12-10-2021

Deep multi-modal U-net fusion methodology of infrared and ultrasonic images for porosity detection in additive manufacturing

Christian E. Zamiela
Mississippi State University, czamiela@gmail.com

Deep multi-modal U-net fusion methodology of infrared and ultrasonic images for porosity detection in additive manufacturing

By

Christian E. Zamiela

Approved by:

Linkan Bian (Major Professor)
Wenmeng Tian
Haifeng Wang
T. C. Falls
Linkan Bian (Graduate Coordinator)
Jason M. Keith (Dean, Bagley College of Engineering)

A Thesis
Submitted to the Faculty of
Mississippi State University
in Partial Fulfillment of the Requirements
for the Degree of Master of Science
in Industrial Engineering
in the Department of Industrial and Systems Engineering

Mississippi State, Mississippi

December 2021

Name: Christian E. Zamiela

Date of Degree: December 10, 2021

Institution: Mississippi State University

Major Field: Industrial Engineering

Major Professor: Linkan Bian

Title of Study: Deep multi-modal U-net fusion methodology of infrared and ultrasonic images for porosity detection in additive manufacturing

Pages of Study: 53

Candidate for Degree of Master of Science

We developed a deep fusion methodology of non-destructive (NDT) in-situ infrared and ex-situ ultrasonic images for localization of porosity detection without compromising the integrity of printed components that aims to improve the Laser-based additive manufacturing (LBAM) process. A core challenge with LBAM is that lack of fusion between successive layers of printed metal can lead to porosity and abnormalities in the printed component. We developed a sensor fusion U-Net methodology that fills the gap in fusing in-situ thermal images with ex-situ ultrasonic images by employing a U-Net Convolutional Neural Network (CNN) for feature extraction and two-dimensional object localization. We modify the U-Net framework with the inception and LSTM block layers. We validate the models by comparing our single modality models and fusion models with ground truth X-ray computed tomography images. The inception U-Net fusion model localized porosity with the highest mean intersection over union score of 0.557.

Key words: Multi-Modal Deep Sensor Fusion, Convolutional Neural Network, Porosity Detection, Infrared, Ultrasonic

TABLE OF CONTENTS

LIST OF TABLES iv

LIST OF FIGURES v

CHAPTER

I. INTRODUCTION 1

II. LITERATURE REVIEW 5

2.1 Metal Part Fabrication and Porosity in AM 5
2.2 Non-destructive Sensing for Porosity Prediction 7
2.3 Machine Learning and Deep Learning in AM 9
2.4 Review of Deep Learning based Image Fusion 12
2.5 Contributions of Our Work on Fusion Model of Ultrasonic and Infrared Data 15

III. METHODOLOGY: SMART FUSION OF INFRARED AND ULTRASONIC IMAGES USING U-NET ARCHITECTURE 16

3.1 Characteristics of Sensor Fusion 16
3.1.1 Pyrometer Data 17
3.1.2 Ultrasonic Data 19
3.1.3 X-Ray Computed Tomography Data 20
3.2 Formation of the Problem 21
3.3 Outline of Algorithm 25
3.4 Details of Algorithm 26
3.4.1 **Step 1:** Create an infrared layer image $\boldsymbol{p}_\ell^*$ from pyrometer melt pool data. 26
3.4.2 **Step 2:** Feed infrared $\boldsymbol{p}_\ell^*$ layer images and ultrasonic $\boldsymbol{a}_\ell$ layer images into separate encoder branches $E_{\ell,n}^v$. 28
3.4.3 **Step 3:** Concatenate the latent space from each branch $E_{\ell,n}$. . . . 30
3.4.4 **Step 4:** Decode latent space $D_{\ell,n}$ and apply skip connections to each modality after each decoder block. 30
3.4.5 **Outcome:** Probability Porosity Image 31

3.5 Inception Block and LSTM Modifications to U-Net Framework 32
3.5.1 Inception Block U-Net 32
3.5.2 Convolutional LSTM Block U-Net 35

IV. CASE STUDY: DEEP MULTI-MODAL FUSION POROSITY DETECTION MODEL FOR TI-6AL-4V MANUFACTURED CYLINDERS 37

4.1 Experimental Setup 37
4.2 Results of Sensor Fusion prediction of porosity for Ti-4AL-4V cylinders . 40
4.2.1 Traditional U-Net vs. Inception and LSTM modified U-Net Models 40
4.2.2 Combined U-Net fusion model vs. thermal and ultrasonic U-Net single modality models 42
4.2.3 Tuning the number of encoder-decoder processing blocks 44

V. CONCLUSION 47

REFERENCES 50

LIST OF TABLES

4.1 Processing parameters for manufactured cylinders 38
4.2 Average metric results combined traditional, inception, and LSTM U-Net models . 41
4.3 Average metric results for combined, thermal single modality, ultrasonic single modality inception U-Net models 43
4.4 Average metric results for inception U-Net fusion model with four, three, and two encoder-decoder blocks 45

LIST OF FIGURES

1.1 Ultrasonic and infrared data fusion process using a two-branch convolutional U-Net encoder-decoder. (A) Image data collection and infrared layer creation. (B) Down sampling encoder. (C) Latent space concatenation. (D) Up sampling decoder with skip connections. (E) Outcome: Probability of porosity feature map. 2

3.1 Signal capturing systems of infrared and ultrasonic data. (a) LENS fabrication chamber with pyrometer camera. (b) Post-processing ultrasonic transducer with function generator and oscilloscope. 17

3.2 A single ultrasonic layer of the fabricated part with fabrication tracks overlay compared to the size of the size of melt pool pyrometer images. (a) Ultrasonic fabrication image with fabrication tracks and melt pool locations. (b) Pyrometer melt pool image 18

3.3 Post-processing XCT System with a high voltage generator, manipulator, X-ray beam tube, and flat panel detector. 21

3.4 Deep multi-Modal U-net fusion framework. (Step 1:) Create an infrared layer image $\boldsymbol{p}_{\ell}^{*}$ from pyrometer melt pool data. (Step 2:) Feed infrared $\boldsymbol{p}_{\ell}^{*}$ layer images and ultrasonic $\boldsymbol{a}_{\ell}$ layer images into separate encoder branches $E_{\ell,n}^{v}$. (Step 3) Concatenate the latent space from each branch $E_{\ell,n}$. (Step 4:) Decode latent space $D_{\ell,n}$ and apply skip connections to each modality after each decoder block. (Outcome:) Probability of porosity map $\hat{\boldsymbol{y}}_{\ell,n}$. 25

3.5 Infrared layer image creation steps. (a) Peak temperature and meltpool locations. (b) K-means smoothed fabrication layer of temperature values. (c) Smoothed edges with Gaussian convolutional signal blur and part mask. 26

3.6 Traditional encoder block for U-Net framework 28

3.7 Traditional decoder block for U-Net framework 28

3.8 Inception Block with four parallel channels of varying length and kernel sizes. . . 33

3.9 U-Net Inception encoder-decoder blocks for activation map downsampling and upsampling. (a) Inception encoder block. (b) Inception decoder block. 34

3.10 Convolutional LSTM Block with the function gate visualization. 36

4.1 Sample images 7, 33, 40, 78, and 124 for U-Net sensor fusion. (Top) Ultrasonic layer images. (Middle) Infrared Layer images. (Bottom) XCT layer images. 39

4.2 Samples 13, 18, and 28 image predictions of traditional, inception, and LSTM U-Net thermal and ultrasonic fusion models. (a) XCT. (b) Traditional U-Net fusion predictions. (c) Inception U-Net fusion predictions. (d) LSTM U-Net fusion predictions. 42

4.3 Samples 13, 18, and 28 image predictions of inception U-Net fusion model, inception U-Net thermal single modality model, and inception U-Net ultrasonic single modality model. (a) XCT. (b) Inception U-Net fusion model. (c) Inception U-Net ultrasonic single modality model. (d) Inception U-Net thermal single modality model. 44
4.4 Samples 13, 18, and 28 image predictions of inception U-Net fusion model with four, three, and two encoder-decoder blocks (a) XCT. (b) Four encoder-decoder blocks inception U-Net fusion model. (c) Three encoder-decoder blocks inception U-Net ultrasonic single modality model. (d) Two encoder-decoder blocks inception U-Net thermal single modality model. 46

CHAPTER I

INTRODUCTION

The objective of this study is to develop a non-destructive testing (NDT) method based on the fusion of ultrasonic structural data and thermal history data for porosity detection in additive manufacturing (AM). Additive manufacturing is a process that joins materials layer by layer to make three-dimensional (3D) objects, which leads to advantages in (1) incorporating complexity of part design, (2) saving time in prototyping, and (3) reducing weight through structural design. Unfortunately, many additively manufactured parts are plagued with structural abnormalities and internal defects. Lack of fusion between successive layers of printed metal in the AM process is still a common issue due to porosity which is undetectable from the outer surface of the fabricated parts [5, 13]. There is a critical need for NDT of porosity because we can determine quality without compromising the structure of the fabricated part. NDT sensing methods can be divided into x-ray computed tomography (XCT), ultrasonic and acoustic, and in-situ thermal imaging. Industrial XCT sensing creates a 3D image to inspect interior features to verify printing turned out as designed [7]. XCT is limited by cost, processing time, versatility, and penetration into larger structures. Ultrasonic and acoustic sensing techniques measure the wave impedance through structures to reveal structural flaws. The ultrasonic and acoustic images are limited in imaging resolution and the wave penetration strength through a structure. Lastly, in-situ thermal images capture the melt

pool history, which is deterministic of microstructure during the printing process [40]. The melt pool behavior is related to the formation of porosity because of the insufficient overlap between successive layers but is limited by only being an indirect representation of the internal structure of AM fabricated part [5, 13]. The gap in the research is extracting reliable information practically and pragmatically from non-destructive in-situ thermal-mechanical dynamics and ex-situ ultrasonic structural data for improved anomaly detection.

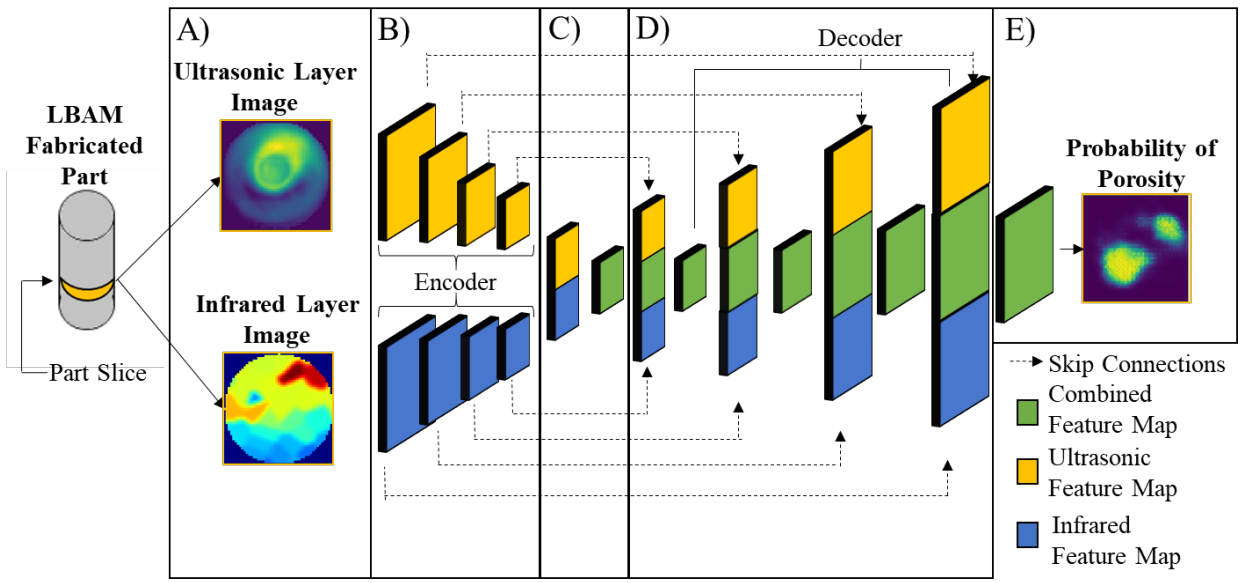


Figure 1.1: Ultrasonic and infrared data fusion process using a two-branch convolutional U-Net encoder-decoder. (A) Image data collection and infrared layer creation. (B) Down sampling encoder. (C) Latent space concatenation. (D) Up sampling decoder with skip connections. (E) Outcome: Probability of porosity feature map.

We propose a new deep learning sensor fusion approach based on the U-Net architecture to combine thermal history images with ultrasonic images to improve the prediction accuracy of

porosity. The use of thermal history and ultrasonic data is advantageous because we can extract valuable information from both in-process and post-process data. The first challenge of fusing data collected by each data collection process is spatially aligning both data streams in the fabricated part. Additionally, the second challenge will be autonomous feature extraction from each modality through the mapping from input data to ground-truth output data rather than handcrafted feature extraction. We propose a new sensor fusion methodology that transforms thermal history data from a small image representation of point to a layer-wise representation of the fabrication layer and then employs a two-branch traditional U-Net encoder-decoder architecture to extract spatial information each sensor modality separately. The steps of this methodology can be seen in Figure 1.1 and are described as follows:

A. We model the thermal history in a new way by converting the melt pool images of each layer into infrared layer images by feeding the highest temperature values in each image through K-means regression and Gaussian convolutional signal smoothing.

B. We are the first to feed in-situ infrared layer images and ex-situ ultrasonic layer images through separate U-Net convolutional encoder (downsampling) block branches to obtain latent space representation.

C. We concatenate the latent space of infrared layer image encoder and ultrasonic layer image encoder.

D. We feed the combined latent space through a U-Net convolutional decoder (upsampling) block while including skip connection between geometrically similar feature maps of the encoder.

E. We obtain the probability of porosity label map representation of each layer mapped to output ground truth XCT layer images.

We modify the U-Net fusion methodology with inception blocks and LSTM layers and compare it with the traditional U-Net fusion model. The inception blocks have been shown to improve localization when objects appear in different sizes within images. The LSTM layer processes 3D

dimensional information as a function of time. The current observations are processed with a specified number of previous observations. The sensor fusion methodology will be compared to thermal and ultrasonic single modality models to show the value in sensor fusion. Practitioners can adapt the proposed sensor fusion methodology of thermal history and ultrasonic data for the NDT of large-scale industrial AM processes to detect abnormalities. This methodology can also be generalized for predictive maintenance in other manufacturing mechanisms. The remainder of the paper is organized as follows. Chapter 2 reviews the background studies in the literature; Chapter 3 provides the mathematical modeling of the proposed methodology; Chapter 4 discusses the case study used to validate the proposed methodology; and Chapter 5 provides concluding remarks and directions of possible future work.

CHAPTER II

LITERATURE REVIEW

This section introduces the related background knowledge of AM processes, non-destructive sensing techniques, machine learning in AM, and deep learning image fusion. Specifically, information about porosity, computed tomography, thermal monitoring, ultrasonic processing, and deep image segmentation are discussed in detail in this section.

2.1 Metal Part Fabrication and Porosity in AM

Laser-based additive manufacturing (LBAM) has become a promising method for developing metallic parts due to its flexibility and ability to create complex parts. Traditional manufacturing methods remove necessary material to achieve the final geometry while AM adds material in a layer-by-layer fashion. LBAM focuses on applying a laser energy source to metallic power to melt the material to fabricate the current layer, and cross-sectional layers [23]. Laser Engineered Net Shaping (LENS) is a directed energy deposition method with feedback control properties that help maintain the product's quality and desired mechanical properties. The parts fabricated by this technique have been shown to have no compositional degradation, and thus LENS has shown to be a promising method of metal additive fabrication [3]. Titanium alloy is fed through multiple tubes that allow mixing of elemental powders, and powder efficiency is reported to be 80 percent for a laser power of 2500-3000 W [3]. Powder deficiency and melt pool distribution are quality issues

that have prevented the universal microstructure of LBAM fabricated parts in many industries [26]. LBAM fabricated parts are prone to defects, and many of these defects are internal. Non-destructive sensors allow the exploration of the defect without the destruction of the part.

Porosity is an important quality issue for AM parts. Porosity is demonstrated as small voids, and micro-cracks within a metal part that can commonly lead to unexpected breaks and increases in residual stress [13]. The objective of porosity detection methods is to improve the quality of printed parts in AM. Lack of fusion and keyholing porosity commonly occurs during printing and can be analyzed through thermal history. Lack of fusion commonly occurs when the melt pool is not large enough. Melt pool size and scan pattern are critical factors in the lack of fusion porosity. For good fusion to take place, there must be a complete overlap between melt pools. Space between semi-circles of heat conduction to the surrounding areas of the part leads to a lack of fusion [5, 39]. Laser power, scan velocity, and laser beam profile are critical factors in keyhole porosity formulation. The keyholing effect results from generating a V-shaped melt pool and vaporizing elements within the melt pool that often leads to entrapped gas [13]. Increases in the energy density cause the melt pool to transition from a conduction mode to a keyhole mode. To avoid keyhole porosity is to reduce laser power or to increase the scanning speed [27]. Keyholing commonly occurs when making a turn in the printing process because the laser slows down, and the weld is very narrow [27]. The infrared images that capture the thermal history during the fabrication process improve AM build quality. While we can generalize how porosity occurs in the fabrication process, there are still limitations to how much the melt pool can reveal about the printing process [21].

2.2 Non-destructive Sensing for Porosity Prediction

In-situ process monitoring is employed to detect defects and abnormalities in the printing process. The melt pool determines the microstructure in the AM parts [40]. Infrared sensors capture a two-dimensional matrix of temperature values in fast succession. The melt pool is the superheated region of molten metal that typically appears as a spherical-shaped droplet that moves at the traverse speed [13]. The melt pool is where the horizontal and vertical fusion between layers/tracks occurs. Heat penetration to lower layers and heat conduction between recently printed metals require the melt pool to reach a high temperature. Melt pool behavior is related to the formation of the porosity in part because of insufficient overlap between successive melt pools [26]. Product quality can significantly deteriorate, and thus optimal processing parameters need to be identified. Thermal history sensors are limited by only being an indirect representation of the AM fabrication process for predicting the internal structure of LBAM parts.

Post-manufacturing sensors are employed to locate structural deformities by projecting waves or determining the level of radiation absorption. Ultrasonic sensors are employed for non-destructive porosity detection to measure the acoustic impedance of waves through a metal part. Measuring the opposition that a system presents to acoustic flow reveals structural damages. The waves can propagate into the material and cause the elastic excitation of cracks and pores that cause the amplitude to increase because there is less material inhibiting the waves [22]. There are advantages in versatility and direct estimation of filling amplitude within a part. Based on different ultrasonic scanning techniques and equipment types, ultrasonic imaging can be mainly classified as A-scan, B-scan, and C-scan [20]. A-scan can be used to calculate the location of the internal defects of a specific material or human body as a function of time. Compared to the A-scan technique,

B-scans provide a cross-sectional view of the test specimen [20], which is generated by using a computational imaging algorithm to display the signal amplitude in grayscale or color scale for the entire cross-section at a specific scanning moment. Similarly, ultrasonic C scan imaging, which is the ultrasonic imaging technique applied in this paper, is the computational rearrangement of signal array datasets that were collected by using the B-mode scanning method at the specimen's surface multiple times in a 2D mesh-grid manner [20]. By applying the C-scan technique, the cross-sectional views along the perpendicular direction to the B-scan can be easily observed. Nowadays, ultrasonic C-scan techniques are of great interest to be applied into the modern NDT field to detect and localize internal material defects [41]. Although, the ultrasonic wave penetration of parts is limited by the strength of the waves and loses velocity as the waves travel through fabricated parts. Also, ultrasonic sensors must be used on objects that have a high amount of open space because the waves must have a path to propagate through a structure [41]. Therefore, sensors that have to observe structural anomalies without losing observation strength and do not require a path to travel are advantageous.

Computed tomography is a computerized x-ray imaging procedure in which a narrow beam of x-ray is aimed at a part and quickly rotated around the part. Cross-sectional images or slices of the part are formed and contains more detail than a single image. Successive slices are stacked together to form a three-dimensional image of the part that allows for easy identification of voids of space. In a CT scan, metal behaves similarly to human tissue in that it absorbs the radiation causing it to appear as the lighter regions in a 3D image, and voided space will appear as the darker regions. Industrial CT is used to inspect interior features to verify printing turn out as designed [7]. Residual powder and internal features can be seen through volume graphics with a CT scan.

However, industrial CT requires more costly higher voltages machines for large dense objects due to the limited radiation penetration at lower voltages [7]. In addition, the absorption properties of a given material can significantly affect the quality. CT X-ray scans are a traditional porosity characterization tool regarded as the ground truth label in this study.

2.3 Machine Learning and Deep Learning in AM

Machine learning has been extensively used to process various data types to improve the additive manufacturing process. Wang in [45] reviews how machine learning is used in AM in the design phase, fabrication process phase, and production phase. The design phase consists of materials and topology, where the design phase is a combination of materials to create a metamaterial that can be challenging and exhaustive to formulate. Machine learning can expedite the discovery of materials that improve AM parts by decreasing weight and improving structural integrity. Improved prediction of material properties and optimization using machine learning [15, 28] allowed for the structural improvements [46] and improved design through metamaterial data generation [19]. Wu in [46] combines performance evaluations with structural evolution to construct meta-atoms with specified properties. Neural network models find the optimal configuration of meta-atoms. Meta-atoms delicately designed subwavelength structures. Furthermore, Ma in [19] represents the relationship and design of metamaterials in the probabilistic model. Data labeling is performed in a semi-supervised learning strategy, and further data collection and generation is performed with a deep generative model to accelerate the design, characterization, even discovery in the domain of metamaterials.

The fabrication process phase can employ machine learning to optimize process parameters [11], powder spreading characterization, and in-process defect monitoring. Process parameters can be optimized through physics-based simulation that reveals the part features that happen during the fabrication process, such as melt pool geometry, keyhole, and microstructures [12]. The growth of AM has led to increasingly more complicated techniques that can be difficult to model. Jiang in [11] employs a deep neural network model which is proposed to predict the connection between paths in different process parameters. Changing process parameters can achieve different goals for the fabricated part, and the paths at which the part is fabricated can be optimized to improve part goals. On the other hand, Liu in [16] uses a physics-informed, data-driven model instead of machine setting parameters to predict porosity levels because many machines are different. We can look at the output of the machine's printing process to better understand porosity.

In-process monitoring can help to reduce AM issues such as cracks, pores, delamination, distortion, rough surface, lack of fusion, and process instability. These problems tend to happen with the fusion of successive layers. The various optical and acoustic sensors can improve real-time process improvement by applying deep learning and transfer learning machine learning models. In-situ monitoring of the melt pool has been extensively studied to determine how the morphology can lead to porosity [13, 33]. Khanzadeh in [13] develops a monitoring tool to detect anomalies in the microstructure of LBAM fabricated parts. The supervised learning methodology dichotomizes the melt pools to determine what characteristics lead to a lack of fusion between successive layers. Rao in [31] and Li in [14] apply multiple in-situ sensor approach to monitoring the AM process. A non-parametric Bayesian Dirichlet process mixture model [31] and new ensemble modeling [14] is compared to traditional ML approaches. Additionally, Shevchik in [35] proposes

highly sensitive acoustic data collected using fiber Bragg grating sensors and machine learning for process monitoring. Wavelet spectrograms of the signal and spectral clustering gather features labels for a deep convolutional neural network for in-situ real-time quality monitoring of AM process that requires minimal modifications to commercially available industrial machines. In-process monitoring thermal sensors are an indirect analysis of structural irregularities in AM fabricated parts and are limited by not considering post-fabrication changes that can be seen by post-manufacturing sensors[13].

Lastly, AM production phase can employ ML product quality control by assisting in the planning to determine the manufacturability of parts. Product quality control is vital to the development of AM because many manufacturing methods are inconsistent from machine to machine. Inconsistency is due to geometrical accuracy, relative density, process stability, and mechanical properties [24]. The challenge with quality control in many AM parts is that many of the defects are internal. In recent years, non-destructive methods of determining quality, such as CT and ultrasonic sensors, have been employed for abnormality detection. In [48], A hierarchical synthesis CNN model is used to improve the sparse-view CT reconstruction. Undersampled reconstructions are converted into high-fidelity estimates to train the neural network. This study examines the global-scale pore evolution in post-processing. Likewise, in [34], a deep neural network is employed for the correlation of in-situ acoustic and optical microscopy with ultrasonic properties for porosity content detection. Micrographic images have also been used for defect detection through the use of machine learning [9]. The limitations of ex-situ post-processing detection of porosity are that it does not consider fabrication process irregularities and cannot make changes during processing to improve fabricated parts [22]. The limitations of in-situ and

ex-situ porosity detection methods with machine learning suggest the need for a study that takes advantage of both in-process and post-processing data for porosity detection. This data would provide valuable information on the printing process and structural integrity.

2.4 Review of Deep Learning based Image Fusion

Image data fusion seeks to combine two-dimensional matrix data from the different image capturing techniques. Image fusion categories are multi-focus, multi-exposure, medical image, visible/infrared, remote sensing [17]. Methods based on super-resolution have been shown to have the most promise in general image fusion. Typical transforms using multi-scale decomposition-based (MSD) fusion include image pyramids, wavelet-based transforms, multi-scale geometric transforms, and spatial filtering-based decomposition. In the fusion phase, the activity level of source images is measured by the decomposed coefficients abed on some specific pixel or window-based approaches and followed by a predesigned fusion rule like choose-max or weighted average are adopted to combine coefficients[17]. The conventional methods reviewed above lack effective image representation that is not characterized by the complex mapping of the relationship between input and target images[17, 47]. Convolutional Neural Networks (CNN) have achieved significant advances in visual recognition problems compared with conventional methods. CNN avoids handcrafted design and is more likely to achieve better performance. The success is mainly due to the strong capability of DL models in characterizing the complex relationship between different signals. A CNN-based method for measuring the local similarity between two image patches demonstrates the superiority of the CNN-based method over traditional methods[17, 47]. The advantages of deep learning are summarized below.

1. Can automatically extract the most useful features from data to overcome the difficulty of manual design.
2. Can characterize very complex relationship between input and targeting data.
3. Community provides some potential image representation approaches which could be useful to the study of image fusion.

Typical convolutional neural networks produce a classification label of an image, but for many problems, the desired output requires localization. Deep learning image segmentation was first introduced in [18] and has shown to be an effective method of partitioning images into multiple segments. The U-Net architecture was an improvement upon [18] by Ronneberger in [32]. Convolutional layers gather spatial information from images in windowing motion and feed into a nonlinear activation function. Sub-sampling or pooling layers are often inserted between each convolution layer for resolution reduction and reduced overfitting. Progressively more refined feature extraction at every layer and less computational overhead [36, 44]. Encoder-decoder networks can use a convolutional layer to encode the image into latent representation to create an output map of the original image. Encoder-decoder-based models such as U-Net for 2D segmentation use convolutional layer to reduce input images and use up-sampling convolutional transpose layer to systematically produce a classification map the same size as the original image [36]. The critical advantage of U-Net architecture is the introduction of skip connections which allows the network to propagate contextual information to higher resolution layers parallel to the decoder layers from the encoder [32]. At all the levels, the convolutional layer's output, before the encoder's downsampling operation, is transferred to the decoder with geometrically similar dimensions. These feature maps are then concatenated with the output of the up-sampling

operation, and the concatenated feature map is propagated to the successive layers. The advantages of U-Net are as follows:

1. Draws on additional information due to skip connections to allow for improved training and improved prediction accuracy.
2. Can process smaller dataset due to dense connections that cause a regularizing effect preventing overfitting.
3. The network does not have any fully connected dense layers and only uses the valid part of each convolution containing the pixels, which provides full context of input images.

Recent studies have proposed changes to U-Net to improve further performance given challenging problems with limited datasets. In [10], a U-Net has been employed for multi-modal biomedical image segmentation with the addition of an inception-like block that captures local and global information by employing multiple parallel convolutional filters on the previous layer. Each convolutional filter has a different kernel size to capture the preferred information from images with uneven distributions of information. The filters are then concatenated before being fed into the next block. Arbella in [1] implements temporal information into the U-net structure with a convolutional LSTM function in the encoder for Microscopy cell segmentation. Spatial-temporal encoding allows the model to make predictions based on the frame-to-frame differences. Past cell appearance can allow the model to track the movement of cells. Additional literature for changes to the U-Net architecture are a gated multi-modal fusion network [2], a pixel transpose convolutional network [8], and hyper-dense connectivity [6].

2.5 Contributions of Our Work on Fusion Model of Ultrasonic and Infrared Data

We contribute a sensor fusion methodology that addresses the challenges of combining data from two distinctly different sensors for quality predictions of LBAM fabricated parts. The critical technical contributions of our work in this regard are summarized as follows

1. Our work develops a layer-wise infrared thermal history image of the fabricated part using the non-parametric K-means regression and Gaussian convolutional smoothing function. This allows for the ultrasonic and infrared thermal history to be spatially aligned within the LBAM fabricated part.
2. Our work is the first to combine in-situ thermal images with ex-situ ultrasonic images for porosity detection in LBAM. Sensor fusion enables practitioners better understanding of the quality of the printing fabricated parts to improve industrial uses of AM.
3. Our methodology extends the traditional U-Net architecture by implementing inception block and convolutional LSTM functions in U-Net architecture for sensor fusion. Inception blocks have advantages in locating objects of different sizes in images, and convolutional LSTM functions use spatial-temporal information as a 3D representation of the fabricated part.
4. In addition to model architecture, our framework localizes porosity within each layer by creating a 3D representation of porosity throughout the entire part. Porosity localization enables our model to better capture the structural integrity of the fabricated part.

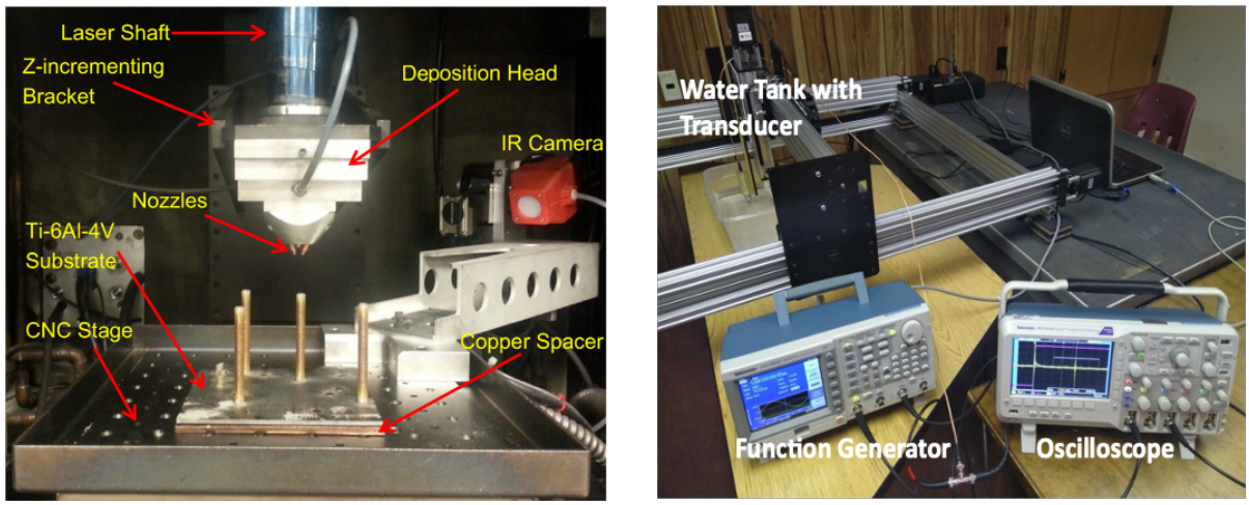
CHAPTER III

METHODOLOGY: SMART FUSION OF INFRARED AND ULTRASONIC IMAGES USING U-NET ARCHITECTURE

We model the fusion infrared and ultrasonic image data for porosity detection in additive manufacturing. By fusing both data types, we can predict the probability of porosity in each layer of the fabricated part. In this section, we discuss the data properties, formulate the problem, and modify U-Net architecture. Inception and LSTM U-Net modification are made to the U-Net structure to improve the localization of objects of different sizes and incorporate three-dimensional information into the current prediction.

3.1 Characteristics of Sensor Fusion

The data capturing systems of the two sensors can be seen in Figure 3.1. The infrared images are captured during the fabrication of LBAM parts by placing a pyrometer camera in the chamber. After fabrication, the ultrasonic images are captured using a transducer connected to a function generator and oscilloscope to send and capture ultrasonic waves as they propagate through the fabricated part. The characteristics of each sensor can be seen in Section 3.1.1 and 3.1.2.



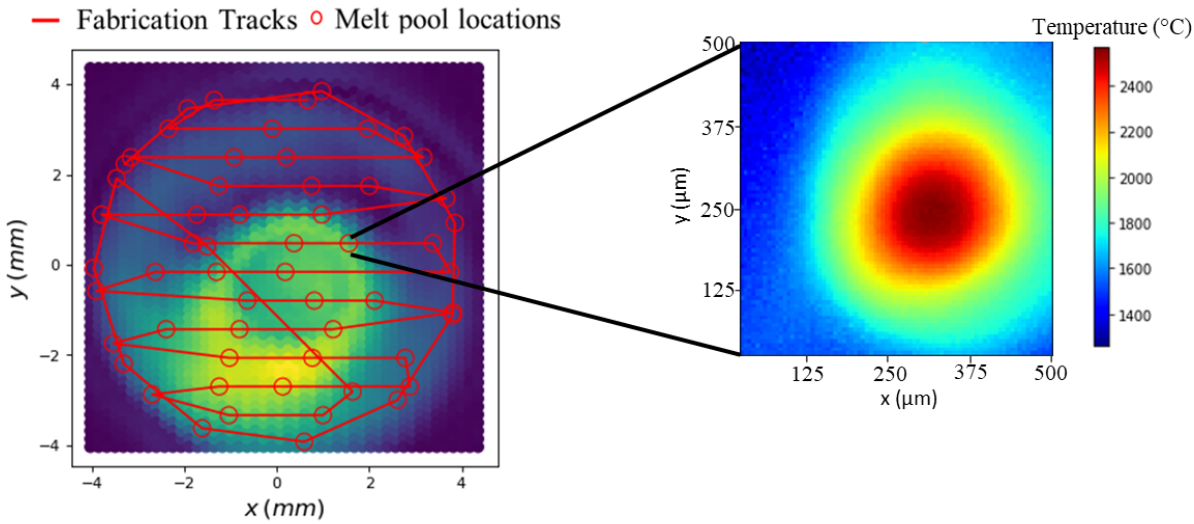
[a] [b]

Figure 3.1: Signal capturing systems of infrared and ultrasonic data. (a) LENS fabrication chamber with pyrometer camera. (b) Post-processing ultrasonic transducer with function generator and oscilloscope.

3.1.1 Pyrometer Data

A co-axial, dual-wavelength pyrometer camera is a non-destructive remote sensing technology that works to capture the emitted radiation from the surface without contact with surface [38]. The pyrometers act as a thermometer to measure the temperature of distant objects. Data acquisition from the pyrometer occurs during AM parts fabrication by placing the camera inside the fabrication chamber seen in Figure 3.1a. The camera is aligned with the laser of the LENS machine to capture the surface temperature of the melt pool of superheated metal alloy. The pyrometer exports 752×480 pixels images with an approximate pixel spacing of 6.45 microns and seven images per second. The pyrometer captures temperatures ranging from 1000 degrees Celsius and above.

Thermaviz software converts each image into a CSV format. The pyrometer images CSV files are cropped to 80 × 80 micros pixel images to contain only the melt pool temperature values greater than 1000 degrees Celsius. The removed temperature values are noise not related to the printing process.



[a] [b]

Figure 3.2: A single ultrasonic layer of the fabricated part with fabrication tracks overlay compared to the size of the size of melt pool pyrometer images. (a) Ultrasonic fabrication image with fabrication tracks and melt pool locations. (b) Pyrometer melt pool image

Figure 3.2a is an example of the melt pool image size in relation to the size of the fabricated part. A melt pool image is captured at various locations in the build process by the pyrometer. The lines and small circles in Figure 3.2a are the fabrication tracts and melt pool locations, respectively.

This data capturing process is limited by not creating an entire thermal history of each fabrication layer.

3.1.2 Ultrasonic Data

An ultrasonic transducer is a non-destructive remote sensing technique that emits and receives ultrasonic waves that pass through a structure. The increased amplitude of waves indicates voided space within a structure because waves can pass through with less resistance. Data acquisition occurs post-processing of fabricated part using an ultrasonic transducer controlled by Tektronix AFG3052C function generator to send 5M Hz sine wave and Tektronix DPO2024B oscilloscope to capture the ultrasonic acoustic waveform of AM sample seen in Figure 3.1b. A C-scan technique is employed to capture the amplitude of ultrasonic waves at each layer. Sine waves are sent periodically through the fabricated part attached to the bottom of a water tank stabilized by acoustic-damping-free plasticine and glue. A computational numerical control (CNC) station moves the transducer in a mesh-grid square of 61×61 scanning points with a spatial interval of $0.25mm$ over a $15mm \times 15mm$ area covers the Ti-6Al-4V sample.

The data signal are collected in a MATLAB program that removes the noise from scanning points with a fast Fourier and inverses fast Fourier algorithm. The ultrasonic wave amplitude is captured in 61x61 two-dimensional (2D) arrays of each layer within the fabricated part. The ultrasonic (2D) arrays are then cropped to 48×48 to remove noise and center array around the fabricated part. Figure 3.2a is an example of the ultrasonic layer image captured by the ultrasonic C-scan process. Contrary to the pyrometer data acquisition process, we capture the entire layer of information during the ultrasonic data acquisition process. However, this data capturing process is

limited by losing momentum as waves pass through the part, and the resolution of data provides minor detail of the fabricated part.

3.1.3 X-Ray Computed Tomography Data

Non-destructive XCT transmits high-energy electromagnetic waves with a very short wavelength. X-rays have enough energy to ionize matter, and electrons are rearranged within an atom. A Nikon X-Ray CT XT H225 machine is used to collect XCT data. Figure 3.3 provides a layout of the X-ray system that has an X-ray bean tube, a high voltage generator, a sample manipulator, and a flat panel detector. The machine uses a kV voltage of 180 and a uA current of 98 for the specimens. X-rays pass through the sample, forming a shadow image on the flat panel detector. The object of interest is placed on a CNC machine and rotated on an axis, which is a beneficial technique for capturing large objects. To create a quality scan of the object of interest, we align the 2D images using beam source, manipulator, and detector to perform a 3D reconstruction with a high-pass filtered back-projection for edge sharpening. Machined components are scanned by rotating the specimen around the central axis [25]. A porosity algorithm scans the gauge section using VGSTUDIO Max analysis and visualization software for industrial computed tomography. The specimen porosity information is converted to a CSV containing porosity location and descriptive variables such as sphericity, compactness, and radius[43]. The exact shape of the pore is unknown; we assume the pore is circular. Pores are reconstructed into a map of the pore within each fabrication layer of the object of interest. The XCT machine is limited by being expensive, slow processing time compared to alternative methods, and is ineffective for larger parts that cannot fit on a rotating axis.

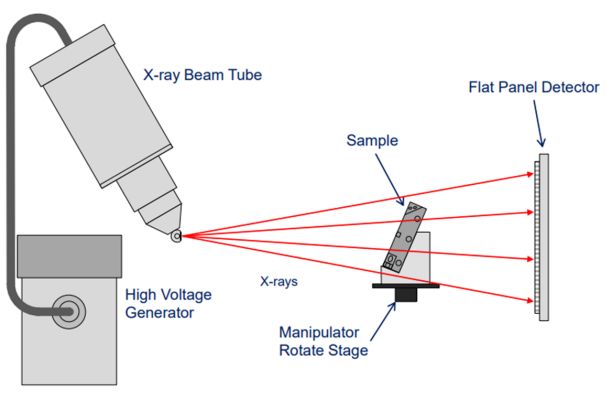


Figure 3.3: Post-processing XCT System with a high voltage generator, manipulator, X-ray beam tube, and flat panel detector.

3.2 Formation of the Problem

We model the peak temperature data fusion in pyrometer images and the entire image of ultrasonic data. Ultrasonic data represents the entire fabrication layer $\boldsymbol{a}_\ell$ but has a lower resolution. The peak temperature is the single value representation of the pyrometer image $\boldsymbol{p}_{\ell,\eta}$ and we smooth the peak value from each pyrometer image η in each layer ℓ to represent the entire fabrication layer. Figure 3.2 show that the pyrometer image covers a small area of the LBAM fabricated part. Maximum temperature in a pyrometer image is indicative of the thermal history in a pyrometer image because high temperatures result in higher penetration to lower layers and increased melt pool spread on current layer [5, 13]. $\boldsymbol{p}_{\ell,\eta}(i,j) = Max(T_{\ell,\eta}(i,j))$ is the maximum temperature for a pyrometer image at ℓth part layer and ηth image in the layer. T is the pyrometer image, and i, j are the coordinates of each pixel in a pyrometer image.

After extracting the maximum temperature from each pyrometer image in a layer we obtain a vector of representing each image ηth in a single layer $\boldsymbol{p}_\ell = (p_{\ell,1}, ..., p_{\ell,\eta})'$. We will then smooth

the peak temperature values using K-means regression and Gaussian convolutional function that can be seen in Section 3.4.1 to obtain a matrix $\boldsymbol{p}_\ell^*$ of thermal layer history seen as follows:

$$\boldsymbol{p}_\ell^* = \begin{bmatrix} p_\ell(0,0) & p_\ell(1,0) & \dots & p_\ell(u,0) \\ p_\ell(0,1) & p_\ell(1,1) & \dots & p_\ell(u,1) \\ \vdots & \vdots & \ddots & \vdots \\ p_\ell(0,v) & p_\ell(1,v) & \dots & p_\ell(u,v) \end{bmatrix} \tag{3.1}$$

The ultrasonic image data captured by the transducer represents the entire layer of fabricated part as a matrix seen as follows:

$$\boldsymbol{a}_\ell = \begin{bmatrix} a_\ell(0,0) & a_\ell(1,0) & \dots & a_\ell(u,0) \\ a_\ell(0,1) & a_\ell(1,1) & \dots & a_\ell(u,1) \\ \vdots & \vdots & \ddots & \vdots \\ a_\ell(0,v) & a_\ell(1,v) & \dots & a_\ell(u,v) \end{bmatrix} \tag{3.2}$$

Now both the infrared thermal history image matrix $\boldsymbol{p}_\ell^*$ and ultrasonic image matrix $\boldsymbol{a}_\ell$ represent the same space in the fabricated part. The next challenge is developing a smart way to fuse the data from each modality. For example, using a simple method such as choosing the maximum value from each modality would fuse the information, but, likely, this information would not indicate porosity. A deep learning convolutional neural network algorithm would map the pyrometer and ultrasonic image data to ground-truth porosity labels ensuring that the relevant information from each modality image is captured. Machine learning models are generally specified under fixed settings to process and map input variables to output. The U-Net convolutional neural network fusion model F for predicting porosity is

$$\hat{\boldsymbol{y}}_\ell = F(\boldsymbol{p}_\ell^*, \boldsymbol{a}_\ell) + \epsilon \tag{3.3}$$

In deep learning, the U-Net neural network layers feed into each other in succession in a feed-forward manner from the input variables $\boldsymbol{p}_{\ell}^{*}, \boldsymbol{a}_{\ell}$ to the n layers of the encoder E to the n layers of the decoder D. Additionally, we consider the case of two modalities, let $E_{\ell,n}^{1}$ and $E_{\ell,n}^{2}$ denote the nth encoder layer and the lth fabrication layer in branch 1 and 2, respectively. In general, the output of the nth network layer in the U-Net encoder branch w can be defined as follows:

$$E_{\ell,n}^{w} = H(E_{\ell,n-1}^{1}, E_{\ell,n-1}^{2}, E_{\ell,n-2}^{1}, E_{\ell,n-2}^{2}, ..., \boldsymbol{p}_{\ell}^{1*}, \boldsymbol{a}_{\ell}^{2}) \tag{3.4}$$

where H denotes the hidden layer of the U-Net encoder-decoder network architecture. After processing of each modality in separate branches of U-Net encoder, the final layers of each branch are concatenated. Processing continues with decoder layers n.

$$D_{\ell,n} = H(D_{\ell,n-1}, D_{\ell,n-2}, ..., [E_{\ell,n}^{1}, E_{\ell,n}^{2}]) \tag{3.5}$$

where n denotes network layer of the U-Net decoder and the square brackets represent concatenation network layers. The hidden layers of the U-Net architecture are further discussed in Step 2, 3, and 4 of the details of the algorithm seen in Section 3.4.2, 3.4.3, and 3.4.4, respectively.

Lastly, it would be more valuable to localize porosity in each layer rather than know if porosity exists in each layer. Deep learning image segmentation strategies could show where porosity is located in each layer through a sensor fusion encoder-decoder architecture. The outputs of image

segmentation strategies produce a matrix of labeled values predicting the probability of porosity in each fabrication layer.

$$\hat{\boldsymbol{y}}_{\ell,n} = \begin{bmatrix} D_{\ell,n}(0,0) & D_{\ell,n}(1,0) & \dots & D_{\ell,n}(u,0) \\ D_{\ell,n}(0,1) & D_{\ell,n}(1,1) & \dots & D_{\ell,n}(u,1) \\ \vdots & \vdots & \ddots & \vdots \\ D_{\ell,n}(0,v) & D_{\ell,n}(1,v) & \dots & D_{\ell,n}(u,v) \end{bmatrix} \tag{3.6}$$

where each location in $\hat{\boldsymbol{y}}_{\ell,n}$ produces a probability that porosity is present in (u, v) location in the current fabrication layer ℓ. The output feature maps compared to the ground truth output feature matrix and incorrect neurons are adjusted through back propagation.

3.3 Outline of Algorithm

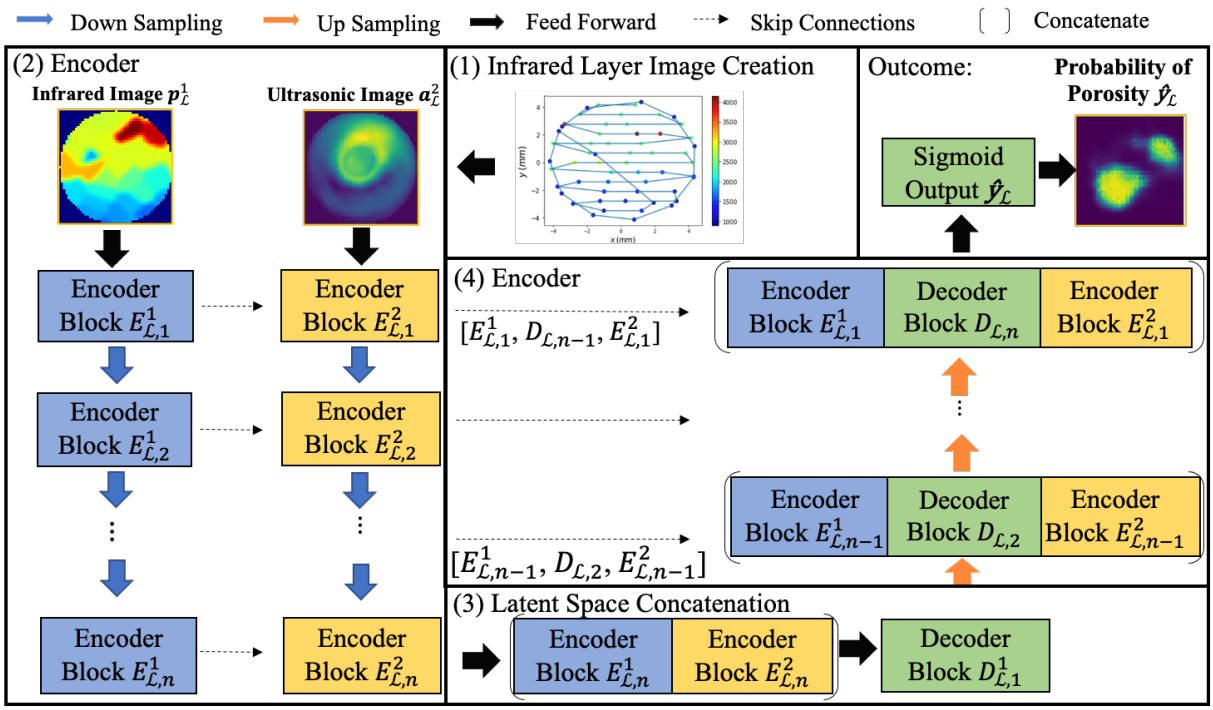


Figure 3.4: Deep multi-Modal U-net fusion framework. (Step 1:) Create an infrared layer image $\boldsymbol{p}^*_\ell$ from pyrometer melt pool data. (Step 2:) Feed infrared $\boldsymbol{p}^*_\ell$ layer images and ultrasonic $\boldsymbol{a}_\ell$ layer images into separate encoder branches $E^v_{\ell,n}$. (Step 3) Concatenate the latent space from each branch $E_{\ell,n}$. (Step 4:) Decode latent space $D_{\ell,n}$ and apply skip connections to each modality after each decoder block. (Outcome:) Probability of porosity map $\hat{\boldsymbol{y}}_{\ell,n}$.

The outline of the proposed methodology is summarized below.

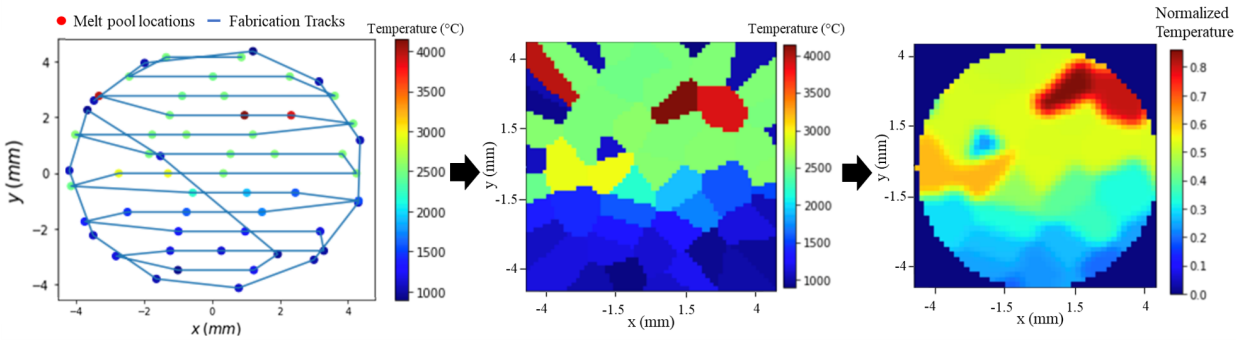
1. Create an infrared layer image $\boldsymbol{p}^*_\ell$ from pyrometer melt pool data.
2. Feed infrared $\boldsymbol{p}^*_\ell$ layer images and ultrasonic $\boldsymbol{a}_\ell$ layer images into separate encoder branches $E^v_{\ell,n}$.
3. Concatenate the latent space from each branch $E^v_{\ell,n}$.

4. Decode latent space $D_{\ell,n}$ and apply skip connections in Equation 3.13 to each modality after each decoder block.

5. Predict the probability of porosity $\hat{\boldsymbol{y}}_{\ell,n}$ compared to XCT layer images.

3.4 Details of Algorithm

3.4.1 Step 1: Create an infrared layer image $\boldsymbol{p}_{\ell}^{*}$ from pyrometer melt pool data.



[a] [b] [c]

Figure 3.5: Infrared layer image creation steps. (a) Peak temperature and meltpool locations. (b) K-means smoothed fabrication layer of temperature values. (c) Smoothed edges with Gaussian convolutional signal blur and part mask.

The transformation of pyrometer image data to infrared thermal history layer image expressed in Section 3.2 was achieved by using non-parametric K-mean regression. First, the maximum temperature of every infrared image in a single layer is extracted, along with the coordinates of each infrared image. An example of the maximum temperature value taken from each infrared image can be seen in Figure 3.5a. The values in Figure 3.5a are then smoothed using non-parametric K-means regression to form a matrix $\boldsymbol{p}_{l}^{*}$ to represent the entire fabrication layer with a coordinate

area of 9x9 with a spatial interval of 0.167. Figure 3.5b is an example of the infrared layer image formulated by the Kmean model. In Equation 3.7, we minimize the Euclidean distance values in a cluster and the nearest centroid.

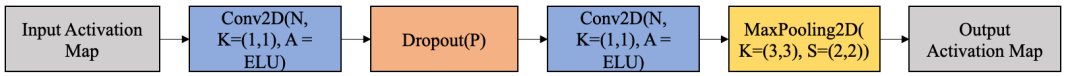
$$argmin \sum_{j=1} \sum_{\eta=1} ||\boldsymbol{p}_{\ell,\eta} - q_j||^2 \tag{3.7}$$

q is the jth centroid, and $\boldsymbol{p}$ is the ηth coordinate in the ℓth layer. After K-means regression, we have a 52×52 mesh grid expressed in Equation 3.1 as $\boldsymbol{p}_\ell$ and then we smooth the edges within the feature map with a Gaussian convolutional kernel function. Smoothing the edges between clusters will add additional noise to data. Equation 3.8 is a two dimensional Gaussian convolutional kernel function $G(i, j)$.

$$\begin{aligned} G(i,j) &= \frac{1}{2\pi\sigma} e^{-\frac{i^2+j^2}{2\sigma^2}} \\ \boldsymbol{p}_l^*(u,v) &= \sum G(i,j) * \boldsymbol{p}_l(u,v) \end{aligned} \tag{3.8}$$

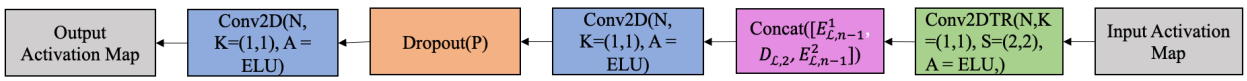
i, j is the size of the Gaussian kernel at which pixel values are combined and $*$ represents the convolution operations. G is the Gaussian kernel filter and $\boldsymbol{p}_\ell$ is the input signal. The Gaussian convolutional filter reduces the size of each dimensions original image by $i \times j$. The function reduces the size of $\boldsymbol{p}_\ell^*$ to a 48×48 mesh grid when the filter size $(i, j) = (2, 2)$. Figure 3.5c is an example of applying a Gaussian convolutional blur to K-Means feature map. This is a novel process of creating a thermal history layer representation in additive manufacturing.

3.4.2 Step 2: Feed infrared p_ℓ^* layer images and ultrasonic a_ℓ layer images into separate encoder branches $E_{\ell,n}^v$.



N = Number of Neurons, K = Kernel Size, A = Activation Function, S = Stride, P = Probability

Figure 3.6: Traditional encoder block for U-Net framework



N = Number of Neurons, K = Kernel Size, A = Activation Function, S = Stride, P = Probability

Figure 3.7: Traditional decoder block for U-Net framework

U-Net encoder processes infrared layer images $\boldsymbol{p}_\ell^*$ and ultrasonic layer images $\boldsymbol{a}_\ell$ separately in a two-branch structure seen in step 2 of Figure 3.4. Equation 3.3 express the deep learning fusion function of infrared and ultrasonic data, and Equation 3.4 expresses the encoding of two separate branches. The U-Net architecture is fully connected deep learning model formulated in Equation 3.4 and Equation 3.5. The input, hidden, and output layers feed into each other in succession. For image processing, a convolutional layer is employed to extract relevant features from the two-dimensional input matrix. Mathematically, the convolutional layer in a 2D-CNN is expressed in Equation 3.9.

$$E_{\ell,n} = b_n + \sum_i \sum_j k(i,j) * E_{\ell,n-1} \tag{3.9}$$

where $E_{\ell,n-1}$ is the output of the previous layer and $k(i,j)$ is the size of the kernel that passes over the output of the previous layer. The bias is a network parameter represented by b_n, and the symbol $*$ denotes the 2D convolution. Notable parameters in a convolutional layer are stride and padding. Stride controls the travel distance of the filter horizontally and laterally over the input, and padding controls the spaces of input not to be mapped by the filter. Common non-linear activation functions are sigmoid, rectified linear units (ReLUs), and exponential linear Units (ELUs). A convolution layer and ELU layer are jointly expressed in Equation 3.10.

$$E_{\ell,n}(u,v) = \begin{cases} b_n + \sum_i \sum_j k(i,j) * E_{\ell,n-1} & E_{l,n-1} \geq 0 \\ \alpha(e^{b_n + \sum_i \sum_j k(i,j) * E_{\ell,n-1}} - 1) & E_{l,n-1} \leq 0 \end{cases} \tag{3.10}$$

where α is a smoothing parameter which is usually between 0.1 and 0.3 and controls the scale of negative outputs. Negative outputs nudge weights and biases in the right direction compared to ReLU function that stops at 0. Convolutional layers are always followed by a non-linear activation layer. Dropout layer that randomly zeroes some of the elements of the input tensor with probability parameter ρ using samples from a Bernoulli distribution can be seen following the activation layer. This is a regularization technique that prevents the co-adaption of neuron in a network. The non-zeroed neuron are up scaled by factor of $\frac{1}{1-\rho}$ making the dropout layer an identity function and no change is made to input shape. A common operation in building a CNN is spatial pooling. It is popular to employ a max-pooling operation, which can bring some desirable invariance into the model. Equation 3.11 max-pooling kernel k operation downsamples the inputs based on the pool size parameter in the layer.

$$k = max(E_{\ell,n-1}(u,v)); u,v = 1,...,m \tag{3.11}$$

u and v represents the pixel location and m is the dimension size of the kernel that passes over each section. The left side U-Net architecture learns features and reduces the size of inputs by stacking layers for downsampling in Figure 3.6. U-Net encoder can be commonly seen with 4 stacks of layers.

3.4.3 Step 3: Concatenate the latent space from each branch $E_{\ell,n}$.

The two encoder branches $E^1_{\ell,n}$ and $E^2_{\ell,n}$ are concatenated to be further processed. In the proposed fusion model, the last layer of encoder operations for each modality is concatenated. Equation 3.5 combines the separate processing branches before we proceed with decoder operations. Combining each branch of U-Net architecture will provide more features for improved porosity prediction.

3.4.4 Step 4: Decode latent space $D_{\ell,n}$ and apply skip connections to each modality after each decoder block.

The decoder upsamples the combined latent space information from each modality. A convolutional transpose layer generates a feature map of spatial dimensions greater than that of the input. Equation 3.9 operation is employed to map to output greater than the input dimensions by filling the space between input values with zeros and fractionally slide kernel function over feature map smoothing output. The size of the output dimension map size γ for a convolutional transpose layer with a stride s greater than 1 with no padding is expressed as

$$\gamma = s(d-1) + k \tag{3.12}$$

k is the kernel size, and d is the input activation map dimensions. Each input feature is expanded to the size of the kernel in the convolution transpose layer. The expanded input feature fills the

shape of the output dimension map expressed in Equation 3.12 based on the stride parameter. Overlapping output feature map values are summed.

Furthermore, skip connections are unique to the U-net architecture for improving image processing. Skip connections concatenate encoder and decoder layers at specific points to add a regularization effect that helps to prevent overfitting of data. We employ skip connections between geometrically similar blocks of layers in the U-Net architecture for porosity detection.

$$D_{l,n} = H([E^1_{\ell,1}, D_{\ell,n-1}, E^2_{\ell,1}], ..., [E^1_{\ell,n-1}, D_{\ell,2}, E^2_{\ell,n-1}]), n = 1, ..., 5 \tag{3.13}$$

Equation 3.13 concatenates skip connection after every convolutional transpose layer. The skip connections can be visualized in Figure 3.4 and the connections create another path for back-propagation to take place. Section 3.4.2 convolutional layer and activation function operations in Equation 3.10 followed by the dropout layer operation are employed after the skip connection concatenation. The upsampling, skip connection concatenation, and feature learning network layers can be visualized in Figure 3.7.

3.4.5 Outcome: Probability Porosity Image

Localization of porosity within a layer is the output of the U-Net architecture. Equation 3.6 expresses the output of U-Net architecture as convolutional label map expressing probability locations of porosity. The last convolutional layer has a sigmoid activation function that activates the different features in map if an object exist in Equation 3.6. The sigmoid is expressed as

$$\hat{\boldsymbol{y}}_{l,n}(m,n) = \frac{1}{1 + e^{-D_{l,n}(m,n)}} \tag{3.14}$$

Equation 3.14 pushes values in the last feature map that are between 0 and 1 closer toward the extremes. Thus, the last feature map contains the probability of porosity from the current fabrication

layer input images. Knowing the likely locations of porosity in each layer is more valuable than know if porosity is located in a layer because we understand how the thermal history information may have affected the quality of the build.

3.5 Inception Block and LSTM Modifications to U-Net Framework

Modifications to the U-Net architecture can be seen in Sections 3.5.1 and 3.5.2. The traditional model can be overfitted to data and only locate objects within a certain number of pixels. The use of some alternate layers in the U-Net structure could show improvements in object localization. The inception blocks create a wider deep learning network with parallel branches processing the same sensor data. The changes in depth and kernel function of parallel branches in the inception block often result in better recognition when the objects appear in different sizes. The LSTM blocks create a three-dimensional representation by processing the previous successive layer information. The number of processing images fed into each layer help the model not to overfit the training data. The modified U-Nets are compared to the original U-Net predictions.

3.5.1 Inception Block U-Net

The inception block changes the encoder-decoder framework of the original model, but the outcome stays the same. Inception blocks employ multiple parallel branches to process the same input activation map. Equation 3.9 2D-CNN, Equation 3.11 max pooling, and batch normalization layers are stacked in different orders. The exploration and development of the inception block are sought after because often, objects being detected appear in a large variety of sizes. The number of pixels being occupied by the object can change, making choosing the kernel size difficult. The kernel sizes in the 2D-CNN layers change in each branch of the inception block for improved

object detection. The outputs of each line are concatenated and fed into the next inception block. The neuron in each block line is activated to predict the output best. Large kernels are for global information, and smaller kernels are for local information.

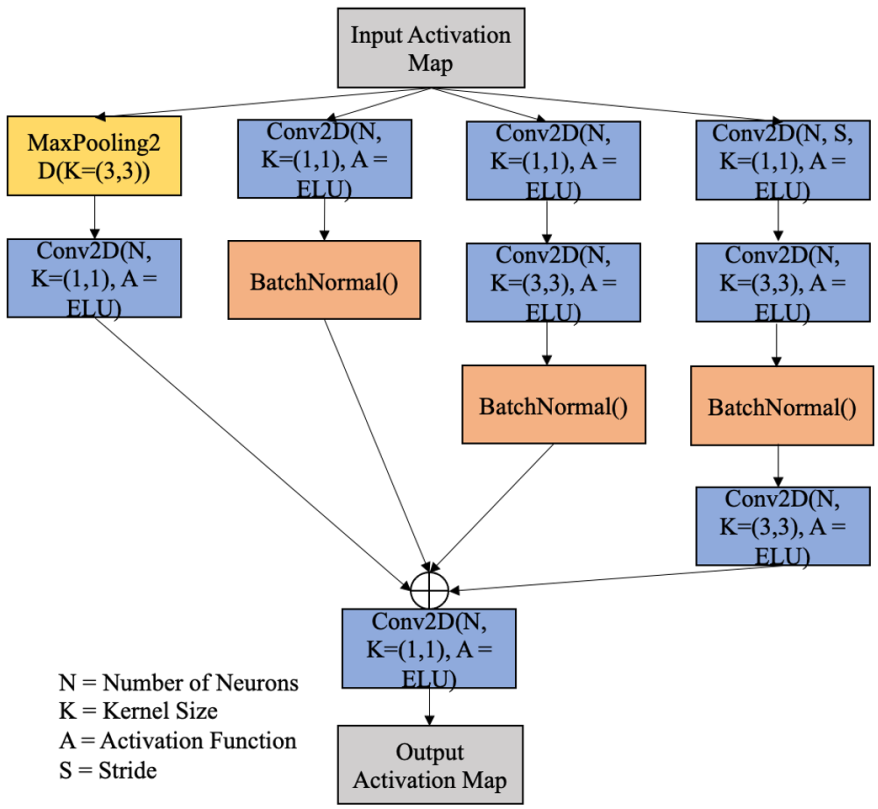


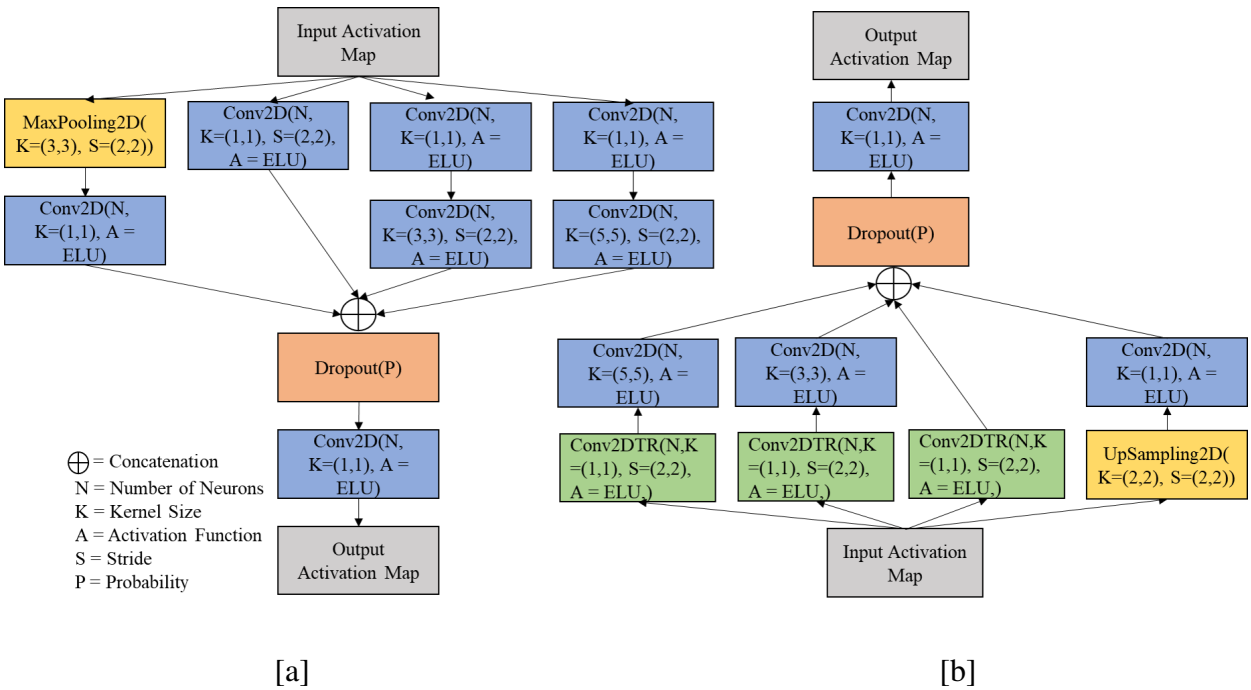
Figure 3.8: Inception Block with four parallel channels of varying length and kernel sizes.

Figure 8 input activation map flow into four separate channels with different varying levels of depth. In this block, the input dimensions are not changing but are being processed by different kernel sizes. Batch Normalization (BatchNormal) provides a regularization effect. Each channel is concatenated and processed by a final 2D-CNN layer before exiting the inception block. Consider

the inception block that processes the input activation map in four separate branches. Let z denote the branches of the inception block $E^1_{\ell,n}$, $E^2_{\ell,n}$, $E^3_{\ell,n}$, and $E^4_{\ell,n}$

$$E^z_{\ell,n} = H(E_{\ell,n-1}, [E^1_{\ell,n-2}, E^2_{\ell,n-2}, E^3_{\ell,n-2}, E^4_{\ell,n-2}], ..., E^1_{\ell,n-3}, E^2_{\ell,n-3}, E^3_{\ell,n-3}, E^4_{\ell,n-3}) \tag{3.15}$$

Square bracket represent concatenation and layers fed into each other in fed forward manner. Each branch of the inception block is processing the input activation map separately in Equation 3.15.



[a] [b]

Figure 3.9: U-Net Inception encoder-decoder blocks for activation map downsampling and upsampling. (a) Inception encoder block. (b) Inception decoder block.

The inception encoder-decoder blocks used in the modified U-Net architecture can be seen in Figure 3.9. Important to note that different channels have varying kernel sizes for improved

processing when objects appear in different sizes. The encoder and decoder block downsample and upsample the input activation maps with a stride s of (2, 2), which either reduces the size of the activation map by half or doubling the size of the activation map, respectively.

3.5.2 Convolutional LSTM Block U-Net

The Convolutional LSTM block changes the encoder framework and the skip connections of the original model, but the outcome stays the same. Also, the framework requires an additional dimension of input to process 3D objects. The convolutional LSTM function replaces one of the 2D-CNN layers (Equation 3.9) in the original U-Net architecture with ConvLSTM2D equations below. The LSTM function captures temporal information, and we regard the temporal information as depth in this study. We have a 3D object feeding the images into the Convolutional LSTM function as 2D frames. We are exploiting depth information to improve image segmentation with a Convolutional LSTM function expressed as:

$$
\begin{aligned}
i_t &= \phi(\theta_{xi} * x_t + \theta_{hi} * h_{t-1} + \theta_{ci} \circ c_{t-1} + b_i) \\
f_t &= \phi(\theta_{xf} * x_t + \theta_{hf} * h_{t-1} + \theta_{cf} \circ c_{t-1} + b_f) \\
c_t &= f_i \circ c_{t-1} + i_t \circ tanh(\theta_{xc} * x_t + \theta_{hc} * h_{t-1} + b_c) \\
o_t &= \phi(\theta_{xo} * x_t + \theta_{ho} * h_{t-1} + \theta_{co} \circ c_t + b_o) \\
\hat{\boldsymbol{y}}_{\ell,n} &= o_t \circ tanh(c_t)
\end{aligned}
$$

Convolutional LSTM layer in ConvLSTM2D is expressed as input gate i_t, forget gate f_t, cell memory c_t, output gate o_t, and output feature map $\hat{\boldsymbol{y}}_{\ell,n}$ and it can be visualized in Figure 3.10. t, ϕ, θ, $h_{(t-1)}$, and c_{t-1} is the current observation location, activation function, weight, previous cell output, and previous cell memory, respectively. $\circ$ represents element-wise multiplication,

and $*$ represents convolution operations. The input gate decides what to update, and forget gate decides what to throw away. These gates help regulate the flow of information in and out of the cell. The cell memory remembers values over arbitrary time intervals to update the current cell state. The output gate decides the next hidden state to make predictions on the output feature map. The placement convolutional LSTM layer at the start of the encoder block to include temporal information can be seen in Figure 3.10.

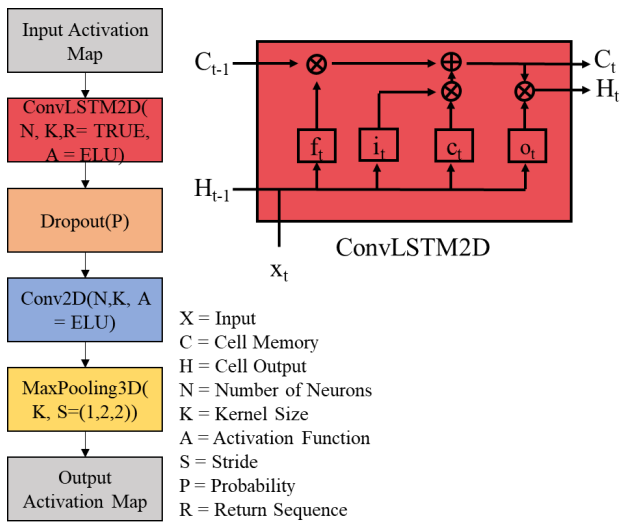


Figure 3.10: Convolutional LSTM Block with the function gate visualization.

CHAPTER IV

CASE STUDY: DEEP MULTI-MODAL FUSION POROSITY DETECTION MODEL FOR TI-6AL-4V MANUFACTURED CYLINDERS

4.1 Experimental Setup

Three tests are performed to determine the ideal structure and number of processing blocks for the U-Net model. The three tests are: (1) Compare the combined traditional U-Net fusion models vs. combined inception and LSTM modified U-Nets fusion models to test if the modification improves the localization of porosity. (2) Compare combined U-Net model vs. single modality U-Net models to test if sensor fusion improves the localization of porosity. (3) Tuning the number of encoder-decoder processing blocks to determine the ideal number of block layers. Finally, each model is validated with three metrics that assess the quality of porosity predictions from U-Net fusion models. The three selected metrics are (1) Mean intersection over union (IOU) Q_I, which measures the amount of mutual information between XCT labels and predicted porosity [30, 42]. (2) Structural similarity index Q_S, which measures the amount of structural information preserved in the fused image [4]. (3) Peak signal to noise ratio Q_P, which measures noise through the efficiency of compressors [29].

An OPTOMEC LENS 750 system that utilizes Ti-6Al-4V powder to manufacture cylinders additively. Ultrasonic and pyrometer data is collected for three Ti-6AL-4V cylinders for our case study to test the performance of U-Net models. The cylinders varied in height, ranging from

23.06mm to 27.12mm, but were printed with a constant height of 8mm. The optimized process parameters that were implemented when manufacturing cylinders are in Table 4.1. Each cylinder did not enter any post-processing heating chambers and was left as built.

Table 4.1: Processing parameters for manufactured cylinders

Processing Parameter	Cylinder 16	Cylinder 19	Cylinder 25
Laser Power (W)	300	350	400
Travel Speed (ipm)	40	30	50
Powder Flow Rate (rpm)	3	2.5	3
Layer Thickness (mm)	0.381	0.635	0.508
Hatch distance (mm)	0.508	0.635	0.508
Scan Pattern	Uni-directional	Zig-zag	Zig-Zag
Hatch Rotation (degrees)	180	180	180

Ultrasonic data is collected every 0.25mm compared to pyrometer data collected at each layer fabrication layer corresponding with layer thickness in Table 4.1 resulting more layers of data. It is necessary to have the same number of layers from each data type. The nearest ultrasonic layer data to fabrication layer are extracted for processing in sensor fusion model. Furthermore, XCT data is collected at various locations within the Ti-6AL-4V cylinders rather than layer by layer for pyrometer and ultrasonic data. It is important to have for ground truth XCT to have the same number of layers as input data. The nearest XCT porosity location to fabrication layer with a

interval of 0.05mm and probability threshold of 1.2 percent are extracted as labels for sensor fusion model. For image segmentation, the XCT data is transformed into a layer image using the radius of each pore. A total of 61, 41, and 54 images from each modality are collected from cylinders 16, 19, 25, respectively. 157 images in total from each modality and sample images can be seen in Figure 4.1. Note the infrared images are after step 1 in sensor fusion methodology, infrared layer creation.

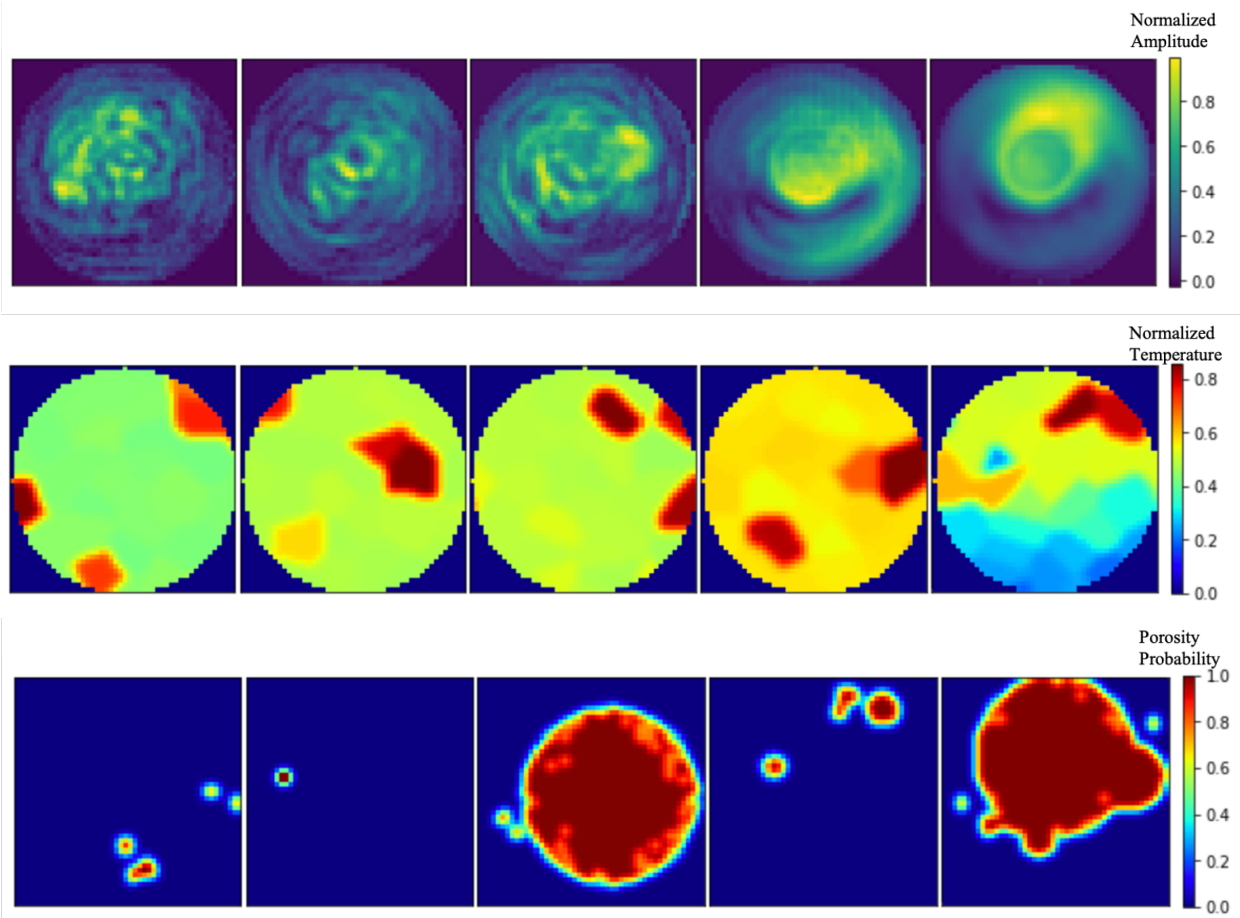


Figure 4.1: Sample images 7, 33, 40, 78, and 124 for U-Net sensor fusion. (Top) Ultrasonic layer images. (Middle) Infrared Layer images. (Bottom) XCT layer images.

4.2 Results of Sensor Fusion prediction of porosity for Ti-4AL-4V cylinders

After fabricating Ti-6AL-4V cylinders and collecting the pyrometer, ultrasonic, and XCT data, we seek to fit ultrasonic and pyrometer data to the sensor fusion model and predict the likelihood of porosity under the mapping of the model to XCT data. The mapping of ultrasonic and pyrometer data to XCT provides a baseline for the model to predict porosity in the subsequent layers of the Ti-4AL-4V cylinder fabrication layers. The models are trained with 100 epochs, a binary cross-entropy loss function, and an Adam optimizer. We perform leave one out fivefold cross-validation where 20 percent of the data is left out of training to test the model's performance. We train the model twice to verify results and then leave the next 20 percent of image data out of training to test performance. We repeat this process until all data has been part of the training and testing process.

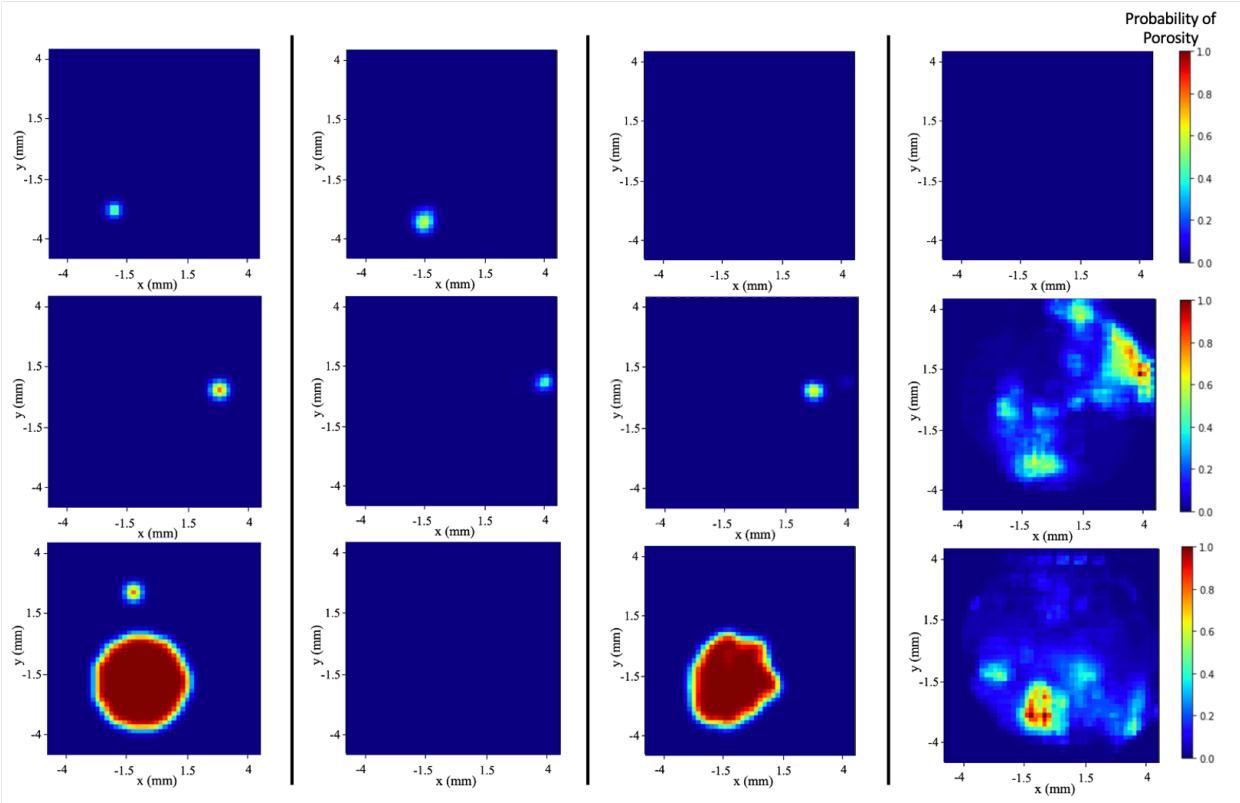
4.2.1 Traditional U-Net vs. Inception and LSTM modified U-Net Models

The quantitative results of combined traditional, inception, and LSTM U-Net models with four encoder-decoder blocks are shown in Table 4.2. Overall, the inception U-Net models exhibit a higher degree of overlap between actual and predicted pixels as seen in the bolded regions of the Q_I metric in Table 4.2. We do see less noise in the traditional U-Net model leading to less incorrect predictions measured by observing the Q_P metric and slightly more structural similarity by observing the Q_S metric. The combined inception model exhibits the highest Q_I, and the traditional model exhibits the highest Q_P value amongst all the models.

Table 4.2: Average metric results combined traditional, inception, and LSTM U-Net models

Model	Q_I	Q_S	Q_P	$Time(sec)$
Combined Traditional U-Net	0.501	0.99883	77.923	971
Combined Inception U-Net	**0.514**	0.98855	74.821	7976
Combined LSTM U-Net	0.482	0.98435	64.332	8612

Figure 4.2 shows the localization quality of the traditional, inception, and LSTM U-Net combined fusion model. Three XCT testing images are paired with the U-Net models to compare the quality of prediction. The XCT sample images can be seen in Figure 4.2a. It can be seen that the inception U-Net fusion model images in Figure 4.2c localize porosity significantly better than the single modality models. The traditional combined model does not perform as well when the pores are significantly larger than one another seen in the last row of Figure 4.2b. In addition, the traditional model achieves a higher Q_I than the LSTM model. We expected the LSTM model to achieve improved Q_I because we are incorporating successive layer data into the current layer as a 3D dimensional representation, but Q_P shows a high degree of noise. When training these models, the inception model and LSTM model training time are significantly higher. Therefore, it is beneficial to run the traditional model if a quick turnaround of results is necessary. Overall, the Q_I metric is the best indicator of the quality localization of an object in an image. The combined inception model performed the best compared to the combined traditional and LSTM U-Net models. The subsequent study will compare the combined inception U-Net to the infrared and ultrasonic single modality U-Net models.



[a] [b] [c] [d]

Figure 4.2: Samples 13, 18, and 28 image predictions of traditional, inception, and LSTM U-Net thermal and ultrasonic fusion models. (a) XCT. (b) Traditional U-Net fusion predictions. (c) Inception U-Net fusion predictions. (d) LSTM U-Net fusion predictions.

4.2.2 Combined U-Net fusion model vs. thermal and ultrasonic U-Net single modality models

We compare the results of the inception U-Net fusion model to inception U-Net thermal and ultrasonic single modality models to verify that sensor fusion produces better results. The qualitative results of training inception U-Net models are shown in Table 4.3. The inception U-Net fusion model out preformed the U-Net ultrasonic single modality model and U-Net thermal single

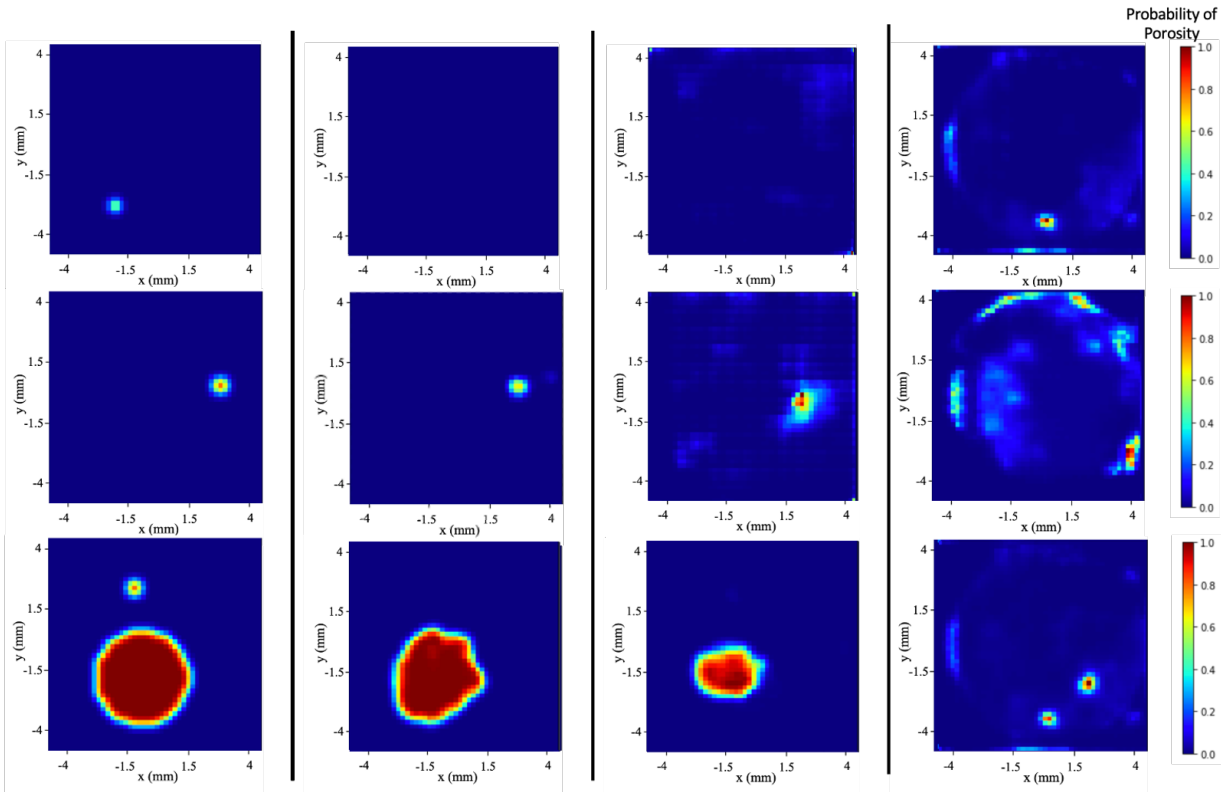
modality model by 0.028 percent and 0.029 percent Q_I, respectively. We would expect the U-Net ultrasonic single modality model to outperform the U-Net thermal single modality model because the ultrasonic data is similar to XCT, in which both sensors test for structural irregularities. The Q_P metric is showing a minor degree of noise in the U-Net fusion model than U-Net single modality models. The Q_S shows a sightly better structural similarity between XCT and U-Net fusion models than U-Net single modality models. Figure 4.3 shows sample image predictions of each model.

Table 4.3: Average metric results for combined, thermal single modality, ultrasonic single modality inception U-Net models

Model	Q_I	Q_S	Q_P	$Time(sec)$
Combined Inception U-Net	**0.514**	0.98855	80.821	7976
Ultrasonic Inception U-Net	0.486	0.98852	78.381	4713
Thermal Inception U-Net	0.485	0.98824	77.756	4318

The U-Net ultrasonic single modality model value over the thermal single modality model can be seen when comparing the prediction images in Figure 4.3c and Figure 4.3d. The ultrasonic single modality prediction shows significantly better localization of porosity than thermal single modality predictions. U-Net fusion model prediction shows better localization with less noise in the sample images. The run time of the U-Net fusion model is nearly double the U-Net single modality models. Fusion of ultrasonic and thermal images produces the best results. In the previous study,

we are tuning the number of processing blocks of the inception U-Net model to produce the best results.



[a] [b] [c] [d]

Figure 4.3: Samples 13, 18, and 28 image predictions of inception U-Net fusion model, inception U-Net thermal single modality model, and inception U-Net ultrasonic single modality model. (a) XCT. (b) Inception U-Net fusion model. (c) Inception U-Net ultrasonic single modality model. (d) Inception U-Net thermal single modality model.

4.2.3 Tuning the number of encoder-decoder processing blocks

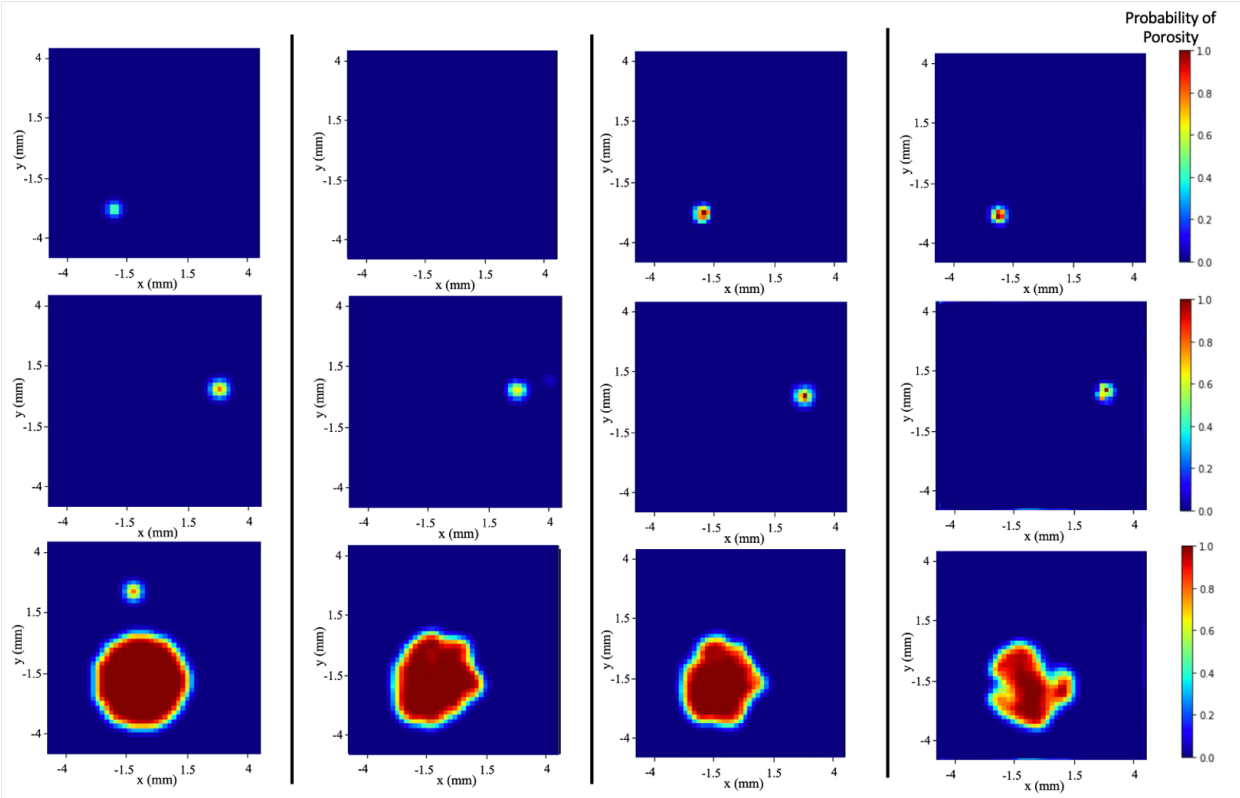
The performance of the fusion model might not be reaching expected accuracy because too small of dimension reduction of input might be resulting in a loss of too much information.

Therefore, we explore using fewer encoder-decoder block layers to tune the inception U-Net fusion model. The qualitative results of training inception U-Net models are shown in Table 4.4. The two encoder-decoder block inception U-Net fusion model produced the high percent Q_I while having a significantly lower run time than the three and four encoder-decoder block models. Additionally, the Q_S showed a higher average structural similarity and higher Q_P noise ratio between XCT and two encoder-decoder block models. Sample image predictions are compared in Figure 4.4.

Table 4.4: Average metric results for inception U-Net fusion model with four, three, and two encoder-decoder blocks

Number of encoder-decoder blocks	Q_I	Q_S	Q_P	$Time(sec)$
Four	0.514	0.98855	80.821	7976
Three	0.555	0.99502	76.332	6219
Two	**0.557**	0.99608	78.137	4445

Overall, the sample image predictions in Figure 4.4 showed similar results. The two and three encoder-decoder block models localize porosity when the pore is smaller in the first row of Figures 4.4c and 4.4d. All the models failed to recognize the second pore in the last image of Figure 4.4a. The two encoder-decoder block inception U-Net fusion model is the ideal number of blocks for improved performance and faster run time.



[a] [b] [c] [d]

Figure 4.4: Samples 13, 18, and 28 image predictions of inception U-Net fusion model with four, three, and two encoder-decoder blocks (a) XCT. (b) Four encoder-decoder blocks inception U-Net fusion model. (c) Three encoder-decoder blocks inception U-Net ultrasonic single modality model. (d) Two encoder-decoder blocks inception U-Net thermal single modality model.

CHAPTER V

CONCLUSION

Sensor fusion in additive manufacturing is a promising technique for improving manufactured components without compromising the structure. The use of various sensors in additive manufacturing can provide distinctly different features that can help illustrate the quality of the printed component. For example, thermal history data provides information about the printing process, while ultrasonic and XCT data provides information on structural integrity. The effective use of timely collected in-situ thermal data and ultrasonic data collected at significantly faster processing time as opposed to XCT can allow professionals to understand the complex relationship between processing parameters and the quality of the build. Taking advantage of the valuable information from each sensor can help manufacture parts with a higher degree of residual stress and print parts as anticipated the first time. Although, often, the sensor data analyzes different geometric locations and comes in different dimensions and formats. The sensor fusion methodology presented in this paper can bridge the gap between combining different sensors in LBAM by transforming data into the same field of view and autonomously extracting relevant features. It does so by efficiently utilizing K-means smoothing of multiple data points and a deep learning U-Net framework that can localize porosity from distinct features in ultrasonic and infrared layer images. To the best of

our knowledge, our framework is the first to bridge the gaps between combining in-situ and ex-situ sensors in LBAM.

Our fusion model method possesses multiple advantages for practitioners and researchers in LBAM.

1. The cost-efficient collection of data and computational run time of sensor fusion porosity detection models are significantly faster than XCT scans. Computationally fast models allow the practitioners and researchers to test the quality of newly fabricated parts quickly.
2. It leverages information from different streams and configures data to the same field of view in an effective manner, and extracts features without hand-crafted fusion specifications. This advantage is beneficial in extracting only relevant features that aim to achieve a specific outcome.
3. The model can effectively work for small datasets (observations > 150) with various parts created with different processing parameters. The addition of data from geometrically different parts created with different processing and structural makeup could take advantage of the low economic and computational cost.

Ultimately, our model can accelerate the use of various sensing techniques that were not clear how to combine in the past. The reduced cost and processing time compared to XCT can lead to the identification of abnormalities in fabricated parts and further advance the adoption of sensor fusion in LBAM and related industries. Additional sensors can be investigated that may yield benefits to further fabricated part quality. Furthermore, more significant complex parts can begin to be created when the quality of the printing process improves with the use of multi-sensor fusion models. Further sensor fusion research may find the correct combination of sensors to suit better the product being created.

Future work includes applying additional sensors to measure geometry and capture a different view of points of the manufactured component. The hope for this research is to determine abnormalities in the printing process at the time of the fault, remove unanticipated metal, and reprint

sections of parts to correctly manufacture parts the first time, saving money and time. However, sensor data collection can be expensive, difficult to obtain, or limited to a small number of machines creating the components. For example, fabrication of Ti-6AL-4V powder parts can cost 300-600 dollars per kilogram [37], can be limited on the number of sensors that can fit in the fabrication chamber, and take over 30 minutes to create parts about the size of a pencil. A consequence of this is that many of the parts created are small, only 1-2 sensors can be placed in the fabrication chamber, and much of the data collected is primarily focused on a few selected configurations. The addition of various parts would significantly improve prediction quality and understanding of the complex relationship between modalities in the sensor fusion model. In addition, future studies involving the (1) transfer learning of experience could be further investigated to bridge the gap in porosity predictions in other fabricated parts. (2) Data generative models could increase the number of training predictions to further use sensor fusion in LBAM. Ultimately, our methodology provides a sensor fusion framework for combining image data to predict porosity that takes steps toward improving LBAM part fabrication.

REFERENCES

[1] A. Arbelle and T. R. Raviv, "Microscopy Cell Segmentation via Convolutional LSTM Networks,", 2019.

[2] E. Balit and A. Ghadli, *GMFNet: Gated Multimodal Fusion Network for Visible-Thermal Semantic Segmentation*, March 2021.

[3] S. N. . U. J. Bian, L., *Laser-Based Additive Manufacturing of Metal Parts: Modeling, Optimization, and Control of Mechanical Properties*, 2017.

[4] D. Brunet, E. R. Vrscay, and Z. Wang, "On the Mathematical Properties of the Structural Similarity Index," *IEEE Transactions on Image Processing*, vol. 21, no. 4, 2012, pp. 1488–1499.

[5] S. Coeck, M. Bisht, J. Plas, and F. Verbist, "Prediction of lack of fusion porosity in selective laser melting based on melt pool monitoring data," *Additive Manufacturing*, vol. 25, 2019, pp. 347 – 356.

[6] J. Dolz, I. B. Ayed, and C. Desrosiers, "Dense Multi-path U-Net for Ischemic Stroke Lesion Segmentation in Multiple Image Modalities,", 2018.

[7] A. du Plessis, S. G. le Roux, and A. Guelpa, "Comparison of medical and industrial X-ray computed tomography for non-destructive testing," *Case Studies in Nondestructive Testing and Evaluation*, vol. 6, 2016, pp. 17–25.

[8] H. Gao, H. Yuan, Z. Wang, and S. Ji, "Pixel Transposed Convolutional Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 5, 2020, pp. 1218–1227.

[9] F. P. García Marquez and C. Q. Gómez Muñoz, "A New Approach for Fault Detection, Location and Diagnosis by Ultrasonic Testing," *Energies*, vol. 13, no. 5, Mar 2020, p. 1192.

[10] N. Ibtehaz and M. S. Rahman, "MultiResUNet : Rethinking the U-Net architecture for multimodal biomedical image segmentation," *Neural Networks*, vol. 121, 2020, pp. 74–87.

[11] J. Jiang and Y. Ma, "Path Planning Strategies to Optimize Accuracy, Quality, Build Time and Material Use in Additive Manufacturing: A Review," *Micromachines*, vol. 11, no. 7, 2020.

[12] J. Jiang, C. Yu, X. Xu, and Y. Ma, "Achieving better connections between deposited lines in additive manufacturing via machine learning," *Mathematical Biosciences and Engineering*, vol. 17, 2020, p. 3382.

[13] M. Khanzadeh, S. Chowdhury, M. Marufuzzaman, M. A. Tschopp, and L. Bian, "Porosity prediction: Supervised-learning of thermal history for direct laser deposition," *Journal of Manufacturing Systems*, vol. 47, 2018, pp. 69 – 82.

[14] Z. Li, Z. Zhang, J. Shi, and D. Wu, "Prediction of surface roughness in extrusion-based additive manufacturing with machine learning," *Robotics and Computer-Integrated Manufacturing*, vol. 57, 2019, pp. 488–495.

[15] R. Liu, A. Kumar, Z. Chen, A. Agrawal, V. Sundararaghavan, and A. Choudhary, "A predictive machine learning approach for microstructure optimization and materials design," *Scientific Reports*.

[16] R. Liu, S. Liu, and X. Zhang, "A Physics-Informed Machine Learning Model for Porosity Analysis in Laser Powder Bed Fusion Additive Manufacturing,", 2021.

[17] Y. Liu, X. Chen, Z. Wang, R. Ward, and X. Wang, "Deep learning for pixel-level image fusion: Recent advances and future prospects," *Information Fusion*, vol. 42, 07 2018, pp. 158–173.

[18] J. Long, E. Shelhamer, and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015.

[19] W. Ma, F. Cheng, Y. Xu, Q. Wen, and Y. Liu, "Probabilistic representation and inverse design of metamaterials based on a deep generative model with semi-supervised learning strategy,", 2019.

[20] R. Maev and F. Seviaryn, *Ultrasonic Imaging in Biomedical Applications*, 01 2018.

[21] 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," *The International Journal of Advanced Manufacturing Technology*, vol. 95, 03 2018.

[22] V. Marcantonio, D. Monarca, A. Colantoni, and M. Cecchini, "Ultrasonic waves for materials evaluation in fatigue, thermal and corrosion damage: A review," *Mechanical Systems and Signal Processing*, vol. 120, 2019, pp. 32 – 42.

[23] G. Marshall, S. Thompson, and N. Shamsaei, "Data indicating temperature response of Ti-6Al-4V thin-walled structure during its additive manufacture via Laser Engineered Net Shaping," *Data in Brief*, vol. 7, 03 2016.

[24] L. Meng, B. McWilliams, W. Jarosinski, H.-Y. Park, Y.-G. Jung, J. Lee, and J. Zhang, "Machine Learning in Additive Manufacturing: A Review," *JOM*, vol. 72, 04 2020.

[25] Nikon, "Series inspection CT: X-ray CT systems: FNikon Metrology,", Oct 2021.

[26] J. Ning, W. Wang, B. Zamorano, and S. Liang, "Analytical modeling of lack-of-fusion porosity in metal additive manufacturing," *Applied Physics A*, vol. 125, 11 2019.

[27] J. Oliveira, A. LaLonde, and J. Ma, "Processing parameters in laser powder bed fusion metal additive manufacturing," *Materials Design*, vol. 193, 2020, p. 108762.

[28] G. Pilania, C. Wang, X. Jiang, S. Rajasekaran, and R. Ramprasad, "Accelerating materials property predictions using machine learning," *Scientific reports*, vol. 3, September 2013, p. 2810.

[29] F. PirahanSiah, S. N. H. S. Abdullah, and S. Sahran, "Adaptive image segmentation based on peak signal-to-noise ratio for a license plate recognition system," *2010 International Conference on Computer Applications and Industrial Electronics*, 2010, pp. 468–472.

[30] M. Rahman and Y. Wang, "Optimizing Intersection-Over-Union in Deep Neural Networks for Image Segmentation," 12 2016, vol. 10072, pp. 234–244.

[31] P. K. Rao, J. Liu, D. Roberson, Z. Kong, and C. B. Williams, "Online Real-Time Quality Monitoring in Additive Manufacturing Processes Using Heterogeneous Sensors," *Journal of Manufacturing Science and Engineering-transactions of The Asme*, vol. 137, 2015, p. 061007.

[32] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *CoRR*, vol. abs/1505.04597, 2015.

[33] L. Scime and J. Beuth, "Using machine learning to identify in-situ melt pool signatures indicative of flaw formation in a laser powder bed fusion additive manufacturing process," *Additive Manufacturing*, vol. 25, 2019, pp. 151–165.

[34] S. C. Seong-Hyun Park and K.-Y. Jhang, "Porosity Evaluation of Additively Manufactured Components Using Deep Learning-based Ultrasonic Nondestructive Testing," *International Journal of Precision Engineering and Manufacturing-Green Technology*, February 2021.

[35] S. A. Shevchik, G. Masinelli, C. Kenel, C. Leinenbach, and K. Wasmer, "Deep Learning for In Situ and Real-Time Quality Monitoring in Additive Manufacturing Using Acoustic Emission," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 9, 2019, pp. 5194–5203.

[36] A. Shrestha and A. Mahmood, "Review of Deep Learning Algorithms and Architectures," *IEEE Access*, vol. 7, 2019, pp. 53040–53065.

[37] T. W. Simpson, "Why Does My 3D-Printed Part Cost So Much?,", May 2018.

[38] Stratonics, "Heat Flow Sensors: Additive Manufacturing: Sensors,", Apr 2016.

[39] M. Tang, P. C. Pistorius, and J. L. Beuth, "Prediction of lack-of-fusion porosity for powder bed fusion," *Additive Manufacturing*, vol. 14, 2017, pp. 39 – 48.

[40] Q. Tian, S. Guo, E. Melder, L. Bian, and W. Guo, "Deep Learning-Based Data Fusion Method for In Situ Porosity Detection in Laser-Based Additive Manufacturing," *Journal of Manufacturing Science and Engineering*, vol. 143, 10 2020, pp. 1–38.

[41] Z. Tian, L. Yu, and C. Leckey, "Rapid guided wave delamination detection and quantification in composites using global-local sensing:," *Smart Materials and Structures*, vol. 25, 08 2016, p. 085042.

[42] F. van Beers, A. Lindström, E. Okafor, and M. Wiering, "Deep Neural Networks with Intersection over Union Loss for Binary Image Segmentation," 02 2019.

[43] VolumeGraphics, "VGSTUDIO Max,", Oct 2021.

[44] A. Voulodimos, N. Doulamis, A. Doulamis, and E. Protopapadakis, "Deep Learning for Computer Vision: A Brief Review," *Computational Intelligence and Neuroscience*, vol. 2018, 02 2018, pp. 1–13.

[45] C. Wang, X. Tan, S. Tor, and C. Lim, "Machine learning in additive manufacturing: State-of-the-art and perspectives," *Additive Manufacturing*, vol. 36, 2020, p. 101538.

[46] L. Wu, L. Liu, Y. Wang, Z. Zhai, H. Zhuang, D. Krishnaraju, Q. Wang, and H. Jiang, "A machine learning-based method to design modular metamaterials," *Extreme Mechanics Letters*, vol. 36, 2020, p. 100657.

[47] T. Zhou, S. Ruan, and S. Canu, "A review: Deep learning for medical image segmentation using multi-modality fusion,", 04 2020.

[48] Y. Zhu, Z. Wu, W. D. Hartley, J. M. Sietins, C. B. Williams, and H. Z. Yu, "Unraveling pore evolution in post-processing of binder jetting materials: X-ray computed tomography, computer vision, and machine learning," *Additive Manufacturing*, vol. 34, 2020, p. 101183.