RESEARCH ARTICLE

Transfer Learning Methods for Fractographic Detection of Fatigue Crack Initiation in Additive Manufacturing

OR HAIM ANIDJAR[1,2,3,4], **RO'I LANG**[5], **AND MOR MEGA**[5]

[1]School of Computer Science, Ariel University, Ariel 40700, Israel
[2]Kinematics and Computational Geometry Laboratory (K&CG), Ariel University, Ariel 40700, Israel
[3]Ariel Cyber Innovation Center (ACIC), Ariel University, Ariel 40700, Israel
[4]Data Science and Artificial Intelligence Research Center, Ariel University, Ariel 40700, Israel
[5]School of Mechanical Engineering and Mechatronics, Ariel University, Ariel 40700, Israel

Corresponding author: Or Haim Anidjar (orhaim@ariel.ac.il)

ABSTRACT Recently, there has been an increased interest in additive manufacturing (AM) for its potential to reduce costs and lighten the weight of manufactured parts. However, materials produced through AM are prone to defects that can significantly impact their fatigue resistance. Identifying fatigue failure sources is crucial for the characterization of critical manufacturing defects, especially for the future use of AM in main load-bearing structural parts. This requires conducting fatigue tests and manually inspecting fracture surfaces. In this research, we introduce an innovative machine-learning model designed to detect the initiation defects causing fatigue cracks in Titanium Ti-6Al-4V samples manufactured by selective laser melting (SLM). The model also measures the distance between the detected fatigue failure source and the surface of the material. Our approach involves initially segmenting out areas without initiation points, and then identifying these points in the remaining areas. We then use established computer vision techniques to calculate their distance from the surface. The results of our study highlight the significant potential of using machine learning and computer vision to automate fractographic analysis. This advancement could greatly improve the speed and efficiency of this process, marking a new phase of productivity in the field. This research not only furthers artificial intelligence by introducing an innovative method but also may possess important applications in engineering.

INDEX TERMS Additive manufacturing, computer vision, fatigue failure, fractography, deep learning.

I. INTRODUCTION

Fractographic analyses are conducted subsequent to material failure to identify the primary causes of failure and extract insights for informing future structural integrity designs. This process is time-intensive and necessitates the expertise of professionals for precise measurements and comprehension. By delineating the underlying reasons for failure, it becomes plausible to anticipate failure through non-destructive methods. To achieve this objective, access to fractographic data is imperative, and the analysis procedure must be streamlined and automated.

The associate editor coordinating the review of this manuscript and approving it for publication was Szidonia Lefkovits.

In this study, we introduce an algorithm designed to facilitate detection and preliminary geometrical characterization. This algorithm represents the initial stride towards autonomous fractography defect characterization.

Section I-A furnishes an introduction to the examined material and underscores the significance of the analysis. Subsequently, Section I-B outlines the contributions of the proposed algorithm. Lastly, in Section I-C, the structure of the paper is delineated.

A. BACKGROUND, RESEARCH GAP, AND SCOPE

Additive Manufacturing (AM), also known as 3D-printing, is a manufacturing technique where parts are fabricated layer by layer as opposed to traditional manufacturing

methods [1], [2]. In traditional methods, subtractive manufacturing (SM) is used [3], [4]. In SM, excess material is removed from a larger block until the desired dimensions are reached. The potential to fabricate customized parts with complex geometries and low buy-to-fly ratios is enabled with AM for all manner of industries, in general, and for the aerospace industry, in particular. In recent years, these techniques have been implemented to produce various parts for aircrafts. However, AM has not been used to manufacture load-bearing elements due to the lack of knowledge regarding the fatigue behavior of these materials. Furthermore, previous studies [5], [6], [7], [8], [9] have shown that AM materials have a high defect concentration which negatively affects their fatigue performances compared to materials produced by traditional methods [10]. Currently, there are no regulatory guidelines for disqualifying faulty AM parts during manufacturing.

Selective laser melting (SLM) is an AM method in which metallic powder is deposited in thin layers onto a work surface and then selectively melted using a laser [11]. This process is repeated until the fabrication of the part is complete. However, the uneven melting, rapid heating and cooling, and other processes involved in AM can result in high concentrations of defects, including porosity, lack of fusion (LOF), and inclusions [12], [13]. These defects can be present on both the surface and interior of AM parts, and they can act as stress concentration sites, which may lead to fatigue crack formation and premature failure of the part. Additionally, microstructural features such as grain boundaries [14] may also act as stress concentration sites, resulting in seemingly random initiation sites for fatigue cracking.

Quality control during manufacturing and monitoring throughout the lifespan of AM parts primarily relies on Non-Destructive Testing (NDT) such as Acoustic emission, ultrasonic, Computer Tomography (CT) etc. [15]. To enhance the understanding of NDT results, Destructive Testing (DT) including fatigue tests followed by fractographic analysis [16] is deemed essential and mandatory. In order to perform fractographic analyses, the fracture surfaces of the tested specimens are scanned with a scanning electron microscope (SEM) [17]. Note that in order to capture SEM images, a microscope operator must manually focus on regions of interest (ROI) on the fractured surface. Moreover, capturing a SEM image is time-consuming. Subsequently, experts in the field of failure of materials are required to examine and measure visual data from the scanned surface of the specimen. Among the measured data are the location and size of the fracture initiation site, as well as the various stages of failure.

The aim of this research is to perform an automated analysis of fatigue DT results for AM specimens. The acquired defects characterization will be utilized for the independent identification of critical defects through NDT. This study comprises multiple phases. The initial phase, elaborated upon in this paper, focuses on extracting essential insights from SEM images of the fractured surface. These insights encompass defect identification, precise geometric analysis, and determination of defect type and its specific location within the specimen. Subsequent phases will incorporate machine learning techniques to establish a more profound comprehension of the interconnection between characteristics derived from SEM images, Micro-CT scans [18], and manufacturing parameters concerning the fatigue lifespan of the specimen.

In this paper, a computer vision model [19] is proposed which may significantly reduce the time required for scanning and analyzing fractographic data. The model was trained for fatigue failure of SLM Ti-6Al-4V specimens. This titanium alloy, known as $\alpha - \beta$ titanium [20], possesses exceptional properties such as high specific strength and remarkable resistance to corrosion. It stands as one of the most widely employed titanium alloys for AM, finding application in numerous fields where low density and outstanding resistance to corrosion are essential requirements. With the proposed model, autonomous detection of the ROI in each specimen will be enabled and used to perform additional SEM scans with a decreased Field of View (FV) for the same fracture surface only. The detection of the ROI will be followed by automatic measuring of attributes in the fractured surface so that only supervision of the process and measurements is required.

The model proposed in this study includes a two-step approach for detecting the fatigue crack initiation site in SEM images. In the first step, ResNet152 [21] architecture was used as a binary classifier to filter out image sections that do not contain the initiation site. In the second step, a YOLOv5s [22] model was trained on the remaining image sections to detect the initiation site and identify the cause of failure. Finally, the distance between the initiation site and the surface of the specimen was computed using standard computer vision techniques.

B. CONTRIBUTION

This paper offers a computer vision approach for performing fractographic analyses of AM materials, namely, detecting, classifying, and measuring the initial cause of failure. The contributions of this paper can be summarized as follows:

- **Increasing amount of conducted studies.** The quicker and more cost-effective analysis would facilitate additional studies and failure analyses of the fatigue behavior of AM materials. This, in turn, would enhance our understanding of the fatigue performance of these materials, enabling the development of a data control system capable of distinguishing between critical and non-critical defects. Such characterization could be utilized for part disqualification using NDT and promote their increased integration into the design process. Moreover, the reduction in expert man-hours required for the fractographic analysis significantly lowers the economic cost associated with the process.

- **Efficient time exploitation.** Performing complete fractographic analyses is a time-consuming process. Capturing SEM images of the fractured surface and measuring relevant attributes is a particularly laborious task. Implementation of an autonomous detection model can effectively reduce the time required for the capture of SEM images. Furthermore, such a model would enable near-instantaneous measurement of attributes, streamlining the analysis process considerably.

- **Reduction of usage in content-experts.** To perform fractographic analyses experts are required. First, a SEM operator with knowledge of the regions of interest in the fractured surface is essential. Additionally, a materials failure expert is needed to classify the image and determine which measurements are necessary. With the implementation of an autonomous model, the need for the involvement of experts in the SEM scanning procedure and in the analysis process will be significantly reduced. The role of the SEM operator would be limited to placing and removing the specimens from the microscope, while the materials failure expert would only need to validate the identification, classification, and measurements of the ROIs as obtained from the model.

- **Enabling repeatable analyses.** One of the significant drawbacks of manually measuring attributes in images is the potential for human bias to influence the measurements. If measurements are conducted by different individuals or even by the same person at different times, variations in the measured values are likely to occur. In contrast, a deep learning model would consistently retrieve the same measurements, enabling easier comparisons between different specimens and experiments. This can help reduce the influence of human bias on the analysis process.

C. PAPER STRUCTURE

The remainder of this paper is structured as follows: In Section II, related work regarding both fatigue failure of AM materials and applications of Deep Learning based computer vision are presented. The datasets and approach employed in order to detect initiation sites and measure their distance from the outer specimen perimeter are described in Section III. The results from each step of the detection process are presented in Section IV. Finally, in Section V, a fundamental discussion, concludes, and summarizes the proposed approach.

For ease of reading, a list of abbreviations is provided in Table 1.

II. RELATED WORK

Additive manufacturing (AM) techniques, which have a history dating back to the 1980s, have experienced a surge in popularity in recent times, primarily attributed to advancements in powder bed technologies [23]. Presently, AM is widely employed across various industries, such as the medical field [23], [24], automotive sector [25], [26], and aerospace domain [27], [28].

Similar to any novel material, comprehending the mechanical properties of AM materials is crucial for their successful integration into the design process, particularly when compared to materials fabricated using conventional techniques. Consequently, numerous studies have been conducted with the aim of gaining a deeper understanding of these materials.

TABLE 1. List of abbreviations.

Abbreviation	Meaning
AM	Additive Manufacturing
DT	Destructive Testing
FP	False Positive
FN	False Negative
HCF	high-cycle fatigue
HIP	Hot Isostatic Pressing
IoU	Intersection over Union
LOF	Lack of Fusion
L-PBF	Laser Beam Powder Bed Fusion
MAP	Mean Average Precision
NDT	Non-Destructive Testing
ROI	Region of Interest
SEM	Scanning Electron Microscope
SLM	Selective Laser Melting
SM	Subtractive Manufacturing
SVR	Support Vector Regression

Edwards and Ramulu [5] found that the tensile properties of as-built AM materials can be comparable to those of as-cast materials. However, the same study found that AM materials exhibit significantly lower fatigue performance. This decrease in fatigue performance was attributed to various factors, including poor surface finish, internal porosity, and residual stresses [29], [30]. Nevertheless, post-manufacturing heat treatments such as Hot Isostatic Pressing (HIP) [31] have been demonstrated by [9] and [32] to mitigate residual stresses and reduce porosity in these materials thereby improving there mechanical properties.

Fatigue performance of materials is typically evaluated through two approaches: fatigue crack propagation [33] and fatigue life assessment [34]. Leuders et al. [6], [8] conducted studies on heat-treated SLM Ti-6Al-4V and observed that its crack propagation performance was comparable to that of wrought materials. However, they found that the fatigue life of SLM Ti-6Al-4V was lower and exhibited higher variance due to the presence of internal pores. Additionally, they established a correlation between the size of defects and their distance from the surface with the fatigue life of the specimen.

The primary approach to better define the variance in fatigue life of AM materials involves examining the defects in materials and establishing criteria for their measurement. This approach, demonstrated by [35] and extensively discussed by [36], plays a significant role in understanding the fatigue behavior. This method proposes an effective area of defect which better characterises its effect on the mechanical properties of the material. However, a notable challenge associated with this approach is the requirement for manual measurement of characteristics, such as area of the defects

and their distance to the surface, as shown in Fig. 1. Manual measurement of these characteristics can be time-consuming. On the other hand, a secondary approach, as demonstrated by [37], focuses on exploring the anisotropic nature of "defect-free" materials. In their probabilistic model, fatigue cracks initiate in "weak" grain boundaries. While valuable, this approach is not as prominently utilized for assessing the fatigue life of AM materials as they contain a significantly higher occurrence of defects relative to SM materials.

The correlation between geometrical features of critical defects and fatigue performance has the potential to establish an algorithmic foundation for the nondestructive fatigue evaluation of additive manufacturing products. In a previous study [38], an integrated data-driven analytical framework was proposed for defect criticality in laser beam powder bed fusion (L-PBF) based on SEM scans of fatigue fractured surfaces. The results demonstrated strong relationships between defect features and fatigue life, achieving a low mean absolute percentage error of 0.101 using kernel support vector regression (SVR). While this framework shows promise in enhancing our understanding of defect-fatigue relationships, it still relies on manual detection and measurements of initiation sites and geometrical features. This study introduces tools that aim to address the measurement challenges associated with the primary approach, offering potential solutions to streamline the analysis of fatigue life in these materials.

Using computer vision models for fractographic analysis is a fairly new field previously employed in [39] and [40] for examining quasi-static fracture. In [39], a VGG16 model in a U-net architecture [40], [41] was used in order to segment intergranular and transgranular fractures in a ceramic material. They were able to achieve a total mean Intersection over Union (IoU) that was > 91% for two different materials. In [42], unsupervised learning [43] was employed on a clustering algorithm in order to classify SEM images based on tungsten composition achieving accuracy that was greater than 90%.

ResNet152 [21] and YOLOv5 [22] are two models that have previously been shown as effective tools for image prediction and object detection [44], [45], [46]. In [44], ResNet152 was used as part of an algorithm that predicts the surface morphology of SEM images. In [45], a ResNet-based RetinaNet model was employed to detect defects in SEM images of semiconductors and concluded with promising results. The YOLOv5 model was used in [46] to detect defects in steel surfaces, as a base for a real-time detector, achieving a mean average precision (mAP) of 0.752.

In this study, a two-step approach is proposed for detecting the fatigue crack initiation site in SEM images. In the initial step, a binary classifier based on the ResNet152 architecture is employed to filter out image sections that do not contain the initiation site. Subsequently, the remaining image sections are used to train a YOLOv5s model, which accurately detects the initiation site and identifies the cause of failure. Finally, standard computer vision techniques are applied to calculate the distance between the initiation site and the specimen's surface.

III. FRAMEWORK

In this section, the dataset and framework of the algorithms used in this study are presented. An illustration of the full framework is shown in Fig. 2. The figure consists of four stages. In the first stage, SEM images of the fractured specimens are divided into an 8 by 8 grid. Subsequently, a binary classification is performed to detect anomalies in the sub-images. In the third stage, YOLOv5 is employed for the detection and classification of each type of ROI using the previously detected images. Finally, the identified anomalies are incorporated into the whole image, and the distance from the center and outer surface of the specimen is calculated.

Between each step in Fig. 2, the data used is prescribed. In Section III-A, an explanation of the pre-processing procedure to obtain the data used, as well as the dataset employed in each step of the analysis, are presented. Then, in Section III-B, the architecture of ResNet152 is discussed. The methods for training of ResNet152 and YOLOv5 models are described, respectively, in Sections III-C, and III-D. Finally, the algorithm used to compute the distance from the initiation site to the outer surface of the specimen is presented in Section III-E.

A. DATASETS

This study utilized SEM images of the entire fractured surfaces of Ti-6Al-4V SLM fatigue specimens. The specimens were manufactured and tested in high cycle fatigue (HCF) in [8], following the ASTM E466-15 standard [47]. Such tests determine S-N curves by measuring fatigue life related to various maximum stresses and stress amplitudes. The specimens tested in [8], were manufactured with varying properties and subjected to a maximum cyclic fatigue stress of 634 MPa at a cyclic stress amplitude of 0.1. The maximum stress applied in the tests was chosen in the expected HCF region of Ti-6Al-4V. The fatigue life measured for each specimen, as well as additional mechanical properties and data regarding the manufacturing process, may be found in [8]. Note that in the case of AM parts, variability in the fatigue life among different specimens is expected, as observed from the results in [8]. To investigate the reasons behind this variability, fractographic analyses are required to identify and characterize the sources of failure in each tested specimen.

It is important to note that to conduct accurate fractographic analyses and pinpoint the failure source, specialized experience and skills are essential. The saturation lines within the model point toward the location of the failure source. Just as an expert would identify the source based on the fractographic surface and differentiate between material inherent defects, which may not be the primary cause of failure, and those that are, the model proposed here enables the detection of defects that were the cause and source of failure.
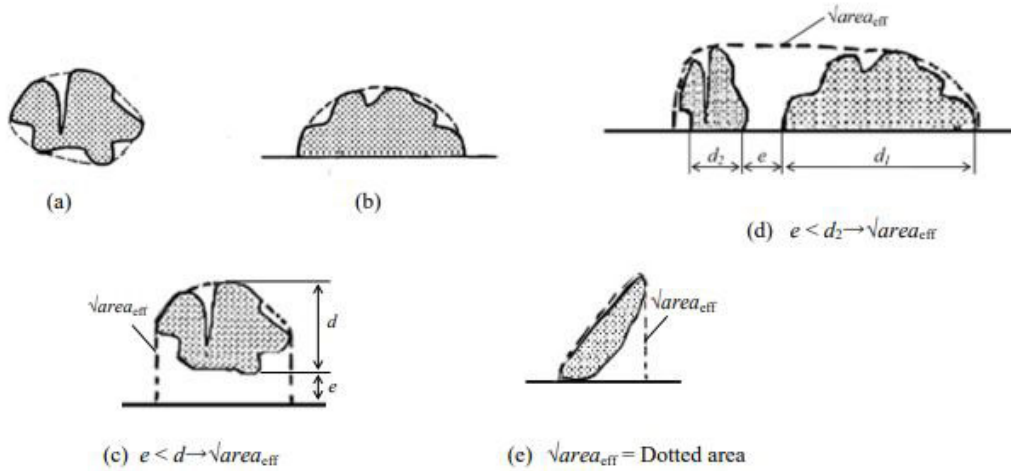
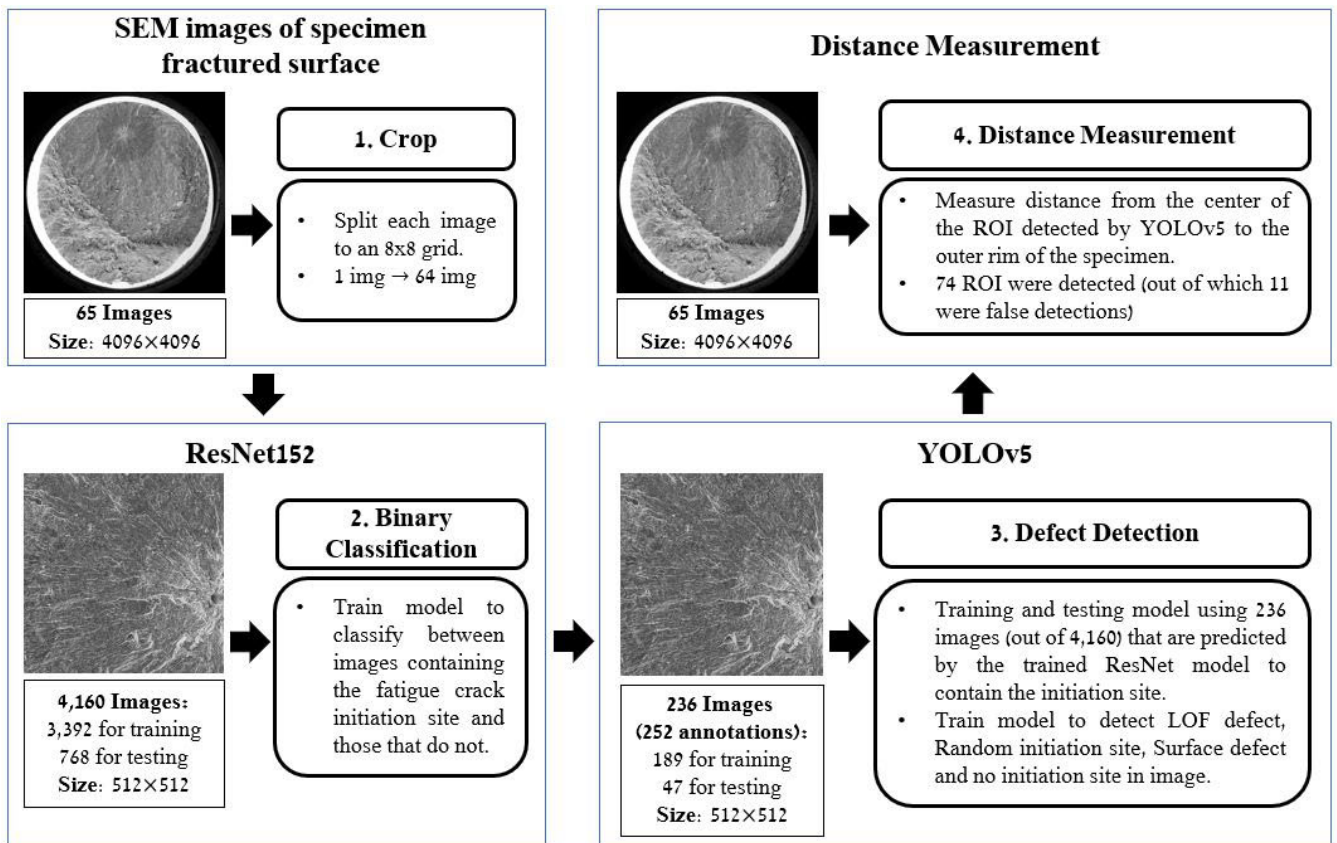FIGURE 1. Method for measuring defect effective area as proposed by [35].

FIGURE 2. Visualization of the entire framework used in this study. The first step (top left) involves splitting each SEM image into 64 images using an 8 by 8 grid. Next (bottom left), ResNet152 is trained as a binary classifier to detect images containing the fatigue crack initiation site. In the third stage (bottom right), YOLOv5 is trained to detect defects in the images predicted by ResNet152 to contain the initiation site. Finally (top right), the distance from the center of the ROI detected by YOLOv5 to the outer rim of the specimen is computed.

A total of 36 specimens were examined with 65 images of the fractured surfaces used. Three specimens were excluded from the training dataset because their entire fractured surfaces exhibited defects resulting from intentionally poor printing parameters used for these specific specimens. Additionally, another image was not utilized due to damage inflicted upon the fractured surface after the fatigue test. The images used had a resolution of 4096×4096 pixels and a view field of 5.5 mm. With an approximate pixel size of 1.34 μm. An example of one SEM image is presented in

Fig. 3a. Additional examples of images of the entire fractured surface are shown in V. The full datasets used in this study will be provided on demand.

To identify the fatigue failure source a comprehensive approach was used including training of ResNet152 and YOLOv5 models using SEM images of fractured SLM Ti-6Al-4V fatigue specimens. The images used for training were classified into distinct categories based on their characteristics. For the ResNet152 model, the classification was between images containing the fatigue crack initiation site and those without. This initial step was crucial in filtering out areas irrelevant to fatigue crack initiation. The subsequent training of the YOLOv5 model further refined this process. Here, images predicted by ResNet152 to contain the initiation site were annotated with four specific labels: ''no initiation'', ''LOF'', ''random'', and ''surface defect''. This labeling system was key to distinguishing between different types of defects and damage.

(1). ''No initiation'' labeled images helped in identifying regions without any fatigue initiation. (2). ''LOF'' (Lack of Fusion) defects were specifically marked to identify a common defect in SLM-manufactured specimens. This category may include pores as well. (3). The ''random'' category encompassed initiation sites that were not identifiable as specific defects at the given resolution, likely indicating weaker grain boundaries rather than explicit defects. (4). ''Surface defect'' was a crucial label for identifying defects on the specimen's surface, including LOF defects and other surface damages.

Through this detailed classification and training process, the models were equipped to discern the fatigue crack initiation sites from other types of defects and damages, including corrosion and impact damage. The combination of high-resolution SEM imaging, precise labeling, and advanced machine learning models enabled accurate identification and differentiation of these various aspects in the specimens. This methodology ensures that the analysis is focused on fatigue crack initiation, providing valuable insights into the fatigue properties of SLM Ti-6Al-4V specimens while effectively accounting for other types of material anomalies.

TABLE 2. ResNet152 dataset composed of 4160 images (8 by 8 grid of the 65 examined specimen interfaces) split into anomaly and no anomaly sets. The number of tagged images used for model training was approximately 81.5% and for testing 18.5%.

| Set/Label	Anomaly	No Anomaly	Total
Train set	69	3,323	3,392
Test set	15	753	768
Total	84	4,076	4,160
%	2	98	100

To train the ResNet152 model, after histogram equalization, the images were split into a training set of 53 images and a test set of 12. Each image was further split into 64 smaller images, each of size 512 × 512 pixels which is depicted in Fig. 3b. In Fig. 3c, an example of one of the split images is presented. The images were then classified into two categories: those containing the fatigue crack initiation site and those without. The training and test datasets are summarized in Table 2. For training, 69 images with the initiation site and 3,323 images without the initiation site were used, while the test dataset comprised 15 images with the initiation site and 753 images without.

For training the YOLOv5 model, 236 images that were predicted by ResNet152 to contain the initiation site were used. These images were obtained from both the training and testing datasets using the best weights obtained by ResNet152. The images were annotated using label-Img software, employing four distinct labels: ''no initiation'', ''LOF'', ''random'', and ''surface defect''. The ''no initiation'' class refers to images that do not exhibit any initiation sites. An example of this class is presented in Fig. 4a. Class ''LOF'' corresponds to Lack of Fusion defects, as demonstrated in Fig. 4b. The ''random'' class, demonstrated in Fig. 4c, represents initiation sites where the cause of initiation could not be discerned from the available image resolution. These sites are likely initiations resulting from ''weak'' grain boundaries rather than being classified as defects. The ''surface defect'' label, shown in Fig. 4d, is assigned to defects located at the surface of the specimen, which can also include LOF defects. Table 3 displays the categorization of 252 anomalies from the 236 images predicted by ResNet, which forms the composition of the dataset utilized for the YOLO5v5 model. Out of the entire dataset, 165 images had no initiation site, 47 images contained a random initiation site and 22 images contained LOF defects, of which two images had two LOF defects each, concluding in 24 LOF defects. Furthermore, among the 24 LOF defects, 14 were also identified as surface defects. Additionally, two images in the dataset contained solely surface defects. Adding up to a total of 16 surface defects. The images were randomly divided into a training set and a test set with an 80-20 split, respectively. The average size of each type of anomaly is shown in table 4.

When examining the composition of the dataset, two primary issues are evident. Firstly, the small dimensions of the initiation sites in relation to the overall fractured surface image. As shown in Table 2, approximately 2% of the smaller images, obtained by dividing each original image into 64 smaller images, contained part of the initiation site. This issue is further highlighted in Table 4, which indicates that, on average, the bounding boxes of random initiation sites span only about 11% of the cropped images, while LOF and Surface defects bounding boxes span less than 1% of the image. Another challenge in analyzing the images is the scarcity of defects in the dataset. As presented in Table 3, the dataset includes only 87 examples of initiation sites (including random, LOF, and surface defects) in total.

Utilizing the datasets outlined in Table 2, and employing a binary classification approach, irrelevant areas were initially removed to refine the dataset. The architecture of the ResNet152 algorithm employed is elucidated in the
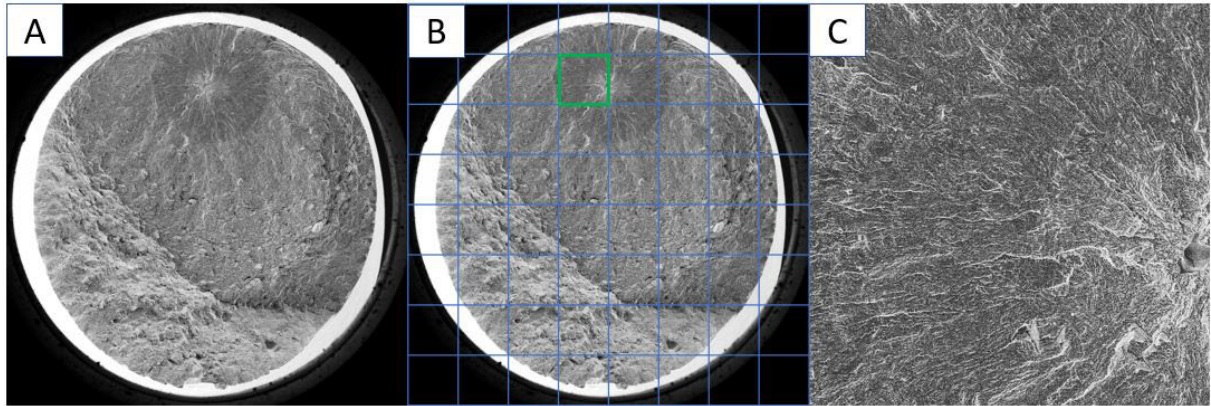
FIGURE 3. (a) SEM scan of the entire fractured surface of a Ti-6Al-4V fatigue specimen (resolution: 4096 × 4096 pixels, view field: 5.5 mm, and pixel size: 1.34 μm) and (b) The 8 × 8 cropping grid of the image where the green cell is (c) a cell from a 8 × 8 grid of the entire SEM specimen containing part of the initiation site (resolution: 512 × 512 pixels).

TABLE 3. YOLOv5 dataset composed of 236 images detected by ResNet152, with 252 anomalies split into four defect types, namely, no initiation, LOF, Random, and Surface Defect. The number of tagged images used for model training was 80% and for testing 20% for each defect type catagory.

Set/Parameter	No Initiation	LOF	Random	Surface Defect	Total
Train set	134	17	38	12	201
Test set	31	7	9	4	51
Total	165	24	47	16	252
%	65.5	9.5	18.7	6.3	100

TABLE 4. Average initiation site bounding box size in YOLOv5 Dataset. Examples for these bounding boxes can be seen in Fig. 4.

Property / Parameter	LOF	Random	Surface Defect	Image
Area (pixels)	2,474	29,407	1,856	262,144
% of image	0.94	11.2	0.7	100

subsequent section. Then, the training procedures employed for the ResNet152 and the Yolov5 are described.

B. RESNET152 ARCHITECTURE

A binary classification approach was employed using ResNet152 [21] to eliminate irrelevant areas in the images of the entire fractured surface of the specimen. This was crucial to improve the training quality of the object detection model. Since the crack initiation site occupies only a small portion of the entire image, as highlighted in Section III-A, there exists a significant imbalance between the sections containing the initiation site and those that do not.

The ResNet152 model, proposed by [21], was designed to classify images in the ImageNet dataset across 1000 classes, with inputs of 224 × 224×3 dimensions. For the binary classification task examined here, the final fully connected layer of the model was replaced with a three-layer sequential model of fully connected layers, with output sizes of 1000, 250, and 2, respectively. The first two layers utilize a ReLU activation function [48] while the final layer uses a Softmax activation function [49]. A visualization of this modified model is shown in Fig. 5. It should be noted that the weights in the convolution block are pre-trained weights from ResNet152, while the weights in the Sequential model are randomly initialized.

Prior to adopting this architecture, a validation process was conducted to compare various configurations of the

model. The training set comprised 40 images of the entire fractured surface, while the validation set contained 6 images. The validation process involved experimenting with image splitting using both an 8 by 8 grid and a 16 by 16 grid, as well as training with and without histogram equalization of the images. Furthermore, the convergence of the model was assessed during 150 epochs of training. After conducting this validation process, the training described in Section III-C was employed. The training phase utilized the optimal configuration determined from the validation process and proceeded with the entire dataset.

C. RESNET152 TRAINING

The dataset used for training the model in Section III-A provides a description of the composition of the datasets that were employed for both the training and testing of the model. To meet the input requirements, each image was transformed from its original size of 512 × 512×1 to a tensor of size 224 × 224×3. Following this, the datasets were shuffled and divided into batches each containing 32 images.

After completing the pre-processing stage, transfer learning was applied to the model by exclusively training the sequential model added to ResNet152. The training process was carried out utilizing the cross-entropy loss function [50] with mean reduction, as illustrated in Equations 1 and 2. As outlined in Section III-A, imbalances within the datasets were addressed by computing and implementing class

weights, which were determined based on Equation 3.

$$l(x, y) = \sum_{n=1}^{N} \frac{l_n}{\sum_{n=1}^{N} w_{y_n}} \qquad (1)$$

$$l_n = -w_{y_n} \cdot log\left(\frac{exp(x_{n,y_n})}{\sum_{c=1}^{C} exp(x_{n,c})}\right) \qquad (2)$$

$$w_{y_n} = \frac{D}{C \cdot d_{y_n}} \qquad (3)$$

In Equation 1, $l(x, y)$ is the total loss of the batch, x and y are the input and target respectively l_n, calculated in Equation 2, is the loss of each item in the batch, and w_{y_n} is the class weight of the class of the target. In Equation 2, $x_{n,c}$ is the logit value for class 'c' in item 'n' of the batch, x_{n,y_n} is the logit value of the same class of the target. The parameter C is the number of classes (in our case 2), N the batch size (in our case 32). In Equation 3, D is the number of images in the dataset, and d_{y_n} is the number of images in the class of the target. For optimization, the model utilized an Adam optimizer [51] with a leaning rate of 10^{-3}.
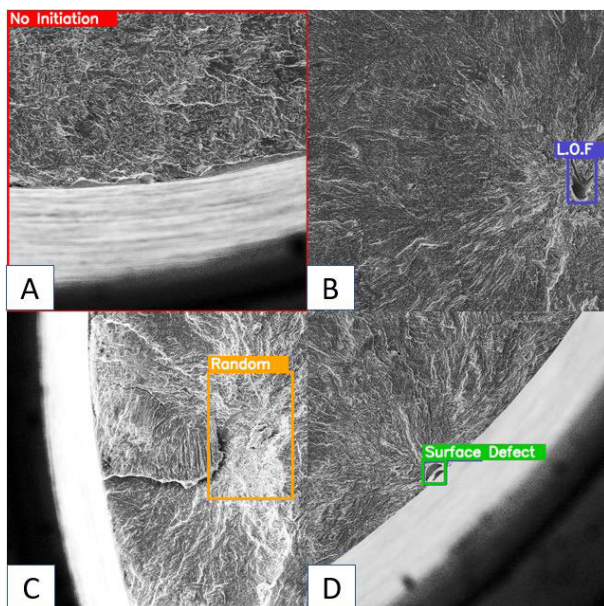
FIGURE 4. Example of annotation of YOLOv5 dataset. (a) an image that does not contain the initiation site (red), (b) an image containing a LOF defect (purple), (c) an image containing a random initiation site (orange) and (d) an image containing a LOF defect which is also a surface defect (green).

D. YOLOV5 TRAINING

In the second stage of the image processing the object detection framework YOLOv5 was utilized to detect the fatigue crack initiation site in images that were identified by the ResNet152 model to contain the initiation site. This framework was obtained from Yolo V5 Github Repository. The dataset utilized for training and validation of the model was split 80-20 respectively, and its composition is detailed in Section III-A. During training, the model analyzed images of $512 \times 512 \times 3$ dimensions, with a batch size of 16, and

initialized with pre-trained YOLOv5 weights. The training was conducted for 300 epochs and evaluated both with and without class weights to determine their necessity.

Upon completion of model training, bounding box coordinates were transformed into coordinates relative to the original image of the entire fractured surface. The bounding boxes that were saved corresponded to the highest probability detection of each small image, as well as any other detection with a probability greater than 0.3. In cases where the initiation site was split between two images and detected separately in each image, the bounding boxes were merged in accordance with Equations 4-7.

$$X = min(x_1, x_2) \qquad (4)$$

$$Y = min(y_1, y_2) \qquad (5)$$

$$W = max(x_1 + w_1, x_2 + w_2) - X \qquad (6)$$

$$H = max(y_1 + h_1, y_2 + h_2) - Y \qquad (7)$$

where, x_i and y_i are the coordinates of the top left corner of each of the bounding boxes, w_i, and h_i are the width and height of each bounding box respectively. The parameters X and Y are the coordinates of the top left corner of the merged bounding box, and W and H are the width and height of the merged bounding box, respectively.

E. DISTANCE MEASUREMENT

The final stage of image processing involved measuring the distance between the detected defect and the outer surface of the specimen. This distance is known as one of the most fundamental characteristics of the fatigue crack initiation site. One of the main objectives of this study is to compute this distance using YOLOv5 and leveraging conventional computer vision techniques. The computation process involved three steps and is visually illustrated in Fig. 6.

Firstly, the center of the specimen was detected. This was performed by initially equalizing the histogram of the image followed by a 3×3 Gaussian blur [52] which was applied to the image to reduce noise and smoothen the edges. Subsequently, the image was converted into an edge representation using the Canny edge detection algorithm [53]. The resulting image is presented in Fig. 6b. To locate the yellow bounding rectangle that encapsulates the specimen, the algorithm identified the outermost tenth white pixel (of the row/column) from each of the four directions (top, bottom, left, and right). These outermost pixels formed the edges of the bounding rectangle shown in Fig. 6c. The center of the specimen was then defined as the center point of this bounding rectangle also shown as a red dot at the specimen center in Fig. 6c.

Secondly, the intersection point was determined by finding the point where a line passing through the center of the specimen and the detected anomaly (whether it is classified as a random anomaly, lack of fusion, or surface defect) intersects with the outer rim of the specimen as shown in Fig. 6e.
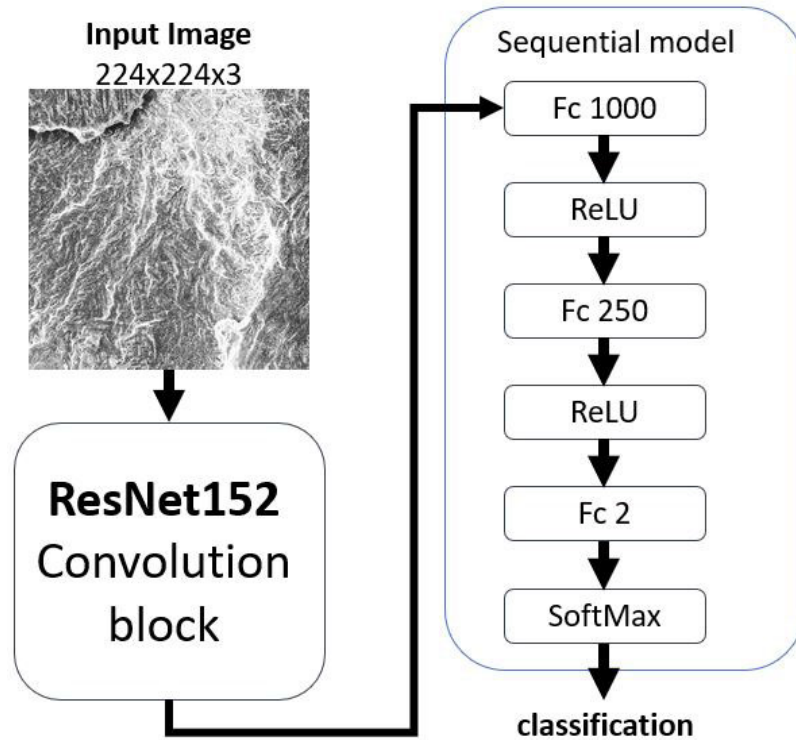
FIGURE 5. Visualization of ResNet152 model used. The input for the model is a 224 × 224 × 3 image which is propagated through the convolution layers of ResNet152. The classification is performed by an added three-layer fully connected sequential model. This model is used for binary classification between images containing the fatigue crack initiation site and images that do not contain it.

To define the outer rim, a multi-step contour detection process was employed. This process aimed to generate a binary masked image of the fractured surface, in which the outer rim of the fractured surface is present. An example of an obtained mask is demonstrated in Fig. 6d. The intersection point was defined as the final black pixel on the line passing through the center of the specimen and the detected anomaly. It is worth noting that this process was not utilized to detect the center of the specimen due to its poor performance in detecting the outer rim located further away from the initiation site in many images. Finally, the Euclidean distance was calculated between the intersection point and the center of the anomaly. This value was then compared to the manual measurement performed.

IV. EXPERIMENTAL EVALUATION AND RESULTS
Following the creation of the dataset and framework as depicted in Section III, the means of evaluation of each Deep Learning model and the results of each phase of this study are presented in this section. In Section IV-A, evaluation methods are discussed and the results obtained are presented in Section IV-B.

A. EVALUATION OF THE DEVELOPED PREDICTIVE MODEL
The means of evaluation of each Deep Learning model are described in this section. In Section IV-A1, the evaluation

method for ResNet152 is presented. The evaluation method for YOLOv5 is discussed in Section IV-A2.

1) RESNET152 EVALUATION
A common metric for evaluating a classification model performance is examining the accuracy in predicting the test set outcomes, as depicted in Equation 8. However, when dealing with imbalanced datasets, it becomes essential to consider alternative metrics that provide a more comprehensive understanding of the model behavior. Precision, Recall, and F1-score serve as important evaluation measures [54], [55], computed by Equations 9-11, respectively. These metrics offer insights into the predictive abilities of the model by considering factors such as true positives, false positives, and false negatives.
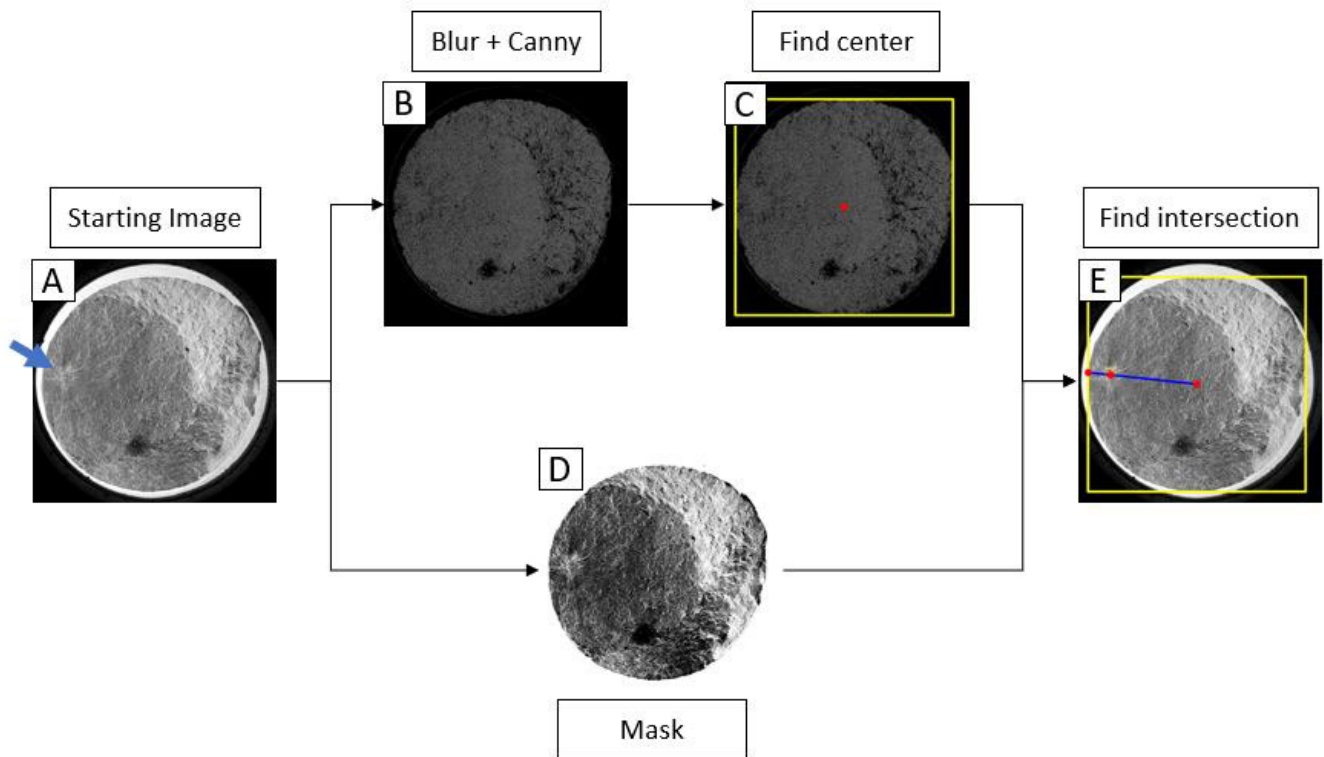
$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{8}$$

$$Precision = \frac{TP}{TP + FP} \tag{9}$$

$$Recall = \frac{TP}{TP + TN} \tag{10}$$

$$F1 - Score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \tag{11}$$

In Equations 8 through 10, TP and FP are both instances where the model predicted a 'positive' class with the former

FIGURE 6. Visual illustration of the measuring process (a) starting image where the arrow points to the initiation site, (b) The image after Blur and Canny algorithms, (c) The bounding box and center of the specimen detected, (d) A binary masked image of the fractured surface and (e) The computed intersection with the outer rim is indicated, with the three marked dots representing, from right to left, the center of the specimen, the center of the detected anomaly, and the computed intersection point.

being a true prediction, and the latter a false one. The parameters TN and FN stand for instances where the model predicted a 'negative' class correctly and incorrectly, respectively. For this study, the 'positive' class is defined as images not containing the initiation site, while the 'negative' class comprises images containing the initiation site.

The primary objective of the model in this study is to accurately identify and filter out sections of the images that do not contain the initiation site. Hence, minimizing the false positive (FP) value, which represents images containing the initiation site being misclassified by the model, is of utmost importance. Additionally, reducing the false negative (FN) value, which corresponds to images not containing the initiation site being misclassified, is of secondary importance. By reducing the FN value, the YOLOv5 model can be trained to focus on images that actually contain the initiation site. This emphasis on minimizing false negatives enables the model to learn more effectively and accurately to detect initiation sites.

To achieve these objectives, the confusion matrix of the test set evaluation was assessed after each epoch of training. By analyzing the confusion matrix, the performance of the model in terms of correctly classifying true positives, true negatives, false positives, and false negatives is evaluated. This iterative evaluation allowed for monitoring and

adjusting the model performance, with a focus on minimizing misclassifications and optimizing the filtering of images based on the presence or absence of the initiation site.

In order to gain a deeper understanding of the model classification behavior, visualizing the classification outcomes is crucial. By displaying the assigned classification for each section on the large images in the test set, a more comprehensive assessment of the model performance can be obtained. Visualizing the classifications provides valuable insights into the severity of mistakes made by the model, allowing for a better understanding of its strengths and weaknesses. This visual assessment enhances the interpretability of the model predictions. Also, it facilitates further analysis while improving the classification capabilities of the model.

2) YOLOV5 EVALUATION

The main focus of this study was to detect the anomaly responsible for initiating fatigue failure, rather than classifying the specific type of anomaly (such as random, lack of fusion, or surface defect). For each image in the test set, the bounding box with the highest probability prediction, as well as any other bounding boxes with a probability greater than 0.3, were visually drawn on the image.

Subsequently, a binary confusion matrix was manually created. The first class in the confusion matrix represented
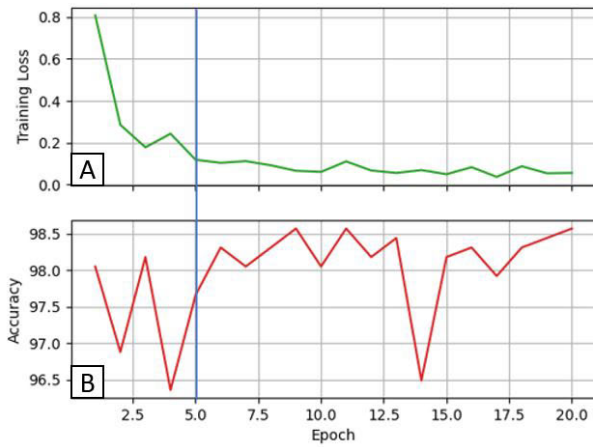
FIGURE 7. ResNet152 evaluation metrics for 20 epoch of training. (a) Average training Loss, (b) Test Accuracy, (c) Precision, (d) Recall and (e) F1-Score. Epoch 5 was selected as the best epoch due to a Precision of 1.0 and higher Accuracy and F1-Score than epoch 4.
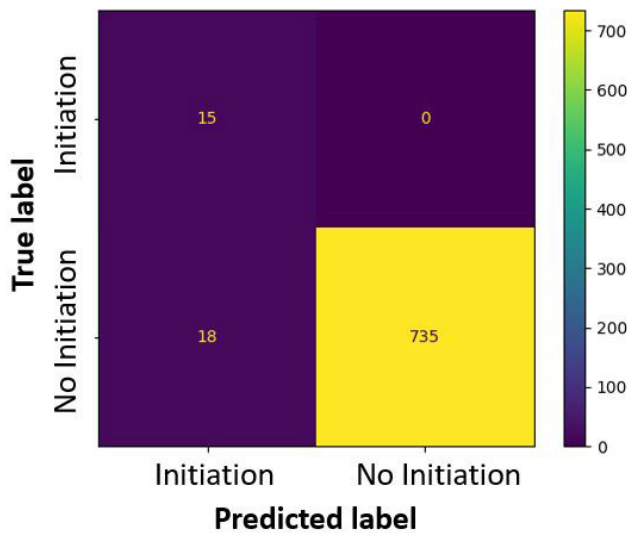
FIGURE 8. Confusion matrix depicting the evaluation results for the test set on epoch 5 of ResNet152 training. It is shown that the model correctly classified all images containing the initiation site while successfully filtering out 735 images that do not contain it, corresponding to an accuracy of 97.6% on the non-initiation site images.

FIGURE 9. Visualization of ResNet152 classification where (a) and (b) are images of the entire fractured surface from the test set (the blue arrows point to the initiation site). (c) and (d) are the respective detection by ResNet152 were the green sections are those predicted to contain the initiation site and the red sections predicted not to contain it.
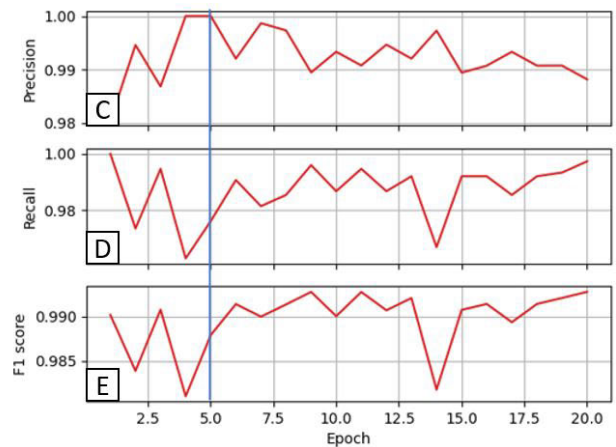
the anomalies, while the second class represented images containing no anomalies. By visually inspecting the drawn bounding boxes and examining the binary confusion matrix, the accuracy, and effectiveness of the model were assessed.

B. RESULTS

In this section, the results of each phase of this study are presented. In Section IV-B1 the results for ResNet152 are described. Results for YOLOv5 are presented in Section IV-B2. Lastly, in Section IV-B3, the results of the distance measuring algorithm are discussed.

1) RESNET152 RESULTS

The findings reported in this study pertain to the training of ResNet152 over a period of 20 epochs. Given its role as

a filter, the primary focus lies on achieving high precision, which entails filtering out only those images that do not contain the initiation site. Secondary metrics such as test accuracy and F1-score provide a general indication of the overall performance of the model. In Fig.7a and 7b, the average training loss and Accuracy for each epoch of training are presented respectively. The values of the Precision, Recall and F1-Score throughout the training process are depicted for each epoch in Figs.7c through 7e, respectively.

It may be observed that for both epoch 4 and epoch 5 a precision score of 1.0 was achieved meaning, no images containing the initiation site in the test set were predicted
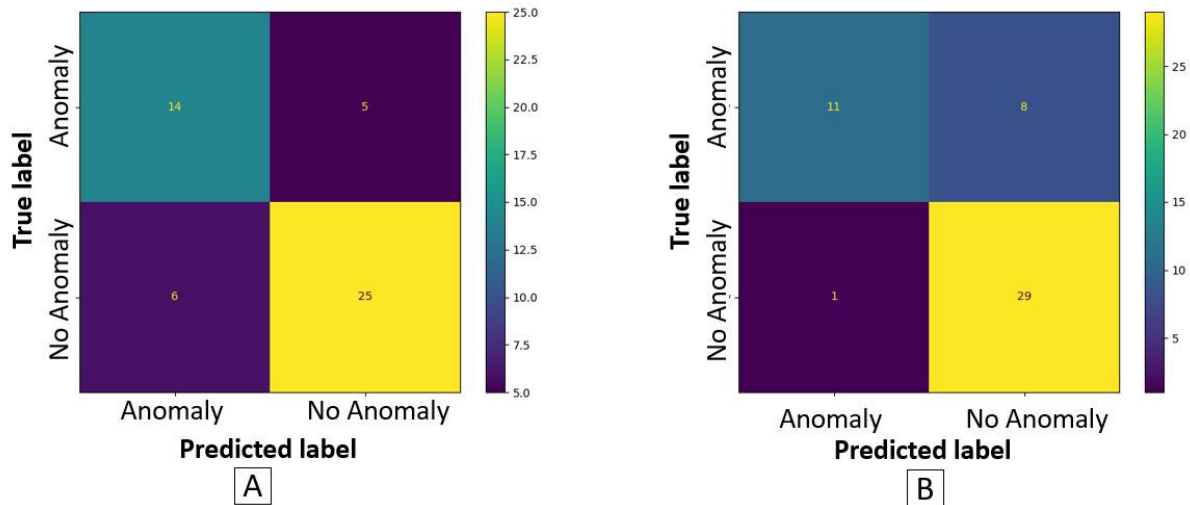
FIGURE 10. Confusion matrices of the best weights obtained with YOLOv5 for (a) Training using class weights and (b) Training without. Training with class weights was deemed the superior option due to higher defect detection rates relative to training without class weights.
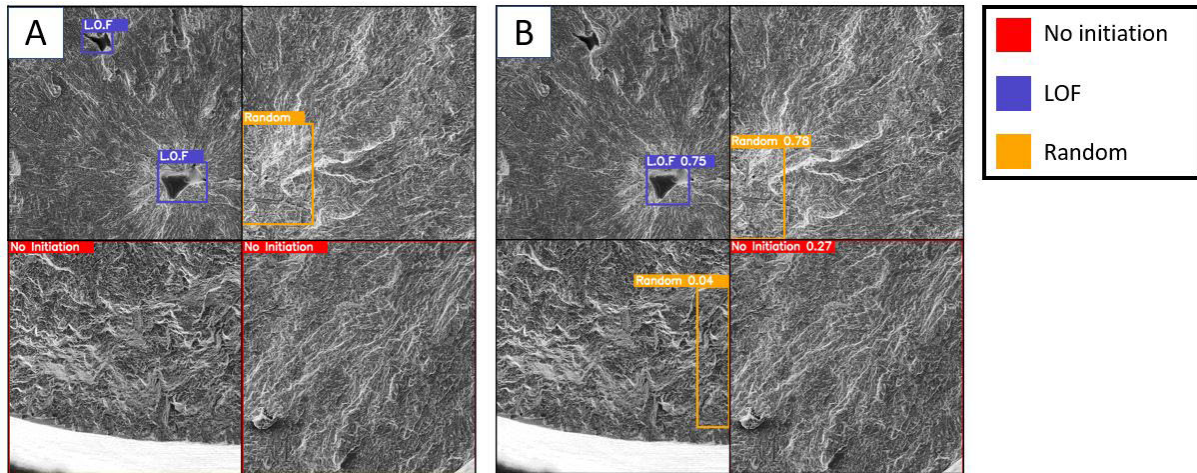

FIGURE 11. Detections made by YOLOv5 best weights trained with class weights. (a) The labels of four images in the test set, and (b) the respective model predictions. In the top-left image, two LOF defects are present. The model successfully detected the main defect with a score of 0.75 where the fatigue crack initiated; however, it did not identify the additional defect where the crack propagated through. In the bottom-left image, the model made a false detection of a random initiation site with a low probability of 0.04. Nevertheless, in both the top-right and bottom-right images, the model made correct detections of the random initiation site with a score of 0.78 and no initiation with a score of 0.27 in the images respectively.

to not contain the initiation site. However, epoch 5 exhibited superior accuracy, and F1-score compared to epoch 4. Thus, rendering it to be the optimal model among the two. It is worth noting that during the fifth epoch of training, two images containing the initiation site were falsely predicted not to contain it, which subsequently led to them being filtered out from the dataset. However, it is important to mention that these two images were visually distinct from the rest of the dataset. One of these images displayed a faint random initiation site, while the other exhibited a significant LOF defect. It is noteworthy that this type of LOF defect was exclusively observed in specimens manufactured using poor printing parameters.

To gain deeper insights into the significance of these metrics, a confusion matrix was constructed for epoch 5, as depicted in Fig. 8. It may be observed from this figure that during epoch 5, 735 of the images that do not contain the initiation site were effectively filtered out of the dataset, corresponding to 97.6% of the images that actually do not contain the initiation site. Consequently, the proportion of images containing the initiation site increased from 1.95% prior to filtering to 45.45% post-filtering.

In Fig.9, a visualization of the model predictions for two full images from the test set are shown. In Fig.9a, and Fig.9b, the two original images with arrows pointing at the initiation site are presented. In Fig.9c, and Fig.9d, the
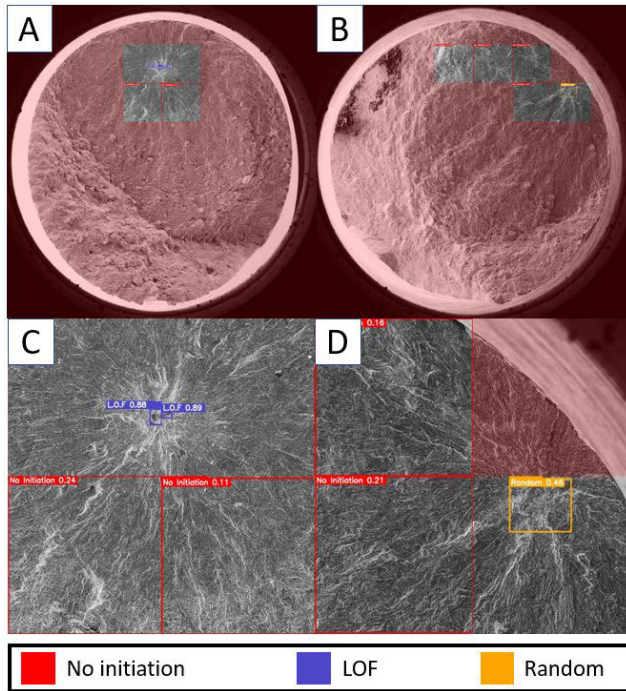
FIGURE 12. Combined results of both models, namely, ResNet152 and YOLOv5. (a) and (b) are images of the entire fractured surface where the red tinted sections were filtered out by ResNet152 and the bounding boxes are YOLOv5 predictions. (c) and (d) are the respective close-up views of the 4 sections surrounding the initiation site.

respective predictions of the models are visualized. Green rectangles represent the sections that were predicted by the model to contain the initiation site, and the red area represents sections that were predicted not to contain initiation. Note that in both images the model detected the initiation site. However, in Fig.9c, no additional sections were falsely predicted, whereas in Fig.9d, there are five falsely identified sections. Hence, further refinement by YOLOv5 to eliminate such false predictions is necessary.

2) YOLOV5 RESULTS

In accordance with the discussion in Section IV-A, the training of YOLOv5 was conducted both with and without the utilization of class weights, and the evaluation was performed in a binary manner. In Fig. 10a and in Fig. 10b the confusion matrix for the best weights obtained during model training with class weights and without class weights, respectively, is shown. Utilizing class weights during training led to higher defect detection rates along with an increased incidence of false positives, where some background was incorrectly identified as defects. Nevertheless, considering the prioritization of avoiding missed anomalies over false detections, it is evident that training with class weights emerges as the superior option, since only 5 missed anomalies occurred whereas without class weights, 8 anomalies were missed. It is noteworthy to emphasize that within the confusion matrix of the training

with class weights shown in Fig. 10a, two missed anomalies were secondary defects that the fatigue crack propagated through and were not the actual locations of fatigue crack initiation.

In Fig. 11a and Fig.11b, the annotations and corresponding predictions of four images from the test set are presented, respectively. In the upper left image two LOF defects are marked, whereas only the main LOF defect was identified by the model with a probability of 0.75. A random defect was marked in the upper right example and identified by the model with a probability of 0.78. In the lower right image the image was correctly identified as not containing the initiation site with a probability of 0.27 and In the lower left image a "Random" initiation site was falsely detected with a probability value of 0.04. These false detections made by the model primarily involved the identification of "Random" initiation sites and were marked with a low probability.

To gain a better understanding of the combined behavior of the models, images depicting the entire fractured surface showcasing the sections filtered out by ResNet152 and the predictions made by YOLOv5 were generated. Examples of such images are presented in Fig. 12 and in V. In Fig. 12a and Fig. 12b the combined results of two full images from the dataset are shown. The red sections were predicted by ResNet152 not to contain the initiation site, and the bounding boxes are YOLOv5 predictions. Fig. 12c and Fig. 12d are enhanced views of the four sections surrounding the initiation site of the images in Fig. 12a and Fig. 12b, respectively. The effectiveness of both models becomes evident in Fig. 12, where their contributions are observable for different types of initiation.

In Fig. 12a and 12b, it is evident that ResNet152 successfully filters out a significant portion of sections that do not contain the initiation site, while YOLOv5 eliminates the remaining sections. In Fig. 12c, the ability of YOLOv5 to detect both halves of a LOF defect from two distinct images is demonstrated. This highlights the ability of the model to overcome the inherent constraint imposed by blindly cropping the images of the entire fractured surface.

3) DISTANCE MEASUREMENT RESULTS

In accordance with the discussion in section III-E, the distance of the initiation site was performed by a two-step process of computing the center of the specimen and computing the intersection with the outer rim. This process was performed on 61 images where YOLOv5 detected anomalies. It is noteworthy that among the four excluded images, one was due to the initiation site being missed by ResNet152, while the remaining three were a result of YOLOv5 failing to detect the initiation site.

To verify the computed measurements, first the accuracy of the computed center location of each specimen was evaluated
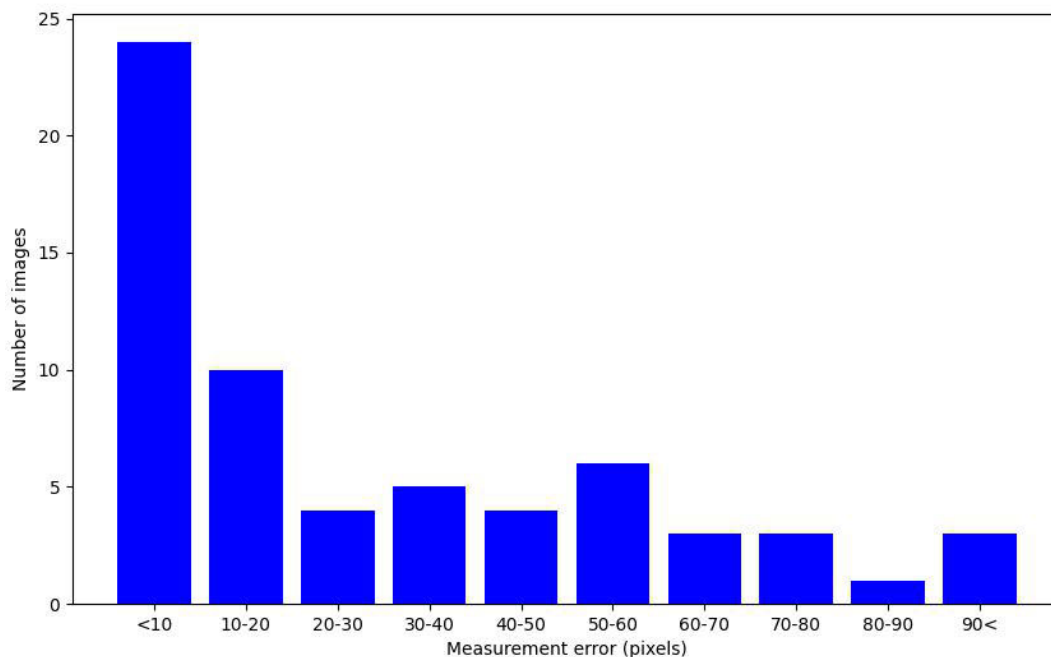
FIGURE 13. Difference in pixels between the manual and autonomous measurements of the distance between the anomalies and the outer rim of the specimen. 63 anomalies were detected in 61 out of 65 original images of the entire fractured surface. Excluding the 3 errors that exceeded 90 pixels the average measurement error of the remaining 60 anomalies was 25.1 pixels with a standard deviation of 24.2 pixels. This corresponds to an average error of 33.7 μm. The average relative error obtained for these measurements was 19.2% with a standard deviation of 21.4%.
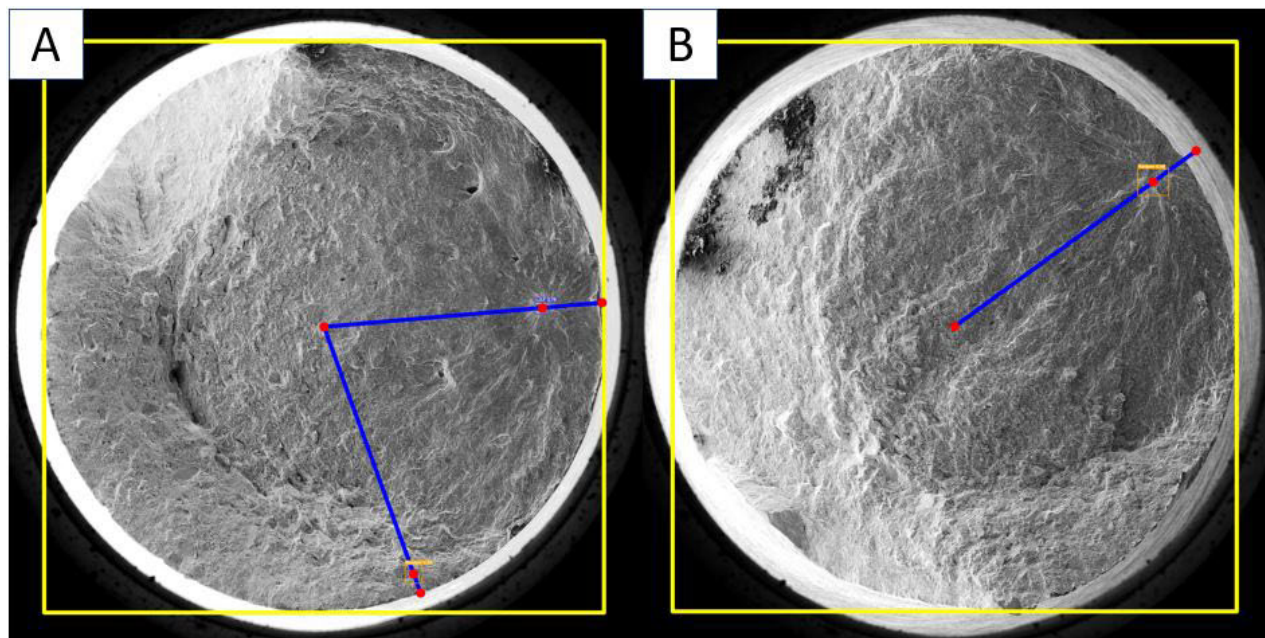

FIGURE 14. Examples of the detected center of the specimen and the intersection with the outer rim. (a) an example image of visually satisfying detection of the specimen center and the outer rim intersection for both the initiation site and false detection by YOLOv5. and (b) An example image of visually inaccurate detection of the intersection point of the outer rim.

with respect to the manually measured value. The average Euclidean distance between the two values was found to be 14.2 pixels with a standard deviation of 12 pixels. Note that among the 61 images examined, there were 7 instances where this distance exceeded 30 pixels. Recall that the full images are of size 4096 × 4096 pixels where the size of each pixel is approximately 1.34 μm. Thus, the discrepancy of 14.2 pixels corresponds to 19 μm.
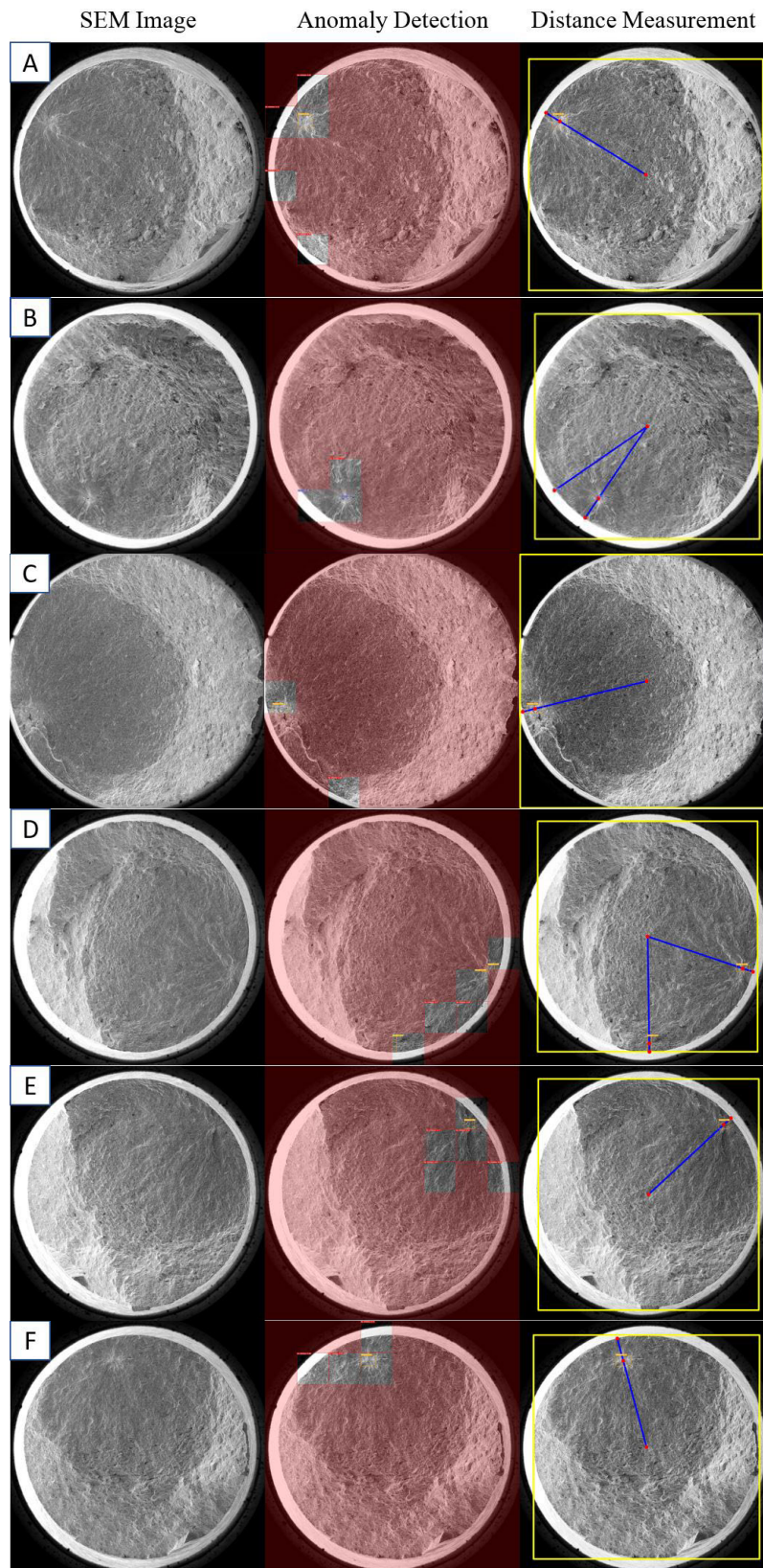
FIGURE 15. Six additional images of the entire fractured surface of the specimens. In the left column is the original SEM image, In the middle column is the resulting detection after cropping the image and using both the ResNet and the YOLO models. In the right column is the corresponding distance measurements between the defect and the outer rim of the specimen.

In Fig. 13, a comparison between manual and autonomous measurements of the distance to the surface is depicted, illustrating the absolute measurement error. It may be observed from the plot that for 24 anomalies the difference between the autonomous and manual measurement was less than 10 pixels, corresponding to less than $13.5\mu m$. In the plot, 3 out of 63 measurements were excluded since the error for these three specimens exceeded 90 pixels. For the remaining anomalies, the average absolute error of the measurements was 25.1 pixels, with a standard deviation of 24.2 pixels. This translates to an average error of $33.7\mu m$ with a standard deviation of $32.5\mu m$. The relative error obtained for these measurements was 19.2% with a standard deviation of 21.4%.

In Fig. 14, an instance of accurately detected intersection points is presented. Notably, this image demonstrates visually satisfactory points identified for both the LOF initiation site and the falsely detected random initiation site. Conversely, Fig.14b provides an example of an inaccurate intersection point detection.

Importantly, it should be emphasized that beyond errors in the detection of the center of the specimen, another key factor contributing to the disparity between manual and autonomous measurements was the variance between the center of the anomaly detected by YOLOv5 and the center manually identified. As a consequence, it is crucial to individually analyze each image to validate the findings of the model. However, it is important to note that this validation process is purely visual and does not entail the need for manual measurements. Hence, the implementation of the proposed model alongside manual validation is notably more time-efficient than performing all measurements manually.

V. CONCLUSION AND FUTURE WORK

This study employed ResNet152 and YOLOv5 models to detect fatigue crack initiation sites in SEM images of fractured surfaces of SLM Ti-6Al-4V specimens. Additionally, standard computer vision techniques were utilized to compute the distance between the initiation site and the specimen surface. Following a thorough analysis of the obtained results, the following conclusions may be drawn:

- ResNet152 was able to filter out approximately 97.6% of the sections from the test set that did not contain the initiation site, effectively isolating and retaining the relevant portions of the images related to the initiation site. This outcome highlights the capability of ResNet152 in efficiently distinguishing and eliminating non-relevant areas, contributing to the accuracy and efficacy of the analysis.
- Even with the utilization of a relatively small dataset for training YOLOv5, as shown in Table 3, the model displayed great potential and delivered promising results. It achieved a detection rate of 82.3% for identifying initiation sites in the test set. This outcome underscores the

effectiveness and promising performance of YOLOv5, despite the limitations imposed by the dataset size.
- The hybrid model employed in this study, combining ResNet152 and YOLOv5, demonstrated successful detection of the initiation site in a majority of the images within the dataset. Specifically, out of the 65 images examined, the model successfully detected the initiation site in 60 images resulting in a success rate of 92.3%. Note that in three images the initiation site was missed by YOLOv5. In two images the initiation site was missed by ResNet, however, in one of these images a related nearby anomaly was detected by YOLOv5.
- A distance measurement algorithm, designed to determine the distance between the initiation site and the outer surface of the specimen, was proposed. This computation was executed on 61 out of 65 images encompassing the complete fractured surface, where a total of 63 accurate ROI were identified. Among them, for three ROI, the discrepancy between manual and autonomous measurements exceeded 90 pixels, rendering the algorithm ineffective for these cases. Consequently, for the remaining 95.4% of measurements, the average difference between manually and autonomously measured distances was 25.1 pixels, with a standard deviation of 24.2 pixels. This equates to an average absolute error of $33.7\mu m$ and a standard deviation of $32.5\mu m$. Additionally, the mean relative error for these measurements stood at 19.2% with a standard deviation of 21.4%.

One may claim that the determination of fatigue crack initiation points is only part of the study of fatigue strength; Material damage tolerance, which encompasses the material's ability to resist growth and propagation of existing cracks under cyclic loading, is indeed an essential aspect of fatigue analysis. This includes understanding how the material behaves once a crack has been initiated and how it tolerates further damage without catastrophic failure.

While this study doesn't directly delve into damage tolerance, the developed methodologies and tools could be adapted for future research in this area. Expanding the dataset, employing image augmentation techniques, and creating models to measure defect characteristics lay the groundwork for obtaining data to establish criteria for critical defects. This could be pivotal in predicting fatigue life based on defect geometrical features and location, with potential application in damage tolerance analysis. High-resolution images and detailed defect analysis offer insights into defects that may develop into cracks and their potential threat to structural integrity. Future research may use information from tests and automated fractographic analyses to characterize the threat level of each defect detected through non-destructive testing before the part enters service. This could lead to a more comprehensive understanding of the fatigue behavior of AM materials and potentially

contribute to the development of more robust and reliable AM products.

In order to enhance the findings of this study, there are potential avenues for improvement. One approach is to expand the dataset by including images of additional specimens with different pixel sizes. Another approach is utilizing image augmentation techniques. Augmentation can be employed to further diversify the dataset and improve the robustness of the model. This may also assist in optimizing the distance-measuring algorithm. Moreover, it may be beneficial to develop additional models specifically tailored to measure different characteristics of defects, such as area and circumference, using higher-resolution images of the initiation site. These models could contribute to a more comprehensive analysis and potentially yield detailed results, thus achieving a detailed autonomous fractographic analysis.

Subsequently, this autonomous tool can be utilized to create a comprehensive dataset for machine learning-based fatigue performance assessment, with a specific focus on critical defect features. Such a tool may provide the ability to predict the fatigue life of AM materials under high-cycle fatigue (HCF) loading using mapped microstructural defect data. In the future, the various facets of critical defects may contribute to establishing the algorithmic foundation for nondestructive testing (NDT) and fatigue life evaluation of additive manufacturing products.

This research proposes the creation of an autonomous tool for detecting fatigue failure sources and determining their location through fractographic analysis. By correlating the failure source's location with additional geometrical characterization and the fatigue life measured during testing, a dataset could be generated in the future to predict fatigue life based on critical defect characterization. The robustness of such machine-learning models, enhanced by a diverse dataset, allows for precise identification and measurement of critical defect location, providing crucial information regarding fatigue life under HCF conditions.

HCF is characterized by a large number of cycles (typically in the range between 10^4 to 10^7 cycles or more) at lower stress levels, where the material experiences elastic deformation. In contrast, low-cycle fatigue (LCF) involves fewer cycles at higher stress levels, leading to plastic deformation. The distinct nature of HCF and LCF is rooted in their different deformation mechanisms and stress levels.

The characterization of HCF life using the tools proposed in this study specifically targets the influence of microstructural defects, commonly found in AM materials, on the initiation and propagation of fatigue cracks under high-cycle, low-stress conditions. Additionally, it's worth noting that ongoing development is underway for an additional tool designed to ascertain the size, location, morphology, and type of critical defect sources identified by the automated

tool proposed in this study. This information plays a crucial role in the classification of critical defects for HCF failure. While the current focus is on HCF, the methodology can be extended or modified for very high cycle fatigue (VHCF) analysis.

APPENDIX
ADDITIONAL IMAGES AND DETECTIONS

This appendix presents six additional images depicting the complete fractured surface of the specimen and corresponding detections. In Fig. 15a, a specimen featuring a random initiation site is displayed. Although the ROI and intersection points were successfully detected, the measurement of the specimen's center is imprecise. Moving to Fig. 15b, a specimen with a LOF defect is showcased. The ROI detection and accurate measurement of the distance to the surface were accomplished. However, an additional low-probability false detection occurred.

Fig. 15c illustrates a specimen with another LOF defect. Despite the LOF section being missed by ResNet, YOLOv5 managed to detect a portion of the ROI. In Fig. 15d, a specimen with a random initiation site is presented. Although the ROI was split during the cropping process, it was identified separately in each section.

Lastly, Fig. 15e and Fig. 15f showcase two additional specimens containing random initiation sites. In both scenarios, the ROI was effectively detected, and the distance to the surface was accurately measured.

ACKNOWLEDGMENT

REFERENCES

[1] K. V. Wong and A. Hernandez, "A review of additive manufacturing," *Int. Scholarly Res. Notices*, vol. 2012, p. 10, Jan. 2012.

[2] S. Gorsse, C. Hutchinson, M. Gouné, and R. Banerjee, "Additive manufacturing of metals: A brief review of the characteristic microstructures and properties of steels, Ti–6Al–4V and high-entropy alloys," *Sci. Technol. Adv. Mater.*, vol. 18, no. 1, pp. 584–610, Dec. 2017.

[3] T. Pereira, J. V. Kennedy, and J. Potgieter, "A comparison of traditional manufacturing vs additive manufacturing, the best method for the job," *Proc. Manuf.*, vol. 30, pp. 11–18, Jan. 2019.

[4] J. K. Watson and K. M. B. Taminger, "A decision-support model for selecting additive manufacturing versus subtractive manufacturing based on energy consumption," *J. Cleaner Prod.*, vol. 176, pp. 1316–1322, Mar. 2018.

[5] P. Edwards and M. Ramulu, "Fatigue performance evaluation of selective laser melted Ti–6Al–4V," *Mater. Sci. Eng., A*, vol. 598, pp. 327–337, Mar. 2014.

[6] S. Leuders, M. Thöne, A. Riemer, T. Niendorf, T. Tröster, H. A. Richard, and H. J. Maier, "On the mechanical behaviour of titanium alloy TiAl6V4 manufactured by selective laser melting: Fatigue resistance and crack growth performance," *Int. J. Fatigue*, vol. 48, pp. 300–307, Mar. 2013.

[7] P. Li, D. H. Warner, A. Fatemi, and N. Phan, "Critical assessment of the fatigue performance of additively manufactured Ti–6Al–4V and perspective for future research," *Int. J. Fatigue*, vol. 85, pp. 130–143, Apr. 2016.

[8] C. Matias, A. Diskin, O. Golan, A. Garkun, and E. Storkin, "Ti–6Al–4V additive manufacturing measure of quality according to fatigue crack initiation vs. crack propagation," in *Proc. Int. Council Aeronaut. Sci.*, 2022, pp. 1–17.

[9] C. Qiu, N. J. E. Adkins, and M. M. Attallah, "Microstructure and tensile properties of selectively laser-melted and of HIPed laser-melted Ti–6Al–4V," *Mater. Sci. Eng., A*, vol. 578, pp. 230–239, Aug. 2013.

[10] S. Bhat and R. Patibandla, "Metal fatigue and basic theoretical models: A review," *Alloy Steel-Properties Use*, vol. 22, pp. 203–236, Jan. 2011.

[11] H. Jia, H. Sun, H. Wang, Y. Wu, and H. Wang, "Scanning strategy in selective laser melting (SLM): A review," *Int. J. Adv. Manuf. Technol.*, vol. 113, nos. 9–10, pp. 2413–2435, Apr. 2021.

[12] E. Malekipour and H. El-Mounayri, "Common defects and contributing parameters in powder bed fusion AM process and their classification for online monitoring and control: A review," *Int. J. Adv. Manuf. Technol.*, vol. 95, nos. 1–4, pp. 527–550, Mar. 2018.

[13] N. Sanaei and A. Fatemi, "Defects in additive manufactured metals and their effect on fatigue performance: A state-of-the-art review," *Prog. Mater. Sci.*, vol. 117, Apr. 2021, Art. no. 100724.

[14] G. S. Rohrer, "Grain boundary energy anisotropy: A review," *J. Mater. Sci.*, vol. 46, no. 18, pp. 5881–5895, Sep. 2011.

[15] I. S. Ramírez, F. P. G. Márquez, and M. Papaelias, "Review on additive manufacturing and non-destructive testing," *J. Manuf. Syst.*, vol. 66, pp. 260–286, Feb. 2023.

[16] S. P. Lynch and S. Moutsos, "A brief history of fractography," *J. Failure Anal. Prevention*, vol. 6, no. 6, pp. 54–69, Dec. 2006.

[17] A. Mohammed and A. Abdullah, "Scanning electron microscopy (SEM): A review," in *Proc. Int. Conf. Hydraul. Pneumatics*, Băile Govora, Romania, 2018, pp. 77–85.

[18] E. L. Ritman, "Current status of developments and applications of micro-CT," *Annu. Rev. Biomed. Eng.*, vol. 13, no. 1, pp. 531–552, Aug. 2011.

[19] K. R. Shetty, V. S. Soorinje, and P. Dsouza, "Deep learning for computer vision: A brief review," *Int. J. Adv. Res. Sci., Commun. Technol.*, vol. 2022, pp. 450–463, Mar. 2022.

[20] C. Veiga and J. P. Davim, "Properties and applications of titanium alloys: A brief review," *Rev. Adv. Mater. Sci.*, vol. 32, no. 2, pp. 133–148, Dec. 2012.

[21] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[22] P. Jiang, D. Ergu, F. Liu, Y. Cai, and B. Ma, "A review of YOLO algorithm developments," *Proc. Comput. Sci.*, vol. 199, pp. 1066–1073, Jan. 2022.

[23] S. L. Sing, J. An, W. Y. Yeong, and F. E. Wiria, "Laser and electron-beam powder-bed additive manufacturing of metallic implants: A review on processes, materials and designs," *J. Orthopaedic Res.*, vol. 34, no. 3, pp. 369–385, Mar. 2016.

[24] M. Javaid and A. Haleem, "Additive manufacturing applications in medical cases: A literature based review," *Alexandria J. Med.*, vol. 54, no. 4, pp. 411–422, Dec. 2018.

[25] P. Bubna, M. P. Humbert, M. Wiseman, and E. Manes, "Barriers to entry in automotive production and opportunities with emerging additive manufacturing techniques," SAE Tech. Paper 2016-01-0329, 2016.

[26] S. Salifu, D. Desai, O. Ogunbiyi, and K. Mwale, "Recent development in the additive manufacturing of polymer-based composites for automotive structures—A review," *Int. J. Adv. Manuf. Technol.*, vol. 119, nos. 11–12, pp. 6877–6891, Apr. 2022.

[27] A. Uriondo, M. Esperon-Miguez, and S. Perinpanayagam, "The present and future of additive manufacturing in the aerospace sector: A review of important aspects," *Proc. Inst. Mech. Eng., G, J. Aerosp. Eng.*, vol. 229, no. 11, pp. 2132–2147, Sep. 2015.

[28] B. Blakey-Milner, P. Gradl, G. Snedden, M. Brooks, J. Pitot, E. Lopez, M. Leary, F. Berto, and A. du Plessis, "Metal additive manufacturing in aerospace: A review," *Mater. Des.*, vol. 209, Jan. 2021, Art. no. 110008.

[29] N. S. Rossini, M. Dassisti, K. Y. Benyounis, and A. G. Olabi, "Methods of measuring residual stresses in components," *Mater. Des.*, vol. 35, pp. 572–588, Mar. 2012.

[30] C. Li, Z. Y. Liu, X. Y. Fang, and Y. B. Guo, "Residual stress in metal additive manufacturing," *Proc. CIRP*, vol. 71, pp. 348–353, Jan. 2018.

[31] N. L. Loh and K. Y. Sia, "An overview of hot isostatic pressing," *J. Mater. Process. Technol.*, vol. 30, pp. 45–65, Jan. 1992.

[32] Z. Tong, X. Ren, J. Jiao, W. Zhou, Y. Ren, Y. Ye, E. A. Larson, and J. Gu, "Laser additive manufacturing of FeCrCoMnNi high-entropy alloy: Effect of heat treatment on microstructure, residual stress and mechanical property," *J. Alloys Compounds*, vol. 785, pp. 1144–1159, May 2019.

[33] R. O. Ritchie, "Mechanisms of fatigue-crack propagation in ductile and brittle solids," *Int. J. Fract.*, vol. 100, no. 1, pp. 55–83, 1999.

[34] S. Xu, J. Chen, W. Shen, R. Hou, and Y. Wu, "Fatigue strength evaluation of 5059 aluminum alloy welded joints considering welding deformation and residual stress," *Int. J. Fatigue*, vol. 162, Sep. 2022, Art. no. 106988.

[35] H. Masuo, Y. Tanaka, S. Morokoshi, H. Yagura, T. Uchida, Y. Yamamoto, and Y. Murakami, "Effects of defects, surface roughness and HIP on fatigue strength of Ti–6Al–4V manufactured by additive manufacturing," *Proc. Struct. Integrity*, vol. 7, pp. 19–26, Jan. 2017.

[36] Y. Murakami, *Metal Fatigue: Effects of Small Defects and Nonmetallic Inclusions*. New York, NY, USA: Academic, 2019.

[37] S. Pommier, "Variability in fatigue lives: An effect of the elastic anisotropy of grains?" *Eur. Struct. Integrity Soc.*, vol. 31, pp. 321–340, Jan. 2003.

[38] A. Li, S. Baig, J. Liu, S. Shao, and N. Shamsaei, "Defect criticality analysis on fatigue life of L-PBF 17–4 PH stainless steel via machine learning," *Int. J. Fatigue*, vol. 163, Oct. 2022, Art. no. 107018.

[39] S. Tsopanidis, R. H. Moreno, and S. Osovski, "Toward quantitative fractography using convolutional neural networks," *Eng. Fract. Mech.*, vol. 231, May 2020, Art. no. 106992.

[40] H. Qassim, A. Verma, and D. Feinzimer, "Compressed residual-VGG16 CNN model for big data places image recognition," in *Proc. IEEE 8th Annu. Comput. Commun. Workshop Conf. (CCWC)*, Jan. 2018, pp. 169–175.

[41] F. Erdem and U. Avdan, "Comparison of different U-Net models for building extraction from high-resolution aerial imagery," *Int. J. Environ. Geoinformatics*, vol. 7, no. 3, pp. 221–227, Dec. 2020.

[42] S. Tsopanidis and S. Osovski, "Unsupervised machine learning in fractography: Evaluation and interpretation," *Mater. Characterization*, vol. 182, Dec. 2021, Art. no. 111551.

[43] H. Larochelle, Y. Bengio, J. Louradour, and P. Lamblin, "Exploring strategies for training deep neural networks," *J. Mach. Learn. Res.*, vol. 10, no. 1, pp. 1–12, Jan. 2009.

[44] H. Na, J. Yoo, and H. Ki, "Prediction of surface morphology and reflection spectrum of laser-induced periodic surface structures using deep learning," *J. Manuf. Processes*, vol. 84, pp. 1274–1283, Dec. 2022.

[45] B. Dey, D. Goswami, S. Halder, K. Khalil, P. Leray, and M. A. Bayoumi, "Deep learning-based defect classification and detection in SEM images," *Proc. SPIE*, vol. 12053, Jun. 2022, Art. no. PC120530Y.

[46] Z. Guo, C. Wang, G. Yang, Z. Huang, and G. Li, "MSFT-YOLO: Improved YOLOv5 based on transformer for detecting defects of steel surface," *Sensors*, vol. 22, no. 9, p. 3467, May 2022.

[47] *ASTM e466-15 Standard practice for Conducting Force Controlled Constant Amplitude Axial Fatigue Tests of Metallic Materials*, Standard ASTM E466-15, Metals—Mechanical Testing; Elevated And Low-temperature Tests; Metallography, West Conshohocken, 2021.

[48] T. Szandała, "Review and comparison of commonly used activation functions for deep neural networks," *Bio-Inspired Neurocomputing*, vol. 903, pp. 203–224, Jan. 2021.

[49] M. Wang, S. Lu, D. Zhu, J. Lin, and Z. Wang, "A high-speed and low-complexity architecture for softmax function in deep learning," in *Proc. IEEE Asia–Pacific Conf. Circuits Syst. (APCCAS)*, Oct. 2018, pp. 223–226.

[50] Q. Wang, Y. Ma, K. Zhao, and Y. Tian, "A comprehensive survey of loss functions in machine learning," *Ann. Data Sci.*, vol. 9, no. 2, pp. 187–212, Apr. 2022.

[51] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.

[52] E. S. Gedraite and M. Hadad, "Investigation on the effect of a Gaussian blur in image filtering and segmentation," in *Proc. ELMAR*, Sep. 2011, pp. 393–396.

[53] L. Ding and A. Goshtasby, "On the Canny edge detector," *Pattern Recognit.*, vol. 34, no. 3, pp. 721–725, Mar. 2001.

[54] S. Kotsiantis, D. Kanellopoulos, and P. Pintelas, "Handling imbalanced datasets: A review," *GESTS Int. Trans. Comput. Sci. Eng.*, vol. 30, no. 1, pp. 25–36, 2006.

[55] J. M. Johnson and T. M. Khoshgoftaar, "Survey on deep learning with class imbalance," *J. Big Data*, vol. 6, no. 1, pp. 1–54, Dec. 2019.

RO'I LANG is currently pursuing the dual B.Sc. degree in mechanical engineering and in mathematics with the Department of Mechanical Engineering and Mechatronics, School of Engineering, Ariel University, Israel.

His academic excellence has earned him the prestigious Top Ariel Scholarship for outstanding students. He conducted his research under the guidance of Dr. Mor Mega at the Fracture and Fatigue Research Laboratory, Ariel University. His research interests include leveraging machine learning algorithms to address fracture and fatigue-related challenges, with a specific emphasis on understanding the mechanical patterns and behavior of additive manufacturing metals.

OR HAIM ANIDJAR received the B.Sc. degree in computer science from Bar Ilan University, Israel, in 2015, and the M.Sc. and Ph.D. degrees in computer science and applied mathematics from Ariel University, Israel, in 2019 and 2021, respectively. In 2022, he was a Postdoctoral Fellow with the Ariel Cyber Innovation Center (ACIC), Ariel University. He forms a Lab Member at the Kinematics and Computational Geometry (K&CG) Laboratory; ACIC; the Data Science and Artificial Intelligence Research Center (DSAIRC); and the Architectural AI Research Laboratory (AAIRL), Ariel University. Formerly, he was the Founder, the CEO, and the Chief Data Scientist of Libonea.AI. He is currently a Lecturer and a Senior Faculty Member with the School of Computer Science, Ariel University. By using deep-learning methods, his research lies at the intersection between natural language processing (NLP) and speech recognition (SR), especially in multidisciplinary methods, such as language models, speaker change detection and diarization, speaker identification, and anomaly detection in acoustic and textual datasets.

MOR MEGA is currently a dedicated Faculty Member with the Department of Mechanical Engineering and Mechatronics, School of Engineering, Ariel University, Israel, holding the position of a Lecturer, which is equivalent to an Assistant Professor. She serves as the Head for the Fracture and Fatigue Research Laboratory (FFRL) and is responsible for leading the classical mechanics specialization course. Her research interests include investigating the fracture and fatigue behavior of composite laminates and additive manufacturing components.

• • •