

# FABRIC DEFECT DETECTION BASED ON IMPROVED OBJECT AS POINT

Yuan He, Xin-Yue Huang and Francis Eng Hock Tay

Department of Mechanical Engineering, National University of Singapore, Singapore

## **ABSTRACT**

*In the field of fabric manufacturing, many factories still utilise the traditional manual detection method. It requires a lot of labour, resulting in high error rates and low efficiency. In this paper, we represent a real-time automated detection method based on object as point. This work makes three attributions. First, we build a fabric defects database and augment the data to training the intelligence model. Second, we provide a real-time fabric defects detection algorithm, which have potential to be applied in manufacturing. Third, we figure out CenterNet with soft NMS will improved the performance in fabric defect detection area, which is considered an NMS-free algorithm. Experiment results indicated that our lightweight network based method can effectively and efficiently detect five different fabric defects.*

## **KEYWORDS**

*Fabric defects detection, Object as point, data augmentation, Deep learning*

## **1. INTRODUCTION**

Fabric is a world widely used material in our daily life, which is made from textile fibers. The defect is a flaw on the fabric surface, occurring in fabric manufacturing process, which determines the quality of the product [1]. The fabric defects can reduce the price of the product by 45% - 65% [2]. Therefore, in the process of fabric production, “defect detection” plays a quite significant role which is a determination process of the position, class and size of the defects occurred on the surface of fabric. Traditionally, the manual inspection method is carried out on the board to contribute the instant correction of the defect, which mainly depends on the professional skills and experience of humans. Also, human inspection has very high error rates as well as the slow detection speed, because of their carelessness [1,3]. Besides of the human error occurring due to fatigue, the small-scale defects can be hard to be detected and optical illusion also has the probability to ignore defects.

A variety of algorithms including statistical analysis methods [4,5], spectrum analysis methods [6,7,8] and model-based methods [9] have been employed in the automation of fabric inspection which is based on the traditional machine vision. Among all the machine vision algorithms developed for detecting fabric defects, the method based on multi-channel Gabor filtering [8] has been considered to be one of the most successful ones. Gabor filter is able to transform the fabric images into the spectrum domain and then detect the defects by some energy criterion whose efficiency depends on the selection of filter banks [3]. However, the parameter selection of Gabor functions in all approaches depends on the dyadic decomposition, which can cause redundancy and correspondingly need excessive storage of data [10]. Also, the network structure cannot be modified since Gabor filter method, a supervised method of weight updating, is usually trained for one special database [11].

In the recent research, deep learning method has made a great success in object detection and image classification. By using advanced algorithm model such as convolutional neural network (CNN) [3], a deep learning approach can provide an unsupervised way in the fabric defects detection and classification. Compared with the traditional methods, it can obviously improve the detection and classification accuracy and speed, even make the real-time defect detection to be feasible [3, 11]. Also, the deep learning algorithm may build an effective model for all kinds of fabric images since it has highly optimized weight and can extract the required features effectively [11]. Nowadays, common object detection algorithms usually use the setting of a priori frame, that is, set a large number of a priori frames on the picture first, and the network's prediction result will adjust the prior frame to obtain the prediction frame. The a priori box greatly improves the detection ability of the network, but it will also be limited by the size of the object and detection speed. The CenterNet's detector uses keypoints estimation to find the center point and returns to other object attributes [12].

In this paper, we modify the CenterNet algorithm to meet the requirement of fabric defects detection. ResNet-50 is used as the feature extraction and the soft non-maximum suppression (Soft-NMS) is used for the expectation for improving the detection accuracy in the subsequent optimization of object detection. First, the training database for supervised learning needs to be built. In order to solve the over-fitting problem due to the lack of fabric defects images sample and improve the stability and anti-jamming ability of the neural network model, we expand the fabric image database by image cropping and rotation, add random real-time data augmentation in training process as well. After data augmentation, there are total 2,000 images in the experiment database which contains 5 kinds of fabric defects: snag, yarn thickness, hole, missing lines, fly yarn.

## **2. MATERIALS AND METHOD**

### **2.1. Fabric Defects Data**

There are more than 70 classes of fabric defects are defined in the textile manufacturing industry. Most of defects are occurring in the same direction of motion or in the direction that is perpendicular to it. According to the standards of quality, the fabric defects can be divided into two classes: surface color change and local texture irregularity [13]. According to the samples we have got, we select the five most numerous defects on pure fabric as shown in the Figure 1. Z

Normally, object detection and image classification based on deep learning rely on a large number image database. In this paper, single-sample database augmentation method, a kind of supervised database augmentation methods is used to expand the database size on the basis of the original data based on the preset data change rules. When augmenting a sample image, all operations are performed around the sample itself [14, 15].

We expand the existing fabric defects database by cropping and rotating each image. Finally, there are 2,000 fabric defect images in the database for detection model training. Besides expanding the database to improve the generalization ability, this operation also can help avoid overestimating and improve the robustness of the model and reduce the sensitivity of the model the images [16].

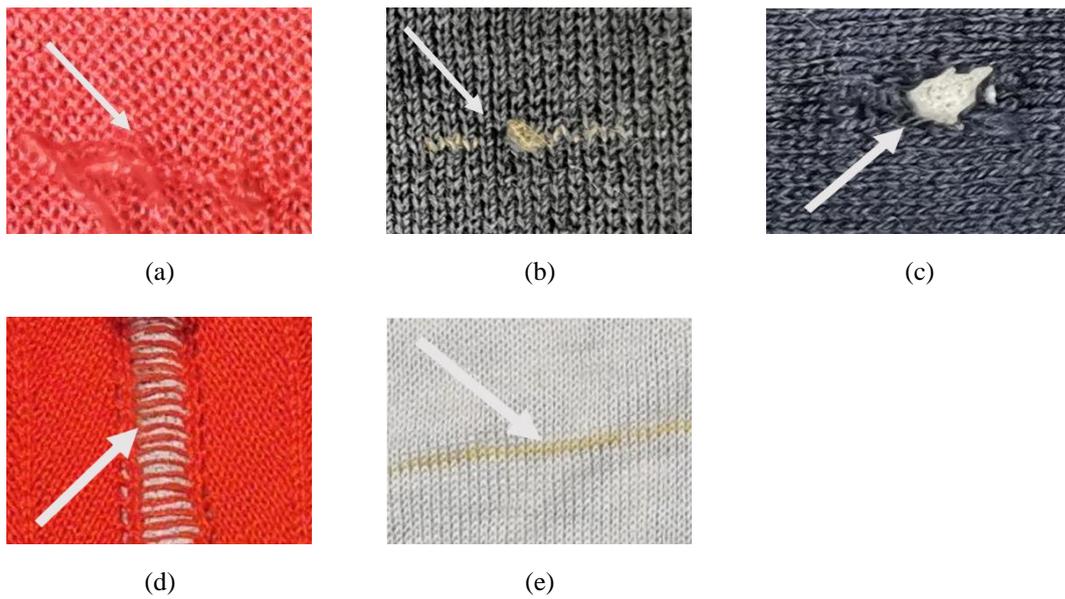


Fig 1. The examples of fabric defects. (a) snag. (b) yarn thickness. (c) hole. (d) missing lines and (f) fly yarn

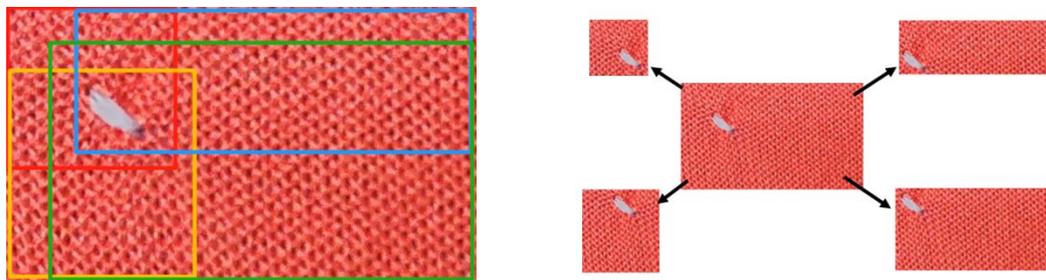


Figure 2. Image cropping and result

Image cropping means selecting a part of region in an image and discarding the rest. Using this technology, we can increase the database by selecting a large number of different regions in each original image. In this paper, as shown in Figure. 2, crop an image according to its four corners and all of the four new images should contain the defect region.

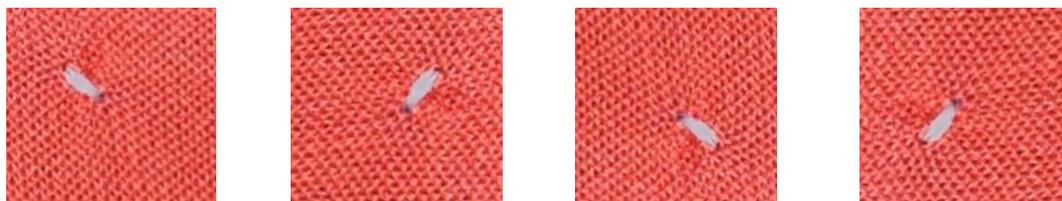


Fig 3. Original image is rotated 0°, 90°, 270° and 180°

Each fabric defects image, including the original image and cropped image, is rotated with 90°, 180° and 270° clockwise rotation in turns, the number of the expanded database is four times as the original database as shown in Figure 3. The images generated by this operation can be similar to the fabric images which are taken under different cameras' position.

There are several types of image noise and filtering technology. In this paper, we have chosen two common kinds of noise: Gaussian noise and salt-and-pepper noise to enhance the image data. Also, a Gaussian blur is applied to filter each image. Gaussian noise refers to a kind of noise whose probability density function obeys Gaussian distribution that is also called normal distribution. We added the Gaussian noise based on the standard normal distribution. Salt-and-pepper noise is also called impulse noise, which is a kind of white or black dots that appear randomly in an image. Gaussian blur is used to reduce the level of detail, the visual effect is like observing the image through a translucent screen.

The main source of Gaussian noise in digital images occurs during acquisition, due to sensor noise caused by poor lighting and/or high temperature. Salt-and-pepper noise may be caused by sudden and strong interference in the image signal, such as image sensors, transmission channels, and decoding processing. Salt-and-pepper noise is often black and white light and dark noise caused by image cutting. Therefore, the images generated by this processing can be considered as the images which are similar to the original ones that are captured under different conditions, such as noise appears under poor illumination environment, or the images taken by camera is not very clear. In this way, we can simulate different industry manufacturing conditions by this processing as well as expand the fabric defects database to increase the stability and reliability of the detection model. The results of adding Gaussian noise and salt-and-pepper noise to an original image and the result after Gaussian blur processing are shown in Figure 4.

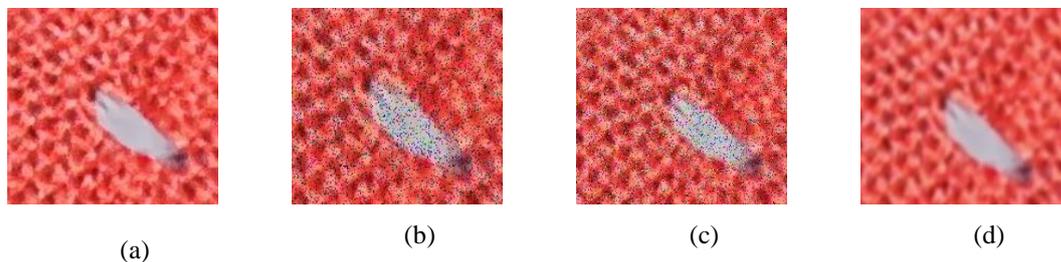


Fig 4. (a) Original image, (b) image with Gaussian noise, (c) image with salt-and-pepper noise and (d) image after Gaussian blur processing.

## 2.2. CenterNet

For object detection, it is common to set large number of priori bounding boxes first. One-stage methods use anchor to set a lot of bounding boxes in the image and score each box directly. Two-stage methods will recalculate the feature maps in those boxes. These sliding window-based detection methods need to enumerate various potential object position and sizes, which leads to the heavy computational cost and time. To address the problem, CenterNet method uses bounding box's center point to represent the object, and then on the basis of the point, other attributes of the object such as size and dimension are regressed. In this way, object detection has become a standard keypoint estimation problem, which greatly improved the detection efficiency [12]. Resnet-50 and the deconvolutional module compose the backbone network together for higher resolution output (final stride=4). The whole framework of the CenterNet used in this paper is shown in Figure 4.

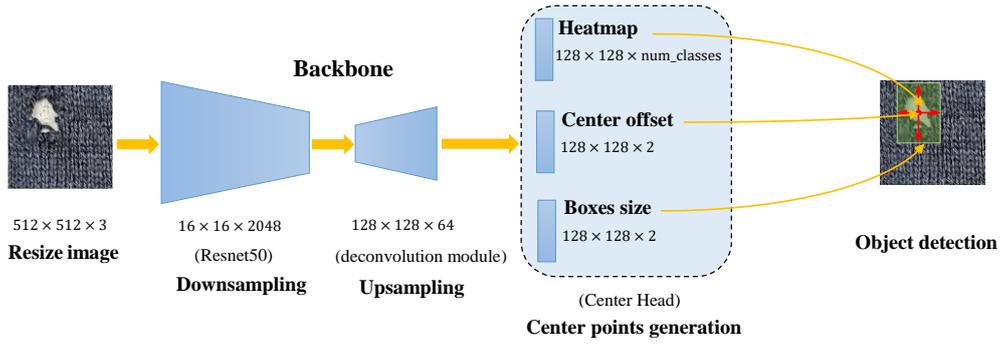


Fig 5. An illustration of CenterNet structure

There are three types of backbone neural networks for feature extraction used in the CenterNet method, normally it can be Hourglass Network, DLANet or Resnet [12]. Hourglass Network is a network which uses multi-scale features for extraction by taking advantages of multiple feature maps at the same time [17]. It is suitable for different gesture recognition of the one object, which is meaningless for fabric defects detection. and the number of Hourglass Network parameters used by CenterNet is too large, there are 190 million parameters. Therefore, we have chosen the Resnet as the backbone neural networks.

### 2.2.1. Heatmap prediction

Heatmap can be used to represent the classification information [12] and each class has its own heatmap. On each heatmap, if there is a center point of the object at a certain coordinate, a keypoint (represented by a Gaussian circle) is generated at that coordinate.  $I \in R^{W \times H \times 3}$  is an input image with width  $W$  and height  $H$  from backbone network. The aim of this step is to generate the keypoint heatmap  $\hat{Y} \in [0,1]^{\frac{W}{R} \times \frac{H}{R} \times C}$ , where  $R$  is output stride=4,  $C$  is the number of object classes and keypoint heatmap value  $\hat{Y}$  is a measure of detection confidence of each class. Therefore, we can get whether there is object in the keypoint, its class and the probability of belonging to each class. Output shape is  $(128,128,num\_classes)$ .

### 2.2.2. Points to Boxes

Center offset calculation is necessary since each keypoint located on top-left corner is responsible for predicting an object whose center is in the bottom-right corner. To extract the peaks in the heatmap for each class separately, we detect all responses whose value is greater or equal to its 8-connected neighbors and keep the top 100 peaks according to detection confidence through adapting a  $3 \times 3$  max pooling operation [12]. This is similar to the NMS method in the anchor-based detection.  $\hat{\mathcal{P}}_c$  is a set of  $n$  detected points of class  $c$ . The location of each point is given by  $(x_i, y_i)$  and the value of  $\hat{Y}_{x_i, y_i, c}$  is the detection confidence of this point in class  $c$ . Points coordinates contribute to determine the location of bounding boxes:

$$(\hat{x}_i + \delta\hat{x}_i - \hat{w}_i/2, \hat{y}_i + \delta\hat{y}_i - \hat{h}_i/2, \hat{x}_i + \delta\hat{x}_i + \hat{w}_i/2, \hat{y}_i + \delta\hat{y}_i + \hat{h}_i/2), \quad (1)$$

where  $(\delta\hat{x}_i, \delta\hat{y}_i) = \hat{O}_{\hat{x}_i, \hat{y}_i}$  is the prediction of center offset of the top-left corner keypoint to the bottom-right corner predicted object's center and  $(\hat{w}_i, \hat{h}_i) = \hat{S}_{\hat{x}_i, \hat{y}_i}$  is the size prediction of the object corresponding to this point. Their output shape both are  $(128,128,2)$ .

### 2.3. Improvements to the Fixed Confidence Threshold

In the final stage of classifying the region proposals which have different confidence of belonging to each class (that is, the probability of belonging to each class). When the confidence of one certain class is higher than the pre-defined confidence threshold, the corresponding region will be judged as the object belonging to this class. If there are multiple classes whose confidence exceeds the predefined threshold, the highest one should be selected.

Comparing the Figure 6( a) and (b), (c) and (d), it can be found that when the object's scale is small or it has been occluded, the confidence of its detection box is relatively low. Normally, we pre-define a fixed confidence threshold, but if the value is not suitable enough, it may lead to increase of false detection. Setting the fixed threshold too high will exclude many real objects, and too low will introduce many false objects. The common approach to define the confidence threshold is to adjust the threshold value constantly and use it to perform many tests on the database. Then, calculate the average accuracy under different threshold, and take the threshold with the largest average accuracy as the final threshold of the model. However, this approach tends to adapt to the certain database, and it is not able to cover all situations, which is still not accurate enough. In this report, adaptive threshold is used to improve the decision-making ability of the model and avoiding adapting to the database meanwhile.

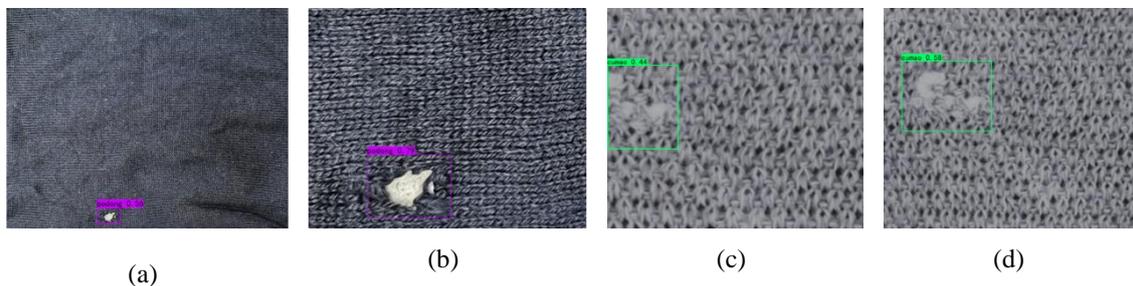


Fig 6. Influence of scale and occlusion of object on confidence.

Both in definition of fixed and adaptive threshold, the threshold setting needs to take consideration of the scores of all the objects in the database. When the detection model is well trained, the confidence of true objects and false objects in the correct detection results is usually one or two orders of magnitude different, and the confidence of true objects is usually above 0.1. Although there is a large gap between the confidence of the real objects, the fake objects will also have the probability of achieving a high confidence of more than 0.1 (as shown in the Figure 7) since some features are similar to the real objects. Therefore, the objects cannot be distinguished from the background simply by only using a fixed pre-defined threshold. According to the obvious gap existing between true and false objects detected, an adaptive threshold can be set through the confidence gap of each region to effectively distinguish between low-confidence true objects and high-confidence false objects, the accuracy can be improved in this way.

The second-order difference method can express the size of the change trend of the discrete array, so it can be used to determine the threshold in a set of confidence.

The first-order difference is the difference between two consecutive adjacent two items in the discrete function. For  $x(k)$ , its first-order difference is:

$$\Delta x(k) = y(k) = x(k + 1) - x(k) \quad (2)$$

and its second-order difference is the first-order difference of  $y(k)$ :

$$\begin{aligned} \Delta(\Delta x(k)) &= \Delta y(k) = \Delta(x(k+1) - x(k)) = \Delta x(k+1) - \Delta x(k) \\ &= y(x+2) - 2y(x+1) + y(x) \end{aligned} \quad (3)$$

Sort the maximum confidence which exceeds 0.1 of each region from largest to smallest, if all the confidence is below 0.1, there is no object. Let the function of the trend of the estimated confidence decrease from large as  $f(\cdot)$  as shown in the formula (4), then the  $C_k$  that make  $f(C_k)$  take the maximum value is used as the confidence threshold of this test image.

$$f(C_k) = \frac{(C_{k+1} - C_k) - (C_k - C_{k-1})}{C_k}, k = 2, 3, \dots, n - 1 \quad (4)$$

In this way, each image has its own confidence threshold that is based on the detection result.

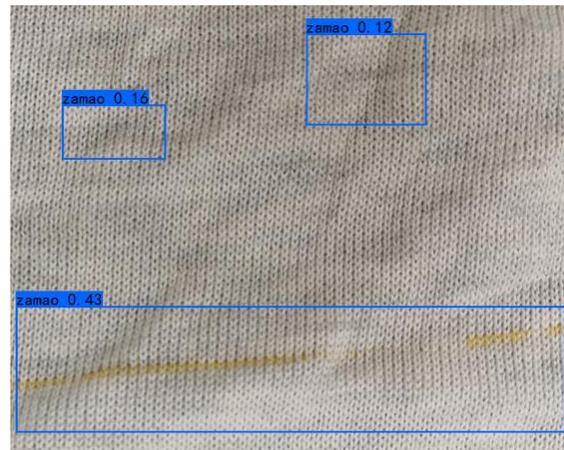


Fig 7. False positive with high confidence.

The second-order difference method can express the size of the change trend of the discrete array, so it can be used to determine the threshold in a set of confidence.

The first-order difference is the difference between two consecutive adjacent two items in the discrete function. For  $x(k)$ , its first-order difference is:

$$\Delta x(k) = y(k) = x(k+1) - x(k) \quad (2)$$

and its second-order difference is the first-order difference of  $y(k)$ :

$$\begin{aligned} \Delta(\Delta x(k)) &= \Delta y(k) = \Delta(x(k+1) - x(k)) = \Delta x(k+1) - \Delta x(k) \\ &= y(x+2) - 2y(x+1) + y(x) \end{aligned} \quad (3)$$

Sort the maximum confidence which exceeds 0.1 of each region from largest to smallest, if all the confidence is below 0.1, there is no object. Let the function of the trend of the estimated confidence decrease from large as  $f(\cdot)$  as shown in the formula (4), then the  $C_k$  that make  $f(C_k)$  take the maximum value is used as the confidence threshold of this test image.

$$(C_k) = \frac{(C_{k+1} - C_k) - (C_k - C_{k-1})}{C_k}, k = 2, 3, \dots, n - 1 \quad (4)$$

In this way, each image has its own confidence threshold that is based on the detection result.

## 2.4. Soft-NMS

As an NMS-free method, CenterNet adapts max pooling operation to search on the heatmap and keep the boxes with the highest confidence in a certain area before decoding, replacing NMS post-processing. However, in the experiments we found that this operation similar to NMS performed well for small-size objects detection. For large-size objects, CenterNet cannot always find the object center correctly, NMS is needed to delete extra boxes.

However, NMS method performs not very well when two defects are quite close to each other and it is difficult to determine a suitable value. In order to solve the problem, Soft-NMS is proposed to improve the detection accuracy [18]. For the detected boxes  $b_i$  which has a high overlap with the maximum confidence detection box  $M$ , we decay their confidence rather than completely delete them. Gaussian penalty function as shown in the formula (5) is applied that is continuous to avoid a sudden decay which can lead to the change of the ranking of all the detection boxes. When the overlap is low, the penalty increases to ensure that there is no effect of  $M$  on other boxes which have a very low overlap with it.

$$s_i = s_i e^{-\frac{iou(M,b_i)}{\sigma}}, \forall b_i \notin D \quad (5)$$

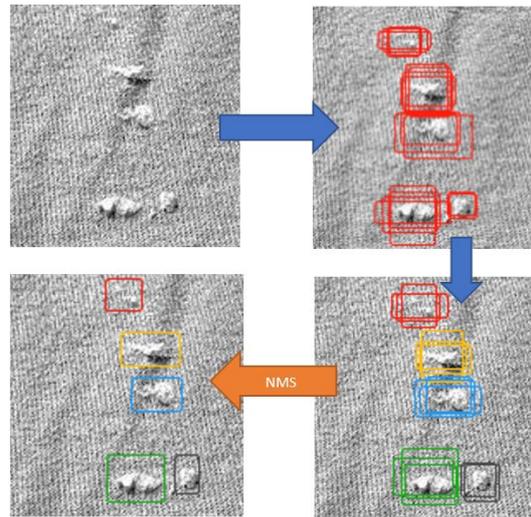


Fig 8. Overview of object detection using NMS method

## 3. EXPERIMENT AND RESULT

In this paper, we used 90% samples in the database as training and validation set in which 90% data for training. To train the model faster, we use a pre-trained model of ResNet-50. Freezing training is adapted to speed up training efficiency as well as prevent weights from being destroyed, since the feature map extracted by the backbone part of the neural network are universal. The backbone network is frozen in the first 50 epochs, and mainly the later detection and classification

part of the training are trained. The last 50 epochs are unfrozen, and the entire network is trained at the same time. The weights of model are saved in each epoch and when the loss of validation is not decrease, the training process will stop early. The learning rate is set as 0.001 initially for freezing training and 0.0001 initially after unfreezing. It will decrease by monitoring validation loss. When there is no change in validation loss for 2 epochs, learning rate will decay by half. The large learning rate can accelerate learning to reach local or global optimal solution quickly, but the loss function fluctuates greatly in the later stage and it is difficult to converge. Decaying learning rate can help be closed to optimal solution. Also, we set  $\sigma$  to 0.5 with the Gaussian weights function for Soft-NMS, the overlap represented by IoU in NMS is 0.5 as well. The threshold of mAP is 0.01.

The test environment is Asus desktop with Inter(R) Core(TM) i7-900K CPU@ 3.60GHz, 8GB RAM and Nvidia GTX 1600 Ti, the simulation software is python 3.6, Tensorflow-GPU 2.3.1 and CUDA 10.1.

Firstly, we evaluate the effect of database augmentation methods and NMS methods on detection accuracy based on mAP (mean average precision) is average precision of all class as shown in the follow formula:

$$mAP = \frac{1}{|Q_R|} \sum_{q \in Q_R} AP(q) \tag{15}$$

where AP (average precision) is the area under *precision – recall* curve. We divide the database augmentation method into two strategies: image cropping and rotation. The results of test dataset detection are shown in the Figure 9. The results of test dataset detection are shown in the Fig. 8-1. (A) is original defect database. (B) is database augmentation by cropping and rotating, the database size is 20 times as the original database. (C) is database expansion by adding noise and blur processing, the size is 80 times as the original database. From the figure below, we can find that as we expand the fabric defect image database step by step, mAP increases. Therefore, adapting database augmentation to increase database size can improve the model detection accuracy. Also, for each strategy, it is clear that NMS is necessary, mAP is quite low without NMS since there are too many extra boxes. mAP of the Soft-NMS method is almost 1% higher than NMS. Therefore, the results show that the Soft-NMS can improve the detection accuracy in CenterNet model.

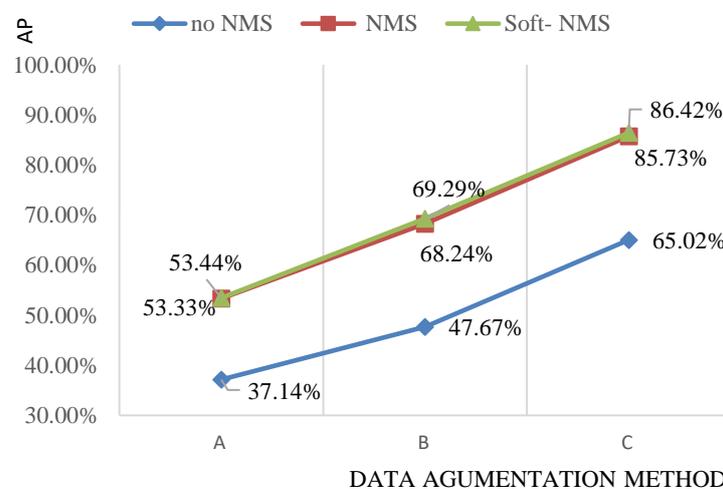


Fig. 9 The mAP of Soft-NMS and NMS with different database augmentation methods

Beside detecting on the test dataset, we also detect the training dataset, the results are shown in the Figure 10. As the database expanded step by step, the over fitting is disappearing gradually, and the precision of both training set and testing set are improved. The experimental results show that these strategies of data augmentation can effectively prevent the over fitting in the small database and improve the accuracy of defect detection.

We also train the Faster R-CNN model to detect the fabric defects, but the model is not trained very well, since it takes too much time to train, much more than that of CenterNet. The result is shown in the follow table. The detection speed of CenterNet is much higher than that of Faster R-CNN, therefore, CenterNet is more suitable for real-time detection, meeting the requirement of real manufacturing.

Table 1. Comparison between Faster R-CNN and CenterNet.

Method	Faster R-CNN	CenterNet
Speed	1.15it/s	17.85it/s
mAP	69.93%	86.42%

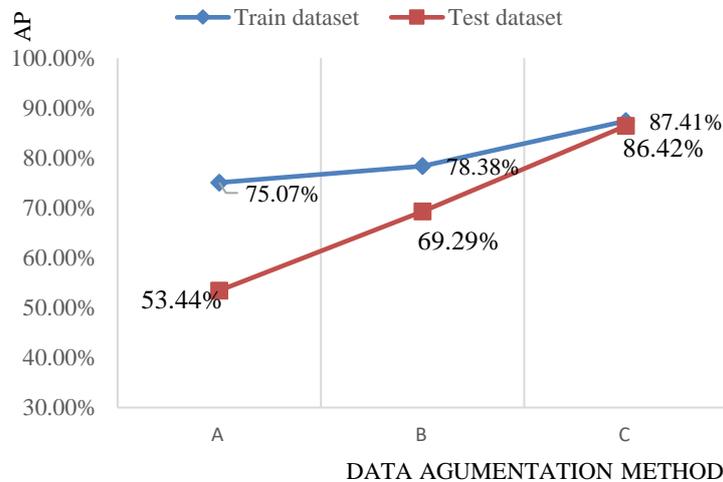


Fig. 10 The mAP of train and test dataset with different database augmentation methods.

In the experiments, we have found the calculation formula of Gaussian circle radius is inconsistent with that in mathematics. After correcting it, the mAP becomes 89.55%, more than 3% higher than the previous 86.42%. The AP of each class is shown in the Figure 7. We can see hole can be detected totally, yarn thickness and missing lines are easy to detect as well. Snag and fly yarn cannot always be detected, since their structure is diverse, and our sample size is not enough.

As shown in the Figure 8, epoch loss of training and validation process are used to monitor the training process. We use validation loss change to determine whether the model is converged. In the two curve, loss value decreases as the epoch increases, meaning the model is learning. After 50 epochs, the loss value increases suddenly, since we adapt freezing training before. In the later 50 epochs, the weights of backbone network will change meanwhile, leading to the increase of the CenterNet loss. The training epoch loss curve is smoother than the validation epoch loss curve.

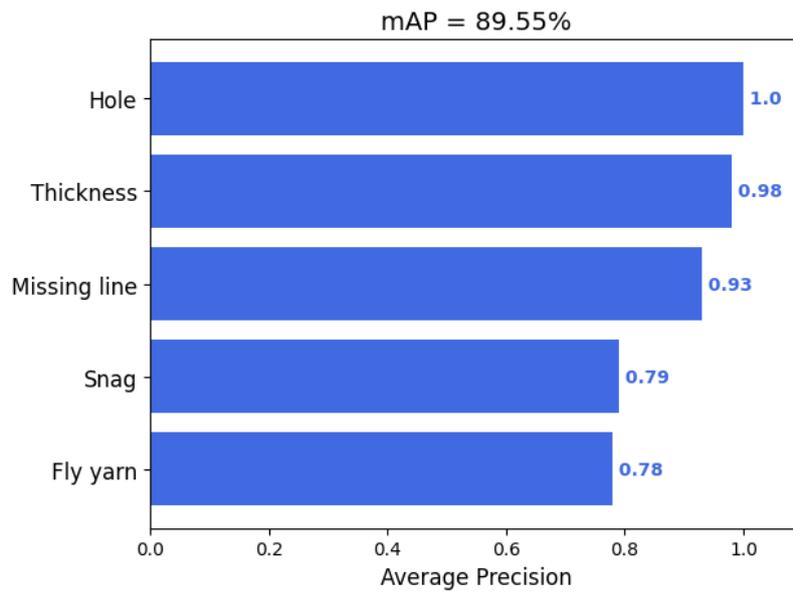


Fig 10. The AP (%) of each class in the final model.

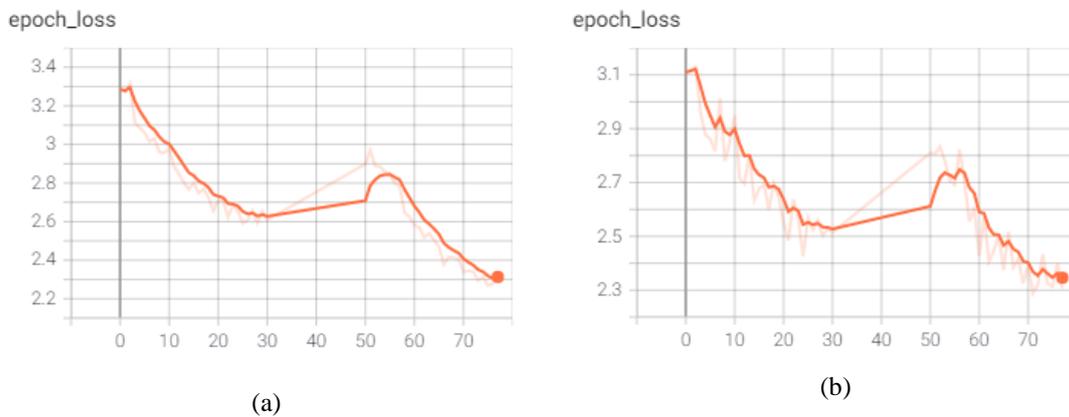


Fig 11. Epoch loss curve of (a) training process, (b) validation process

## 4. CONCLUSIONS

In this paper, we have proposed an effective fabric defect detection method based on CenterNet algorithm. Our objective is to detect the certain fabric defects on pure textile automatically in real time. To achieve this, we use a modified Resnet-50 to extract the feature map. And then three different convolutional layers are used to determine the object as point with classification information and its center offset as well as box size. After prediction, Soft-NMS is involved to improve the detection accuracy. CenterNet is a simple neural network, but it has both high accuracy and high detection speed. It takes both detection precision and speed into consideration compared with Faster R-CNN. Therefore, it is suitable for fabric detection in real manufacturing, reducing human source sharply. As a lightweight network, it may even be able to be applied to a small computing platform like mobile.

In future research, we will continue to improve the accuracy of the model and increase the types of detectable defects. Existing algorithm and our method have difficulties in the detection of complex

pattern textile. The reason is the lack of sufficient defect-free and defect samples to train reliable models. These difficulties also result in manufacturing processing cannot apply this technic to inspect the quality of fabric. We hope in the future, this real-time method can be installed in the manufacturing line to improve production efficiency and collect more data through manual assistance. This allows other art-to-date deep learning technique to be experimented and applied in fabric defect detection.

In term of improving our method, there are three research gap that should be resolved: 1) This method is difficult to detect small defects. When checking the negative sample of snag detection, we noticed if the defects area is small, this method cannot recognize. 2) If the textile structure is similar, and the color is wrong (fly yarn defects), the accuracy of this method will be relatively poor. Part of the situation stems from the inability of the method to determine the center point of the defect. 3) The Soft-NMS technology can solve part of the overlapping situation. Normally overlapping defects will not be considered to have the same center point. However, if the centers of two objects overlap in the process of extracting feature map, they will be trained as one object, leading to some prediction error. We will focus on solving these disadvantages in our future work.

## REFERENCES

- [1] Ngan, H. Y., Pang, G. K., & Yung, N. H. (2011). Automated fabric defect detection—a review. *Image and vision computing*, 29(7), 442-458.
- [2] Srinivasan, K., Dastoor, P. H., Radhakrishnaiah, P., & Jayaraman, S. (1992). FDAS: A knowledge-based framework for analysis of defects in woven textile structures. *Journal of the textile institute*, 83(3), 431-448.
- [3] Liu, Z., Liu, X., Li, C., Li, B., & Wang, B. (2018, April). Fabric defect detection based on faster R-CNN. In *Ninth International Conference on Graphic and Image Processing (ICGIP 2017)* (Vol. 10615, p. 106150A). International Society for Optics and Photonics.
- [4] YUAN, L. H., Fu, L., Yang, Y., & Miao, J. (2009). Analysis of texture feature extracted by gray level co-occurrence matrix [J]. *Journal of Computer Applications*, 29(4), 1018-1021.
- [5] Youting, Z. X. D. Z. R. (2012). Detection of Fabric Defect Based on Weighted Local Entropy [J]. *Cotton Textile Technology*, 10.
- [6] Arivazhagan, S., & Ganesan, L. (2003). Texture classification using wavelet transform. *Pattern recognition letters*, 24(9-10), 1513-1521.
- [7] ZHOU, S., ZHANG, F. S., & LI, F. C. (2013). Image Feature Extraction of Fabric Defect Based on Wavelet Transform [J]. *Journal of Qingdao University (Engineering & Technology Edition)*, 2.
- [8] Kumar, A., & Pang, G. K. (2002). Defect detection in textured materials using Gabor filters. *IEEE Transactions on industry applications*, 38(2), 425-440.
- [9] Xiaobo, Y. (2013). Fabric defect detection of statistic aberration feature based on GMRF model. *Journal of Textile Research*, 4, 026.
- [10] Mak, K. L., Peng, P., & Lau, H. Y. K. (2005, December). A real-time computer vision system for detecting defects in textile fabrics. In *2005 IEEE International Conference on Industrial Technology* (pp. 469-474). IEEE.
- [11] Jeyaraj, P. R., & Nadar, E. R. S. (2019). Computer vision for automatic detection and classification of fabric defect employing deep learning algorithm. *International Journal of Clothing Science and Technology*.
- [12] Zhou, X., Wang, D., & Krähenbühl, P. (2019). Objects as points. *arXiv preprint arXiv:1904.07850*.
- [13] Hanbay, K., Talu, M. F., & Özgüven, Ö. F. (2016). Fabric defect detection systems and methods—A systematic literature review. *Optik*, 127(24), 11960-11973.
- [14] Chaitanya, K., Karani, N., Baumgartner, C. F., Becker, A., Donati, O., & Konukoglu, E. (2019, June). Semi-supervised and task-driven data augmentation. In *International conference on information processing in medical imaging* (pp. 29-41). Springer, Cham.
- [15] Leng, B., Yu, K., & Jingyan, Q. I. N. (2017). Data augmentation for unbalanced face recognition training sets. *Neurocomputing*, 235, 10-14.

- [16] DeVries, T., & Taylor, G. W. (2017). Improved regularization of convolutional neural networks with cutout. arXiv preprint arXiv:1708.04552.
- [17] Newell, A., Yang, K., & Deng, J. (2016, October). Stacked hourglass networks for human pose estimation. In European conference on computer vision (pp. 483-499). Springer, Cham.
- [18] Bodla, N., Singh, B., Chellappa, R., & Davis, L. S. (2017). Soft-NMS--improving object detection with one line of code. In Proceedings of the IEEE international conference on computer vision (pp. 5561-5569).

## AUTHORS

**Yuan He** is a Ph.d. Candidate in National University of Singapore, Singapore. He received the B.Sc. degree in Northeastern University, P.R. China. His research interests include defects detection, computer vision and image processing.



**Dr. Francis** Tay is currently an Associate Professor with the Department of Mechanical Engineering, Faculty of Engineering, National University of Singapore. Dr. Tay is the Deputy Director (Industry) for the Centre of Intelligent Products and Manufacturing Systems, where he takes charge of research projects involving industry and the Centre.



**Xin-Yue Huang** is a Master Candidate in University of Singapore, Singapore. She received the B.Sc. degree in Sichuang University, P.R. China. Her research interests include computer vision and image processing.

