

An Image Segmentation Framework for In-Situ Monitoring in Laser Powder Bed Fusion Additive Manufacturing

Jason Xie* Tianyu Jiang* Xu Chen*

** Department of Mechanical Engineering, University of Washington,
Seattle, WA 98105 USA
(e-mail: jasonx1e@uw.edu, tjiang19@uw.edu, chx@uw.edu)*

Abstract: Laser powder bed fusion (LPBF) is an additive manufacturing technique that offers a number of competitive advantages over conventional manufacturing methods; however, challenges in assuring final part quality hinder its broader industrial adoption. As a step towards closing the loop in quality control for LPBF, this paper presents an image segmentation framework for extracting process characteristics from in-situ image data collected by a coaxial imaging system in the visible spectrum. Since manual data annotation can be time-consuming and error prone, we present an image processing pipeline for generating precise segmentation labels semi-automatically. The resulting dataset is used to train a machine learning model for segmentation of processed zone boundaries for in-situ monitoring. Experimental results validate the effectiveness of our approach for process monitoring in LPBF across a variety of challenging illumination and process conditions.

Keywords: additive manufacturing, laser powder bed fusion, in situ monitoring, image processing, image segmentation, machine learning, deep learning

1. INTRODUCTION

Laser powder bed fusion (LPBF) is a popular class of additive manufacturing (3D printing) processes whereby three-dimensional objects are constructed from digital specifications by sequentially depositing and sintering / melting thin layers of material powders via one or more laser sources. Common LPBF processes include selective laser sintering (SLS), direct metal laser sintering, and selective laser melting. LPBF offers a number of competitive advantages over conventional manufacturing methods: it supports a wide range of materials often used in engineering applications (e.g. polymers, metals, ceramics, composites) and allows for the direct fabrication of unsupported complex geometries (e.g. overhangs) while minimizing material waste and production costs [Gibson et al. (2015)].

Although these attributes make LPBF well-suited for a wide range of emerging commercial applications, issues with quality assurance, part reliability, and process reproducibility limit broader adoption of this technology, particularly in safety-critical applications such as those in the automotive [Clarke (2017)], aerospace [Bullis (2013)], and medical devices [Grunewald (2016)] industries. One reason for this challenge is due to the complex dynamics between a large number of highly correlated process parameters, many of which cannot be directly measured [Spears and Gold (2016)]. In order to realize effective control strategies for mitigating reliability issues during the manufacturing process, it is crucial i) to identify process signatures that

correlate with the quality of the final build and ii) to develop of in-situ monitoring methods that can generate online feedback signals from these process signatures.

Since melt pool and heat-affected zone (HAZ) signatures provide critical information about the uniformity of part quality and the presence of defects [Grasso and Colosimo (2017)], various sensing strategies have been proposed for sensing and monitoring melt pool signatures, such as optical imaging, thermography, spectrometry, optical coherence tomography, and acoustic techniques [Grasso et al. (2021)]. Visible light cameras are of particular interest because they can provide detailed spatial information about the LPBF process complementary to sensors operating in other ranges of the electromagnetic spectrum (e.g. infrared cameras or pyrometers) at a high data bandwidth and comparatively low cost [Jiang (2021)].

One promising means for extracting process signatures from camera data is through image segmentation. Image segmentation is a well studied problem in the image processing and computer vision communities with a vast body of academic literature [Szeliski (2022)] which seeks to partition an image into distinct regions of interest. The resulting segmentation can then be used to directly compute relevant metrics describing process characteristics - e.g. melt pool size and geometry - in a straightforward manner [Jiang et al. (2021)]. Prior work on LPBF monitoring has demonstrated that statistical pattern recognition can be applied to image segmentations for extracting process characteristics and detecting anomalies [Scime et al. (2018), Grasso et al. (2016)]. In these works, image segmentations were obtained via global intensity thresh-

* This work was partially supported through National Science Foundation Grant CMMI-1953155 and UW CoMotion Innovation Gap Fund S-2021-8088-Chen.

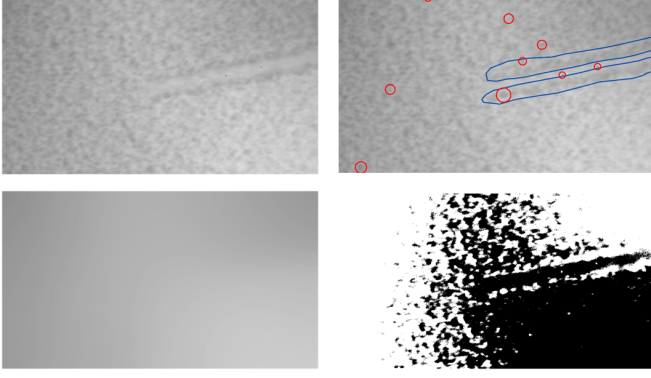


Fig. 1. (a) Sample image captured during the build process by the coaxial imaging system. (b) Optical contaminants are circled in red, and a coarse annotation of the laser-processed zones is contained within the blue polygon. (c) Non-uniform illumination of the powder bed, which challenges (d) global thresholding algorithms such as Otsu's method [Otsu (1979)].

olding. However, because digital images are sensitive to illumination conditions, contaminated optics, and poor contrast between fused and raw materials, segmentation via global intensity thresholding alone struggles to generalize across more challenging experimental variations, particularly for the case of polymer laser sintering. Moreover, there is a dearth of labeled datasets for the problem of LPBF process monitoring, and manual annotation of LPBF data can be difficult due to subtle and imprecisely defined boundaries surrounding the laser-processed zones. Fig. 1 provides a representative example of an image captured during polymer SLS and illustrates some of these difficulties.

In this paper, we present an image segmentation framework for in-situ monitoring of laser-processed zones during LPBF from a coaxial visible light video stream. To overcome the absence of publicly available labeled LPBF datasets, we develop an image processing pipeline for the semi-automatic generation of processed zone boundary segmentation labels from videos collected from an in-house LPBF testbed platform (Fig. 2) introduced in [Jiang et al. (2021)]. The data annotation procedure only makes simple assumptions about the LPBF build process and image parameters that can be easily obtained upon inspection. These labels are then used to train a deep learning model for the task of processed zone boundary segmentation that generalizes across a variety of experimental parameters. We apply the proposed method to the task of image segmentation of coaxial video streams collected during polymer LPBF. Our results validate that the proposed framework can generate reliable segmentations from images captured across different experimental instances containing a variety of scan track geometries. The learned segmentation model also provides a promising means of correlating in-process observations with final part quality and serves as a step towards enabling online feedback control of the LPBF process.

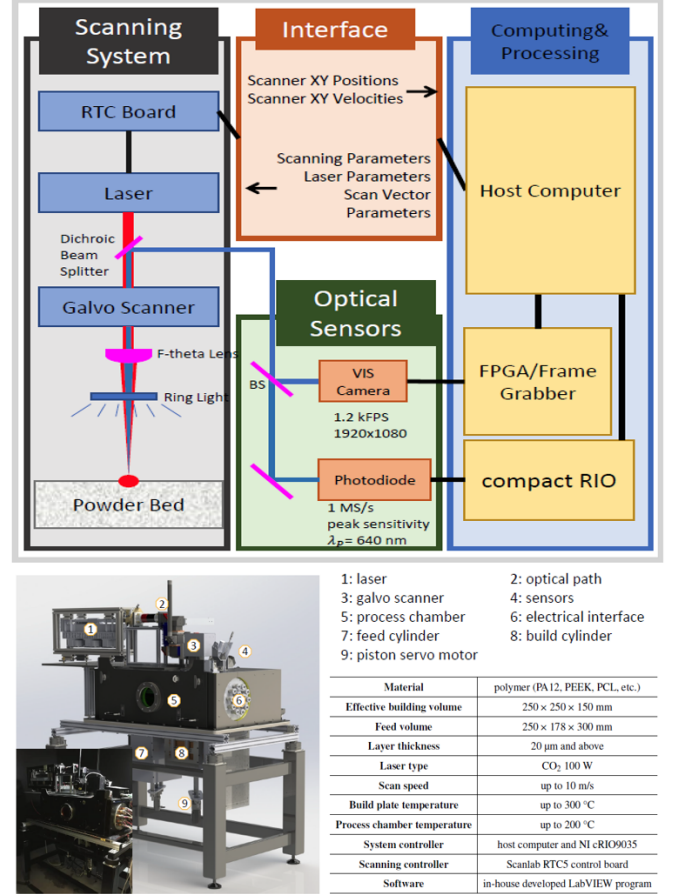


Fig. 2. (a) Block diagram of the coaxial monitoring and control system. (b) Overview of the experimental testbed. (c) General machine technical specifications.

2. PROBLEM FORMULATION

2.1 Image Segmentation for Processed Zone Identification

The problem of image segmentation can be formulated as a binary classification problem where each pixel in the image is assigned a label from a predetermined, often mutually exclusive set of categories. Formally, given an image I_t captured at timestep t with dimensions $H \times W$, the objective is to assign one of K classes c_i drawn from $C = \{c_i\}_{i=0}^{K-1}$ to each pixel $I_t(m, n)$ where $m \in \{0, \dots, H\}$, $n \in \{0, \dots, W\}$ such that pixels belonging to the same category should receive the same class label. Since each pixel in the original input image $I_t(m, n)$ is assigned a label, we can construct an image segmentation mask S with dimensions $H \times W$ where each $S(m, n)$ takes on a value from C corresponding to the label assigned to it. Interested readers are invited to consult [Szeliski (2022), Minaee et al. (2021)] for more details on general-purpose image segmentation algorithms.

Since the framework presented in this paper primarily focuses on processed zone segmentation, in our problem setup the $K = 2$ classes correspond to whether a pixel should belong to the processed zone boundary (or not). Although the image segmentation problem formulation is sufficiently general to accommodate more specific labels relevant to process monitoring (e.g. melting pool, fumes,

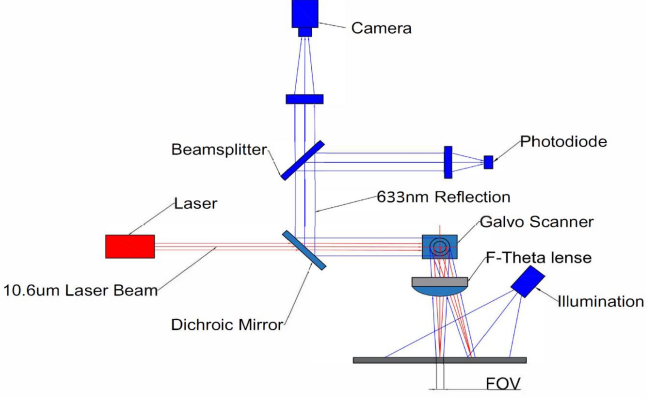


Fig. 3. Illustration of the optical path of the coaxial imaging system.

spatters, defects), we focus primarily on segmentation of laser-processed zones in this work.

2.2 Experimental Setup

We collect video data for our experiments on an in-house developed LPBF testbed. The experimental platform uses a 100-Watt CO₂ laser as the power source and a SCAN-LAB intelliSCAN14 galvanometer scanner with a $f = 400$ mm F-theta lens as the beam positioning system. A ZnSe dichroic beam splitter at an incidence angle of 45° degrees with 98.5% reflectance at 633 nm propagates approximately 70% of visible light radiation from the powder bed to a Cyclone 2 Machine Vision camera in a coaxial configuration, which allows for direct top-down observation of the laser-material interaction dynamics on the powder bed surface (Fig. 3). A Coaxlink Quad CXP-12 FPGA (Field-Programmable Gate Array)-based frame grabber triggers the monitoring camera, providing up to 5,000 MB/s data bandwidth and allowing acquisition of full resolution HD images ($1080 \text{ pixels} \times 1920 \text{ pixels}^1$) at 2158 fps. Image and motion data is streamed to and processed by a host computer that controls the process parameters through a RTC (real-time control) board. We refer interested readers to [Jiang et al. (2021)] for more details regarding the testbed’s mechatronic design.

3. METHODOLOGY

Deep learning has proven successful on a number of problems in image processing and computer vision [Krizhevsky et al. (2012), LeCun et al. (2015)], owing to their capacity to automatically learn strong nonlinear feature representations directly from high-dimensional images. In particular, many deep learning models have been developed for the task of image segmentation across a variety of problem domains [Minaee et al. (2021)]. However, these successes often rely on the availability of large, diverse, labeled datasets lacking in the LPBF problem domain.

One common approach to labeling data for image segmentation is to draw a polygon that bounds a particular region of interest (ROI) within an image and to assign

¹ Following the $H \times W$ convention of expressing image dimensions adopted in the rest of this paper

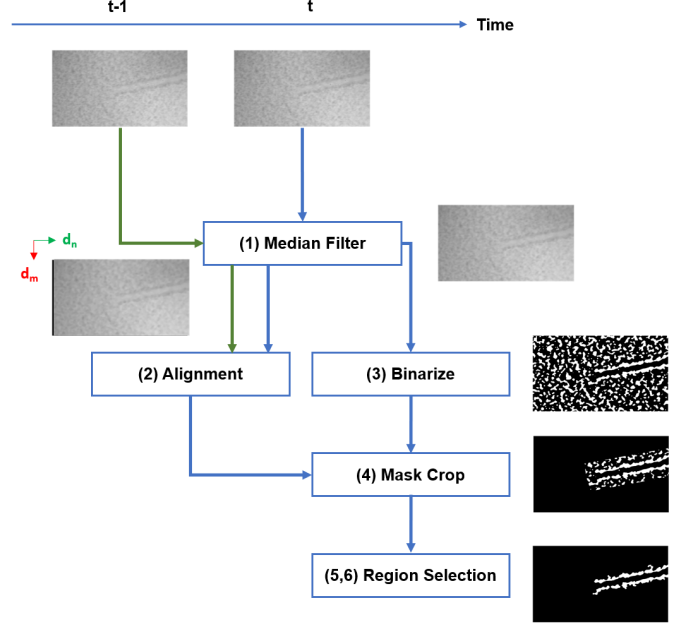


Fig. 4. Overview of data annotation pipeline.

a label to all pixels within the closure of the polygon [Wada (2016)]. This data annotation strategy can potentially be prohibitively time-consuming, as a large amount of data is collected during the LPBF process due to the necessity of a high-speed camera for capturing the rapidly changing laser-material interactions. Additionally, because processed zone boundaries are often smooth and ambiguous, the use of coarse polygons during the data annotation process may lead to imprecise labels which can negatively affect the performance of the trained machine learning model. Instead, we propose an image processing pipeline that exploits geometric regularities in sequentially collected images to generate segmentation labels with minimal manual intervention.

3.1 Acquiring Segmentation Labels via Image Processing

Our proposed data annotation strategy assumes that images collected when the laser follows a linear scan track and that two tracks are not being fused within the image. We also assume *a priori* approximations of the melt pool radius in pixels r_{melt} and the location of the melt pool center $\mu = (m_c, n_c)$ in the image plane, which remains largely unchanged throughout the LPBF build process due to the coaxial optical design of the experimental testbed. We note that these variables can be easily estimated via visual inspection; thus, the proposed annotation strategy does not require access to exact system parameters. Given a stream of sequentially captured images taken during the LPBF process, the data annotation procedure (Fig. 4) is described as follows:

- (1) Apply a median filter to the current image I_t to remove optical contamination artifacts, characterized by small dark particles. The filter kernel size is chosen to be the smallest odd integer greater than the pixel diameter of the largest particle.
- (2) Estimate the 2D translation vector of the scan $d_{scan} = (d_m, d_n)$ in the image plane between the current image frame I_t and the previous frame I_{t-1}

via image alignment. Since most of the powder bed is unchanged between consecutive frames, the pixels in $I_{t-1}(m, n)$ that correspond to the unchanged powder bed surface in I_t will be displaced in the opposite direction of the scan translation vector:

$$I_{t-1}(m, n) = I_t(m + d_m, n + d_n) \quad (1)$$

This displacement vector also allows us to iteratively accumulate the length of the laser-processed zone at time t :

$$l_t := l_{t-1} + \sqrt{d_m^2 + d_n^2} \quad (2)$$

In our experiments, we compute this displacement vector by optimizing with respect to the enhanced correlation coefficient (ECC) criterion [Evangelidis and Psarakis (2008)].

- (3) Binarize the median-filtered image such that the contour of the laser-processed zone is preserved. To account for non-uniformity in powder bed illumination, we apply adaptive Gaussian thresholding to generate a binary image that distinguishes pixels between the processed zone boundary and the unfused material based on their intensity relative to nearby pixels. Morphological dilations are used to close any gaps within the binarized regions. Although we use adaptive thresholding due to its simplicity and fast computation speed, we note that alternative methods (e.g. [Jiang et al. (2021), Chan and Vese (1999)]) can be substituted in this step without requiring adjustments to the rest of the pipeline.
- (4) Based on geometry and projection of the laser beam from the rotating galvo scanner to the powder bed surface, isolate the scan track by masking out regions greater than r_{melt} pixels from the laser scan track. We compute a rectangular bounding box B centered at the laser-processed zone with length, width, orientation, and center coordinate and $[B_l, B_w, B_\theta, B_c]$ that extends from the center of the melt pool across the laser-processed zone l_t pixels in the direction of the scan vector with fixed width r_{melt} :

$$\begin{aligned} B_\theta &= \theta_{scan} = \arctan 2(d_m, d_n) \\ B_l &= r_{melt} + l_t \\ B_w &= 2r_{melt} \\ p_{forward} &= \mu + r_{melt} \begin{bmatrix} \cos(B_\theta) \\ \sin(B_\theta) \end{bmatrix} \\ p_{edge,1} &= \mu + r_{melt} \begin{bmatrix} \cos(B_\theta + \pi/2) \\ \sin(B_\theta + \pi/2) \end{bmatrix} \\ p_{edge,2} &= \mu + r_{melt} \begin{bmatrix} \cos(B_\theta - \pi/2) \\ \sin(B_\theta - \pi/2) \end{bmatrix} \\ B_c &= p_{forward} - \mu + \frac{p_{edge,1} + p_{edge,2}}{2} \\ &\quad + \frac{B_l}{2} \begin{bmatrix} \cos(B_\theta + \pi) \\ \sin(B_\theta + \pi) \end{bmatrix} \end{aligned} \quad (3)$$

The bounding box is used to generate a boolean image mask B_{mask} where pixels corresponding to the closure of B are assigned a value of 1 and all other pixels are assigned a value of 0. The mask is used to crop the binary image obtained from step (3) to only include the laser-processed zones.

- (5) Distinguish unique connected regions within the cropped binary image via connected components labeling, excluding the background (which can be iden-

tified as the largest connected component in the image).

- (6) Uniformly sample L line segments of length B_l across the width of B and compute the percentage of pixels across these line segments that intersect with a non-background region. The endpoints $s_{i,1}, s_{i,2}$ corresponding to line segment $s_i = \{(1-t)s_{i,1} + ts_{i,2} \mid t \in (0, 1)\}$ can be directly computed from the parameters of B as follows:

$$\begin{aligned} s_{0,c} &= B_c + \frac{B_w}{2} \begin{bmatrix} \cos(B_\theta + \pi/2) \\ \sin(B_\theta + \pi/2) \end{bmatrix} \\ s_{L,c} &= B_c + \frac{B_w}{2} \begin{bmatrix} \cos(B_\theta - \pi/2) \\ \sin(B_\theta - \pi/2) \end{bmatrix} \\ s_{i,c} &= s_{0,c} + \frac{i}{L} s_{L,c} \\ s_{i,1} &= s_{i,c} + \frac{B_l}{2} \begin{bmatrix} \cos(B_\theta) \\ \sin(B_\theta) \end{bmatrix} \\ s_{i,2} &= s_{i,c} + \frac{B_l}{2} \begin{bmatrix} \cos(B_\theta + \pi) \\ \sin(B_\theta + \pi) \end{bmatrix} \end{aligned} \quad (4)$$

With the endpoint values of each line segment $s_{i,1}, s_{i,2}$, we can index the pixels across the line segment in the image and compute the percentage that overlap with the regions identified in (5). Since regions corresponding to the processed zone edges have greater color continuity than that of the raw powder, we keep only the regions that overlap with the two unique line segments that intersect with the least amount of background pixels.

- (7) Return the final binary image, where the filtered regions from (6) comprise a training label for processed zone segmentation of input image I_t .

Floating-point numbers are used in all computations above and quantized to the nearest integer when a pixel location is referenced (bounded by image dimensions $H \times W$). With minimal tuning of the rectangular mask size and image binarization parameters, the proposed method is capable of generating precise labels for linear scan track observations when there is only one melted track, when the melt pool is clearly visible in the observed image, and when the scan track is mostly unbroken.

3.2 Deep Learning for Processed Zone Segmentation

To generalize beyond the assumptions described in the aforementioned data labeling procedure, we train a deep learning model for the task of processed zone boundary segmentation.

In recent years, convolutional neural networks (CNNs) [Krizhevsky et al. (2012)] have been successfully applied to a wide variety of image processing and analysis tasks, ranging from recognition of a diverse category of objects to classification of defects in manufacturing. In contrast to classical approaches in image processing and computer vision which often rely on human-designed models or features characterizing salient aspects of the task at hand, CNNs can automatically learn visual feature representations for a given task from image data. Since these features are learned at the pixel level, CNNs are not subject to the limitations of the image processing pipeline described in the prior section.

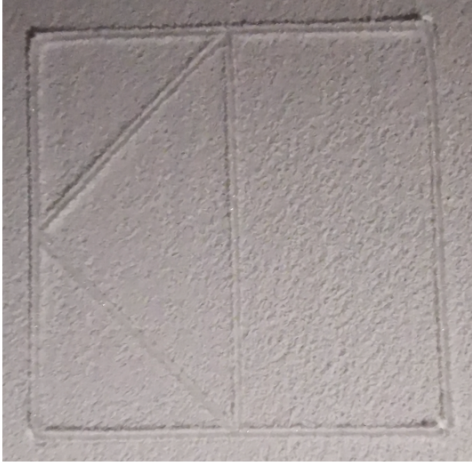
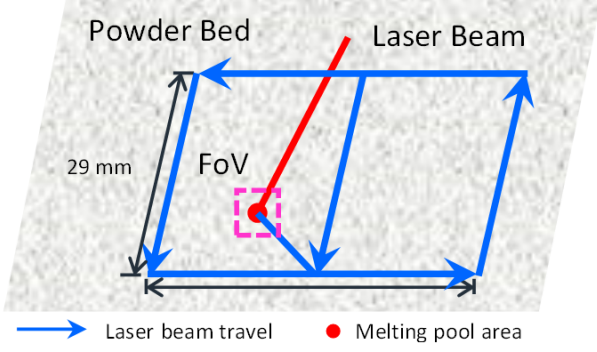


Fig. 5. Polymer SLS experiment setup. (top) Schematic diagram of the fusion geometry (bottom) Solidified parts from case study

A common network architecture used in pixel-wise semantic segmentation is the fully convolutional network (FCN) [Long et al. (2015)]. FCNs spatially downsample the high dimensional input image with a CNN and up-sample the image to generate a dense output of the same spatial resolution as the input, which is decoded via the softmax operator $\sigma(z)_i = e^{z_i} / \sum_{j=0}^{K-1} e^{z_j}$ to a probability distribution over K possible classes for each pixel in the input image, thus producing an output of $H \times W \times K$ dimensions with values between (0, 1). During training, supervision is provided at the pixel level for each (image, segmentation label) pair. At inference, each pixel in the final segmentation $S(x, y)$ is typically assigned the label that corresponds to the class with the highest probability.

We adopt DeepLabv3, which has achieved state-of-the-art results on several semantic segmentation benchmarks [Chen et al. (2017)] as our deep learning model. To resolve the issue of spatial resolution loss from downsampling convolutional layers, DeepLabv3 modifies the basic FCN architecture by instead using successive upsampled convolutional filters ("atrous convolutions"). The upsampled filters are generated by zero padding between filter taps at multiple image scales.

Leveraging the successes of transfer learning [Minaee et al. (2021)], we fine-tune a DeepLabv3 model with a ResNet-50 [He et al. (2015)] feature extraction backbone pre-trained on ImageNet [Deng et al. (2009)]. Since the feature

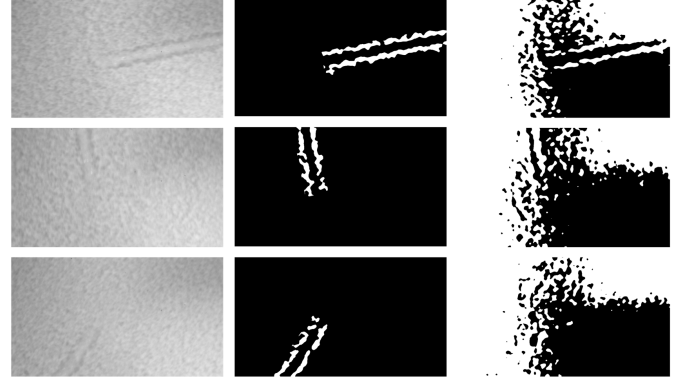


Fig. 6. Results from running the data annotation pipeline. (left) Original image captured by imaging system (middle) Output of the proposed image processed pipeline (right) Naive thresholding using [Otsu (1979)]

extractor is trained on RGB images and expects a 3-channel input, we repeat the 1-channel grayscale 3 times across the channel dimension before feeding the input into the network. Input images are normalized with respect to the pre-computed mean and variance of the ImageNet dataset and downsampled to 224 pixels $\times$ 398 pixels prior to being passed into the network. The network's output layers are replaced for the task of binary classification, and the model is trained using binary cross entropy loss via stochastic gradient descent with momentum and weight decay. We train on a dataset of 3000 images using a learning rate of 0.02 for 15 epochs with a batch size of 16. For comparison, we also test using a MobileNet v3-based feature extractor [Howard et al. (2019)], a more lightweight feature extractor optimized for performance on mobile CPUs, and with applying histogram equalization to the raw image to normalize the illumination profile of the inputs to the network.

LPBF monitoring data is highly correlated due to the temporal relationship between successive observations. To improve robustness to different process conditions, we apply standard data augmentation techniques to all training labels. Left-right flipping, vertical flipping, and random rotations in the range $[-60, 60]$ degrees are applied at regular intervals to increase geometric diversity of the training set.

4. CASE STUDY AND EXPERIMENTAL RESULTS

We evaluate our proposed framework on coaxial video streams collected by processing a single-layer geometry consisting of a square ($29\text{mm} \times 29\text{mm}$) and a triangle inscribed within the square using white Polyamide 12 (PA12 / Nylon 12) powder material (Fig. 5). To investigate the melt pool behavior during the fusion of an unsupported overhang, the fusion is performed on a thick powder bed such that the melt pool never reaches the underlying solid substrate. Before printing, the powder bed is preheated to 150°C . The camera is set to capture video at 50 fps to allow for adequate exposure, and the laser scanning speed is fixed at 6mm/s with a duty cycle of 0.8%.

Fig. 6 shows the qualitative results of generating segmentation labels via our proposed image processing pipeline.

White regions in the binary image correspond to the boundary of the laser-processed zones, while black regions correspond to the background (i.e. everything else). Upon visual inspection, it is clear that the contour of the laser-processed zone is ambiguous and poorly defined, and attempting to manually annotate the data using a polygonal selection tool would be time-consuming and likely result in labels that do not have precise boundaries. In contrast, our proposed method is sensitive to local geometric variations in the laser-processed zone and can reliably generate annotations for a full linear scan track using a single set of parameters. As a baseline, we also include results from applying simple global thresholding.

Fig. 7 shows the trained deep network’s predicted segmentation masks, and Table 1 reports the mean pixel accuracy (mPA) and mean intersection-over-union (mIOU) of the segmentation model on a test dataset of 500 images, with the outputs of the image processing pipeline described in Section 3.1 treated as ground truth. In our experiments, the learned image segmentation model significantly outperforms Otsu’s method, a global thresholding strategy adopted in prior work on extracting process signatures from LBPF image data. Furthermore, the learned model accurately segments a number of images that our proposed data labeling procedure cannot directly handle - e.g. the images with multiple laser-processed tracks from previous material fusion, images in which the laser-processed zone consists of a small proportion of the image, and images in which the laser changes scan direction. These results demonstrate that the network is capable of learning an adaptable internal feature representation for image segmentation that generalizes beyond the linear geometries provided in the training dataset and can thus be used to extract process metrics from unlabeled data with minimal retraining.

Table 1 also reports the mean runtime needed to process a single image. Our framework, implemented in Python / PyTorch / OpenCV / scikit-image, runs at >12 Hz on a modern desktop computer (Intel® Core™ i5-7600K, NVIDIA GTX1080), and is thus suitable for adoption in an online control scheme.

Table 1. Summary of experimental results

Method	mPA	mIOU	Speed (ms)
Otsu’s	76.0%	29.7%	0.3
DeepLab v3 (MobileNetv3)	79.7%	66.1%	37.7
DeepLab v3 (ResNet-50)	94.6%	79.4%	82.2
+CLAHE	94.6%	77.6%	82.2

5. CONCLUSIONS

We present a framework for the segmentation of laser-processed zones from in-situ coaxial image data captured during the LPBF process. We describe in detail our image processing pipeline for acquiring ground truth labels with minimal assumptions about experimental parameters, which can then be used to train a more general deep learning model for image segmentation. Our proposed framework is able to reliably segment the boundaries of laser-processed zones across a variety of conditions not present in the training data. Future work will focus on correlating in-situ process metrics extracted from the segmentation masks to final build quality and optimizing the

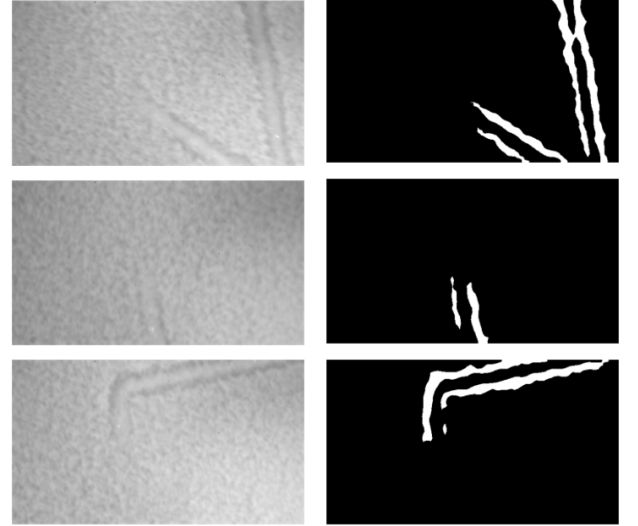


Fig. 7. Results of applying the learned segmentation model to a representative sample of challenging cases to directly annotate. (a) Multiple fused tracks (b) Small melt pool (c) Change in laser scan direction

proposed network architecture for real-time inference in feedback control.

REFERENCES

- Bullis, K. (2013). A more efficient jet engine is made from lighter parts, some 3-d printed. *MIT Technology Review*.
- Chan, T. and Vese, L. (1999). An active contour model without edges. *Scale-Space Theories in Computer Vision*, 141–151.
- Chen, L., Papandreou, G., Schroff, F., and Adam, H. (2017). Rethinking atrous convolution for semantic image segmentation. *CoRR*, abs/1706.05587.
- Clarke, C. (2017). University of nottingham’s flac project makes fuel efficient 3d printed car components. *3D Printing Industry*.
- Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., and Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. 248–255.
- Evangelidis, G.D. and Psarakis, E.Z. (2008). Parametric image alignment using enhanced correlation coefficient maximization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(10), 1858–1865.
- Gibson, I., Rosen, D.W., and Stucker, B. (2015). *Additive Manufacturing Technologies*. Springer, 2nd edition.
- Grasso, M., Remani, A., Dickens, A., Colosimo, B.M., and Leach, R.K. (2021). In-situ measurement and monitoring methods for metal powder bed fusion: an updated review. *Measurement Science and Technology*, 32(11), 112001.
- Grasso, M. and Colosimo, B.M. (2017). Process defects and in situ/monitoring methods in metal powder bed fusion: a review. *Measurement Science and Technology*, 28(4), 044005.
- Grasso, M., Laguzza, V., Semeraro, Q., and Colosimo, B.M. (2016). In-Process Monitoring of Selective Laser Melting: Spatial Detection of Defects Via Image Data Analysis. *Journal of Manufacturing Science and Engineering*, 139(5). 051001.

- Grunewald, S.J. (2016). Stryker’s spine division to debut 3d printed tritanium posterior lumbar cage spinal implant. *3DPrint.com*.
- He, K., Zhang, X., Ren, S., and Sun, J. (2015). Deep residual learning for image recognition. *arXiv preprint arXiv:1512.03385*.
- Howard, A., Sandler, M., Chu, G., Chen, L.C., Chen, B., Tan, M., Wang, W., Zhu, Y., Pang, R., Vasudevan, V., Le, Q.V., and Adam, H. (2019). Searching for mobilenetv3. *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*.
- Jiang, T. (2021). Mechatronic design and feedback control for quality-assured laser powder bed fusion additive manufacturing. PhD Thesis, University of Washington, Seattle, WA, USA.
- Jiang, T., Leng, M., and Chen, X. (2021). Control-oriented mechatronic design and data analytics for quality-assured laser powder bed fusion additive manufacturing. *2021 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*, 1319–1324.
- Krizhevsky, A., Sutskever, I., and Hinton, G.E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 25, 1097–1105.
- LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature*, 521, 436–444.
- Long, J., Shelhamer, E., and Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 3431–3440.
- Minaee, S., Boykov, Y.Y., Porikli, F., Plaza, A.J., Kehtarnavaz, N., and Terzopoulos, D. (2021). Image segmentation using deep learning: A survey. *IEEE transactions on pattern analysis and machine intelligence*.
- Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1), 62–66.
- Scime, L., Fisher, B., and Beuth, J. (2018). Using coordinate transforms to improve the utility of a fixed field of view high speed camera for additive manufacturing applications. *Manufacturing Letters*, 15, 104–106. Industry 4.0 and Smart Manufacturing.
- Spears, T.G. and Gold, S.A. (2016). In-process sensing in selective laser melting (slm) additive manufacturing. *Integrating Materials and Manufacturing Innovation*, 5, 16–40.
- Szeliski, R. (2022). *Computer vision: algorithms and applications*. Springer, Cham, 2nd edition.
- Wada, K. (2016). Labelme: Image polygonal annotation with python.