Cognitive Computation (2025) 17:36
https://doi.org/10.1007/s12559-024-10397-8

RESEARCH

A Novel Depth-Connected Region-Based Convolutional Neural Network for Small Defect Detection in Additive Manufacturing

Yiming Wang[1] · Zidong Wang[1] · Weibo Liu[1] · Nianyin Zeng[2] · Stanislao Lauria[1] · Camilo Prieto[3] · Fredrik Sikström[4] · Hui Yu[5] · Xiaohui Liu[1]

Received: 4 September 2024 / Accepted: 8 December 2024

Abstract
Defect detection on the computed tomography (CT) images plays an important role in the development of metallic additive manufacturing (AM). Although some deep learning techniques have been adopted in the CT image-based defect detection problem, it is still a challenging task to accurately detect small-size defects in the presence of undesirable noises. In this paper, a novel defect detection method, namely, the depth-connected region-based convolutional neural network (DC-RCNN), is proposed to detect small defects and reduce the influence of noises. In particular, a saliency-guided region proposal method is first developed to generate small-size region proposals with the aim to accommodate the small defects. Then, the main architecture of DC-RCNN is proposed to extract and connect the consistent features across multiple frames, thereby reducing the influence of randomly distributed noises. Moreover, the transfer learning technique is utilized to improve the generalization ability of the proposed DC-RCNN. In order to verify the effectiveness and superiority, the proposed method is applied to the real-world AM data for defect detection. The experimental validations show that the proposed DC-RCNN is able to detect the small-size defects under noises and outperforms the original RCNN method in terms of detection accuracy and running time.

Keywords Defect detection · Additive manufacturing · Region based convolutional neural network · Region proposals · Depth connectivity

Introduction

Over the past several decades, the additive manufacturing (AM) technology, also known as three-dimensional (3D) printing, has attracted ever-increasing research attention owing primarily to its successful applications in various communities such as biomaterials, aerospace, and transportation [9]. As is well known, the AM technology allows for the fabrication of 3D objects with complex structures/geometries by layering the materials one by one, thereby providing a rapid and flexible way to produce the custom-tailored components. Compared with the traditional subtractive manufacturing, the AM exhibits distinct advantages which include, but are not limited to, rapid prototyping, eco-friendliness, high adaptiveness, and sustainability. On the other hand, the AM is actually a sophisticated multistage process, which may create certain unqualified products with undesirable defects due to the improper operations and stochastic disturbances [37]. Such defects are usually invisible and may potentially affect the mechanical properties of products. In this sense, it is practically meaningful to develop an effective method to automatically detect the invisible defects in the fabricated component and hence maintain the manufacturing quality.

The non-destructive testing (NDT) has been well recognized as an effective technique to detect the surface/interior flaws without causing secondary damage. Up to now, various NDT methods have been proposed with examples including the thermography testing, eddy current testing, 3D scanning testing, and ultrasonic testing [9, 22]. In particular, as one of the 3D scanning methods, the X-ray computed tomography (CT) technique, which generates a large number of volumetric images to visualize the internal morphology of an object, has been widely utilized to detect the defects. Accordingly, many knowledge-based methods have been developed to analyze the collected high-resolution CT scan images, such as the domain-specific rule-based methods, statistical methods, and handcrafted feature-based methods [4,

Extended author information available on the last page of the article

Published online: 24 December 2024

13, 26, 35, 61]. Among others, the support vector machine has been employed for defect identification based on the local 3D root mean squared features extracted within a small voxel size [18]. Some handcrafted feature-based methods have been developed for CT inspection, see, e.g., geometric feature analysis, statistical image feature analysis, and the filter-based methods [13, 35].

With the revolution of artificial intelligence, the deep learning methods have been successfully applied to the field of defect detection in recent years. In comparison with the traditional image processing methods, the deep learning methods, following the end-to-end learning paradigm, are able to directly acquire knowledge from the data without using any handcrafted features. As one of the popular deep learning methods, the so-called convolutional neural network (CNN) has been widely exploited for the image-based industrial product quality inspection, see, e.g., [15, 19, 25, 27, 44, 62] and the references therein. Some representative CNN-based object detection methods are the region-based convolutional neural network (RCNN) [17, 41], you only look once (YOLO) methods [39], and the RetinaNet [30].

In the context of the CNN-based object detection, a notable fact is that a large-scale image dataset is required for the training of CNN-based deep model. Accordingly, the transfer learning technique has been adopted in the field of computer vision to enable the effective learning on small-scale dataset [34]. The main idea of transfer learning technique is that the knowledge of a pre-trained deep model (on a large-scale source dataset) can be applied to another customized task by fine-tuning this model on the small-scale target dataset.

Recently, transfer learning techniques with pre-trained CNN models has received considerable attentions due to their promising performance in object detection. For example, a transfer learning technique has been employed in [42] to improve faster RCNN performance in surface defect recognition. In [51, 55], pre-trained CNN models have been utilized for powder bed defect detection in selective laser sintering systems. In [16], a comparison of several popular object detection methods has been made to evaluate the transfer ability between different sets of X-ray images, where the pre-trained faster RCNN model demonstrates superiority over YOLO and RetinaNet methods in terms of detection accuracy and transfer ability.

Generally speaking, there are mainly three challenges in dealing with the defect detection problem of AM components: (1) the size of the defects is relatively small; (2) the CT images often contain noises; and (3) the number of CT images with labelled defects is usually small. In this regard, it makes practical sense to investigate the small-size defect detection problem, i.e., the small object detection problem with target object occupying less than 20% of the image. Unfortunately, it should be mentioned that most existing CNN-based methods cannot achieve a satisfactory performance in handling such an issue [36]. On the other hand, the noise, looking similar to the defects, may appear in the CT images due to a variety of reasons (e.g., computational error and fluctuation of X-ray radiation dose [5]), which renders additional difficulties for defect detection and also explains why it is still a challenging task to accurately detect small-size defects in the presence of noise.

In order to tackle these identified challenges, a recurring research attention has been devoted to the RCNN owing to its outstanding performance in processing the small-size defects. For example, the RCNN combined with the multi-scale initialization of default candidate regions [41] or the integrated feature pyramid networks (FPN) [29] has proven to be effective in defecting the small-size objects. To leverage multi-scale feature fusion, most existing small defect detection methods reply FPN and its variants [56, 62]. Note that these approaches still employ traditional region proposal techniques, where default region proposals are uniformly distributed across the entire image, as illustrated in the left column of Fig. 1b. Such uniform distribution may result in significant computational inefficiencies. Notably, none of these variant FPN methods embeds an efficient region proposal strategy which is capable of quickly and effectively generating proposals tailored to small-size defects. In practice, the desired region proposals for AM specimens are often small and concentrated along the outlined salient regions of the component. To address this specific challenge, as depicted in the right column of Fig. 1b, a natural and effective solution is to develop a new region proposal method that focuses on generating small-size candidate regions aligned with the salient areas of interest.

For the noises in CT images, it can be observed in Fig. 1a that the defects usually lie across multiple frames, while the randomly distributed noise regions lack the cross-frame consistency. Such features make it possible to reduce the impact of noises by detecting the cross-frame associations within a set of sequential CT images. To date, a variety of RCNN-based object tracking methods have been developed by taking into account the cross-frame association. These existing methods, however, require large-scale datasets for both pre-training and fine-tuning, and hence are inapplicable to the case of AM defect detection with small-scale dataset. In this context, it is of practical significance to develop a new RCNN-based method that enables the cross-frame feature extraction and is compatible with the transfer learning technique.

Motivated by the above discussions, in this paper, we endeavor to propose a novel depth-connected RCNN (DC-RCNN) method by (1) generating small and sparsely distributed region proposals at the salient regions to adapt for

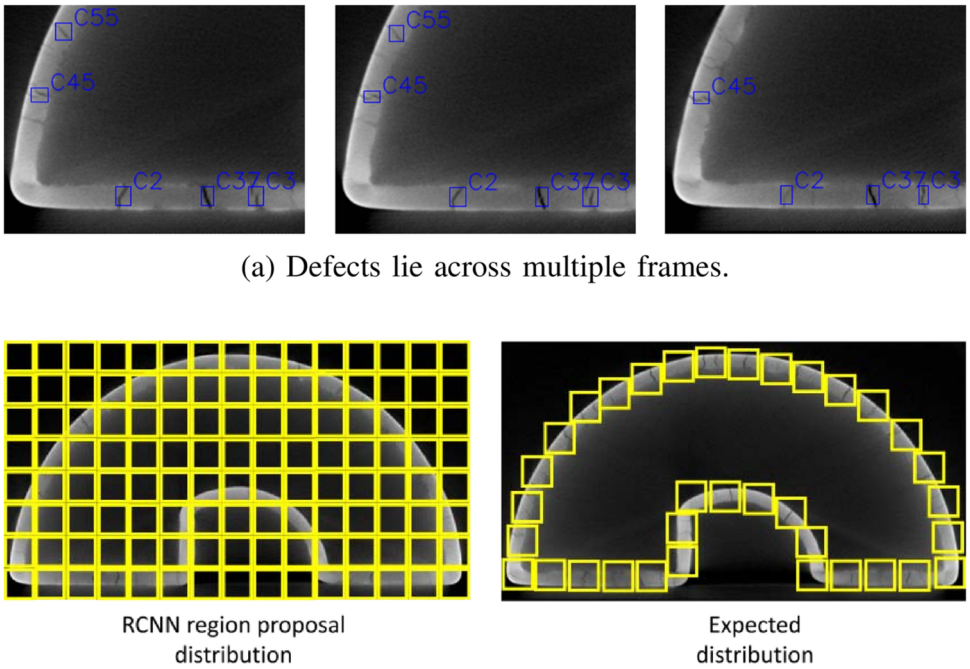


Fig. 1 Illustration for the problems in AM defect detection

(a) Defects lie across multiple frames.

(b) Region proposal distribution.

the small defects and (2) extending the frame-by-frame detection scheme to the cross-frame depth-connected one with aim to reduce the influence of noises. The main contributions of this paper can be summarized as follows.

1) Unlike existing FPN-based small defect detection methods, a newly designed saliency-guided region proposal approach is developed to effectively generate small-scale region proposals.
2) A novel DC-RCNN method is proposed for AM defect detection, featuring a unique architecture with parallel convolutional blocks for depth-connected feature extraction, which is capable of mitigating the influence of noises.
3) The established network architecture is compatible with the transfer learning technique, and a pre-trained faster RCNN model is employed to improve the detection performance and training efficiency.
4) The proposed method is successfully applied to the small defect detection for the real-world AM specimens, and the experimental results show that the proposed method outperforms the standard faster RCNN in terms of both accuracy and efficiency.

The remainder of this paper is organized as follows. The "Related Work" section introduces the background of the NDT technologies and the state-of-the-art defect detection methods. The "Methodology" section gives the technical details of the proposed DC-RCNN method. The utilized dataset is described in the "Experiment" section, where the experimental results and quantitative analysis are also presented. Finally, this paper is concluded in the "Conclusion" section.

Related Work

In order to assess the product quality, it is important to perform the post-fabrication testing and inspection on an AM product. As is well known, the NDT is capable of identifying or locating the defects in an AM component without resulting in secondary defects. In this context, various NDT methods have been developed, among which the CT-based 3D scanning has attracted particular research interest in recent years. In this section, we first introduce the general knowledge of the AM process and the NDT techniques. Then, we present the state-of-the-art CT image-based defect detection methods and discuss the advantages and disadvantages of these defect detection methods.

AM Process and NDT Techniques

During the past few decades, the AM process has been successfully applied to a wide range of industrial fields such as aerospace, medical & dental, automotive, and entertainment.

Roughly speaking, an AM process is composed of four steps, i.e., 3D model design, simulation, manufacturing, and post-processing. Specifically, the first step is the design of 3D model by using the computer-aided design (CAD) software. Then, the simulation step is performed to simulate the designed 3D model and determine some critical parameters before manufacturing. The third step is to print this 3D model through a binder jetting process that deposits the materials layer by layer to form a physical 3D object. Finally, the post-processing step is carried out to inspect the quality of the produced AM component. It should be pointed out that, due to the complicated control procedure and the unpredictable noises, the AM process may produce defective products with many small and internal defects [37]. Therefore, it is of great significance to implement the post-processing quality assessment.

In the context of quality assessment for an AM component, the so-called NDT technique has gained a persistent research interest due to its prominent capability in inspecting internal geometric structures without damaging the AM component. Up to now, a variety of NDT methods have been developed, e.g., thermography testing [38], electromagnetic testing [43], 3D CT testing [13, 26], and ultrasonic testing [22]. Among others, the 3D CT scanning method has stood out since it is capable of visualizing the fine and internal structures of the object being tested. In the CT scanning process, the radiation beam is employed to penetrate an AM component and create a group of volumetric images displaying the internal structure (possibly including the originally small and invisible defects) of the component. Based on the high-resolution CT scans, a large number of semi-automatic or fully automatic defect detection methods have been developed, see, e.g., [13, 18, 26]. Moreover, it is worth mentioning that the state-of-the-art defect detection methods have been mainly developed by resorting to various deep learning techniques (which are specifically designed for object detection and image segmentation) [12, 42].

Defect Detection

Generally speaking, the CT image-based defect detection can be regarded as a branch of the object detection. It is well known that the object detection, whose aim is to detect and localize all the objects in an image, is one of the most important tasks in the computer vision field due to the extensive applications ranging from security, face recognition, autonomous driving to robotic vision [3, 8]. Specifically, the location of an object in the image can be described by a rectangle bounding box formed by four coordinates (x, y, w, h), where (x, y) is the coordinate of top-left corner of the bounding box, w and h represent, respectively, the height and width of the box. In this sense, the localization problem becomes a regression task that learns to map the default anchor boxes to the target boxes. Correspondingly, the object detection is converted into the problems of the object categorical classification and the box regression.

So far, much effort has been made to develop the deep learning-based object detection methods, and some representatives are the RCNN family (e.g., the fast RCNN [17] and the faster RCNN [41]) and the single-stage rapid detection methods (e.g., the YOLO methods [39], the single shot multibox detector [31], the RetinaNet [30]) and the Detection Transformer (DETR) [1]. In the context of the NDT, the CNN-based methods have stirred remarkable interest and many research results have been reported in the literature by resorting to different NDT techniques, e.g., the ultrasonic imaging, the infrared thermal imaging and the CT images [38]. Among others, in electromagnetic NDT, the CNN has been utilized to identify the weld defects based on the magneto-optical images [21].

It is worth mentioning that the aforementioned methods require large-scale datasets for training. Nevertheless, it is often the case in the AM defect detection that only small-scale datasets are available, which implies that the above methods are no longer effective. To this end, the so-called transfer learning technique has been developed by transferring the knowledge obtained by pre-training a deep model (on a large-scale dataset) to the target task (with a small-size dataset). Recently, the transfer learning-based deep learning methods have been widely employed in defect detection with appealing performance [12, 16]. For example, a typical transfer learning-based defect detection method has been presented in [20], where the detection network is pre-trained on a large-scale publicly available dataset.

In practical engineering, the noises in the NDT images and the small-size defects render it really challenging to detect the possible defects in AM components. Several studies have addressed small-size defect detection [56, 62]. For instance, in [62], an improved FPN method has been proposed to achieve multi-scale feature fusion, enhancing the extracted features in potential small-size defect regions. Similarly, a new region proposal selection method has been introduced in [56], where the size of candidate boxes is optimized to improve small defect detection. However, these approaches still rely on traditional region proposal techniques. Until now, little attention has been devoted to the generation of small-size region proposals. In this paper, for the purpose of achieving the noise-free small-size defect detection on the CT images, a novel RCNN-based detection framework is developed with a domain-specific small-size region proposal method and a depth-connected neural network (whose aim is to reduce the influence of noises).

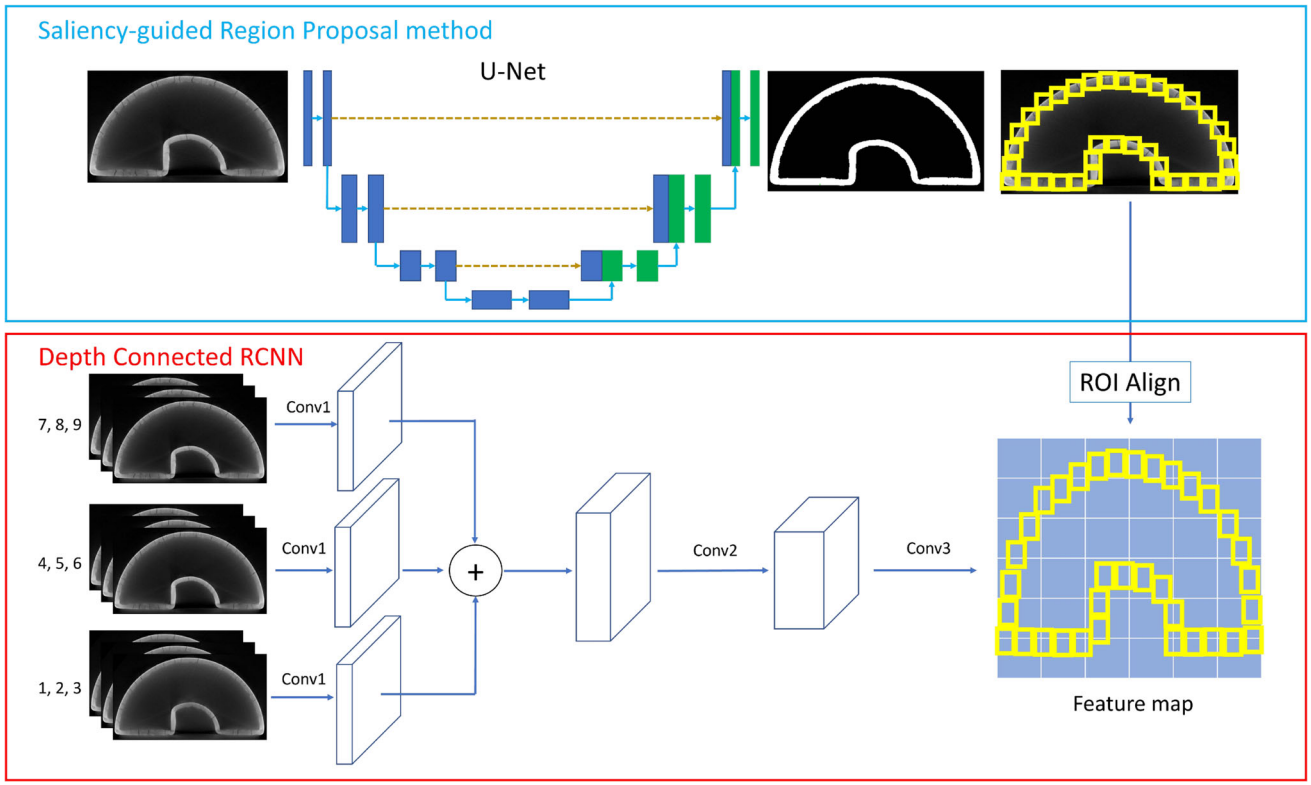


Fig. 2 Overview of the proposed DC-RCNN method

Methodology

In this section, a novel DC-RCNN method, as shown in Fig. 2, is proposed to detect the small defects on the noise-corrupted CT images. Specifically, a saliency-guided region proposal method is first introduced to realize the fast generation of domain-specific and small-size region proposals. Then, a depth-connected RCNN architecture is developed to detect the cross-frame continuous defects and hence reduce the influence of randomly distributed noises.

Region Proposals

A CT image can be generally divided into two regions: the darker background regions and the brighter foreground regions (which display the AM component). Considering that the defects only located in the salient foreground regions, a natural yet reasonable idea is to generate a group of region proposals along the salient regions. Accordingly, a saliency-guided region proposal generation method is proposed with the following two steps: (1) saliency detection for semantically segmenting the entire foreground regions in pixel level and (2) region proposal generation that produces a group of sparsely distributed small boxes along the detected foreground regions.

Saliency Detection

The goal of saliency detection is to segment the foreground regions from a CT image. Recently, the U-Net has been widely adopted in developing the automatic saliency detection methods. Note that the U-Net is an encoder-decoder neural network that directly learns the mapping between the input images and the saliency maps. In other words, in the training process of a U-net model, each image needs to be annotated with a black-and-white saliency map (which segments the salient regions from the background in pixel level). Nevertheless, such saliency maps are usually not readily available in an AM dataset. On the other hand, it is clear from Fig. 4c that there are sharp changes of grey-scale intensities at the boundaries between the foreground and background regions in a grey-scale CT image. Therefore, the edge detection methods can be utilized to mark the salient foreground regions. In this paper, three different image filtering methods (i.e., Gamma correction, Sobel operation, and Laplacian edge-sharpening method) are first employed to roughly mark the saliency region, and then the saliency map is manually annotated pixel by pixel. The technical details are given as follows.

Gamma correction is an effective image filtering method for contrast enhancement [23], whose main idea is to adjust the image contrast by remodelling the saturation that maps

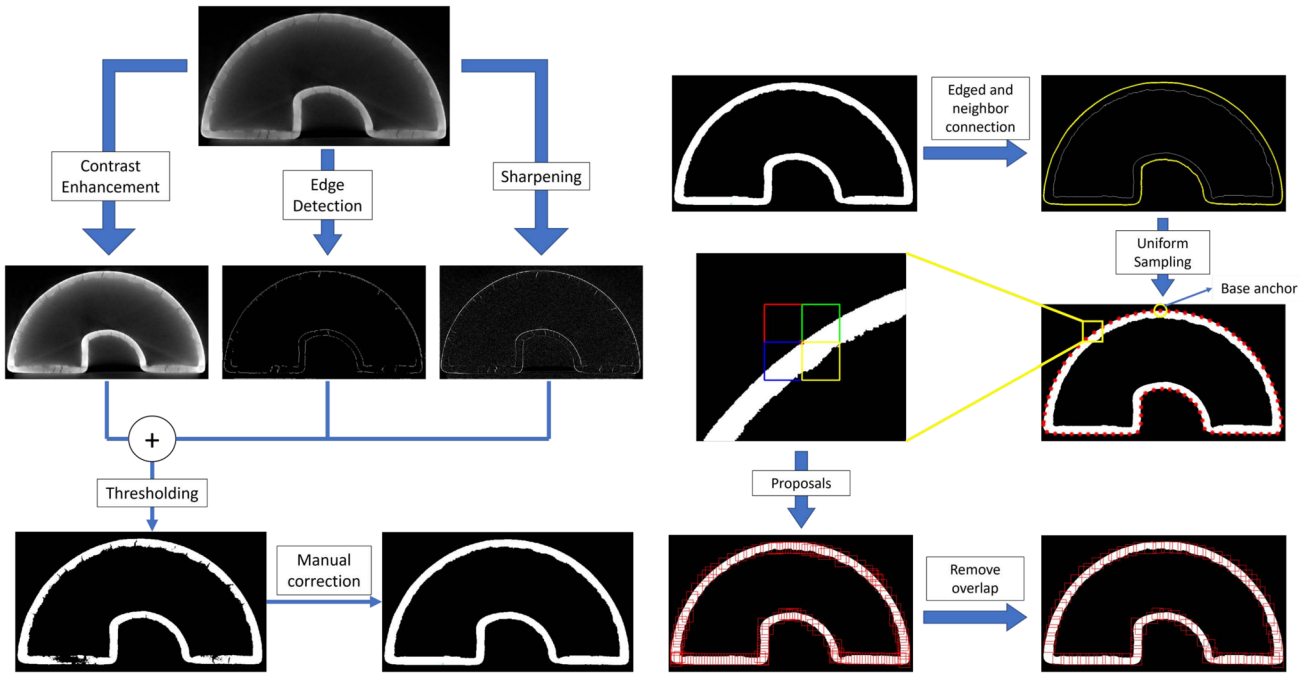
low-intensity pixels to the bottom and high-intensity pixels to the top. By default, 1% of the pixels from bright (or dark) regions are saturated at bottom (or top) intensities. Intuitively, such an operation makes the dark regions darker and the bright regions brighter. As shown in Fig. 3a, the brightness of the salient regions of AM component is indeed magnified by using the Gamma correction, which would facilitate the subsequent region segmentation.

The Sobel method (which is commonly used for edge detection) is a typical gradient-based filtering method that magnifies the local intensity changes by computing the first-order derivatives of an image in both the vertical and horizontal directions. More specifically, the Sobel method convolves the image with two 3×3 kernels, i.e., $[1, 2, 1; 0, 0, 0; -1, -2, -1]$ and $[1, 0, -1; 2, 0, -2; 1, 0, -1]$. As shown in Fig. 3a, the output is a map of the normalized gradient magnitude, which illuminates the edges of the input image. The third method is the so-called Laplacian filtering, which looks for the sharp and discontinuous intensity transitions in an image by computing the second-order derivatives and convolving the image with a kernel $[0, -1, 0; -1, 4, -1; 0, -1, 0]$. The output of Laplacian filtering is an edge-sharpened map.

It is clear that by using the above three image processing methods, we are able to obtain a contrast-adjusted image, an edge-filtered map, and an edge-sharpened map from an input image. Then, a single edge-enhanced map can be generated by pixel-wise summation. Based on such a map, a threshold with the average intensity value is set to distinguish the salient foreground regions and the background regions, thereby outputting a black-and-white saliency map. Nevertheless, it is often the case that the outputted saliency map cannot perfectly annotate the entire salient regions. To this end, the saliency map is manually corrected pixel-by-pixel, and the final saliency map is illuminated as shown in Fig. 3a. It is worth mentioning that the annotated saliency map can be utilized as the ground truth for training a U-Net model, which in turn enables the accurate and fully automatic saliency detection.

Region Proposals Generation

Based on the detected saliency map, a saliency-guided region proposal method is presented here to generate a group of small-size region proposals that are sparsely distributed along the saliency map. Similar to the traditional anchor box generation methods, we first generate a number of fixed-position anchors that are strictly distributed along the saliency map, and then generate the small-size default anchor boxes. To be more specific, the anchors are the fixed-position pixel points that are uniformly distributed along the outer edge of the saliency map. The base anchor is set to be the fiducial pixel point at the top of the saliency map, which is marked with a yellow circle in Fig. 3b. Starting from the base anchor, the rest anchors are one-by-one created every 40-pixel distance (i.e., the average width/height of the defects) along the outer



(a) Saliency region annotation.

(b) Region proposal generation.

Fig. 3 Illustration of the region proposals

edge of the saliency map. It is clear that all the anchors are uniformly distributed along the salient region.

In the following, the anchor boxes are generated based on the aforementioned anchors. For each anchor, noting that the average size of defects is nearly 40×40, four 40×40 anchor boxes are, respectively, generated at the top left, top right, bottom left and bottom right regions. Then, a large number of anchor boxes are generated around the saliency map. Nevertheless, it should be pointed out that there are some useless anchors boxes (lying outside the saliency regions) and redundant anchors boxes (with high overlapping rate). To remove such boxes, the following operations are performed: (1) a box is removed if its area covered by the saliency map is less than 50% and (2) one of the box pair is removed if their overlapping rate is larger than 50%. Consequently, as shown in Fig. 3b, only 75 boxes are left. Compared with the traditional methods (which usually generate over 1000 default boxes in different scales), the proposed saliency-guided region proposal method is able to significantly reduce the number of anchor boxes. Furthermore, the generated anchor boxes are always small and close to the positions of the real defects, which makes it easier for the RCNN method to learn the mapping between the anchor boxes and the target boxes.

Depth-Connected RCNN

In this subsection, the DC-RCNN method is put forward based on the generated anchor boxes. The overview of the proposed DC-RCNN method is depicted in Fig. 2. Different from the traditional faster RCNN method that only processes an individual frame, the proposed DC-RCNN method takes into account the between-frame connectivity with the aim to detect the defects appeared in several continuous frames and reduce the influence of the discontinuous noises. The network architecture of the proposed DC-RCNN consists of a depth-connected backbone network, an additional region proposal network (RPN), and a detection network. Specifically, the backbone network is regarded as a feature extractor that squeezes an input image into a stack of compacted feature maps. As one of the most popular backbone networks, the ResNet is usually referred to as the gold-standard architecture in the field of computer vision [52]. In this paper, there is no need to integrate the FPN into the backbone network since the proposed method only focuses on the small defects and the FPN-based multi-scale feature extraction strategy is not necessary.

To reduce the influence of noises and address the between-frame connectivity, the ResNet is extended to a depth-connected architecture, where three parallel convolutional blocks (without parameter sharing) are presented at the first computational stage. To be specific, each convolutional block receives a stacked three-channel image merged by three sequential grey-scale CT images, and then these three convolutional blocks take nine sequential frames as the input and output three individual feature maps. These feature maps are delivered to a pixel-wise summation calculator, which outputs a single feature map. As such, the backbone network of the DC-RCNN is able to connect nine sequential frames and extract the depth-connected features of all these frames. To adapt for the small boxes, the ResNet is further modified by following the cross-scale box prediction strategy in YOLOv3 [40]. Specifically, the ResNet is modified by only maintaining the three early convolutional layers. In this case, the output feature map still has a high resolution, which is more suitable for the subsequent small box prediction.

The network architecture of the proposed DC-RCNN is displayed in Table 1. It is worth mentioning that the backbone network architecture of the DC-RCNN is different from the ResNet, which implies that a pre-trained ResNet model cannot be directly utilized to transfer knowledge to the backbone network of the DC-RCNN. In this paper, to enable the transfer learning, the parameters of the first convolutional block from a pre-trained ResNet model are duplicated three times and assigned to the three parallel blocks in the proposed DC-RCNN. For the second and third convolutional blocks, the backbone network of the DC-RCNN and the original ResNet share the same architecture such that the parameters from these blocks can be transferred in a direct way. It is obvious that the parameters of a pre-trained ResNet model can be well conveyed to the backbone network of the DC-RCNN, and these parameters can be fine-tuned on the target dataset.

Following the aforementioned backbone network, the first-stage prediction is performed by the RPN, which is a shallow neural network (attached to the backbone network) with the aim to map the predefined anchor boxes to the ground-truth boxes. Different from the traditional two-layer RPN, the RPN in the proposed DC-RCNN is a three-layer network with an additional region of interest (ROI) align layer. Specifically, the ROI align layer is deployed in the first layer of the RPN, which is exploited to extract a fixed-size feature map for each anchor box by using the image warping method. The following two layers are the same as the traditional RPN with a 3×3 convolutional layer and a fully connected layer. In fact, the RPN has two tasks: (1) the classification task (predicting whether an anchor box contains a defect) and (2) the regression task (mapping the anchor box to the target box). Clearly, the output of the RPN is a group of ROIs that possibly contain defects.

Based on the ROIs and the feature map, the second-stage prediction is conducted by the detection network. The architecture of the detection network is nearly the same as that of the RPN. To be specific, the detection network is also attached to the backbone network with a ROI align layer (for ROI feature extraction) and several fully connected layers (for defect recognition). With such a detection network, the

Table 1 Architecture of the proposed DC-RCNN

Layer name	Architecture		
Conv1	kernel 7 × 7, channel 64, stride 2	kernel 7 × 7, channel 64, stride 2	kernel 7 × 7, channel 64, stride 2
	Pixel-wise summation		
	MaxPooling 3 × 3, stride 2		
Conv2	[kernel 1 × 1, channel 64 kernel 3 × 3, channel 64 kernel 1 × 1, channel 256] × 3		
Conv3	[kernel 1 × 1, channel 128 kernel 3 × 3, channel 128 kernel 1 × 1, channel 512] × 4		
Region Proposal Network	[ROI aligned with generated anchor boxes Convolutional layer: kernel 3 × 3, channel 256 Fully Connected Layer]		
Detection Network	[ROI aligned with region proposals Convolutional layer: kernel 3 × 3, channel 256 Fully Connected Layer]		

location of each ROI is further refined, which gives rise to a more accurate prediction.

Loss Function

As discussed in the "Depth-Connected RCNN" section, both the RPN and the detection network need to perform the prediction task. Therefore, it is necessary to consider two independent loss functions, i.e., the RPN loss and the detection loss. Considering that these two loss functions are in the same format, only the RPN loss is introduced as follows:

$$\mathcal{L}(p_i, t_i) = \frac{1}{N_{cls}} \sum_i \mathcal{L}_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i^* \mathcal{L}_{reg}(t_i, t_i^*) \quad (1)$$

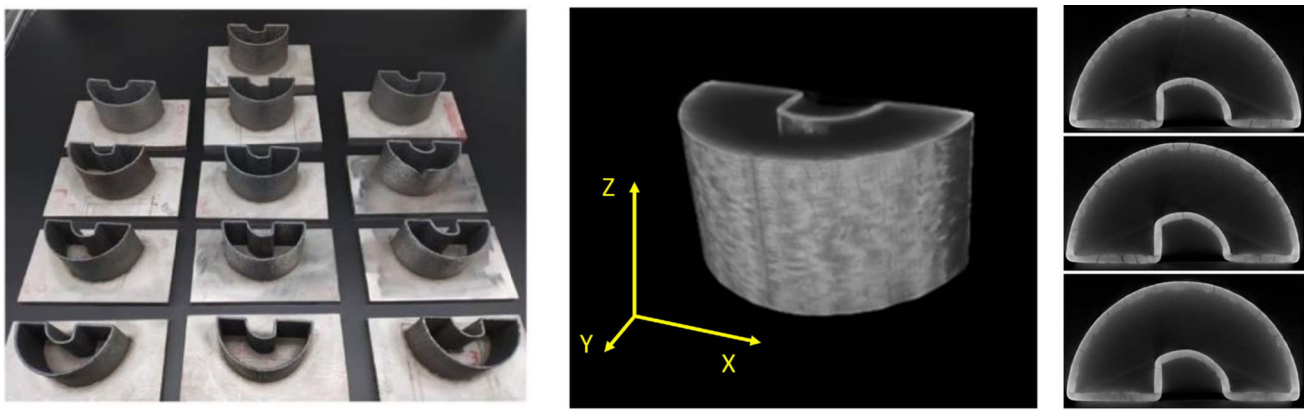
where λ is a balancing parameter ($\lambda = 10$ by default), i is the index of a region proposal, p_i is the predicted probability of the ith box containing a defect, and p_i^* is the ground-truth label. N_{cls} denotes the batch size and N_{reg} is the number of the predefined anchor boxes. The classification loss $\mathcal{L}_{cls}$ is the log loss over two categories: defects and non-defects. The regression loss $\mathcal{L}_{reg}$ is the smooth L_1 loss defined in [17]. For the box regression, the parameterized coordinates t_i and t_i^* represent, respectively, the predicted result and the ground truth.

Based on these two independent loss functions, the alternating training strategy in [41] is employed in this paper to alternatively train the RPN and the detection network. It should be pointed out that the loss function of the proposed DC-RCNN is slightly different from that of the standard faster RCNN since a different format of the bounding box is utilized. On the other hand, noting that the DC-RCNN processes nine sequential frames simultaneously, the single-frame 2D boxes should be extended to the 9-frame 3D boxes in order to meet the input requirement of the DC-RCNN. In this case, a defect marked with a 3D box is represented by six coordinates (x, y, z, w, h, d), where the coordinates (x, y, w, h) are introduced in the "Defect Detection" section, z is the starting frame of the defect and d is the corresponding length along the depth direction. Specifically, following the line of [41], the parameterization of the coordinates (x, y, z, w, h, d) is defined as follows:

$$\begin{aligned} &t_x = (x - x_a)/w_a,\ t_y = (y - y_a)/h_a, \\ &t_z = (z - z_a)/d_a,\ t_w = \log(w/w_a), \\ &t_h = \log(h/h_a),\ t_d = \log(d/d_a), \\ &t_x^* = (x^* - x_a)/w_a,\ t_y^* = (y^* - y_a)/h_a, \\ &t_z^* = (z^* - z_a)/d_a,\ t_w^* = \log(w^*/w_a), \\ &t_h^* = \log(h^*/h_a),\ t_d^* = \log(d^*/d_a) \end{aligned} \quad (2)$$

where x, x_a and x^* denote, respectively, the x-axis coordinates of the predicted box, anchor box, and ground-truth box. y, z, w, h and d follow the same settings.

The main difference between the DC-RCNN loss and Faster RCNN loss is the integration of a pair of depth parameters (z, d). Compared with the Faster RCNN that only detects single-frame 2D boxes, the DC-RCNN loss extends the z-axis coordinates, which enables annotating the 3D boxes. In this situation, the defects across multiple frames can be recog-



(a) Products with the "C" shape. (b) 3D model. (c) Projected 2D images.

Fig. 4 AM component dataset

nized as the same defect rather than independent defects at different frames.

Implementation Details

In the preprocessing phase, each input image is resized to guarantee that the maximum size of the shorter edge of the image is 800 pixels, and the longer edge does not exceed 1333 pixels [29, 41]. The resultant image is of the size 1333×745. Then, the U-Net is used to obtain the saliency map, based on which 75 anchor boxes are generated using the saliency-guided region proposal method described in the "Region Proposals" section.

The backbone network of the proposed DC-RCNN takes nine frames as the input and outputs a 84×48 feature map. Then, based on the feature map and the generated anchor boxes, the RPN performs a classification step (identifying the boxes that may contain defects) and a regression step (locating the identified boxes). The detection network makes the final decision on the exact position of each defect.

In the training phase, the transfer learning technique is used where a fast RCNN model pre-trained on the COCO database is downloaded. As described in the "Depth-Connected RCNN" section, the parameters of the backbone network and the RPN convolutional layer are transferred to the corresponding parts of the proposed DC-RCNN. In the testing phase, the DC-RCNN takes nine sequential frames as the input and outputs a 3D box, which can be easily projected back to the 2D boxes by simply assigning the detected boxes (with values of (x, y, w, h)) to the frames ranging from z to $z + d$.

In the post-processing phase, the non-maximum suppression (NMS) method is employed to remove the redundant overlapping bounding boxes. To begin with, the overlapping ratio of each pair of boxes is computed based on the intersection over union (IoU). For a box pair with the IoU value larger than 30%, the box with a lower confidence level (classification score) will be removed. Then, the remaining boxes are regarded as the final detection results.

Experiment

Dataset

In this experiment, as shown in Fig. 4a, the AM product is a "C" shape industrial component. To enable the NDT, the CT imaging technique is utilized to create cross-sectional slides. Then, these CT scans are registered with the CAD software to reconstruct an entire 3D model of the component. The grey-scale CT images are obtained by projecting the 3D model into 2D plane, where both the within-frame pixel distance and between-frame distance are set to be 0.1 mm. The dataset consists of two well-annotated components, which include, respectively, 486 and 481 frames of grey-scale images. Basic information can be found from Table 2.

Table 2 Detailed information of the dataset

	Frame number	Resolution	Defect type	Defect size (average)	Total number of defects	Pixel value range
"C" shape object_1	481	895×477	metal casting defects	37×33	6105	$0 \sim 65,535$
"C" shape object_2	486	889×497	metal casting defects	37×33	5176	$0 \sim 65,535$

The CT images are shown in Fig. 4c, and the defects are manually annotated. The characteristics of the CT images are summarized as follows.

1) The amount of the CT image data is small, and the annotation of defect regions is labor intensive. Therefore, it is difficult to obtain a large-scale dataset.
2) The size of the defect is small. Compared with the image size (around 800 × 400), the average defect size (around 40 × 40) accounts for only 5% of the image size. Such a percentage is much smaller than the commonly accepted standard (i.e., 20%) in the small object detection problem.
3) The quality of the CT images might be degraded by the undesirable noises, which are usually induced by the beam-hardening effect, computational error, and fluctuation of radiation dose. The noise regions in a CT image are often small and dark. Meanwhile, it is often the case that the pixel values and shapes of the noise areas and the defect areas are similar, which makes it really difficult (even for human experts) to distinguish them.

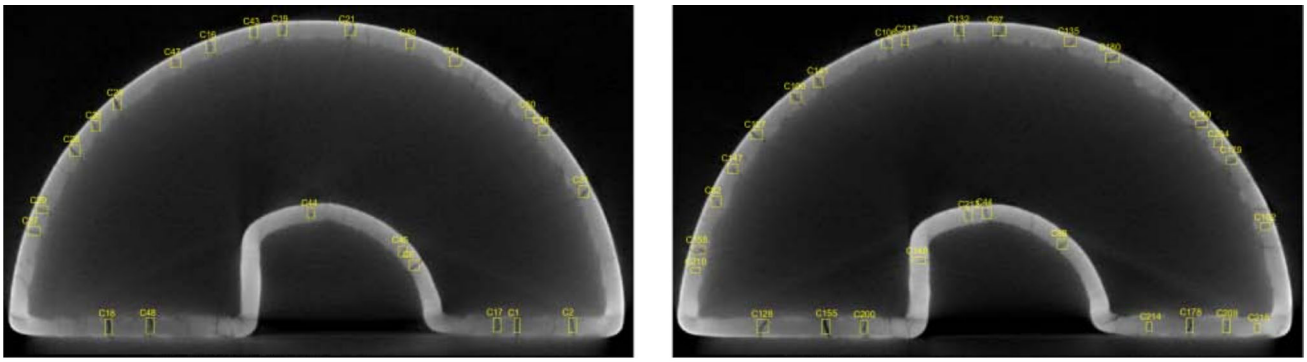
The primary limitation of this research work lies in the small-scale dataset, which only contains the "C" shape components. In this situation, the established deep learning models may be biased to "C" shape components. Despite these limitations, the dataset remains unique and valuable for advancing research in defect detection on AM components.

Defect Detection Results

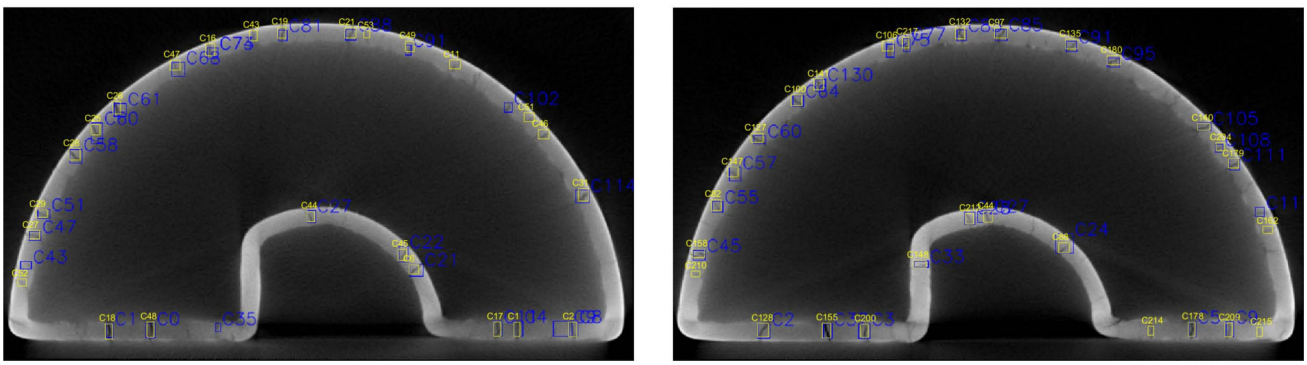
Based on the described characteristics of the dataset, a two-fold cross-validation strategy is employed to evaluate the performance of the proposed method. In this evaluation strategy, one subset is used for training the model, and the other one is employed for testing, and vice versa.

The visualization results of defect detection are shown in Fig. 5. The yellow boxes denote the detected defects, and the blue boxes are the ground truth. It can be seen that there is a high overlapping rate between the yellow boxes and the blue boxes, which indicates that the proposed DC-RCNN can effectively detect the defects.

The comparative results across several sequential frames obtained by the DC-RCNN and the faster RCNN are also provided. In Fig. 6a, the results of the DC-RCNN show that the defects across several frames are well detected. However, as shown in Fig. 6b, the performance of the faster RCNN is unstable. Obviously, some noise regions are also recognized



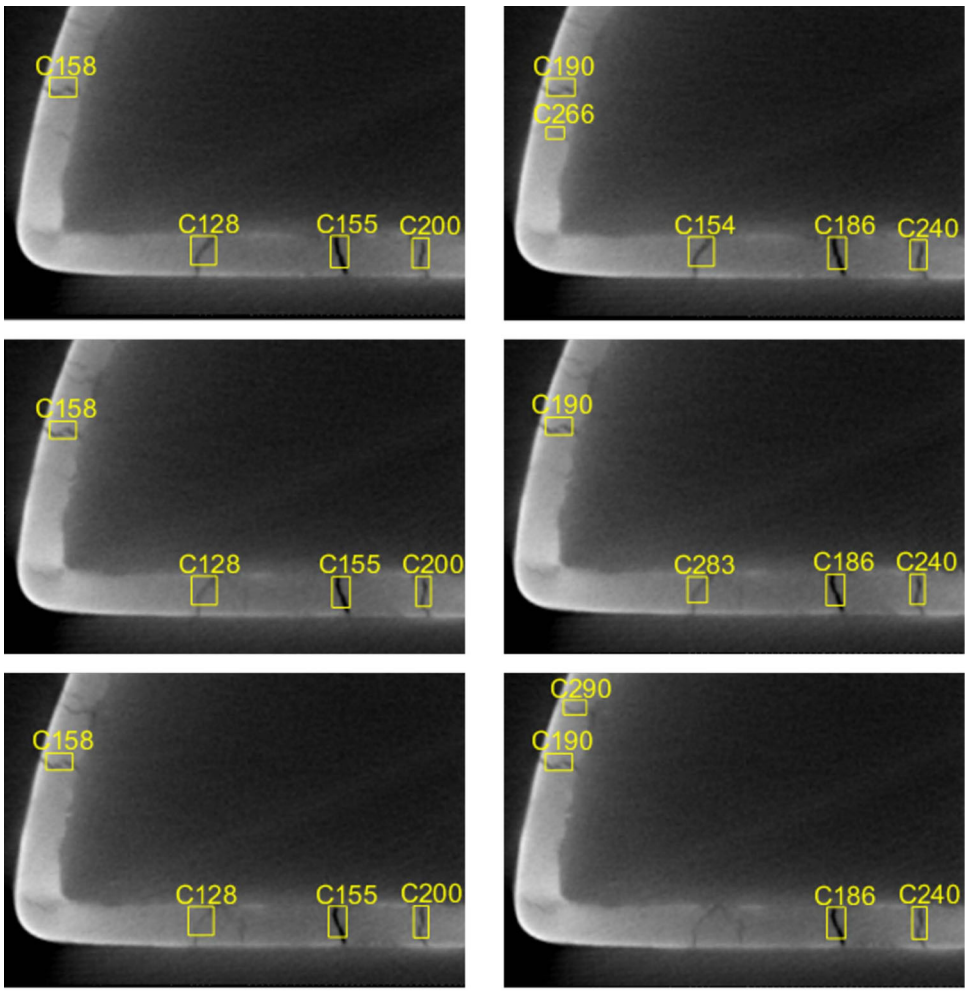
(a) Visualization results of the DC-RCNN.



(b) Comparison between the DC-RCNN and the ground truth.

Fig. 5 Defect detection results of the DC-RCNN

Fig. 6 Continuous results



(a) Continuous results of the DC-RCNN.

(b) Continuous results of the faster RCNN.

as the defects. Meanwhile, a defect continuously appeared in these three frames is not well detected. The above results confirm that the proposed DC-RCNN is able to preserve the cross-frame connectivity and is more robust to the noises when compared with the faster RCNN.

Quantitative Analysis

In this subsection, three metrics, including precision, recall and mean average precision (mAP), are utilized to evaluate the performance of the proposed DC-RCNN. Specifically, the recall metric measures the number of real defects that have been detected. The precision metric measures the accuracy of the detected defects. The mAP is a comprehensive metric that combines the precision and recall. Moreover, the computational complexity is also introduced to measure the speed of the DC-RCNN.

Comparison

The proposed DC-RCNN includes two hyperparamters, which are the batch size and the learning rate. Notably, the batch size for DC-RCNN is set to 1, meaning that a group of 9 sequential images is loaded into a mini-batch. To justify the sensitivity of the model hyperparameters, a hyperparameter tuning experiment is conducted to examine how the mAP changes with different hyperparameter settings. When the batch size is set to 1, 4, 8, and 16, the resulting mAP@0.5 scores are 73.3%, 72.4%, 72.5%, and 67.1%, respectively. Similarly, when the learning rate is set to 0.1, 0.01, 0.001, and 0.0001, the mAP@0.5 scores are 72.2%, 73.3%, 73.0%, and 73.3%, respectively. It is evident that variations in hyperparameters will not significantly influence the detection results. Therefore, the hyperparameters are set by the default values.

To validate the superiority of the DC-RCNN, four popular object detection methods (including DETR, single shot

Table 3 Performance comparison between the DC-RCNN and the popular object detection methods

	mAP@0.5(%)	mAP@0.5:0.95(%)	Precision(%)	Recall(%)
DETR	28.8	9.0	20.6	35.88
SSD	30.7	10.6	25.2	35.3
YOLO-v5s	37.4	13.3	24.6	46.1
YOLO-v5l	40.7	16.6	30.2	50.3
YOLO-v8	43.2	18.8	54.1	37.0
Faster RCNN	59.9	22.8	54.9	81.2
DC-RCNN	73.3	43.1	78.9	78.5

multibox detector (SSD), YOLO, and faster RCNN) are employed for performance comparison. The hyperparameters are specified as follows: (1) the number of epochs during training is set to 25 for DETR and 15 for the other methods; (2) the learning rate is 0.01 with a weight decay of 0.0001; and (3) the batch size follows the default settings, specified as 4, 8, 8, 1, and 1 for DETR, SSD, YOYO, faster RCNN, and DC-RCNN, receptively. It is worth noting that the batch size of DC-RCNN is 1, which means that a group of 9 sequential images is loaded into a mini-batch for model training. The experiments are conducted on a computer equipped with an Intel Core i7 CPU at 2.6 GHz and a NVIDIA GeForce RTX 2060 GPU, running Windows 10 as the operating system.

The performance comparison of the chosen object detection methods is presented in Table 3. DETR integrates CNN backbone with transformer encode-decoder structure, which has been shown to outperform YOLO in many object detection scenarios [66]. However, in this experiment, the performance of DETR is unsatisfactory, which is likely due to the limited scale of the dataset. It has been validated that CNN-based methods (e.g., SSD, YOLO and RCNN families) outperforms transformer-based methods (e.g., DETR) on small datasets due to their inductive bias [6]. CNNs are inherently more data-efficient because their architectures incorporate prior knowledge about images, reducing the need for large-scale datasets to learn these properties. Consequently, in this study, CNN-based methods are better suited than transformer-based methods.

Among CNN-based methods, SSD shows inferior performance compared to YOLO and RCNN families, indicating its reduced proficiency for detecting small defects. This is likely due to SSD's limited ability in processing small-scale objects using only low-level features without the assistance of the FPN. In contrast, the state-of-the-art YOLO and RCNN methods incorporate FPNs, which results in superior performance in detecting small-size objects.

Regarding YOLO methods, YOLO-v5 is a high-performance and user-friendly detection method, offering multiple architectures. YOLO-v5s and YOLO-v5l are two representative models for small-scale and large-scale configurations, respectively. YOLO-v8, the latest YOLO version, enhances both accuracy and speed compared to YOLO-v5. Nevertheless, as shown in Table 3, the performance of YOLO methods across all evaluation metrics is worse than that of the RCNN-based methods. Meanwhile, DC-RCNN outperforms Faster RCNN in terms of both mAP and precision.

As shown in Fig. 4d, the shape of the component remains consistent along the vertical direction (z-axis). Despite the success of defect detection on the "C" shape components, the proposed DC-RCNN has the potential to be generalized to components with structures that either remain consistent or exhibit smooth variations along a given direction.

To evaluate the efficiency of the detector, the model scale and runtime performance are presented in Table 4. Here, the number of parameters represents the scale of the model. The FLOPs are denoted by the number of floating point operations of a model, which indicates the computational cost. The frame per second (FPS) measures how many frames the model can process per second. As a two-stage detection method, Faster RCNN lags in efficiency. DETR and YOLO outperform RCNN-based methods in terms of FLOPs and FPS, with YOLO-v8 demonstrating significant reductions in model size and FLOPs, achieving real-time performance. Among the RCNN-based methods, the proposed DC-RCNN is approximately three times faster than that of the Faster RCNN, demonstrating its superior detection efficiency. In summary, the proposed DC-RCNN exhibits better defect detection performance and computational efficiency compared to the traditional Faster RCNN.

Table 4 Comparison of efficiency

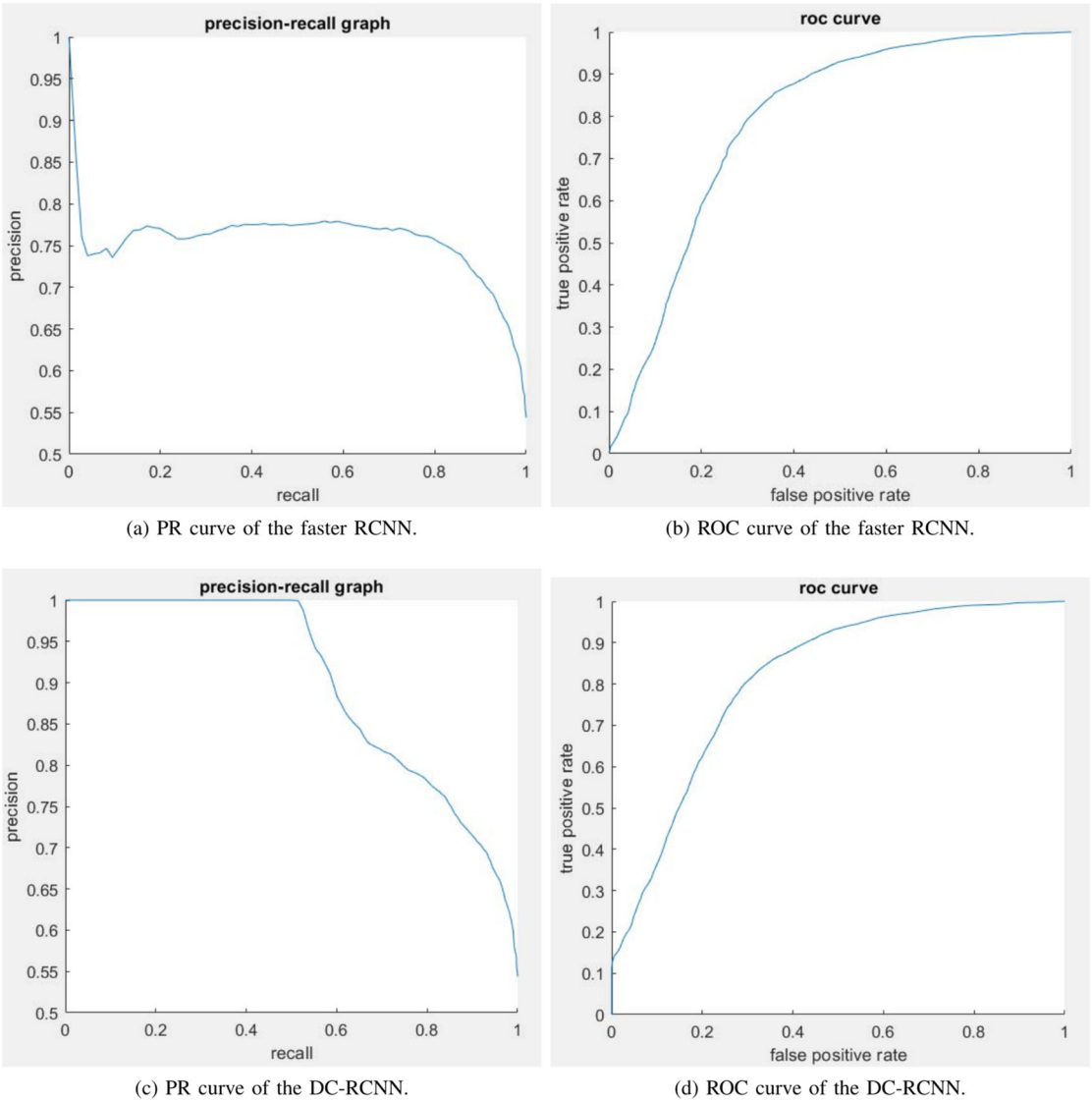
	Parameters (M)	GFLOPs (G)	FPS
DETR	41.3	86.0	8
YOLO-v5l	46.1	107.6	7
YOLO-v8	3.0	8.1	66
Faster RCNN	41.8	242.5	2
DC-RCNN	22.6	113.6	6

Table 5 Performance comparison between the DC-RCNN and the popular object detection methods

	mAP@0.5(%)	mAP@0.5:0.95(%)	Precision(%)	Recall(%)
Faster RCNN	59.9	22.8	54.9	81.2
DC-RCNN w/o DC	60.8	23.1	56.2	79.1
DC-RCNN w/o RP	72.2	40.2	76.2	78.0
DC-RCNN	73.3	43.1	78.9	78.5

Ablation Study

An ablation study has been conducted to validate the effectiveness of the DC-RCNN. The performance of the selected methods is presented in Table 5, where DC-RCNN w/o RP denotes DC-RCNN without saliency-guided region proposal method, and DC-RCNN w/o DC represents DC-RCNN without depth connectivity. Clearly, while the recall of the Faster RCNN method is the highest, its precision is only 54.97%. This indicates that although the Faster RCNN detects a large number of defects, nearly half of these detections are false positives. Compared with the faster RCNN, the DC-RCNN w/o DC method slightly improves the performance, suggesting that the proposed saliency-guided region proposal method outperforms the traditional region proposal method. Meanwhile, DC-RCNN w/o RP significantly



(a) PR curve of the faster RCNN.

(b) ROC curve of the faster RCNN.

(c) PR curve of the DC-RCNN.

(d) ROC curve of the DC-RCNN.

Fig. 7 PR curve and ROC curve

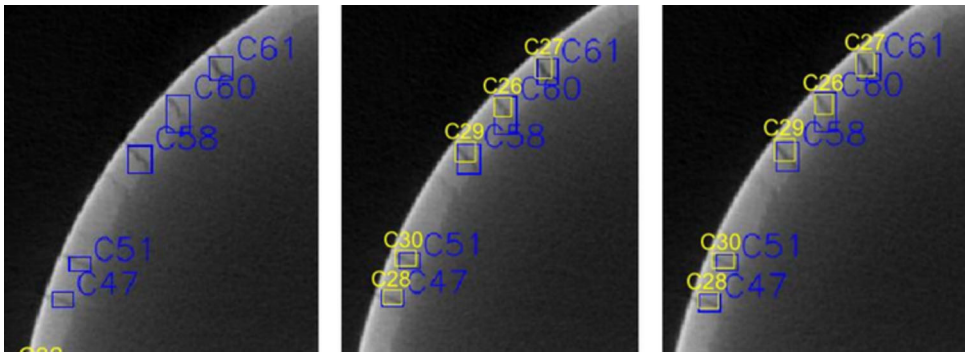


Fig. 8 Failure detections

enhances performance in terms of mAP and precision, indicating that the depth-connected feature extraction method effectively improves the detection accuracy. Finally, the proposed DC-RCNN achieves the highest mAP and precision, demonstrating its superior ability to accurately detect defects and its stronger robustness against noises compared to the Faster RCNN method.

To further justify the effectiveness of the proposed DC-RCNN, the Precision-Recall (PR) curve and the Receiver Operating Characteristic (ROC) curve are used. As shown in Fig. 7, the DC-RCNN achieves higher Area Under Curve (AUC) than the faster RCNN on both the PR and ROC curves, which indicates that the DC-RCNN outperforms the faster RCNN. Faster RCNN and DC-RCNN show similar ROC curves, indicating comparable performance in defect identification.

In contrast, the PR curves emphasize the correct identification of objects. As shown in Fig. 7a, the PR curve of the faster RCNN shows a sharp drop at the top-left corner, which indicates that recall increases much faster than precision. In other words, a large number of *fake* defects are detected by the faster RCNN. On the other hand, the PR curve for the DC-RCNN remains much steadier, demonstrating its outstanding performance in detecting real defects correctly.

Failure and Limitation Analysis

Although the DC-RCNN achieves promising defect detection results, it is still necessary to perform failure analysis. A case study of detection failure is illustrated in Fig. 8 which displays three sequential CT images. The yellow boxes denote the detected defects, and the blue boxes represent the ground truth. It can be seen that while the DC-RCNN fails to detect the defects in the first frame, it can detect them in the subsequent frames successfully. This failure may be attributed to the depth-connected operation. Since the DC-RCNN processes nine frames together and makes a single decision for all of them subject to the majority features, it is possible that some frames with defects are incorrectly grouped with the majority that contains no defect. This is a limitation of the proposed DC-RCNN defect detection approach, which will be addressed in the future.

Conclusion

In this paper, a novel DC-RCNN method has been proposed for the AM defect detection with CT images. To detect the small-size defects, a saliency-guided region proposal method has been put forward to generate small-size region proposals. Then, in order to reduce the impact of the randomly distributed noises, a depth-connected backbone network has been constructed to extract and connect the consistent features across multiple frames. Accordingly, the proposed method is able to effectively detect the small-size defects and reduce the influence of noises. The experimental results have demonstrated that the proposed DC-RCNN can achieve outstanding detection performance on the noise-corrupted CT images. It is worth pointing out that the utilized dataset in this study contains only images of "C" shape components, which may bias the established defect detector toward this specific shape of the component, thus limiting the model generalization ability to other types of components. Future research would focus on the following targets: (1) employing advanced machine learning techniques to improve the DC-RCNN for defect detection [11, 47, 48, 53, 59, 60]; (2) employing evolutionary computation methods to tune the hyperparameters of the DC-RCNN [10, 57, 58]; (3) integrating latest network designs (e.g., the encoder-decoder structure and adversarial learning) to improve the model generalization ability [7, 63, 64]; (4) extending the current single-modality (image-based) approach to a multi-modality framework which embeds control stream data into the system to reduce inductive biases inherent to image-only modalities [2, 46]; and (5) employing state-of-the-art signal processing methods (e.g., the attention mechanism, Kalman filtering, and state estimation) to further enhance the depth connectivity [14, 24, 32, 33, 45, 49, 54, 65, 67, 68].

Author Contributions Yiming Wang: conceptualization, formal analysis, investigation, methodology, software, visualization, writing—original draft; Zidong Wang: conceptualization, formal analysis, methodology, resources, supervision, writing—review and editing; Weibo Liu: formal analysis, data curation, methodology, writing—original draft, writing—review and editing; Nianyin Zeng: conceptualization, validation, review and editing; Stanislao Lauria: conceptualization, project administration, resources, supervision, writing—review and editing; Camilo Prieto: project administration, resources, supervision; Fredrik Sikström: project administration, resources, supervision. Hui Yu: conceptualization, methodology, supervision, writing—review and editing. Xiaohui Liu: conceptualization, methodology, supervision, writing—review and editing.

Funding This work was supported in part by the European Union's Horizon 2020 Research and Innovation Programme under Grant 820776 (INTEGRADDE), the Royal Society of the UK, the BRIEF funding of Brunel University London, and the Alexander von Humboldt Foundation of Germany.

Data Availability The data that support the findings of this study are not openly available due to data privacy and are available from the corresponding author upon reasonable request.

Declarations

Conflict of Interest The authors declare no competing interests.

References

1. Carion N, Massa F, Synnaeve G, Usunier N, Kirillov A, Zagoruyko S. End-to-end object detection with transformers. In: Proceedings of European Conference on Computer Vision. 2020. pp. 213–229.
2. Chen H, Chen Q, Shen B, Liu Y. Parameter learning of probabilistic Boolean control networks with input-output data. Int J Netw Dyn Intell. 2024;3(1):100005.
3. Chen X, Yang R, Xue Y, Yang C, Song B, Zhong M. A novel momentum prototypical neural network to cross-domain fault diagnosis for rotating machinery subject to cold-start. Neurocomput. 2023;555:126656.
4. Chen Y, Yang R, Huang M, Wang Z, Liu X. Single-source to single-target cross-subject motor imagery classification based on multisubdomain adaptation network. IEEE Trans Neural Syst Rehabil Eng. 2022;30:1992–2002.
5. Diwakar M, Kumar M. A review on CT image noise and its denoising. Biomed Signal Process Control. 2018;42:73–88.
6. Dosovitskiy A. An image is worth 16x16 words: transformers for image recognition at scale. 2020. arXiv:2010.11929.
7. Dou J, Song Y. An improved generative adversarial network with feature filtering for imbalanced data. Int J Netw Dyn Intell. 2023;2(4):100017.
8. Du G, Wang K, Lian S, Zhao K. Vision-based robotic grasping from object localization, object pose estimation to grasp estimation for parallel grippers: a review. Artif Intell Rev. 2021;54(3):1677–734.
9. Fang J, Wang Z, Liu W, Chen L, Liu X. A new particle-swarm-optimization-assisted deep transfer learning framework with applications to outlier detection in additive manufacturing. Eng Appl Artif Intell. 2024;131:107700.
10. Fang W, Shen B, Pan A, Zou L, Song B. A cooperative stochastic configuration network based on differential evolutionary sparrow search algorithm for prediction. Syst Sci Control Eng. 2024;12(1):2314481.
11. Feng S, Li X, Zhang S, Jian Z, Duan H, Wang Z. A review: state estimation based on hybrid models of Kalman filter and neural network. Syst Sci Control Eng. 2023;11(1):2173682.
12. Ferguson MK, Ronay A, Lee Y-TT, Law KH. Detection and segmentation of manufacturing defects with convolutional neural networks and transfer learning. Smart Sustain Manuf Syst. 2018;2:1–43.
13. Fuchs P, Kröger T, Garbe CS. Defect detection in CT scans of cast aluminum parts: a machine vision perspective. Neurocomput. 2021;453:85–96.
14. Gao X, Deng F, Shang W, Zhao X, Li S. Attack-resilient asynchronous state estimation of interval type-2 fuzzy systems under stochastic protocols. Int J Syst Sci. 2024;55(13):2688–700.
15. Garland AP, White BC, Jared BH, Heiden M, Donahue E, Boyce BL. Deep convolutional neural networks as a rapid screening tool for complex additively manufactured structures. Addit Manuf. 2020;35:101217.
16. Gaus YFA, Bhowmik N, Akcay S, Breckon T. Evaluating the transferability and adversarial discrimination of convolutional neural networks for threat object detection and classification within X-ray security imagery. In: Proceedings of 18th IEEE International Conference on Machine Learning and Applications. Boca Raton, FL, USA; 2019. pp. 420–425.
17. Girshick R. Fast R-CNN. In: Proceedings of the IEEE International Conference on Computer Vision. Boston, MA, USA; 2015. pp. 1440–1448.
18. Gobert C, Reutzel EW, Petrich J, Nassar AR, Phoha S. Application of supervised machine learning for defect detection during metallic powder bed fusion additive manufacturing using high resolution imaging. Addit Manuf. 2018;21:517–28.
19. Han C, Li G, Liu Z. Two-stage edge reuse network for salient object detection of strip steel surface defects. IEEE Trans Instrum Meas. 2022;71:5019812.
20. Han F, Liu S, Zou J, Ai Y, Xu C. Defect detection: defect classification and localization for additive manufacturing using deep learning method. In: Proceedings of 2020 21st international conference on electronic packaging technology. Guangzhou, China; 2020.
21. He X, Wang T, Wu K, Liu H. Automatic defects detection and classification of low carbon steel WAAM products using improved remanence/magneto-optical imaging and cost-sensitive convolutional neural network. Meas. 2021;173:108633.
22. Honarvar F, Varvani-Farahani A. A review of ultrasonic testing applications in additive manufacturing: defect evaluation, material characterization, and process control. Ultrason. 2020;108:106227.
23. Huang SC, Cheng FC, Chiu YS. Efficient contrast enhancement using adaptive gamma correction with weighting distribution. IEEE Trans Image Process. 2013;22(3):1032–41.
24. Jin F, Ma L, Zhao C, Liu Q. State estimation in networked control systems with a real-time transport protocol. Syst Sci Control Eng. 2024;12(1):2347885.

25. Ke L, Zhang Y, Yang B, Luo Z, Liu Z. Fault diagnosis with synchrosqueezing transform and optimized deep convolutional neural network: an application in modular multilevel converters. Neurocomput. 2021;430:24–33.
26. Khosravani MR, Reinicke T. On the use of X-ray computed tomography in assessment of 3D-printed components. J Nondestruct Eval. 2020;39:7.
27. Lema DG, Pedrayes OD, Usamentiaga R, Venegas P, García DF,. Automated detection of subsurface defects using active thermography and deep learning object detectors. IEEE Trans Instrumen Meas. 2022;7:4503213.
28. Lian J, Jia W, Zareapoor M, Zheng Y, Luo R, Jain DK, Kumar N. Deep-learning-based small surface defect detection via an exaggerated local variation-based generative adversarial network. IEEE Trans Ind Inf. 2019;16(2):1343–51.
29. Lin TY, Dollár P, Girshick R, He K, Hariharan B, Belongie S. Feature pyramid networks for object detection. in Proceedings of the IEEE conference on computer vision and pattern recognition. Honolulu, Hawaii, USA; 2017. pp. 2117–2125.
30. Lin TY, Goyal P, Girshick R, He K, Dollár P. Focal loss for dense object detection. In: Proceedings of the IEEE international conference on computer vision. Honolulu, Hawaii, USA; 2017. pp. 2980–2988.
31. Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu CY, Berg AC. SSD: single shot multibox detector. In: Proceedings of European Conference on Computer Vision. Amsterdam, The Netherlands; 2016. pp. 21–37.
32. Liu Y, Coombes M, Liu C. Mesh-based consensus distributed particle filtering for sensor networks. IEEE Trans Signal Inf Process over Netw. 2023;9:346–56.
33. Liu Y, Wang Z, Liu C, Coombes M, Chen W-H. A novel algorithm for quantized particle filtering with multiple degrading sensors: degradation estimation and target tracking. IEEE Trans Ind Inf. 2023;19(4):5830–8.
34. Ma G, Wang Z, Liu W, Fang J, Zhang Y, Ding H, Yuan Y. Estimating the state of health for lithium-ion batteries: a particle swarm optimization-assisted deep domain adaptation approach. IEEE/CAA J Autom Sin. 2023;10(7):1530–43.
35. Maskery I, Aboulkhair NT, Corfield MR, Tuck C, Clare AT, Leach RK, Wildman RD, Ashcroft IA, Hague RJM. Quantification and characterisation of porosity in selectively laser melted Al-Si10-Mg using X-ray computed tomography. Mater Charact. 2016;111:193–204.
36. Nguyen ND, Do T, Ngo TD, Le DD. An evaluation of deep learning methods for small object detection. J Electr Comput Eng. 2020;2020:3189691.
37. Pereira T, Kennedy JV, Potgieter J. A comparison of traditional manufacturing vs additive manufacturing, the best method for the job. Procedia Manuf. 2019;30:11–8.
38. Qu Z, Jiang P, Zhang W. Development and application of infrared thermography non-destructive testing techniques. Sensors. 2020;20(14):3851.
39. Redmon J, Farhadi A. YOLO9000: better, faster, stronger. In: Proceedings of the IEEE conference on computer vision and pattern recognition. Honolulu, Hawaii, USA; 2017. pp. 7263–7271.
40. Redmon J, Farhadi A. YOLOv3: an incremental improvement. 2018. arXiv:1804.02767.
41. Ren S, He K, Girshick R, Sun J. Faster R-CNN: towards real-time object detection with region proposal networks. Adv Neural Inf Process Syst. 2015;28.
42. Sun X, Gu J, Huang R, Zou R, Giron Palomares B. Surface defects recognition of wheel hub based on improved faster R-CNN. Electr. 2019;8(5):481.
43. Tian L, Wang Z, Liu W, Cheng Y, Alsaadi FE, Liu X. An improved generative adversarial network with modified loss function for crack detection in electromagnetic nondestructive testing. Complex Intell Syst. 2022;8:467–76.
44. Wang C, Wang Z, Dong H. A novel prototype-assisted contrastive adversarial network for weak-Shot learning with applications: handling weakly labeled data. IEEE/ASME Trans Mechatron. 2024;29(1):533–43.
45. Wang D, Wen C, Feng X. Deep variational Luenberger-type observer with dynamic objects channel-attention for stochastic video prediction. Int J Syst Sci. 2024;55(4):728–40.
46. Wang L, Huang M, Yang R, Liang H, Han J, Sun Y. Survey of movement reproduction in immersive virtual rehabilitation. IEEE Trans Vis Comput Graph. 2022;29(4):2184–202.
47. Wang R, Liang J. The effect of multiscale parameters on the spiking properties of the morphological neuron with excitatory autapse. Syst Sci Control Eng. 2024;12(1):2313865.
48. Wang W, Ma L, Rui Q, Gao C. A survey on privacy-preserving control and filtering of networked control systems. Int J Syst Sci. 2024;55(11):2269–88.
49. Wang Y, Wen C, Wu X. Fault detection and isolation of floating wind turbine pitch system based on Kalman filter and multi-attention 1DCNN. Syst Sci Control Eng. 2024;12(1):2362169.
50. Wang YA, Shen B, Zou L, Han QL. A survey on recent advances in distributed filtering over sensor networks subject to communication constraints. Int J Netw Dyn Intell. 2023;2(2):100007.
51. Westphal E, Seitz H. A machine learning method for defect detection and visualization in selective laser sintering based on convolutional neural networks. Addit Manuf. 2021;41:101965.
52. Wightman R, Touvron H, Jégou H. ResNet strikes back: an improved training procedure in timm. 2021. arXiv:2110.00476.
53. Wu D, Li Z, Yu Z, He Y, Luo X. Robust low-rank latent feature analysis for spatiotemporal signal recovery. IEEE Trans Neural Netw Learn Syst. in press. 2023. https://doi.org/10.1109/TNNLS.2023.3339786
54. Wu P, Li H, Hu L, Ge J, Zeng N. A local-global attention fusion framework with tensor decomposition for medical diagnosis. IEEE/CAA Journal of Autom Sin. 2024;11(6):1536–8.
55. Xiao L, Lu M, Huang H. Detection of powder bed defects in selective laser sintering using convolutional neural network. Int J Adv Manuf Technol. 2020;107:2485–96.
56. Xiao L, Wu B, Hu Y. Missing small fastener detection using deep learning. IEEE Trans Instrum Meas. 2020;70:1–9.
57. Xue J, Shen B. A survey on sparrow search algorithms and their applications. Int J Syst Sci. 2024;55(4):814–32.
58. Xue Y, Li M, Arabnejad H, Suleimenova D, Jahani A, Geiger BC, Boesjes F, Anagnostou A, Taylor SJE, Liu X, Groen D. Many-objective simulation optimization for camp location problems in humanitarian logistics. Int J Netw Dyn Intell. 2024;3(3):100017.
59. Yan S, Xia Y, Zhai DH. Iterative learning of output feedback stabilising controller for a class of uncertain nonlinear systems with external disturbances. Int J Syst Sci. 2024;55(13):2780–95.
60. Yang Q, Zhou J, Wei Z. Time perspective-enhanced suicidal ideation detection using multi-task learning. Int J Netw Dyn Intell. 2024;3(2):100011.
61. Yue X, Chen J, Zhong G. Metal surface defect detection based on Metal-YOLOX. Int J Netw Dyn Intell. 2023;2(4):100020.
62. Zeng N, Wu P, Wang Z, Li H, Liu W, Liu X. A small-sized object detection oriented multi-scale feature fusion approach with application to defect detection. IEEE Trans Instrumen Meas. 2022;71:3507014.
63. Zhang R, Liu H, Liu Y, Tan H. Dynamic event-triggered state estimation for discrete-time delayed switched neural networks with constrained bit rate. Syst Sci Control Eng. 2024;12(1):2334304.
64. Zhang W, Fan Y, Song Y, Tang K, Li B,. A generalized two-stage tensor denoising method based on the prior of the noise location and rank. Expert Syst Appl. 2024;255:124809.

65. Zhang Y, Zou L, Liu Y, Ding D, Hu J. A brief survey on nonlinear control using adaptive dynamic programming under engineering-oriented complexities. Int J Syst Sci. 2023;54(8):1855–72.
66. Zhao Y, Lv W, Xu S, Wei J, Wang G, Dang Q, Liu Y, Chen J. DETRs beat YOLOS on real-time object detection. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. Seattle, Washington, USA; 2024. pp. 16965–16974.
67. Zou L, Wang Z, Shen B, Dong H. Encryption-decryption-based state estimation with multi-rate measurements against eavesdroppers: a recursive minimum-variance approach. IEEE Trans Autom Control. 2023;68(12):8111–8.
68. Zou L, Wang Z, Shen B, Dong H. Moving horizon estimation over relay channels: dealing with packet losses. Autom. 2023;155:111079.

Authors and Affiliations

Yiming Wang[1] · Zidong Wang[1] · Weibo Liu[1] · Nianyin Zeng[2] · Stanislao Lauria[1] · Camilo Prieto[3] · Fredrik Sikström[4] · Hui Yu[5] · Xiaohui Liu[1]

✉ Weibo Liu
Weibo.Liu2@brunel.ac.uk

Yiming Wang
yiming.wang@port.ac.uk

Zidong Wang
Zidong.Wang@brunel.ac.uk

Nianyin Zeng
zny@xmu.edu.cn

Stanislao Lauria
Stasha.Lauria@brunel.ac.uk

Camilo Prieto
camilo.prieto@aimen.es

Fredrik Sikström
fredrik.sikstrom@hv.se

Hui Yu
Hui.Yu@glasgow.ac.uk

Xiaohui Liu
Xiaohui.Liu@brunel.ac.uk

1 Department of Computer Science, Brunel University London, Uxbridge, Middlesex UB8 3PH, UK

2 Department of Instrumental and Electrical Engineering, Xiamen University, Fujian 361005, China

3 AIMEN Technology Centre, E36418 O Porriño, Spain

4 Department of Engineering Science, University West, 461-32 Trollhättan, Sweden

5 School of Psychology and Neuroscience, University of Glasgow, G12 8QB Glasgow, UK