Neutrosophic Enhancement of YOLO-MD Algorithm for Automated Metal Surface Micro Defect Detection

Li Jiao [1,2], Muhammad Irsyad Abdullah [3*]

[1] School of Graduates Studies of Management and Science University, Selangor, Malaysia
[2] Mianyang Polytechnic, Mianyang, Sichuan, China
[3] Software Engineering and Digital Innovation Center, Management and Science University, Selangor, Malaysia

Emails: lijiao@mypt.edu.cn ; irsyad@msu.edu.my

Abstract

To achieve automation of defect detection, the metal surface micro defect detection algorithm YOLO-MD is proposed. From the perspective of object detection, YOLOv5s is selected as the backbone algorithm and the SPD-Conv module is added to reduce feature loss caused by ordinary convolutional downsampling, improve the adaptability of low-resolution images, and improve the accuracy of small object detection. Using the MPDIoU loss function to accelerate model convergence and improve detection accuracy. Considering the small size of the dataset, data augmentation methods were adopted. After model training, mAP50-95 improved by 0.02 compared to YOLOv5, which has high real-time and robustness and can more effectively detect metal surface micro defects.

Keywords: Object detection; Metal surface micro defect detection; YOLO

1. Introduction

With the development of the manufacturing industry, metal workpieces are becoming increasingly refined and complex, and people's requirements for the surface quality of metal workpieces are also increasing. However, during the forging process, due to factors such as production environment, raw materials, and processing technology, small cracks and pits are inevitable on the surface of the workpiece. These defects not only affect the aesthetics of the workpiece, but also lead to a decrease in fatigue resistance, survivability, and corrosion resistance, posing significant safety hazards. Conducting research on metal surface micro defect detection and achieving automation of defect detection can reduce false or missed detections caused by manual screening.

This research from the perspective of object detection, focuses on the task of metal surface micro defect detection and proposes the metal surface micro defect detection algorithm YOLO-MD based on the YOLOv5s.

YOLOv5s is selected as the backbone algorithm, mainly for the following reasons:

YOLOv5 is a one-stage object detection algorithm, with a detection speed of 140FPS in its YOLOv5s version, which greatly improves the speed of object detection while ensuring accuracy.Moreover,YOLOv5 is a high-performance, flexible, and scalable object detection algorithm that provides an open hyperparameter configuration and code integration environment for developers to quickly turn innovation into reality.YOLOv5 is also a user-friendly algorithm, users and improvers can easily load and train their own datasets. Also, YOLOv5 is developed based on the Pytorch framework, and its weight file can be easily converted into Android and IOS mobile application formats, making deployment convenient and simple.

On YOLOv5s, It replaces the downsampling convolution module with the SPD-Conv module to reduce feature loss caused by ordinary convolution downsampling, improves the adaptability of low-resolution images, and improves the accuracy of small object detection. Using MPDIoU loss function instead of CIoU loss function to

DOI: https://doi.org/10.54216/IJNS.230225
Received: July 21, 2023 Revised: September 21, 2023 Accepted: December 26, 2023

accelerate model convergence and improve detection accuracy. By Compare with existing algorithms, YOLO-MD can effectively distinguish defects from normal areas, eliminates non-defect interference such as metal oxidation and watermarks, and has better detection ability for micro defects on metal surfaces. Due to inheriting the real-time advantage of YOLO series algorithms, it can meet the requirements of real-time detection on industrial production lines. Neutronsophic algebraic structures have been studied recently and has established the groundwork for an entire family of novel mathematical theories, including a neutrosophic set theory, that generalize both its fuzzy and classical equivalents. This work offer an enhancement which is aim to gratly improve the ability of the algorithm to accurately detect and classify micro defects on metal surfaces, especially in scenarios where the information is vague or uncertain

2. Literature Review

Scholars have conducted extensive research on the detection of metal surface defects. It can be roughly divided into traditional machine learning direction [1, 2, 3] and deep learning direction [4, 5, 6]. Traditional machine learning methods require manual construction of image features, and problems such as noise interference and background confusion often limit their application scope. Deep learning methods continuously learn visual and semantic features such as colors, textures, and shapes, resulting in strong abstraction and detection abilities. The metal surface defect detection methods of deep learning can be further divided into image classification, image segmentation, and object detection：

- **Image classification.** Shang et al. [7] propose a method for identifying rail defects， the algorithm first crops the orbit region and then uses the Inception V3 network for classification. Chen et al. [8] proposed an ensemble approach method of steel surface defect detection, it uses 3 different DCNN models and applies their output combinations to defect classification. Akram et al. [9] used a Convolutional Neural Network (CNN) architecture to automatically detect defects in photovoltaic cells in electroluminescent images， based on simple network structures such as convolution, pooling, and fully connected layers, feature vectors for defect classification were obtained.
- **Image segmentation.** Bao et al. [10] proposed the Triplet-Graph Reasoning Network (TGRNet), to transform the few-shot metal surface defect segmentation problem into a few-shot semantic segmentation problem of defect area and background area. Neven et al. [11] proposed a multi-task model that performs defect segmentation and severity estimation. and implemented and trained a multi-task U-Net segmentation network to detect steel sheet surface defects. Aslam et al. [12] proposed an automatic segmentation and quantization method using a U-Net structure for detecting defects in digital images of titanium-coated metal surfaces.
- **Object detection.** Cheng et al. [13] used RetinaNet as the basic network and added differential channel attention and adaptive spatial feature fusion for steel surface defect detection. Kou et al. [14] developed an end-to-end defect detection model based on YOLO-V3, replacing the anchor structure with anchor-free and incorporating dense convolutional blocks for surface defects of the steel strip detection. Aiming at real-time detection of steel strip surface defects, Ren et al. [15]proposed a slighter Faster R-CNN. Replacing convolutional layers with depthwise separable convolutions, using center loss to improve the network's ability to distinguish different types of defects.

Among the three methods, image classification is an early solution that requires cropping and positioning before classification, which has a significant disadvantage compared to the latter two. Image segmentation and object detection are currently common technical directions.

3. Related Work

A. Related technologies

a) YOLOv5s

In June 2020, YOLOv5 was officially released, inheriting the real-time and precision advantages of the YOLO series algorithm. YOLOv5 was widely used in the industry and achieved great success once it was launched. Currently, 7 versions have been updated. YOLOv5 is divided into five types: n, s, m, l, and x, with an increasing number of parameters and improved detection performance, **Figure 1**. Considering the need for model deployment on embedded platforms, YOLOv5s was selected as the backbone model for improvement.

DOI: https://doi.org/10.54216/IJNS.230225
Received: July 21, 2023 Revised: September 21, 2023 Accepted: December 26, 2023

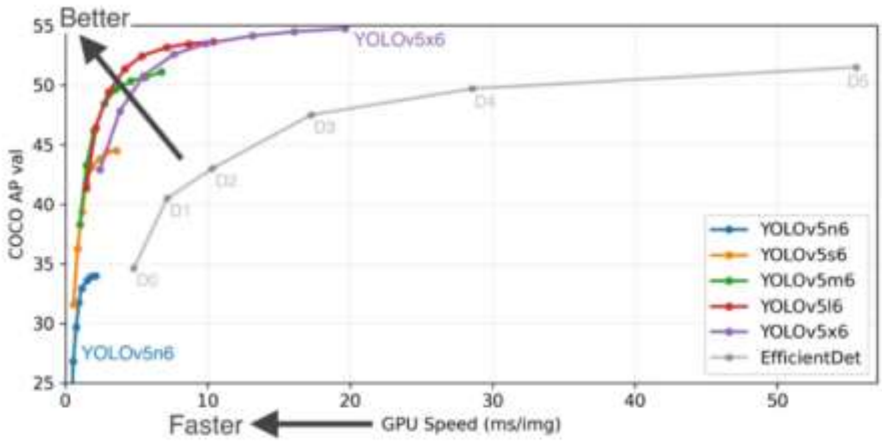


Figure 1: The performance of the 5 scales

The YOLOv5 is mainly divided into the following parts:

i. Data enhancement: It uses Mosaic data augmentation, which involves arbitrarily concatenating 4 training set images to increase sample diversity, including small object samples.

ii. Backbone: Using the modified CSPDarknet53 [16] as the backbone network, it integrates basic structures such as C3 and SPPF. The C3 module utilizes the extraction and diversion ideas of CSPNet, combined with the idea of residual structure, to fuse features and enrich the semantic information of features. SPPF is an improved version of the SPP [17] structure, which undergoes three rounds of pooling with the same size for result concatenation. The model size decreases but the number of channels increases, and the feature map retains rich semantic information.

iii.Neck: Adopting the PANet [18] structure for bidirectional feature fusion can better fuse the features of objects at different scales and improve the detection performance of each detection head.

iv.Head. To improve the performance of small object detection, YOLOv5 adopts three detection heads. The Loss function mainly consists of three parts, the first is the bounding box regression loss (L_{CIOU_j}), the second is object loss (L_{obj_j}), and the third is classification loss(L_{cls_j}). The calculation formula is as follows:

$$L_{total} = \sum_{i}^{n} (\lambda_1 \sum_{j}^{B_i} L_{CIOU_j} + \lambda_2 \sum_{j}^{s_i \times s_i} L_{obj_j} + \lambda_3 \sum_{j}^{B_i} L_{cls_j}) \qquad (1)$$

b) SPD-CONV

In the YOLOv5, due to the need for feature fusion of PANET, the entire process involves 7 downsampling, which may lead to the loss of fine-grained information, such as small object features, during the downsampling process. Detection performance for small objects decreases. To address this issue, Sunkara et al. [19] proposed a new CNN module called SPD-Conv. It consists of SPD layers and 1 * 1 convolutional layers, which transform feature maps from space to depth. For any size of feature map X, a series of sub-feature maps can be extracted from left to right and from top to bottom. The process is represented as follows:

$$f_{0,0} = X[0:S:scale, 0:S:scale], f_{1,0} = X[1:S:scale, 0:S:scale], \ldots,$$
$$f_{scale-1,0} = X[scale-1:S:scale, 0:S:scale] \qquad (2)$$

$$f_{0,1} = X[0:S:scale, 1:S:scale], f_{1,1} = X[1:S:scale, 1:S:scale], \ldots,$$
$$f_{scale-1,1} = X[scale-1:S:scale, 0:S:scale] \qquad (3)$$

$$\begin{aligned} f_{0,scale-1} &= X[0:S:scale, scale-1:S:scale], f_{1,scale-1} \\ &= X[1:S:scale, scale-1:S:scale], \ldots, f_{scale-1,scale-1} \\ &= X[scale-1:S:scale, scale-1:S:scale] \qquad (4) \end{aligned}$$

Taking scale=2 as an example, the original feature map (a) is downsampled twice to obtain the four feature maps ($f_{0,0}, f_{0,1}, f_{1,0}, f_{1,1}$) shown in **Figure 2**：

DOI: https://doi.org/10.54216/IJNS.230225
Received: July 21, 2023 Revised: September 21, 2023 Accepted: December 26, 2023

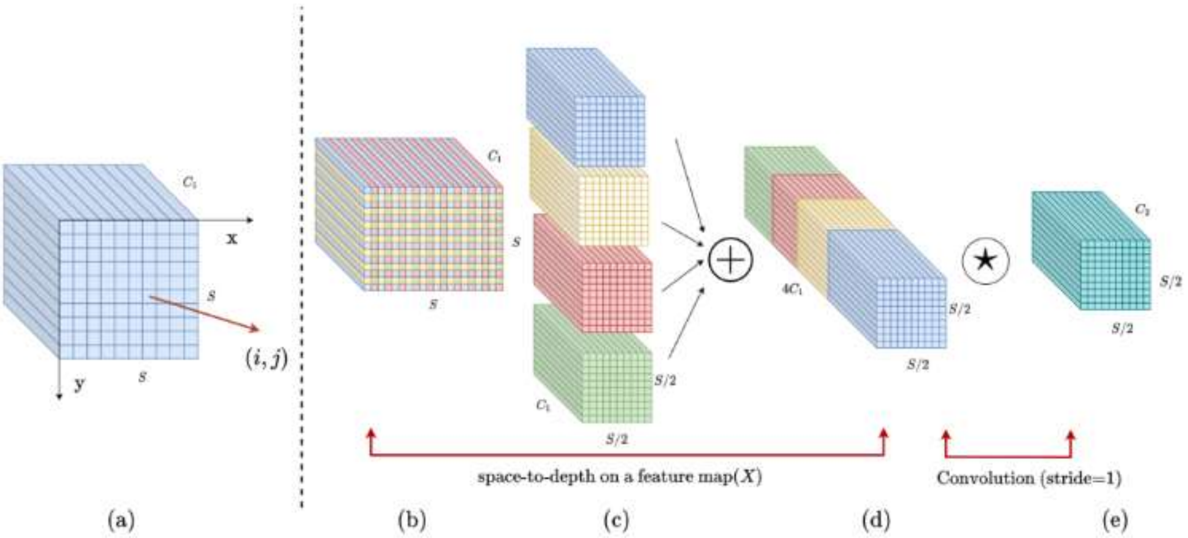
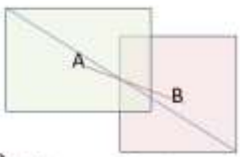


Figure 2: Illustration of SPD-Conv when scale=2

Next, concatenate the four feature maps in the channel dimension to obtain a feature map with a half spatial dimension and twice channel dimension. Finally, using 1 * 1 convolutional blocks performs channel transformation.

c) MPDIoU

YOLOv5 defaults to using CIoU to calculate the bounding box regression loss. It is proposed based on the minimum standard distance between the centre points of two detection boxes in DIoU and taking into account the aspect ratio between the predicted box and the target box. The calculation method is shown in **Figure 3**.



$$v = \frac{4}{\pi^2}(arctan\frac{\omega^{gt}}{h^{gt}} - arctan\frac{\omega}{h})^2$$

$$\alpha = \frac{v}{(1 - IoU) + v}$$

$$CIoU = IoU - \frac{\rho^2}{c^2} - \alpha v$$

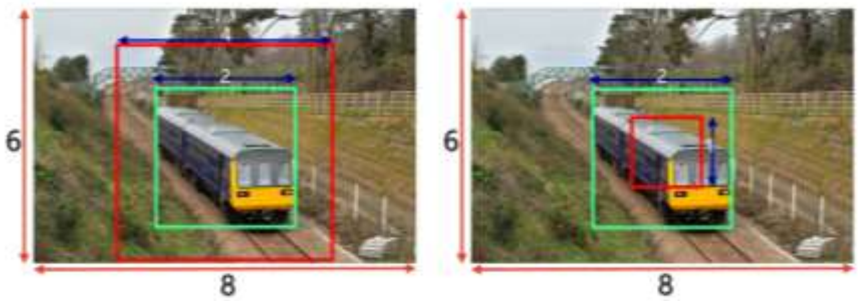
$$L_{CIoU} = 1 - CIoU$$

ρ Distance of center point

c Diagonal length of minimum bounding box

Figure 3: Schematic diagram of CIoU loss calculation

However, as shown in **Figure 3**, when the aspect ratio of the red prediction box and the green ground truth box is consistent, the width and height are different, and the centre points overlap, the CIoU loss cannot converge.



(a) $\mathcal{L}_{GIoU} = 0.75$, $\mathcal{L}_{DIoU} = 0.75$, $\mathcal{L}_{CIoU} = 0.75$, $\mathcal{L}_{EIoU} = 1.25$, (b) $\mathcal{L}_{GIoU} = 0.75$, $\mathcal{L}_{DIoU} = 0.75$, $\mathcal{L}_{CIoU} = 0.75$, $\mathcal{L}_{EIoU} = 1.25$,

Figure 4: The situation of CIoU failure

DOI: https://doi.org/10.54216/IJNS.230225
Received: July 21, 2023 Revised: September 21, 2023 Accepted: December 26, 2023

To address this issue, Siliang et al. [20] proposed MPDIoU, which measures the distance between the top left and bottom right points of the prediction box and the ground truth box and minimizes it. The calculation formula as follows:

$$MPDIoU = IoU - \frac{d_1^2}{h^2 + w^2} - \frac{d_2^2}{h^2 + w^2} \quad (5)$$

$$L_{MPDIoU} = 1 - MPDIoU \quad (6)$$

In the formula, d_1 d_2 represents the distance between the top left and bottom right corners of the two boxes, w,h represents the width and height of the image. After using MPDIoU, the differences of lossing values of the two pictures shown in **Figure 4.** When the two boxes completely overlap, MPDIoU=0.

B. Framework

Using YOLOv5s as the backbone network, replace the downsampling convolution module with the SPD-CONV module, and use MPDIoU to calculate the bounding box regression loss. The framework of YOLO-MD is shown in **Figure 5.**

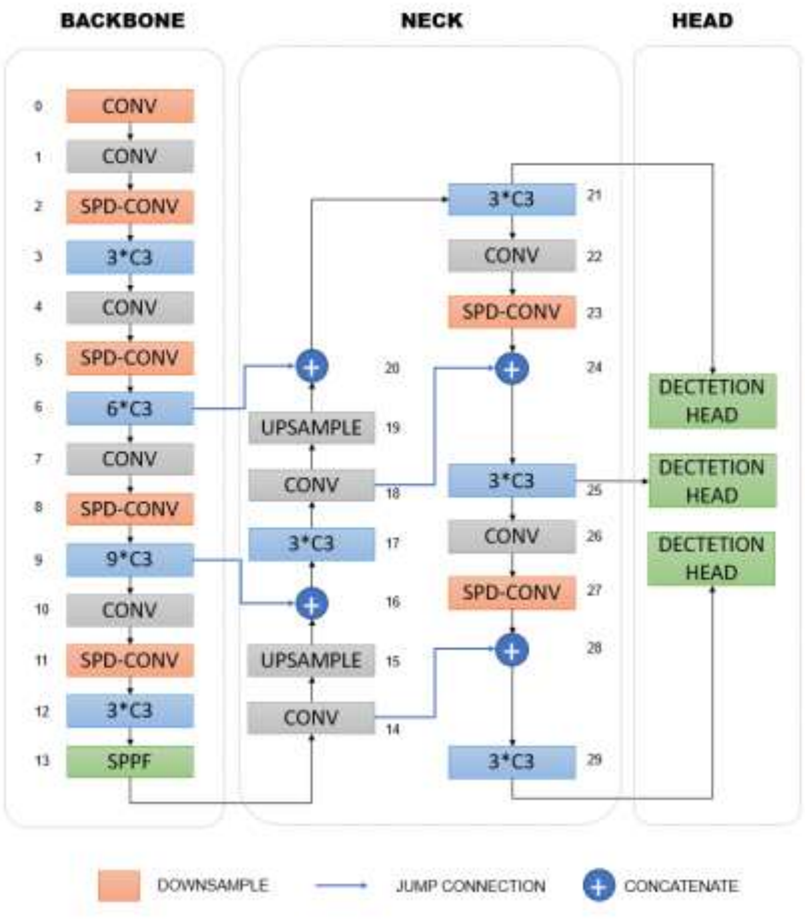


Figure 5: YOLO-MD Framework

C. Coding and Training Environment

This research uses an Intel Xeon Silver 4216 CPU, and 4 NVIDIA Tesla T4 GPU as the main hardware environment. The server operation system is CentOS7.9.2009, installing CUDA12.2, Python3.10.11 and PyTorch v1.13.

DOI: https://doi.org/10.54216/IJNS.230225
Received: July 21, 2023 Revised: September 21, 2023 Accepted: December 26, 2023

4. FINDING

a) Data Augmentation

The research used the publicly available dataset KolektorSDD [21]. The dataset includes 399 electronic commutator images, of which 52 are images with obvious defects and 347 are images without defects. These images are captured from steel surfaces in real industrial environments, and the defects mainly manifest as small damages, cracks, etc., with pixel-level annotations.

Next, annotating images by the labelimg tool and enhancing them with rotation, translation, scaling, and cropping. A total of 468 training images, 52 validation images, and 52 test images were generated, forming a dataset in 9:1:1 ratio.

b) Training

Train 300 epochs using SGD, set lr0=0.01, lrf=0.01, momentum=0.937, weight_ decay=0.0005. Set warmup_ epochs=3, warm up_ momentum=0.8, warmup_ bias_ lr=0.1. The learning rate within the first three epochs is relatively small, and the model can gradually stabilize. After the model is relatively stable, it automatically adopts a pre-set learning rate for training, which leads to faster convergence speed and better performance.

c) Study Results

After 300 epochs of model training, analyze the output results of the validation set， Loss、P、R、mAP50 and mAP50-95 curve as shown in **Figure 6**. Under the same operating environment, the performance comparison with YOLOv5 before improvement is shown in Table 4.1.

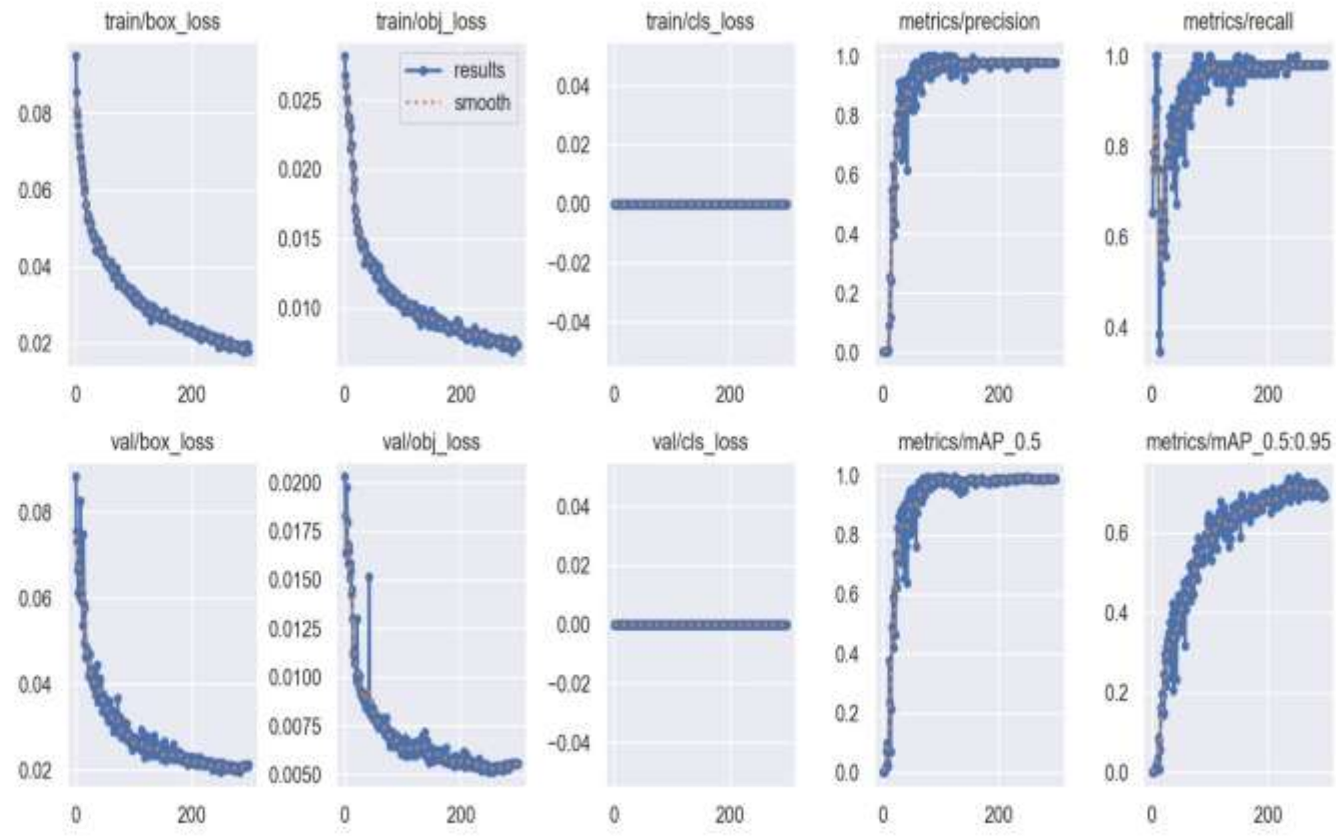


Figure 6: Analysis chart of experimental results

DOI: https://doi.org/10.54216/IJNS.230225
Received: July 21, 2023 Revised: September 21, 2023 Accepted: December 26, 2023

Table 4.1:Performance comparison between YOLO-MD and YOLOv5

Algorithm	mAP50	mAP50-95	Weight(M)
YOLOv5	0.99	0.72	14.1
YOLO-MD	0.99	0.74	17.1

Based on the above results analysis, mAP50-95 of YOLO-MD improved by 0.02 compared to YOLOv5, and the model size did not significantly increase, which can meet the efficient operation on embedded devices. The detection effect is shown in the **Figure 7.**

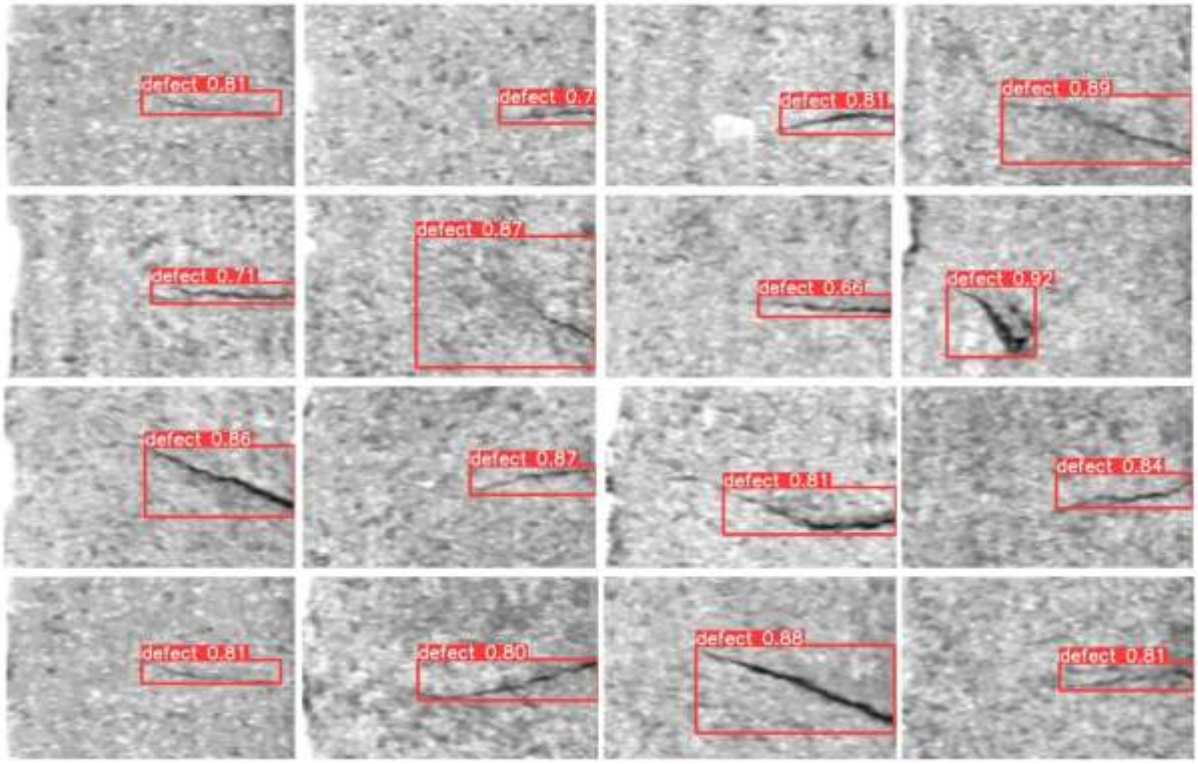


Figure 7: The detection effect

5. DISCUSSION

Good object detection algorithms are often trained on a large amount of data,otherwise the model performance may not reach the ideal state and overfitting may occur. Due to insufficient training datasets, the model may not be able to fully learn various types of defect features, resulting in poor performance and insufficient generalization ability when facing new data. Therefore, in order to further enrich the performance of training data and improve the robustness of the model, data augmentation is a low-cost and effective solution.

There are only 52 training images in the KolektorSDD dataset, the research only expanded through data augmentation. In future research, techniques such as transfer learning can be used to assist training by utilizing training datasets from other related fields, to compensate for the lack of training data for metal defects. In addition,in terms of excluding non-defect interference such as metal oxidation and watermarks, no false detection issues have been found based on existing test images. However, in the research, the design of the model has not sufficiently considered the problem, and there is still room for optimization.

DOI: https://doi.org/10.54216/IJNS.230225
Received: July 21, 2023 Revised: September 21, 2023 Accepted: December 26, 2023

6. Summary and Conclusion

The Research has improved YOLOv5s and proposed the metal surface micro defect detection algorithm YOLO-MD. To reduce feature loss caused by ordinary convolutional downsampling and improve the accuracy of small object detection, the SPD-Conv module has been added. To accelerate model convergence and improve detection accuracy, the MPDIoU loss function was adopted. Experimental results have shown that using the KolektorSDD public dataset for training, compared to YOLOv5s, mAP50-95 improves by 0.02, It has high real-time and robustness, and can more effectively detect micro defects in metal surfaces.

In future research, more effective image enhancement methods and training methods can be considered to expand the scale of training images. Furthermore, the model can be further optimized to address issues such as non-defect interference and micro defect classification, comprehensively enhancing its ability to detect defects.

7. Acknowledgment

The authors extend their sincere appreciation to the Faculty of Information Sciences and Engineering and the Software Engineering and Digital Innovation Centre at Management and Science University, Malaysia, for their invaluable assistance and support during the course of this study.

References

[1] K. Song and Y. Yan, "A noise robust method based on completed local binary patterns for hot-rolled steel strip surface defects," Applied Surface Science, pp. 858-864, 2013.

[2] G. K. Nand and N. Neogi, "Defect detection of steel surface using entropy segmentation," in In 2014 Annual IEEE India Conference (INDICON) , 2014.

[3] Y. Liu, K. Xu and D. Wang, "Online surface defect identification of cold rolled strips based on local binary pattern and extreme learning machine," Metals, p. 197, 2018.

[4] S. Wang, X. Xia, L. Ye and B. Yang, "Automatic detection and classification of steel surface defect using deep convolutional neural networks," Metals, p. 388, 2021.

[5] Z. Guo, C. Wang, G. Yang, Z. Huang and G. Li, "Msft-yolo: Improved yolov5 based on transformer for detecting defects of steel surface," Sensors, p. 3467, 2022.

[6] R. Tian and M. Jia, "DCC-CenterNet: A rapid detection method for steel surface defects," Measurement, p. 110211, 2022.

[7] L. Shang, Q. Yang, J. Wang, S. Li and W. Lei, "Detection of rail surface defects based on CNN image recognition and classification," in In 2018 20th International Conference on Advanced Communication Technology (ICACT) , 2018.

[8] W. Chen, Y. Gao, L. Gao and X. Li, "A new ensemble approach based on deep convolutional neural networks for steel surface defect classification," Procedia CIRP, pp. 1069-1072, 2018.

[9] M. W. Akram, G. Li, Y. Jin, X. Chen, C. Zhu, X. Zhao and A. Ahmad, "CNN based automatic detection of photovoltaic cell defects in electroluminescence images," Energy, p. 116319, 2019.

[10] Y. Bao, K. Song, J. Liu, Y. Wang, Y. Yan, H. Yu and X. Li, "Triplet-graph reasoning network for few-shot metal generic surface defect segmentation," IEEE Transactions on Instrumentation and Measurement, pp. 1-11, 2021.

[11] R. Neven and T. Goedemé, "A multi-branch U-Net for steel surface defect type and severity segmentation," Metals, p. 870, 2021.

[12] Y. Aslam, N. Santhi, N. Ramasamy and K. Ramar, "Localization and segmentation of metal cracks using deep learning," Journal of Ambient Intelligence and Humanized Computing, pp. 4205-4213, 2021.

[13] X. Cheng and J. Yu, "RetinaNet with difference channel attention and adaptively spatial feature fusion for steel surface defect detection.," IEEE Transactions on Instrumentation and Measurement, pp. 1-11, 2020.

[14] X. Kou, S. Liu, K. Cheng and Y. Qian, "Development of a YOLO-V3-based model for detecting defects on steel strip surface," Measurement, p. 109454, 2021.

[15] Q. Ren, J. Geng and J. Li, "Slighter Faster R-CNN for real-time detection of steel strip surface defects," in

DOI: https://doi.org/10.54216/IJNS.230225
Received: July 21, 2023 Revised: September 21, 2023 Accepted: December 26, 2023

2018 Chinese Automation Congress (CAC) , 2018.

[16] A. Bochkovskiy, C. Y. Wang and H. Y. M. Liao, "Yolov4: Optimal speed and accuracy of object detection," 2020. [Online]. Available: arXiv:2004.10934.

[17] K. He, X. Zhang, S. Ren and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," in IEEE transactions on pattern analysis and machine intelligence, 2015.

[18] S. Liu, L. Qi, H. Qin, J. Shi and J. Jia, "Path aggregation network for instance segmentation," in In Proceedings of the IEEE conference on computer vision and pattern recognition, 2018.

[19] R. Sunkara and T. Luo, "No more strided convolutions or pooling: A new CNN building block for low-resolution images and small objects," in In Joint European Conference on Machine Learning and Knowledge Discovery in Databases , 2022.

[20] M. Siliang and X. Yong, "MPDIoU: A Loss for Efficient and Accurate Bounding Box Regression.," 2023. [Online]. Available: arXiv:2307.07662.

[21] D. Tabernik, S. Šela, J. Skvarč and D. Skočaj, "Segmentation-based deep-learning approach for surface-defect detection," Journal of Intelligent Manufacturing, pp. 759-776, 2020.

[22] R. Vorobel, I. Ivasenko, O. Berehulyak and T. Mandzii, "Segmentation of rust defects on painted steel surfaces by intelligent image analysis.," Automation in Construction, p. 103515, 2021.

DOI: https://doi.org/10.54216/IJNS.230225
Received: July 21, 2023 Revised: September 21, 2023 Accepted: December 26, 2023