

Research Paper

Multi-modal sensor fusion with machine learning for data-driven process monitoring for additive manufacturing

Jan Petrich ^{*}, Zack Snow, David Corbin, Edward W. Reutzel

Applied Research Laboratory, Pennsylvania State University, University Park, PA 16804, USA



A B S T R A C T

This paper presents a complete concept and validation scheme for potential inter-layer flaw detection from in-situ process monitoring for powder bed fusion additive manufacturing (PBFAM) using supervised machine learning. Specifically, the presented work establishes a meaningful statistical correlation between (i) the multi-modal sensor footprint acquired during the build process, and (ii) the existence of flaws as indicated by post-build X-ray Computed Tomography (CT) scans. Multiple sensor modalities, such as layerwise imagery (both pre and post laser scan), acoustic and multi-spectral emissions, and information derived from the scan vector trajectories, contribute to the process footprint. Data registration techniques to properly merge spatial and temporal information are presented in detail. As a proof-of-concept, a neural network is used to fuse all available modalities, and discriminate flaws from nominal build conditions using only in-situ data. Experimental validation was carried out using a PBFAM sensor testbed available at PSU/ARL. Using four-fold cross-validation on a voxel-by-voxel basis, the null hypothesis, i.e. absence of a defect, was rejected at a rate corresponding to 98.5% accuracy for binary classification. Additionally, a sensitivity study was conducted to assess the information content contributed by the individual sensor modalities. Information content was assessed by evaluating the resulting correlation as classification performance when using only a single modality or a subset of modalities. Although optical imagery contains the highest amount of information for flaw detection, additional information content observed in other modalities significantly improved classification performance.

1. Introduction

While the geometric design freedoms enabled by additive manufacturing (AM) promise to expand available design spaces and facilitate component optimization, process repeatability and quality control remains one of the largest inhibitors to widespread industry adoption of the technology [1]. Dynamic mechanical properties, such as fatigue life and fatigue crack growth rates, suffer from variability not seen in their wrought counterparts [2]. Several works have attributed this variability to the presence of processing flaws, e.g., gas porosity, keyhole pores, and lack-of-fusion [3–6]. To address this, various in situ monitoring methodologies have been explored in order to identify AM flaws in real time. Such a process monitoring system could contribute to qualification and certification, and could further enable feedback control systems to automatically correct detected flaws in real time, thereby improving AM component quality and decreasing variability in mechanical performance.

Several common powder bed fusion additive manufacturing (PBFAM) flaws relevant to mechanical performance are of particular note. For example, keyhole pores are vapor-filled depressions commonly associated with high power and low scan speed conditions that lead to porosities during AM. Keyhole formation for PBFAM processes was

studied in [7], where keyhole instabilities and the associated pressure gradient were found to generate acoustic waves. Therefore, if those acoustic emissions could be detected in-situ, it is conceivable to properly identify acoustic fluctuations using machine learning in order to detect and identify the formation of keyhole porosities. Similarly, the authors have also demonstrated the ability of custom-made multispectral sensor, which uses filtered photodiodes to monitor key emissions spectra generated by the melt pool plume during laser processing, to detect both keyhole pores and lack-of-fusion flaws in directed energy deposition (DED) and laser-based powder bed fusion (L-PBF) AM processes [8–10]. The ability to detect spatter particles, which have been hypothesized to result in lack-of-fusion flaws [11–15], has also been demonstrated through the use of high-resolution layerwise images [16].

The goal of this work is to demonstrate the application of machine learning methods to such in situ sensor data to automatically detect flaws, facilitating process qualification and enabling future efforts focused on feedback control and automated flaw repair. A recent overview of machine learning techniques as applicable to AM is given in [17]. Benefits, utilities and use cases are discussed for supervised, unsupervised, semi-supervised, and reinforced learning are discussed. Current ML applications to AM range from design, part quality, and process optimization to in-situ monitoring for quality control. The latter

^{*} Corresponding author.

E-mail address: jup37@arl.psu.edu (J. Petrich).

is most relevant to this paper, and several studies exist that assess individual sensor modalities for flaw correlation and detection [18–21]. In [18], methodologies for monitoring the melt pool for PBFAM are discussed. Specifically, high speed video is used to extract melt pool signatures. Process parameters such as beam velocity and power are varied throughout the build to artificially create regions with higher likelihood for flaws to form, i.e. via keyholing, under-melting, or balling. Using computer vision techniques in concert with machine learning, the authors correlate the extracted melt pool signature with process parameters in order to determine the likelihood of flaw formation. Acoustic emissions and their utility for process monitoring are discussed in [19], while image-based solutions are proposed in [20] and [21].

This paper expands upon existing research by introducing multiple sensor modalities (i.e., high-resolution layerwise images, acoustic and multispectral spectra, and commanded laser scan vectors) and data fusion to the problem at hand. Using this methodology, complimentary information content can be shared between sensors to further increase the correlation to actual flaws and thus pave the way for high confidence flaw detection in the future. In addition, we conduct a sensitivity study to assess information content for each modality, which may guide future sensor selection and integration efforts. Ground truth labels, i.e. the actual location of flaws, is available through X-ray CT scanning, automated defect recognition (ADR) tools [21], and proper data registration, see [22] and [23]. We therefore attempt to correlate in-situ data directly to flaws on a voxel-by-voxel basis.

The paper is organized as follows. Section 2 describes the technical approach, including the PBFAM test bed and data acquisition, data registration and rasterization, label and feature extraction for supervised machine learning, the data fusion and classification framework, and the performance assessment strategy. Experimental validation using existing PBFAM test coupons is carried out in Section 3. Several performance metrics for data correlation are discussed and trade-offs for potential flaw detection are studied. Section 4 concludes the paper, while Section 5 outlines future work.

2. Technical approach

This section outlines the technical approach as a complete tool chain that ingests the raw, in-situ sensor data as well as high resolution CT

data in order to establish a statistically meaningful correlation between physical flaws and the corresponding multi-modal sensor footprint. As shown in a high level overview in Fig. 1, the AM build is monitored in real-time, which generates a complete set of in-situ sensor data. Post build, the AM part is CT scanned at a high resolution in order to identify and localize actual flaws with high confidence. Data registration techniques then allow binary ground truth labels, e.g. flaw or no flaw, to be assigned to the corresponding multi-modal footprint. Data correlation is then formally established by tasking a Neural Network to discriminate potential flaws based solely on their in-situ footprint.

This section is organized as follows. Section 2.1 introduces the PBFAM test bed and overall data acquisition strategy. Available sensor modalities are discussed in detail including their respective frame rates and/or resolutions. Data registration and the rasterization of temporal modalities is outlined in Section 2.2. Data curation between spatial and temporal modalities prior to data fusion is highlighted. Section 2.3 focusses on the extraction of ground truth labels from high resolution CT scans, which are needed for supervised machine learning. In addition, 3-D feature extraction within the in-situ domain is discussed. The data fusion and classification framework is presented in Section 2.4. Again, the objective here is to establish a statistical correlation between the in-situ sensor footprint, i.e. the extracted features, and the existence of flaws within the part as indicated by CT scans. Section 2.5 outlines the overall approach to assess correlation and introduces strategies and metrics commonly used for binary classification.

2.1. PBFAM test bed and data acquisition

In-situ process monitoring data was acquired using the test bed shown Fig. 2 and first presented in [24] and [25]. The test bed is integrated into a 3D Systems ProX 320 machine, a commercially available L-PBF system [26], and outfitted with an array of sensors for collecting relevant processing information, including:

- **Layerwise electro-optical (EO) imagery.** Six layerwise images are collected for every build layer: three after recoating but before laser scanning and three after the laser processing. The three images before and after laser processing each use different flash modules, labeled *front*, *side* and *top*, to provide different lighting conditions of

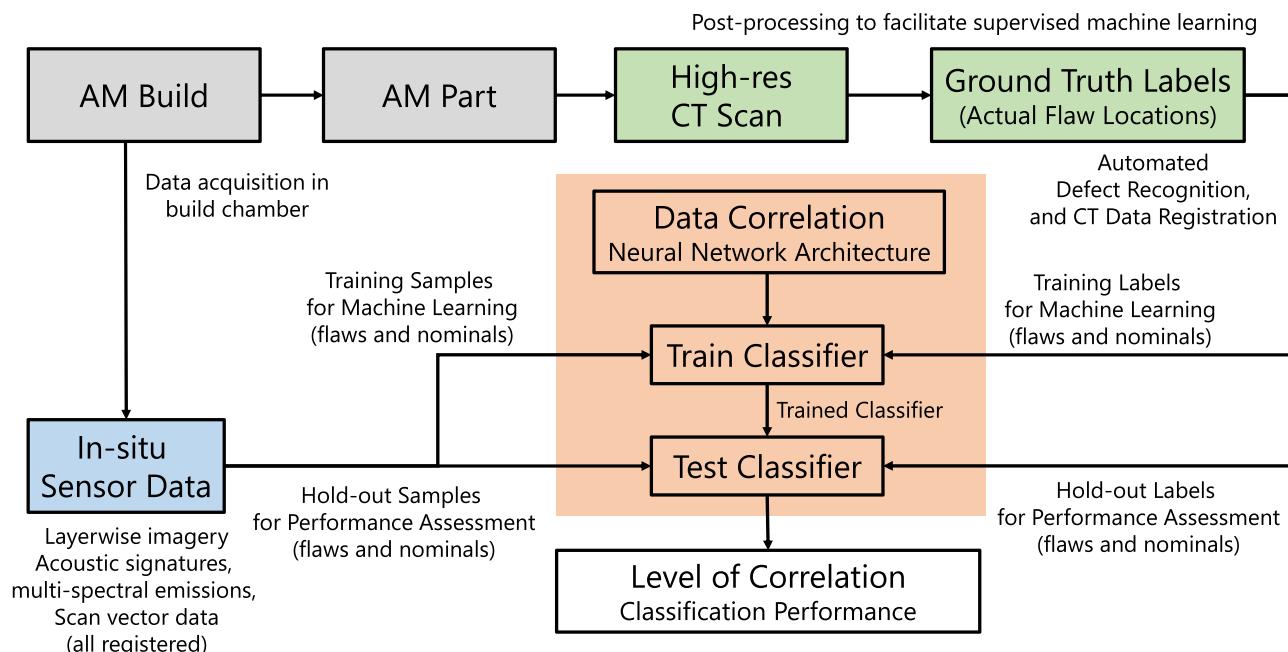


Fig. 1. Strategy for correlating in-situ sensor data to flaws generated during the PBFAM process.

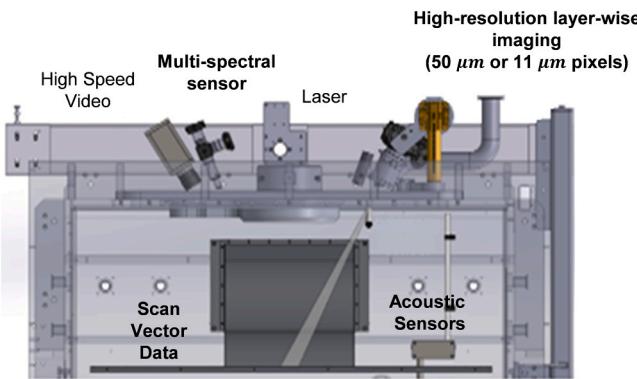


Fig. 2. Available PBFAM test bed including sensor modalities from [24] and [25]. The test bed is integrated into a 3D System ProX 320 [26].

the powder bed (before laser scan) and the build layer (after laser scan). Image resolution is $10 \mu\text{m}$ per pixel, corresponding to a $10 \mu\text{m} \times 10 \mu\text{m} \times 60 \mu\text{m}$ voxel in the image domain (due to $60 \mu\text{m}$ layer thickness).

- **Acoustic monitoring.** One microphone was placed in the build chamber to monitor acoustic emissions. With a sampling frequency of 192 kHz , the microphone covers frequencies inside and outside the human audible range.
- **Multi-spectral (MS) emissions.** Multi-spectral emissions were captured at 520 nm and 532 nm at a sampling frequency of 50 kHz .

The two wavelengths represent the line and continuum emissions for Ti6Al4V, as discussed in [27–29].

- **Scan Vectors.** The ProX 320 machine provides access to the machine logs, including the laser scan trajectory and power as function of time. Scan vectors are logged at 50 kHz . Raw scan vector information may then be converted into process metrics, such as hatch-to-contour angles, scan vector direction relative to the inert gas flow, or hatch length.

Fig. 3 shows a sample of the collected layerwise imagery for one layer. Three images are taken before laser scanning (top row) and after laser scanning (bottom row). Different lighting conditions are clearly visible when using the three flash modules, here Front, Side and Top (left to right). Given an image resolution of $10 \mu\text{m}$ per pixel, the actual field of view shown here is about $6.5 \text{ mm} \times 6.5 \text{ mm}$. The diameter of the cylindrical part is 5 mm .

A sample of the scan vector data is shown as a point cloud in **Fig. 4**. Scan vectors are logged by the machine (a 3D Systems Pro X 320) and contain information regarding the AM build strategy, including (i) time, (ii) laser position, (iii) laser power, and (iv) laser speed. From the raw scan vector data, several process metrics believed to impact the build quality may be calculated. Those are shown in **Fig. 5** and include:

1. Distance to contour for each xy location measured in mm or μm
2. Distance to point cloud for raster, i.e. the distance of each raster cell (or pixel) to the closest point in the scan vector point cloud in mm or μm (see also data rasterization in **Section 2.2**)

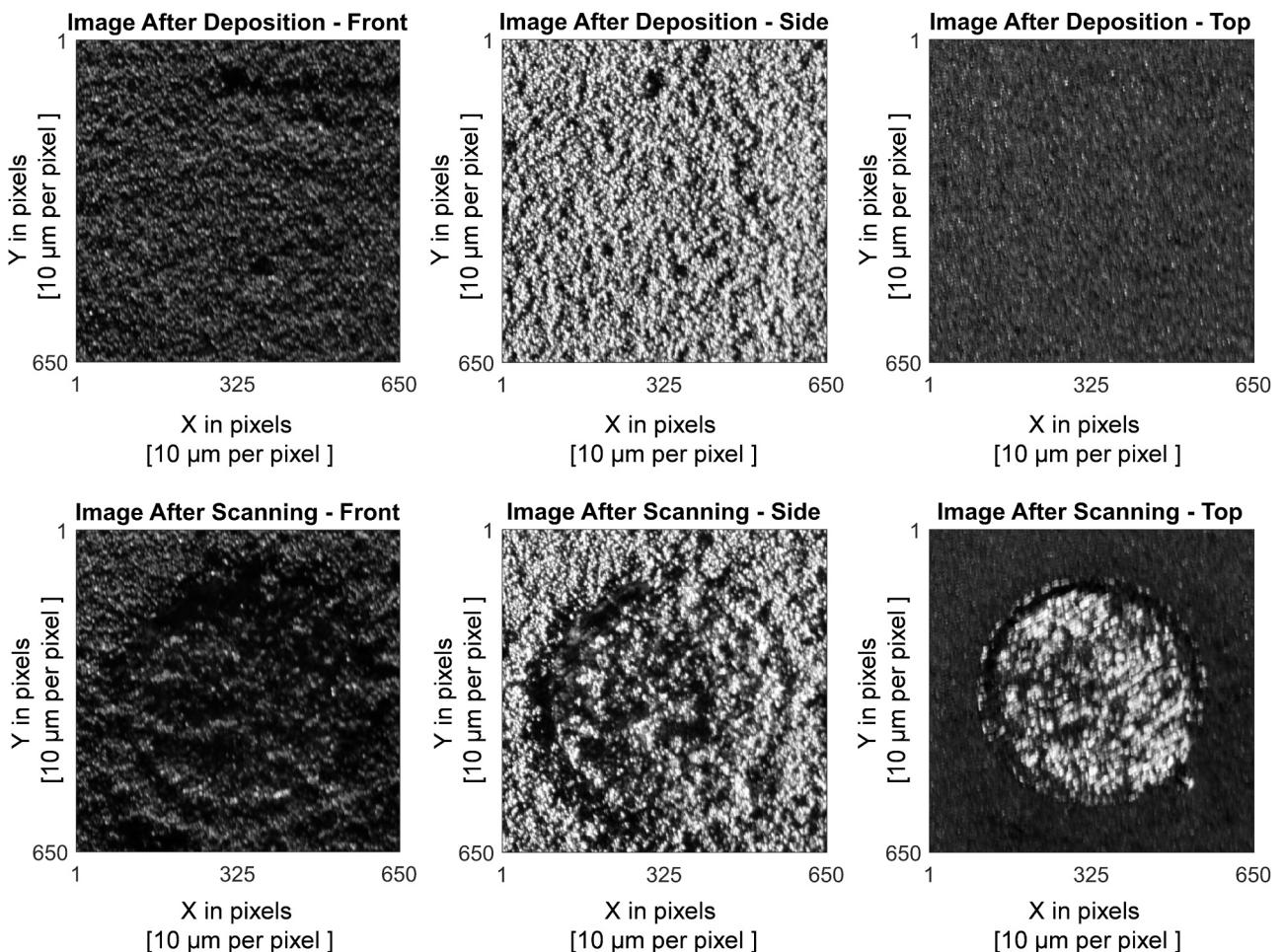


Fig. 3. Layerwise imagery collected during the PBFAM build process, showing images taken before laser scanning, i.e. after powder deposition, (top row) and after laser scanning (bottom row). Different lighting conditions are clearly visible when using the three flash modules, here Front, Side and Top.

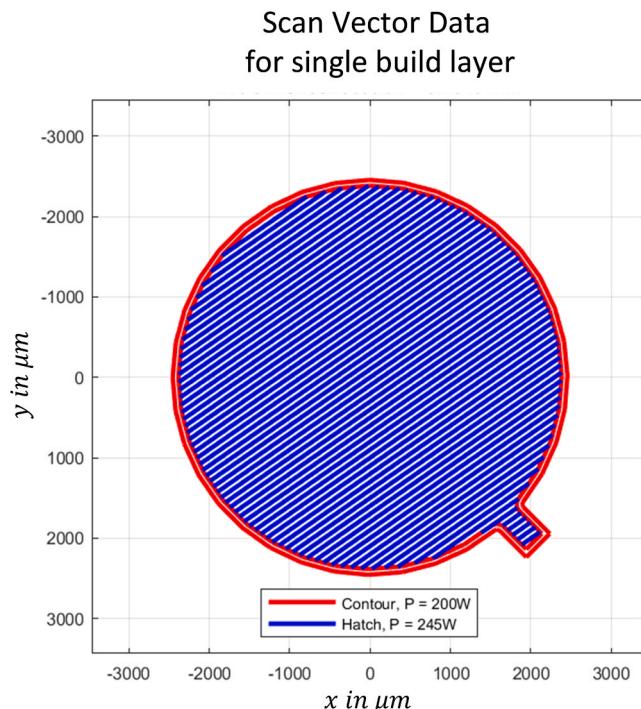


Fig. 4. Scan vector sample data for single build layer. Here, laser power is color-coded showing hatches (245 W) in blue and two contours (200 W) in red. (For interpretation of the references to colour in this figure, the reader is referred to the web version of this article.)

3. Hatch-to-contour angle in deg, here 0 – 90 deg
4. Hatch angle relative to the build plate (using \cos and \sin to void discontinuity at 0–360 deg or +/- 180 deg wrap around)
5. Hatch length in mm or μm
6. Hatch sequence as integer

As outlined in [27] and [28], multi-spectral and acoustic emissions also hold valuable information regarding the build process. For the former, the wavelengths considered for process monitoring can be selected in accordance with the line emissions specific to the material used [30]. The test bed in Fig. 2 is equipped with two Avalanche photodiodes that are tuned via optical filtering to 532 nm and 520 nm to capture continuum and line emissions for Ti6Al4V, respectively. Similarly, an ICP prepolarized microphone records acoustic emissions in the build chamber at 250 kHz. All temporal data is synchronized with the machine log files and stored together with the vector data for ease of data alignment. Data registration and rasterization will be discussed in Section 2.2.

In addition to in-situ sensor modalities, high-resolution Computed Tomography (CT) scans are collected post build. Those data are utilized to generate ground truth for supervised machine learning, similar to the work in [20] and [21]. Here, ground truth refers to the existence of flaws and their actual, physical locations relative to the build plate.

Registration of all sensor modalities as well as the rasterization of temporal modalities, here acoustic and MS, are discussed in the next section.

2.2. Data registration and rasterization

For the work presented in this paper, build plate coordinates as defined by the scan vectors serve as the reference coordinate frame. Therefore, all process monitoring data modalities outlined in the previous section, as well as the post-build CT data, need to be properly registered to the scan vectors before data fusion and training may occur. A framework to register in-situ sensor data and post-build CT scans was developed in [22] and [23]. In general, data registration is performed by aligning the respective part surfaces, as extracted from different modalities, using a homography (or coordinate transformation). The parameters that define the homography, mainly translation, rotation, scaling and skewing, can be deduced iteratively using a least squares optimization. Fig. 6 shows the result of the registration with the scan vectors as reference in the center and registered CT data and layerwise

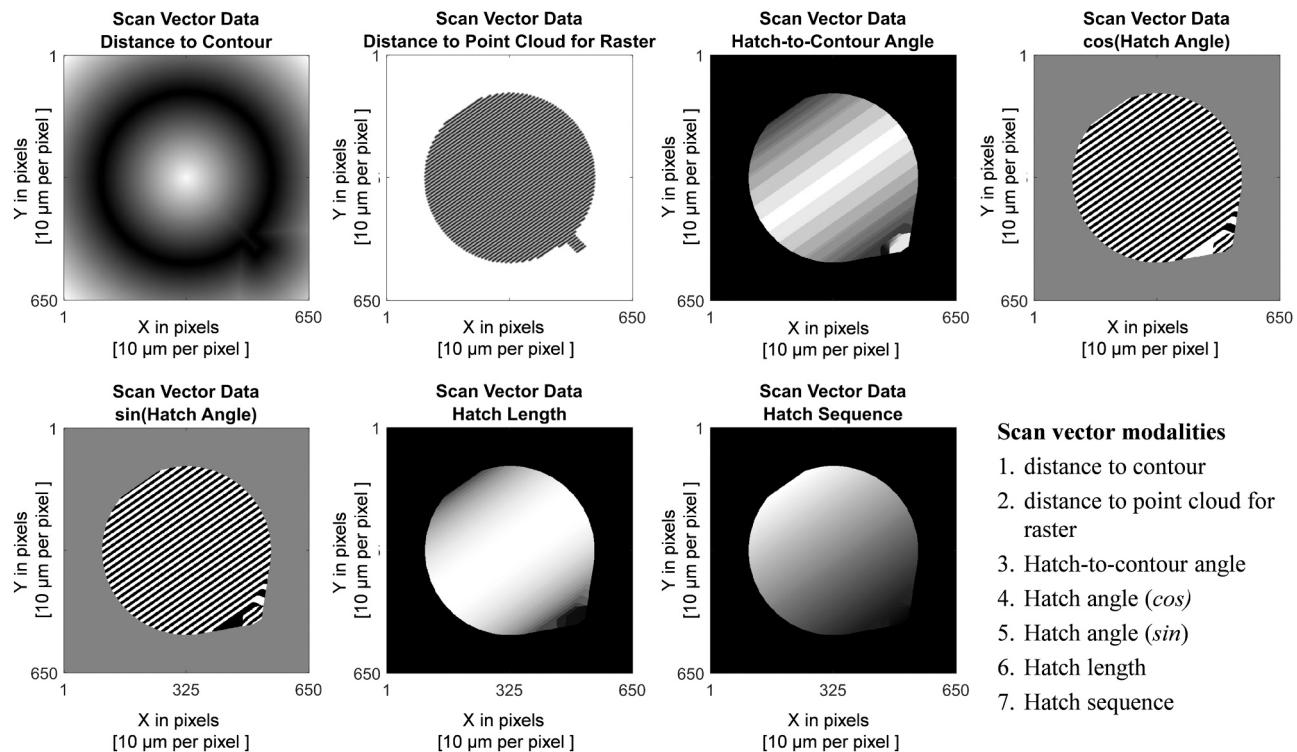


Fig. 5. Additional modalities derived from the scan vectors.

Scan vector modalities

1. distance to contour
2. distance to point cloud for raster
3. Hatch-to-contour angle
4. Hatch angle (\cos)
5. Hatch angle (\sin)
6. Hatch length
7. Hatch sequence

EO imagery to the left and right, respectively.

Temporal data, i.e., the multispectral and acoustic data, are also registered to machine scan vectors. While the former is recorded synchronously with the scan vector data, the latter needs to be compensated for the time of flight (ToF). ToF compensation for acoustic emissions inside the PBFAM chamber is an area of active research. However, early findings indicate that a spike in acoustic energy can be correlated directly with the laser on/off power cycles. Therefore, given a sufficiently close initial estimate for ToF, one can maximize the correlation between acoustic energy and laser on/off cycles to determine the ToF more precisely and to improve data registration.

Once registered to the vector data, acoustic and multi-spectral emissions are available as a point cloud with corresponding time t , layer number l , and location x,y for each data sample. Subsequent rasterization transforms both point cloud into a raster (or image) similar to the layerwise imagery. During rasterization, voxel values are assigned based on the closest data point. Adjacent data points may be averaged if multiple points fall into the same voxel. Fig. 7 shows the resulting rasterized images for MS (left) and acoustic (right) emissions. Here, both registered point clouds have been projected onto the same xy grid used for the layerwise EO imagery from Fig. 3.

Currently, the Digital Imaging and Communications in Medicine (DICOM) standard [30] is utilized to store metadata information, such as the position, orientation, and resolution of all imagery.

2.3. Label and feature extraction for supervised machine learning

Within the field of machine learning, one commonly distinguishes between supervised and unsupervised approaches. While the latter derives the target classes from the data structure itself, the former uses prior knowledge on what target classes exist in the data and assumes a label (or ground truth) to be given for each data sample. For the application at hand, such ground truth labels, i.e. “flaw” or “no flaw”, can be assigned to specific xyz voxels in the build coordinate system using registered post-build CT scans. Again, CT scans are properly registered (see Section 2.2) such that the xyz-coordinates are seamlessly transferrable from the CT domain to the build plate reference frame.

Automated Defect Recognition (ADR) techniques for CT scans have previously been developed and validated (see [20] and [21]). In short, the ADR algorithm employs machine perception and visual saliency techniques to identify local extrema within the gray scale intensity (or voxel value) of the CT scan that are often indicative of actual flaws. CT voxels are flagged individually as either nominal or anomalous. Adjacent anomalous voxels in 3D space are merged into clusters to form the flaw. Fig. 8 shows an example of a detected flaw consisting of 746 individual CT voxels at a resolution of $(10\mu\text{m})^3$, which represents a spherical equivalent diameter of roughly $100\mu\text{m}$. The flaw (here it's a gas pore or other void) manifests itself through a very low CT intensity as compared to its local neighborhood. Similarly, if CT return intensities are close to uniform over a local region, then that region can be labeled as fully dense or nominal, i.e. no flaw exists.

Using the ADR software and transferring the respective xyz locations

of all detected flaws into the build plate reference frame, one can assign ground truth labels to each voxel in the registered imagery. Note that, due to image registration, labels apply similarly to all rasterized images regardless of modality. More formally, we define the ground truth label for data sample i as

$$Y_i \in \{0, 1\} \quad (1)$$

where $Y_i = 0$ indicates a nominal build condition (i.e. no observable flaw exists), and $Y_i = 1$ indicates a flaw.

Let N_1 denote the number of data samples, or voxels, that have been labeled as flaws. For training purposes, four nominal samples are selected for each of those voxels using the following two criteria. First, we randomly select two nominal data samples from the same layer. Second, we randomly select two additional data samples from the entire data stack. This strategy was chosen to maintain a class balance of 1:4, and to preserve the spatial distribution of data samples, thereby preventing the classifier from homing in on a particular region, or layer, within the build. Note that flaw detection in AM is inherently a highly imbalanced classification problem, with nominal voxels outnumbering flaws by several orders of magnitude. In order to avoid the need for significant class balancing for training and testing, a ratio of roughly 1:4 was chosen. Consequently, the number of negative data samples, i.e. nominal conditions, is $N_0 \approx 4N_1$.

Features for supervised machine learning are extracted for each of those $N = N_0 + N_1$ samples. Specifically, for any registered sensor modality, we consider the voxels in a three-dimensional (3D) local neighborhood around the labeled voxel. Based on anticipated flaw sizes (determined empirically from analysis of the CT data) and residual registration errors, the volume of the local neighborhood was chosen to be $(300\mu\text{m})^3$ ³. An example of such a feature extraction volume is illustrated on the left in Fig. 9. Given the resolution of the collected in-situ imagery of $10\mu\text{m}$ in the xy-direction and $60\mu\text{m}$ in z-direction (i.e. build height), the feature vector for a single sensor modality contains $31 \cdot 31 \cdot 5 = 4805$ elements or dimensions. It will be denoted as $\tilde{X}_{i,m} \in \mathbb{R}^{4805}$, where $i = 1, \dots, N$ indicates the data sample and $m = 1, \dots, M$ indicates the sensor modality, with M representing the number of all available sensor modalities, here $M = 15$.

In order to reduce the number of features used for machine learning and to prevent potential overfitting, the xy-resolution was reduced to $30\mu\text{m}$ using image blurring and resizing. The reduced feature vector $X_{i,m}$ then has a dimension of $11 \cdot 11 \cdot 5 = 605$. Using Ψ to denote the projection within the feature space that yields the desired feature reduction, we find

$$\Psi \cdot \tilde{X}_{i,m} \in \mathbb{R}^{4805} \rightarrow X_{i,m} \in \mathbb{R}^{605} \quad (2)$$

Since all data has been properly registered, feature reduction via Ψ is independent of sensor modality m .

2.4. Data fusion and classification framework

For data fusion and classification, we leverage and modify existing

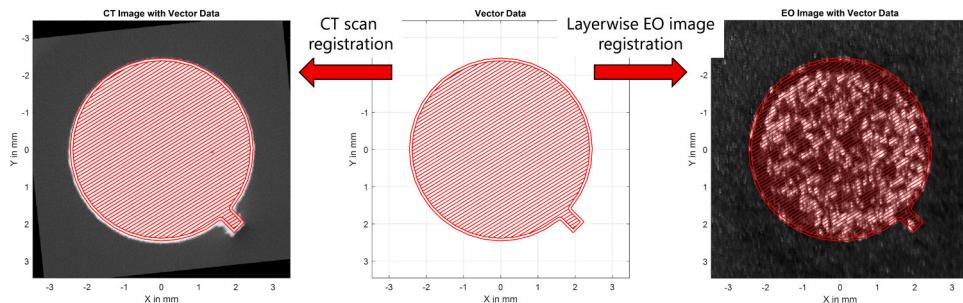


Fig. 6. Data registration using scan vector (center plot) data as reference frame. Registered CT data and layerwise EO data is shown on the left and right, respectively.

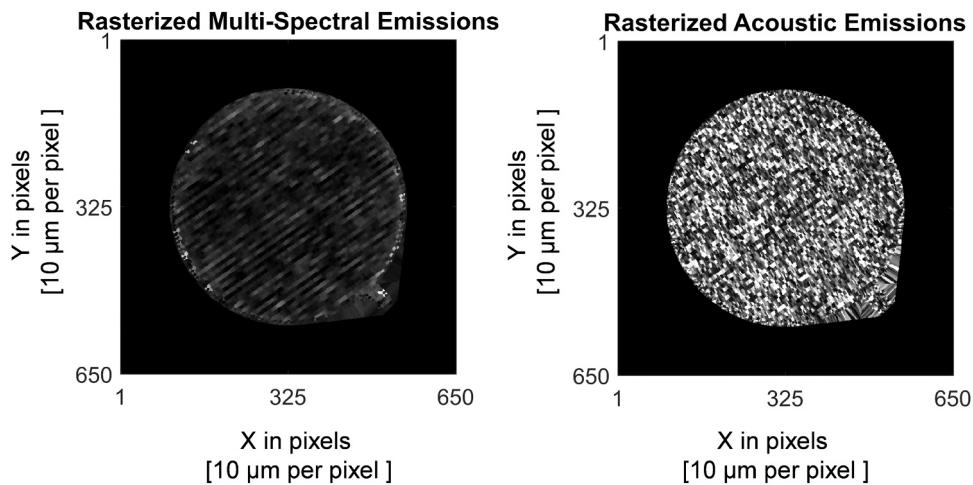


Fig. 7. Rasterized images of registered MS data (left) and acoustic data (right). Both point clouds have been projected onto the same xy grid used for the layerwise EO imagery shown in Fig. 3.

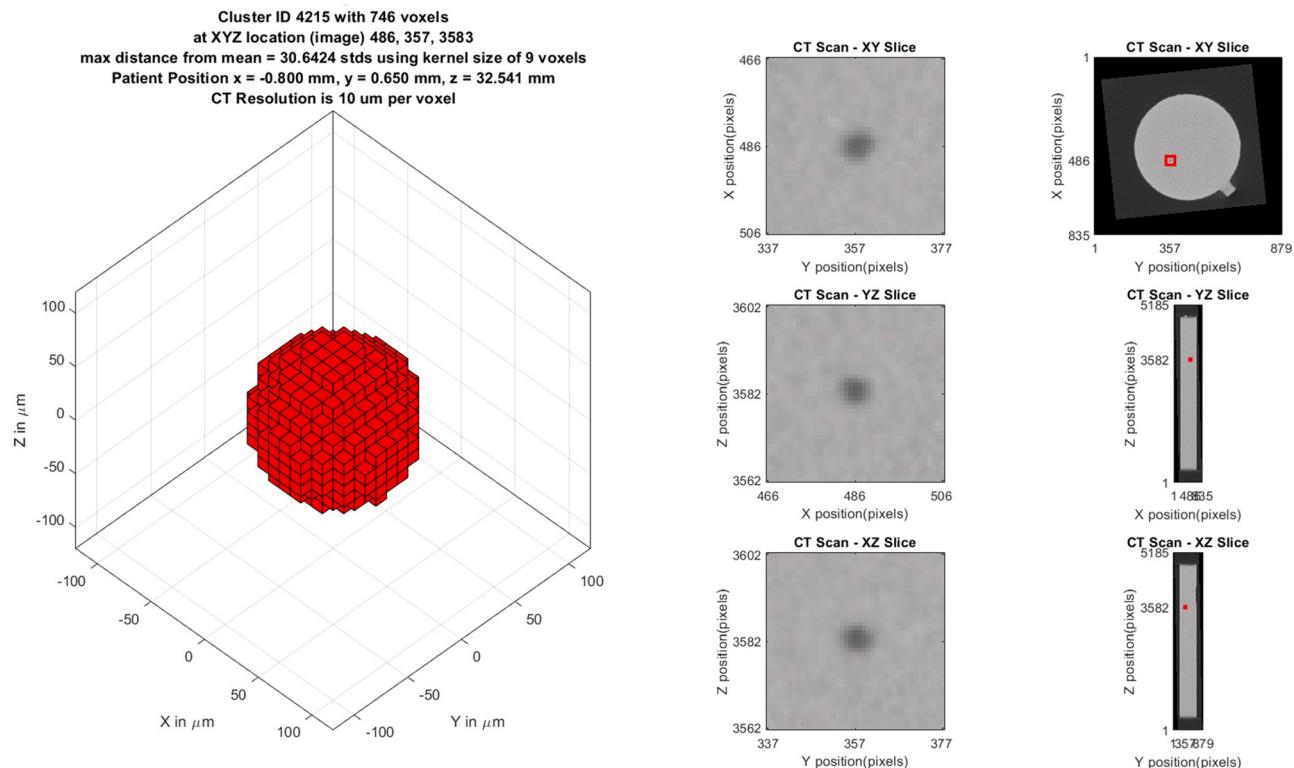


Fig. 8. Physical AM flaw (roughly 100 μm in diameter) as captured via CT scanning and subsequently identified using ADR software. Here, the flaw (pore) manifests itself through a very low CT return intensity as compared to its local neighborhood.

binary classifiers from [21]. Specifically, we propose the neural network (NN) architecture from Fig. 10. The selected architecture utilizes a hidden layer with states $H_{i,m} \in [0, 1]^{24}$, one for each sensor modality m , where i indicates the data sample index and $m \in [1, M]$ indicates the sensor modality. The dimension of $H_{i,m}$ is a design parameter trading computational complexity for overall fidelity of the underlying data representation. For the application at hand, the dimension for the hidden layer was chosen to be 24 in order to achieve similar reduction in dimensionality for the signal path from input to hidden layer (i.e. 605:24) and hidden to output layer (i.e. 24:1). Assuming a fully connected neural network for each sensor modality m , the signal path between input layer and hidden layer, for each individual modality, is then

given by

$$H_{i,m} = \sigma(W_{1,m} \cdot X_{i,m} + b_{1,m}) \quad \text{with } m \in [1, M] \quad (3)$$

using the weight matrices $W_{1,m} \in \mathbb{R}^{24 \times 605}$, bias terms $b_{1,m} \in \mathbb{R}^{24}$, and the nonlinear activation function σ , chosen to be the sigmoid function, a common choice for NN applications [31]. Note that sensor information from each modality m enters the hidden layer separately in order to keep the number of weights at the input layer tractable. Therefore, weight matrices and bias parameters are different for each modality as is indicated by the subscript m .

Sensor fusion is established at the hidden layer by concatenating the reduced order representations $H_{i,m} \in [0, 1]^{24}$ for each sensor modality m .

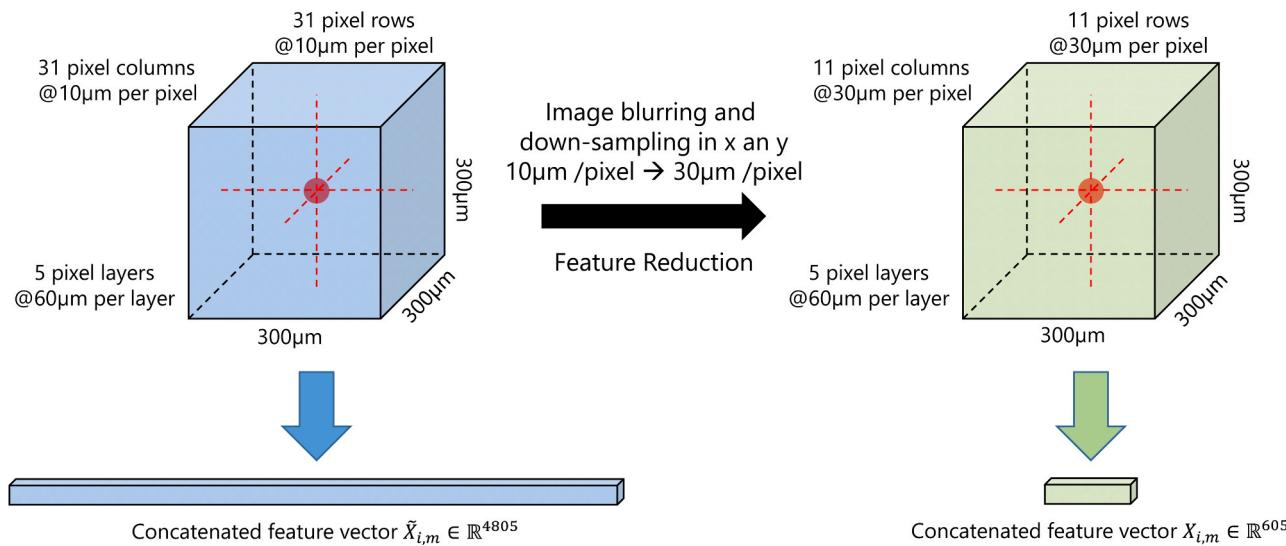


Fig. 9. Feature extraction and reduction for supervised machine learning, showing 3D neighborhood around labeled voxel and concatenated feature vector on the left, and reduced feature vector (using image blurring) on the right.

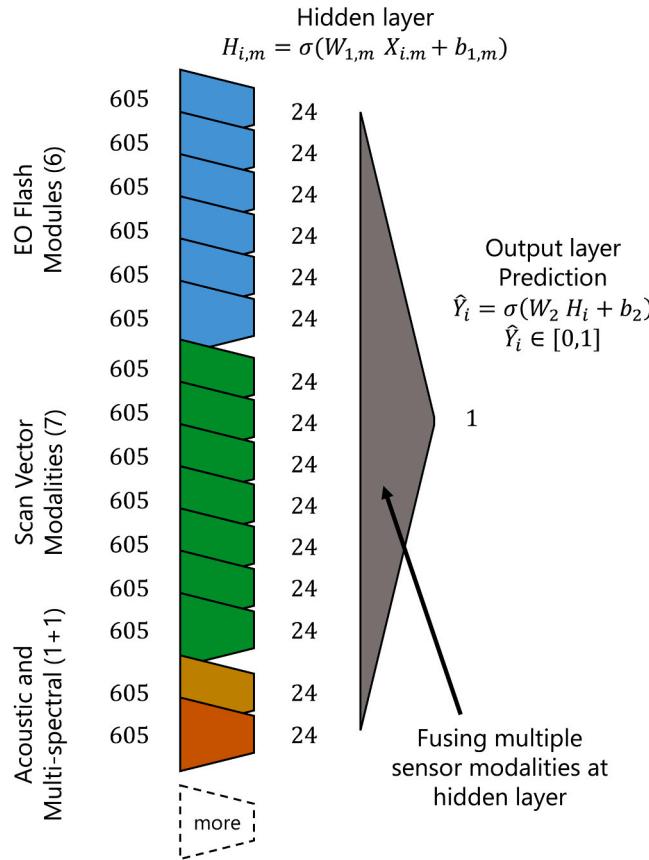


Fig. 10. Neural network architecture using one hidden layer to fuse all sensor modalities.

Formally, we define the state of the hidden layer given the input sample i as concatenation of hidden layer states $H_{i,m}$ from (3)

$$H_i = [H_{i,1}, H_{i,2}, \dots, H_{i,M}] \in [0, 1]^{24M} \quad (4)$$

A fully connected output layer then transforms the fused information, here H_i , into a prediction. Similar to (3), the prediction \hat{Y}_i for data sample i is constructed via

$$\hat{Y}_i = \sigma(W_2 H_i + b_2) \in [0, 1] \quad (5)$$

using the weight matrix $W_2 \in \mathbb{R}^{1 \times 24M}$, and the bias term $b_2 \in \mathbb{R}$.

Assuming that N_T out of the $N = N_0 + N_1$ available data samples are used for training, we can construct the vector of predictions $\hat{Y} = [\hat{Y}_1 \dots \hat{Y}_{N_T}] \in [0, 1]^{N_T}$, one for each sample, that is generated at the output of the NN. The objective for machine learning is to train the NN, i.e. estimate weight matrices and bias terms, in order to best approximate the vector of true labels Y from (1). Following guidelines and best practices from [31], a cross entropy cost function is utilized to quantify the difference between Y and \hat{Y} . The cross entropy for the collection of true and predicted labels is defined as

$$H(Y, \hat{Y}) = -\frac{1}{N_T} \sum_{i=1}^{N_T} [Y_i \ln \hat{Y}_i + (1 - Y_i) \ln (1 - \hat{Y}_i)] \quad (6)$$

The cross entropy effectively resembles the loss function between two probability distributions, and it is thus frequently used for optimization and probability estimation. Despite being highly nonlinear, the cross entropy improves and accelerates achievable learning rates during NN training. As shown in [32] this is especially the case when using gradient-descent-based optimization techniques and sigmoid-activation functions.

The proposed NN architecture was implemented in Python/TensorFlow,¹ and NN training was carried out using built-in backpropagation routines [31]. Backpropagation is an iterative, gradient-descent-based optimization scheme to successively update the weight matrices $W_{1,m}$ and W_2 as well as bias vectors $b_{1,m}$ and b_2 such that a pre-defined cost function, here $H(Y, \hat{Y}_k)$ from (6), will be minimized over time. Regularization terms may be added to prevent network parameters from growing too large during training [32].

It is important to note that a simple, shallow NN architecture was chosen intentionally in order to establish a proof-of-concept for sensor fusion and data correlation. Deeper NNs offer more flexibility, i.e. more weight parameters, and may thus outperform the proposed NN. However, the quantity of available training data may be insufficient, which can lead to overfitting, and a tradeoff between the number of data samples and network complexity is often required. Identifying the best

¹ <https://www.tensorflow.org/>

network architecture for the application at hand as well as the available training and test data is beyond the scope of this paper but may be revisited in the future.

Another way to reduce network complexity would be to identify and eliminate those data modalities that offer essentially the same information as another modality or those that offer no additional information benefitting the categorization, i.e. redundant modalities. This topic is revisited in [Section 3.2](#).

2.5. Performance assessment strategy

A four-fold cross validation is used to assess the strength of the data correlation and ultimately the performance of the classification framework. Therefore, one classifier, i.e. weights and bias term, will be trained on 75% of the data while the remaining 25% percent of the data are used as previously unseen test samples. Repeating this approach four times with randomly selected data in a round robin fashion allows us to include every data sample once for testing and three times for training. Correlation measures are then established using common performance metrics derived for binary classification problems, such as accuracy, recall, and precision, as well as false positive and false negative rates.

3. Experimental validation

The proposed algorithm for correlating in-situ sensor modalities with ex-situ X-ray CT scans as applied to PBFAM was formally evaluated on 1 AM part, a Ti6Al4V cylinder, using data collected throughout the build, see [Fig. 11](#). The cylinder ($\varnothing 5\text{ mm} \times 42\text{ mm}$ tall) was built on the instrumented 3D Systems ProX 320 system using powder sourced from GKN conforming to specifications for Grade 5 Ti-6Al-4 V [33]. Build parameters, e.g. laser speed, laser power, scan strategies, etc., were modified throughout the build to study the impact on build parameters on flaw formation. However, all data sample used in this study were extracted in regions that were built with OEM recommended parameters. Those regions are marked by the overhang visible in [Fig. 11](#).

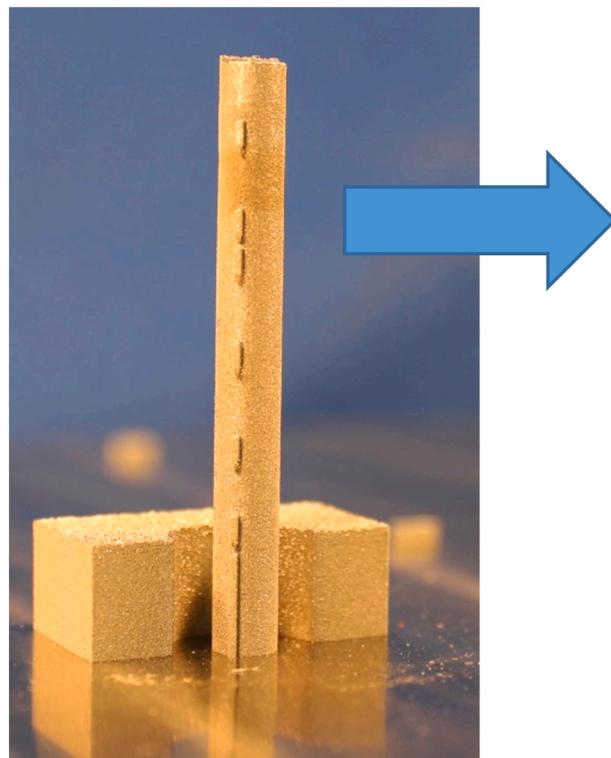


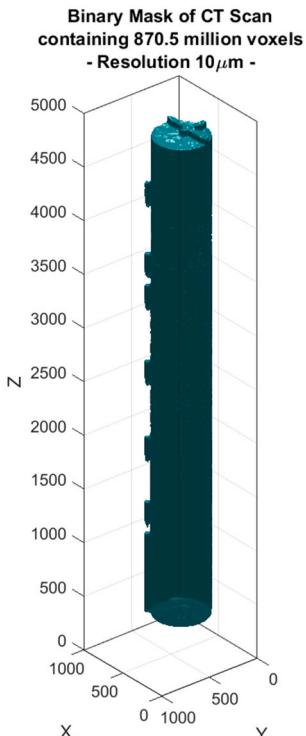
Fig. 11. Ti-6Al-4 V test specimen for experimental validation (left) and binary mask derived from CT scan (right).

[Fig. 12](#) summarized the outcome of the automated defect recognition (ADR) software for the given test specimen. Note that all flaws that fall within $500\mu\text{m}$ of the part surface are disregarded for this analysis. In other words, we focus on embedded flaws only, for which the 3D image chip from [Fig. 9](#) does not contain any part boundary or background. The ADR algorithm identified 1156 embedded flaws, which have all been verified manually. Once mapped from the CT domain into the in-situ domain, those 1156 flaws collectively occupy 32,352 voxels. After down-sampling as shown on the right side of [Fig. 9](#), individual voxels in the in-situ domain measure $30\mu\text{m} \times 30\mu\text{m} \times 60\mu\text{m}$. The left hand side of [Fig. 12](#) shows the distribution of flaws throughout the build using color to indicate size. Note that only a subset of layers was used for this study (i.e. OEM recommended build parameters), which creates gaps in the ADR results that are noticeable along the Z axis. The center plot visualizes the random selection of nominal voxels, at a ratio of roughly 4:1, yielding $N_0 = 136,222$ nominal voxels for machine learning (ML). Though "sample" is typical terminology used in ML circles the authors have elected to refer to "voxels", as providing a stronger link to the physical world more familiar to many additive manufacturing engineers.

Again, for each anomaly voxel, we randomly select four nominal voxels, two of which are taken from the same layer as the anomaly voxel. Histograms of flaw sizes, as represented by the equivalent spherical diameter, are shown on the right hand side. Frequency is measured by flaw, i.e. a cluster of adjacent anomalous voxels, (upper right) or by voxel (lower right). Flaw sizes of interest are bounded below by the image resolution, or $\sim 30\mu\text{m}$ spherical equivalent. The biggest flaws found in this data set are around $120\mu\text{m}$ in equivalent diameter. Although anomaly voxels are generally clustered, nominal voxels are not clusters here in order to expose the classifier to a wide variety of nominal build conditions.

3.1. Data correlation and binary classification

After NN training via four-fold cross validation, the four instances of



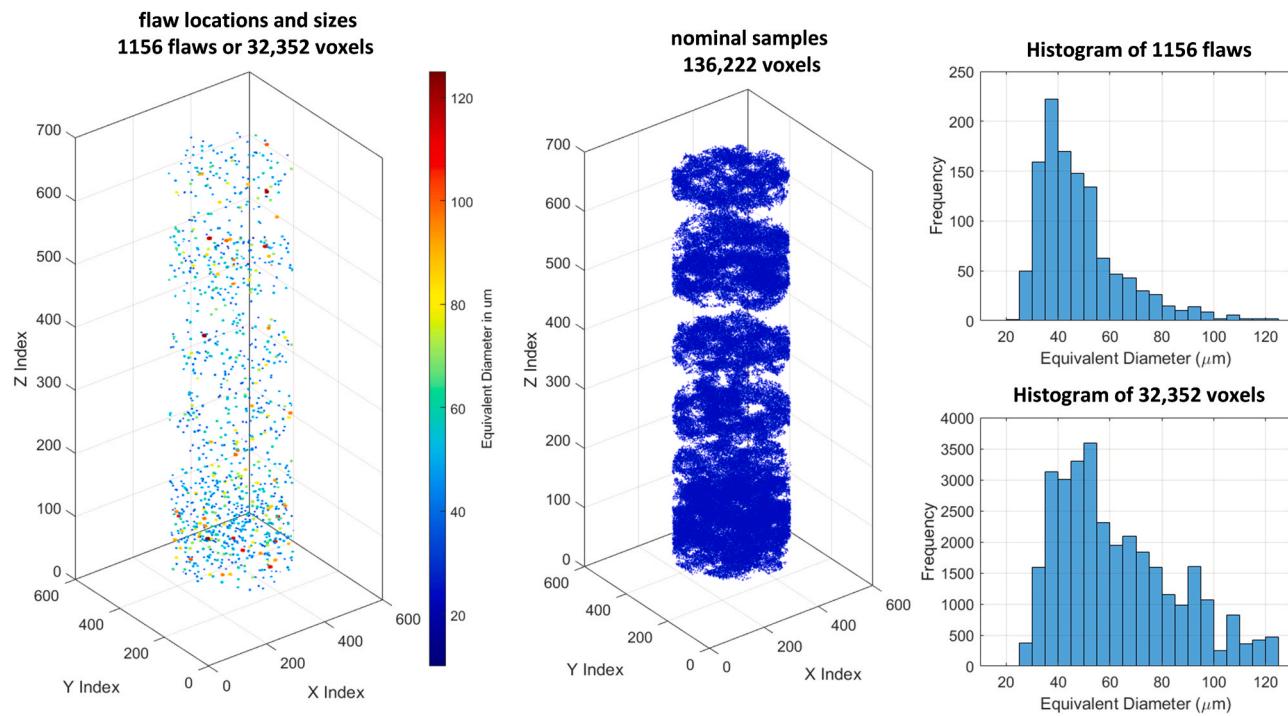


Fig. 12. Flaw locations, i.e. positive samples (left), versus randomly selected nominal locations, i.e. negative data samples (center). Histograms of all flaws as a function of equivalent diameter are shown on the right. Frequency is measured as number of flaws (upper right) and as number of voxels (lower right). (For interpretation of the references to colour in this figure, the reader is referred to the web version of this article.)

the classifier (one for each cross-validation) collectively create one prediction \hat{Y}_i for each voxel i , see (5). Training the NN on all 15 modalities of 136,222 nominal voxels and 32,352 flaw voxels using 10,000 epochs per cross-validation took roughly 20 hr on a Quadro P1000 GPU. The maximum number of epochs was determined based on the decay of the cost function (6). After 10,000 training epochs, the decline in the cost function appeared to be stagnant.

The resulting histogram of all $\hat{Y}_i \in [0, 1]$ with $i = 1, \dots, N$ is shown in Fig. 13. The predictions are color-coded for nominal voxels, $Y_i = 0$, in green and anomalous voxels, $Y_i = 1$, in red. For better comparison and

to offset the class imbalance, the frequencies for the latter are multiplied by a factor of four. A very strong correlation between ground truth labels and prediction is noticeable, indicating that “flaws” and “no-flaws” are very separable using only in-situ data.

In fact, using a classification threshold of 0.5 for the NN prediction \hat{Y}_i , one can construct a confusion matrix and formally evaluate the correlation using performance metrics commonly used for binary classification. The resulting confusion matrix when using a classification threshold of 0.5 is shown in Table 1. Out of the entire 168,574 voxels evaluated in the layerwise image domain, only 2461 voxels were misclassified, 2150 false positives (FPs) and 311 false negatives (FNs).

From Table 1, performance metrics commonly used for binary classification can be derived to quantify the level of correlation. They include accuracy, precision, recall, F1-score, false positive rate (FPR) and false negative rate (FNR). All metrics are summarized in Table 2.

Fig. 14 shows a comparison between NN predictions $\hat{Y}_i \in [0, 1]$ (left), ground truth flaw locations $Y_i \in \{0, 1\}$ (center) and the corresponding X-ray CT scan (right) for a single layer. Here, the layer with the most flaws was chosen. The CT and ground truth data indicates the presence of eight flaws spanning 295 pixels in this build layer. The nominal build conditions, i.e. no flaws, cover 113,065 pixels. This class imbalance of 1:380 is very common for otherwise very mature PBFAM processes.

The binary classification using a classification threshold of 0.5, as indicated in Fig. 13 is shown on the left in Fig. 15. The classifications are color-coded as follows: true negatives (green), true positive (red), false positives (blue) and false negatives (yellow). The resulting confusion

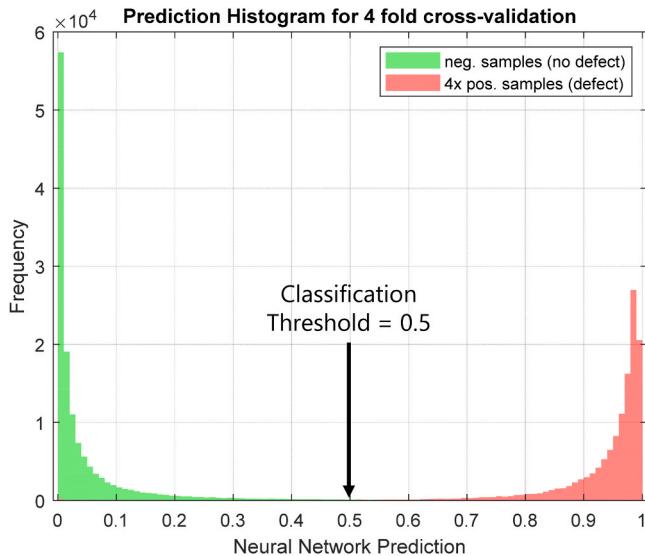


Fig. 13. Histogram of NN predictions \hat{Y}_i using four-fold cross validation color-coded by label: nominal (neg. samples) in green and anomalies/flaws (pos. samples) in red. (For interpretation of the references to colour in this figure, the reader is referred to the web version of this article.)

Table 1

Confusion matrix derived from NN predictions \hat{Y}_i shown in Fig. 13 using a classification threshold of 0.5.

	Classification threshold = 0.5		Prediction
	No Flaw	Flaw	
Ground Truth	No Flaw	134,072	2150
	Flaw	311	32,041

Table 2

Performance metrics for binary classification derived from the confusion matrix shown in [Table 1](#).

Accuracy	0.9854
Precision	0.9371
Recall	0.9904
F1-Score	0.9630
FPR	0.0158
FNR	0.0096

Note that precision and recall metrics are not dependent on the large number of true negatives which is typical for highly unbalanced data. They could thus be interpreted as object detection metrics.

matrix including classification performances such as accuracy, precision, recall, f1-score, false positive rate (FPR), and false negative rate (FNR), is provided in the top two rows in [Table 3](#). As one can see, flaw detection ability appears to be robust. However, due the large quantity of nominal samples, even a small false positive rate of less than 0.8% now generates 891 false positives.

It is notable that many false positives and false negatives on this layer appear to be immediately adjacent to actual flaws, so it is likely that application of a tolerance or clustering strategy would reduce both counts.

As with all binary classification tasks, one can generally trade false

positives for false negatives by adjusting the classification threshold applied to \hat{Y}_i . [Fig. 15](#) shows the resulting binary classification for a classification threshold of 0.7 (center) and 0.9 (right). Confusion matrices and classification performances for both scenarios are also included in [Table 3](#). As can be inferred from the histogram in [Fig. 13](#), increasing the classification threshold reduces the number of detections thereby reducing the number of false detections, i.e. false positives. In fact, the number of false positives reduces to 280 when using a classification threshold of 0.7, which results in a better classification at the center of [Fig. 15](#), see also [Table 3](#). Increasing the classification threshold to 0.9 further reduces the number of false positives to 52, see right plot in [Fig. 15](#) and bottom of [Table 3](#). However, the classifier now missed 9 anomalous voxels, i.e. false negatives.

Additional post-processing step may be taken to properly filter false positive as function of both size and confidence. However, those will need to be developed in concert with potential mitigation options that may be triggered after a flaw has been detection. However, those steps are beyond the scope of this work.

Although the derived performance metrics indicate a very strong correlation between in-situ process data and the existence of flaws within the build as indicated by X-ray CT scans, the ability to perform actual in-situ flaw detection will further require this level of correlation to extrapolate to other data sets. The extrapolation of the NN classifier towards other AM parts and/or builds will be the subject of further investigation.

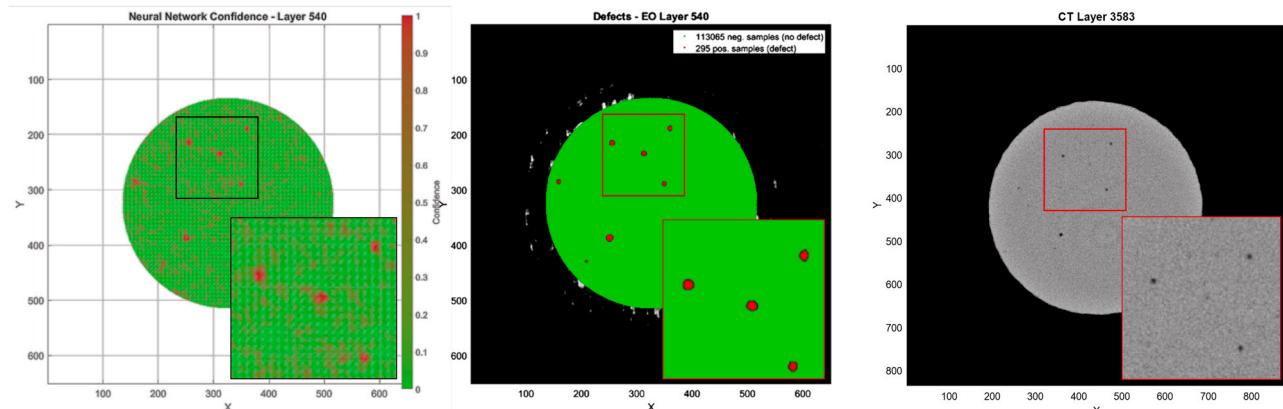


Fig. 14. NN prediction \hat{Y}_i for a single layer (left) indicating high confidence of “flaw”/“no flaw” in red/green, respectively, ground truth flaw locations (center), and corresponding 2d CT slice (right). (For interpretation of the references to colour in this figure, the reader is referred to the web version of this article.)

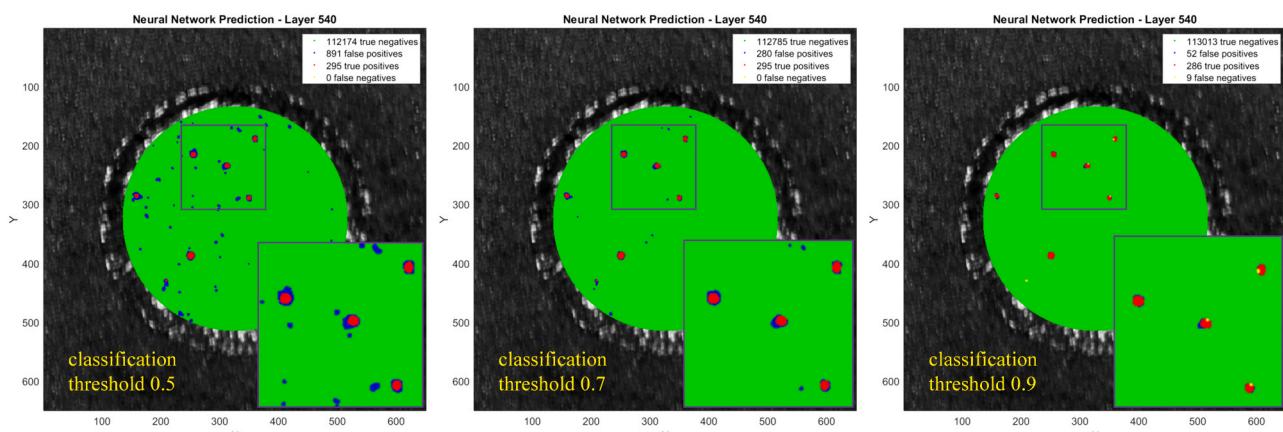


Fig. 15. Binary classification using predictions $\hat{Y}_i \in [0, 1]$ and a classification threshold of 0.5 (left), 0.7 (center) and 0.9 (right). The classification is color-coded as follows: true negatives (green), true positive (red), false positives (blue) and false negatives (yellow). (For interpretation of the references to colour in this figure, the reader is referred to the web version of this article.)

Table 3

Binary classification performance using predictions $\hat{Y}_i \in [0, 1]$ and a classification threshold of 0.5, 0.7, and 0.9 from Fig. 15.

Classification Threshold	0.5	Ground Truth	NN prediction		Correlation / Performance metrics					
			No Flaw	Flaw	Accuracy	Precision	Recall	F1-Score	FPR	FNR
0.7	0.7	Ground Truth	No Flaw	112,174	891	0.9921	0.2487	1.0000	0.3984	0.0079
			Flaw	0	295					0.0000
0.9	0.9	Ground Truth	No Flaw	112,785	280	0.9975	0.5130	1.0000	0.6782	0.0025
			Flaw	0	295					0.0000
		Ground Truth	No Flaw	113,013	52	0.9995	0.8462	0.9695	0.9036	0.0005
			Flaw	9	286					0.0305

3.2. Sensitivity study and information content

In an effort to study the viability of the proposed ML framework for technology transition and integration, identifying relevancy and redundancy within the data is crucial. Clearly, sensor modalities that correlate strongly to flaws carry higher priority for future investigations and investment. Therefore, a sensitivity study was conducted to assess overall information content (i) for each sensor modality individually, and (ii) for subsets of sensor modalities. Specifically, the NN architecture from Fig. 10 was truncated to ingest only a subset of input signals. Similar to the analysis in Section 3.1, we measure correlation using performance metrics for the resultant binary classification.

Fig. 16 shows the Receiver Operating Characteristics (ROC) curves for all single modalities and subsets of modalities. ROC curves are chosen here because they represent general classification performance over all classification thresholds. By definition, ROC curves plot the true positive rate (TPR) over the false positive rate (FPR). For a classification threshold of 0, all voxels are predicted as flaws, implying $TPR = 1$ and $FPR = 1$ (which is at the upper right corner in the ROC curve). Similarly, for a classification threshold of 1, all voxels are predicted as nominal, i.e. no flaws, implying $TPR = 0$ and $FPR = 0$ (with is at the lower left corner in the ROC curve). Sliding the classification threshold from 0 to 1 using small increments creates the ROC curve. The area under the ROC curve (AuC) is then commonly used to measure classification performance across all classification thresholds.

Fig. 17 shows a bar graph of the achievable AuC for each scenario

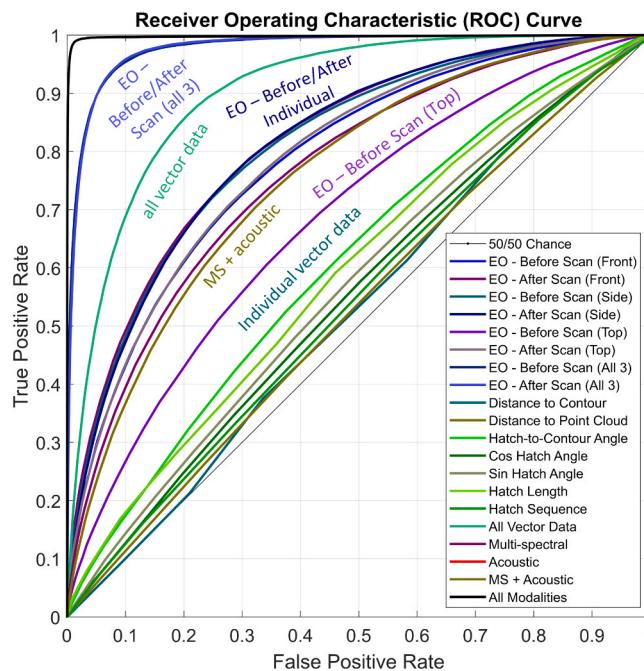


Fig. 16. Classification performance as measured by the Area under the ROC Curve for individual sensor modalities and subsets of sensor modalities.

under consideration. The bar graph is partitioned into sensor modalities relating to layerwise imagery (blue), scan vector-based metrics (green) and emissions (magenta). As can be seen, layerwise EO modalities (blue) appear to carry the most information with AuCs well above 0.75. The exception is the EO – before laser scan - top flash for which AuC drops below 0.7. This can be attributed to the limited powder bed topography information that can be extracted from an overhead flash. An example of all EO modalities, before and after laser scan, was shown in Fig. 3. Once all layerwise modalities are fused, the classification performance jumps to $AuC > 0.97$.

Interestingly, individual vector data modalities (green) only achieve AuCs of 0.52–0.61, suggesting little relevant information content. After sensor fusion, however, AuC reaches above 0.9, see all vector data. Scan vector modalities therefore seem very complementary and in their entirety correlate surprisingly well with actual flaw location. Scan vector data may in fact generate a meaningful assessment for where flaws are more likely to form during the PBFAM process before the process even begins, which could inform and improve scanning strategies. Additional investigation will be needed to validate this suspicion.

Lastly, using multi-spectral and/or acoustic data, we find AuC of around 0.75 with no significant increase when fusing both modalities. This could indicate high relevance for both but potential redundancy. When using all modalities (black), overall correlation increases to an AUC of close to 1. This case was discussed in detail in Section 3.1.

4. Conclusion

This paper presents and validates a complete solution for correlating

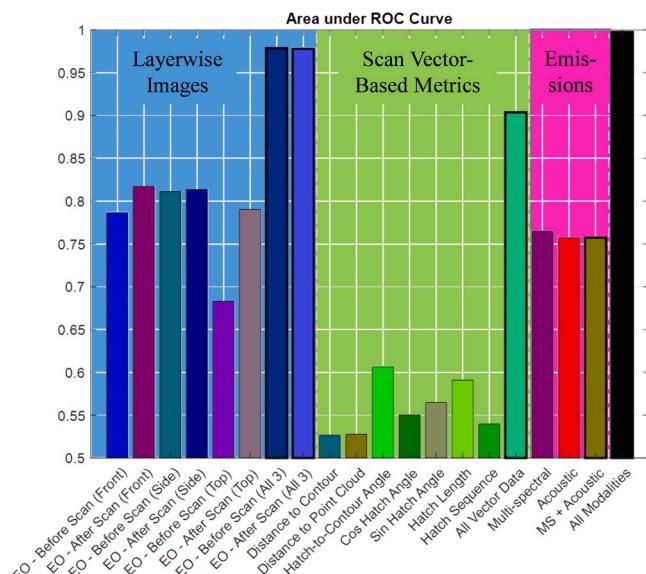


Fig. 17. Bar graph visualizing the area under the ROC curve for each sensor modality and subset of sensor modalities from Fig. 16. (For interpretation of the references to colour in this figure, the reader is referred to the web version of this article.)

in-situ sensor data acquired during the PBFAM process with actual flaw location as extracted using post-build X-ray CT scans. The level of correlation is assessed using performance metrics commonly used for binary classification. As such, the proposed framework may lay the foundation for in-situ flaw detection in the future. The overall strategy for data correlation as well as the available test bed for data collection are described in detail. Several data modalities including (i) layerwise imagery, pre and post laser scan, (ii) acoustic and multi-spectral emissions, as well as (iii) scan vector data are captured and subsequently registered to a common reference frame, here the build plate. Temporal modalities are synchronized with the scan vector data provided by the AM machine and rasterized to the same *xyz* grid that defines the layerwise imagery. Ground truth labels, i.e. flaw / no-flaw, for supervised machine learning is provided using automated defect recognition (ADR) techniques applied to post-build X-ray CT scans. Training and test samples are extracted from in-situ data, in the form of a three-dimensional image chips, positive samples represent flaws whereas negative samples represent nominal build conditions. A shallow NN architecture executes the data fusion between modalities and establishes data correlation between in-situ data and ground truth labels. The modular and flexible NN architecture allows for the ingestion of one, some or all sensor modalities for sensitivity analysis. Additional modalities may seamlessly be added at a future stage.

Experimental validation is carried out using a Ti-6Al-4 V specimen. The automated defect recognition (ADR) software identified 1156 flaws yielding 32,352 positive training sample, i.e. in-situ voxels that are part of a flaw. The training data set was supplemented with 136,222 negative samples representing voxels with nominal or fully dense build condition. Although, flaw detection for AM is a highly imbalanced classification problem, with porosities well below 1% volume, a ratio of 4:1 was chosen here to keep class balance manageable. A four-fold cross validation revealed a strong correlation between the in-situ, multimodal sensor footprint and the associated label, i.e. flaw / no flaw. In fact, the correlation represents a test accuracy of 98.5% when measured as binary classification performance. This indicates the possibility for in-situ flaw detection in the future. Clearly, generalizability of the established correlation and associated binary classifier will have to be validated across multiple parts and/or builds.

The decision threshold applied to the classifier's confidence level can be adjusted to further reduce the residual false positive rate (FPR). This is especially important considering the severe class imbalance, i.e. many negative or nominal voxels for just a few positive samples or flaw voxels. Using just one layer of the build with a class imbalance of 1:380, the impact of the decision threshold was illustrated, and the inherent tradeoff between false positive versus false negative misclassification was discussed.

Finally, a sensitive study was conducted to assess the information content of each sensor modality as well as subsets of sensor modalities. For this, the NN was truncated and trained only on those modalities, and conclusion were drawn based on the resulting classification performance. The area under the ROC (AuC) curve was used in order to establish the classification performance across all possible decision thresholds. Layerwise imagery was found to contribute the highest information content. Individual scan vector modalities appear to provide little information for discriminating flaws from nominal conditions. However, once fused, the collection of scan vector modalities alone achieves very good separability. Additional sensitivity analysis could help eliminate unneeded modalities, decrease implementation cost, reduce network complexity and the potential for overfitting, shorten training time, and improve performance. It may also guide future investment in sensor research.

5. Future work

Future work will include the verification of the correlation and classification framework for different parts and/or builds. Lighting

conditions as well as acoustic and multi-spectral background noise may differ slightly between builds, thereby subjecting the classifier to previously unseen, statistical deviations within the sensor data. Overall generalizability and robustness is essential for migrating and integrating in-situ process monitoring approaches.

Complexity of the NN architecture also plays a key role when addressing potential overfitting. Highly dimensional features may often contain redundant information but nonetheless require large amounts of weights within the fully connected NN layers to be processed. Future work will improve upon the current feature extraction technique with the goals of substantial dimensionality reduction. Conceivable approaches include convolution operations, as well as multi-dimensional Fast Fourier or wavelet transformations.

Extending the sensitivity study will further help eliminate redundant sensors and guide future sensor development, and integration efforts. Improving existing sensor technology for the application at hand will provide cleaner data sets and improve the signal-to-noise ratio. Once data acquisition, available data sets, and sensor calibration protocols are deemed mature enough, a prototype will follow to demonstrate how this strategy can be used to accelerate part quality assessment and ultimately certification. Relationships to conventional post-build inspection techniques, such as CT scanning, will be established to guide the certification process. If successful, the prototype will migrate into an integrated framework that exhibits how in a production setting, additional data (sensor data and post-build CT scans) can be used to continually train/update/improve the neural network, enabling continuous improvement (similar to the continuously evolving CAPTCHA test).

Flaw mitigation and interlayer repair are beyond the scope of this paper. However, adequate strategies need to be developed in the near future. The trained neural network including the derived predictions may not only inform the system about the need for interlayer-repair but may also trigger the most promising mitigation strategy. In addition to actual flaw locations, part performance and fatigue properties may also be predicted via in-situ monitoring. Statistical correlations can be validated via, e.g., DARWIN fatigue simulation code.

Future work will also include an extension of this work towards other materials, other AM systems, including multi-laser system, and other AM techniques, such as Direct Energy Deposition.

Disclaimer

The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsement, either expressed or implied, of the Air Force Research Laboratory, the Naval Air Systems Command, or the U.S. Government.

Authorship statement

All persons who meet authorship criteria are listed as authors, and all authors certify that they have participated sufficiently in the work to take public responsibility for the content, including participation in the concept, design, analysis, writing, or revision of the manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The data analytics work was funded by PSU/ARL's Internal Science and Technology Program (ISTP). The authors would like to thank Dr. Abdalla Nassar and Mr. Griffin Jones from ARL and Dr. Jared Blecher and Mr. Ryan Overdorff from 3D Systems, Inc., for support in designing

and executing the experiments, in data acquisition, and in providing CT scanning expertise and data. This work utilizes experimental data generated under a prior effort supported by the National Center for Defense Manufacturing and Machining under the America Makes Program entitled “Understanding Stochastic Powder Bed Fusion Additive Manufacturing Flaw Formation and Impact on Fatigue” sponsored by Air Force Research Laboratory under agreement number FA8650-16-2-5700. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation thereon.

References

- [1] T. DebRoy, T. Mukherjee, J.O. Milewski, J.W. Elmer, B. Ribic, J.J. Blecher, W. Zhang, Scientific, technological and economic issues in metal printing and their solutions, *Nat. Mater.* 18 (10) (2019) 1026–1032, <https://doi.org/10.1038/s41563-019-0408-2>.
- [2] Zackary Snow, Abdalla Nassar, Edward W. Reutzel, Review of the formation and impact of flaws in powder bed fusion additive manufacturing, *Addit. Manuf.* 36 (2020), 101457.
- [3] M. Seifi, A. Salem, D. Satko, J. Shaffer, J.J. Lewandowski, Defect distribution and microstructure heterogeneity effects on fracture resistance and fatigue behavior of EBM Ti-6Al-4V, *Int. J. Fatigue* 94 (2017) 263–287, <https://doi.org/10.1016/j.ijfatigue.2016.06.001>.
- [4] T. Merdes et al., Additively Manufactured MV-22B Osprey Flight Critical Components: Production Data for Witness Coupons and Test Specimens, Pennsylvania State University Applied Research Laboratory, Technical Report #20-001, Jan. 2020.
- [5] P. Edwards, A. O’Conner, M. Ramulu, Electron beam additive manufacturing of titanium components: properties and performance, *J. Manuf. Sci. Eng.* 135 (6) (2013), 061016, <https://doi.org/10.1115/1.4025773>.
- [6] J. Beuth, N. Klingbeil, The role of process variables in laser-based direct metal solid freeform fabrication, *JOM* 53 (9) (2001) 36–39, <https://doi.org/10.1007/s11837-001-0067-y>.
- [7] Cang Zhao, Niranjan D. Parab, Xuxiao Li, Kamel Fezzaa, Wenda Tan, Anthony D. Rollett, Tao Sun, Critical instability at moving keyhole tip generates porosity in laser melting, *Science* 370 (6520) (2020) 1080–1086, <https://doi.org/10.1126/science.abd1587>.
- [8] Christopher B. Stutzman, Abdalla R. Nassar, Edward W. Reutzel, Multi-sensor investigations of optical emissions and their relations to directed energy deposition processes and quality, *Addit. Manuf.* 21 (2018) 333–339.
- [9] Christopher B. Stutzman, Wesley F. Mitchell, Abdalla R. Nassar, Optical emission sensing for laser-based additive manufacturing—what are we actually measuring? *J. Laser Appl.* 33 (2021), 012010.
- [10] Alexander J. Dunbar, Abdalla R. Nassar, Assessment of optical emission analysis for in-process monitoring of powder bed fusion additive manufacturing, *Virtual Phys. Prototyp.* 13 (2018) 14–19.
- [11] A.R. Nassar, M.A. Gundermann, E.W. Reutzel, P. Guerrier, M.H. Krame, M. J. Weldon, Formation processes for large ejecta and interactions with melt pool formation in powder bed fusion additive manufacturing, *Sci. Rep.* 9 (1) (2019) 5038, <https://doi.org/10.1038/s41598-019-41415-7>.
- [12] G. Ladewig, M. Schlick, V. Fisser, Schulze, U. Glatzel, Influence of the shielding gas flow on the removal of process by-products in the selective laser melting process, *Addit. Manuf.* 10 (2016) 1–9, <https://doi.org/10.1016/j.addma.2016.01.004>.
- [13] B. Ferrar, L. Mullen, E. Jones, R. Stamp, C.J. Sutcliffe, Gas flow effects on selective laser melting (SLM) manufacturing performance, *J. Mater. Process. Technol.* 212 (2) (2012) 355–364, <https://doi.org/10.1016/j.jmatprotec.2011.09.020>.
- [14] A. Bin Anwar, Q.C. Pham, Study of the spatter distribution on the powder bed during selective laser melting, *Addit. Manuf.* 22 (2018) 86–97, <https://doi.org/10.1016/j.addma.2018.04.036>.
- [15] R. Esmaeilizadeh, U. Ali, A. Keshavarzkermani, Y. Mahmoodkhani, E. Marzbaniad, E. Toyserkani, On the effect of spatter particles distribution on the quality of Hastelloy X parts made by laser powder-bed fusion additive manufacturing, *J. Manuf. Process.* 37 (2019) 11–20, <https://doi.org/10.1016/j.jmapro.2018.11.012>.
- [16] Z. Snow, B. Diehl, E.W. Reutzel, A. Nassar, Toward in-situ flaw detection in laser powder bed fusion additive manufacturing through layerwise imagery and machine learning, *J. Manuf. Syst.* 59 (2021) 12–26.
- [17] G.D. Goh, S.L. Sing, W.Y. Yeong, A review on machine learning in 3D printing: applications, potential, and challenges, *Artif. Intell. Rev.* 54 (2021) 63–94, <https://doi.org/10.1007/s10462-020-09876-9>.
- [18] Luke Scime, Jack 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, <https://doi.org/10.1016/j.addma.2018.11.010>.
- [19] K. Wasmer, T. Le-Quang, B. Meylan, S.A. Shevchik, In situ quality monitoring in AM using acoustic emission: a reinforcement learning approach, *J. Mater. Eng. Perform.* 28 (2) (2018) 666–672.
- [20] E.W. Reutzel, J. Petrich, C. Goert, A.R. Nassar, S. Phoha, Application of Supervised Machine Learning for Defect Detection During Metallic Powder Bed Fusion Additive Manufacturing Using High Resolution Imaging, in Elsevier/AM, Vol 21, May 2018.
- [21] J. Petrich, E.W. Reutzel, C. Goert, A.R. Nassar, S. Phoha, Machine Learning for Defect Detection for PBFAM using High Resolution Layerwise Imaging coupled with Post-Build CT Scans, 29th Annual Intern. Solid Freeform Fabrication Symp., Austin, TX, Aug. 13–15 2017.
- [22] Jacob P. Morgan, John P. Morgan, Jr., Donald J. Natale, Robert W.M. Smith, Wesley F. Mitchell, Alexander J. Dunbar, Edward W. Reutzel, Selection and Installation of High Resolution Imaging to Monitor the PBFAM Process, and Synchronization to Post-Build 3D Computed Tomography”, 29th Annual International Solid Freeform Fabrication Symposium, Austin, TX, August 13–15 2017.
- [23] Jacob P. Morgan, Data Fusion for Additive Manufacturing Process Inspection, M. Sc. Thesis, PSU/Electrical Engineering, 2019.
- [24] B. Foster, E. Reutzel, A. Nassar, B. Hall, S. Brown, C. Dickman, Optical, layerwise monitoring of powder bed fusion. In Solid Freeform Fabrication Symposium, Austin, TX, Aug, pp. 10–12. 2015.
- [25] Mostafa Abdelrahman, Edward W. Reutzel, Abdalla R. Nassar, Thomas L. Starr, Flaw detection in powder bed fusion using optical imaging, *Addit. Manuf.* 15 (2017) 1–11.
- [26] 3D Systems, ProX DMP 320, https://www.3dsystems.com/3D-printers/metal_2020.
- [27] A.R. Nassar, T.J. Spurgeon, E.W. Reutzel, Sensing defects during directed-energy additive manufacturing of metal parts using optical emissions spectroscopy. 278–287. 25th Annual International Solid Freeform Fabrication Symposium, SFF 2014, Austin, United States.
- [28] Alexander J. Dunbar, Abdalla R. Nassar, Assessment of optical emission analysis for in-process monitoring of powder bed fusion additive manufacturing, *Virtual Phys. Prototyp.* 13 (1) (2018) 14–19.
- [29] C. Suplee, Atomic Spectra Database NIST, 2009, <https://www.nist.gov/pml/atomic-spectra-database> (Accessed 19 October 2017).
- [30] NEMA PS3 / ISO 12052, Digital Imaging and Communications in Medicine (DICOM) Standard, National Electrical Manufacturers Association, Rosslyn, VA, USA (<http://medical.nema.org/>).
- [31] Laurene Fausett, *Fundamentals of Neural Networks: Architectures, Algorithms And Applications*, first ed., Pearson Education, 1994.
- [32] Michael Nielsen, *Neural Networks and Deep Learning*, Determination Press, 2015. Chapter 3.
- [33] ASTM, F2924 – 14, Standard Specification for Additive Manufacturing Titanium-6 Aluminum-4 Vanadium with Powder Bed Fusion, 02 January 2014.