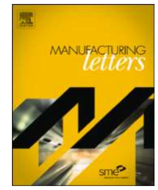
Available online at www.sciencedirect.com

Manufacturing Letters

Manufacturing Letters 41 (2024) 1526–1534




52nd SME North American Manufacturing Research Conference (NAMRC 52, 2024)

# A Deep Learning Framework for Automated Anomaly Detection and Localization in Fused Filament Fabrication

Sakib S Avro[a], S M Atikur Rahman[b], Tzu-Liang (Bill) Tseng[b], Md Fashiar Rahman[b]*

[a]*Aerospace and Mechanical Engineering, The University of Texas at El Paso, El Paso, TX, 79968, USA*
[b]*Industrial, Manufacturing and Systems Engineering, The University of Texas at El Paso, El Paso, TX, 79968, USA*

* Corresponding author. Tel.: +1-915-747-6903; fax: +1-915-747-7184. *E-mail address:* mrahman13@utep.edu

**Abstract**

Under-extrusion poses a common challenge in filament 3D printing, necessitating precise adjustments to printing parameters for optimal resolution. Swift identification of this flaw is crucial for implementing timely corrective measures. In response, we present a novel machine learning approach designed to detect anomalies in 3D printing, specifically in the fused filament fabrication process (FFF). Our framework utilizes "You Only Look Once" (YOLO), a real-time object detection system, and "Visual Geometry Group-16" (VGG-16), a Convolutional Neural Network (CNN) model for image recognition, to accurately identify and localize under-extrusion events. Initially, models such as VGG-16, VGG-19, and ResNet-50 were trained without the inclusion of YOLO to establish baseline accuracies. Subsequently, an automated image pre-processing phase employs YOLO to discern the nozzle head, facilitating subsequent cropping around the region of interest for retraining the models. This inclusion of the nozzle head detection approach significantly improved the accuracy of all models, with the combined application of YOLO and VGG-16 demonstrating the most substantial enhancement, boosting detection accuracy to 97%. Moreover, Gradient-weighted Class Activation Mapping (Grad-CAM) technique, another CNN-based method, is employed to effectively highlight predicted areas in under-extrusion scenarios. While our method significantly advances the automatic detection of printing anomalies, it primarily serves as a diagnostic tool, signaling the need for intervention. Our rigorous testing on images of varying complexities confirms the model's robustness, evaluated using comprehensive metrics such as precision, recall, and F1 score.



*Keywords:* Fused Filament Fabrication; Process Monitroing; Anamoly Detection; Deep Learning

## 1. Introduction

The advent of 3D printing, also known as additive manufacturing (AM), has revolutionized the design and production landscape by facilitating the creation of complex geometries and ushering in an era of manufacturing that prioritizes both precision and sustainability [1, 2]. Despite its transformative potential, AM faces inherent challenges, with under-extrusion being a critical flaw that undermines the structural integrity and aesthetic quality of printed objects [3]. Characterized by the insufficient deposition of material, under-extrusion leads to parts with weak structural points and surface irregularities, significantly compromising the functionality and durability of the final product [4]. This issue often arises from factors such as suboptimal printer parameters, inconsistent filament flow, and variable environmental conditions that impact the printing process [5]. FFF stands as the most widely used AM technology, holding more than 70% of the market. Its popularity stems from a straightforward setup and operational ease, coupled with the ability to work with a diverse range of thermoplastic filaments. This process involves the precise extrusion of melted filament material layer by layer to construct 3D objects, enabling detailed customization and rapid prototyping across industries [6].

Addressing under-extrusion is crucial, particularly due to its significant impact on the safety and performance of components in critical sectors such as aerospace, automotive, and healthcare [7, 8]. FFF's role in manufacturing critical components, like aerospace turbine blades, automotive structural parts, and medical implants, highlights the need for stringent quality standards in these industries. The work of researchers like Ram´ırez et al., and Pienaar et al., further emphasizes the importance of precision in producing these vital components, which range from intricate dental prostheses to lightweight yet durable parts for spacecraft [9, 10]. This underscores the imperative of resolving printing flaws to ensure the reliability and efficacy of products across these high-stakes fields. Any compromise in quality can lead to severe consequences, positioning the resolution of printing flaws like under-extrusion not only as a technical challenge but also as a matter of public safety [11, 12]. Moreover, in the rapidly evolving global manufacturing landscape, the agility and quality assurance that FFF technology offers represent crucial competitive advantages [13, 14]. Beyond the implications for quality, the economic repercussions of under-extrusion, including material waste and the costs associated with reprints and extended quality assurance processes, highlight the need for effective solutions to locate these issues [15].

Ensuring the quality of 3D printed products is paramount, and various innovative methods have been developed for identifying abnormalities. Selot et al., for example, explored acoustic monitoring, employing acoustic sensors to detect irregularities in the printing process [16]. Another notable approach is thermal imaging, utilized by Hossain et al., which focuses on detecting inconsistencies in the printing process, such as under-extrusion or overheating [17]. Yeong et al., introduced an onsite monitoring system that leverages computer vision and object detection models, to detect defects in real-time during FFF processes [18]. This system, combining camera technology with sophisticated algorithms, has demonstrated high accuracy in pinpointing anomalies, thereby facilitating immediate corrective actions. Among these varied techniques, image analysis methods, particularly those employing ML algorithms, have shown exceptional promise [19].

The research conducted by Liu et al. and Xu et al. serves as a testament to the adaptability and potential of ML in scrutinizing the manufacturing process of FFF printed products [20, 21]. Their work not only showcases the application of convolutional neural networks (CNNs) in detecting deficiencies within the FFF printing process but also highlights the broader applicability of CNNs for in-process quality control across various AM techniques. This underscores the practical relevance of CNNs in real-world manufacturing settings, illustrating their critical role in enhancing quality control measures and ensuring the reliability of printed products across the AM spectrum [22].

In the past few years, advancements in image processing and deep learning, especially through CNNs, have significantly advanced the fields of image classification and object recognition. These innovations have been effectively applied for identifying defects in the FFF process as well [23, 24]. Nonetheless, the process of training CNN models from the ground up is resource-intensive, requiring vast datasets, considerable computing power, and specialized knowledge, which particularly challenges Small and Medium-sized Enterprises (SMEs) due to the competitive nature of manufacturing and the scarcity of resources [25, 26]. To circumvent these hurdles, the strategy of deep transfer learning has emerged as a promising solution, enabling the enhancement of model accuracy with limited data by repurposing knowledge from one domain to address related challenges in another [27, 28]. Furthermore, the advent of Industry 4.0 has not fully permeated the SME sector, as well as among individual enthusiasts and hobbyists, who face barriers in adopting advanced technologies due to limitations in workforce, funding, and infrastructure. Recent efforts aim to bridge this gap by introducing cost-effective, user-friendly smart manufacturing solutions, such as smart factory implementations and retrofitting initiatives. These strategies are designed to facilitate the integration of digital technologies into existing manufacturing frameworks, promoting the digital transformation of production environments [29, 30].

In this study, we introduce a systematic deep transfer learning approach leveraging a collected dataset of FFF process images for the identification of printing anomalies. Initially, we assembled a dataset featuring images of defective extrusions and employed image augmentation techniques to generate training and validation images. This process was designed to accommodate variations in part shape and visual monitoring conditions. We then selected three widely recognized ImageNet pre-trained CNN architectures—VGG-16, VGG-19, and ResNet-50—for feature extraction purposes. Additionally, we devised an innovative nozzle head detection technique to automatically refine the initial dataset images. During the evaluation phase, we adjusted fine-tuning configurations to systematically optimize performance and computational efficiency. We also evaluated the computing time for each model, aiming to identify the most suitable models for different scenarios—specifically, those offering a balance between high performance and low computational demand.

## 2. Related work

### *2.1. Monitoring of part quality in FFF using CNNs*

The advancement of deep learning has sparked numerous studies focused on vision-based, end-to-end applications in FFF, primarily utilizing CNNs for detecting anomalies through image samples under normal and under-extrusion conditions [31]. A significant hurdle in these endeavors is the precise identification of filament and lamina conditions during the printing process. Jin et al. devised a real-time monitoring system by integrating a CNN model with a vision system, achieving over 96% accuracy in detecting under or over-extrusion and interlayer defects [27]. Banadaki et al. also developed a real-time part quality monitoring system, employing a deep CNN trained on images of printed layers to categorize surface quality into five levels with an average accuracy of 94%. Nevertheless, these studies focused on simple zig-zag patterns, omitting complex engineering features such as holes, fillets, and diverse surfaces [23]. Further efforts include using CNNs and vision modules to monitor part failures under unforeseen conditions. Kim et al. utilized a webcam-based system and a 19-layer CNN to classify spaghetti-shape errors in the FDM process with 96% accuracy [28]. Similarly, Paraskevoudis et al. presented a single-shot defect detection model that identified stringing errors with 92% accuracy in real-time [24]. Despite these successes, the studies rely on custom-built models trained on specific datasets, posing challenges in model replication without access to similar datasets. The additive manufacturing field often lacks comprehensive, organized datasets of process images, highlighting the necessity for a methodical approach to developing deficiency monitoring systems via deep transfer learning. Such systems would address dataset size, computing power, and expertise constraints prevalent in actual manufacturing environments.

### 2.2. Transfer learning approach using CNNs

Transfer learning is a key technique in ML where knowledge gained from solving one problem is reused to address a similar, but new, problem [32]. Specifically, in deep learning, which involves using CNNs, transfer learning helps overcome the difficulties of training networks with multiple layers on new datasets. This method is effective because it utilizes the 'weights' (the learned features) from models that were originally trained on large and diverse datasets, such as ImageNet. This approach is akin to inductive learning, where the model applies known information to new tasks [33]. The main benefit here is that these pre-trained models can rapidly and efficiently recognize a broad spectrum of features that are crucial for analyzing new, target data. The typical methodology for transfer learning in CNNs starts with using these pre-trained models to effectively identify features in the new data. The process then involves 'fine-tuning', which means making slight adjustments to the model so it better fits the specific needs of the new dataset. Deep transfer learning has proven to be extremely effective, particularly in fields where data is scarce or limited. Its successful applications range from classifying medical images [34], to inspecting the quality of manufactured parts [35], and even monitoring construction environments [36]. A pertinent example in the field of FFF is the study by Jin et al. [37]. They utilized transfer learning with the VGG-16 and Xception models, incorporating pre-trained weights from ImageNet. This application is of interest to our study as it demonstrates the potential of transfer learning in enhancing defect detection in FFF processes. In their research, VGG-16, armed with this technique, achieved an impressive 95.8% accuracy in classifying printing anomalies.

## 3. Methodology

This section describes the details of the proposed framework to detect defects in the fused filament fabrication. The overview of the framework is shown in Fig. 1. The proposed framework consists of four basic steps 1) Data collection and annotation, 2) Image pre-processing and augmentation using traditional and automated approach, 3) transfer-learning using pre-trained deep learning models, and 4) Defect detection result and visualization. The details of each of the steps are described in the following sub-sections.

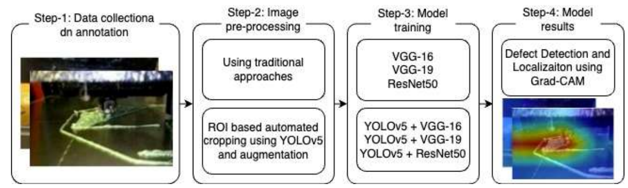



Fig. 1. The overview of the proposed framework

### 3.1. Data collection and automation

The study is performed on "kaggle-early-detection-of-3dprinting-issues" dataset, which is publicly accessible via GitHub (https://github.com). This dataset is renowned for its extensive use in various digital imaging disciplines, encompassing areas such as image processing, binary classification, industrial quality assurance, and the enhancement of ML methodologies within practical manufacturing contexts. Comprising images sourced from seven distinct 3D printers, each contributing a range of 6 to 20 prints, the dataset offers a comprehensive view of the FFF process. The captured images, taken in quick succession at intervals of approximately 0.5 seconds from cameras mounted on the printer nozzles, provide a dynamic perspective of the printing activity. The primary objective of utilizing this dataset was to classify the images into two distinct categories: those indicating 'under-extrusion' and those representing 'not under-extrusion'. This classification was achieved using binary labeling, where a label of '1' denotes under-extrusion, and a label of '0' signifies a normal, defect-free print. In our study, Fig. 2 is employed to visually demonstrate the dichotomy of our collected dataset before applying any image augmentation process. Fig. 2 (a) depicts scenarios of 'under-extrusion', where the imperfections are evident in the form of disrupted print parts and lines. Conversely, Fig. 2 (b) presents instances of optimal printing, characterized by flawless print lines and nozzle areas, indicative of 'not under-extrusion'. The images selected for this were chosen at random to demonstrate the data contained within the dataset. It is noteworthy that the dataset presented a relatively balanced distribution of images across the two categories, thus obviating the necessity for extensive data rebalancing efforts in our analysis.

### 3.2. Data pre-processing and augmentation using traditional approaches

Data pre-processing and augmentation are foundational steps in ML that enhance the diversity and quality of training data, leading to more robust models. Pre-processing standardizes input data, ensuring that the model learns from cleanly formatted and normalized datasets. Augmentation artificially expands the dataset by generating varied and plausible training samples, thereby enriching the model's experience and preventing it from learning noise and irrelevant details. The methodologies and specifics of data augmentation are comprehensively discussed by Shorten and Khoshgoftaar [38], and within the Keras library documentation [39]. In our approach, data augmentation served as a critical control measure to preclude the model's premature overfitting to the training data, which could detrimentally affect its performance on novel data. The data augmentation process involved implementing various modifications in real-time to the training dataset. These modifications include rescaling factor of 1/255 for image data normalization, horizontal flipping of images, zoom range for the image area set to 0.15, shifts in width and height by 0.20, shear rate of 0.15, Utilizing "nearest" as the fill mode, and random image rotation within a range of ± 10 degrees.


Fig. 2. Demonstration of the dataset dichotomy (a) depicts scenarios of "under-extrusion," with disrupted print parts and lines and (b) presents instances of good printing, characterized by flawless print and nozzles areas.

### 3.3. Data pre-processing using automatic cropping of region of interests (ROIs)

We aimed to enhance defect detection in FFF processes by focusing on the printer head, utilizing the advanced detection and localization capabilities of the YOLOv5 model [40]. Targeting the printer head was predicated on the rationale that concentrating the analysis on this crucial area would substantially reduce background noise and extraneous visual data, optimizing the dataset for more effective anomaly detection. This strategy was designed to create a more streamlined and focused dataset and set the stage for a comprehensive exploration of the YOLOv5 methodology, showing a deliberate move towards precise analysis in enhancing quality control.

YOLOv5 represents a significant advancement in object detection technology, optimized for exceptional speed and accuracy essential for real-time applications [41]. For each detection task, input images are standardized to a 640 x 640 dimension and converted to RGB values. The architectural backbone of YOLOv5 is composed of Cross Stage Partial Networks (CSPNet), featuring five convolutional layers, one Spatial Pyramid Pooling layer, and four CSP bottleneck blocks that consist of multiple convolution and pooling layers as shown in Fig. 3. This backbone is integral for extracting feature maps. The head part of the architecture includes a Path Aggregation Network, which processes deep layer features and relays them to the 2D convolution layer for the actual detection process. Notably, YOLOv5s, the smaller variant in the YOLO series, requires the training of over 7 million parameters (7,022,326 to be exact). This pre-trained model has been trained on the COCO dataset, a large-scale object detection, segmentation, and captioning dataset, making it robust and versatile for a wide range of detection scenarios [42]. The integration of these components within YOLOv5's framework allows for the effective pinpointing of objects, proving its utility and adaptability in high stakes environments where accuracy and speed are paramount.

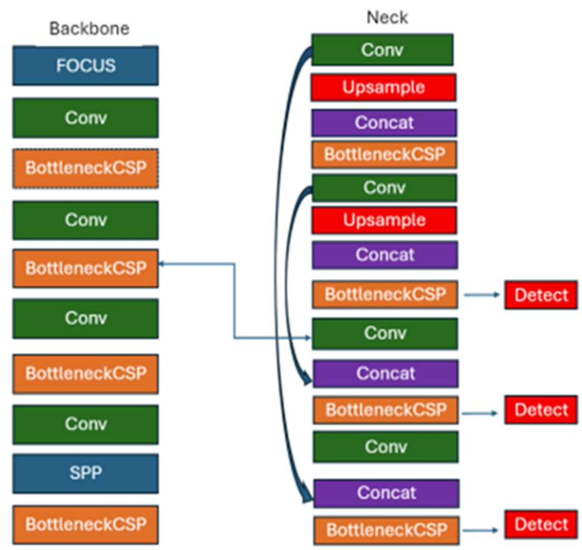



Fig. 3. Architecture of YOLOv5 network

.

### 3.4. Data pre-processing using automatic cropping of region of interests (ROIs)

VGGNet, developed by the Visual Geometry Group at the University of Oxford, has garnered significant attention in the field of deep learning, particularly following its achievement in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2014 [39]. Known for its architectural simplicity and robust performance, VGGNet has been widely adopted for diverse image processing applications. The core design principle of VGGNet revolves around assessing the impact of network depth on its performance. This is exemplified by its use of multiple 3x3 convolution filters with stride settings, coupled with image resizing through pooling layers. Theoretically, stacking two 3x3 convolution filters can effectively replicate the receptive field of a single 5x5 filter. The architectural configuration offers the distinct advantage of accommodating a greater number of non-linear rectified linear units (ReLUs) as activation functions, which serves to augment performance. This enhancement is clearly delineated in Fig. 4 (a). Moreover, this approach reduces the total number of parameters, cutting down computational demands. In this study, we explored both VGG-16 and VGG-19 models, leveraging their pre-trained weights for our specific application in anomaly detection within FFF processes. Both models, differing primarily in their depth, were assessed for their efficacy in defect detection. VGG-16, with its 16-layer architecture, and VGG-19, with 19 layers, were chosen for their balance of depth, computational efficiency, and proven track record in image recognition tasks [45].

(a)

(b)

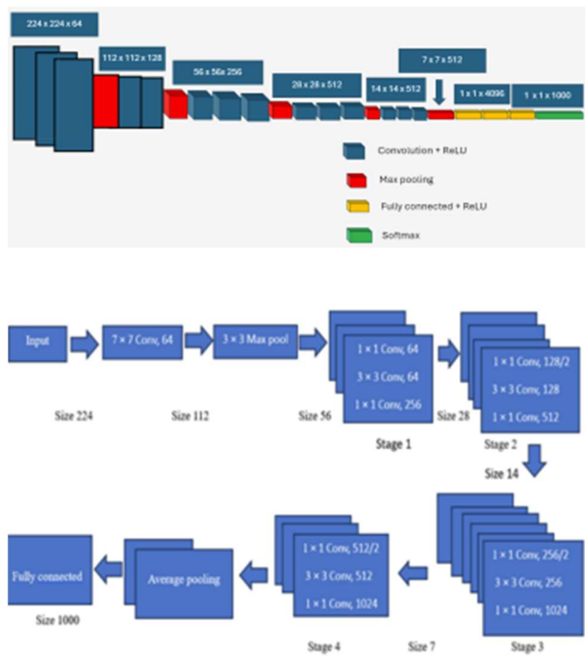



Fig. 4. Model architecture used for the transfer learning (a) VGG-16 and (b) ResNet50

ResNet, developed by Microsoft researchers, achieved significant recognition as the winner in several categories, including image classification, localization, and detection at the ILSVRC 2015 [22]. ResNet-50 V2, a specific variant, features 50 convolutional layers tailored for efficient feature extraction. It processes 224 × 224 px RGB images using multiple 3 × 3 filter convolutional layers, structured in four stages of residual blocks. These blocks integrate skip connections, effectively addressing the vanishing gradient problem common in deep networks. Distinct from VGG-16 and VGG-19, ResNet-50 does not use Max-Pooling layers within its residual blocks. Instead, it employs convolutional layers with increased stride for down sampling, culminating in a global average pooling layer followed by a single fully connected layer, leading to a SoftMax layer for classification [35]. This design as shown in Fig. 4 (b) facilitates the effective training of deep networks, making it particularly adept at complex tasks such as anomaly detection in Fused Filament Fabrication processes. Table 1 presents a detailed matrix outlining the deep transfer learning process used for fine-tuning each pre-trained model.

Table 1. Architecture of the pre-trained models: VGG-16, VGG-19, and ResNet-50.

| VGG-16 | VGG-19 | ResNet-50 |
|---|---|---|
| $Conv(3,64,1) \times 2$ | $Conv(3,64,1) \times 2$ | $Conv(3,64,1) + Maxpool(3,64,2)$ |
| $Conv(3,1281) \times 2$ | $Conv(3,128,1) \times 2$ | $\begin{bmatrix} Conv(1,64,1) \\ Conv(3,64,1) \\ Conv(1,256,1) \end{bmatrix} \times 3$ |
| $Conv(3,256,1) \times 3$ | $Conv(3,256,1) \times 4$ | $\begin{bmatrix} Conv(1,128,1) \\ Conv(3,128,1) \\ Conv(1,512,1) \end{bmatrix} \times 4$ |
| $Conv(3,512,1) \times 3$ | $Conv(3,512,1) \times 4$ | $\begin{bmatrix} Conv(1,256,1) \\ Conv(3,256,1) \\ Conv(1,1024,1) \end{bmatrix} \times 6$ |
| $Conv(3,512,1) \times 3$ | $Conv(3,512,1) \times 4$ | $\begin{bmatrix} Conv(1,64,1) \\ Conv(3,64,1) \\ Conv(1,256,1) \end{bmatrix} \times 3 + AvgPool(7,2048,1$ |

### 3.5. Data pre-processing using automatic cropping of region of interests (ROIs)

Grad-CAM, developed by Selvaraju et al. in 2017, a powerful method was used for generating visual explanations of decision-making processes within CNNs. This technique leverages the gradients of any target concept (like a classification decision) flowing into the final convolutional layer of a CNN to produce a heatmap or localization map [47]. This map vividly highlights the areas of the input image that are pivotal for the network's decision, providing insights into which features are important. Specifically, for a given input image, Grad-CAM first computes the output score for a target class and captures the feature maps from the final convolutional layer during the forward pass. It then evaluates the gradient of the class score with respect to these feature maps through backpropagation [48]. By pooling these gradients across the spatial dimensions of the feature map, Grad-CAM effectively weights the contributions of each channel to the target class, creating a coarse localization map that illuminates the critical regions influencing the CNN's prediction.

Table 2. Hyperparameters of the CNN models

| Parameter | Value |
|---|---|
| Cost function | Binary cross-entropy |
| Learning rate ($L_r$) | $1 \times 10^{-4}$ |
| Optimizer | Adam ($\beta = 0.9$) |
| Epochs | 150 |
| Batch size | 32 |
| $L_r$ Decay | 5 |
| Early stopping | 20 |

## 4. Experimental setup and results

### 4.1. Experimental design

We conducted all experiments using the Python API on a desktop with an Intel (R) Alienware Graphical Processing Unit (GPU) machine at 2.40GHZ and 128GB RAM. The result was analyzed thoroughly using six experimental setups; among them, three are based on the standard deep learning architecture (VGG-16, VGG-19, and ResNet-50), and the other three are the extensions of these three architectures paired with the YOLOV5 architecture (VGG-16 + YOLOV5, VGG-19 + YOLOV5, and ResNet-50 + YOLOV5). As mentioned earlier, the of pairing YOLOV5 is to identify and crop the surrounding area of the printer head (ROIs). In addition, to overcome the overfitting issue, the model parameters were set as shown in Table. 2. One of the primary objectives of performing these experiments is to observe the effectiveness of the setups on defect detection in FFF. Before running the experiment, we split the dataset into 80% for training and 20% for testing. Each of the experiments were run for 150 epochs. Following the training, the data was analyzed to measure the performance based in the evaluation metrics as described at Table. 2.

### 4.2. Evaluation metrics

The performance of each model employing transfer learning was assessed through a confusion matrix, which encompassed the F1-score based on precision, recall, and accuracy. In the context of evaluating a model's performance, various metrics provide insight into different aspects of accuracy and relevance. Precision serves as an indicator of the relevance of the results produced by the model, recall quantifies the proportion of actual relevant outcomes correctly identified by the model, and accuracy measures how closely the model's predictions align with the actual values. The F1-score is particularly useful for providing a single metric that balances both precision and recall, especially important in scenarios where both false positives and false negatives carry significant consequences. Precision, Recall, Accuracy, and F1 score are defined by Equations (1), (2), (3), and (4).

$$Precision = \frac{True\ postive}{True\ positive + False\ positive} \quad (1)$$

$$Accuracy = \frac{True\ postive}{True\ positive + False\ negative} \quad (2)$$

$$Recall = \frac{True\ postive}{True\ positive + False\ positive} \quad (3)$$

$$F1\ score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

Additionally, to assess the practicality of deploying these transfer learning models, we evaluated the computing requirements. This evaluation included the number of parameters, the floating-point operations per second (FLOPS), and the actual computing time for each pre-trained model. Such metrics are essential for understanding the feasibility and efficiency of implementing the proposed deep transfer learning method in real-world manufacturing settings.

(a) (b) (c)

Fig. 5. Demonstration of printer head detection and automated ROI cropping (a) represents the initial dataset image, (b) identification of printer head using YOLOv5 algorithm, and (c) automated cropping of ROI around printer head.

### 4.3. Performance analysis of the trained model

Incorporating YOLOv5 for printer head detection significantly enhances our methodology by reprocessing our initial dataset for our specific FFF process. Fig. 5 illustrates these processes and outcomes: beginning with the original dataset image (a), our trained YOLOv5 model precisely identifies and delineates the printer head with a bounding box (b), thereby focusing the analysis on crucial areas. This critical step eliminates extraneous background elements, enabling concentrated analysis on areas prone to potential anomalies. The procedure culminates with an automated crop around the printer head (c), distinctly isolating the area of interest for further examination. This refined image preprocessing technique significantly bolsters the dataset's readiness for retraining our CNN models—VGG-16, VGG-19, and ResNet-50. It facilitates a comparative evaluation of the models' efficacy, both before and after applying our printer head detection strategy.

(a) (b)

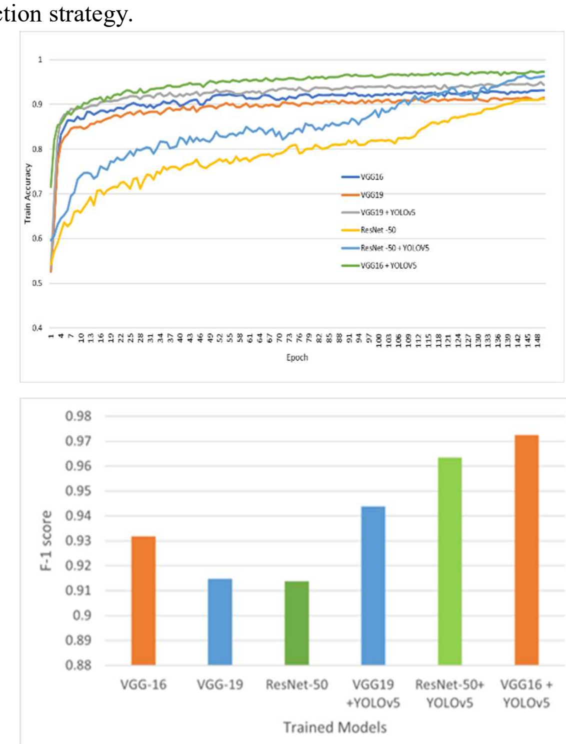



Fig.6. Comparison of Trained CNN models, (a) presents accuracy vs epoch and (b) illustrates the F-1 score corresponding to each trained model.

In this study, we set a definitive performance baseline by training the pre-trained VGG-16, VGG-19, and ResNet-50 models on the original dataset, achieving initial accuracy rates of 93.1%, 91.4%, and 91.3%, respectively. This established a foundation for comparison to assess the impact of our printer head detection approach using YOLOv5. Each model exhibited its distinct learning curve over the course of 150 epochs. It was observed that integrating YOLOv5 with VGG-16, VGG-19, and ResNet-50 significantly enhances image recognition accuracy as shown in Fig. 6 (a). Specifically, VGG-16 combined with YOLOv5 achieved a remarkable accuracy of nearly 97.3% after 150 epochs, surpassing the standalone VGG-16's stabilization at 93.1%. Similarly, VGG-19 + YOLOv5 outperformed the standalone VGG-19, reaching an accuracy of 94.4% versus 91.4%. ResNet-50's improvement to 91.3% pales in comparison to the ResNet-50 + YOLOv5 combination, which soared to 96.3% accuracy. These enhancements underscore the importance of cropping the images around the ROIs (printer head) using the YOLOv5 algorithm. Moreover, the increased starting accuracy rates after installing YOLOv5 are particularly noteworthy, as it underscores the substantial improvements facilitated by the integration of YOLOv5. These enhancements are attributed to the "printer head approach's" efficiency, which not only elevates model accuracy but also instills a level of robustness, enabling the models to counteract the effects of learning variability and achieve a swift and consistent stabilization of performance across epochs. Furthermore, the observed downward trends in accuracy across epochs for highlight the challenges of overfitting, underfitting, or learning rate adjustments. Yet, the resilience of the hybrid models against these fluctuations' underscores YOLOv5's role not only in enhancing accuracy but also in stabilizing the learning process more swiftly and effectively.

Additionally, we delve into the analysis of the F1 score, a more nuanced performance metric that balances precision and recall. Fig. 6 (b) illustrates the F1 scores corresponding to each pre-trained model across the fine-tuning sections, both with and without the application of YOLOv5 for printer head detection. It can be observed that the integration of YOLOv5 consistently enhances the F1 score across all models and fine-tuning steps. VGG-16, when combined with YOLOv5, exhibits a significant improvement in the F1 score, maintaining a high level of performance throughout the fine-tuning. The VGG-19 model also shows a marked enhancement in conjunction with YOLOv5, particularly in the later sections. ResNet-50, while benefiting from the addition of YOLOv5, displays a more varied pattern of F1 scores across the sections, suggesting a sensitivity to the fine-tuning process. Moreover, The F1-score comparison also confirms that VGG-16 showed the best performance among the evaluated pre-trained models. The performance evaluation results are summarized in Table 3.

Table 3. Performance metrics of pre-trained models with and without image pre-processing using YOLOv5

| Trained Model | Accuracy (%) | Precision | Recall | F1-Score |
|---|---|---|---|---|
| VGG-16 | 93.1 | 0.94 | 0.92 | 0.93 |
| VGG-16 + YOLOv5 | 97.9 | 0.98 | 0.97 | 0.972 |
| VGG-19 | 91.4 | 0.91 | 0.91 | 0.915 |
| VGG-19 + YOLOv5 | 92.2 | 0.92 | 0.91 | 0.925 |
| ResNet-50v2 | 91.3 | 0.90 | 0.91 | 0.91 |
| ResNet-50v2 + YOLOv5 | 95.7 | 0.94 | .95 | 0.96 |

### *4.4. Computational demand and efficacy analysis*

Computational efficiency of anomaly detection varies significantly across the models tested, as demonstrated by the processing times detailed in Table 4. VGG-16, known for its simpler yet effective architecture, demonstrated a processing time of 0.052 seconds. The introduction of YOLOv5 preprocessing reduced this time to 0.049 seconds, highlighting the impact of targeted image cropping on computational demand. This reduction suggests that focusing on relevant regions within the image can streamline the detection process without compromising accuracy. VGG-19, with a slightly more complex architecture than VGG-16, initially required 0.061 seconds for processing. Interestingly, the application of YOLOv5 did not alter the processing time for VGG-19, remaining at 0.061 seconds. This stability in processing time, despite the additional preprocessing step, may indicate that the complexity of VGG-19's deeper network offsets the potential efficiency gains from image cropping, maintaining a consistent computational load. ResNet-50, leveraging residual connections for enhanced performance with less computational overhead, posted the fastest processing time of 0.045 seconds among the original models. With the YOLOv5 enhancement, this time marginally improved to 0.044 seconds.

Table 4. Scanning time per image to detect defect.

| Model | Detection time (Second per image) |
|---|---|
| VGG-16 | 0.052 |
| VGG-19 | 0.061 |
| ResNet-50 | 0.045 |
| VGG-16+YOLOv5 | 0.049 |
| VGG-19+YOLOv5 | 0.056 |
| ResNet-50+YOLOv5 | 0.044 |

## 5. Discussion

The results highlight the effectiveness of the deep transfer learning strategy employed across various pre-trained models. The analysis confirms that even with a limited dataset from FFF processes, the models are trainable and can deliver systematic and reliable results. A key insight from this study is the balance between model performance and computational load, with considerations given to memory usage and processing time. Moreover, the results, highlight the effectiveness of the deep transfer learning strategy employed across various pre-trained models. The analysis confirms that even with a limited dataset from FFF processes, the models are trainable and can deliver systematic and reliable results. A key insight from this study is the balance between model performance and computational load, with considerations given to memory usage and processing time.

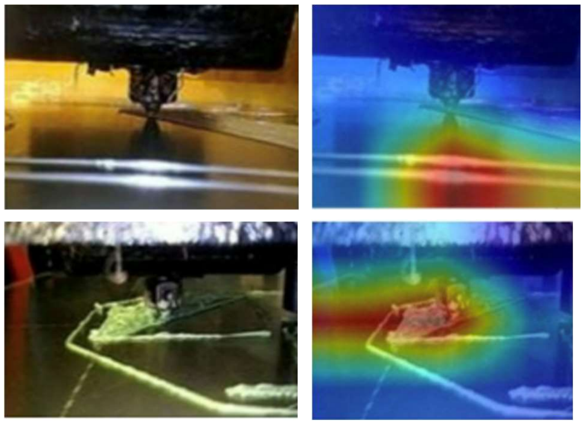
(a)
(b)


Fig.7. Visualization of anomaly detection using VGG16+YOLOv5 and Grad-CAM (a) presents a test image with an identified anomaly alongside its corresponding Grad-CAM heatmap and (b) illustrates another instance of anomaly detection, showing a test image with a different anomaly and its Grad-CAM heatmap

The application of the printer head detection method yielded improvements in all models, but the VGG-16 model, when paired with YOLOv5, achieved the highest accuracy. In contrast, the ResNet-50 model stood out for its computational thriftiness, boasting the shortest classification times and lowest memory demands, though this came with a trade-off in terms of achieving the lowest accuracy in the group. For real-world applications where high accuracy and swift prediction are critical—surpassing 90% accuracy and under a second in prediction time—VGG-16 integrated with YOLOv5 is the preferred choice. The study also demonstrates the adaptability of the approach to other, potentially newer, CNN architectures that may enhance performance, although this could require more advanced expertise to implement effectively.

Further, practical tests including visual verifications were conducted, reinforcing the chosen method's applicability. The use of Grad-CAM for visual analysis confirmed the VGG-16 through nozzle head detection model's ability to localize anomaly patterns accurately, as shown in Fig. 7. This visual validation lends credence to the model's utility in a manufacturing setting, where such precision is paramount.

## 6. Conclusion

In the manufacturing sector, the challenge of limited process datasets necessitates the use of pre-trained deep learning models to monitor defects, even with small datasets. This study introduced a deep transfer learning approach tailored for a small dataset of FFF printing images, focusing on extrusion errors, utilizing well-known CNN models like VGG-16, VGG-19, and ResNet-50. Through image augmentation, we expanded the variety within the original dataset, enabling more effective deep transfer learning. The experimental results underscore the effectiveness of our method, particularly when employing the VGG-16 model combined with YOLOv5, demonstrating consistent and reliable performance in anomaly detection. The integration of VGG-16 and YOLOv5 not only achieved high performance but also met computing efficiency criteria. Validation through Grad-CAM and practical demonstrations further confirmed the model's capability for real-world monitoring tasks, offering valuable insights for manufacturing professionals on leveraging existing pre-trained CNN models for quality monitoring of FFF printing processes and beyond.

This current study was based on a dataset that lacked layer specific anomaly data for FFF printing. Future research will delve into layer-based quality monitoring and test other datasets with our proposed model. Additionally, this research lays the groundwork for a diagnostic tool in detecting printing anomalies. Future explorations will include the development of hardware for visual monitoring in conjunction with FFF printers. This will not only diagnose but potentially correct errors, providing a more comprehensive solution to enhance the practicality and effectiveness of our proposed method.

## Acknowledgements

We acknowledge funding by the MSEIP and FIPSE Modelling and Simulation (MSP) programs of the U.S. Department of Education under Award No. P120A220044 and P120A180101, respectively.

## References

[1] Bavoria, Rehan Kumar, 2023. A Study on Enablers and Barriers of Additive Manufacturing Adoption in Industry 4.0: An AHP Approach.

[2] Ninduwezuor-Ehiobu, Nwakamma, et al. ”Exploring Innovative Material Integration In Modern Manufacturing For Advancing US Competitiveness In Sustainable Global Economy.” Engineering Science & Technology Journal 4.3 (2023): 140-168

[3] Erokhin, Kirill, Naumov, Sergei, and Ananikov, Valentine, 2023. ”Defects in 3D Printing and Strategies to Enhance Quality of FFF Additive Manufacturing. A Review.”

[4] Sola, Antonella, and Adrian Trinchi. Fused Deposition Modeling of Composite Materials. Elsevier, 2022.

[5] Bruere, V. M., et al. ”Under-extrusion challenges for elastic filaments: The ` influence of moisture on additive manufacturing.” Progress in Additive Manufacturing.

[6] Fu Y, Downey A, Yuan L, Pratt A, Balogun Y. ”In situ monitoring for fused filament fabrication process: a review.” Additive Manufacturing 2021;38:101749.

[7] Buj-Corral, Irene, Tejo-Otero, Aitor, and Fenollosa-Artes, Felip. ”Use of ´ FDM Technology in Healthcare Applications: Recent Advances.” Fused Deposition Modeling Based 3D Printing 2021; pages 277–297. Springer.

[8] Mohanavel, V., Ali, KS Ashraff, Ranganathan, K., Jeffrey, J. Allen, Ravikumar, MM., and Rajkumar, S. ”The roles and applications of additive manufacturing in the aerospace and automobile sector.” Materials Today: Proceedings 2021;47: pages 405–409. Elsevier.

[9] Ramirez, Alberto Sanchez, Marcos, Manuel Enrique Islan, Haro, Fernando ´ Blaya, D'Amato, Roberto, Sant, Rodolfo, and Porras, Jose. ”Application of ´ FDM technology to reduce aerodynamic drag.” Rapid Prototyping Journal 2019;25(4): pages 781–791. Emerald Publishing Limited.

[10] Pienaar, T. ”Viability of 3D fused filament fabrication for aerospace applications.” PhD dissertation, North-West University (South Africa), 2021.

[11] Fico, D., Rizzo, D., Casciaro, R., Esposito Corcione, C., 2022. ”A Review of Polymer-Based Materials for Fused Filament Fabrication (FFF): Focus

on Sustainability and Recycled Materials." Polymers (Basel) 14(3):465. doi: 10.3390/polym14030465. PMID: 35160455; PMCID: PMC8839523.

[12] Fogliata, A., Garcia, R., Knoos, T., Nicolini, G., Clivio, A., Vanetti, E., Khamphan, C., Cozzi, L., 2012. "Definition of parameters for quality assurance of flattening filter free (FFF) photon beams in radiation therapy." Med Phys 39(10):6455-64. doi: 10.1118/1.4754799. PMID: 23039680.

[13]Suiker, Akke SJ, 2022. "Effect of accelerated curing and layer deformations on structural failure during extrusion-based 3D printing." Cement and Concrete Research 151: 106586.

[14] Kumar, Rahul, Kanwarpreet Singh, and Sanjiv Kumar Jain, 2022. "Agility enhancement through agile manufacturing implementation: a case study." The TQM Journal 34.6: 1527-1546.

[15] Ding, Jin, et al. "The economics of additive manufacturing: Towards a general cost model including process failure." International Journal of Production Economics 237 (2021): 108087.

[16] Selot, A., Dwivedi, R.K., 2023. "Machine learning and sensor-based approach for defect detection in MEX additive manufacturing process- A Review." J Braz. Soc. Mech. Sci. Eng. 45, 535.

[17] Hossain, Rifat-E-Nur, Lewis, Jerald, Moore, Arden L., 2021. "In situ infrared temperature sensing for real-time defect detection in additive manufacturing." Additive Manufacturing, 47, 102328. ISSN 2214-8604.

[18] Goh GD, Hamzah NMB, Yeong WY, 2023. "Anomaly Detection in Fused Filament Fabrication Using Machine Learning." 3D Print Addit Manuf. 10(3):428-437.

[19] Goh, G.D., Sing, S.L., Yeong, W.Y., 2021. "A review on machine learning in 3D printing: applications, potential, and challenges." Artif Intell Rev 54, 63–94.

[20] Liu, Chenang, Law, Andrew Chung Chee, Roberson, David, and Kong, Zhenyu (James), 2019. "Image analysis-based closed loop quality control for additive manufacturing with fused filament fabrication." Journal of Manufacturing Systems, 51, pp. 75-86. ISSN 0278-6125.

[21] Xu, Ke, Lyu, Jiaqi, and Manoochehri, Souran, 2022. "In situ process monitoring using acoustic emission and laser scanning techniques based on machine learning models." Journal of Manufacturing Processes, 84, pp. 357- 374. ISSN 1526-6125.

[22] He, K., Zhang, X., Ren, S., and Sun, J., 2016. "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, pp. 770-778.

[23] Banadaki, Y., Razaviarab, N., Fekrmandi, H., Sharifi, S., 2020. "Toward enabling a reliable quality monitoring system for additive manufacturing process using deep convolutional neural networks." arXiv preprint arXiv:200308749.

[24] Paraskevoudis, K., Karayannis, P., Koumoulos, EP., 2020. "Real-time 3D printing remote defect detection (stringing) with computer vision and artificial intelligence." Processes, 8(11):1464.

[25] Ramezankhani, M., Crawford, B., Narayan, A., Voggenreiter, H., Seethaler, R., Milani, AS., 2021. "Making costly manufacturing smart with transfer learning under limited data: a case study on composites autoclave processing." Journal of Manufacturing Systems, 59:345–54.

[26] Noor, A., Zhao, Y., Koubaa, A., Wu, L., Khan, R., Abdalla, FYO., 2020. "Automated sheep facial expression classification using deep transfer learning." Computers and Electronics in Agriculture, 175:105528.

[27] Jin, Z., Zhang, Z., Gu, GX., 2020. "Automated real-time detection and prediction of interlayer imperfections in additive manufacturing processes using artificial intelligence." Advanced Intelligent Systems, 2(1):1900130.

[28] Kim, H., Lee, H., Kim, J-S., Ahn, S-H., 2020. "Image-based failure detection for material extrusion process using a convolutional neural network." Journal of Advanced Manufacturing Technology, 111(5):1291–1302.

[29] Jung, W-K., Kim, D-R., Lee, H., Lee, T-H., Yang, I., Youn, BD., et al., 2021. "Appropriate smart factory for SMEs: concept, application and perspective." International Journal of Precision Engineering and Manufacturing, 22(1):201–15.

[30] Guerreiro, BV., Lins, RG., Sun, J., Schmitt, R., 2018. "Definition of smart retrofitting: first steps for a company to deploy aspects of industry 4.0." Advances in Manufacturing. Cham: Springer International Publishing.

[31] Razvi, S.S., Feng, S., Narayanan, A., Lee, Y-TT., Witherell, P. (Eds.), 2019. "A review of machine learning applications in additive manufacturing." Proceedings of the ASME 2019 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. V001T02A040.

[32]Pan, SJ., Yang, Q., 2010. "A survey on transfer learning." IEEE Transactions on Knowledge and Data Engineering, 22(10):1345–1359.

[33]Ren, Z., Fang, F., Yan, N., Wu, Y., 2021. "State of the art in defect detection based on machine vision." International Journal of Precision Engineering and Manufacturing - Green Technology.

[34]Rahman, Md Fashiar, et al., 2022. "Improving lung region segmentation accuracy in chest X-ray images using a two-model deep learning ensemble approach." Journal of Visual Communication and Image Representation, 85:103521.

[35] Ferguson, MK., Ronay, A., Lee, Y-TT., Law, KH., 2018. "Detection and segmentation of manufacturing defects with convolutional neural networks and transfer learning." Smart and Sustainable Manufacturing Systems, 2.

[36]Kolar, Z., Chen, H., Luo, X., 2018. "Transfer learning and deep convolutional neural networks for safety guardrail detection in 2D images." Automation in Construction, 89:58–70.

[37]Jin, Z., Zhang, Z., Ott, J., Gu, GX., 2021. "Precise localization and semantic segmentation detection of printing conditions in fused filament fabrication technologies using machine learning." Additive Manufacturing, 37:101696.

[38]Johnson, J.M., Khoshgoftaar, T.M., 2019. "Survey on deep learning with class imbalance." J. Big Data, 6:27.

[39]Mustaquim, S. M. "Utilizing remote sensing data and ArcGIS for advanced computational analysis in land surface temperature modeling and land use property characterization.

[40] Jiang, Peiyuan, Ergu, Daji, Liu, Fangyao, Cai, Ying, and Ma, Bo. "A Review of Yolo algorithm developments." Procedia Computer Science 2022;199: pages 1066–1073. Elsevier.

[41] Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A., 2016. "You Only Look Once: Unified, Real-Time Object Detection." In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788.

[42] Thuan, Do, 2021. "Evolution of Yolo algorithm and Yolov5: The State-of the-Art object detention algorithm."

[43] Gulli, Antonio and Pal, Sujit, 2017. "Deep Learning with Keras." Packt Publishing Ltd.

[44] Manaswi, Navin Kumar, 2018. "Understanding and working with Keras." In Deep Learning with Applications Using Python: Chatbots and Face, Object, and Speech Recognition with TensorFlow and Keras, pp. 31–43. Springer.

[45] Ding, Xiaohan, Zhang, Xiangyu, Ma, Ningning, Han, Jungong, Ding, Guiguang, Sun, Jian, 2021. "Repvgg: Making vgg-style convnets great again." In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 13733–13742.

[46] Shin, H., et al., 2016. "Deep Convolutional Neural Networks for Computer Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning." IEEE Transactions on Medical Imaging, vol. 35, no. 5, pp. 1285-1298, May 2016.

[47]Selvaraju, Ramprasaath R., Cogswell, Michael, Das, Abhishek, Vedantam, Ramakrishna, Parikh, Devi, and Batra, Dhruv, 2017. "Grad-cam: Visual explanations from deep networks via gradient-based localization." Proceedings of the IEEE International Conference on Computer Vision, pages 618– 626.

[48]Selvaraju, Ramprasaath R., Das, Abhishek, Vedantam, Ramakrishna, Cogswell, Michael, Parikh, Devi, and Batra, Dhruv. "Grad-CAM: Why did you say that?" arXiv preprint arXiv:1611.07450, 2016