In Situ Nondestructive Fatigue-Life Prediction of Additive Manufactured Parts by Establishing a Process–Defect–Property Relationship

Seyyed Hadi Seifi, Aref Yadollahi, Wenmeng Tian, Haley Doude, Vincent H. Hammond, and Linkan Bian∗

The presence of process-induced internal defects (i.e., pores, microcracks, and lack-of-fusions) significantly deteriorates the structural durability of parts fabricated by additive manufacturing. However, traditional defects characterization techniques, such as X-ray CT and ultrasonic scanning, are costly and time-consuming. There is a research gap in the nondestructive evaluation of fatigue performance directly from the process signature of laser-based additive manufacturing processes. Herein, a novel two-phase modeling methodology is proposed for fatigue life prediction based on in situ monitoring of thermal history.

Phase (I) includes a convolutional neural network designed to detect the relative size of the defects (i.e., small gas pores and large lack-of-fusions) by leveraging processed thermal images. Subsequently, a fatigue-life prediction model is trained in Phase (II) by incorporating the defect characteristics extracted from Phase (I) to evaluate the fatigue performance. Estimating defect characteristics from the in situ thermal history facilitates the fatigue predicting process.

1. Introduction

The additive manufacturing (AM) domain contains novel fabrication techniques, which are defined as the process of joining the material together in a layer-by-layer manner, to make 3D objects.[1] AM enables fabricating complex geometries where most of them are not feasible in the domain of conventional manufacturing techniques. The significant design flexibility offered by AM techniques can save a noticeable amount of money and time if applied correctly. However, the broader adoption of AM processes has faced challenging issues, especially at satisfying quality standards and process repeatability.[2] The compromised structural performance resulted from process-induced defects is the main challenge against the continued adoption of AM in different industries.[3,4]

Among different modes of mechanical failures, fatigue failure, that is, failure under cyclic loading, is the dominant failure mode in mission-critical applications.[5] This is due to the fact that fatigue is a local phenomenon; thus, it is more directly affected by the microstructural features.[3] 62% of aircraft structures have had failures due to fatigue, where only 14% of them were because of mechanical overload.[6] Meeting fatigue and durability requirements has proven to be a challenging task for AM parts.[3,4] Process and design parameters have shown a significant impact on the microstructure and defect properties of AM parts, and thus largely determine their fatigue life behavior.[1,3,5] In the absence of voids and inclusions, which typically serve as crack initiation sites, slip bands usually drive the crack initiation in material.[7] Slip length and slip planarity are the two most important factors in determining fatigue properties of titanium alloys.[8] Slip length can be controlled by microstructural features including α laths, phase boundaries and colony boundaries, while slip planarity is governed by oxygen and aluminum content as well as secondary precipitation (Ti3Al).[7] Since fatigue cracks tend to initiate at the longest existing crystallographic slip bands in the microstructure, reducing the maximum dislocation slip length is critical for enhancing the material resistance against fatigue crack initiation (i.e., HCF strength). Crystallographic orientations of the adjacent grains as well as the grain size and geometry act as a barrier for

S. H. Seiﬁ, W. Tian, L. Bian
Department of Industrial and Systems Engineering
Mississippi State University
479-2 Hardy Road, 260 McCain Hall, Mississippi State, MS 39762, USA
E-mail: bian@ise.msstate.edu
A. Yadollahi
Department of Mechanical Engineering
University of South Alabama
150 Student Services Dr, Mobile, AL 36688, USA
H. Doude
Center for Advanced Vehicular Systems
Mississippi State University
200 Research Blvd, Starkville, MS 39759, USA
V. H. Hammond
Weapons and Materials Research Directorate
US Army Research Laboratory
Aberdeen Proving Ground, Maryland 21005, USA

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/aisy.202000268.

© 2021 The Authors. Advanced Intelligent Systems published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1002/aisy.202000268
Fatigue behavior of AM metallic materials is currently dominated by process-induced defects, such as gas-containing pores and lack of fusion (LOF) defects. When the microstructure is nearly homogeneous and the residual stresses in an AM part are limited—for instance, by post-fabrication heat treatments or preheating the build plate—variations in size, location, and spacing of defects are the main reasons for the uncertainty and scatter in the AM fatigue data.\textsuperscript{[13–15]} Therefore, defect characteristics (i.e., size, shape, location, and distribution) are leading influential parameters affecting fatigue life performance.\textsuperscript{[16]} In this article, all the specimens underwent a homogenizing and stress relieving heat treatment to minimize the effects of residual stresses and microstructural variations, yet microstructural features, including the grain size, grain orientation, and misorientation angles, dictated by the manufacturing and post-manufacturing heat treatment processes, are taken into account for fatigue-life prediction. Several studies have demonstrated that one of the main issues in the fatigue-life description of AM parts is due to scattering in the high-cycle fatigue (HCF) regime, mostly resulting from the presence of process-induced defects.\textsuperscript{[3,16,17]} Thus, being able to predict the fatigue behavior of AM materials based on the defect characteristics would be an important step towards improving the reliability of AM parts.\textsuperscript{[18]} Fatigue performance of an AM part may not necessarily be dominated by the most massive defect, preferably a combination of microstructural features, including grain size and defect characteristics (i.e., size, shape, location, and distribution) typically compete together for initiating cracks.\textsuperscript{[18]} It has been shown that the crack initiating defect for AM Ti–6Al–4V is rarely simply the largest defect; rather, various aspects affect the tendency for a defect to initiate a crack, including proximity to the surface and to other pores.\textsuperscript{[19]} This indicates that the fatigue life prediction based on the largest of the defects found in multiple specimens can be conservative in some cases. As a result, without considering the defect location and spacing, it is not possible to discriminate critical defects from harmless ones. Thus, a successful in situ qualification and certification algorithm requires the ability to identify the formation of defects during the build process and link the characteristics of detected anomalies to the performance and durability of the fabricated part.

Developing proper applications for fracture mechanics-based damage tolerance principles provide a mechanism for effective risk mitigation in dealing with inherent material flaws, such as manufacturing-induced defects.\textsuperscript{[20]} In this concept, that is, crack-growth-based and defect-tolerant modeling of fatigue, each anomaly is treated as a sharp crack from the first mission cycle.\textsuperscript{[15,21]} This methodology is an appropriate tool for expressing the relationship between fatigue resistance and material quality and is widely used in aerospace industries.\textsuperscript{[20], [22–24]} One of the practical methods to predict the “fatigue” behavior of materials has been proven to be the use of the “crack propagation” concept since the mid-1980s by the Advisory Group for Aerospace Research and Development (AGARD) Structures and Materials Panel technical reports (R-732, R-767) on studying “small-crack” behavior.\textsuperscript{[25–27]} Evidently, AM process-induced defects have a different formation mechanism and a disparate impact on fracture-critical properties.\textsuperscript{[29]} In the presence of large lack of fusion (LOF) defects, crack-growth-based modeling of fatigue—that is, approximating LOFs as the preexisting sharp cracks—is a promising method for analyzing the fatigue-life.\textsuperscript{[14]} However, for gas porosity, the assumption of preexisting sharp cracks may produce too conservative results.\textsuperscript{[30,31]} In addition, for AM parts, due to the formation of a large number of defects with various sizes and locations, several defects may have the potential to serve as the crack initiation site under cyclic loading.\textsuperscript{[13]} Using crack-growth-based concepts to predict the fatigue life for multiple initiations of cracks and their coalescence, which is prevalent in AM parts, is challenging.\textsuperscript{[14]} The microstructure-based multistage fatigue (MSF) model, developed initially for cast Al alloy, directly involves the effects of microstructural features and defect characteristics, that is, defect size, nearest neighbor distance (NND), and void volume fraction in the fatigue-life prediction.\textsuperscript{[29]} MSF model characterizes fatigue life in three stages: incubation, microstructurally small crack growth, and long crack growth; based on microstructural features. MSF model used the local constrained cyclic microplasticity phenomenon to address the inclusions in fatigue crack incubation as well as microstructurally small crack growth, which also includes the impact of crystallographic orientation on crack tip displacement as the driving force.

In literature, casual relationships such as process–structure–property linkages have been established by relying mainly on observations made from cleverly designed control experiments and physics-based models. According to vision setting documents, it has come clear to the experts in material science and engineering that data science and analytics have provided a third set of distinct, powerful tools for establishing the mentioned casual relationships.\textsuperscript{[30–32]} Therefore, a data-driven methodology is proposed in this article to characterize the fatigue life from thermal history by creating the process–defect–structure relationship.

The traditional fatigue-life prediction models are using X-ray CT scans or similar post-manufacturing technologies to derive defect features such as defect size or nearest neighbor distance. The issue of nondestructive evaluation methods (X-ray CT scans) is the expense and unavailability of these machines for industrial sections. This study bridges this gap by proposing a model to predict the defect features from the process signature (i.e., thermal history). Many studies have aimed at building defect models based on morphological characteristics of a melt pool. A melt pool is an expression given to a superheated area of molten metal in proximity to the laser–material interface.\textsuperscript{[11]} Melt pools are captured as high-resolution thermal images during the build process using a powerful co-axial thermal camera. These images carry the most critical information of the AM process. Morphological characteristics of a melt pool can refer to its depth, size, temperature distribution, and other geometrical features. For example, Tan et al. used melt pool dimensions extracted from a coaxially aligned high-speed camera in a welding process.\textsuperscript{[33]} Song et al. proposed a predictive model to control the melt pool temperature for a high-power diode laser cladding process.\textsuperscript{[34]} Clijsters et al. developed an anomaly detection model based on in situ melt pool dimensions for the selective laser melting (SLM) process.
Moreover, Lane et al. proposed a model based on high-speed thermographic measurements to monitor the melt pool region (characterized by contours and melt pool depth) in a laser powder bed fusion (LPBF) process. Heigel et al. computed the melt pool length using a high-speed thermal camera in an LPBF process. Finally, Fisher et al. studied the correlation between melt pool dimensions with temperature information gathered using a visible-light camera in an LPBF process. Additionally, some studies have worked on qualitative and quantitative observation of melt pool dynamics where none of them directly worked on detecting defect formation. Many studies, similar to the current one, have incorporated powerful techniques such as statistical analysis methods, machine learning (ML), or computer vision (CV) techniques to extract higher-order and yet prominent features of melt pools instead of morphological ones. For example, a neural network model based on data collected from high-speed cameras is developed by Lou et al. to identify the relationship between the process parameters and keyhole defect formation. Off-nominal defects are detected by statistically comparing the pixel-wise trans-layer emitted light intensity profiles by Grasso et al. Khanzadeh et al. proposed a clustering-based methodology to analyze the melt pool morphologies extracted from a radial base function using the melt pool boundary captured by a thermal camera. Dongsen et al. proposed a model based on both simple statistical and convolutional neural network descriptors to discriminate between melt pools generated under three different process parameters. Finally, Scime and Beuth incorporated computer vision and unsupervised machine learning based on a scale-variant description of the melt pool to detect key-holing porosity and balling instabilities.

The aforementioned studies have mainly focused on methodologies to predict the occurrence of defects within the parts, and their limitation is that they have not characterized the magnitude of predicted defects. Having a method to predict the size of defects paves the way toward extracting defect features that play a critical role in describing the fatigue life. Therefore, we propose a methodology to predict the relative size of defects (small or large in terms of significant axis length) based on observed thermal history, that is, melt pool images.

In this article, a multiphase data-driven methodology is proposed to predict the final fatigue life based on in situ process signatures, that is, real-time thermal history. The proposed method is a two-phase process. The objective of Phase (I) methodology is to study and extract the size and location of defects by establishing the thermal–defect relationship within the additive manufacturing processes (Figure 1a,b). Traditional defect characterization techniques, such as X-ray computed tomography (CT) and ultrasonic scanning, are time-consuming and expensive. Moreover, X-ray CT scans provide inferior results for bulk specimens and cannot serve as the main tool for real-time modeling or monitoring for AM processes. Defect prediction is one of the significant areas of study in the AM community. So far, the main objective of the previous studies was predicting the defect occurrence rather than defect size characterization. To move the current state-of-the-art approaches one step forward, the relative size of the defect is predicted based on processed thermal images. After data preprocessing steps (data transformation), a convolutional neural network (CNN) is designed to discriminate the size of anomalies effectively.

The main objective of Phase (II) is to establish the defect–fatigue relationship by modeling the fatigue behavior based on defect properties and microstructural characteristics (Figure 1c); therefore, completing the process–defect–property relationship. The proposed methodology considers four parameters (i.e., defect size, nearest neighbor distance, grain size, and grain misorientation) to predict the fatigue life of Ti–6AL–4V parts. Furthermore, grain size and grain misorientation are added to the model due to their noticeable influence on fatigue life and the fact that they cannot be predicted from thermal history. The main purpose of this study is the identification of defect features from thermal history and their impact on the fatigue behavior. Therefore, the feature extraction of the microstructure

![Figure 1](https://www.advancedsciencenews.com)
is not discussed in this article. Grain features can be learned from historical experiments on Ti–6Al–4V parts. The proposed methodology provides a parametric representation of fatigue life where each fatigue curve is characterized using a set of parametric coefficients. Two regression models, stepwise linear regression (SLR) and Gaussian process regression (GPR), are trained to establish the defect–fatigue relationship by predicting the parametric fatigue coefficients from defect and grain features.

The remainder of the article is organized as follows. Section 2 reviews the background information alongside the overview of the methodology. Sections 3 and 4 provide Phase (I) characterization of establishing the thermal–defect relationship and Phase (II) methodology for determining the defect–fatigue relationship, respectively. Section 5 includes the uncertainty analysis. Section 6 discusses the modeling results and validations, and Section 7 states the concluding remarks.

2. Process–Defect–Property Methodology

2.1. Fatigue Life and Microstructure-based Fatigue Modeling (MSF)

Fatigue life is determined through a process where a material undergoes cycle-by-cycle damage of fluctuating stresses or strains. In fatigue experiments, the applied stress is not large enough to break the specimen instantly (immediate failure); instead, the material stays under a certain number of load fluctuations until a failure, that is, when the accumulated damage reaches a certain threshold.

In fatigue life prediction, the main idea is to predict the number of cycles for specific stress or strain range. In this article, the proposed fatigue model is aimed at predicting the number of cycles for a particular strain range. In the following paragraphs, microstructure-based fatigue modeling (MSF) is discussed thoroughly since the proposed model is developed based on MSF modeling.

In MSF, fatigue life is decomposed into three successive stages based on the microstructural properties of crack growth, as follows:

\[
N_{\text{TOTAL}} = N_{\text{INC}} + N_{\text{MSC}} + N_{\text{LC}}
\]

where \(N_{\text{TOTAL}}\) stands for the total number of cycles to failure, \(N_{\text{INC}}\) stands for the number of cycles to fatigue crack incubation, \(N_{\text{MSC}}\) stands for the number of cycles for microstructurally small crack propagation, and \(N_{\text{LC}}\) stands for the number of cycles for long crack propagation. The basis for the MSF was initially founded upon multiscale finite element simulations, and only microstructural data was applied toward the end of the development. The incubation stage of the MSF model is based upon a modified Coffin–Manson law. The MSC stage of the MSF model is based upon crack tip displacement (CTD). The long crack (LC) stage is based upon a modified Paris law. The governing equations in the multistage fatigue (MSF) model are listed here.

Generalized incubation model

\[
\frac{l}{D} = \eta_{\text{lim}} \left( \frac{\epsilon_{\text{th}} - \epsilon_{\text{per}}}{\epsilon_{\text{lim}} - \epsilon_{\text{th}}} \right), \quad \frac{l}{D} > \eta_{\text{lim}}
\]  

The total fatigue life (Equation (1)) includes three parts, namely incubation, microstructurally small crack (MSC) growth, and long crack (LC) growth. Equations (2) and (3) describe the ratio of the plastic zone to the particle size or defect (\(l/D\)). The ratio of the plastic zone to the particle size is described as a function of strain amplitude using \(\epsilon_{\text{th}}, \epsilon_{\text{per}}, \eta_{\text{lim}}\), and \(r\) parameters. The incubation life is modeled in Equation (4) based on a modified Coffin–Manson law using the cyclic plastic deformation (microplasticity). As in the Coffin–Manson law, \(C_{\text{INC}}\) and \(a\) are the coefficient and exponent fit to experimental data, respectively. The \(\beta\) term is a nonlocal maximum plastic shear strain amplitude located around the particle or defect. To model multiaxial fatigue, a variation of the nonlocal maximum plastic shear strain amplitude expression is presented in Equation (5) where \(\sigma_0\) is the yield stress and \(k\) are determined from experiments. The parameters \(r_f, r_s, b_0, c_0\) are torsion fatigue parameters for both strain-life and stress-life, and \(\mu\) is the shear modulus. Equation (6) describes the microstructural small crack growth regime, where \(\Delta \text{CTD}\) is the crack tip displacement range, \(\chi^*\) is a material constant, and \(\Delta \text{CTD}_{\text{th}}\) is the threshold for crack tip displacement. Equation (7) defines the crack tip displacement as a function of remote loading, where \(\chi, C_{\text{II}}, \xi, \eta, \tau, \varepsilon\) constants are determined through small crack growth experiments. The GS parameter captures the effect of grain size on the microstructure small crack behavior. Long crack growth behavior is presented in Equation (8). The effective stress intensity factor is defined as \(\Delta K_{\text{eff}} = K_{\text{max}} - K_{\text{op}}\), where \(K_{\text{max}}\) is the maximum stress intensity factor and \(K_{\text{op}}\) is the opening stress intensity factor. Finally, the faster of the small crack and long crack calculated growth rates is employed for constant amplitude fatigue life prediction (9). According to Bu and Stephens, physically-short cracks have a size less than 1.0 mm, where physically long cracks are in the range of several millimeters. Therefore, it should be noted that for the purpose of this study,
according to the fact that experiments are conducted on small laboratory specimens, the $N_{LC}$ is not considered in calculating the final fatigue life, and predicting fatigue life at the very high cycle (VHCF) is not within this work’s scope.

As it is evident, MSF uses a thorough and complex combination of mathematical and mechanical calculations to predict the fatigue life of a particular part. Additionally, the number of parameters involved in these calculations is significantly large (over 40), which includes mechanical properties, microstructure parameters, material incubation, small crack, and long crack constants. Although these parameters cover a majority of fatigue modeling aspects, it complicates the simple fatigue predicting task where it is required to input all the parameters correctly to acquire a better estimation. This study provides a specific-purpose fatigue prediction of Ti–6Al–4V material from hand-picked parameters of the MSF model, such as defect size, nearest neighbor distance, grain size, and grain misorientation. Moreover, the proposed methodology is nondestructive and directly predicts the total fatigue life $N_{total}$ for a given range of strain amplitude. It is essential to mention that a data-driven surrogate modeling procedure enables the reverse optimization of the process to derive an understanding of the thermal history and design inputs based on fatigue life.

2.2. Proposed Framework Overview

The objective of this article is to present a two-phase modeling approach to characterize the thermal–defect–fatigue relationship for nondestructive fatigue life prediction. In Phase (I), the thermal–defect relationship is established for identifying the location of defect occurrence and size characterization. The major challenge in Phase (I) is the high dimensionality and complex spatial–temporal correlation in the thermal image stream. To account for the dimensionality issue, prior to the CNN, a spherical transformation technique is used to reduce the noise and dimensionality of the thermal images. Predicting the relative size of the defect enables us to characterize the defect features to get on step closer in estimating the fatigue life, while the existing methods only predict the defect occurrence without further characterizing its size.

In Phase (II), the defect–fatigue relationship is described by leveraging the MSF model and functional regression analysis. The main challenge concerned with this phase is the lack of experimental samples. Therefore, the MSF model is leveraged to estimate a range of fatigue lives for Ti–6Al–4V. To directly predict the fatigue curve, a parametric representation of the curve is introduced using a polynomial function. The polynomial coefficients can be estimated using the defect and grain features (as shown in Figure 2). Two methods, namely stepwise linear regression (SLR) and Gaussian process regression (GPR) are used to correlate the defect and grain features to the fatigue polynomial coefficients. The fatigue curve can be reconstructed using the predicted coefficients and the known strain amplitude range. Further details of the methodologies are discussed in the following sections.

3. Phase (I) Modeling: Thermal–Defect Relationship

Phase (I) builds a model to predict the relative size of defects from in situ thermal history. The size of a defect is a continuous variable that naturally brings out the idea of a regression model. Unfortunately, the resolution of X-ray data is limited, and therefore, regression learning is a poor choice in this case. For the issue of fatigue life, the size of the defects plays a critical role in reducing the number of cycles to failure, and therefore, detecting large defects is more critical than the smaller ones. Thus, defects are divided into two groups based on a predefined threshold, creating a multiclass classification problem. This procedure is divided into three steps: 1) data transformation, 2) CNN design, and 3) defect prediction (model I), which are discussed in detail within the following sections.

3.1. Data Transformation

One way to represent in situ thermal history is by using Spatio-temporal melt pool images captured by a high-resolution thermal
camera during the build process. The melt pool images have different shapes and sizes. One practical idea in literature is to transform the images into the spherical domain where all images have identical domains. Moreover, this transformation helps to decrease the size of images slightly. Each pixel in the melt pool has three attributes, $x$-location ($x_{\text{index}}$), $y$-location ($y_{\text{index}}$), and temperatures ($z_{\text{temp}}$). The spherical transformation is transforming cartesian coordinate ($x_{\text{index}}, y_{\text{index}}, z_{\text{temp}}$) to spherical coordinate ($r, \theta, \phi$) where $r$ is the radial distance, $\theta$ is the polar angle, and $\phi$ is the azimuthal angle, using the following set of formulations

$$
r = \sqrt{x_{\text{index}}^2 + y_{\text{index}}^2 + z_{\text{temp}}^2} \quad (10)$$

$$\theta = \arccos \frac{z_{\text{temp}}}{\sqrt{x_{\text{index}}^2 + y_{\text{index}}^2 + z_{\text{temp}}^2}} = \arccos \frac{z_{\text{temp}}}{r} \quad (11)$$

$$\phi = \arctan \frac{y_{\text{index}}}{x_{\text{index}}} \quad (12)$$

Assuming that each image has a size $I_1 \times I_2$, the $x$-axis and $y$-axis of each image is in the range of $[0, I_1]$ and $[0, I_2]$, respectively. The method used by Seifi et al. does not centralize the ranges before the transformation. Using their methodology causes the melt pool images to lose valuable information at the peak of melt pools. The reason behind this fact is the different thermal distribution of cylinder thermal images compared to the thin wall. One example of the thermal image transformed by both methods is illustrated in Figure 3, where there is an evident data loss at the peak of the melt pool in Figure 3a. In the current study, the range of each melt-pool is changed to $[-\frac{I_1}{2}, \frac{I_1}{2}]$ and $[-\frac{I_2}{2}, \frac{I_2}{2}]$ for the $x$-axis and $y$-axis, respectively. After data transformation, a nonparametric surface interpolation is used to create continuous data out of discrete data points. The interpolation function helps to extract a smaller grid of information and smooth the data to remove the thermal noise, especially at lower temperatures.

### 3.2. Convolutional Neural Network Design

Melt pool images are a set of high dimensional captures carrying the essential information about the AM process. This information is not visible to the naked eye or the simple analytical tools. Therefore, a higher-order technique is required to extract the quality relevant information from melt pool images and deal with the curse of dimensionality. The primary information that a melt pool image carries is the thermal information in the form of temperatures. Thus, melt pool images are considered single-channel images, meaning that they are similar to grayscale images. A grayscale image contains the magnitude of each datapoint with a value between 0 (black) and 1 (white). For CNN, melt pool images are converted into grayscale images using max–min global standardization.

A convolutional neural network (CNN) is a type of deep neural network with the frequent use of analyzing visual imagery that is mainly used for image and video recognition, image

![Figure 3. The comparison of the two spherical methods: a) the method used by Seifi et al., b) the method used in the current study.][2]
classification, etc. CNN uses convolution, a specialized kind of linear operation, instead of the general matrix multiplication in at least one of the layers of the network.\(^{[53]}\) The CNN consists of several layers to learn the output from input images and prevent overfitting. The main layers are the input layer, output layer, and hidden layers. The hidden layers consist of the convolution layer, pooling layer, activation function layer, normalization layer, fully connected layer, softmax layer, and a classification layer. These concepts are briefly explained in the following sections.

### 3.2.1. Convolutional Layer

A convolutional layer abstracts the input image tensor to a feature map by applying convolutional filters to the input images. The output shape depends on several convolutional layer parameters, such as filter size, stride (step size), and padding.\(^{[54]}\) The output width and height of the convolutional layer is

\[
\text{output size} = (\text{input size} - \text{filter size} + 2 \times \text{padding}) / \text{stride} + 1
\]

The output of the convolutional layer is passed to the next layer.

### 3.2.2. Pooling Layer

A pooling layer aims at reducing the spatial size of the feature map and removing unnecessary spatial information by the down-sampling operation. Ordinary pooling operation is class max-pooling that returns the maximum of the rectangular region of input based on pooling size.\(^{[55]}\) In the case of overlapping regions, the output of the pooling layer is

\[
\text{output size} = (\text{input size} - \text{pool size} + 2 \times \text{padding}) / \text{stride} + 1
\]

### 3.2.3. Activation Function Layer

The activation function is a linear or nonlinear function that maps the set of inputs to a node to an output. The most common activation function is called rectified linear unit (ReLU) layer that performs a threshold operation, as follows\(^{[58]}\)

\[
f(x) = \begin{cases} 
  x, & x > 0 \\
  0, & x < 0 
\end{cases}
\]

This threshold operation does not change the size of the input.

### 3.2.4. Batch Normalization Layer

A batch normalization layer is used to normalize the input based on the mean and variance of a mini-batch.\(^{[59]}\) The normalized values are

\[
x_i = \frac{x_i - \mu_B}{\sqrt{\sigma^2_B + \epsilon}}
\]

where \(x_i\) is the input variable, \(\mu_B\) and \(\sigma^2_B\) are the mean and the variance of the mini-batch, and \(\epsilon\) is the property epsilon that improves the numerical stability in case that the variance is near zero.

### 3.2.5. Fully Connected Layer

A fully connected layer connects all of its neurons to the neurons in the preceding layer, similar to a fully connected neural network.

### 3.2.6. Softmax Layer

A softmax layer applies a softmax function to the input to get a number between 0 and 1, wherein the case of classification can be interpreted as the probability of class 1 for the corresponding input. The standard (unit) softmax function is

\[
S_j(x) = \frac{\exp x_j}{\sum_{j=1}^K \exp x_j}
\]

where \(K\) is the number of classes, \(0 \leq S_j(x) \leq 1\) and \(\sum_{j=1}^K S_j = 1\). The softmax function is usually considered as the multiclass generalization of the sigmoid function.\(^{[60]}\)

### 3.2.7. Classification Layer

A classification layer calculates the cross-entropy loss function for a multiclass classification problem where classes are mutually exclusive. The classification layer uses the output of the softmax layer to assign each input to one of the \(K\) mutually exclusive categories. The cross-entropy function is\(^{[60]}\)

\[
\text{loss} = -\sum_{i=1}^N \sum_{j=1}^K t_{ij} \ln S_{ij}
\]

where \(N\) is the total number of samples, \(K\) is the total number of classes, \(t_{ij}\) is an indicator variable that shows the true value if \(i\)th sample belongs to the \(j\)th class, and \(S_{ij}\) is the output of the softmax layer for sample \(i\) and class \(j\).

### 3.3. Model (I): Defect Prediction

There are many ways to design a CNN, and it is the designer’s choice on how to design the layers and assign the parameters. In this article, a simple design of CNN is used that provides reasonable accuracy. The objective of CNN is to learn the defective behavior from processed melt pool images. In this section, at first, the design of the CNN used in this article is provided, and afterward, the possible training options are mentioned. By using the different layers defined in the previous section, the design mentioned in Table 1 is used in this article. The details of the CNN design are mentioned in the modeling results section.

The next step is training the network based on the available samples. Before the training, several options are required to be defined, such as the solver for training the network, initial
learning rate, the maximum number of epochs, and validation procedure. The final output of the CNN is the label for each image, whether the image is healthy, small anomaly, or large anomaly. This classification provides the information required to build the fatigue prediction model.

4. Phase (II) Modeling: Defect–Fatigue Relationship

Phase I helped to identify the defect distribution within an AM part using thermal history. In Phase II, the objective is to extract the defect features from predicted defect distribution and use it to predict the fatigue life of the part. This phase is divided into four steps: 1) extracting defect characteristics, 2) fatigue curve representation, and 3) fatigue life prediction (models II and III). These topics are elaborated upon in the following sections.

4.1. Extracting Defect Related Features

Defects are voids within a manufactured part that seriously affect its mechanical behavior. To present defects’ intensity and spatial distribution, two features, defect size, and nearest neighbor distance, are utilized. These features are further used to predict fatigue life.

4.1.1. Defect Size

The defect size is the first feature demonstrating the size of the largest defect in the material at the scale of a micrometer. Several works have shown that the main life-limiting failure mechanism of AM parts is due to cracks initiating from the defects located near the surface. Therefore, the most dominating defect (the largest/closest to the surface) determines the HCF life, where a single defect leads to a crack that propagates until failure. Thus, the information of the largest defect (mostly lack-of-fusion defects) is crucial in characterizing the long-life regimes (i.e., HCF), and the effect of smaller effects can be negligible. This variable is mentioned by the name $p_{size}$ throughout the article. It is highly probable for the crack to initiate at the location of the largest defect; that is the reason one of the features is allocated to the size of the largest defect.

**Calculation:** The output of the defect prediction model (model (I)) is a discrete value, such as small or large defects. Therefore, the exact magnitude of the largest defect is unavailable. The approximate approach to determine the $p_{size}$ is to choose the average size of the large defects seen in the current samples as the maximum size of the defect in the material (400 μm). In case that there is no large defect, the mean value of the size of the small defects is sufficient (100 μm). This approach is identical to choosing an upper bound for $p_{size}$ which leads to a lower bound for fatigue life. Future research can focus on how to accurately predict the defect size instead of small or large classification.

4.1.2. Nearest Neighbor Distance

For short-life regimes (LCF), large defects may not dominate the fatigue life of the parts. Instead, a combination of defect features (i.e., size, shape, location, and distribution of all types of defects), and microstructural characteristics may compete for initiating cracks. Nearest neighbor distance (NND) is a feature to characterize the spatial distribution of the defects. This feature is mentioned by the name $p_{nnd}$ throughout the article. The more defects are located near each other, the less value of the NND, leading to poor fatigue life. Additionally, the distance of a defect from the surface plays a critical role in measuring NND. A near-surface defect is more likely to reduce fatigue life; hence, decreasing the value of NND. The approximate approach to calculate NND involves defining a neighboring criterion. In this article, the defects, within the $4D_s$, radius of defect $i$, are assumed to be neighbors of defect $i$. Having neighboring criteria in hand, neighbors of each defect can be determined ($N_j$). By measuring the distance between each defect and its neighbors ($DD_j; j \in N_i$), the average neighboring distance between defects $\bar{DD}_i$ can be calculated. Similarly, the distance to the nearest surface of each defect ($DS_i$) is calculated as well. The minimum value for both $\min_{NN} DS_i$ and $\min_{NN} \bar{DD}_i$ is calculated, and their minimum is $\min(\bar{DD}_{\min}, DS_{\min})$ is set to be the approximate value of $p_{nnd}$.
Calculation: As mentioned for \( p_{\text{size}} \), the output of the model I defines each defect as small or large. Other than defect features, two microstructural characteristics are used in this article to account for the grain effect of fatigue life prediction. Grain size \( (g_{\text{size}}) \) and grain misorientation \( (g_{\text{mis}}) \) are considered as design parameters and can be determined from historical experiments. Having these four features in hand, the fatigue life prediction model can be developed.

### 4.2. Fatigue Curve Representation

Predicting a curve from multiple parameters is a challenging task. One effective approach is to represent a curve with a parametric function. Predicting coefficients of a parametric function is a much more effective approach where the number of parameters is much smaller than the number of data points within the fatigue curve. Additionally, the number of cycles \( (N_{\text{total}}) \) for different strains in a fatigue curve are scattered between 1 and \( 10^4 \). Fitting a parametric function to this data is extremely inefficient and has a large fitting error. To solve this issue, the natural logarithm of the \( N_{\text{total}} \) is used, which makes it suitable for parametric modeling.

In this article, a family of polynomial functions is used to fit the fatigue curves as follows

\[
f(\varepsilon_q) = a_0 + a_1\varepsilon^1 + \ldots + a_Q\varepsilon^Q + \varepsilon = \sum_{q=0}^{Q} a_q \times \varepsilon^q + \varepsilon \quad (19)
\]

where \( \varepsilon_q \) is the given strain range, \( \varepsilon \) is the error term, \( a_q \) is the coefficient of the polynomial term \( \varepsilon^q \), \( Q \) is the total degree of polynomial function, and \( f(\varepsilon_q) \) is the function of predicting fatigue life, which is equal to the logarithm of \( N_{\text{total}} \). The degree of the polynomial vastly affects the output of parametric modeling. Choosing a small \( Q \) creates a function lacking the power to represent the fatigue curve truly, and for a large \( Q \), the function becomes unnecessarily complex and may overfit the curve. Thus, the best value for \( Q \) is the minimum value that satisfies a high adjusted R-squared and aptly correlated the functional fatigue parameters to the input features. After representing the fatigue curve with functional parameters, the next step is to build a model to predict them from defect features and grain characteristics.

### 4.3. Fatigue Life Prediction Models

In this section, two models are represented to predict the fatigue curve parameters, introduced in the previous section, from defect features and grain characteristics. The input to both models is from feature space, and the output is from the polynomial coefficient space (fatigue curve parameters). The first model is a simple stepwise linear regression (SLR) model where efficient (the criterion is discussed later in the article) features from feature space are added to the model in a step-by-step manner. The second model is a probabilistic model named Gaussian process regression (GPR), where all the features from feature space are considered for training the model. First, the feature space, both input, and output, is defined, and later, both models are discussed in detail.

The main variables in feature space are defect size \( (p_{\text{size}}) \), nearest neighbor distance \( (p_{\text{nn}}) \), grain size \( (g_{\text{size}}) \), and grain misorientation \( (g_{\text{mis}}) \). The input feature space \( (X) \) contains the solo, interactive, and quadratic effect of features, as follows:

1. **Main effect**: \( p_{\text{size}} \times p_{\text{nn}} \times g_{\text{size}} \times g_{\text{mis}} \)
2. **Interactive effect**: \( p_{\text{size}} \times p_{\text{nn}} \times g_{\text{size}} \times g_{\text{mis}} + \ldots \times g_{\text{size}} \times g_{\text{mis}} \)
3. **Quadratic effect**: \( p_{\text{size}}^2 \times p_{\text{nn}}^2 \times g_{\text{size}}^2 \times g_{\text{mis}}^2 \)

The output space contains the polynomial coefficients \( \{a_1, \ldots, a_Q\} \) extracted from Section 4.2. The general formulation of this methodology is represented as the following equations

\[
a_q = G_q(X) + \varepsilon \quad \forall q \in 0, 1, \ldots, Q
\]

where \( a_q \) is the \( q \)-th polynomial coefficient, \( X \) is the variables from feature space (main, interactive, and quadratic effect), \( \varepsilon \) is the error, and \( G_q \) is the function that models the relationship between the feature space and the \( q \)-th order polynomial coefficient. Overall, it is required \( Q \) functions to finalize the model. \( G_q \) can be either SLR or GPR. The following two sections will discuss the details of these two models to predict the polynomial coefficients from input space features.

#### 4.3.1. Model (II.1): Stepwise Linear Regression (SLR)

Stepwise linear regression (SLR) is a model that automatically adds or removes the explanatory variables based on a predefined criterion. The criterion can be chosen from multiple choices such as a change in the sum of squared error, an increase in the value of R-squared, or an increase in the value of adjusted R-squared. The criterion chosen in this article is adjusted R-squared with a determined threshold. If the increase in adjusted R-squared is greater than the threshold (1%), the candidate variable is added to the model. The step-by-step algorithm is shown in Figure 4. For each polynomial coefficient \( \{a_1, \ldots, a_Q\} \), one SLR is trained based on steps defined in Figure 4. Final regression models are examined in the experimental results section.

#### 4.3.2. Modell (II.2): Gaussian Process Regression (GPR)

Linear regression used in the previous section offers models with virtues like the simplicity of implementation and interpretation. The main drawback is that it has limited flexibility where the relationship between input and output variables cannot be approximated as a linear function. Gaussian process regression (GPR) is a powerful nonparametric probabilistic model that offers flexibility to model nonlinear relationships. Assume that the training set contains the data in the form of \( \{(x_i, y_i); i = 1, 2, \ldots, n\} \) where \( x_i \in \mathcal{R}^d \) is from feature space and \( a_q \in \mathcal{R}^j \) is one of the polynomial coefficients.

A GPR explains the response using latent variables \( l(x_i); i = 1, \ldots, n \), from a Gaussian process (GP), and basis functions \( h \). The covariance function between the latent variables captures the smoothness of the response variable, and the basis function \( h \) projects the input variables \( x_i \) into a \( p \)-dimensional space. A Gaussian process (GP) is a set of random variables where every finite set of them have joint Gaussian distribution.

If \( f(x) \) is a GP, for finite observation set \( \{x_1, x_2, \ldots, x_n; x_i \in \mathcal{R}^d\} \),
random variables \( l(x_1), l(x_2), \ldots, l(x_n) \) have joint Gaussian distribution. Every GP is determined by its mean \( m(x) \) and covariance function \( k(x, x') \). Therefore, if \( l(x) \) is a GP, then \( E[l(x)] = m(x) \) and \( \text{Cov}[l(x), l(x')] = E[(l(x) - m(x))(l(x') - m(x'))] = k(x, x') \). Now the GPR model can be represented by \( h(x) = \beta + l(x) \) where \( l(x) \) are from a GP with mean zero and covariance \( k(x, x') \), \( h(x) \) are the set of basis functions that transform every observation \( x \) from original space \( R^d \) into a new feature vector space \( R^p \) and \( \beta \) is a \( 1 \times p \) vector of coefficients for basis function. This model represents a GPR, and one instance of response can be modeled as

\[
P(a_i | l(x_i), x_{ij}) \sim N(a_i | h(x_i)^T \beta + l(x_i), \sigma^2)
\]

which demonstrates the probabilistic nature of GPR. More details of GPR can be found in a book written by Williams and Rasmussen.\(^{[63]}\) All the hyperparameters of the GPR model are optimized to minimize the cross-validation loss. These hyperparameters are basis function, Kernel function, sigma (noise standard deviation), and standardization. A limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) algorithm is used to optimize the aforementioned model parameters. More details for L-BFGS can be found in Chapter 7 of the numerical optimization book.\(^{[68]}\)

### 5. Uncertainty Analysis

In this section, the goal is to model the effect of Phase (I) error on the output of Phase (II), fatigue life. In other words, how would misclassification of Phase (I) classifier affect the defect-related features, and how much would it change the final fatigue life.

Assume that the number of actual large and small defects are shown with \( N^L \) and \( N^S \), respectively. \( P_i^L \) is the probability that large defect \( i \) is classified as large. Similarly, \( P_i^S \) is the probability that small defect \( i \) is classified as small. If there exists a large defect inside the part, the presence of small defects does not affect the \( p_{size} \). The probability that neither of the large defects is classified as large is \( \prod_{i=1}^{N^L} (1 - P_i^L) \). For the proposed defect prediction model, if this probability is near zero, it means that the model is classifying at least one large defect correctly. In case that there is no large defect within the material, a similar approach can be applied with small defect probabilities \( P_i^S \). Moreover, at the presence of the condition \( \prod_{i=1}^{N^L} (1 - P_i^L) \approx 0 \), the calculation of \( p_{size} \) is robust to misclassifications. The values assigned for \( p_{size} \) along with their probabilities are represented in Table 2.

A similar approach is applied for \( p_{nnnd} \) using the same notation. Two tables are required to calculate \( p_{nnnd} \): the first one is for distance to the surface of the defects, and the second one is for the clustered average distance between neighboring defects. The probabilities are calculated separately for both tables. Assume that \( P_i^L = P_i^L \times P_i^S \) is the probability of predicting defect \( i \). Comparing the value of \( P_i^L \) with \( P_i^S \) defines the diameter of the defect for neighboring criterion \( P_i^L > P_i^S \rightarrow D_{i} = 400 \). Table 3 is created for distance to the surface of every predicted defect where defects are sorted based on their \( D_{S_i} \) in an ascending order.

If defect \( i \) is not detected from thermal analysis, the algorithm moves to the second defect with the smallest distance to the surface. In this case, the \( p_{nnnd} \) increases by \( D_{S_{i+1}} - D_{S_i} \), without considering the table for the average neighboring distance between defects. Table 4 is built for the average neighboring distance between defects where they are sorted by \( D D_{i} \) in an ascending order. It is important to mention that only defects with at least

| \( p_{size} \) | Probability |
|-----------------|-------------|
| 400             | 1 - \( \prod_{i=1}^{N^L} (1 - P_i^L) \) |
| 100             | \( \prod_{i=1}^{N^L} (1 - P_i^L) \times (1 - \prod_{i=1}^{N^S} (1 - P_i^S)) \) |

| \( p_{nnnd} \) | Probability |
|-----------------|-------------|
| \( D_{S_1} \)   | \( P_1 \)    |
| \( D_{S_2} \)   | \( P_2 (1 - P_1) \) |
| \( D_{S_3} \)   | \( P_3 (1 - P_2)(1 - P_1) \) |
| \vdots          | \vdots       |
| \( D_{S_i} \)   | \( P_i \prod_{1<i}^{i-1} (1 - P_i) \) |

| \( D_{nnnd} \) | Probability |
|-----------------|-------------|
| \( D D_{i} \)   | \( \prod_{i=1}^{N^L} (1 - P_i^L) \) |
where the algorithm of NND, at least 5 of its neighbors should be identified. Thus, the overall probability of assigning the value $\mathbf{D}_{i}$ to $p_{\text{nn}}$ is $P_{i} \prod_{j=1}^{k}(1 - P_{j})$. The objective of this section is to examine the performance of the proposed multiphase methodology to see if the defect prediction based on thermal history can be used to predict fatigue life accurately. The validation procedure is demonstrated in Figure 5, where rounded blue rectangles are the proposed methodology, and the red boxes are the traditional procedure. According to Figure 5, two separate validation procedures are used to examine the performance of the proposed method. Following are the description of the datasets used for the validation.

### 6. Modeling Results and Validation

The objective of this section is to examine the performance of the proposed multiphase methodology to see if the defect prediction based on thermal history can be used to predict fatigue life accurately. The validation procedure is demonstrated in Figure 5, where rounded blue rectangles are the proposed methodology, and the red boxes are the traditional procedure. According to Figure 5, two separate validation procedures are used to examine the performance of the proposed method. Following are the description of the datasets used for the validation.

#### 6.1. Thermal Images

Thermal images are collected from a direct laser deposition (DLD) process where an OPTOMEC LENS 750 machine (with a 1 kW Nd: YAG laser) was used to fabricate Ti–6Al–4V cylinders, and the machine is equipped with a thermal imaging camera, one co-axial pyrometer camera to capture melt pools (Stratronics, Inc.), and one infrared camera to capture the global heat flow (Sierra-Olympic Technologies, Inc. Viento320). Seifi et al. have provided the LENS machine and camera setup details. This dataset is a collection of thermal images captured by a co-axial pyrometer camera, carrying the main information about the build process. The laser has used a unidirectional scan finishing contour path. There is a hatch rotation of 180° between consecutive layers. The experimental setup of the laser machine is as follows: 1) Laser power: 300 (W). 2) Laser travel speed: 40 (rpm). 3) Powder flow rate: 3 (rpm). 4) Layer thickness: 0.015 (in). 5) Hatch distance: 0.02 (in).

The captured thermal images are projected into a 752 × 480 size frame and further cropped to a smaller frame with more concentration on the peak of melt pools. The cylinder is finished by depositing 69 layers on top of each other, where the thermal camera collected 2827 none-noise images of the fabrication process.

![Figure 5](image-url)
6.2. Defect Data

The fabricated cylinder is scanned using a Nikon X-ray CT XT H225 (https://www.nikonmetrology.com/en-us/x-ray-ct/x-ray-ct-systems-225kv-and-320kv-ct-inspection-and-metrology) machine with a maximum voltage of 225 kV. This machine is suitable for inspecting small to medium parts. The voltage and the current of the X-ray machine are set to 180 kV and 98 μA to scan the cylinder. The scanned X-ray file is processed using the myVGL application, and defect information is extracted in an excel sheet. The primary columns are: 1) Probability: a software assigned value to emphasize the accuracy of a detected void to be pore. 2) Radius (mm): The approximate radius of the detected pore. 3) Diameter (mm): The approximate diameter of the detected pore. 4) Center x (mm): The coordinate value of the center of pore in the x-direction. 5) Center y (mm): The coordinate value of the center of pore in the y-direction. 6) Center z (mm): The coordinate value of the center of pore in the z-direction. 7) Voxel: The number of data points in the 3D grid.

The probability column is used to remove the less significant defects. The diameter represents the size of the defects, and center x-y-z data demonstrates the location of the defects and helps to match them with thermal images. The average and standard deviation of all the defects are 150 and 120 μm. The histogram of the size of defects is shown in Figure 6. This figure contains all the defects, regardless of their probabilities. In this article, only high probability defects are considered in the modeling. While most defects are small and large defects rarely happen during the build, the threshold is set to 200 μm to distinguish between frequent small defects and rare large defects.

6.3. Fatigue Experiments

Fatigue experiments are highly costly and time-consuming; therefore, the MSF model is calibrated based on experimental data on Ti–6Al–4V. Afterward, the MSF model is used to estimate a range of possible fatigue lives based on different combinations of defect size, NND, grain size, and grain misorientation. These experiments are used to validate the defect–fatigue model (Phase (II)).

6.4. Phase (I) Validation: Thermal–Defect Model

6.4.1. Dataset Description

Thermal history carries the primary information about the process signature, and here, it is used to predict the relative size of defects. Thermal history is a set of high-resolution images of melt pools captured by a thermal camera, called the pyrometer camera. These melt pool images have different shapes and sizes, which makes the comparison of melt pool images extra hard and less meaningful. To amend this issue, melt pool images are transformed from a cartesian to a spherical coordinate system. The pyrometer camera captures the images in a frame with size 752 × 480, which can be cropped to smaller images with more concentration on the peak of melt pools. The current cylinder dataset contains 69 layers of 2827 melt pool images, each with a size of 80 × 80 pixels. The location of melt pools is known with three parameters, x-location, y-location, and layer number. Out of the 2827 melt pools, 2634 are flagged normal, 110 are flagged with a small defect, and 83 are flagged with a large defect. The threshold to divide defects into small and large is 200 μm, comparing the major axis length. All the transformed melt pools are converted into single-channel (grayscale) images.

Table 5. The detailed design of the convolutional neural network.

| #  | Layer name                  | Details                                                                 |
|----|----------------------------|-------------------------------------------------------------------------|
| 1  | Input image layer           | 25 × 25 × 1                                                             |
| 2  | Convolutional layer 1       | Filter size: 3 × 3<br>The number of filters: 8<br>Stride: 1 × 1          |
| 3  | Batch normalization layer   |                                                                         |
| 4  | ReLU layer                  |                                                                         |
| 5  | Max-pooling layer 1         | Pool size: 2 × 2<br>Stride: 2 × 2                                      |
| 6  | Convolutional layer 2       | Filter size: 3 × 3<br>The number of filters: 16<br>Stride: 1 × 1        |
| 7  | Batch normalization layer   |                                                                         |
| 8  | ReLU layer                  |                                                                         |
| 9  | Max-pooling layer 2         | Pool size: 2 × 2<br>Stride: 2 × 2                                      |
| 10 | Convolutional layer 3       | Filter size: 3 × 3<br>The number of filters: 32<br>Stride: 1 × 1        |
| 11 | Batch normalization layer   |                                                                         |
| 12 | ReLU layer                  |                                                                         |
| 13 | Fully connected layer 1     | Output size: 6                                                          |
| 14 | Fully connected layer 2     | Output size: 3                                                          |
| 15 | Softmax layer               |                                                                         |
| 16 | Classification layer        |                                                                         |

Figure 6. The histogram of the size of pores extract from the myVGL application.
6.4.2. Model (I): Defect Prediction by CNN

The specific architecture of CNN is mentioned in Table 5. The model is using a solver named stochastic gradient descent with momentum (SGDM) to train the network. The initial learning rate is 0.01, the maximum number of epochs is ten, and the validation frequency is every ten iterations. To prevent CNN from overfitting, 20% of the training dataset is used to validate the network based on predefined validation frequency. Due to the unbalanced number of samples in classes, a stratified data split approach is incorporated to preserve the percentage of samples for each category in sub-sets.

Sixfold cross-validation is used to examine the performance of the proposed model by using metrics such as precision, recall, and f-score that are defined as follows:

\[
\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \tag{24}
\]

\[
\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \tag{25}
\]

\[
F - \text{Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{26}
\]

True or false terms illustrated the correct and incorrect prediction of the classification algorithm, respectively; additionally, positive and negative terms represent the objective class of the algorithm at the time (healthy, small defect, or large defect) and the opposing classes, respectively. The averaged confusion matrix alongside averaged accuracy metrics is reported in Table 6.

Regarding the performance of the CNN in the test set, out of 13 large defects in the fabricated part, 2 (15%) are misclassified as healthy (nonexistent), and 2 (15%) are misclassified as a small defect. Additionally, out of 18 small defects in the test set, 2 (11%) are misclassified as large, and 3 (16%) are misclassified as healthy (nonexistent). These numbers are trivial for healthy samples, but the sample size is larger as well. The averaged f-score for healthy, small defect, and large defect classes are 99%, 74%, and 72%, respectively. Overall, considering the small number of defect samples, the accuracy of the model is acceptable. The uncertainty analysis is used to account for the fact that the model can misclassify the images, especially the large ones.

To further investigate the accuracy of the model (I), the receiver operating characteristic (ROC) curve of the classification model is plotted. The ROC curve is a graphical measure that demonstrates the diagnostic ability of a binary classifier by varying the discrimination threshold. The ROC curve can be created by drawing the true positive rate (TPR) against the false positive rate (FPR) for different threshold values. TPR is known as the recall factor, introduced earlier, and FPR is known as the probability of false alarm (type I error).[70] The ROC graph depicts the tradeoffs between TPR (benefits) and FPR (costs). The best possible classification model would result in a ROC curve with a point in the upper left corner (100% recall).

Additionally, a random guess classifier would yield a diagonal line from the bottom left to the top right corner, which is called the “no-discrimination” or “chance” line. The ROC curve of a fair classification model should lie between the random and the best possible curves. Moreover, the area under the curve (AUC) is a measure that calculates the area under the ROC curve. AUC demonstrates the probability of ranking a randomly selected positive instance higher than a randomly selected negative example by the classifier. Greater AUC means better average performance.[70] The ROC response of sixfold cross-validation is shown in Figure 7. The AUC of different folds for all of the classes is near one.

6.5. Phase (II) Validation: Fatigue Life Prediction

In this section, the fatigue prediction models defined earlier are implemented using generated experiments. Prediction accuracy, as well as model performances, are analyzed. In the following section, experiment generation is discussed, and both proposed models are implemented along with their results.

6.5.1. Experiment Generation

Even though fatigue experiments are extremely expensive and time-consuming, it is highly unlikely to fabricate specimens with similar defects and microstructure properties regarding the AM process’s uncertain nature. Additionally, there are not enough fatigue specimens in hand from an additive manufacturing process to train a reliable prediction model. This uncertain nature of AM processes and lack of sufficient samples are the main reasons behind incorporating the MSF model as the backbone of this study. The MSF model is calibrated based on experimental data on Ti–6Al–4V, and different combinations of defect size, NND, grain size, and grain misorientation are used to estimate a range for the fatigue life of Ti–6Al–4V material. For this purpose, the acceptable range of these features is defined based on the defect statistics determined through fractography and microstructural examination of the fatigue specimens. These ranges are shown in Table 7.

Figure 8 illustrates the maximum change in fatigue life for each feature while keeping the other three features at their mean value. After defining the features range, a full factorial design of

| Feature | Range | Unit |
|---------|-------|------|
| $p_{\text{ax}}$ | [5 – 750] | Microns |
| $p_{\text{rad}}$ | [5 – 400] | Microns |
| $g_{\text{ax}}$ | [5 – 200] | Microns |
| $g_{\text{m}}$ | [1 – 30] | Degree |

Table 7. The acceptable range of defect and microstructural feature.
the experiment is incorporated to model the feature space effectively. For the full factorial design, three levels for each feature are chosen, leading to $3^k$ training samples. Each training sample is generated using the microstructure-based multistage fatigue (MSF) model developed by McDowell et al. [29]. The remaining parameters of the MSF model are kept at a constant level, specifically for the Ti–6Al–4V parts.

In addition to the training samples aforementioned, a testing sample set is required to examine the performance of the proposed methodologies. While a full factorial design is used to generate training samples, a near-random method, named Latin hypercube sampling (LHS), is used to generate the testing samples. The LHS was described by researchers from Los Alamos National Laboratory in 1979. [71] LHS uses the idea behind the Latin square and generalizes it to an arbitrary number of dimensions. A Latin square is a $n \times n$ matrix containing $n$ distinct values where each value occurs exactly once in each row and each column.

Assume there exist $N$ variables, and the range of each is divided into $M$ equally probable intervals. Then, $M$ samples points while satisfying the Latin hypercube requirement, where each specific row and column only contains one sample. This requirement forces each variable to have the same number of divisions, $M$. One of the main advantages of the LHS is that

---

**Figure 7.** The sixfold cross-validation of ROC response for a) healthy label, b) small defect label, and c) large defect label for CNN classifier.

---

**Figure 8.** The maximum change in fatigue life for a) defect size, b) nearest neighbor distance, c) grain size, and d) grain misorientation.
the sampling scheme does not require more samples for more dimensions (variables). For the testing set, each variable range is divided into 40 intervals leading to 40 samples.

6.5.2. Fatigue Models

Each fatigue experiment is represented with a polynomial function. The degree of the polynomial function is a hyper-parameter and should be chosen based on the training set. Therefore, a 10-fold cross-validation method is incorporated using the training set to select the degree of the polynomial for fatigue curve modeling comparing three measures. The first measure is the root mean square error (RMSE) for predicting the polynomial coefficients.

The second measure is called the error percentage. Error percentage is derived by dividing the magnitude of error in fatigue life prediction to the maximum range of fatigue life within the training dataset. Maximum fatigue life is achieved when $p_{\text{size}}$, $g_{\text{size}}$, and $g_{\text{mod}}$ are at their minimum, and $p_{\text{end}}$ is at its maximum. Minimum fatigue life is achieved oppositely. Each fatigue curve is created from 101 data points where for each strain value, fatigue life (number of cycles) is recorded. At each data point, the error percentage is calculated, and afterward, the average error percentage for each sample is recorded.

The third measure is called the dissimilarity score. The dissimilarity score calculates the resemblance of the predicted fatigue curve to the actual one, and it is defined as the standard deviation of error percentage at 101 data points of a fatigue curve. Both error percentages and dissimilarity scores should be near zero. According to Figure 9, $Q = 4$ is the optimum degree of the polynomial function to represent each fatigue curve. Subsequently, the accuracy of the model for predicting the polynomial coefficient and fatigue curves is examined using the testing set. First, the root-mean-square error (RMSE) of both regression models (SLR and GPR) for predicting the polynomial coefficients are represented. Second, the predicted coefficients are used to construct the fatigue curves based on the known strain range.

Before using SLR or GPR, both input features and output coefficients are normalized to compare performance measures. The RMSE resulting from predicting output coefficients is illustrated in Table 8, where SLR performs slightly better.

After predicting the coefficients, the fatigue curves are constructed and compared with the original curves. One sample output, along with the boxplot of the error-percentage for all 40 testing samples, is shown in Figure 10. According to Figure 9, 2 out of 40 for SLR and 3 out of 40 samples for GPR have an error percentage near 20%, and the remaining samples have less than 10% error in prediction. Overall, GPR has a slightly better performance compared to the SLR due to low error variance.

6.6. Numerical Analysis of the Uncertainty

In this section, a numerical analysis of the uncertainty in defect prediction and its effect on fatigue life is provided. For this purpose, the dataset is randomly divided into train, validation, and test sets, and the prediction probabilities of the samples in the test set are calculated. The test set has $N_S = 16$ small defects and $N_L = 13$ large defects. The probability table for $p_{\text{size}}$ is shown

![Figure 9. Illustration of the results of three measures with different degrees of the polynomial function.](image-url)
The probability of classifying at least one defect is 1, meaning that the model (I) can classify at least one large defect and therefore, the estimated $\hat{\beta}_{\text{size}}$ is 400 μm. Additionally, during the condition $\prod_{i=0}^{N}(1 - p_i^2) \approx 0$ holds, the model is robust to misclassification of the small defects.

The probability table for $p_{\text{nnd}}$ based defects’ distance to the surface is provided in Table 10. The probability table for defects’ distance to distance is not delivered due to the scatteredness of the defects. Therefore, $D_D$ is insignificant compared to $D_S$, and Table 10 can be used for an approximation of $p_{\text{nnd}}$. According to Table 10, while the probability of the first defect is less than 0.5, the second defect is used for NND calculation. Thus, the estimated $\hat{p}_{\text{nnd}}$ is equal to 26 μm.

The actual values of $p_{\text{size}}$ and $p_{\text{nnd}}$ are 477 and 21 μm, respectively, for the test set. The changes in these two variables are $\Delta_{\text{size}} = \hat{p}_{\text{size}} - p_{\text{size}} = 400 - 477 = -77$ μm and $\Delta_{\text{nnd}} = \hat{p}_{\text{nnd}} - p_{\text{nnd}} = 26 - 21 = 5$ μm. The fatigue life resulted from these two sets of defect parameters is calculated where grain features are kept constant at their mean values ($g_{\text{size}} = 100$, $g_{\text{mo}} = 15$).

The results are illustrated in Figure 11, where the average fatigue life error is 298 and 276 cycles for GPR and SLR, respectively.

### Table 10. The probability table for $p_{\text{nnd}}$ based on distance to the surface of classification results.

| $i$ | $D_S$ [mm] | $p_{\text{size}}$ | Probability of defect $i$ |
|-----|------------|-------------------|--------------------------|
| 1   | 0.021      | 400               | 1.92E-01                 |
| 2   | 0.026      | 100               | 8.05E-01                 |
| 3   | 0.028      | 400               | 1.27E-03                 |
| 4   | 0.030      | 400               | 1.99E-03                 |
| 5   | 0.033      | 400               | 8.55E-07                 |
| 6   | 0.034      | 100               | 1.03E-07                 |
| 7   | 0.036      | 100               | 3.58E-12                 |
| 8   | 0.036      | 100               | 1.57E-16                 |
| 9   | 0.037      | 400               | 3.79E-16                 |
| 10  | 0.037      | 400               | 2.66E-18                 |
| 11  | 0.039      | 400               | 7.51E-20                 |
| 12  | 0.117      | 100               | 6.93E-21                 |
| 13  | 0.134      | 100               | 1.28E-21                 |
| 14  | 0.199      | 400               | 8.68E-23                 |
| 15  | 0.221      | 100               | 2.12E-25                 |
| 16  | 0.227      | 400               | 1.55E-27                 |
| 17  | 0.402      | 400               | 3.54E-28                 |
| 18  | 0.404      | 100               | 2.45E-30                 |
| 19  | 0.507      | 100               | 1.52E-33                 |
| 20  | 0.521      | 100               | 2.51E-35                 |
| 21  | 0.561      | 100               | 1.18E-37                 |
| 22  | 0.728      | 100               | 3.28E-38                 |
| 23  | 0.766      | 400               | 4.46E-39                 |
| 24  | 0.835      | 400               | 4.19E-40                 |
| 25  | 0.910      | 400               | 2.31E-41                 |
| 26  | 0.925      | 100               | 7.79E-43                 |
| 27  | 0.959      | 400               | 1.12E-43                 |
| 28  | 1.767      | 400               | 1.53E-44                 |
| 29  | 1.857      | 100               | 1.96E-45                 |

The model fails to identify 1.82E-47

### Figure 10. a) One sample output where the solid green curve is the actual fatigue life, SLR prediction is shown with the blue dotted curve, and the GPR prediction is shown with the red dash-dotted curve. b) The boxplot representation of error percentage for 40 testing samples.

7. Conclusion

One of the main challenges that AM processes are facing is the uncertainty in the structural performance of AM parts. This issue has been a significant barrier in front of the broader adoption of AM by industries. This uncertainty can be seen in different modes of failure, especially fatigue failure, the dominant failure mode in various sectors. The amount and distribution of defects alongside microstructural features play a crucial role in characterizing the fatigue life of AM parts. Thus, it is necessary to develop a successful in situ qualification and certification methodology not only to identify the formation of defects but also to characterize the fatigue life based on detected defects.

In this article, a novel data-driven multiphase nondestructive in situ fatigue prediction methodology is developed to predict the fatigue resistance based on defect characteristics. The proposed method is developed in two phases. The objective of Phase (I) is
to characterize the defect features based on in situ thermal history. Thermal history, that is, melt pool images, carries the most critical information about the build process; therefore, these images are processed and utilized to predict the relative defect size. The model (I) demonstrates this procedure and wraps it up by training a classification model that predicts the relative size of the defects (small or large). The objective of Phase (II) is to build the relationship between the defect properties extracted from Phase I and the fatigue behavior based on the MSF model. MSF model was used to identify the defect-related input features. Each fatigue curve is represented with a parametric polynomial function, and polynomial parameters are predicted from defect properties and microstructural features. Stepwise linear regression (SLR) and Gaussian process regression (GPR) approaches are used to predict the final fatigue parameters.

The performance of Phase (I) methodology is validated using a cylinder fabricated by a direct laser deposition process, and Phase (II) model is validated based on MSF fatigue lifetimes. Considering the small number of samples, the proposed methodologies have noticeable accuracies. Although MSF is a potent physics-based prediction model for fatigue lives with useful utilities, it depends on a handful of parameters to provide a prediction. Additionally, the AM parts are prone to defects, and identifying the defect-related parameters using the post-manufacturing methods is expensive and time-consuming; therefore, the proposed methodology significantly accelerates the qualification and certification of additive manufacturing processes.

Some exciting topics remain open for future research. First, the proposed methodology can be extended to add a new phase to account for the design–thermal relationship. Second, an inverse robust design exploration can be developed to generate satisfying design solutions.

**Acknowledgements**

The research was sponsored by the Army Research Laboratory and was accomplished under Cooperative Agreement Number W911NF-15-2-0025. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

**Conflict of Interest**

The authors declare no conflict of interest.

**Data Availability Statement**

Research data are not shared.

**Keywords**

additive manufacturing, fatigue modeling, thermal history

Received: November 30, 2020
Revised: September 13, 2021
Published online: October 29, 2021

[1] S. M. Thompson, L. Bian, N. Shamsaei, A. Yadollahi, *Addit. Manuf.* 2015, 8, 36.
[2] I. A. Okaro, S. Jayasinghe, C. Sutcliffe, K. Black, P. Paoletti, P. L. Green, *Addit. Manuf.* 2019, 27, 42.
[3] A. Yadollahi, N. Shamsaei, *Int. J. Fatigue* 2017, 98, 14.
[4] J. Gockel, J. Beuth, K. Taminger, *Addit. Manuf.* 2014, 1, 119.
[5] R. I. Stephens, A. Fatemi, R. R. Stephens, H. O. Fuchs, *Metal Fatigue in Engineering*, 2nd ed., John Wiley & Sons, New York 2000.
[6] S. Findlay, N. Harrison, *Mater. Today* 2002, 5, 18.
[7] A. L. Pilchak, J. C. Williams, *Metall. Mater. Trans. A* 2009, 40, 2603.
[8] G. Lütjering, J. C. Williams, Springer Science & Business Media, Berlin 2007.
[9] J. A. Hall, *Int. J. Fatigue* 1997, 19, 23.
[10] T. Zhai, A. J. Wilkinson, J. W. Martin, *Acta Mater.* 2000, 48, 4917.
[11] I. Bantounas, T. C. Lindley, D. Rugg, D. Dye, *Acta Mater.* 2007, 55, 5655.
[12] A. L. Pilchak, R. E. A. Williams, J. C. Williams, *Metall. Mater. Trans. A* 2010, 41, 106.
[13] A. Yadollahi, N. Shamsaei, S. M. Thompson, A. Elwany, L. Bian, *Int. J. Fatigue* 2017, 94, 218.
[14] A. Yadollahi, M. Mahmoudi, A. Elwany, H. Doude, L. Bian, J. C. Newman Jr, *Fatigue Fract. Eng. Mater. Struct.* 2019, 43, 831.
[15] A. Yadollahi, M. J. Mahtabi, A. Khalili, H. R. Doude, J. C. Newman, *Fatigue Fract. Eng. Mater. Struct.* 2018, 41, 1602.
[16] S. Romano, A. Brückner-Foit, A. Brandão, J. Gumpinger, T. Ghidini, S. Beretta, *Eng. Fract. Mech.* 2018, 187, 165.
[17] P. A. Kobryn, S. L. Semiatin, in *2001 Int. Solid Freeform Fabrication Symp.*, 2001.
[18] A. Bagheri, A. Yadollahi, M. J. Mahtabi, M. Horstemeyer, *Int. J. Fatigue* 2019, unpublished.
[19] S. Tammas-Williams, P. J. Withers, I. Todd, P. B. Prangnell, *Sci. Rep.* 2017, 7, 1.
[20] M. Gorelik, *Int. J. Fatigue* 2017, 94, 168.
[21] S. Beretta, S. Romano, *Int. J. Fatigue* 2017, 94, 178.
[22] J. Newman, X. Wu, S. Venneri, C. Li, A NASA/CAE Cooperative Program, NASA RP-1309, 1994.
[23] J. C. Newman, B. S. Annigeri, *J. Eng. Gas Turbines Power* 2012, 134, 032501.
[24] J. C. Newman, K. Kota, T. E. Lacy, *Eng. Fract. Mech.* 2018, 187, 211.
