Title:
**Deep Learning-based Automatic Damage Recognition and Spatial Localization for Remanufacturing/Repair**

Authors:
Yufan Zheng, yufan6@ualberta.ca, University of Alberta
Harshavardhan Mamledesai, mamledes@ualberta.ca, University of Alberta
Habiba Imam, imam@ualberta.ca, University of Alberta
Rafiq Ahmad, rafiq.ahmad@ualberta.ca, University of Alberta

Keywords:
Deep learning; Remanufacturing; Repair; Automatic inspection; Mask-RCNN; Spatial localization

DOI: 10.14733/cadconfP.2020.381-385

Introduction:
In contemporary manufacturing, the increasing developments and over-exploitation of resources result in numerous "end-of-life" products. However, the products have not been used thoroughly, and their product life-cycle can be extended by remanufacturing or repair process. It has been widely emphasised because it enables the remanufactured product to be sold as a new product and also maintains the intrinsic energy of the "end-of-life" product without creating redundant energy. It is reported that remanufacturing reduces cost by 50%, energy by 60%, material by 70% and air pollution by 80% as compared to a conventional manufacturing process [5].

Although significant benefits can be gained from remanufacturing/repair, there are still numerous challenges to implement it in industry. One of the reasons is that, compared to manufacturing process, stochastic returns of used part and their uncontrollable quality condition result in a high degree of uncertainty for remanufacturing process [6]. The uncertainty surrounding the return of the parts complicates the remanufacturing process. Recently, great efforts have been devoted to the remanufacturing process plan optimization with uncertainties [5,8]. These optimization frameworks are initialized with characterized and quantified fault features (e.g. crack, dent, scratch, abrasion). Visual or manual inspection determines the fault feature characterization which indicates damage type, damage location and damage degree. These three factors play a key role to generate an optimal process plan with different additive operations (e.g. chromium plating, arc welding, cold welding, laser cladding, thermal spraying) and subtractive operations (e.g. milling, grinding) with heuristic algorithms. The current visual or manual inspection methods require extensive human intervention, and quality of process is hard to be stable. Therefore, an automated inspection approach for remanufacturing is urgently demanded. For this reason, an increasing level of interest in research on the automated or semi-automated inspection for remanufacturing or repair has been witness over recent years [2,8,9,10]. By summarizing these research results, to the best of the authors' knowledge, an automatic approach which enables damage recognition and spatial localization simultaneously for remanufacturing has not been discovered. In this study, a deep learning-based damage recognition and spatial localization method is proposed, which can classify different damage features and localize in the global three-dimensional coordinate.

Main Idea:
The main objective of this study is to automatically detect damages from a remanufacturing part. The study proposes a detection strategy based on a deep-learning technique to recognize and localize

damages. The flowchart is shown in Fig. 1. There are four main steps of the process: (1) Data acquisition for the RGB image and depth data by a depth camera; (2) the damage recognition and segmentation using a Mask-RCNN-based method, providing damage segments with recognized damage type; (3) the localization of the damage determined by the integration of damage segments and point cloud from the depth data.

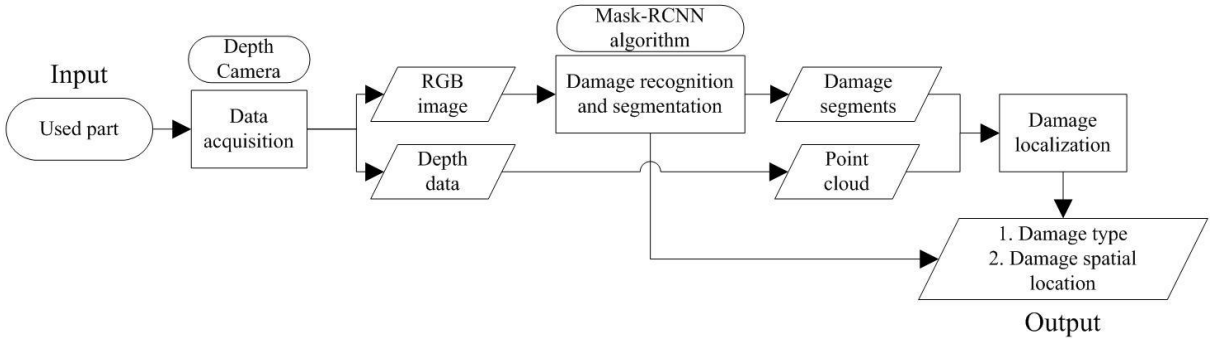


Fig. 1: The flowchart of the proposed method.

In this study, the damage recognition and segmentation method is based on a Mask-RCNN architecture [3]. The proposed damage recognition and segmentation method is illustrated in Fig. 2. As shown, it is composed of four modules: (1) Input the original image to be processed into a pre-trained convolutional backbone to extract features and to obtain a feature map; (2) the region proposal network (RPN) proposes region of interest (RoI) in the feature map with a set of rectangular object proposals; (3) each RoI generates a fixed size feature map by RoIAlign layer; (4) the fixed size feature map goes through two branches of layers for objective classification, frame regression and pixel segmentation.

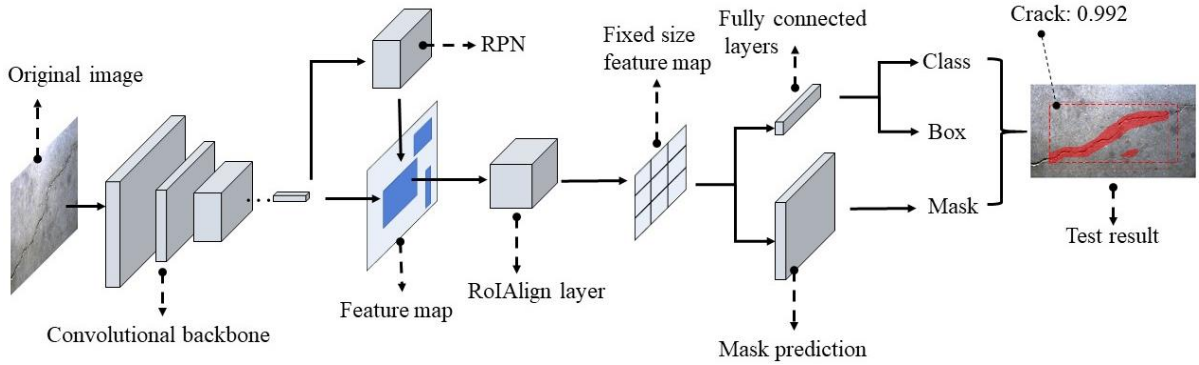


Fig. 2: The neural network architecture of the proposed damage recognition and segmentation method.

Convolutional backbone is composed of a series convolutional neural network (CNN) to extract feature maps from the image. The properties of a neural network backbone are characterized by selection and arrangement of different layers. Deeper networks generally allow to extract more complicate features from the input image, meanwhile stacking more layers will result issues for training, due to degradation problem. The residual network (ResNet) was designed to address this problem in deeper neural networks (up to 152 layers) [4] by reformulating its layers as residual learning function with reference to the layer input. A block of the ResNet can be mathematically represented in Eq. (1).

$$y_l = x_l + F\ x_l, W_l$$
$$x_{l+1} = f\ y_l \tag{1}$$

where $x_l$ and $x_{l+1}$ are the *l*-th layer's input and output respectively; $F()$ represents the residual mapping to be learned.

From *l*-th layer to *L*-th layer for the feature learning can be represented as:

$$x_L = x_l + \sum_{i=l}^{L-1} F\ x_i, W_i \tag{2}$$

The gradient for the loss function at *l*-th layer can be derived by chain rule of backpropagation:

$$\frac{\partial Loss}{\partial x_l} = \frac{\partial Loss}{\partial x_L} \cdot \frac{\partial x_L}{\partial x_l} = \frac{\partial Loss}{\partial x_L} \cdot \left(1 + \frac{\partial}{\partial x_l} \sum_{i=l}^{L-1} F\ x_i, W_i\right) = \frac{\partial Loss}{\partial x_L} + \frac{\partial Loss}{\partial x_L} \cdot \frac{\partial}{\partial x_l} \sum_{i=l}^{L-1} F\ x_i, W_i \tag{3}$$

Eq. (2) and (3) imply that the signal can be directly propagated from any unit to another, both forward and backward.

Generally, Mask RCNN model adopts ResNet101 as the backbone. It is a very deep network with 101 layers and approximately 27 million parameters. In this study, because the damage category is simple and the dataset is limited, a smaller backbone ResNet50 is used to improve the running speed for training. Feature pyramid network (FPN) [7] uses a top-down architecture with lateral connections to build an in-network feature pyramid, which address the multi-scale object recognition problem. Overall, this study uses the combination of ResNet50 and FPN as the backbone for feature extraction.

The second module in the proposed damage detection and recognition is RPN. The original image passes through the ResNet50 and FPN convolutional network and outputs a set of convolutional feature maps. Then, a sliding window runs over the feature maps. In each sliding windows, multiple region proposals are predicted based on a predefined number of anchor box. An anchor box is a reference box with a set of scales and aspect ratios and in centered at the sliding windows. A proposed area can generate a large number of anchors with different sizes and aspect ratios, and they overlap to cover as many areas as possible. In this study, the algorithm uses 9 different sizes of anchors as (128*128, 256*256, 512*512) with aspect ratios of (1:1, 1:2, 2:1). Positive or negative anchors are computed by considering the interest-over-union (IoU) between the analyzed anchor and ground-truth bounding boxes on the image. The IoU is calculated by Eq. (4). In this paper, positive anchors are those that have an IoU is greater or equal to 0.7 ith any ground-truth object and negative anchors are those have IoU is smaller or equal to 0.3. The anchors with IoU between 0.3 and 0.7 are not considered for the training objective. The positive anchors are then process to the proposal classification.

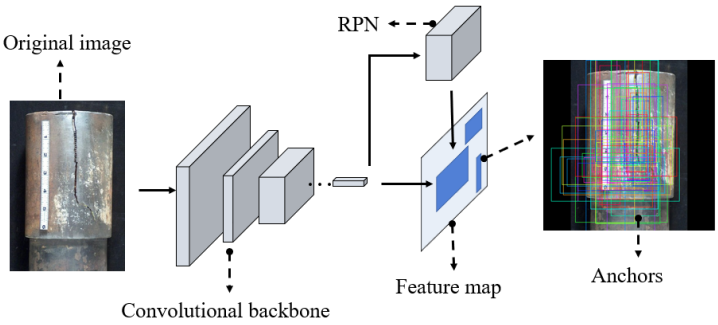


Fig. 3: Region Proposal Network (RPN) of the proposed method.

$$\mathrm{IoU} = \frac{A_{overlap}}{A_{union}} \tag{4}$$

where $A_{overlap}$ is the area of overlap and $A_{union}$ is the area of union.

In the Mask-RCNN model, the mask branch must determine whether a given pixel is inside the target in a pixel level accuracy. However, the original image has been already convolved and pooled which results in size changes. Therefore, the interest-region alignment layer (RoIAlign) [7] is used to help image target be accurately positioned. In RoIAlign, the bi-linear interpolation method is implemented to calculate the exact position of the sample points in each unit, retaining its decimal, and the uses the maximum pooling or average pooling operation to output the last fixed-size ROI. RoIAlign calculate the value of each sampling point by bi-linear interpolation method from the nearby grid points on the feature map. Finally, the maximum pooling or average pooling operation is performed to obtain the feature map of fixed size.

The multi-tasking loss function of the Mask R-CNN training process is defined in Eq. (5), where $L$ is the training total loss; $L_{cls}$ is the classification loss, $L_{box}$ is the bounding-box loss, and $L_{mask}$ is the mask loss.

$$L = L_{cls} + L_{box} + L_{mask} \tag{5}$$

The variables for $L_{cls}$ and $L_{box}$ are defined in [1], as shown in Eq. (6). Each training RoI is labeled with a ground-truth class $u$ and a ground-truth bounding-box regression target $v$.

$$L_{cls} + L_{box} = L_{cls}\ \ p, u\ \ + \lambda \left[u \geq 1\right] L_{loc}\ \ t^{u}, v \tag{6}$$

where $u$ is the label of each training RoI with a ground-truth class; $v$ is a label of each RoI with a ground-truth bounding-box regression target; $t^{u}$ specifies a scale-invariant translation and log-space height/width shift relative to $u$ class; $p = p_0, \ldots, p_K$ represents the probability distribution over $K+1$ categories; $\left[u \geq 1\right]$ denotes the Iverson bracket indicator function that evaluates to 1 when $u \geq 1$ and 0 otherwise.

The $L_{mask}$ is calculated by taking the average cross entropy of the all pixels on the RoI, as:

$$L_{mask} = -\frac{1}{N}\left[y_i \ln a_i + \ 1 - y_i\ \ln\ 1 - a_i\ \right] \tag{7}$$

where $x_i$ and $b_i$ are the prediction value and true value of the *i*-th pixel in the positive RoI, respectively; $N$ indicates the number of pixels in the positive RoI.

In this study, spatial localization of the damage area is achieved by finding the mapping relations between the 2D coordinates in image and 3D spatial coordinates by the depth sensor model, as shown in Eq. (8).

$$z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{f_x}{dx} & 0 & u_0 \\ 0 & \frac{f_y}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R_z R_y R_z & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{8}$$

where $u$ and $v$ are the 2D image coordinates; $u_0$ and $v_0$ are the origin of the 2D coordinate system; $f_x$ and $f_y$ are the focal length along $x$ and $y$ direction, respectively; $R_z R_y R_z$ and $T$ are the rotation matrix and translation matrix from the camera coordinate system to the global coordinate system, $X$,

$Y$ , $Z$ are the 3D coordinates under global coordinate. To simplify this problem, the authors coincides the camera coordinate system and the global coordinate system. Then, the 3D coordinates of the damage area can be calculated as: $X = \frac{u - u_0 \cdot z_c \cdot dx}{f_x}$; $Y = \frac{v - v_0 \cdot z_c \cdot dy}{f_y}$; $Z = z_c$

Conclusions:

Remanufacturing has been considered as an eco-industry, demonstrating environmental and economic benefits. Damage feature inspection is a critical and step in remanufacturing, which establishes the connection between used part and process planning. However, the current inspection method for remanufacturing heavily relies on manual operations. In this study, a deep learning-based damage recognition and spatial localization method is developed. The damage recognition method is based on a Mask-RCNN model to output damage type, 2D damage segments. By mapping the 2D pixel coordinate to 3D global coordinate system, the spatial coordinate of damage is calculated. With identifying and positioning damages, further automatic repairing/remanufacturing processes can be operated based on these results.

Acknowledgements:

The authors would like to acknowledge NSERC (Grant No. RGPIN-2017-04516 and CRDPJ 537378-18) for funding this project.

References:

[1] Girshick R.: Fast R-CNN, Proceedings of the IEEE International Conference on Computer Vision, 2015, 1440-1448. https://doi:10.1109/ICCV.2015.169.

[2] Hascoët J.; Touzé S.; Rauch M.: Automated identification of defect geometry for metallic part repair by an additive manufacturing process, Welding in the World, 62(2), 2018, 229–241. https://doi.org/10.1007/s40194-017-0523-0.

[3] He K.; Gkioxari G.; Dollar P.; Girshick R.: Mask R-CNN, Proceedings of the IEEE International Conference on Computer Vision, 2017, 2980-2988. https://doi:10.1109/ICCV.2017.322.

[4] He K.; Zhang X.; Ren S.; Sun J.: Deep residual learning for image recognition, Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition; 2016, 770-778. https://doi:10.1109/CVPR.2016.90.

[5] Jiang Z.; Zhou T.; Zhang H.; Wang Y.; Cao H.; Tian G.: Reliability and cost optimization for remanufacturing process planning, Journal of Cleaner Production, 135, 2016, 1602–1610. http://doi:10.1016/j.jclepro.2015.11.037

[6] Lee C.-M.; Woo W.-S.; Roh Y.-H.: Remanufacturing: Trends and issues. International Journal of Precision Engineering and Manufacturing - Green Technology, 4(1), 2017, 113–125. http://doi:10.1007/s40684-017-0015-0.

[7] Lin T.-Y.; Dollár P.; Girshick R.; He K.; Hariharan B.; Belongie S.: Feature pyramid networks for object detection, Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 936-944. https://doi:10.1109/CVPR.2017.106.

[8] Wang H.; Jiang Z.; Zhang X.; Wang Y.; Wang Y.: A fault feature characterization-based method for remanufacturing process planning optimization, Journal of Cleaner Production, 161, 2017; 708–719. http://doi:10.1016/j.jclepro.2017.05.178.

[9] Zhang X.; Li W.; Liou F.: Damage detection and reconstruction algorithm in repairing compressor blade by direct metal deposition, The International Journal of Advanced Manufacturing Technology, 95, 2017, 2393-2404. http://doi:10.1007/s00170-017-1413-8.

[10] Zheng Y.; Liu J.; Liu Z.; Wang T.; Ahmad R.: A primitive-based 3D reconstruction method for remanufacturing, The International Journal of Advanced Manufacturing Technology, 103, 2019, 3667–3681. http://doi:10.1007/s00170-019-03824-w.