, S. Turk, D. Donnelly, Q. Liu, Detectability by X-ray micro computed tomography of fatigue initiating void defects in laser powder-bed additively manufactured Ti-6Al-4V coupons, Addit. Manuf. 81 (2024) 103986, http://dx.doi.org/10.1016/J.ADDMA.2024.103986.

[65] Standard Practice for Radiographic Examination, ASTM, 2018, http://dx.doi.org/10.1520/E1742{_}E1742M-18.

[66] Standard Test Method for Measurement of Computed Tomography (CT) System Performance, ASTM, 2020.

[67] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: Machine learning in python, J. Mach. Learn. Res. 12 (2011) 2825–2830, URL http://scikit-learn.org.

[68] J.R. Gardner, G. Pleiss, D. Bindel, K.Q. Weinberger, A.G. Wilson, GPyTorch: Blackbox matrix-matrix Gaussian process inference with GPU acceleration, Adv. Neural Inf. Process. Syst. 2018-December (2018) 7576–7586, http://dx.doi.org/10.48550/arxiv.1809.11165.

[69] D. Milios, R. Camoriano, P. Michiardi, L. Rosasco, M. Filippone, Dirichlet-based Gaussian processes for large-scale calibrated classification, Adv. Neural Inf. Process. Syst. 2018-December (2018) 6005–6015, http://dx.doi.org/10.48550/arXiv.1805.10915.

[70] T. Chen, C. Guestrin, XGBoost: A scalable tree boosting system, in: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 13-17-August-2016, Association for Computing Machinery, 2016, pp. 785–794, http://dx.doi.org/10.1145/2939672.2939785.

[71] C.A. Harding, G.R. Hugo, Statistical analysis of probability of detection hit/miss data for small data sets, AIP Conf. Proc. 657 (1) (2003) 1838–1845, http://dx.doi.org/10.1063/1.1570352.

[72] S. Van Der Walt, J.L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J.D. Warner, N. Yager, E. Gouillart, T. Yu, Scikit-image: Image processing in python, PeerJ 2014 (1) (2014) e453, http://dx.doi.org/10.7717/PEERJ.453/FIG-5.

[73] G. Repossini, V. Laguzza, M. Grasso, B.M. Colosimo, On the use of spatter signature for in-situ monitoring of laser powder bed fusion, Addit. Manuf. 16 (2017) 35–48.