Glass Wool Defect Detection Using an Improved YOLOv5

Yizhou Jin*, Yu Lu*, Gang Zhou, Qingjie Liu†, Yunhong Wang
State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing, China
Hangzhou Innovation Institute, Beihang University
{yizhou.jin, qingjie.liu, yhwang}@buaa.edu.cn, {luyu, zhougang}@bhhzcxyjy.wecom.work

Abstract

Glass wool defect detection is a key part of product quality assessment in the glass wool production process, yet few studies have been reported in this area. We propose a glass wool defect dataset named GWD, and also use the YOLOv5s model embedded in the GSConv and the CBAM modules for both Gap and Glueless defects in this dataset. The experimental results show that the performance of the improved YOLOv5s on the GWD dataset is superior to other compared methods and achieves a relatively good level on other publicly available datasets. Compared to the vanilla YOLOv5s, the mAP50 increased by 3.7% to 84.1%, the recall increased by 4.2% to 84.4%, and the number of parameters decreased by 0.42 MB to 6.27 MB of the improved YOLOv5s model on the GWD dataset. Speed-wisely, the improved YOLOv5s achieves a 97 FPS on RTX 2080Ti, thus making it practical to be applied in the industry of glass wool defect detection. The research on the GWD dataset is likely to contribute to breakthroughs in research on other datasets of the same type as well. The GWD dataset can be obtained by contacting us via email.

1. Introduction

Glass wool is a porous material that is widely used in construction, aerospace, and other industrial production due to its corrosion resistance, stable chemical properties, good thermal insulation, and effective sound absorption properties [1]. In chemistry, glass wool is also widely employed as a filter, packing material in gas chromatography columns, purge traps, and adsorbent beds [2]. In recent years, studies have shown that glass wool can act as catalyst support for many different reactions, and it offers a promising alternative as an adsorptive material for virus concentration [2, 3]. As one of the most common insulating materials in the world [4], glass wool plays an important role in industry and daily life.

Glass wool defect detection is the identification of where and what kind of defects are present in a piece of glass wool. Its goal is to monitor the production process of glass wool, but it also helps glass wool products meet market standards and identify inferior products before quality checks. In recent years, researchers have made a lot of efforts to improve the effectiveness of neural networks in industrial vision defect detection. Deep learning-based image processing models that can classify and locate defects more accurately have been applied in industrial production [5–7]. However, little research has been reported on the detection of defects in glass wool.

The main challenges in glass wool defect detection are the wide range of defect sizes and the high inference speed required of the model. According to the images we collected, the area of the defects in the labeled glass wool defects can range from 0.1% to 94% of the total image. In the actual production environment, glass wool is inspected on the assembly line for quality, and its residence time is very short, so the model must complete defect detection on glass wool in real time under the condition of meeting the production requirements of the assembly line.

Focusing on the study of the glass wool defect detection problem, we collected a glass wool defect dataset in an actual production environment in a factory and then improved on the YOLOv5 [8] model to propose a method applicable to glass wool defect detection. The method mainly uses the GSConv module [9] and the convolutional block attention module (CBAM) [10] to replace the native convolution module in the YOLOv5, thus improving the performance of the vanilla YOLOv5 on this problem.

In summary, our contributions are as follows: (1) we constructed a glass wool defect detection dataset named GWD, which to the best of our knowledge is the first dataset on glass wool defects; (2) we investigated the performance of YOLOv5n, YOLOv5s, and YOLOv5m models on the proposed dataset, respectively, and discovered that these models struggled to detect defects in glass wool and pin-

*These author contribute equally to this work and should be considered co-first author.
†Corresponding author.

pointed the primary problems; (3) an improved YOLOv5 model applicable to glass wool defect detection is proposed, which achieves 84.1% mAP50 and 84.4% recall on the GWD dataset, with inference speed up to 97 FPS measured under RTX 2080Ti GPU. The proposed model also achieves relatively good results on other publicly available datasets.

2. Related Work

Since the glass wool used in this study is very similar to the fabric in appearance from the images, this section focuses on the work related to the detection of fabric defects.

One-Stage Fabric Defect Detection: As a specific implementation of the one-stage detector for fabric defect detection, the one-stage algorithm inherits both the benefits and drawbacks of the one-stage detector. One-stage detectors typically attempt to classify and regress bounding boxes directly by either densely classifying a set of predefined anchor boxes or by densely searching for geometric entities. The detection speed of one-stage detectors is fast enough to satisfy online detection needs, but their detection accuracy often falls short of expectations [11].

The one-stage detectors are pioneered by YOLO, and some fabric defect detectors are based on the improved YOLO algorithms. By enhancing the YOLOv2 deep convolutional neural network's super parameters, Zhang et al. [12] refined the model, which was then applied to the detection of defects in the yarn-dyed fabric. Jing et al. [13] combined fabric defect size and k-means algorithm for dimensional clustering of target frames based on YOLOv3 to determine the number and size of previous frames; then combined low-level features with high-level information to add detection layers on feature maps of different sizes. Dlamini et al. [14] used the pre-trained model of YOLOv4 for retraining to localize defects in the textile. Zheng et al. [15] suggest a Squeeze-and-Excitation(SE)-module-based YOLOv5 (SE-YOLOv5), which adds the SE module to the YOLOv5 backbone and substitutes the ActivateOrNot (ACON) activation function for the Leaky Rectified Linear Unit (ReLU) activation function of the YOLOv5 cross stage partial (CSP).

Two-Stage Fabric Defect Detection: The majority of two-stage detectors propose class-agnostic bounding boxes with preset sizes and aspect ratios using a region proposal network (RPN), which excludes a large number of negative locations. These bounding boxes are combined into the region of interest (RoI) features, and a multilayer perceptron (MLP) is used to categorize them [11]. Although the two-stage algorithm has a high detection accuracy, it struggles to meet the real-time demands of the algorithm in practical applications due to its slow detection speed [16].

As a remarkable two-stage target detector, Faster R-CNN has been improved in several studies for fabric defect detection. Zhou et al. [17] combined several effective techniques, including Feature Pyramid Network (FPN), Deformable Convolution (DC) network, and distance IoU loss function into the vanilla Faster R-CNN to enhance the accuracy and speed of fabric defect detection. Chen et al. [18] improved the Faster R-CNN model by embedding Gabor kernels. A two-stage training method based on Genetic Algorithm (GA) and back-propagation were designed to train the modified Faster R-CNN model. Wu et al. [19] enhance the feature extraction capability of the vanilla Faster R-CNN by designing a dilated convolution module. Additionally, convolution kernel decomposition and bottleneck techniques are applied to streamline the feature extraction networks, and high-level semantic features are fused with bottom-level detail features to produce multi-scale fusion features.

By summarizing the previous studies mentioned above, the main improvement directions on fabric defect detection include: (1) improving the detection effect of the model for small-sized defects; (2) enhancing the performance of the model for images with multiple noises, strong illumination, blurring, etc.; (3) improving the performance of the model in the case of large span of defect sizes and an unbalanced distribution of defect types.

3. Method

This section describes the proposed method for producing the glass wool defect dataset (GWD) and the improved method for YOLOv5. In the improvement method of YOLOv5, we describe the YOLOv5 model, the GSConv convolution technique, and the CBAM module in detail, as well as the source of inspiration for improvement.

3.1. Dataset

The proposed glass wool defect dataset, named GWD, used in this experiment contains 964 glass wool images of size 1024×1200, which were collected from the production environment of a glass wool factory using a 2D camera. The images were first annotated by one practiced industry expert using labelImg annotation software to exploit the annotation format of the Pascal VOC dataset [20], then tested and reviewed by two experienced industry experts. The dataset contains 892 and 313 examples of each of two types of defects, Gap caused by excessive tension and Glueless due to lack of glue. These two kinds of defects are the most common defects in the production of glass wool process. The size of the Gap is generally on a small scale, and the size of the Glueless is more variable, with the area of the whole image ranging from 0.1% to 94%. The relatively small number of defects is due to the rarity of the defects themselves, and we collected nearly 5,000 glass wool images in total to produce the GWD dataset. Figure 1 shows a few sample images from the collected images and the annotated dataset. The main characteristics of the GWD dataset are weak texture and a wide range of defect size variations. Datasets often used for fabric defect studies are also characterized by a wide range of defect sizes [21, 22], indicating that the GWD dataset has similar properties to them and the study of the GWD dataset may also contribute to breakthroughs in the study of other datasets.

3.2. Improved YOLOv5

The YOLOv5 detection algorithm: The YOLOv5 proposed by Jocher et al. has the characteristics of simplicity, speed, and portability [23]. One of the typical models of the YOLOv5 series models is YOLOv5s. The main components of the traditional YOLOv5s structure are Bakbone, PANet [24], and Output. The CSPDarknet [25] is used as the backbone network for extracting features. The Focus structure and the CSP structure are applied in the CSPDarknet. The Focus structure is an image-slicing index operation that converts image spatial dimensions information into channel dimensions and can obtain a double downsampled feature map without information loss, which is believed to increase the inference speed of the model. The CSP structure borrows the design model of the CSPNet network, which allows the model to learn more features, and it solves the problem of over-computation in inference from the perspective of model structure design. PANet is a bottom-up enhancement structure based on FPN, which employs a strategy of bi-directional fusion of feature layers that enhances the transmission of localization information and enables prediction on three different scales of the feature layer. The Output part can output three different scales of features with dimensions of $80 \times 80 \times 256$, $40 \times 40 \times 512$, and $20 \times 20 \times 1024$ simultaneously, assuming the input image size is $640 \times 640 \times 3$ (height $\times$ width $\times$ channels).

The GSConv convolution technique: The GSConv aims to cooperate Standard Convolution (SC) with Deeply Separable Convolution (DSC) since the feature maps generated by DSC are still "depth-wise separated" [9]. The GSConv uses a unified blending strategy to permeate the information generated by the SC into every part of the information generated by the DSC. The approach blends the SC information into the DSC output by uniformly exchanging local feature information from different channels. The GSConv has demonstrated the ability to achieve significant accuracy gains without additional operations, and its low memory and computational requirements make it particularly suitable for edge computing devices. A new design paradigm using the GSConv, called Slim-Neck, is also proposed in the original study.

The Convolutional Block Attention Module (CBAM): It has been demonstrated that attention mechanisms can increase the effectiveness and performance of detectors, particularly light detectors [26]. The CBAM is a compact hybrid attention mechanism that improves feature maps in both spatial and channel dimensions. The channel attention module is arranged first, followed by the spatial attention module. Utilizing the inter-channel and the inter-spatial relationships of features are the goals of channel and spatial attention, respectively. The channel attention of a feature map concentrates on "what" is meaningful given an input image because each channel is a feature detector. In contrast to channel attention, spatial attention focuses on "where" is an informative part, which is complementary to channel attention. The CBAM enhances the model performance with a slight increase in computation and the number of parameters [10].

Inspired by the glass wool image in the GWD has very simple texture features, which implies that the model needs to pay more attention to the lower-level features, so we use the CBAM at the beginning of the backbone network for the first two downsampling layers to extract features. The CBAM enables the model to adaptively adjust the importance of each spatial location and channel to improve the algorithm's attention to the target information in the feature map and suppress the influence of irrelevant information. Meanwhile, for industrial defect detection, where speed and accuracy are equally important, we introduce Slim-Neck through GSConv to reduce the complexity of the model while maintaining its performance of the model.

4. Experiments

In this section, we first present the implementation details of the experiment. Afterward, in Section 4.2, we perform an extensive ablation study on design choices for this work. In Section 4.3, we compare the proposed model to existing approaches on different datasets. Finally, we launch the discussions in Section 4.4.

4.1. Implementation Details

The CBAM is used to replace the native convolution module at the beginning of the backbone network, and the GSConv is used to replace the native convolution module in the PANet. Figure 2 shows the network structure of the improved YOLOv5s. Before training, the best recall of the dataset annotation information against the default anchor frame is automatically calculated using an adaptive anchor frame. During training, 70% of all images are used for training and 30% for validation. The official pre-training model is used, and the data is pre-processed and enhanced: the image scaling operation is first performed to scale the image to 640×640, and then the mosaic data enhancement is implemented to obtain the processed image and the corresponding labels. The color enhancement, image mirroring, and center cropping are carried out after the mosaic data enhancement, and then the images are input into the neural network. Figure 3 shows a few sample images after data augment. The SGD algorithm is applied to update the pa-

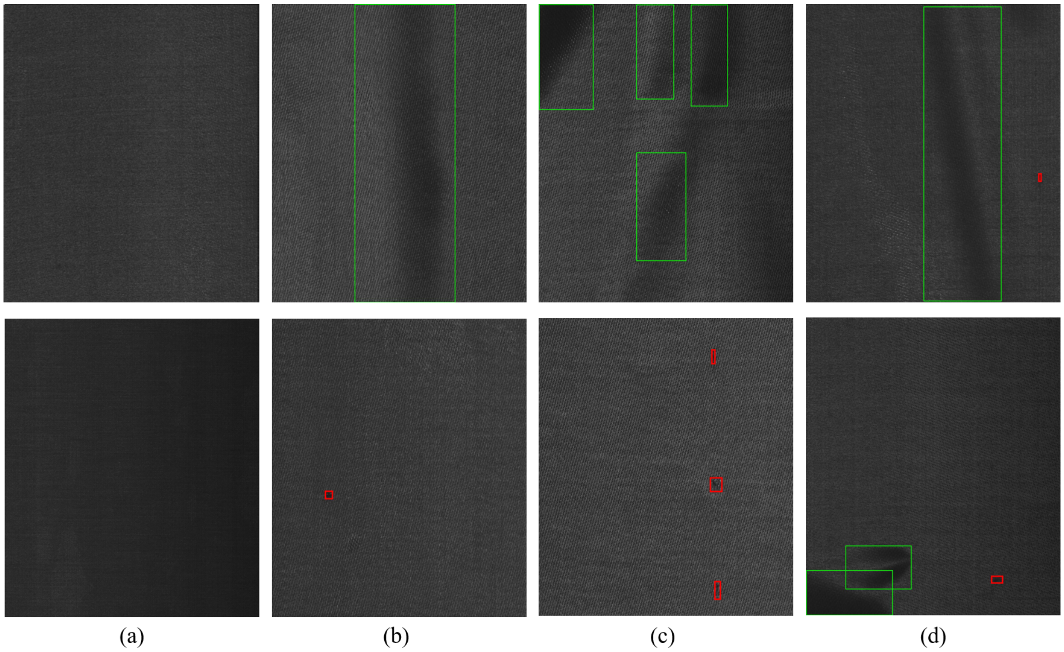

Figure 1. Sample images from the collected images and the annotated dataset, the Gap and the Glueless defects are colored **red** and green, respectively. Image contains (a) no defect (b) only one defect (c) multiple defects of the same kind (d) multiple defects of different kinds.

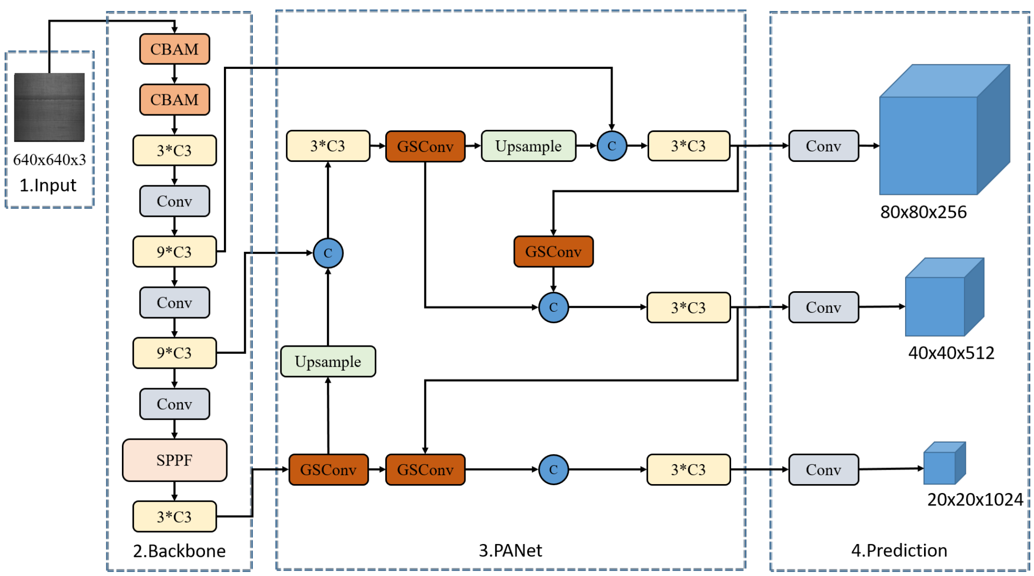

Figure 2. Network structure of the improved YOLOv5s.

rameters during training, and the batch size is 16. The hardware environment of all the experiments is Intel Xeon Gold 5218 CPU @ 2.30GHz, 32G RAM, and Four RTX2080Ti GPUs used during training. The total training epoch is 400 with an initial learning rate of zero, and the learning rate is increased to 0.001 in the first six epochs with a warm-up strategy. The model with the best performance on the validation set is reported.

4.2. Ablation Studies

We perform ablation studies based on the vanilla YOLOv5n, YOLOv5s, and YOLOv5m, respectively. The test results are displayed in Tab. 1. As illustrated in Tab. 1, the three vanilla models perform mediocrely on the GWD dataset, which we believe is mainly due to the lack of learning ability of the vanilla models for defects with large size variations. Compared to the baseline models, both the

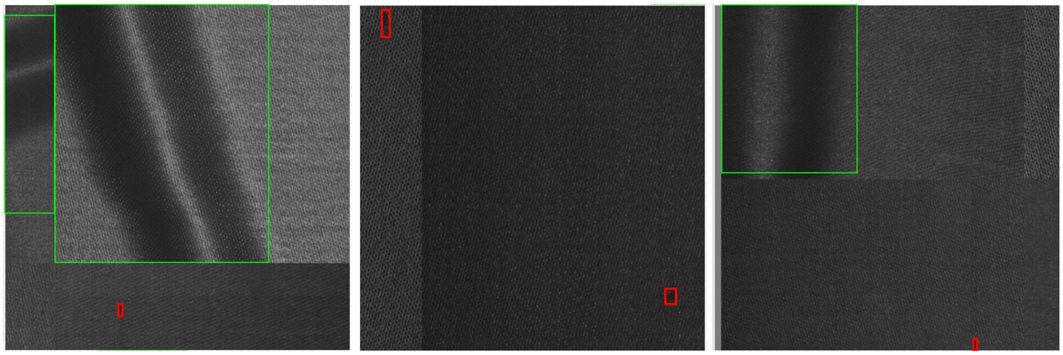

Figure 3. Sample images after data augment, the Gap and the Glueless defects are colored **red** and green, respectively.

GSConv module and the CBAM bring some performance improvements. The absolute value of mAP50 and recall increased by 2.0% and 2.1%, respectively, for the YOLOv5n; by 3.7% and 4.2%, respectively, for the YOLOv5s; and both increased by 2.4% for the YOLOv5m when the GSConv and the CBAM are used simultaneously. It is observed that for the mAP50, the simultaneous usage of the GSConv and the CBAM resulted in the greatest improvement for all three baseline models. However, for the recall, only YOLOv5s achieves the greatest improvement. As illustrated in Tab. 1, the YOLOv5s achieves the best results in both the mAP50 and recall with the simultaneous usage of the GSConv and the CBAM. Figure 4 shows a few sample results of the improved YOLOv5s.

Additionally, the parameter column indicates that using the CBAM module alone slightly increases the number of model parameters, adopting the GSConv module alone decreases the number of parameters, and adding both at once decreases the number of parameters, which consistently with the features of these two modules. After simultaneously adding both modules, YOLOv5s has the maximum reduction in the number of parameters, reaching 6.3% (0.42 MB), while YOLOv5n and YOLOv5s are 6.0% (0.1 MB) and 4.8% (0.95 MB), respectively. Tests result shows that the improved YOLOv5s model can process an image of the GWD dataset on RTX 2080Ti at a speed of 97 FPS. Although it is slightly slower than the vanilla YOLOv5s, it also fully meets the requirements of real-time defect detection in glass wool production.

Table 2 shows in detail the different performances of the model on the Gap and the Glueless defects after using the GSConv module and the CBAM on the YOLOv5s. As illustrated in Tab. 2, the model performs better on the Gap than on the Glueless before and after applying the GSConv module and the CBAM. Using the GSConv module alone improves both the Gap and the Glueless on both mAP50 and recall while using the CBAM alone does not bring this effect, as will be discussed in Section 4.4. Overall, the improved model outperformed the vanilla model on both the Gap and the Glueless defects after adopting the GSConv module and the CBAM.

4.3. Comparison with other methods

A series of comparative experiments are conducted on the proposed GWD and other datasets to compare the proposed model with various other models. The compared methods include one-stage methods: SSD [27], SSDlite [28], YOLOv7 [29], YOLOXs [30], and two-stage methods: Cascade R-CNN [31] and Faster R-CNN [32], and the results are shown in Tab. 3. Our proposed improved YOLOv5s model outperforms any other model on both mAP50 and recall and is the only method among all the compared single-stage methods that exceed 84% on both metrics.

Furthermore, the recently proposed YOLOv7, which are single-stage detection algorithms of interest, do not exactly outperform the YOLOXs and the vanilla YOLOv5s on the GWD dataset on both mAP50 and recall, which implies that for the glass wool defect detection problem, choosing the model that performs best on other publicly available object detection datasets is not necessarily the most suitable.

A benchmark dataset is built by the 2018 Tainchi Guangdong Industrial Smart Manufacturing Big Data Innovation Competition [33] and is composed of a total number of 10 aluminum surface defect categories of 3004 images with a resolution of 2560 × 1920, aiming to compare the performance of the models on images that are not similar to the fabric images. We call it the aluminum defect dataset. The example images of the aluminum defect dataset are shown in Fig. 5. The results of the comparison experiments on this dataset are shown in Tab. 4. As illustrated in Tab. 4, our proposed improved Yolov5s and Yolov7 achieved the best results on mAP50 and recall, respectively. Moreover, the improved Yolov5s model achieved the second-best results on recall. The effectiveness of the improved Yolov5s model is verified on images that are not similar to the fabric images.

Another benchmark dataset is built by the 2019 Guang-

Model	CBAM	GSConv	mAP50(%)	Recall(%)	Params(MB)
YOLOv5n			81.1	79.2	**1.68**
	✓		82.3	**82.3**	1.69
		✓	**82.7**	*83.2*	1.58
	✓	✓	*83.1*	82.0	*1.58*
YOLOv5s			80.4	80.2	6.69
	✓		81.4	78.8	6.70
		✓	**83.7**	**82.8**	**6.30**
	✓	✓	*84.1*	*84.4*	*6.27*
YOLOv5m			81.5	76.5	**19.91**
	✓		**82.1**	*82.8*	19.91
		✓	81.6	75.1	18.96
	✓	✓	*83.9*	**78.9**	*18.96*

Table 1. Ablation studies on the proposed GWD dataset with YOLOv5n, YOLOv5s, and YOLOv5m respectively. Best and second-best results, per model, are colored *red* and **blue**, respectively.

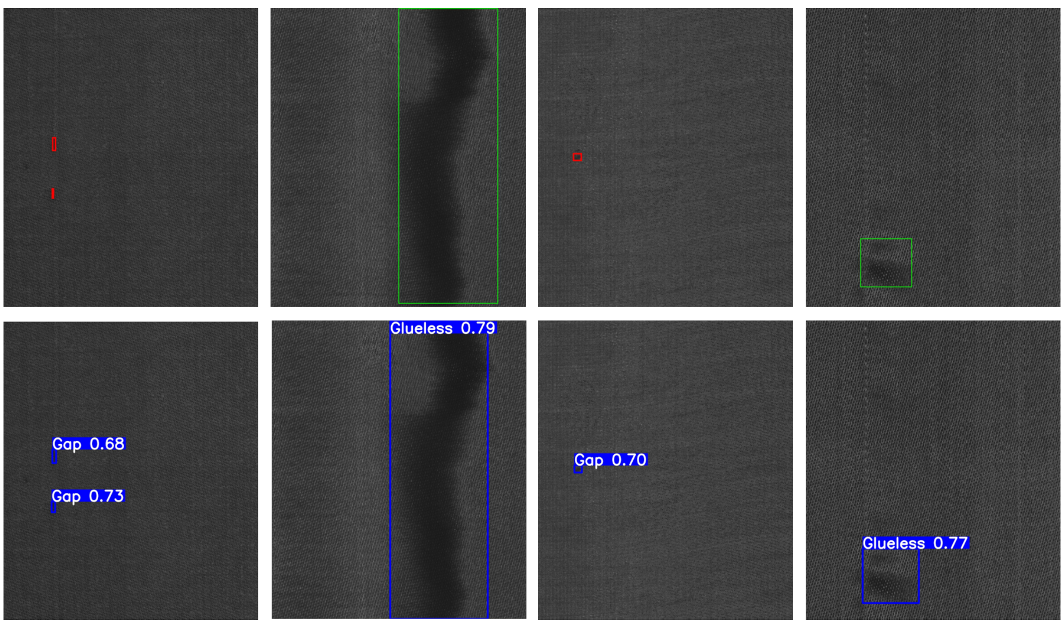

Figure 4. Sample results of the proposed improved YOLOv5s on the GWD dataset. First and second rows show the results of the ground truth labels and our proposed model, respectively.

dong Industrial Intelligent Manufacturing Innovation Competition [21] and is composed of a total number of 16 fabric defect categories of 3101 images with a resolution of 2446 × 2000. We call it the fabric defect dataset. The example images of the fabric defect dataset are shown in Fig. 6. The results of the comparison experiments on this dataset are shown in Tab. 5. As illustrated in Tab. 5, our proposed improved Yolov5s and Yolov5s achieved the best results on mAP50 and recall, respectively. The unremarkable recall performance of the improved Yolov5s model shows there is still potential for improvement of the model in the fabric defect detection study.

4.4. Discussions

Experimental results demonstrate the effectiveness of the GSConv module and the CBAM in enhancing the performance of YOLOv5s. However, as described in Section 4.2, the model performs better on the Gap than the Glueless both before and after the addition of the GSConv module and the CBAM, which we think is due to the higher number of Gap (about three times more than the Glueless) and their less size variation. In addition, with the application of the GSConv module, the model achieved the best or second-best improvement in mAP50 and recall metrics compared to the other three ablation experiments for both the Gap

Algorithm	mAP50(%)			Recall(%)		
	Gap	Glueless	All classes	Gap	Glueless	All classes
YOLOv5s	81.6	79.2	80.4	84.8	75.6	80.2
YOLOv5s-C	81.5	81.3	81.4	85.2	72.5	78.8
YOLOv5s-G	*85.1*	**82.3**	**83.7**	*88.9*	**76.7**	**82.8**
YOLOv5s-CG	**83.6**	*84.6*	*84.1*	**86.3**	*82.6*	*84.4*

Table 2. Differential performance of the model on the Gap and the Glueless defects in ablation studies. Best and second-best results, per column, are colored *red* and **blue**, respectively.

Model	Backbone	Input size	mAP50(%)	Recall(%)
Cascade R-CNN [31]	Resnet50	640×640	**82.0**	**81.3**
Faster R-CNN [32]	Resnet50	640×640	77.2	76.5
SSD [27]	VGG16	512×512	80.3	76.3
SSDlite [28]	Mobilenetv2	640×640	72.1	72.7
YOLOv7 [29]	CSPDarknet	640×640	75.7	77.4
YOLOXs [30]	CSPDarknet	640×640	78.4	79.3
YOLOv5s [8]	CSPDarknet	640×640	80.4	80.2
Ours	CSPDarknet	640×640	*84.1*	*84.4*

Table 3. Performance of our improved YOLOv5s compared with other algorithms on the GWD dataset. Best and second-best results for mAP50 and recall are colored *red* and **blue**, respectively.

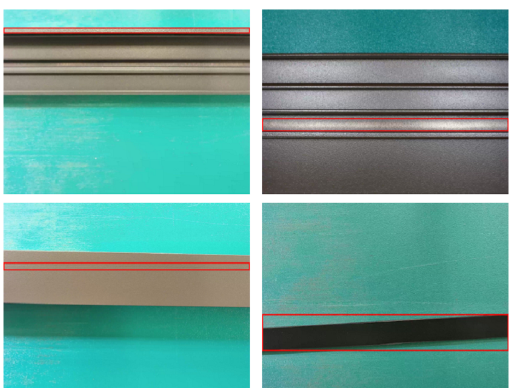

Figure 5. Sample images of the aluminum defect dataset, defects are labeled with red box.

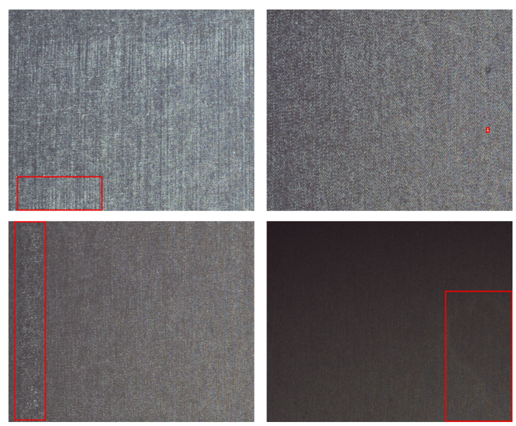

Figure 6. Sample images of the fabric defect dataset, defects are labeled with red box.

and the Glueless defects, indicating the effectiveness of the GSConv module. The addition of the CBAM to the model with the GSConv module resulted in an overall improvement in the mAP50 and recall, although there is a decrease in the mAP50 and recall for the Gap, but a greater improvement in both mAP50 and recall for the Glueless. We consider this to be caused by the fact that the spatial attention module in the CBAM module is more effective for targets with large size variations when augmented by the GSConv module.

5. Conclusion

In this work, we propose a glass wool defect dataset named GWD, which is the first glass wool defect dataset to the best of our knowledge, and achieve the precise localization and classification of glass wool defects by improving the YOLOv5 algorithm. The YOLOv5s model is chosen as the baseline model by experiments, and the GSConv and the CBAM modules are applied to perform the ablation studies. The GSConv module can better balance the accuracy and speed of the model, and for this experiment, it reduces the model parameters while improving the model performance. The CBAM module, augmented by the GSConv module,

Model	Backbone	Input size	mAP50(%)	Recall(%)
Cascade R-CNN	Resnet50	640×640	**74.0**	72.5
Faster R-CNN	Resnet50	640×640	69.6	54.5
SSD	VGG16	512×512	70.8	63.6
SSDlite	Mobilenetv2	640×640	67.7	67.9
YOLOv7	CSPDarknet	640×640	72.1	*75.8*
YOLOXs	CSPDarknet	640×640	71.3	68.6
YOLOv5s	CSPDarknet	640×640	73.8	67.4
Our	CSPDarknet	640×640	*74.9*	**72.9**

Table 4. Performance of different models on the aluminum defect dataset. Best and second-best results for mAP50 and recall are colored *red* and **blue**, respectively.

Model	Backbone	Input size	mAP50(%)	Recall(%)
Cascade R-CNN	Resnet50	640×640	**64.4**	61.7
Faster R-CNN	Resnet50	640×640	63.4	**64.4**
SSD	VGG16	512×512	61.2	60.1
SSDlite	Mobilenetv2	640×640	55.8	55.1
YOLOv7	CSPDarknet	640×640	58.4	64.1
YOLOXs	CSPDarknet	640×640	62.5	61.2
YOLOv5s	CSPDarknet	640×640	63.5	*65.4*
Our	CSPDarknet	640×640	*65.3*	61.6

Table 5. Performance of different models on the fabric defect dataset. Best and second-best results for mAP50 and recall are colored *red* and **blue**, respectively.

works better for Glueless defects with large-size variations. The comparison results with other algorithms show that the present improved YOLOv5s algorithm achieves the best results on the proposed GWD dataset, which verifies the effectiveness of the presented method. The GWD dataset has similar properties to other fabric defect detection datasets, which implies that the study on the GWD dataset can also contribute to the breakthrough of the study on other datasets of the same type.

For future work, we will continue to optimize the GWD dataset and the proposed method in this paper. On the one hand, for the GWD dataset, the number of images and defect types of the samples need to be increased. On the other hand, for the improved YOLOv5s method proposed in this paper, there is a need to increase the detection effect of the model for large-size range defects, as well as the recall metric which is very important in production, while keeping the detection speed of the model unchanged.

Acknowledgement

This work is funded by Hangzhou Science and Technology Development Plan (No.20201203B133) and the National Key Research and Development Program of China (No.2021ZD0110503).

References

[1] Zhaofeng Chen and Tianlong Liu. Development and application status of glass wool, rock wool, and ceramic wool. In *Thermal Insulation and Radiation Control Technologies for Buildings*, pages 129–161. Springer, 2022. 1

[2] Ayda Elhage, Bowen Wang, Nancy Marina, M Luisa Marin, Menandro Cruz, Anabel E Lanterna, and Juan C Scaiano. Glass wool: a novel support for heterogeneous catalysis. *Chemical Science*, 9(33):6844–6852, 2018. 1

[3] Elisabetta Lambertini, Susan K Spencer, Phillip D Bertz, Frank J Loge, Burney A Kieke, and Mark A Borchardt. Concentration of enteroviruses, adenoviruses, and noroviruses from drinking water by use of glass wool filters. *Applied and Environmental Microbiology*, 74(10):2990–2996, 2008. 1

[4] Rajeswari Ramaswamy, Juho Yliniemi, Tero Luukkonen, Ilkka Vesavaara, and Mirja Illikainen. Dissolution studies of glass wool and stone wool at alkaline ph. *Multidisciplinary Digital Publishing Institute Proceedings*, 34(1):5, 2019. 1

[5] Yu He, Kechen Song, Qinggang Meng, and Yunhui Yan. An end-to-end steel surface defect detection approach via fusing multiple hierarchical features. *IEEE Transactions on Instrumentation and Measurement*, 69(4):1493–1504, 2019. 1

[6] Ssu-Han Chen and Chia-Chun Tsai. Smd led chips defect detection using a yolov3-dense model. *Advanced engineering informatics*, 47:101255, 2021. 1

[7] Chao Li, Jun Li, Yafei Li, Lingmin He, Xiaokang Fu, and Jingjing Chen. Fabric defect detection in textile manufacturing: a survey of the state of the art. *Security and Communication Networks*, 2021:1–13, 2021. 1

[8] Xingkui Zhu, Shuchang Lyu, Xu Wang, and Qi Zhao. Tph-yolov5: Improved yolov5 based on transformer prediction head for object detection on drone-captured scenarios. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 2778–2788, 2021. 1, 7

[9] Hulin Li, Jun Li, Hanbing Wei, Zheng Liu, Zhenfei Zhan, and Qiliang Ren. Slim-neck by gsconv: A better design paradigm of detector architectures for autonomous vehicles. *arXiv preprint arXiv:2206.02424*, 2022. 1, 3

[10] Sanghyun Woo, Jongchan Park, Joon-Young Lee, and In So Kweon. Cbam: Convolutional block attention module. In *Proceedings of the European conference on computer vision (ECCV)*, pages 3–19, 2018. 1, 3

[11] Prannay Kaul, Weidi Xie, and Andrew Zisserman. Label, verify, correct: A simple few shot object detection method. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 14237–14247, 2022. 2

[12] Hong-wei Zhang, Ling-jie Zhang, Peng-fei Li, and De Gu. Yarn-dyed fabric defect detection with yolov2 based on deep convolution neural networks. In *2018 IEEE 7th data driven control and learning systems conference (DDCLS)*, pages 170–174. IEEE, 2018. 2

[13] Junfeng Jing, Dong Zhuo, Huanhuan Zhang, Yong Liang, and Min Zheng. Fabric defect detection using the improved yolov3 model. *Journal of engineered fibers and fabrics*, 15:1558925020908268, 2020. 2

[14] Sifundvolesihle Dlamini, Chih-Yuan Kao, Shun-Lian Su, and Chung-Feng Jeffrey Kuo. Development of a real-time machine vision system for functional textile fabric defect detection using a deep yolov4 model. *Textile Research Journal*, 92(5-6):675–690, 2022. 2

[15] Liaomo Zheng, Xiaojie Wang, Qi Wang, Shiyu Wang, and Xinjun Liu. A fabric defect detection method based on improved yolov5. In *2021 7th International Conference on Computer and Communications (ICCC)*, pages 620–624. IEEE, 2021. 2

[16] Yingfeng Cai, Tianyu Luan, Hongbo Gao, Hai Wang, Long Chen, Yicheng Li, Miguel Angel Sotelo, and Zhixiong Li. Yolov4-5d: An effective and efficient object detector for autonomous driving. *IEEE Transactions on Instrumentation and Measurement*, 70:1–13, 2021. 2

[17] Hao Zhou, Byunghyun Jang, Yixin Chen, and David Troendle. Exploring faster rcnn for fabric defect detection. In *2020 Third International Conference on Artificial Intelligence for Industries (AI4I)*, pages 52–55. IEEE, 2020. 2

[18] Mengqi Chen, Lingjie Yu, Chao Zhi, Runjun Sun, Shuangwu Zhu, Zhongyuan Gao, Zhenxia Ke, Mengqiu Zhu, and Yuming Zhang. Improved faster r-cnn for fabric defect detection based on gabor filter with genetic algorithm optimization. *Computers in Industry*, 134:103551, 2022. 2

[19] Jun Wu, Juan Le, Zhitao Xiao, Fang Zhang, Lei Geng, Yanbei Liu, and Wen Wang. Automatic fabric defect detection using a wide-and-light network. *Applied Intelligence*, 51(7):4945–4961, 2021. 2

[20] Antonio Torralba and Alexei A Efros. Unbiased look at dataset bias. In *CVPR 2011*, pages 1521–1528. IEEE, 2011. 2

[21] Jin Wang, Jingru Yang, Guodong Lu, Cheng Zhang, Zhiyong Yu, and Ying Yang. Adaptively fused attention module for the fabric defect detection. *Advanced Intelligent Systems*, page 2200151, 2022. 3, 6

[22] Ying Wang, Zhengyang Hao, Fang Zuo, and Shanshan Pan. A fabric defect detection system based improved yolov5 detector. In *Journal of Physics: Conference Series*, volume 2010, page 012191. IOP Publishing, 2021. 3

[23] Guijuan Lin, Keyu Liu, Xuke Xia, and Ruopeng Yan. An efficient and intelligent detection method for fabric defects based on improved yolov5. *Sensors*, 23(1):97, 2022. 3

[24] Kaixin Wang, Jun Hao Liew, Yingtian Zou, Daquan Zhou, and Jiashi Feng. Panet: Few-shot image semantic segmentation with prototype alignment. In *proceedings of the IEEE/CVF international conference on computer vision*, pages 9197–9206, 2019. 3

[25] Diganta Misra. Mish: A self regularized non-monotonic activation function. *arXiv preprint arXiv:1908.08681*, 2019. 3

[26] Qilong Wang, Banggu Wu, Pengfei Zhu, Peihua Li, Wangmeng Zuo, and Qinghua Hu. Eca-net: Efficient channel attention for deep convolutional neural networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 11534–11542, 2020. 3

[27] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C Berg. Ssd: Single shot multibox detector. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14*, pages 21–37. Springer, 2016. 5, 7

[28] Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, and Liang-Chieh Chen. Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4510–4520, 2018. 5, 7

[29] Chien-Yao Wang, Alexey Bochkovskiy, and Hong-Yuan Mark Liao. Yolov7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. *arXiv preprint arXiv:2207.02696*, 2022. 5, 7

[30] Zheng Ge, Songtao Liu, Feng Wang, Zeming Li, and Jian Sun. Yolox: Exceeding yolo series in 2021. *arXiv preprint arXiv:2107.08430*, 2021. 5, 7

[31] Zhaowei Cai and Nuno Vasconcelos. Cascade r-cnn: Delving into high quality object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 6154–6162, 2018. 5, 7

[32] Ross Girshick. Fast r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pages 1440–1448, 2015. 5, 7

[33] Wen Zhang, Shibao Sun, and Huanjing Yang. Research on aluminum defect classification algorithm based on deep learning with attention mechanism. *Frontiers in Computing and Intelligent Systems*, 2(1):101–105, 2022. 5