



Improved Plate Defect Detection Algorithm Based on YOLOv5

Zijie Wang^{1,2}, Lan Wang^{1(✉)}, and Sihui Zheng³

¹ College of Electronics and Information Engineering, Shenzhen University, Shenzhen, China

wanglan@szu.edu.cn

² Research Institute of Tsinghua University in Shenzhen (RITS), Shenzhen, China

³ Shenzhen International Graduate School, Tsinghua University, Shenzhen, China

Abstract. Furniture plates, being a crucial raw material in furniture manufacturing, often exhibit various defects during their production. These defects potentially compromise the quality of the finished furniture products and inflate production costs. Traditional methods for detecting plate defects face challenges, particularly in identifying less distinct features and handling surface noise, leading to suboptimal detection results. To address these limitations, this study introduces a specialized dataset named the “Furniture Plate Defect Dataset” for evaluating and improving defect detection algorithms more comprehensively. Furthermore, the study employs an enhanced version of the YOLOv5 algorithm, augmented with a small object detection head and incorporated with a Convolutional Block Attention Module (CBAM) to specifically optimize for plate defects. Experimental results demonstrate that with extensive training and fine-tuning on the newly constructed dataset, the enhanced YOLOv5 algorithm exhibits significant improvements in defect detection in furniture plates. The upgraded algorithm is adept at accurately identifying both texture-related and shape-related defects thereby substantially improving the detection’s accuracy and robustness. In summary, the refined YOLOv5 algorithm excels in defect detection, reaching an mAP50 of 81.6%, indicating its considerable potential for application.

Keywords: Defect detection · YOLOv5 · CBAM · Small object detection

1 Introduction

Furniture sheets, as one of the essential raw materials in the furniture manufacturing industry, have a significant demand and broad application prospects in the market. However, during the manufacturing process, various defects often appear on the surface of the sheets, such as textural and shape defects. These issues directly affect the quality and lifespan of the furniture products. To ensure product quality, there is an urgent need for sheet defect detection technology in industrial production lines. However, the complexity of furniture sheet surface defects makes their definition ambiguous, and defect

The work is supported by the Shenzhen Natural Science Foundation under Grant JCYJ20200109143016563.

© ICST Institute for Computer Sciences, Social Informatics and Telecommunications Engineering 2025

Published by Springer Nature Switzerland AG 2025. All Rights Reserved

X. Chen et al. (Eds.): IoTaaS 2023, LNICST 585, pp. 371–384, 2025.

https://doi.org/10.1007/978-3-031-70507-6_28

features are hard to unify. Currently, there are two primary defect detection methods applied to furniture sheets: supervised and unsupervised methods. Supervised methods [1] typically use neural networks to learn defect features, achieving high-accuracy defect detection. However, this approach requires a substantial amount of defect sample data for training, and obtaining sufficient defect data is often a challenging task in real-world scenarios. Another method is unsupervised defect detection [2], which detects differences between generated fake images and defect images through generative networks. Although unsupervised methods avoid the need for vast sample data, its image segmentation process makes it challenging to apply on ultra-high-resolution and non-uniformly sized sheet images. Moreover, its extended processing time limits its application in actual production environments.

In recent years, the application of deep learning convolutional neural networks (CNN) in defect detection has made remarkable progress. CNN is a deep learning model used for image classification and object detection, which can automatically learn image features and perform classification and detection. In the field of defect detection, CNN have been widely applied to surface defects [3], welding defects [4], crack defects [5], and various other types of defect detection tasks. This method can directly extract features from raw images and perform classification and detection without the need for manual feature extraction and classifier design.

Moreover, object detection methods could be broadly categorized based on the requirement of region proposals into two classes: two-stage algorithms [6–8] and one-stage algorithms [9–11].

Two-stage detectors, exemplified by algorithms like R-CNN [12], Fast R-CNN [13], and Faster R-CNN [14], typically comprise two steps: proposing object regions and then classifying these regions. This methodology initially generates a set of region proposals that might contain objects and then classifies these regions to specific object categories. Though these methods often deliver high accuracy, they can be computationally demanding due to the two-stage nature of the approach.

In contrast, one-stage detectors, such as YOLO [15] and SSD [16], eliminate the preliminary step of proposing object regions. Instead, they predict bounding boxes and class probabilities directly from the input images in one go. This streamlined approach not only improves efficiency, making it apt for real-time applications, but also maintains competitive accuracy. Among these, YOLO is particularly noteworthy. Over time, several iterations of YOLO have been released, each optimizing and refining the model, resulting in a detector known for its speed and accuracy.

While advancements in CNN for defect detection have been rapid, each industry's defect datasets typically possess their unique features and patterns. Therefore, when applying CNN to defect datasets from different industries, it is crucial to consider the specificities of each dataset and make the necessary modifications and optimizations to achieve optimal performance.

The defect dataset in this study pertains to defects in manufactured boards. Manufacturing defect datasets often include flaws found on various mechanical parts and workpieces, such as surface cracks and damages. These defects vary in shapes and sizes and usually appear against intricate backgrounds, necessitating specialized image enhancement techniques. Given the distinctiveness of defect datasets across industries,

CNN can't be directly replicated. When applying CNN to defect datasets of different industries, adjustments in network architecture, training strategies, image enhancement, and segmentation techniques are imperative for desired outcomes.

Addressing these challenges, this paper introduces an innovative algorithm for furniture plate defect detection, aiming to achieve efficient and accurate defect detection with a limited number of defect samples, even in complex plate image scenarios. To counteract the imbalance in defect sample data, we've integrated effective data balancing strategies. By leveraging transfer learning [17] and data augmentation techniques [18], we make the most out of the available defect samples, enhancing the algorithm's accuracy and robustness.

To bolster the research and evaluation of our algorithm, we've curated a comprehensive plate defect database that encapsulates various types and severities of defect samples. This serves as a robust foundation for future research and algorithm comparisons. The presented research, with its limited sample data training, manifests effective applications in defect detection.

In this paper, we'll primarily discuss the proposed algorithm and validate its effectiveness in furniture plate defect detection through comparative experimental results. Our research will substantially support the quality control and production efficiency of the furniture manufacturing industry. Furthermore, it will serve as a significant reference in addressing the limited sample challenges across various domains. Looking ahead, we aim to further optimize the algorithm's performance, explore more real-world applications, and continuously refine our defect database, driving the evolution and application of furniture plate defect detection technology.

The rest of the paper is structured as follows: Sect. 2 discusses related work. Section 3 delves into dataset collection and processing. Section 4 presents the research methodology. Section 5 presents the training process and results. Finally, Sect. 6 summarizes the conclusions of this paper.

2 Related Work

In the domain of defect detection, early methodologies relied heavily on image processing algorithms. These algorithms worked by extracting surface features like textures and shapes from the plates and subsequently used combined rules and thresholds to detect defects. Commonly adopted methods included edge detection, texture analysis, and morphological processing.

Luo et al. [19] proposed an image binarization optimization algorithm based on a local thresholding method, specifically targeting the non-uniform background issues in wood defect images. Experimental results demonstrated that their method substantially improved wood defect image segmentation under complex backgrounds, achieving an impressive 92.6% accuracy rate. Pahlberg et al. [20] and colleagues explored the fusion of two feature detection methods to automatically match and identify individual Scots pine boards. The first approach involved a block matching method utilizing normalized sum of squared differences to detect corners and subsequently match the surrounding square regions. The second approach was anchored on the Accelerated Robust Feature (SURF) matching method. The fusion of these two detection techniques significantly

boosted the recognition rate of wooden boards, pushing the matching accuracy beyond 90%.

As the wave of CNN took over, defect detection methods founded on machine learning began to gain traction. Such techniques revolved around algorithms that learned defect characteristics from a plethora of samples, facilitating automated defect detection. Compared to traditional image processing methods, machine-learning-oriented defect detection approaches augmented both detection accuracy and robustness. However, these methods are often contingent on having an extensive defect dataset for training.

Building upon VGG16, HE et al. [21] introduced the Mix-FCN, a hybrid fully convolutional neural network. While ensuring automated defect classification, this model also accelerated the detection speed, processing 50 images in just 0.368 s. Yang et al. [22] and team presented a deep extreme learning machine model that married deep learning feature extraction techniques with an extreme learning machine classification method, achieving an outstanding 96.72% accuracy rate on a wood defect dataset.

Shi et al. [23] proposed the glance multiple channel mask region convolution neural network (R-CNN). By employing neural network architecture search techniques and genetic algorithm optimization, they achieved a commendable 98.70% overall classification accuracy and 95.31% mean average precision. Xu et al. [24] modified the YOLOv5n and YOLOv5m models by incorporating the SimAM attention model and utilizing Ghost convolution to reduce model parameters. This enhancement led to a 1.5% uplift in map0.5 while concurrently slashing the parameters by over 50%.

Another frontier in defect detection is the use of reconstruction network methods. The hallmark of these methods is that they don't rely on a vast defect dataset. Instead, they necessitate training on normal data to develop a network adept at reconstructing data. Schlegl [25] was a trailblazer in utilizing deep convolutional Generative Adversarial Networks (GAN) for medical image defect detection, primarily focusing on disease prediction. The approach involved learning the features of normal images and eventually detecting defects. Building on AnoGAN, Han et al. [26] introduced enhancements by incorporating an encoder to generate latent variables, eliminating the need for extensive time-consuming prefix variable generation. Zavrtnik [27] adopted autoencoders to reconstruct images and, to address the limitations associated with autoencoders' inability to fully reconstruct anomalous regions, employed random deletion of parts of the image area.

In summary, from the rudiments of image processing to the sophistication of neural networks and reconstruction methods, the landscape of defect detection has witnessed a remarkable transformation over the years.

3 Sheet Defect Dataset

This section provides an overview of the sheet defect database composition and the associated preprocessing steps.

3.1 Image Collection

In actual furniture sheet production lines, defective sheets are rare. They can only be collected over extended periods and through active manual inspection. For this project, an

industrial digital camera, LA-CM-08K-08A-00-R, was employed for image acquisition. A collection of 313 defect data images was acquired from the production line of a certain Company's wood plate manufacturing facility. These images were categorized into three classes: edge breakage, scratches, and damaged holes. Due to the wide range of surface patterns found on furniture boards, which often include intricate textures and patterns resembling board defects, our plate defect detection model faces significant challenges. In particular, certain plate patterns and surface markings bear a striking resemblance to actual defects, making it difficult for our model to accurately distinguish between them. This presents a major obstacle in achieving reliable plate defect detection, and part of the defect pictures in the dataset are shown in Fig. 1.

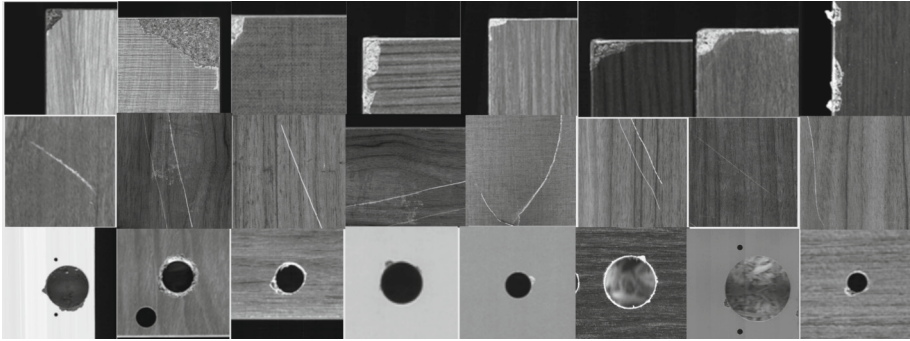


Fig. 1. Part of the sheet defect dataset image.

3.2 Dataset Processing

To streamline the training process and facilitate accurate labeling, the original sheet images collected exhibited high pixel values and were resized to a length of 1280 pixels. The Lanczos interpolation method was employed for resizing, with the shorter side downscaled proportionally. The annotation process involved utilizing the labeling tool to obtain precise annotation data.

The resulting dataset consisted of 567 instances of edge breakages, 456 instances of scratches, and 324 instances of damaged holes. While the number of hole defects was relatively lower compared to the other two categories, the unique characteristics of hole defects rendered them more distinguishable. Conversely, edge breakages and scratches were more susceptible to the interference of sheet patterns, making the identification of holes comparatively easier to identify. The distribution of defects on the training dataset and the height and width of the defects accounted for in the image are shown in Fig. 2.

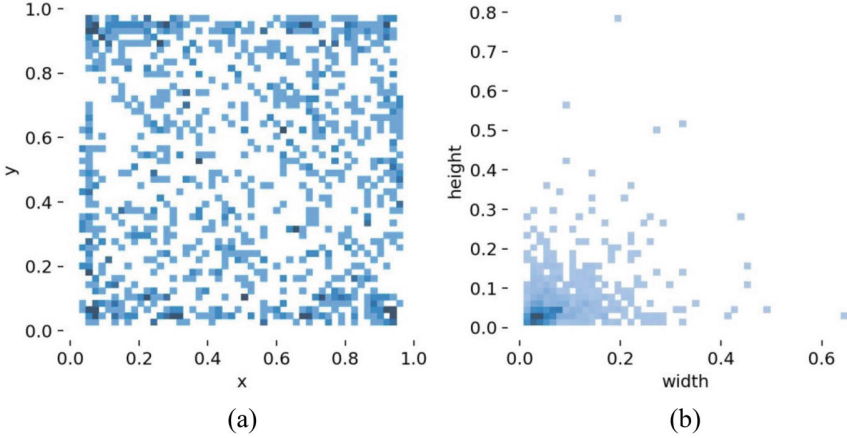


Fig. 2. (a) The distribution of defects' center point coordinates on the training dataset. (b) The distributions of the defects' width and length on the training dataset.

3.3 Incorporating Negative Samples

Owing to the scarcity of defect data, the trained network had difficulty discerning actual defects. The missed detection rate for holes stood at 0.39, and the false positive rate for scratches was as high as 0.45. Observing the excessive false positive rates for defects like holes, 210 non-defective sheet samples were added to the dataset. This intervention reduced the missed detection rate for holes to 0.24 and the false positive rate for scratches to 0.22. This shows that when training, the model learns to distinguish objects of interest from the background and assign them appropriate class labels. In the inference stage, the model can better distinguish the detected target objects from the background and provide corresponding detection boxes and category labels.

4 Network Method

4.1 Improved YOLOv5 Structure

The primary goal of this work is to develop a network that offers enhanced identification capabilities for our plate defect dataset.

Although YOLOv7 [28] has demonstrated better performance than YOLOv5 on the COCO dataset, YOLOv5 presents stronger potential for improvement while maintaining similar detection results. Therefore, we have opted to use YOLOv5 as the baseline model for our enhancements. Addressing the stated objective, we have crafted a deep learning network based on YOLOv5, striving to elevate the recognition capabilities for the plate defect dataset. YOLOv5 is considered among the advanced object detection algorithms and is lauded for its efficiency, precision, and user-friendly training approach. Building on the YOLOv5 foundation, we have introduced CBAM [29], an attention mechanism, aiming to augment the network's feature extraction prowess and target recognition accuracy. Also, bearing in mind the presence of tiny objects within the plate

defect dataset, a small object detection head was added to the original YOLOv5, ensuring enhanced recognition and precise localization of minute defects [30]. Such refinements enable our network to aptly cater to the plate defect dataset, offering more precise and trustworthy defect identification outcomes, the overall structure of the proposed network is shown in Fig. 3.

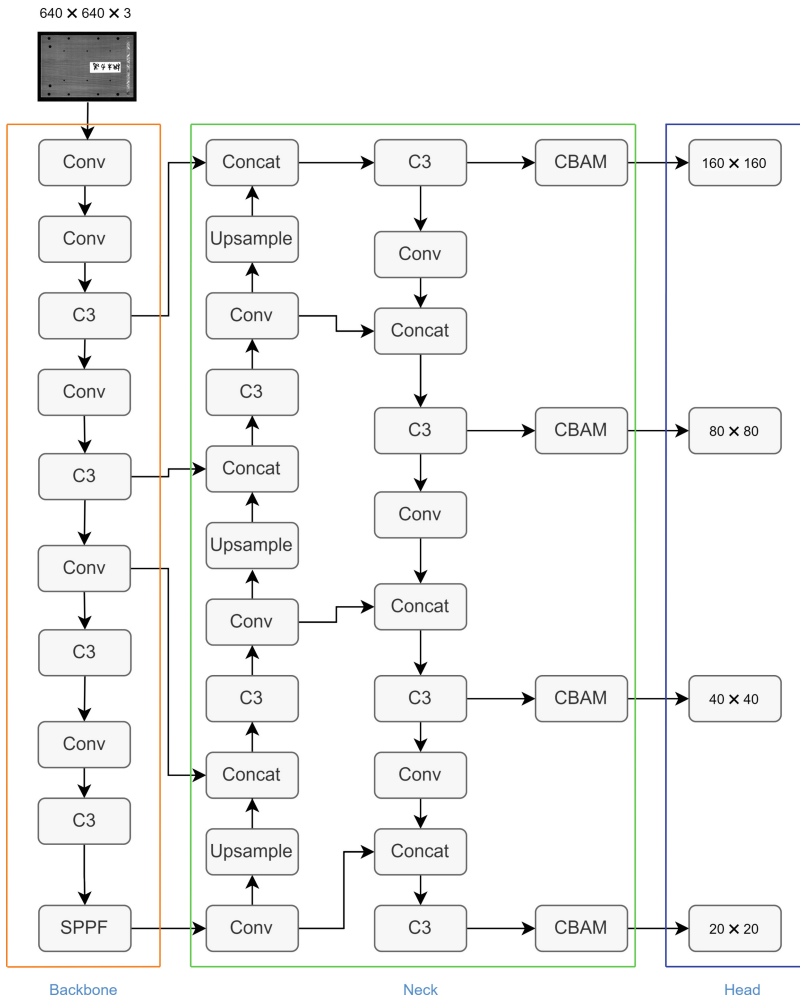


Fig. 3. The overall structure of the proposed network.

4.2 Add CBAM Module

CBAM is a lightweight attention module capable of deploying Attention both in channel and spatial dimensions. The backbone of YOLOv5 is already adept at feature extraction,

reducing its modifications aids in our transfer learning. Defects often have small sizes and subtle texture changes, requiring models that can accurately focus on and capture these detailed features. The integration of CBAM into YOLOv5 is designed to amplify the influence of limited features on detection outcomes. The schematic diagram of the CBAM structure is shown in Fig. 4.

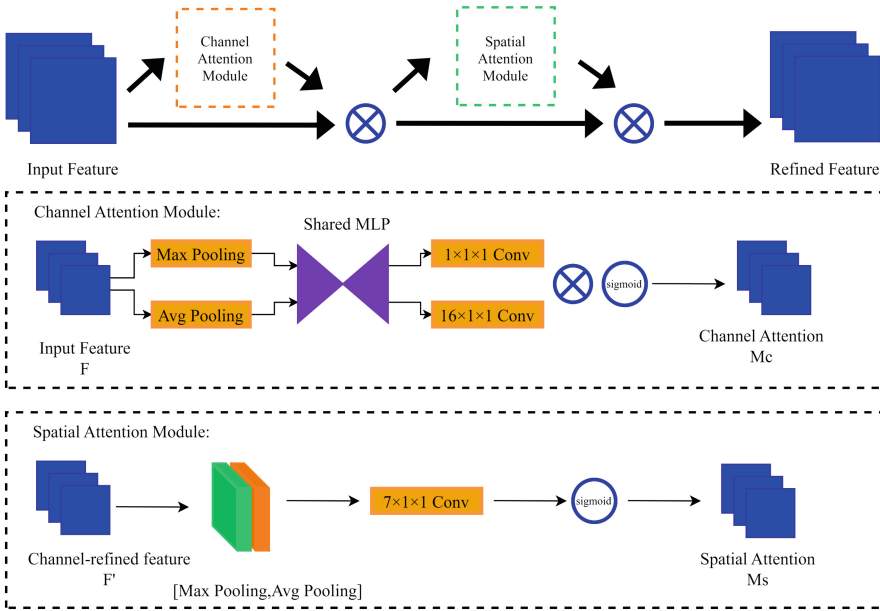


Fig. 4. CBAM attention module.

The CBAM consists of two independent sub-modules: Channel Attention Module and Spatial Attention Module. They perform channel and spatial attention operations, respectively. In the Channel Attention Module, the input feature map undergoes global average pooling and global max pooling in width and height. These pooled features are passed through Multi-Layer Perceptrons (MLP) and summed element-wise. Subsequently, a sigmoid activation function generates the final channel attention feature map. The mathematical representation of the Channel Attention Module is expressed as:

$$M_c(F) = \sigma(\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{MaxPool}(F))) \tag{1}$$

The output feature map from the Channel Attention Module serves as the input for the Spatial Attention Module. Initially, global average and max pooling operations are executed across channels to produce two distinct feature maps. These maps are then concatenated channel-wise, followed by a convolutional operation to reduce the dimensions to a single channel. Finally, a sigmoid activation function generates the spatial attention feature. The output features of the Spatial Attention Module are element-wise multiplied by the input features, generating the final output feature map of the

CBAM. The formulation of the Spatial Attention Module is as follows:

$$Ms(F) = \sigma(f^{7 \times 7}([\text{AvgPool}(F); \text{MaxPool}(F)])) \quad (2)$$

Incorporating CBAM at the end of the neck region in the YOLOv5 architecture substantially enriches the model's feature representation capabilities during the final stage of feature fusion. It adaptively selects important channels and spatial locations, thereby bolstering feature expressiveness and perceptual abilities. By integrating CBAM, the YOLOv5 model shows better in its performance for defect detection tasks.

4.3 Small Object Detection Head

As can be inferred from the Fig. 2, The overall structure of the proposed net-work, small objects are a significant part of the training data. The fundamental YOLOv5 has three detection heads, performing detections on feature maps sized 20×20 , 40×40 , and 80×80 . Assuming that a defect target, when resized to 640 pixels in the input image, has a pixel value of 8, it would vanish post-convolution on an 80×80 feature map. Our dataset has a high proportion of small targets, and there are many defects with pixel values less than 8. Therefore, it is necessary to add a small Object detection head.

In our revised approach, we continued from the 80×80 feature map concatenation point of the original YOLOv5. It undergoes a C3 structure, convolution, up-sampling, and concatenation with the backbone's first C3. This sequence is followed by another C3 and CBAM to yield a 160×160 feature map.

5 Experimental Process and Analysis of Results

YOLOv5 offers four different variants, differentiated by their level of complexity: YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x. After careful consideration, YOLOv5l was selected for modification due to its balanced trade-off between detection accuracy and training difficulty.

The training dataset consists of 264 images, and the testing dataset comprises 50 images. Given the relatively small size of the training set and the absence of pre-trained weights for the modified architecture, the model was trained for an extended duration of 600 epochs. The batch size was set to 32, and all training images were resized to a uniform dimension of 640×640 . The hardware and software environment configuration used for all experiments in this article is shown in Table 1.

In the domain of object detection, commonly used evaluation metrics include Average Precision (AP), Mean Average Precision (mAP), Precision (P), and Recall (R). These metrics provide a comprehensive assessment of a model's effectiveness in terms of both its discriminatory power and its ability to correctly identify positive samples. True Positive (TP): The number of actual positive samples correctly predicted as positive. False Negative (FN): The number of actual positive samples incorrectly predicted as negative. False Positive (FP): The number of actual negative samples incorrectly predicted as positive. True Negative (TN): The number of actual negative samples correctly predicted

Table 1. Experimental environment configuration.

Operating system	Centos 7
CPU	Intel(R) Xeon(R) CPU E5-2680 v3 @ 2.50 GHz
GPU	NVIDIA GTX 1080Ti 11G
Memory	125 G
Programing language	Python 3.7
AIgorithm framework	Pytorch 1.7.0
Development environment	Anaconda

as negative. The metrics Precision (P), Recall (R), Average Precision (AP), and Mean Average Precision (mAP) can be defined as follows:

$$P = \frac{TP}{TP + FP} \quad (3)$$

$$R = \frac{TP}{TP + FN} \quad (4)$$

$$AP = \frac{P_1 + P_2 + \dots + P_n}{n} \quad (5)$$

$$mAP = \frac{AP_1 + AP_2 + \dots + AP_n}{n} \quad (6)$$

In the experimental process, four CBAM modules were appended to the end of the Neck, along with the inclusion of a small object detection head, aiming to enhance the model's capability in detecting plate defects. In order to validate the effectiveness of the attention modules and the small object detection head within the plate defect detection model, comparative experiments were conducted. The detection performance of the model with the integration of CBAM positions and the small object detection head in YOLOv5 was compared, and the comparative results are presented in Table 2.

Table 2. The comparison of models.

Model	CBAM	Little Head	P/%	R/%	mAP@0.5/%
YOLOv5l	×	×	82.1	76.4	80.1
Improve1	3	×	80.8	75.1	80.4
Improve2	4	×	82.0	71.4	78.9
Improve3	×	√	80.2	71.8	79.7
Improve4	5	√	78.4	76.5	81.3
YOLOv5l-ours	4	√	83.3	71.2	81.6

From the experimental data in Table 2, it can be seen that compared with Improve 1 where CBAM was not added at the end of the backbone, adding CBAM to the backbone in Improve 4 that did not lead to an increase in detection performance, but a drop in map50 from 80.4% to 78.9%. This outcome can be attributed to the fact that the original main feature extraction network of YOLOv5 is already proficient, and the relatively small size of our dataset is insufficient to effectively train the newly introduced CBAM-enhanced main network. However, the introduction of CBAM to the Neck of YOLOv5 has shown beneficial effects in terms of fusing features of various scales more effectively, thereby enhancing the model's perceptual capability and accuracy.

We selected Faster R-CNN, SSD, and YOLOv8 as comparative experiments in the context of object detection. All these algorithms were trained on pre-trained base models. This approach not only significantly reduces training time but also enhances the detection performance of the networks. The final training outcomes and evaluations of each algorithm are summarized in Table 3.

Table 3. The comparison of different algorithms.

Model	Backbone	P/%	R/%	mAP@0.5/%
Faster R-CNN	ResNet-50	39.4	78.5	62.6
SSD	ResNet-50	88.9	39.0	71.5
YOLOv8	CSPDarkNet-53	71.34	73.9	77.2
YOLOv5s	CSPDarkNet-53	79.0	70.8	78.9
YOLOv5l	CSPDarkNet-53	82.1	76.4	80.1
YOLOv5x	CSPDarkNet-53	83.1	73.9	80.3
YOLOv5l-ours	CSPDarkNet-53	83.3	71.2	81.6

Compared to the original YOLOv5l model, our proposed method achieved an approximately 1.5% improvement in mAP. Furthermore, when compared to other algorithms, our method demonstrated superior performance on the test dataset. This leads to the conclusion that the detection approach proposed by the author, involving the incorporation of a small object detection head and CBAM attention modules, significantly enhances the efficacy of plate defect detection. Figure 5 shows the different sheet defects predicted by our algorithm model.

In order to verify the effectiveness of this algorithm, we introduced the NEU steel defect dataset, which is similar to our plate defect dataset, for comparative experiments. The results and evaluations before and after improvement for each defect dataset are summarized in Table 4.

Judging from the results of the NEU steel defect dataset, our algorithm has improved defect detection.

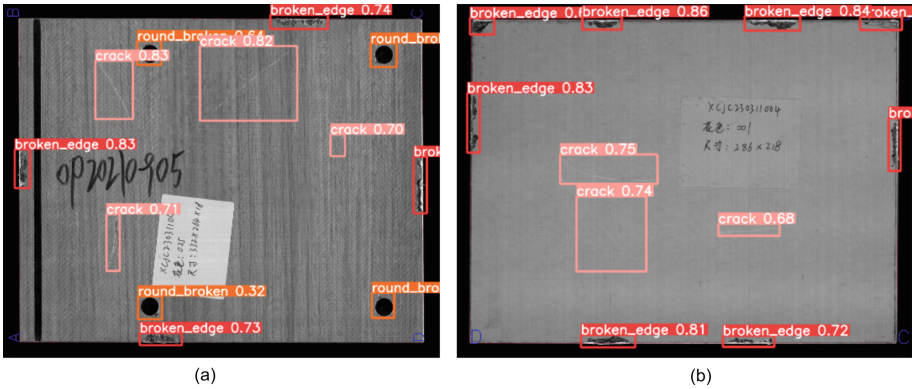


Fig. 5. Predicted plate defect images.

Table 4. The comparison of different datasets.

Model	Improve	P/%	R/%	mAP@0.5/%
NEU-DET	×	73.4	77.6	80.3
	✓	73.4	82.7	85.7
Our dataset	×	82.1	76.4	80.1
	✓	83.3	71.2	81.6

6 Conclusion

Based on the plate defect dataset provided by a certain intelligence company, this paper addresses challenges related to defect feature complexity and the small size of defect targets. By enhancing the YOLOv5 object detection algorithm, notable improvements have been achieved in its detection performance on this dataset.

Experimental results underscore the effectiveness of this approach in enhancing plate defect detection. The proposed model demonstrates an approximate 1.5% increase in AP compared to the original YOLOv5l model.

To simulate a real-life scenario in plate production, we have curated a novel wood plate defect dataset characterized by a high proportion of small-sized targets, effectively reflecting the actual defects found in production. This dataset provides valuable material for an objective evaluation of the model’s performance.

References

1. Cai, J., Zhang, L.-B.: Apple surface defect detection based on deep learning. In: 2021 International Conference on Electronic Information Technology and Smart Agriculture (ICEITSA), pp. 491–495 (2021)

2. Gao, Y., Gao, L., Li, X.: A generative adversarial network based deep learning method for low-quality defect image reconstruction and recognition. *IEEE Trans. Ind. Inf.* **17**, 3231–3240 (2021). <https://doi.org/10.1109/TII.2020.3008703>
3. Xie, Q., Zhou, W., Tan, H., Wang, X.: Surface defect recognition in steel plates based on improved faster R-CNN. In: 2022 41st Chinese Control Conference (CCC), pp. 6759–6764 (2022)
4. Melakhsou, A.A., Baton-Hubert, M., Casoetto, N.: Computer vision based welding defect detection using YOLOv3. In: 2022 IEEE 27th International Conference on Emerging Technologies and Factory Automation (ETFA), pp. 1–6 (2022)
5. Rehana, K.C., Remya, G.: Road damage detection and classification using YOLOv5. In: 2022 Third International Conference on Intelligent Computing Instrumentation and Control Technologies (ICICT), pp. 489–494 (2022)
6. Lin, C.-H., Ho, C.-W., Hu, G.-H., Kuo, P.-C., Hu, C.-Y.: Alloy cast product defect detection based on object detection. In: 2021 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS), pp. 1–2 (2021)
7. Li, L., Jiang, Z., Li, Y.: Surface defect detection algorithm of aluminum based on improved faster RCNN. In: 2021 IEEE 9th International Conference on Information, Communication and Networks (ICICN), pp. 527–531 (2021)
8. Luo, L., Deng, C., Wu, Z., Wang, S., Ye, T.: Automobile airbag defect detection algorithm based on improved faster RCNN. In: 2021 International Conference on Computer Engineering and Artificial Intelligence (ICCEAI), pp. 281–285 (2021)
9. Liu, Y., Wang, J., Yu, H., Li, F., Yu, L., Zhang, C.: Surface defect detection of steel products based on improved YOLOv5. In: 2022 41st Chinese Control Conference (CCC), pp. 5794–5799 (2022)
10. Kuan-Ying, S., Ming-Fei, C., Po-Cheng, T., Cheng-Han, T.: Establish a dynamic detection system for metal bicycle frame defects based on YOLO object detection. In: 2022 IET International Conference on Engineering Technologies and Applications (IET-ICETA), pp. 1–2 (2022)
11. Guan, S., Wang, X., Wang, J., Yu, Z., Wang, X., Zhang, C., Liu, T., Liu, D., Wang, J., Zhang, L.: Ceramic ring defect detection based on improved YOLOv5. In: 2022 3rd International Conference on Computer Vision, Image and Deep Learning & International Conference on Computer Engineering and Applications (CVIDL & ICCEA), pp. 115–118 (2022)
12. Girshick, R., Donahue, J., Darrell, T., Malik, J.: Rich feature hierarchies for accurate object detection and semantic segmentation. In: 2014 IEEE Conference on Computer Vision and Pattern Recognition, pp. 580–587 (2014)
13. Girshick, R.: Fast R-CNN (2015). <http://arxiv.org/abs/1504.08083>
14. Ren, S., He, K., Girshick, R., Sun, J.: Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks (2016). <http://arxiv.org/abs/1506.01497>
15. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: unified, real-time object detection. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779–788 (2016)
16. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S.E., Fu, C.-Y., Berg, A.C.: SSD: single shot multibox detector. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) *Computer Vision - ECCV 2016 - 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I*, pp. 21–37. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46448-0_2
17. Li, B., Ren, F., Ni, H., Kang, X., Lv, S., Hao, Z.: Classification method of surface defects of aluminum profile based on transfer learning. In: 2022 International Conference on Machine Learning and Intelligent Systems Engineering (MLISE), pp. 1–5 (2022)

18. Kaur, P., Khehra, B.S., Mavi, Er.B.S.: Data augmentation for object detection: a review. In: 2021 IEEE International Midwest Symposium on Circuits and Systems (MWSCAS), pp. 537–543 (2021)
19. Luo, W., Sun, L.: An improved binarization algorithm of wood image defect segmentation based on non-uniform background. *J. For. Res.* **30**, 1527–1533 (2019)
20. Pahlberg, T., Hagman, O., Thurley, M.: Recognition of boards using wood fingerprints based on a fusion of feature detection methods. *Comput. Electron. Agric.* **111**, 164–173 (2015)
21. He, T., Liu, Y., Xu, C., Zhou, X., Hu, Z., Fan, J.: A fully convolutional neural network for wood defect location and identification. *IEEE Access.* **7**, 123453–123462 (2019)
22. Yang, Y., Zhou, X., Liu, Y., Hu, Z., Ding, F.: Wood defect detection based on depth extreme learning machine. *Appl. Sci.* **10**, 7488 (2020)
23. Shi, J., Li, Z., Zhu, T., Wang, D., Ni, C.: Defect detection of industry wood veneer based on NAS and multi-channel mask R-CNN. *Sensors* **20**, 4398 (2020)
24. Xu, J., Yang, H., Wan, Z., Mu, H., Qi, D., Han, S.: Wood surface defects detection based on the improved YOLOv5-C3Ghost with SimAm module. *IEEE Access* **11**, 105281 (2023)
25. Schlegl, T., Seeböck, P., Waldstein, S.M., Schmidt-Erfurth, U., Langs, G.: Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. In: Niethammer, M., et al. (eds.) *Information Processing in Medical Imaging*, pp. 146–157. Springer International Publishing, Cham (2017). <https://doi.org/10.1007/b137723>
26. Han, X., Chen, X., Liu, L.-P.: GAN ensemble for anomaly detection. *Proc. AAAI Conf. Artif. Intell.* **35**, 4090–4097 (2021)
27. Zavrtnik, V., Kristan, M., Skočaj, D.: Reconstruction by inpainting for visual anomaly detection. *Pattern Recogn.* **112**, 107706 (2021)
28. Wang, C.-Y., Bochkovskiy, A., Liao, H.-Y.M.: YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors (2022). <http://arxiv.org/abs/2207.02696>
29. Woo, S., Park, J., Lee, J.-Y., Kweon, I.S.: CBAM: convolutional block attention module. In: *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 3–19 (2018)
30. Benjumea, A., Teeti, I., Cuzzolin, F., Bradley, A.: YOLO-Z: Improving Small Object Detection in YOLOv5 for Autonomous Vehicles. <https://arxiv.org/abs/2112.11798v4>. Accessed 28 July 2023