Additive Manufacturing 37 (2021) 101696

Contents lists available at ScienceDirect

Additive Manufacturing

journal homepage: www.elsevier.com/locate/addma

Research Paper

Precise localization and semantic segmentation detection of printing conditions in fused filament fabrication technologies using machine learning

Zeqing Jin, Zhizhou Zhang, Joshua Ott, Grace X. Gu *

Department of Mechanical Engineering, University of California, Berkeley, CA 94720-1740, USA

ARTICLE INFO

Keywords:
Additive manufacturing
Machine learning
Fused filament fabrication
Localization detection
Semantic segmentation

ABSTRACT

Although a vast array of anomaly detection methods has been developed in fused filament fabrication, a widely-applied additive manufacturing technology, acquiring in-situ detailed spatial information of the defects within the detection field remains a significant challenge in actual processing. In this paper, machine learning algorithms are proposed to realize precise localization and semantic segmentation detection of the in-plane printing conditions including over-extrusion and under-extrusion in both local and global frameworks. Results visualization and evaluation methods are conducted to demonstrate the high performance of the models. Our results show that detection latency is also improved by successfully recognizing the transitions between print quality conditions within a single raster. This advanced detection system is able to provide comprehensive defect information for real-time assessment and has great potential for further automated control as well as correction of additive manufacturing systems.

1. Introduction

The face of manufacturing has transformed over the last decades due to the advent of advanced computing algorithms [1–3]. Specifically, computing algorithms have shaped the field of additive manufacturing (AM) which has offered many advantages such as rapid prototyping and complex design possibilities. In recent studies, the capability of AM has been further extended to create multi-material, multi-scale, and multi-functional products by combining various AM methods [4–7]. Additionally, much research has been focused on using computer algorithms to solve various problems during the manufacturing process such as porosity and residual stress in laser powder bed fusion (LPBF), fracture of components in stereolithography (SL), and warpage in fused filament fabrication (FFF) [8–12]. From these studies, it has shown that incorporating computer algorithms is a promising method to overcome challenges in current additive manufacturing technologies [13].

As one of the most widely adopted additive manufacturing methods, fused filament fabrication (FFF) is an extrusion-based printing process using polymer filament to build prints layer by layer. Considerable studies have been conducted to optimize the manufacturing process and improve printing quality through many different aspects and approaches. Some of the various approaches include analyzing how printing parameters (layer thickness, road width, and speed of deposition) affect the quality of prototypes [14–16], real-time and in-situ monitoring of the manufacturing process [17,18], and assessment of the mechanical properties using non-destructive methods [19,20]. These studies focus on a macroscopic perspective of the problems that may occur in the FFF process such as catastrophic failure of the prints including shifting and breaking of parts or inaccurate dimensions after the printing process. However, most macro issues are often caused by the accumulation of local anomalies such as in-plane defects (under or over-extrusion in partial regions) or inter-plane delamination and warping that usually occurs at the corner of a print [21,22]. These small defects could lead to significant imperfections and create barriers for full industrial adoption of AM [23,24]. This paper focuses on discovering local anomalies in-situ to avoid the occurrence of significant imperfections including undesired functional performance and dimensional inaccuracy. Hence, the capability of detecting regional defects along with their spatial and temporal information is critical for industrial applications to effectively reduce material consumption and create parts of high quality. Traditional methods to correct for imperfections involve experienced operators who will manually and iteratively adjust process parameters such as printing speed, flow rate, or fan speed during printing.

* Corresponding author.
E-mail address: ggu@berkeley.edu (G.X. Gu).

https://doi.org/10.1016/j.addma.2020.101696
Received 26 April 2020; Received in revised form 29 October 2020; Accepted 3 November 2020
Available online 12 November 2020

However, this trial-and-error approach can be time-consuming and inefficient. With the help of new methods such as embedded acoustic sensors, computer vision, and machine learning, the newly integrated system can learn the internal relationship between the sensor signals and working condition information. These systems are then able to automate analysis and conduct corrections during the manufacturing process. For instance, Wu et al. developed an in situ and real-time FFF monitoring system via acoustic emission sensors and hidden semi-Markov models to detect the condition of FFF machines [25–27]. Narayanan et al. applied principal component analysis (PCA), support vector machine (SVM), and convolutional neural network (CNN) machine learning models to distinguish defective parts from good quality parts based on camera images [28]. These methods provide promising results in successfully detecting imperfections during the FFF process. However, current challenges lie in localizing the defects both accurately and efficiently. These two aspects are very critical to understanding the spatial information of the defects to develop an in-situ auto-correction system that can operate without trial-and-error processes.

In this paper, detailed spatial information of in-plane anomalies is studied based on real-time images captured by a camera attached to the extruder. Two detection methods including precise localization and semantic segmentation are conducted in both local and global frameworks. Specifically, a local framework refers to the angle of view from the camera and a global framework refers to the coordinate system based on the printer (print bed). Localization uses bounding boxes to outline the loose location of the defects while semantic segmentation applies masks to shade the exact area of interest. For the localization problem, solutions are inspired by object detection methods applied in the autonomous vehicle industry. A modified YOLO (you only look once) v3 algorithm is conducted to detect the previous layer and current layer defects using bounding boxes. YOLO is a state-of-the-art object detection system, where one-stage learning is developed to provide end-to-end detection. Compared to the two-stage detectors where a set of the potential area is predicted by the first model and a second classifier is then trained based on the region candidates, one-stage model directly predicts the bounding boxes and their confidence in a faster and more efficient way. The result of YOLO outperforms many other detectors such as deformable part models (DPM) and region based convolutional neural networks (R-CNN) based on standard image datasets (Picasso and People-Art Dataset) [29,30]. In terms of semantic segmentation, the aim of the task is to understand the image at pixel-wise level and segment the areas for different categories. Here, a modified DeepLabv3 architecture is applied to realize semantic segmentation for the previous layer and current layer conditions using colored masks. The first version of DeepLab innovatively uses atrous convolution structure to set a new achievement on the PASCAL VOC-2012 Challenge (Dataset) [31,32]. A pre-trained model is used in each task since it keeps the feature extraction at the front part of the model and only the end part of it is required for training, where the method has been verified in many studies [33,34]. After the semantic segmentation model is trained, predictions of pixel-level quality distribution are made on every input frame of images. Layer-wise images in the global framework can be rebuilt based on coordinate transformations between the local framework and the printer coordinate system. Hence, a layer-wise quality profile for the printing process is achieved in the global framework. Our work has the following contributions: 1) Development of a multi-scale framework capable of detecting defects locally in real-time printing; 2) Realization of an efficient anomaly detection model that can decrease the response time of defects recognition; 3) Integration of layer-wise quality profiles that can aid in the monitoring and correction of additive manufacturing systems. The paper is organized as follows. Section 2 discusses the methods used for localization and semantic segmentation tasks in detail as well as three result visualization approaches. Section 3 shows the results of the whole model generating process including data collection, model training, and evaluation. Additionally, typical local and global images are displayed to visualize the predictions and discussion of the results are also presented in this section. Section 4 summarizes the work, raises current challenges, and proposes future plans.

2. Material and methods

In this section, two modified machine learning algorithms will be analyzed in detail on their working principles. Moreover, three visualization methods of localization results are introduced in this part including labeling the bounding box prediction, highlighting the high-performance area, and sensitivity analysis.

2.1. Localization model - you only look once (YOLO)

The YOLO method provides real-time object detection capability that uses images as input and produces predictions through bounding boxes. YOLOv3 performs feature extraction using architecture Darknet-53, which can be seen in Fig. 1(a). At the back of each detection output (image with grids), the algorithm will predict two maps. The first one has the information about the location and size of the bounding box (four variables) plus a confidence score (one variable) describing overlap of the predicted bounding box against the labeled one (ground truth). For each grid, three bounding boxes can be predicted. The second map is the distribution of class probability, as each grid will be classified into a category [29]. Therefore, the predictions can be expressed as an $N \times N \times (3*5+C)$ tensor, where N is the number of grids on each side of the square image, and C is the number of classes. In our localization problem, N has three values ($N = 14$, 28, 56 at model layer number 82, 94, and 106 correspondingly) to detect objects with different sizes through either finer or looser grids. Here, C equals three which refers to three classes in the localization problem: previous layer, over-extrusion in the current layer, and under-extrusion in the current layer. In this paper, 'Previous-layer', 'Over-extrusion', and 'Under-extrusion' are used to stand for these three categories. Printing conditions in the previous layer are not included in this method because they will be covered by the current layer. Therefore, only previous layer detection is focused on to distinguish it from the current prints.

2.2. Localization results visualization

In this work, three visualization methods are applied to present the localization results. The first one directly shows the predicted bounding boxes based on the model output tensor mentioned in the previous section. To interpret the prediction tensor as a final result, the non-maximum suppression (NMS) method is applied to remove the redundant and inaccurate bounding boxes based on a threshold [35]. In this study, the threshold is set to 0.5, which is the standard. The second method colors the grid based on an integrated score, which is defined as the element-wise product of confidence and class probability obtained from two prediction maps of the model. The integrated score represents both how well the predicted bounding box matches the ground truth and the probability of the class existing in the box. The grids at the boundary of the detected features could have a high score for correctly predicting the classes, yet have a low confidence value, which leads to a modest integrated score. In this case, grids in the middle of the features would have higher integrated scores because they have better performance in both values. The third method is known as sensitivity analysis, which evaluates the model's local gradient and determines the location of the input image that is most sensitive [36]. The gradient is evaluated at each pixel of the input image through backpropagation of the model output [37]. By setting a threshold for the gradient, high sensitivity areas can be displayed on the input images.

2.3. Semantic segmentation model - DeepLabv3

Using the localization method introduced in the previous

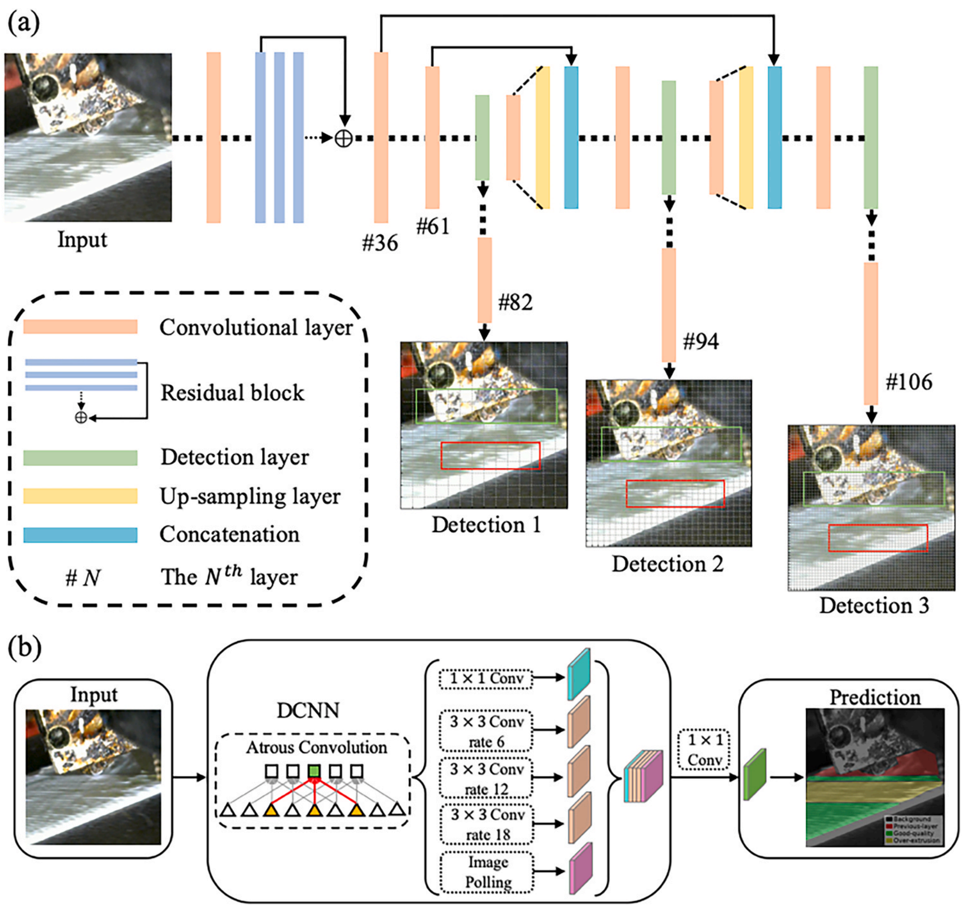

Fig. 1. (a) The schematic architecture for the modified you only look once (YOLO) machine learning model. The model has an input image with size 448 × 448 pixels and detections under three resolutions at model layer 82, 94, and 106. Critical model layers (layer 36 and layer 61) showing concatenation are also specially labeled in the image. Model layers with different functionalities are shown in different colors with the legend in the lower left quarter. (b) The flowchart diagram of the semantic segmentation DeepLabv3 model. The input image with 448 × 448 pixels size is fed into the deep convolutional neural network (DCNN) with atrous convolution structures, which have five different kinds of processing layers shown in the brace. The output of the model is a matrix predicting the category at pixel-wise resolution.

paragraphs, rough spatial information of the defects can be obtained. However, the bounding box still does not provide accurate detection of the area of interest. Therefore, the semantic segmentation method is applied to recognize the exact content of the image including the condition (good quality or anomalies), the location, and the number of defects. The basic idea of semantic segmentation involves using an encoder network such as fully convolutional networks (FCN) or deep convolutional neural networks (DCNN) to generate a pixel-wise prediction and then decode it to a masked segmentation image. The architecture applied in this paper is a new state-of-the-art method called DeepLabv3 shown in Fig. 1(b). The pooling and down-sampling operations in the normal DCNN would cause a serious reduction in spatial information. Therefore, atrous convolution with different sampling rates are proposed to re-scale the vision field and maintain the resolution [31, 32]. The output of the model is a matrix with a pixel-wise prediction of the category. Each element in the matrix represents the classification result of the corresponding pixel in the image using a normalized array (summation of the elements in the array equals to 1) via the Softmax operation, indicating the possibility of belonging to each class. The output classes are set to five in this problem, which are 'Background', 'Previous-layer', 'Good-quality', 'Over-extrusion', and 'Under--extrusion'. It is important to note that in the semantic segmentation problem 'Background' category is added due to the working principle of the task to classify the image into pixel level and class. An additional category of 'Good-quality' is augmented, compared to the categories of the localization problem, to detect the change of printing conditions in the current layer.

3. Results and discussion

3.1. Image data set preparation

In this paper, the PRUSA i3 MK3 FFF 3D-printer and polylactic acid (PLA) filament are used to fabricate testing samples. In order to collect image data with different types of anomalies, a USB-camera is first attached to the extruder through a 3D-printed camera mount, with more details shown in Fig. 2(a). The setup provides a fixed filming view (highlighted in yellow shadow) to capture and monitor the printing process in real-time. A cropped physical display of the field of view can be referred to the input image in Fig. 1. In this view, the print bed (background) as well as the previous and current layers of the prints are visible. To create uniform mechanical properties of the final products, two orthogonal printing directions are set up with one direction for odd layers and the other for even layers. After that, different flow rates are manually adjusted to create two types of quality conditions (over-extrusion or under-extrusion) in both even and odd layers. The first type starts from good-quality, then turns into over-extrusion and finally changes back to good-quality, while the second type follows the same process with the under-extrusion defect. Although over and under-extrusion defects can form as a result of multiple process parameters such as print speed or layer height, it is mainly caused by the improper rate of material extrusion (flow rate). Due to the fact that defects can be more local and not uniformly distributed, the extent of intra-layer defects in our training data is changed by adjusting the flow rate. Therefore, our method is capable of mimicking the actual defects that may occur in real-time printing conditions. By printing several rectangular sheets $(100\ mm \times 25\ mm \times 0.6\ mm,\ 3\ \text{layers})$ under the above procedures, 20000 images are recorded during the process. Among them, 1400

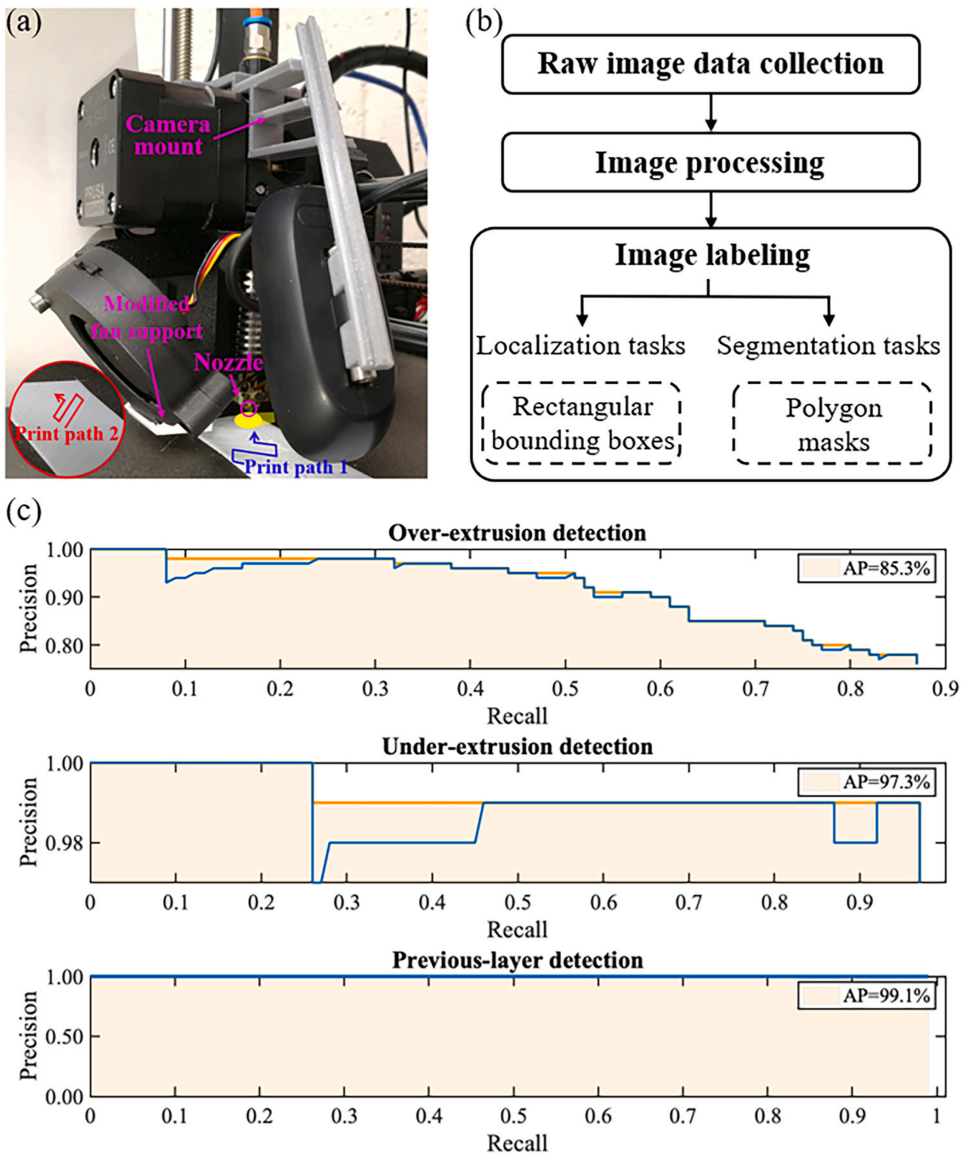

Fig. 2. (a) Experimental setup of the anomaly detection system. A USB camera is attached to the extruder through the camera mount. The camera has a filming angle of 30 degrees downwards with respect to the horizontal line. Two kinds of print paths for odd and even layers are labeled in blue and red arrows. (b) The flow diagram summarizing the process of image data set preparation. The image processing step involves locating the area of interest and cropping into desired image size. (c) The Precision-Recall curve under three detection conditions. Raw data is shown in the blue curve and interpolated versions are marked in orange making the curve monotonically decreasing. The area below the orange curve is shaded which represents the accuracy precision (AP) of the model in particular cases.(For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

images containing quality transitions are selected as our entire image data set. To be identified as a quality transition, an image must contain two or more printing quality conditions. To roughly locate the quality transition images, image files are named with a sequence number and real-time flow rate value. After locating the image file based on the changed flow rate, a more accurate determination will be made based on visual inspection since there will be a delay between the appearance of defects and printing parameter changes. With the selected images, a fixed square window is applied to extract 448×448 pixels images with the print nozzle located in the center. Since there are two orthogonal printing directions in the process (a single printing direction for all the odd layers and an orthogonal direction for all the even layers), the square window has two orientations during the extraction of input images for odd or even layers. Lastly, 20% of the whole image data set is randomly chosen as testing data, and the remaining images are treated as training data.

After preparing the whole raw data set, images are then labeled by using an open-source python software (labelme) [38] for localization and segmentation tasks in two different ways. For the localization problem, rectangular bounding boxes are used to label the three types of conditions of interest. Here, the coordinates of the bounding box and their categories are recorded. For the segmentation problem, polygons are drawn on the images to create masks for the five categories and each pixel will be allocated a category. A flowchart summarizing the image data set preparation process is shown in Fig. 2(b). It is worth noting that the labeling process can be time-consuming especially sketching the outlines of the polygon masks in the segmentation task at the beginning of the labeling procedure. With more experience, the time needed for the labeling process can be reduced. During the labeling process, constant judgment based on image features is kept to distinguish the anomalies from good quality and all the labeling task is conducted by one person to minimize the random noise created in the labeling procedure. Additionally, following the same image data preparation procedure, augmented data sets can be created using different colors of filament or infill patterns in future work to expand the input feature of the machine learning model.

3.2. Precise localization detection

To locate the position of the defect in the image, a bounding box is used to mark the information. Therefore, the task of the machine learning model is to predict the location and size of the bounding box from the input image. As mentioned in the introduction section, a modified YOLOv3 network and a pre-trained model are applied to this task. The training process continues for 9000 epochs until the loss converges. The training results are evaluated through mean accuracy precision (mAP), which measures the performance of an object detection method [39]. In order to get mAP, accuracy precision (AP) [40,41] for each class is needed. The term is further calculated as the area under the interpolated precision-recall curve shown in Fig. 2(b). Note, the orange curve is the interpolated smoothing of the blue curve (raw data), and the shaded area below the orange curve represents the accuracy precision. The precision (P) and recall (R) can be calculated based on the following equations: $P = TP/S_1$ and $R = TP/S_2$. Here, S_1 stands for the total number of detected bounding boxes and S_2 represents the total number of bounding boxes in the dataset. True positive (TP) represents the number of correctly detected bounding boxes, which have the right detection of a category and IOU larger than 50%. Intersection of union (IOU) refers to the overlap area of the predicted bounding box divided by the union of ground truth and prediction. A threshold of 50% is chosen according to the general rule of bisection and consistency of the threshold to compare among other models. A higher threshold baseline means a more restricted condition for a correctly detected case. Finally, the accuracy precision for three categories based on the testing data set are shown in Fig. 2(b) with its value shown in the legend. Therefore, the model's mAP can be calculated as the mean value of the AP results, achieving a performance of 93.9%. A summary of the training results for the localization model is also presented in Table 1.

The evaluation metric mAP explored in the last paragraph shows the overall performance of the applied model. In order to give an intuitive and clear display, typical images under different categories are analyzed and shown in Fig. 3. In the 'Bounding box prediction' row, both the ground truth and the predicted bounding box are labeled with different colors on the example images. Each color represents one category and detailed information is shown in the caption. The 'Result visualization' row shows how well the model predicts the category and location of the cases. As mentioned in the second section, the model will conduct an output at each YOLO detection layer. Here, the first prediction at model layer 82 with a stride of 32 (14×14 grids) is studied for its appropriate resolution for visualization. The matrices are further normalized to (0, 1) interval for standardization. As mentioned in the previous paragraph, the integrated score is defined as the product of confidence and class probability. The maximum product among the three predictions determines the value of each grid. Since both confidence and class probability are viewed as correctly predicted if the value exceeds 0.5, a threshold of the integrated score is set as $0.5 \times 0.5 = 0.25$. Grids with a value larger than 0.25 will be colored, which means the grid has a higher possibility to correctly predict both the class and the location of the bounding box. As shown in the 'Result visualization' row of Fig. 3, the colored grids are all located around the center of the bounding box giving a consistent match between results and theory.

Sensitivity analysis is applied to visualize the high gradient location on input images. Note, each pixel has three gradient values for the image. The average of the three R, G, B channels determines the overall gradient of the pixel. The image is further divided into 14×14 grids to maintain display clarity and consistency. The gradient of the grid is defined as the mean value of all the pixels within that grid. The grid will be colored in grey if the value exceeds a certain threshold as shown in the last row of Fig. 3. Sensitivity analysis highlights the area based on the gradient evaluation at each pixel of the input image. A pixel with a larger gradient indicates that it has a higher sensitivity in the classification task rather than a larger probability in a certain category. Therefore, there are no governing rules on setting the threshold. Hence, an appropriate and fixed value of a threshold that satisfies the majority of the testing images is applied in the visualization of the sensitivity analysis. Based on the sensitivity analysis result studied on MNIST database shown in the literature [36], it can be concluded that the heatmaps representing the sensitivity are spatially discrete and distributed at the boundary of the digits. Similarly, the high sensitivity grids (shaded in grey) also lie in the background, outside of the prints boundary, as well as at the intersection line between the previous and current layer. The distribution of the shaded grids implies that these areas have a greater impact on the output of the model, which is consistent with the individual's subjective perception, that edges and boundaries are normally the priority area to focus on.

Table 1

Summary of the training results of precise localization and semantic segmentation detection models. AP_1, AP_2, and AP_3 stand for the accuracy precision of over-extrusion, under-extrusion, and previous-layer case, respectively. mAP is the mean value of all the AP results. PA and mIOU represent the pixel accuracy and mean intersection over union for the semantic segmentation model.

	Precise localization				Semantic segmentation	
Metric	AP_1	AP_2	AP_3	mAP	PA	mIOU
Performance	85.30%	97.30%	99.10%	93.90%	97.60%	92.10%

3.3. Semantic segmentation detection

Although the localization detection is able to provide a loose location of the defect, large amounts of useless information such as the print bed background and the print nozzle are retained due to the rectangular bounding box shape. Therefore, semantic segmentation is applied to obtain a more precise detection of the defect using polygons to mask the desired area. As mentioned in the introduction section, a modified DeepLabv3 architecture and a pre-trained model are applied in this section. With the pre-trained model, all the weight matrices except for the last layer will be fixed and the training results converge quickly after 15 epochs shown in Fig. 4(a). Two metrics are studied to evaluate the performance of the segmentation model, which are pixel accuracy (PA) and mean intersection over union (mIOU). Since semantic segmentation focuses on pixel-wise classification, pixel accuracy is the percentage of correctly predicted pixels over the total number of pixels (448×448). To obtain mIOU, intersection over union (IOU), which is the overlap area of the predicted mask and the ground truth divided by the union of the two areas, is first calculated for each class. The average value over all five classes is the mIOU of the model. After 30 epochs of training, both PA and mIOU converge and reach 97.6% and 92.1% respectively as shown in Fig. 4(a). The training results of the semantic segmentation model are also listed in Table 1. The training results show that the semantic segmentation model is able to detect the category correctly and accurately at the pixel level. To present the results in a more intuitive sense, four typical images with different defect information are shown in Fig. 5. In the first row, the ground truth is labeled on the original image with different colors representing each category. In the second row, the predicted outputs are masked using the same color scheme and show a significant match compared to the ground truth. As mentioned in the introduction section, one major objective of the semantic segmentation is to reduce the time latency of detecting the defect. Therefore, more attention is focused on the single-raster case (the rightmost column of Fig. 5) that is present at the transition period. Specifically, it refers to the single raster that is currently being printed has a different quality from the previous rasters in the same layer. Multiple-rasters refer to images with quality change over several lines of raster ('Under-extrusion' case in the even layer condition of Fig. 5).

Among the testing data set, 42 images are under the single-raster condition and the remaining have multiple-rasters case. As shown in Fig. 4(b), the IOU performance on single-raster cases primarily lies in the range of 0–75%; however, the results shift to a higher percentage (50%–100%) for multiple-raster cases. This implies that the model has lower confidence in accurately locating the position of the single-raster; however, during the actual testing, even if a low percentage of IOU is detected, the prediction still signals that quality transition is occurring and actions should be taken if the IOU keeps increasing. The detection of existence has a response time within one second and judgment can be

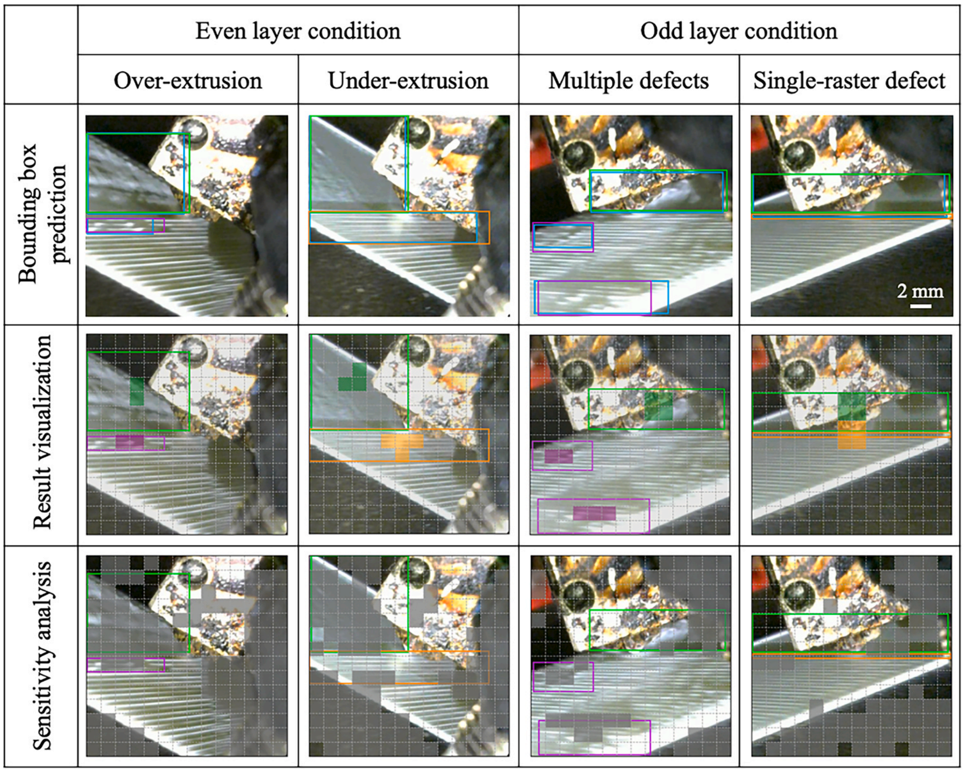

Fig. 3. Three types of results visualization and analysis (bounding box prediction, results visualization, and sensitivity analysis) are displayed for both even and odd layers with four conditions of defects: general over and under-extrusion, multiple defects, and single-raster defect. Color of bounding boxes: blue-ground truth, green-predicted 'Previous-layer', purple-predicted 'Over-extrusion', and orange-predicted 'Under-extrusion'.(For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

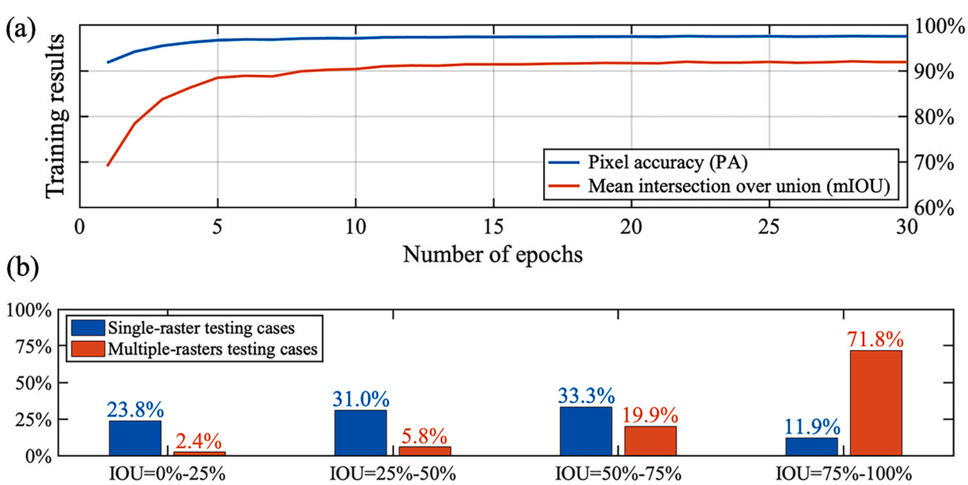

Fig. 4. (a) The performance of semantic segmentation model within 30 epochs of training. The blue curve converging at 97.6% represents pixel accuracy (PA) and the orange one reaching 92.1% shows the result of mean intersection over union (mIOU). (b) The statistical diagram shows the distribution of intersection over union (IOU) performance for single-raster testing cases (blue) and multiple-rasters testing cases (orange).(For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

validated in the following two or three seconds. This fast reaction capability provides more efficient feedback to realize the automatic control of the printing system.

3.4. Layer-wise quality profile

Both methods mentioned above show the capability of detecting defects locally in the view of the camera. In order to analyze and fix the defects, exact global coordinates in the actual prints are more practical for further correction and examination. Therefore, the image stitching method is applied to fulfill the global recombination of the images. During the fabrication process within one layer for complex print geometries, the printing is conducted in multiple regions which leads to the extruder nozzle jumping across the piece, which in turn creates discontinuities in the images. Thus, image alignment methods based on consecutive images and feature extraction are not applicable. Instead, with the known global coordinates of the nozzle, affine transformations can be calculated and applied to every local image, mapping them to the global frame. A schematic diagram illustrating the affine transformation is shown in Fig. 6(a). A standard right triangle is printed with the right-angle side L mm and parallel to the edge of the print bed. Since the information of the nozzle location can be obtained from the printer, we can assume its global coordinates are denoted as $O(O_X, O_Y)$, the other two vertex coordinates can be expressed as $A(O_X - L, O_Y), B(O_X, O_Y + L)$. Meanwhile, the same three points can also be written as $o(o_x, o_y)$, $a(a_x, a_y)$, and $b(b_x, b_y)$ in local image coordinates. Therefore, the affine

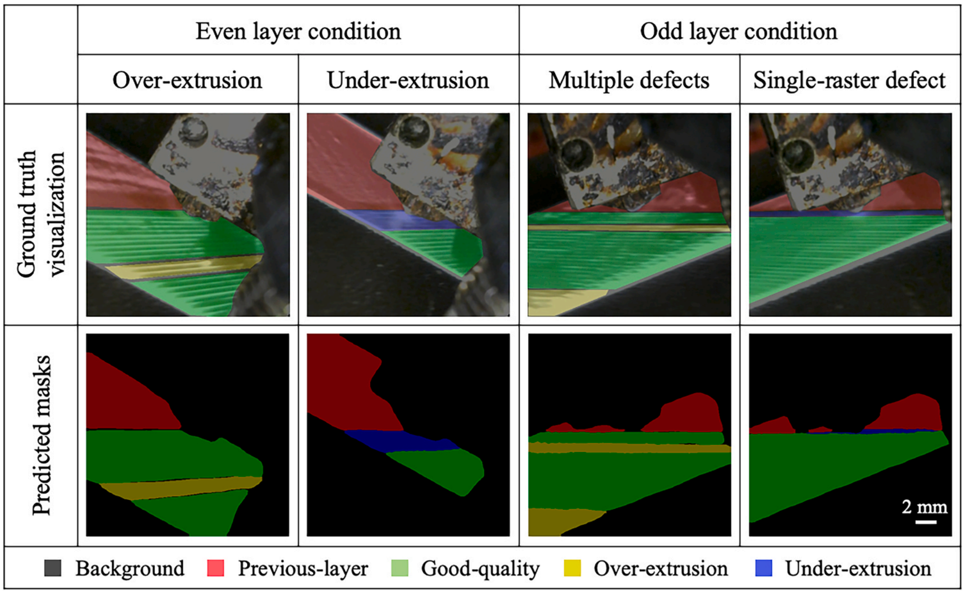

Fig. 5. Ground truth and predicted masks visualization in both even and odd layers with four types of defects. Categories are masked in specific colors: black-background, red-'Previous-layer', green-'Good-quality', yellow-'Over-extrusion', and blue-'Under-extrusion'.(For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

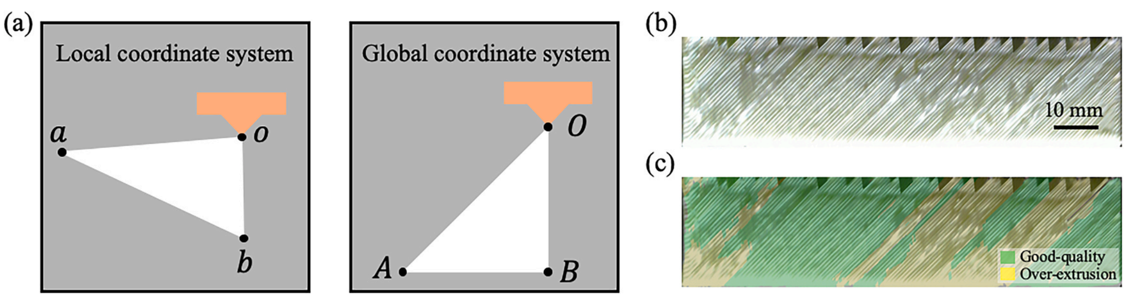

Fig. 6. (a) Diagram depicting affine transformation from the local to the global coordinate system. Triangle vertices with lower-case letters in the local framework map to the same capitalized letters in the global framework. (b) Local image stitching in global view via affine transformation. (c) Semantic segmentation prediction is visualized in the global framework with a green mask representing 'Good-quality' and yellow mask standing for 'Over-extrusion'.(For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

transformation can be expressed in the following equation:

$$\begin{bmatrix} O_X \\ O_Y \\ A_X \\ A_Y \\ B_X \\ B_Y \end{bmatrix} = \begin{bmatrix} O_X \\ O_Y \\ O_X - L \\ O_Y \\ O_X \\ O_Y + L \end{bmatrix} = \begin{bmatrix} A & \underline{0} & \underline{0} \\ \underline{0} & A & \underline{0} \\ \underline{0} & \underline{0} & A \end{bmatrix} \begin{bmatrix} o_x \\ o_y \\ a_x \\ a_y \\ b_x \\ b_y \end{bmatrix} + \begin{bmatrix} B \\ B \\ B \end{bmatrix}, \tag{1}$$

where $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, B = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$, and $\underline{0} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ Since all six points are neither collinear nor coincident and bijection mapping is followed, there exists a unique solution to the affine matrix A and bias B. Thus, any point (x, y) in the local image frame with a new nozzle coordinate $O'(O_X', O_Y')$ can be written in the global coordinate (x', y') as follows:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} + \begin{bmatrix} O_X' - O_X \\ O_Y' - O_Y \end{bmatrix} \tag{2}$$

Since the global view can be directly applied to the top layer of the prints afterward, attention is primarily focused on an internal layer in this section. After getting the affine matrix, an internal even layer extrusion sample is analyzed as follows. The raw image is displayed by stitching the transformed images together in the global frame shown in Fig. 6(b). For the overlapping area, images behind this area would cover the previous features. The indented triangles at the top of the image are shadows created by the sensors near the nozzle. The local images at the top edge are all affected by this problem. Adding a more uniform light source is believed to remedy this deficiency. With the trained semantic segmentation model, predictions can be obtained from every local image and follow the same affine transformation procedure. However, in terms of overlapping, the prediction of the pixel is determined by the highest category probability among the overlapped images, where the probability of each category at one pixel is mentioned in Section 2. The final prediction results in the global view are presented in Fig. 6(c) showing the 'Over-extrusion' region in yellow and the 'Good-quality' area in green. Since the post-processing layer-wise quality profile aims to provide an efficient visualization of segmentation results in the global framework, labeling the ground truth of thousands of local input images are not practical in the actual printing process. Moreover, for internal layer cases, labeling ground truth on a rebuilt global image can involve errors and inconsistent labeling settings. In the current study, the efficacy of the method is evaluated based on the detection of features from visual inspection. In this case, the prediction of semantic segmentation model on layer-wise quality profile matches the features of the over-extrusion defect such as bulges or raised rasters located in the middle part and corner of the prints. In the future, intermediate steps will be added to pause the printing process during the layer transition for image capturing. Another camera will be equipped on the print bed to capture the global layer-wise image, which will be further labeled as the ground truth. Overall, global visualization provides a comprehensive understanding of the physical location of the defects, especially internally, which makes the examination and correction easier to locate the area of interest. The global prediction mapping to the physical prints rebuilt from local images bridges the gap between the model output (plain numbers) and actual defects, thereby realizing a real-time and layer-wise anomaly detection in the quality control of additive manufacturing processes.

4. Conclusions

In summary, two advanced real-time detection methods are developed in this paper to precisely locate the area of interest during the printing process both in the local and global frameworks. Mean average precision reaches 93.9% for the localization model and mean intersection over union achieves 92.1% for semantic segmentation. Typical images are also displayed using three visualization methods to verify and understand the high accuracy of prediction results. The layer-wise quality profile rebuilds the local images in their entirety and is able to assess the in-plane defects internally. Additionally, a single-raster condition in the segmentation task is exclusively studied and analyzed to demonstrate the capability of detecting new printing conditions in-situ and in real-time. Although the response time for detection is curtailed, the time latency within the printing hardware (from sending the order to actually changing the printing parameter) remains a challenge. Future work may include improving the printing quality and predicting the mechanical properties of the prints based on the stacked spatial information of defects. Moreover, since process parameters may play a critical role in dictating the quality of printed parts, it is important to understand the dominant factors and coupling relationships among these process parameters (print speed, layer height, printing temperature) on the printed products. An interesting next step will be developing principal component analysis (PCA) and clustering techniques to reveal the effects of process parameters for unsupervised monitoring of a wide variety of defects in the additive manufacturing process. Last but not least, utilizing the framework established in our work, other solid infill printing situations including different printing materials or printing directions can be utilized through augmenting the data set with newly defined settings and re-training the machine learning model. Recent developments in transfer learning methods [42–45] are believed to be a useful approach to accelerate the re-training process.

CRediT authorship contribution statement

Zeqing Jin: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing - original draft. **Zhizhou Zhang**: Conceptualization, Data curation, Investigation, Methodology, Writing - original draft. **Joshua Ott**: Data curation, Investigation, Methodology, Writing - original draft. **Grace X. Gu**: Conceptualization, Investigation, Methodology, Writing - original draft.

Declaration of Competing Interest

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

Acknowledgements

The authors acknowledge support from the Extreme Science and Engineering Discovery Environment (XSEDE) at the Pittsburgh Supercomputing Center (PSC) by National Science Foundation grant number ACI-1548562. Additionally, the authors acknowledge support from Amazon, Chau Hoi Shuen Foundation, Johnson & Johnson, and NVIDIA.

References

[1] D. Lehmhus, T. Wuest, S. Wellsandt, S. Bosse, T. Kaihara, K.-D. Thoben, M. Busse, Cloud-based automated design and additive manufacturing: a usage data-enabled paradigm shift, Sensors 15 (12) (2015) 32079–32122.
[2] L. Wang, P. Orban, A. Cunningham, S. Lang, Remote real-time CNC machining for web-based manufacturing, Robot. Comput. Integr. Manuf. 20 (6) (2004) 563–571.
[3] G.X. Gu, S. Wettermark, M.J. Buehler, Algorithm-driven design of fracture resistant composite materials realized through additive manufacturing, Addit. Manuf. 17 (2017) 47–54.
[4] D.J. Roach, C.M. Hamel, C.K. Dunn, M.V. Johnson, X. Kuang, H.J. Qi, The m4 3D printer: a multi-material multi-method additive manufacturing platform for future 3D printed structures, Addit. Manuf. 29 (2019), 100819.
[5] G.X. Gu, C.-T. Chen, D.J. Richmond, M.J. Buehler, Bioinspired hierarchical composite design using machine learning: simulation, additive manufacturing, and experiment, Mater. Horiz. 5 (5) (2018) 939–945.
[6] M.A. Skylar-Scott, J. Mueller, C.W. Visser, J.A. Lewis, Voxelated soft matter via multimaterial multinozzle 3D printing, Nature 575 (7782) (2019) 330–335.
[7] G.X. Gu, M. Takaffoli, M.J. Buehler, Hierarchically enhanced impact resistance of bioinspired composites, Adv. Mater. 29 (28) (2017), 1700060.
[8] N.T. Aboulkhair, N.M. Everitt, I. Ashcroft, C. Tuck, Reducing porosity in AlSi10Mg parts processed by selective laser melting, Addit. Manuf. 1 (2014) 77–86.
[9] D. Pang, H. Thomas, D. Jacobs, F. Paul, StereoLithography 1993: QuickCast TM, 1993 International Solid Freeform Fabrication Symposium, 1993.
[10] A. Armillotta, M. Bellotti, M. Cavallaro, Warpage of FDM parts: experimental tests and analytic model, Robot. Comput. Integr. Manuf. 50 (2018) 140–152.
[11] L. Cheng, X. Liang, J. Bai, Q. Chen, J. Lemon, A. To, On utilizing topology optimization to design support structure to prevent residual stress induced build failure in laser powder bed metal additive manufacturing, Addit. Manuf. 27 (2019) 290–304.
[12] Z. Zhang, K.G. Demir, G.X. Gu, Developments in 4D-printing: a review on current smart materials, technologies, and applications, Int. J. Smart. Nano. Mater. 10 (3) (2019) 205–224.
[13] L. Scime, J. Beuth, Anomaly detection and classification in a laser powder bed additive manufacturing process using a trained computer vision algorithm, Addit. Manuf. 19 (2018) 114–126.
[14] R. Anitha, S. Arunachalam, P. Radhakrishnan, Critical parameters influencing the quality of prototypes in fused deposition modelling, J. Mater. Process. Technol. 118 (1) (2001) 385–388.
[15] L. Rebaioli, P. Magnoni, I. Fassi, N. Pedrocchi, L.M. Tosatti, Process parameters tuning and online re-slicing for robotized additive manufacturing of big plastic objects, Robot. Comput. Integr. Manuf. 55 (2019) 55–64.
[16] V. Kishore, C. Ajinjeru, A. Nycz, B. Post, J. Lindahl, V. Kunc, C. Duty, Infrared preheating to improve interlayer strength of big area additive manufacturing (BAAM) components, Addit. Manuf. 14 (2017) 7–12.
[17] E.C. Balta, D.M. Tilbury, K. Barton, A Digital Twin Framework for Performance Monitoring and Anomaly Detection in Fused Deposition Modeling, 2019 IEEE 15th International Conference on Automation Science and Engineering (CASE), IEEE, 2019, pp. 823–829.
[18] J. Li, H. Xie, K. Ma, In-situ monitoring of the deformation during Fused Deposition Modeling process using CGS method, Polym. Test. 76 (2019) 166–172.
[19] J.L. Bartlett, B.P. Croom, J. Burdick, D. Henkel, X. Li, Revealing mechanisms of residual stress development in additive manufacturing via digital image correlation, Addit. Manuf. 22 (2018) 1–12.
[20] A. Kelkar, N. Kumbhar, A. Mulay, Surface roughness measurement of parts manufactured by FDM process using light sectioning vision system, J. Inst. Eng. Ser. C 99 (4) (2018) 429–433.
[21] Z. Jin, Z. Zhang, G.X. Gu, Autonomous in-situ correction of fused deposition modeling printers using computer vision and deep learning, Manuf. Lett. 22 (2019) 11–15.
[22] Z. Jin, Z. Zhang, G.X. Gu, Automated real-time detection and prediction of interlayer imperfections in additive manufacturing processes using artificial intelligence, Adv. Intell. Syst. (2019), 1900130.
[23] K. Fayazbakhsh, M. Movahedi, J. Kalman, The impact of defects on tensile properties of 3D printed parts manufactured by fused filament fabrication, Mater. Today Commun. 18 (2019) 140–148.
[24] J.M. Gardner, K.A. Hunt, A.B. Ebel, E.S. Rose, S.C. Zylich, B.D. Jensen, K.E. Wise, E. J. Siochi, G. Sauti, Machines as craftsmen: localized parameter setting optimization for fused filament fabrication 3D printing, Adv. Mater. Technol. 4 (3) (2019), 1800653.
[25] H. Wu, Z. Yu, Y. Wang, A new approach for online monitoring of additive manufacturing based on acoustic emission, ASME 2016 11th International Manufacturing Science and Engineering Conference, American Society of Mechanical Engineers Digital Collection, 2016.
[26] H. Wu, Z. Yu, Y. Wang, Real-time FDM machine condition monitoring and diagnosis based on acoustic emission and hidden semi-Markov model, Int. J. Adv. Manuf. 90 (5–8) (2017) 2027–2036.
[27] H. Wu, Y. Wang, Z. Yu, In situ monitoring of FDM machine condition via acoustic emission, Int. J. Adv. Manuf. 84 (5–8) (2016) 1483–1495.
[28] B.N. Narayanan, K. Beigh, G. Loughnane, N. Powar, Support vector machine and convolutional neural network based approaches for defect detection in fused filament fabrication, SPIE, 2019.
[29] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, You only look once: Unified, real-time object detection, Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 779–788.

[30] J. Redmon, A. Farhadi, Yolov3: Increm. Improv., (2018) arXiv preprint arXiv:1804.02767.
[31] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, A.L. Yuille, Deeplab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs, IEEE Pami. 40 (4) (2017) 834–848.
[32] L.-C. Chen, G. Papandreou, F. Schroff, H. Adam, Rethink. Atrous Convolution Semant. Image Segm. (2017) arXiv preprint arXiv:1706.05587.
[33] A. Canziani, A. Paszke, E. Culurciello, Anal. Deep Neural Netw. Models Pract. Appl. (2016) arXiv preprint arXiv:1605.07678.
[34] W. Rawat, Z. Wang, Deep convolutional neural networks for image classification: a comprehensive review, Neural Comput. 29 (9) (2017) 2352–2449.
[35] R. Rothe, M. Guillaumin, L. Van Gool, Non-maximum suppression for object detection by passing messages between windows. Computer Vision – ACCV 2014, Springer Cham, 2015, pp. 290–306.
[36] G. Montavon, W. Samek, K.-R. Müller, Methods for interpreting and understanding deep neural networks, Digit. Signal. Process. 73 (2018) 1–15.
[37] C.M. Bishop, Neural Networks for Pattern Recognition, Oxford University Press, 1995.
[38] K. Wada, labelme: Image Polyg. Annot. Python (2016). ⟨https://github.com/wkentaro/labelme⟩.
[39] M. Everingham, L. Van Gool, C.K. Williams, J. Winn, A. Zisserman, The pascal visual object classes (voc) challenge, Int. J. Comput. Vis. 88 (2) (2010) 303–338.
[40] G. Salton, M.J. McGill, Introduction to Modern Information Retrieval, McGraw-Hill, Inc, 1986.
[41] M. Ivašić-Kos, M. Krišto, M. Pobar, Human detection in thermal imaging using YOLO, Proceedings of the 2019 5th International Conference on Computer and Technology Applications, 2019, pp. 20–24.
[42] W. Dai, Q. Yang, G.-R. Xue, Y. Yu, Boosting for transfer learning, Proceedings of the 24th international conference on Machine learning, 2007, pp. 193–200.
[43] L. Torrey, J. Shavlik, Transfer learning, handbook of research on machine learning applications and trends: algorithms, methods, and techniques, IGI Glob. (2010) 242–264.
[44] S.J. Pan, Q. Yang, A survey on transfer learning, IEEE Trans. Knowl. data Eng. 22 (10) (2009) 1345–1359.
[45] K.K. Kim, I. Ha, M. Kim, J. Choi, P. Won, S. Jo, S.H. Ko, A deep-learned skin sensor decoding the epicentral human motions, Nat. Commun. 11 (1) (2020) 2149.