

Research Article

Automatic Fabric Defect Detection Based on an Improved YOLOv5

Rui Jin  ^{1,2} and Qiang Niu  ¹

¹School of Computer Science and Technology, China University of Mining and Technology, Xuzhou, China

²Changzhou Vocational Institute of Textile and Garment, Changzhou, China

Correspondence should be addressed to Qiang Niu; niuq@cumt.edu.cn

Received 7 August 2021; Revised 6 September 2021; Accepted 9 September 2021; Published 30 September 2021

Academic Editor: Paolo Spagnolo

Copyright © 2021 Rui Jin and Qiang Niu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Fabric defect detection is particularly remarkable because of the large textile production demand in China. Traditional manual detection method is inefficient, time-consuming, laborious, and costly. A deep learning technique is proposed in this work to perform automatic fabric defect detection by improving a YOLOv5 object detection algorithm. A teacher-student architecture is used to handle the shortage of fabric defect images. Specifically, a deep teacher network could precisely recognize fabric defects. After information distillation, a shallow student network could do the same thing in real-time with minimal performance degeneration. Moreover, multitask learning is introduced by simultaneously detecting ubiquitous and specific defects. Focal loss function and central constraints are introduced to improve the recognition performance. Evaluations are performed on the publicly available Tianchi AI and TILDA databases. Results indicate that the proposed method performs well compared with other methods and has excellent defect detection ability in the collected textile images.

1. Introduction

The textile industry is the traditional advantageous industry in China's economic development and is an important livelihood industry. The quality of textiles has a great influence on the textile industry. Fabric defects would reduce the cost and profit by 45%–65% [1]. Therefore, defect detection plays an important role in the control of textile quality. Traditional textile defect testing is usually achieved by training skilled operators with high training costs, and the manual detection efficiency is low (the detection speed is less than 20 m/min). The error and leakage rates are high due to personnel fatigue or other subjective factors. Hence, how to detect fabric defects by automatic means has become an engaging, difficult research spot in the field of textile industry and machine vision.

The core of machine-vision-based fabric defect detection is extracting the characteristics related to defects from the textile images. A detailed review of the machine-vision-based fabric defect detection methods could be found in References [2, 3]. Thomas and Cattoen [4] used the gray-scale means of image rows and columns as defect-related characteristics,

which are sensitive to illumination changes. Ye [5] presented the fuzzy inference based on image histogram statistical variables, which is robust to defects' rotation and translation. However, handling complex image texture is difficult. For complex texture images, researchers proposed methods based on edges [6], local binary patterns [7, 8], contour waves [9], and gray co-occurrence matrix [10, 11]. These methods perform well in identifying defective images but have difficulty recognizing specific fabric defects. Moreover, several researchers used the characteristics of the high-frequency parts, such as Fourier transform methods [12, 13], Gabor filter methods [7, 14], and wavelet transform methods [15, 16]. Compared with fabric defect detection in the spatial domain, it has more space-time overheads in the frequency domain.

Deep learning has been widely used in the fields of computer vision [17–19]. Researchers designed deep neural networks to realize fabric defect detection in a data-driven manner. Liu et al. [20] proposed using multistage GAN the detection of fabric defects through unsupervised data reconstruction. Hence, it could overcome the challenges of diversified fabric defects. Mei et al.

[21] introduced a multiscale convolutional denoising autoencoder to learn the reconstruction of textile images. The reconstruction errors are utilized to realize automatic defect detection. Xian et al. [22] studied the problem of metallic surface defect detection that is similar to fabric defect detection. Convolutional neural network-based segmentation is used to detect and recognize defect regions. Wei et al. [23] used faster-RCNN to detect fabric defects automatically. It achieves satisfied detection performance benefiting from faster-RCNN's strong feature engineering ability. However, faster-RCNN has large space-time complexity due to its two-stage object detection scheme. Jing et al. [24] improved YOLOv3, which is a single-stage object detection method with real-time detection performance. Then, it could better detect fabric defects.

In addition, several researchers studied the model-driven fabric defect detection methods, such as Markov random field [25], autoregression [26], and sparse dictionary [27, 28]. After effective training, these methods could identify small-region defects. However, they are vulnerable to external signals such as noise and light.

In conclusion, many researchers have proposed different methods to study how to detect fabric defects. However, detecting fabric defects is still challenging owing to many kinds of defects with large differences and uneven distributions. These problems lead to the difficulty of designing an effective system to detect and localize the fabric defects automatically. Moreover, the proposed system is required to operate faster and could be realized in an intelligent edge device platform.

According to the above requirements, a lightweight fabric defect detection method is proposed by improving YOLOv5 [29] based on the special needs of the defect detection system. It could detect and recognize special fabric defects in real time. The main contributions of this article are as follows.

A teacher-student architecture is introduced to detect fabric defects. The deep teacher network could precisely recognize fabric defects. After information distillation, a shallow student network could do the same thing in real time with minimal performance degeneration. The student network could be deployed in the edge equipment because of its low space-time overheads.

To solve the problems of many kinds of fabric defects that are difficult to be distinguished, a multitask learning strategy is proposed to detect ubiquitous and specific defects simultaneously. Such a strategy could fully utilize the complementary between ubiquitous and specific defects. Moreover, an attention mechanism is used to enhance the defect-related features.

To handle data imbalance and small-region defects better, the focal loss function [30] is employed to mitigate data imbalance. The center loss is introduced as a constraint to increase the interclass distance while reducing the intraclass distance, hence improving the recognition performance of specific defects.

The proposed method is evaluated on the publicly available Tianchi AI and TILDA databases. The results reveal

its ability to detect and recognize specific fabric defects. To verify the generalization capability of the proposed algorithm, it is tested on self-collected fabric defect images and achieves good results.

2. Related Technologies

2.1. Convolutional Neural Networks. Convolutional neural networks (CNNs) are widely used in computer vision tasks [31]. CNN is a kind of feed-forward neural network that contains convolutional computation and deep structure. It has the representation learning ability to learn structured and translation-invariant information from input images. Compared with fully connected operations, CNN has the advantage of small computational overhead. A common CNN-based computer vision system consists of the following parts:

Input layer: it performs gray processing, normalization, and data augmentation on the input images.

Convolutional layer: it performs convolutional operations in each layer to ensure the forward and backward transmission of the information. The feature map of the l th layer is derived from that of the $l-1$ th layer using the convolutional operation, as follows:

$$y^{l(i,j)} = K_i^l * x^{l(r^j)}, \quad (1)$$

where K_i^l is the weight of i th convolutional kernel in the l th layer, and $x^{l(r^j)}$ represents the j th local region being calculated in the l th layer.

Activation layer: it always follows the convolutional layer to introduce nonlinearity. Hence, the network could have better representation learning ability. Commonly used activation functions contain Sigmoid, Tanh, ReLu, and their variants. Figure 1 shows the curves of three different activation functions.

- (1) Pooling layer: it is used to subsample the feature maps to decrease computational overheads. It could also mitigate the overfitting phenomenon. Commonly used pooling functions consist of the average and maximum pooling strategies.
- (2) Output layer: it presents various structures according to different computer vision applications. For classification tasks, the SoftMax function is often used in the output layer to calculate the probability that the input belongs to each category, thus obtaining the classification results.

The five components above are used in the improved YOLOv5. They would not be introduced in detail in the following sections.

2.2. Object Detection Algorithm. Object detection is one of the essential issues in the field of computer vision. It enables the computer to discover and locate targets of interest from images automatically, such as flaws in the fabric. Deep learning-based object detection algorithms have achieved great successes recently. Commonly used methods include RCNN [32], fast-RCNN [33], faster-RCNN [34], SDD [35],

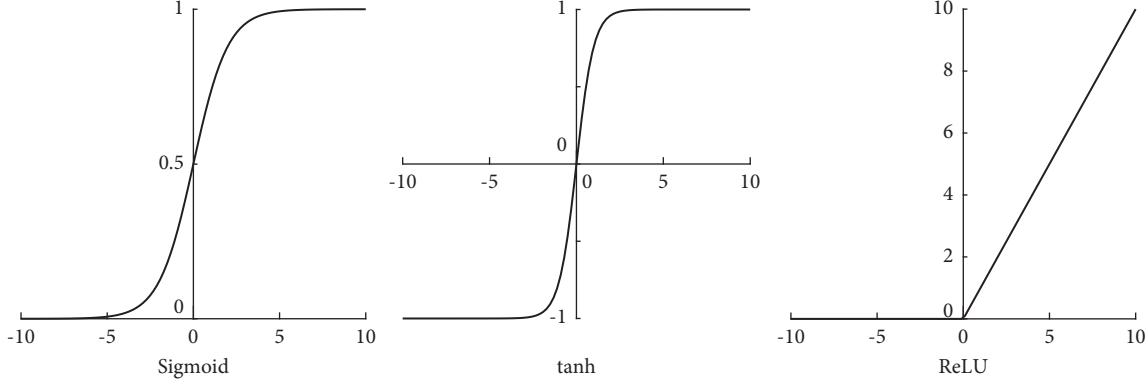


FIGURE 1: Curves of the Sigmoid, Tanh, and ReLU activation functions.

and YOLO [36]. However, the above methods have difficulty meeting the real-time requirements of the fabric defect detection system because they have high computational overheads. To balance precision and speed, a lightweight object detection network, named YOLOv5, is used in this work. The traditional YOLOv5 is improved based on the characteristics of the fabric defects, such that it can be applied to the fabric defect detection system.

Figure 2 demonstrates the structure of the traditional YOLOv5, which mainly includes Backbone, PANet, and Output. Backbone is used to perform feature engineering from input images. PANet could obtain visual features robust to scale changes due to the used pyramid structure. The positions are output, and the regions of interest are classified simultaneously. Assuming the input image size as $608 * 608 * 3$ (height * width * channels), the Output part could output three different scales of features with dimensions of $76 * 76 * 255$, $38 * 38 * 255$, and $19 * 19 * 255$. Specific details of the YOLOv5 network could be found in [29].

2.3. Attention Mechanism. The attention mechanism draws on human's selective attention characteristic. Specifically, a human being could quickly scan the global image and concentrate on the regions of interest. Then, detail information of these regions are obtained, and useless information is suppressed. Based on different applications, attention mechanism could be divided into temporal attention, spatial attention, and channel attention. Temporal attention [37] could assign different weights to sequence features. Then, the model could automatically focus on important sequence features, thus enhancing the ability to process sequence data without increasing the computational costs. Spatial attention [38] transforms the spatial information in the original image into another space and retains the key information, thereby identifying the substantial areas and increasing the attention on these areas. Channel attention [39] excavates effective features from the channel dimension and suppresses task-independent features, thus improving network performance.

For fabric defect detection, temporal attention cannot be used because the input is a static image. Considering that the

defects may occupy a small proportion in the overall image, spatial attention can be used to pay more attention to small-region defects. Moreover, channel attention is used to refine features and improve the algorithm performance. Figure 3 shows that input feature F is initially processed by max-pooling (MaxPool) and average-pooling (AvgPool). Then, channel and spatial attention realize feature transformation with the shared three-layer MLP and convolutional operation, respectively. Finally, the sigmoid activation function is used to calculate different attention weights.

3. Proposed Method

Figure 4 illustrates the flow chart of the overall algorithm: (1) training stage: fabric images after data augmentation are sent to the teacher network to detect specific fabric defects. Then, the defect-related knowledge is distilled from the teacher network to the lightweight student network. (2) Testing stage: the student network is used to detect specific fabric defects in real-time performance and with minimal performance degradation. The testing stage requires deploying the student network on the NVIDIA JETSON TX2 platform based on TensorRT, which is used to accelerate the student network.

3.1. Teacher Network Structure. The structure of the proposed teacher network is shown in Figure 5. The feature extraction part and multiscale information extraction part of the teacher network are implemented using Backbone and PANet of the YOLOv5 network. Their specific structures have been introduced in Section 2.2 and are not repeated here. Two improvements are presented to perform better fabric defect detection.

- (1) Attention enhancement mechanism: the defect areas may occupy small regions in the overall textile image. Extracting defect-related features from these small regions is still a problem, even if PANet could extract the context information. Hence, the attention enhancement mechanism is introduced to mitigate the problem. First, spatial attention is used to enhance the network's sensitivity to small defect areas. Then, channel attention is used to suppress the

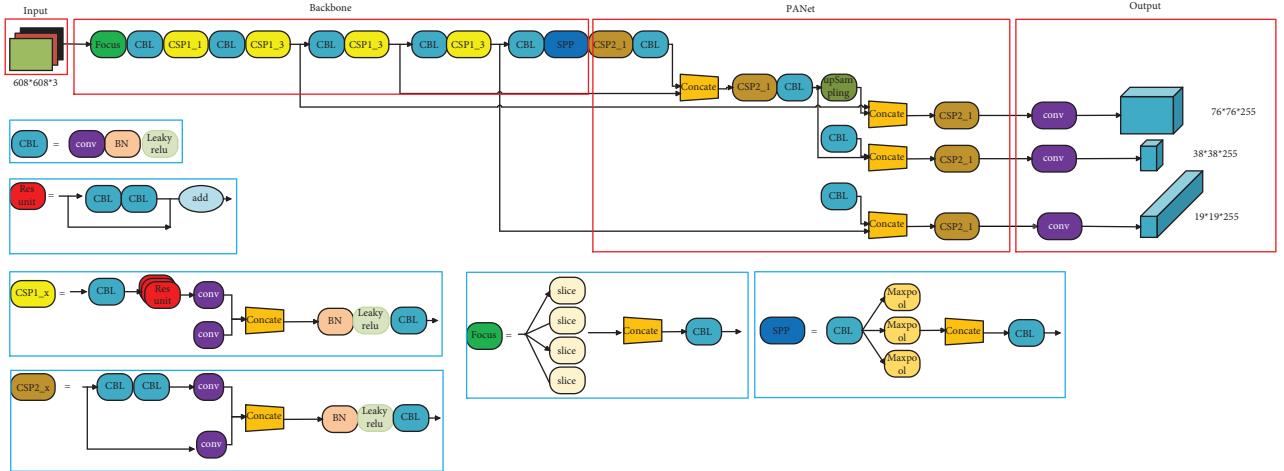


FIGURE 2: Pipeline of YOLOv5.

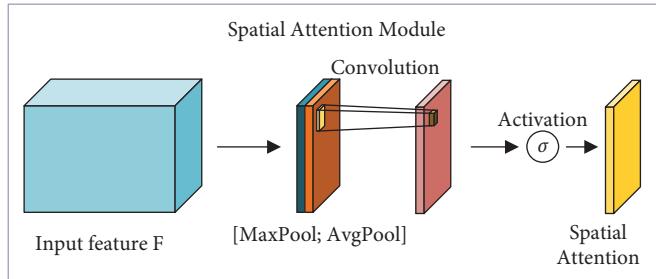
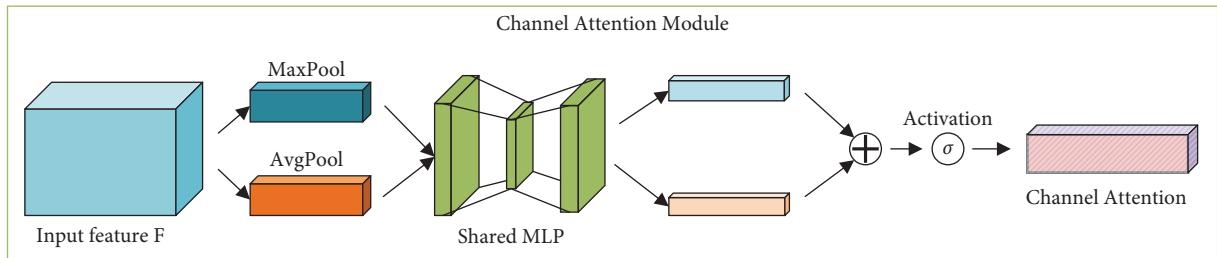


FIGURE 3: Illustration of spatial and channel attention mechanisms.

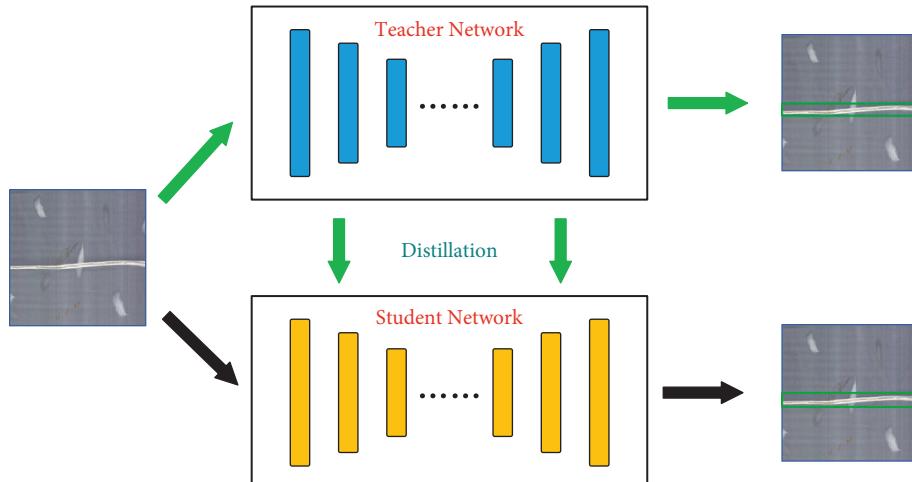


FIGURE 4: Flow chart of the overall algorithm.

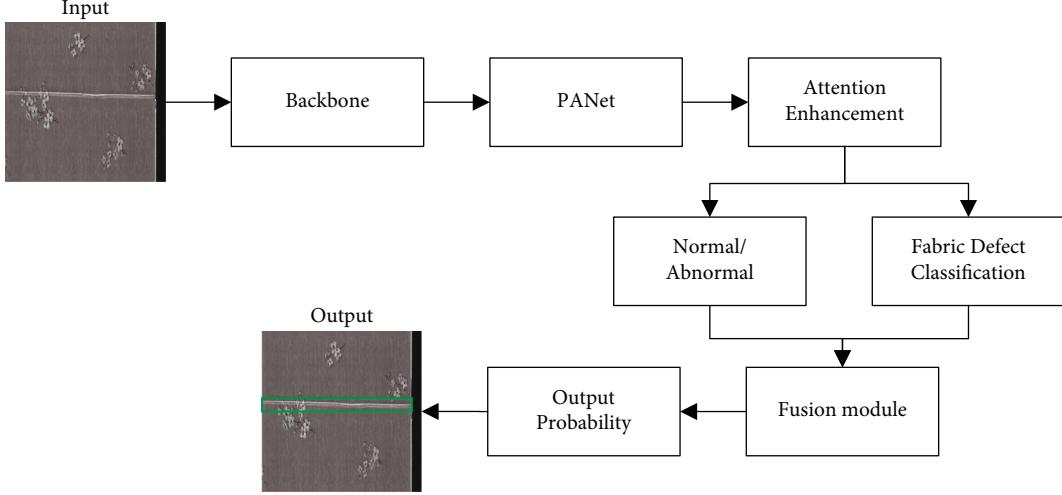


FIGURE 5: Structure of the proposed teacher network.

nondefective features, thus highlighting the defective features. Assuming that the output of PANet is F , spatial attention weight $A_s(F)$ and channel attention weights $A_c(F)$ could be calculated as follows:

$$\begin{aligned} A_c(F) &= \sigma(\text{MLP}(F_{\text{avg}}^s) + \text{MLP}(F_{\text{max}}^s)), \\ A_s(F) &= \sigma(\text{Conv}(F_{\text{avg}}^s; F_{\text{max}}^s)), \end{aligned} \quad (2)$$

where $\text{MLP}()$ represents a shared multilayer perceptron (three layers, the number of neurons is m , $m/4$, m , respectively; m represents the channel dimension of F) and $\text{Conv}()$ represents a convolution operation with the kernel size $7 * 7$. The attention enhancement mechanisms used in this work are defined as follows:

$$\begin{aligned} F' &= A_c(F) \otimes F, \\ F'' &= A_s(F') \otimes F'. \end{aligned} \quad (3)$$

- (2) Multitask learning strategy: the fabric defect detection task is usually divided into ubiquitous defect detection and specific defect recognition. Complementarity exists between these two tasks. Hence, the multitask learning strategy is introduced to utilize the complementarity fully. Specifically, two detection heads are designed to detect ubiquitous defects and recognize specific defects. A fusion model is then proposed to fuse the outputs of two detection heads to predict a more accurate defect recognition probability. Details are as follows:
 - (1) For the detection head to detect ubiquitous defects, the defective probability of the ROI with the largest defective probability is defined as P_A . Then, the normal probability of the given fabric image is defined as $P_N = 1 - P_A$.
 - (2) For the detection head to recognize specific defects, the defective probability of each ROI is defined as P_j ($j = 1, \dots, M$), where M indicates the number of ROIs.

- (3) P_N and all P_j are concatenated, and the concatenated vector is then sent into the SoftMax activation function for normalization. Then, the probability that the given fabric image belongs to a normal sample or a certain defect could be obtained.

3.2. Student Network Structure. Figure 6 exhibits the structure of the proposed student network. Different from the teacher network, the student network performs the following lightweight processing:

- (1) The backbone part is thin. Specifically, only two sets of BottleNeckCSP modules are preserved in the new backbone part. Details of the BottleNeckCSP module could be found in [29].
- (2) The PANet is removed to reduce the space-time complexity. The student network relies on the knowledge distilled from the teacher network to extract multiscale features.

The rest of the student network, including the attention enhancement, multitask learning strategy, and information fusion, are the same with the teacher network.

3.3. Loss Functions. The network is trained in a multitask learning manner, and a weighted combined loss function is presented to optimize the network. The loss functions used consist of the following sections:

- (1) The ubiquitous defect detection is termed as a binary classification problem. A cross-entropy loss function L_T is used and defined as follows:

$$L_T = -\frac{1}{N} \sum_{i=1}^N [y_i \ln p_i + (1 - y_i) \ln (1 - p_i)], \quad (4)$$

where y_i represents the sample label, p_i represents the output probability of the ubiquitous defects detection head, and N represents the number of samples.

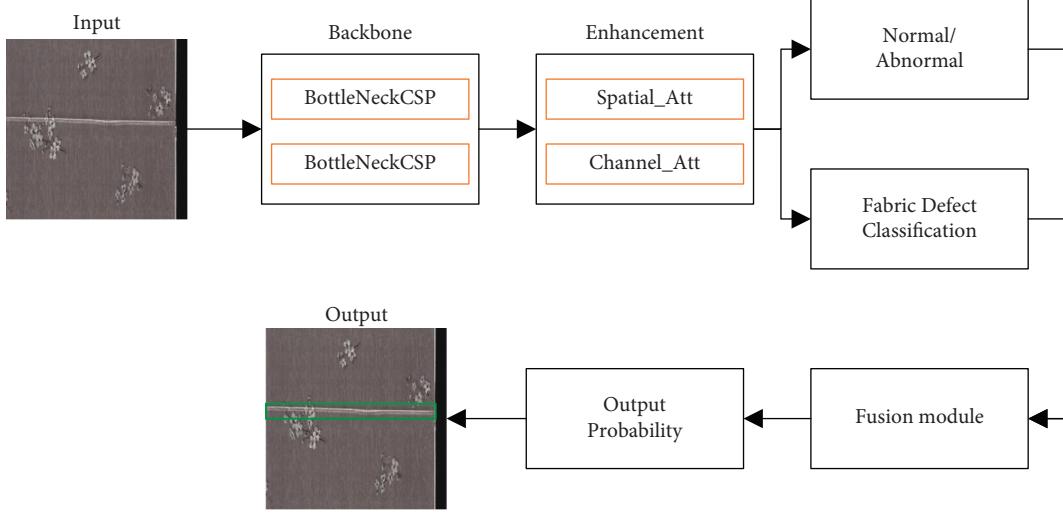


FIGURE 6: Structure of the proposed student network.

- (2) The specific defect detection is termed as a multiclass problem. A SoftMax loss function L_s is used and defined as follows:

$$L_s = - \sum_{i=1}^K \tilde{y}_i \log s_i, \quad (5)$$

where K represents the kinds of specific defects, \tilde{y}_i represents the one-hot encoding of the ground truth label, and s_i indicates the probability that the sample belongs to the i^{th} defect.

- (3) Considering the sample imbalance in the ubiquitous defect detection head, focal loss function L_F is used to mitigate the problem. L_F is defined as follows:

$$L_F = -\alpha(1-p_i)^{\gamma} \log(p_i), \quad (6)$$

where the hyperparameters α and γ are used to alleviate the imbalance problem of positive and negative samples and difficult samples, respectively.

- (4) To improve the feature discriminability in the specific defect detection head, central loss function L_C is employed to increase the interclass distances while reducing the innerclass distances of learned features. L_C is defined as follows:

$$L_C = \frac{1}{2} \sum_{i=1}^N \|x_i - c_{y_i}\|_2^2, \quad (7)$$

where x_i represents the sample encoding, and c_{y_i} is the center of the corresponding category, which x_i belongs to.

The final loss function of the proposed method is calculated in a weighted manner as follows:

$$L = w_1 L_T + w_2 L_S + w_3 L_F + w_4 L_C, \quad (8)$$

where the weights w_1, w_2, w_3 , and w_4 are set to 0.4, 0.4, 0.1, and 0.1, respectively. Settings of different weights are obtained based on the crossvalidation on the publicly available databases.

4. Experimental Results

4.1. Databases. One public database comes from the Xuelang Tianchi AI Challenge. It contains 3,331 labeled images with the rectangular locations to label the defects. The number of normal pictures is 2,163, and the number of defective pictures is 1,168. It has 22 kinds of defects, including jumps, knots, stains, puncture holes, and lacking warp. The data distribution on the database shows an unbalanced state in which the number of normal pictures is much higher than the number of defective pictures. Using the same experimental protocol as [19], the specific defect category is reintegrated into puncture hole, knots, rubbing hole, thin spinning, jumps, hanging warp, lacking warp, brushed hole, stains, and others. In experiments, 70% of the entire database is taken as the training set, and the remaining 30% are the test set. Several training samples and their labels are shown in Figure 7.

Another used public database is TILDA, a well-known fabric texture database containing eight kinds of representative fabric categories. Seven error classes and a correct class are defined according to the textile atlas analysis. Similar to [40], 300 fabric images are chosen and are divided into six categories, such as holes, scratch, knots, stain, carrying, and normal. Each class consists of 50 fabric images, and each image is resized to 256×256



FIGURE 7: Illustrations of some training samples and their labels on the Tianchi AI database.

pixels. In experiments, 70% of the entire database is taken as the training set, and the remaining 30% are the test set. Figure 8 demonstrates several samples and their labels.

4.2. Evaluation Metrics. The defect detection algorithm proposed in this work could distinguish between normal and defect images and identify specific fabric defects. Therefore, area under the ROC curve (AUC) and mean average precision (mAP) are used as metrics for evaluation. The former reflects the algorithm's ability to distinguish between normal and defective fabric images, whereas the latter reflects the algorithm's ability to recognize specific fabric defects. To calculate AUC and mAP, precision (P) and recall (R) are calculated initially, as follows:

$$P = \frac{TP}{(TP + FP)}, \quad (9)$$

$$R = \frac{TP}{(TP + FN)},$$

where TP (true positive) represents the number of samples whose labels are positive, and the actual forecasts are positive. FP (false positive) indicates the number of samples whose labels are negative, and the actual forecasts are positive. FN (false negative) represents the number of samples whose labels are positive, and the actual forecasts are negative. Based on the calculated P and R , the P-R curve could be obtained. Then, the ROC curve could be obtained. The cover area of the ROC curve is AUC.

mAP represents the mean of different APs, where AP represents the area under the P-R curve. mAP is calculated as follows:

$$mAP = \frac{1}{K} \sum_{i=1}^K (R_i - R_{i-1})P_i, \quad (10)$$

where k represents the number of categories.

4.3. Qualitative Analysis. A qualitative analysis of the proposed method is performed from three aspects: (1) the ability of the proposed teacher network to detect specific defects on public databases is evaluated, and OurNet is used for comparison; (2) the accuracy of the proposed teacher network to locate the defect areas is evaluated, and the improved YOLOv3 proposed by Jing et al. [24] is used for comparison; and (3) comparisons between the teacher and student networks are performed on self-collected fabric images to verify the generalization performance of the proposed method. Quantitative comparisons between the teacher and student networks will be introduced in the following section.

Figure 9 demonstrates the comparisons between the proposed teacher network and OurNet in detecting specific defects on the Tianchi AI database. The results show that our method successfully recognizes different defect types benefiting from the used multitask learning, focal loss function, and the center loss constraint. By contrast, OurNet fails to identify the puncture hole defects. It also mistakes the brushed hole and thin spinning defects for others and jumps defects, respectively.

Figure 10 shows the location results between the proposed teacher network and the improved YOLOv3 proposed by Jing et al. [24] on the Tianchi AI database. Types of specific defects are labeled under each subfigure for a clearer view. In the subfigure, the green box represents the real defect area, the red box is the positioning result of the proposed teacher network, and the yellow box is the positioning result of the improved YOLOv3. Figure 10 shows that the defect regions predicted by the proposed method are more accurate than those predicted by the improved YOLOv3. Such superiority may be a benefit from the strong YOLOv5 and our improvements. The improved YOLOv3 suffers from positioning small defect areas, although it could detect most defects. For example, it fails to detect the hanging warp and jump defects.

Figure 11 compares the teacher and student networks on self-collected fabric images, specifically, their

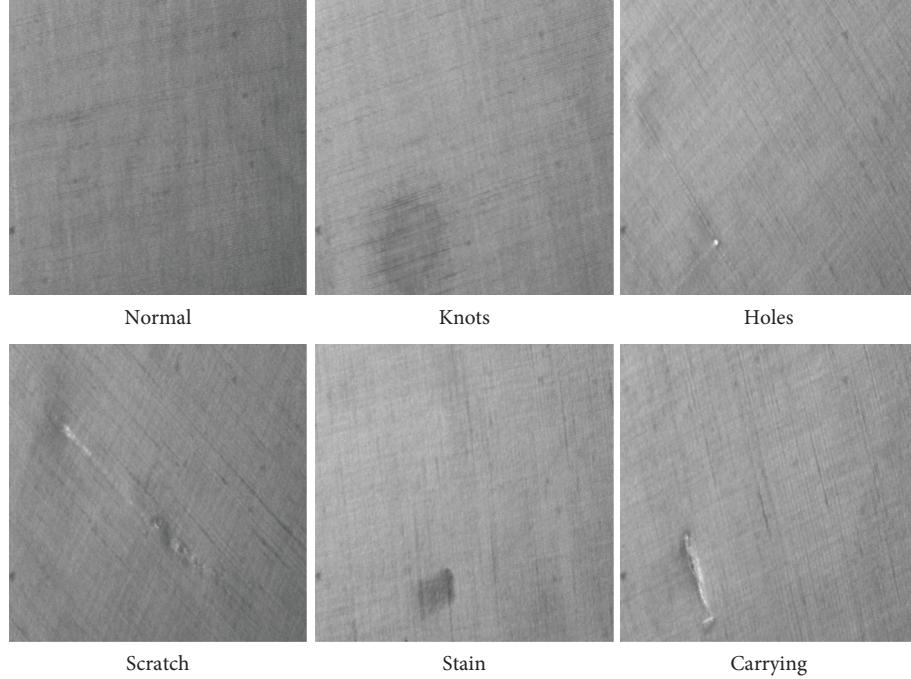


FIGURE 8: Illustrations of several samples and their labels on the TILDA database.

performance in positioning defect areas. In each sub-figure, the green box represents the real defect area, the red box is the positioning result of the teacher network, and the yellow box is the positioning result of the student network. The teacher network could more accurately identify the defect areas. The defect detection performance of the student network is slightly weaker than that of the teacher network. However, the student network has lower space-time overheads; thus, it is more suitable to be arranged for embedded systems.

4.4. Quantitative Analysis Results. An ablation study is performed on the Tianchi AI database to verify the effects of different improvement methods, including multitask learning, focal loss, and central loss constraints. The results are presented in Table 1. The ablation study of the teacher network shows that the student network has similar results.

Table 1 shows that the teacher network is degraded into traditional YOLOv5 when none of the improvements is used. Compared with the YOLOv5-based detection method, the introduced attention module could lead to an improved performance with increased AUC and mAP. Then, AUC and mAP are further improved by simultaneously detecting ubiquitous and specific defects with the proposed multitask learning strategy because of the complementarity between different tasks. Based on the multitask learning strategy, the introduction of the focal loss function and central loss constraint could further improve the defect detection results. Simultaneously using all improvements achieves the best performance on the Tianchi AI database, which verifies the effects of different improvement methods.

A quantitative comparison between the teacher and student networks is presented in Table 2. The identification times are tested on an Nvidia JETSON TX2. The table shows that the student network could still meet the needs of fabric defect detection, despite the performance degradation observed compared with the teacher network. More importantly, the identification time of the student network is approximately half of the teacher network. Its identification time guarantees the real-time performance on embedded devices.

Finally, comparisons with other mainstream methods are performed to verify the effectiveness of the proposed method. The improved YOLOv3 [24] and the pretrained deep CNN [40] are selected as the fabric defect detection algorithms. Faster-RCNN [34] and YOLOv5 [29] are selected as the universal object detection methods. The comparison results are presented in Table 3.

The above table shows that the original OurNet based on AlexNet has poor defect detection performance because it fails to handle small defect areas well. Two variants of OurNet, namely, OurNet-VGG16 and OurNet-ResNet, obtain better performance benefit from extracting better features with deeper structures. Jing et al. [24] achieves better defect detection performance using improved YOLOv3 networks. A pretrained CNN is also beneficial in boosting the defect detection performance as proposed by Jing et al. [40]. YOLOv5 and faster-RCNN achieve similar defect detection performance benefiting from their strong power in object detection. Both methods are superior to the student network proposed in this work, but the time overhead is relatively large. The proposed teacher network achieves the best fabric defect detection performance,



FIGURE 9: Comparisons between the teacher network and OurNet in detecting specific defects.

whereas the student network provides an alternative to detect fabric defects with acceptable accuracy on embedded devices.

Table 4 presents the comparisons between different methods on the TILDA database. OurNet [41] and its variants perform much better than on the Tianchi AI database because the TILDA database contains fewer categories and equal samples per category. Improved YOLOv3

[24] proposed by Jing et al. [40] achieve similar performance due to the reason discussed above. Similar to the comparisons on the Tianchi AI database, two state-of-the-art detectors, YOLOv5 [29] and faster-RCNN [34], obtained higher AUC and mAP compared with that of the proposed student network. The proposed teacher network still achieves the best defect detection performance, which verifies the accuracy of the proposed method.

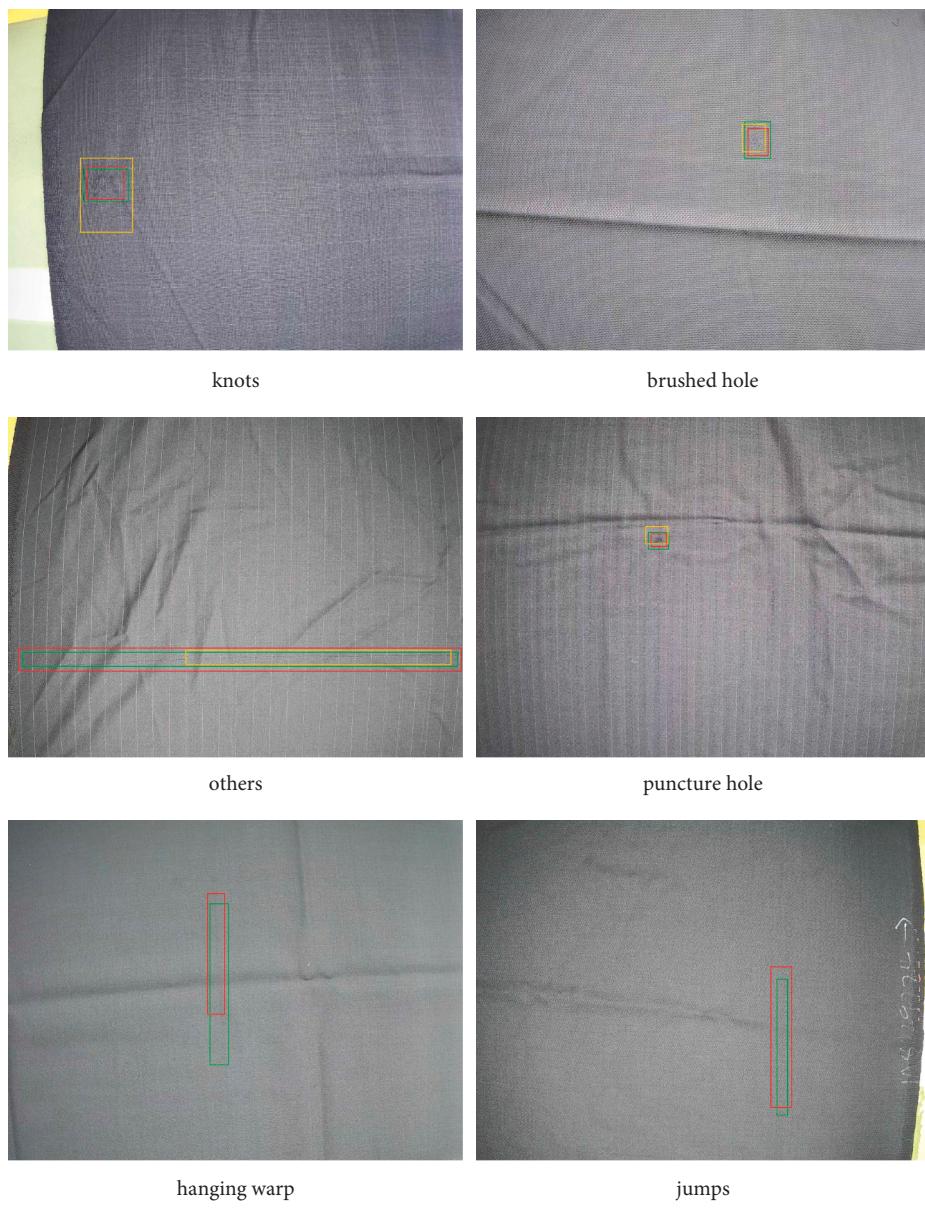


FIGURE 10: Comparisons between the teacher network and the improved YOLOv3 in positioning defect areas.

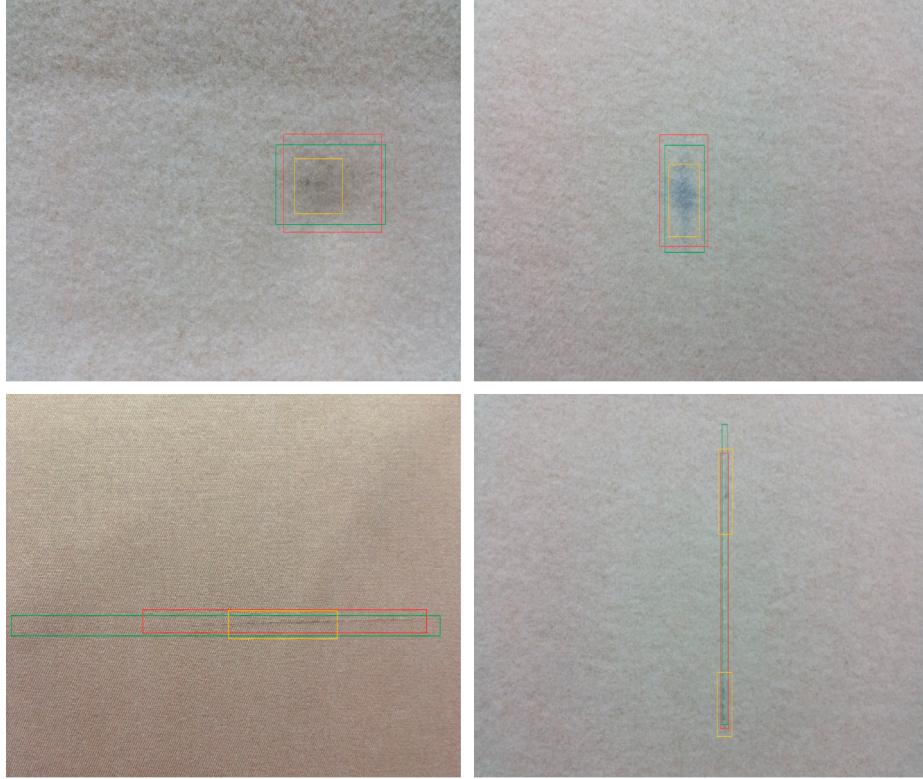


FIGURE 11: Comparisons between the teacher and student networks on self-collected fabric images.

TABLE 1: Ablation study of the teacher network on the Tianchi AI database.

Attention module	Multitask learning	Focal loss	Central loss constraints	AUC	mAP
✓				0.938	0.403
✓	✓			0.957	0.412
✓	✓			0.965	0.431
✓	✓	✓		0.971	0.441
✓	✓		✓	0.973	0.442
✓	✓	✓	✓	0.981	0.447

TABLE 2: Quantitative comparisons between the teacher and student networks on the Tianchi AI database.

	AUC	mAP	Identification times (ms)
YOLOv5	0.957	0.412	32
Teacher network	0.981	0.447	35
Student network	0.952	0.406	16

TABLE 3: Comparisons of different fabric defect detection algorithms on the Tianchi AI database.

	AUC	mAP
OurNet [41]	0.787	0.104
OurNet-VGG16	0.848	0.288
OurNet-ResNet	0.882	0.311
Improved YOLOv3 [24]	0.927	0.372
Jing et al. [40]	0.932	0.382
YOLOv5 [29]	0.957	0.412
Faster-RCNN [34]	0.956	0.413
Student network	0.952	0.406
Teacher network	0.981	0.447

TABLE 4: Comparisons of different fabric defect detection algorithms on the TILDA database.

	AUC	mAP
OurNet [41]	0.866	0.301
OurNet-VGG16	0.912	0.346
OurNet-ResNet	0.926	0.382
Jing et al. [40]	0.958	0.411
YOLOv5 [29]	0.970	0.442
Faster-RCNN [34]	0.972	0.443
Student network	0.965	0.428
Teacher network	0.988	0.451

5. Discussion and Conclusion

An automatic fabric defect detection method based on YOLOv5 is proposed because of the considerable role of fabric defect detection in the textile industry. A teacher-student architecture is used in considering the real-time requirements of the fabric defect detection. The deep teacher network could precisely detect specific fabric defects. After knowledge distillation, the shallow student network could perform fabric defects in real time with an acceptable accuracy. A multitask learning strategy is introduced to detect ubiquitous and specific defects simultaneously, and better utilize the complementarity between different tasks. Focal loss and center loss constraints are introduced for better defect detection performance. Evaluations are performed on the public databases and self-collected fabric images. Comparisons with other mainstream methods indicate that the proposed method is applicable to the automatic detection task of textile defects, which can greatly improve the accuracy and efficiency of defect detection and enhance the automation level of the textile industry.

Data Availability

The Xuelang Tianchi AI Challenge dataset is publicly available.

Conflicts of Interest

The authors declare no conflicts of interest.

Authors' Contributions

All authors have read and agreed to the published version of the manuscript.

Acknowledgments

This research was funded by the National Natural Science Foundation of China under grant no. 51674265.

References

- [1] K. Srinivasan, P. H. Dastoor, and P. Radhakrishnaiah, “FDAS: a knowledge-based framework for analysis of defects in woven textile structures,” *Journal of the Textile Institute Proceedings and Abstracts*, vol. 83, no. 3, pp. 431–448, 1990.
- [2] A. Rasheed, B. Zafar, and A. Rasheed, “Fabric defect detection using computer vision techniques: a comprehensive review,” *Mathematical Problems in Engineering*, vol. 2020, Article ID 8189403, 24 pages, 2020.
- [3] A. Latif, A. Rasheed, and U. Sajid, “Content-based image retrieval and feature extraction: a comprehensive review,” *Mathematical Problems in Engineering*, vol. 2019, Article ID 9658350, 21 pages, 2019.
- [4] T. Thomas and M. Cattoen, “Automatic inspection of simply patterned material in the textile industry,” in *Proceedings of SPIE: Society of Photo-Optical Instrumentation Engineers*, pp. 2–12, Bellingham, WA, USA, 1994.
- [5] Y. Ye, “Fabric defect detection using fuzzy inductive reasoning based on image histogram statistic variables,” in *Proceedings of the 6th International Conference on Fuzzy Systems and Knowledge Discovery*, pp. 191–194, Tianjin, China, August 2009.
- [6] X. Jia, “Fabric defect detection based on open source computer vision library OpenCV,” in *Proceedings of the 2010 2nd International Conference on Signal Processing Systems*, Dalian, China, July 2010.
- [7] J. Jing, H. Zhang, J. Wang, P. Li, and J. Jia, “Fabric defect detection using Gabor filters and defect classification based on LBP and Tamura method,” *Journal of the Textile Institute*, vol. 104, no. 1, pp. 18–27, 2013.
- [8] M. Hao, J. Junfeng, and S. Zebin, “Patterned fabric defect detection based on LBP and HOG feature,” *Journal of Electronic Measurement and Instrument*, vol. 32, no. 4, pp. 95–102, 2018.
- [9] D. Yapi, M. S. Allili, and N. Baaziz, “Automatic fabric defect detection using learning-based local textural distributions in the contourlet domain,” *IEEE Transactions on Automation Science and Engineering*, vol. 15, no. 3, pp. 1014–1026, 2017.
- [10] N. T. Deotale and T. K. Sarode, “Fabric defect detection adopting combined GLCM, gabor wavelet features and random decision forest,” *3D Research*, vol. 10, no. 1, p. 5, 2019.
- [11] M. A. Shabir, M. U. Hassan, and X. Yu, “Tyre defect detection based on GLCM and gabor filter,” in *Proceedings of the 2019 22nd International Multitopic Conference (INMIC)*, pp. 1–6, Islamabad, Pakistan, November 2019.
- [12] G. Liu and X. Zheng, “Fabric defect detection based on information entropy and frequency domain saliency,” *The Visual Computer*, vol. 37, pp. 1–14, 2020.
- [13] C. Chi-Ho Chan and G. K. H. Pang, “Fabric defect detection by Fourier analysis,” *IEEE Transactions on Industry Applications*, vol. 36, no. 5, pp. 1267–1276, 2000.
- [14] L. Jia, C. Chen, J. Liang, and Z. Hou, “Fabric defect inspection based on lattice segmentation and Gabor filtering,” *Neurocomputing*, vol. 238, pp. 84–102, 2017.
- [15] X. Yang, G. Pang, and N. Yung, “Discriminative training approaches to fabric defect classification based on wavelet transform,” *Pattern Recognition*, vol. 37, no. 5, pp. 889–899, 2004.

- [16] S. Sadaghiyanfam, "Using gray-level-co-occurrence matrix and wavelet transform for textural fabric defect detection: a comparison study," in *Proceedings of the 2018 Electric Electronics, Computer Science, Biomedical Engineering Meeting (EBBT)*, pp. 1–5, Istanbul, Turkey, April 2018.
- [17] B. Yang, G. Yan, and P. Wang, "A novel graph-based trajectory predictor with pseudo-oracle," 2021, <https://arxiv.org/abs/2002.00391>.
- [18] J. Wang, P. Fu, and R. X. Gao, "Machine vision intelligence for product defect inspection based on deep learning and Hough transform," *Journal of Manufacturing Systems*, vol. 51, pp. 52–60, 2019.
- [19] B. Yang, W. Zhan, and P. Wang, "Crossing or not? Context-based recognition of pedestrian crossing intention in the urban environment," *IEEE Transactions on Intelligent Transportation Systems*, 2021.
- [20] J. Liu, C. Wang, and H. Su, "Multistage GAN for fabric defect detection," *IEEE Transactions on Image Processing*, vol. 29, pp. 3388–3400, 2019.
- [21] S. Mei, Y. Wang, and G. Wen, "Automatic fabric defect detection with a multi-scale convolutional denoising autoencoder network model," *Sensors*, vol. 18, no. 4, p. 1064, 2018.
- [22] T. Xian, D. Zhang, and W. Ma, "Automatic metallic surface defect detection and recognition with convolutional neural networks," *Applied Sciences-Basel*, vol. 8, no. 9, 2018.
- [23] B. Wei, K. Hao, X.-S. Tang, and L. Ren, "Fabric defect detection based on faster RCNN," in *Proceedings of the International Conference on Artificial Intelligence on Textile and Apparel*, Hong Kong, China, June 2018.
- [24] J. Jing, D. Zhuo, and H. Zhang, "Fabric defect detection using the improved YOLOv3 model," *Journal of Engineered Fibers and Fabrics*, vol. 15, 2020.
- [25] P. M. Mahajan, S. R. Kolhe, and P. M. Patil, "A review of automatic fabric defect detection techniques," *Advances in Computational Research*, vol. 1, no. 2, pp. 18–29, 2009.
- [26] J. Cao, J. Zhang, Z. Wen, N. Wang, and X. Liu, "Fabric defect inspection using prior knowledge guided least squares regression," *Multimedia Tools and Applications*, vol. 76, no. 3, pp. 4141–4157, 2017.
- [27] X. Kang and E. Zhang, "A universal and adaptive fabric defect detection algorithm based on sparse dictionary learning," *IEEE Access*, vol. 8, pp. 221808–221830, 2020.
- [28] J. Zhou, D. Semenovich, A. Sowmya, and J. Wang, "Dictionary learning framework for fabric defect detection," *Journal of the Textile Institute*, vol. 105, no. 3, pp. 223–234, 2014.
- [29] A. Kuznetsova, T. Maleva, and V. Soloviev, "Detecting apples in orchards using YOLOv3 and YOLOv5 in general and close-up images," in *Proceedings of the International Symposium on Neural Networks*, Cairo, Egypt, October 2020.
- [30] T. Y. Lin, P. Goyal, and R. Girshick, "Focal loss for dense object detection," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2980–2988, Venice, Italy, October 2017.
- [31] B. Yang, W. Zhan, N. Wang, X. Liu, and J. Lv, "Counting crowds using a scale-distribution-aware network and adaptive human-shaped kernel," *Neurocomputing*, vol. 390, pp. 207–216, 2020.
- [32] S. Ren, K. He, and R. Girshick, "Object detection networks on convolutional feature maps," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 7, pp. 1476–1481, 2016.
- [33] R. Girshick, "Fast r-cnn," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1440–1448, Santiago, Chile, December 2015.
- [34] S. Ren, K. He, and R. Girshick, "Faster r-cnn: towards real-time object detection with region proposal networks," *Advances in Neural Information Processing Systems*, vol. 28, pp. 91–99, 2015.
- [35] W. Liu, D. Anguelov, D. Erhan et al., "Ssd: single shot multibox detector," in *Proceedings of the European Conference on Computer Vision*, Amsterdam, The Netherlands, October 2016.
- [36] J. Redmon, S. Divvala, R. Girshick et al., "You only look once: unified, real-time object detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 779–788, Las Vegas, NV, USA, June 2016.
- [37] A. Vaswani, N. Shazeer, and N. Parmar, "Attention is all you need," 2017, <https://arxiv.org/abs/1706.03762>.
- [38] F. Locatello, D. Weissenborn, and T. Unterthiner, "Object-centric learning with slot attention," 2020, <https://arxiv.org/abs/2006.15055>.
- [39] S. Woo, J. Park, J.-Y. Lee, and I. S. Kwon, "Cbam: convolutional block attention module," in *Proceedings of the European Conference on Computer Vision (ECCV)*, Munich, Germany, September 2018.
- [40] J. F. Jing, H. Ma, and H. H. Zhang, "Automatic fabric defect detection using a deep convolutional neural network," *Coloration Technology*, vol. 135, no. 3, pp. 213–223, 2019.
- [41] Z. Wu, Y. Zhuo, J. Li, Y. Feng, B. Han, and S. Liao, "A Fast monochromatic fabric defect Fast detection method based on convolutional neural network," *Journal of Computer-Aided Design & Computer Graphics*, vol. 30, no. 12, p. 2262, 2018.

Copyright © 2021 Rui Jin and Qiang Niu. This is an open access article distributed under the Creative Commons Attribution License (the “License”), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. Notwithstanding the ProQuest Terms and Conditions, you may use this content in accordance with the terms of the License. <https://creativecommons.org/licenses/by/4.0/>