International Journal of Advanced Engineering Research and Science (IJAERS)
Peer-Reviewed Journal
ISSN: 2349-6495(P) | 2456-1908(O)
Vol-12, Issue-2; Feb, 2025
Journal Home Page Available: https://ijaers.com/
Article DOI: https://dx.doi.org/10.22161/ijaers.122.4

A Welding Defect Detection Algorithm Based on Deep Learning

Yi Chen[1], Yan Zuo Chang [3*], Lin Po Shang [1], Ze Feng Lin[2], Yong Qi Chen[4], Liu Yi Yu [1], Wan Ying Wu[1], Jun Qi Liu[1]

[1]Department of Energy and Power Engineering, Process Equipment and Control Engineering major, Guangdong University of Petrochemical Technology, China
[2]Department of Computer Science, Computer Science and Technology major, Guangdong University of Petrochemical Technology, China
[3]Department of Energy and Power Engineering, Guangdong University of Petrochemical Technology, China
[4]Department of Automation, Automation major, Guangdong University of Petrochemical Technology, China
*Corresponding author: cy14767703430@outlook.com

Received: 03 Jan 2025,

Received in revised form: 08 Feb 2025,

Accepted: 13 Feb 2025

Available online: 22 Feb 2025

***Keywords*— Deep learning, SCConv, Weld defect, YOLOv8**

Abstract*— *In order to meet the needs of process inspection technology for industrial equipment, image recognition technology based on deep learning has shown great potential in the field of welding defects. In this paper, an improved YOLOv8 algorithm is proposed to improve the welding defect identification ability of the workpiece. Through experimental verification on selected data sets in kaggle, this study evaluates the detection performance of YOLOv8 improved algorithm that integrates SCConv in C2f module at Backbone level. The experimental results show that the improved YOLOv8 has improved the accuracy of welding defect detection compared with the traditional version, and has certain application potential.

I. INTRODUCTION

In the field of industrial equipment, welding process is an important technical requirement, good welding technology can ensure the solidity and integrity of welded joints, avoid product quality problems and safety hazards caused by welding defects, at the same time, improve production efficiency, reduce production costs in large engineering projects, welding process can significantly affect the cost and schedule of the project. Some welding processes can achieve automatic welding, improve production efficiency and product quality, reduce labor costs and energy consumption, and in the welding process due to material properties, process parameters, equipment status and other factors, there may be cracks, pores, slag and other welding defects, if these defects are not found and treated in time, It may lead to product failure or even safety accidents during use. Through welding defect detection, these defects can be found and dealt with in time, thereby preventing potential quality problems and safety hazards.

The traditional detection process is mainly manual detection after the equipment is welded by the workers, and this method is not only time-consuming and labor-intensive, but also leads to workers' inattention during a long time working, resulting in some defects not being discovered [1]. Some researchers use relevant physical information to detect welding defects, such as Droubi [2] et al. Defects in carbon steel welds can be detected and identified by evaluating information such as peak amplitude and RMS value. Y [3] et al. studied pulsed induction thermal imaging (PIT) for detecting hidden defects in stainless steel welds, while Bebiano [4] et al. proposed a new detection mode, which simulated the data of relevant welding defects by using interference arc. The disturbance generated by the simulated arc is captured by the spectrometer, and then the correlation detection algorithm is used to indicate the existence and location of these defects.

With the continuous development of deep learning technology, many researchers begin to use deep learning technology for non-destructive testing of welding defects, especially in computer vision and algorithms. For example, (H [5], Bing Zhu [6], Yang [7], R [8]) and others use relevant neural network models to detect welding images after X-ray processing. Liu et al. proposed an improved and optimized fast multipath vision transformer (FMPVIT) framework for welding defect detection and identification [9], Li [10] et al. designed a welding defect detection method based on cross-layer feature fusion, and used an irregular long weld extraction algorithm based on drift gauss to improve efficiency and accuracy. Oh et al, [11] proposed a FAST R-CNN automatic detection method for welding defects based on deep learning. In 2016, Joseph Redmon et al. proposed a one-stage object detection network [12], which has the advantage of fast detection speed and can process about 45 frames of images per second. The author named it You Only Look Once, Therefore, the first generation of YOLO algorithm was born [13]. The core idea of YOLO algorithm is to transform the object detection task into a regression problem, and use convolutional neural network to infer images directly to achieve real-time object detection. Now YOLO series algorithms have been updated for many generations and are widely used in the field of target detection. For object surface defect detection, Hatab [14] et al proposed a steel surface defect detection system based on YOLO algorithm, and Zhao [15] et al proposed a model named LDD-YOLO based on YOLOv5 for steel surface defect detection. M [16] et al. designed a YOLO-HMC network, which realized more accurate and efficient identification of micro-sized PCB board defects with fewer model parameters. Gao [17] et al. improved the YOLOv5 algorithm. RepVGG module, NAM and lightweight uncoupled head RD-Head are introduced to improve the detection performance of the algorithm and are applied to weld feature detection. While Light-YOLO-Welding, a new type of lightweight detector based on improved YOLOv4 developed by L [18] et al., is used to detect weld feature points. In this study, we proposed an improved YOLOv8 welding defect detection algorithm based on YOLOv8 target detection algorithm. The experimental results show that the detection accuracy of the improved algorithm is improved, and the performance of various indexes is good.

II. INTRODUCTION TO EXPERIMENTAL MODEL

YOLOv8 is a new generation of object detection algorithm introduced by ultralytics for real-time object detection. Based on the previous YOLO version, YOLOv8 introduced new features and optimizations, including a more complex network architecture, a more optimized training flow, and a more powerful feature extraction capability [19].

Its general architecture is composed of Backbone (backbone network), Neck (neck network) and Head (head network), and its network structure is shown in Figure 1:

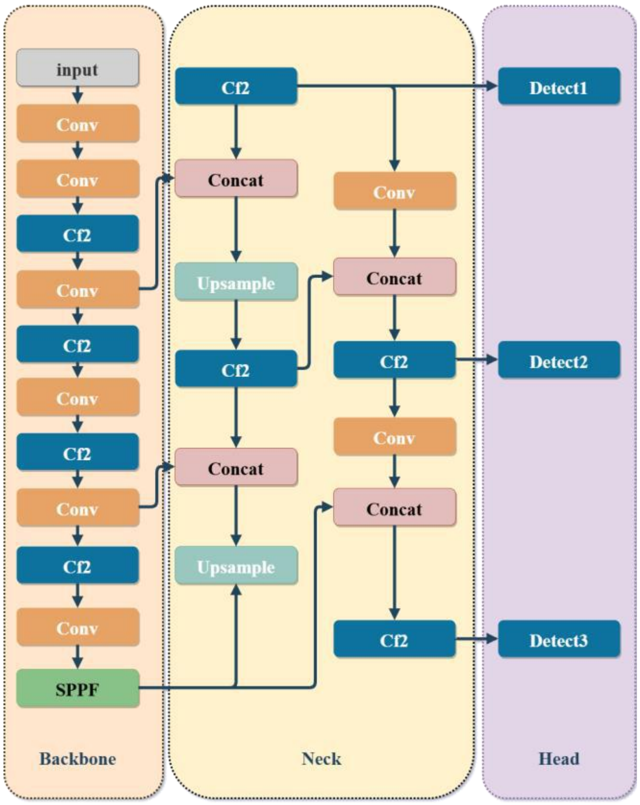

Fig.1: YOLOv8 network structure diagram

Backbone is responsible for feature extraction. It adopts a series of convolutional and deconvolution layers to extract feature information of different levels from input images. Using ResNet' s idea for reference, Backbone uses residual connections to reduce the size of the network, reduce the computational complexity and the number of parameters, and improve the running speed and efficiency of the model [20]. In this part, the C2f module is adopted as the basic constituent unit, and new structures and improved technologies are introduced, such as Depthwise Separable Convolution and DilatedConvolution, etc., to further optimize the capability of feature extraction and make the extracted features more representative and differentiated. So the target detection task can be carried out better. Compared with C3 module of YOLOv5, C2f module has fewer parameters and better feature extraction capability. It is the foundation and key component of the whole model, which provides strong support for the subsequent Neck and Head network parts, and jointly realizes the efficient and accurate target detection task.

Through multi-scale feature fusion, SPPF module in YOLOv8 pooled, splicing and fused feature maps of different scales, effectively expanded the sensitivity field and extracted rich information. By optimizing the algorithm, the computational load was reduced, the accuracy was improved by using large kernel convolution, and the detection ability of targets of different sizes was enhanced, thus improving the performance and robustness of the model.

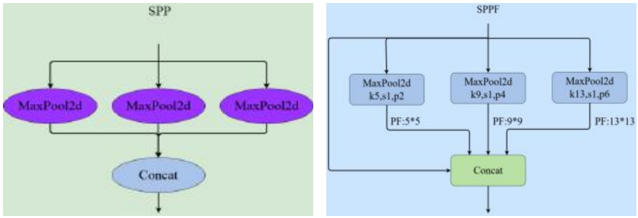

Fig.2: Comparison between SPP module and SPPF module

SPP works by concatenating inputs with three parallel MaxPool layers. The three MaxPool tiers typically use 5*5, 9*9, and 13*13 kernels. This allows features at different scales to be captured at the same time, thereby reducing identification errors due to changes in the scale of the input image. SPPF, on the other hand, connects three MaxPool layers with 5*5 kernel in series, and then combines them together through residual connection, which reduces the redundant calculation between feature graphs and improves the reasoning speed of the model [21].

The Neck part is responsible for multi-scale feature fusion, which enhances the feature representation ability by fusing feature maps from different stages of Backbone. The deep feature map carries stronger semantic features and weaker positioning information. Shallow feature maps carry strong positional information and weak semantic features [22]. Yolov8 uses PAN-FPN (Path Aggregation network-feature Pyramid Network) as its Feature Pyramid Network, and fuses feature maps from different stages of the backbone network to form feature maps with more semantic information and multi-scale perception. Through bottom-up and top-down path aggregation, feature maps of different scales are fused to achieve cross-scale information transfer, which further enhances the representation of multi-scale features and improves the detection ability of the model for objects of different sizes and shapes [23]. At the same time of feature fusion, the Neck network can further process and fuse the splicing features through modules and structures such as convolutional layer and residual connection, and extract more representative feature information, so that the model can identify and locate the target more accurately and improve the accuracy of target detection [24]. Hierarchical feature selection and fusion mechanisms are also introduced, such as the application of the latest lightweight neck network structures such as HS-FPN, which further reduces the number of model parameters and computational complexity, while improving the model performance.

Compared with YOLOv5 and YOLOv6, YOLOv8 removes the convolutional layer structure in the up-sampling stage of PAN-FPN and directly feeds the features output in different stages of Backbone into the up-sampling operation. The C3 module and RepBlock were also replaced with the C2f module [25]. The number of parameters and computational complexity of the model are reduced, and the computational efficiency is improved.

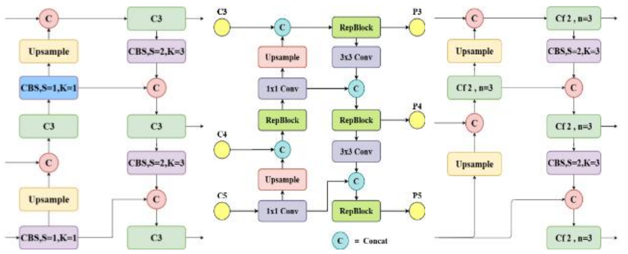

Fig.3: Comparison of YOLOv5, YOLOv6 and YOLOv8 Backbone

The Head part is responsible for further processing and integration of multi-scale features after neck network fusion, that is, the final target detection and classification task. The anchor-based is improved to Anchor-free, which does not rely on predefined Anchor frames, and flexibly processes targets of different sizes and shapes by directly predicting the position and shape of the target on the image or feature map [26]. The number of box prediction is reduced, the training process is simplified, and the flexibility and accuracy of detection are improved. The detection head consists of a series of convolutional and deconvolution layers that are used to generate detection results. It can accurately locate objects in the input image, predict the boundary box regression value of each anchor box and the confidence of the existence of the target, so that the model can accurately determine the location range of objects in the image [27]. The classification head uses Global Average Pooling to process each feature map, reduce the dimension of the feature map, analyze the extracted features, judge the probability that the target in the image belongs to each category, and achieve accurate classification of the target. Finally, after non-maximum suppression processing, the threshold is adjusted by adaptive adjustment. The boundary box with the highest confidence is retained, the duplicate detection box is removed, the false detection and missing detection are reduced, and the detection accuracy is effectively improved.

III. ALGORITHM IMPROVEMENT AND OPTIMIZATION

YOLOv8 mainly uses convolutional neural networks as the main means of feature extraction. The key information of the target detection task, such as the edge, texture, shape and other features of the object, is extracted from the input image, so as to carry out accurate target recognition and positioning. This process converts the original high-dimensional image data into low-dimensional feature vectors, which not only simplifies the complex data, but also reduces the amount of computation, significantly improves the training and inference efficiency of the model, and enhances the generalization ability of the model. By learning more representative and general features, Yolov8 can adapt to work on different scenarios and data sets, so that it can maintain excellent performance in different environments and achieve efficient and accurate detection results.

C2f module is a convolutional neural network module for feature extraction. By fusing feature maps from different levels, the model can make use of both details and semantic information, so as to better capture complex features in images, improve the accuracy of target detection, and retain rich gradient flow information [28]. However, stacking operation may cause the problem of channel information redundancy, and the use of a general convolution kernel may affect the detection of the receptor field, which may lead to the omission of target detection in complex scenes, especially when there is a lot of background interference, multiple detection targets or occlusion.

Spatial and Channel Reconstruction Convolution (SCConv) is a new convolutional neural network (CNN) module, which aims to improve the compression efficiency and feature representation capability of CNN by reducing the space and channel redundancy of features [29]. The module is composed of space reconstruction unit (SRU) and channel reconstruction unit (CRU). SRU supplants spatial redundancy through separation and reconstruction operations, while CRU adopts split-transform-fusion strategy to reduce channel redundancy [30].

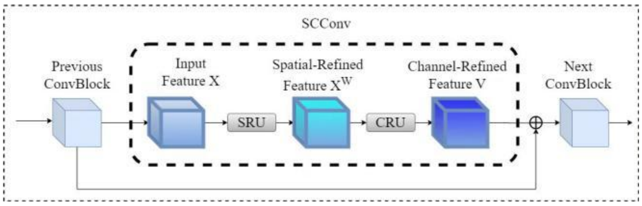

Fig.4: SCConv module structure diagram

SRU first receives input feature X, reduces the scale difference between feature maps through group normalization, then generates weights to weight features, separates features into multiple subsets, and finally generates refined spatial feature output through transformation and recombination [31]. CRU is responsible for segmentation and compression of SRU output features, using 1×1 convolution kernel to reduce the number of channels, then extracting "rich features" through GWC and PWC operations, and finally using SKNet method to adaptively merge these features to obtain channel extracted feature Y [32]. These two units work together to not only reduce redundant features, but also improve the performance of the model and the efficiency of feature characterization.

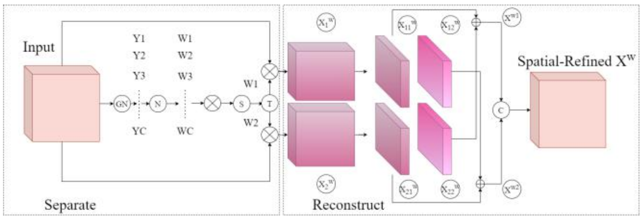

Figure 5: SRU structure

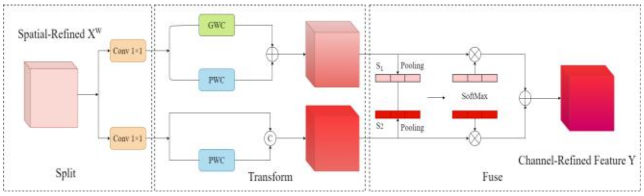

Fig.6: CRU structure diagram

SCConv as a plug-and-play architecture unit, we replace the general convolution in the Cf2 module with the SCConv module as shown in Figure 7:

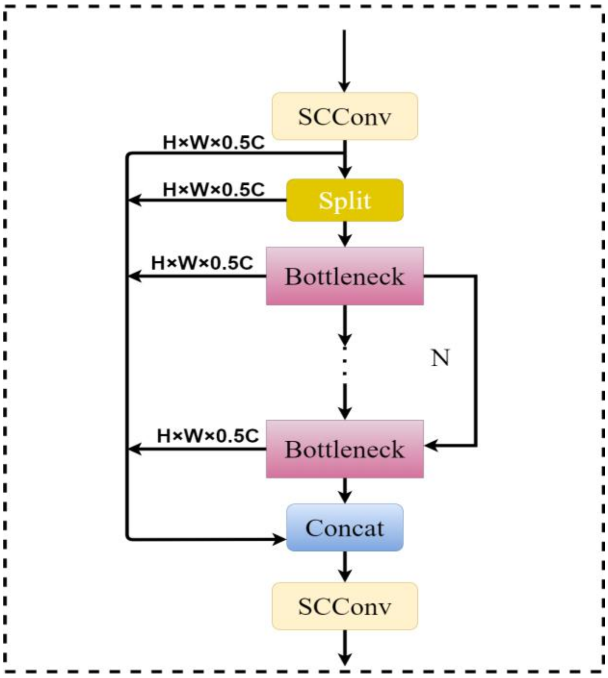

Fig.7: Improved Cf2 module structure

General convolution unified processing of all areas and features of the image, it is difficult to distinguish important differences, may pay too much attention to insignificant features and ignore welding defect related features, resulting in redundant information. SCConv, with its spatial and channel reconstruction capabilities, can extract welding defect related features in a targeted manner, highlight important information, make the network more focused on key features, and improve detection accuracy [33].

In the field of welding defect detection, general convolution has many limitations. In the face of weld defects of different types, scales and environments, it is unstable, difficult to adapt to changes, and the generalization ability is limited. For example, in actual industrial scenarios, the shape, size and ambient light conditions of welding defects are different, and general convolution cannot effectively deal with these complex and changeable situations. Moreover, the features extracted by general convolution are relatively simple, and it is difficult to capture the fine and complex structural information of weld defects. It is easy to be ignored when detecting some tiny cracks or defects hidden in complex background, resulting in the accuracy of detection is affected. The SCConv module, with its special structural design, can better adapt to various welding defects. By learning more representative and robust feature representations, the detection effect of changing weld defects is good, and the reliability of the model in practical industrial applications is improved. At the same time, SCConv can better capture key information such as shape, texture and edge of weld defects by simultaneously reconstructing space and channel dimensions, and form a more discriminating feature representation, which is conducive to accurate detection and identification of welding defects. In addition, in terms of computational efficiency, general convolution processing of high-resolution welding images requires a large amount of computation and is slow. SCConv can effectively decompose and reassemble feature maps, reduce unnecessary computation operations, improve computing resource utilization while maintaining attention to important features, and realize faster detection speed, which is conducive to real-time detection in industrial production.

IV. EXPERIMENT AND RESULT ANALYSIS

4.1 Experimental environment and experimental data

The open data set was selected from the kaggle website and the selected images were annotated by the labelimg tool. The data set contained more than 3000 images, and two types of data -good welding and bad welding - were divided into training set, verification set and prediction set in proportion. Among the selected data, welds such as edge bite, burn through, wrong edge, crater, porosity and crack are classified as bad weldinng, while the rest are good welding. Part of the data to be tested is shown in FIG.

Fig.8: part of the data image to be detected

The research model development language is Python, the deep learning frame is Pytorch1.4.0, the CUDA version is 11.18.

4.2 Experiment and result analysis

To verify the improved performance of YOLOV8 model, Precision, Recall, Average Precision (AP, Average Precision) and average precision (mAP, average precision) were selected in this experiment. mean Average Precision) as a key evaluation indicator. Among them, the accuracy rate directly reflects the correct proportion of the model prediction, which is a direct reflection of the overall accuracy. The recall rate focuses on the sensitivity of the model to positive instances and reveals the ability of the model to capture the target. By synthesizing the accuracy of different recall rates, the performance of the model under various thresholds was evaluated. The mean average accuracy provides a uniform standard for measuring the average performance of the model across all categories, thus more fully reflecting the overall effect of the model. Although these four indicators have different focuses, they are interrelated and provide strong support for the optimization of the model and the improvement of the prediction accuracy.

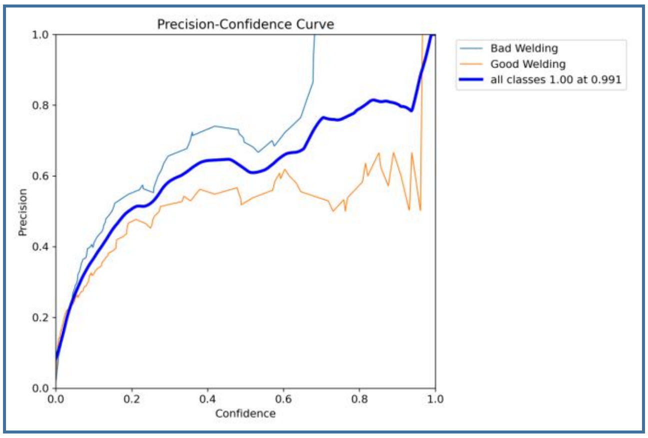

Fig.9: Feedback curve of the improved Precision-confidence curve

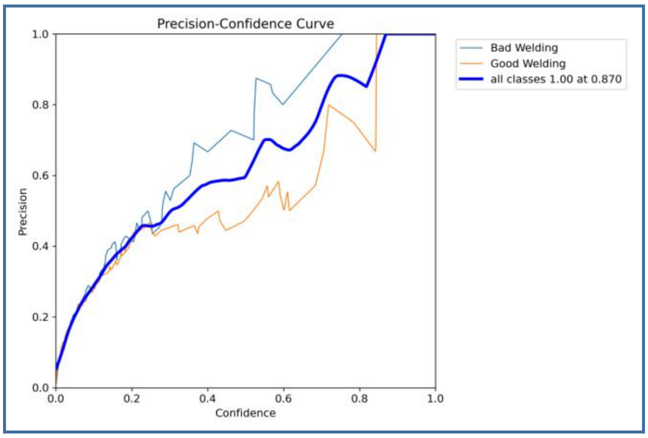

Fig.10: Precision-confidence curve Feedback curve before improvement

Table 1: Comparison of relevant experimental data

Arithmetic	Precision ratio	Recall rate	Map
YOLOv8	0.398	0.392	0.212
YOLOv8（SCConv）	0.627	0.5	0.471

V. CONCLUSION

In this experiment, we carried out the operation in strict accordance with the standard process. A suitable model environment is built to ensure the accuracy and stability of each parameter and configuration. All the required data sets are obtained legally from the Internet, and the data sets are from reliable and authoritative sources. When constructing relevant data sets, in order to reflect the diversity and representativeness of data, we selected experimental data sets with a large number of references, which could almost fully reflect various situations in practical application scenarios, and then trained the YOLOv8 detection algorithm after the introduction of SCConv based on these data sets. After data training, we evaluate the performance of the model. The results show that the precision of YOLOv8 detection algorithm after the introduction of SCConv reaches 0.627, which is indeed improved to a certain extent compared with the accuracy of 0.398 before the improvement, indicating that the introduction of SCConv can optimize the detection ability of the algorithm to a certain extent. However, although we used a large number of experimental data sets in data selection, the actual experimental results show that the improved algorithm does not perform well in welding trace detection. This may mean that there are still some problems in data processing and data selection that have not been discovered or solved. In view of this situation, in the following research, we will focus on the two perspectives of data processing and data selection, in-depth analysis of possible problems, and explore corresponding optimization strategies to further improve the performance of the algorithm in welding trace detection, so that it can better meet the needs of practical applications.

ACKNOWLEDGEMENTS

This work described in this paper was supported by The Research Funding Of GDUPT, Lagrangian Coherent Structure Analysis of Heat Sink for Heat Transfer Performance with Piezoelectric Fan, Guangdong Science and Technology Program, (Project No. 2024A0505050022), and the Research Funding of

GDUPT, Horizon Patrol - Machine Inspection Expert Driven by YOLOV8 Engine (No. 71013513124014).

REFERENCES

[1] Pan, K., Hu, H., & Gu, P. (2023). Wd-yolo: A more accurate yolo for defect detection in weld x-ray images. Sensors, 23(21), 8677.

[2] Yang, L., Fan, J., Huo, B., & Liu, Y. (2021). Inspection of welding defect based on multi-feature fusion and a convolutional network. Journal of Nondestructive Evaluation, 40, 1-11.

[3] Zhang, R., Liu, D., Bai, Q., Fu, L., Hu, J., & Song, J. (2024). Research on X-ray weld seam defect detection and size measurement method based on neural network self-optimization. Engineering Applications of Artificial Intelligence, 133, 108045.

[4] Li, L., Ren, J., Wang, P., Lü, Z., Di, R., Li, X., ... & Zhao, X. (2023). Defect detection method for high-resolution weld based on wandering Gaussian and multi-feature enhancement fusion. Mechanical Systems and Signal Processing, 199, 110484.

[5] Droubi, M. G., Faisal, N. H., Orr, F., Steel, J. A., & El-Shaib, M. (2017). Acoustic emission method for defect detection and identification in carbon steel welded joints. Journal of constructional steel research, 134, 28-37.

[6] Cheng, Y., Bai, L., Yang, F., Chen, Y., Jiang, S., & Yin, C. (2016). Stainless steel weld defect detection using pulsed inductive thermography. IEEE Transactions on Applied Superconductivity, 26(7), 1-4.

[7] Bebiano, D., & Alfaro, S. C. (2009). A weld defects detection system based on a spectrometer. Sensors, 9(4), 2851-2861.

[8] Zhang, H., Chen, Z., Zhang, C., Xi, J., & Le, X. (2019, August). Weld defect detection based on deep learning method. In 2019 IEEE 15th international conference on automation science and engineering (CASE) (pp. 1574-1579). IEEE.

[9] Zhu, B., Wang, Z., Ma, Y., Gao, W., & Wang, S. (2024). Welding defects extraction method by fusing saliency information of mid-level and underlying level images. Heliyon, 10(20).

[10] Liu, Y., Yuan, K., Li, T., & Li, S. (2023). NDT Method for Weld Defects Based on FMPVit Transformer Model. IEEE Access, 11, 61390-61400.

[11] Oh, S. J., Jung, M. J., Lim, C., & Shin, S. C. (2020). Automatic detection of welding defects using faster R-CNN. Applied Sciences, 10(23), 8629.

[12] Redmon, J. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition.

[13] Shang, L. P., Chang, Y. Z., Chen, Y., & Ou, Y. S. (2024). Fire detection algorithm based on the fusion of YOLOv8 and Deformable Conv DCN. International Journal of Advanced Engineering Research and Science, 11(8), 8-15.

[14] Hatab, M., Malekmohamadi, H., & Amira, A. (2021). Surface defect detection using YOLO network. In Intelligent Systems and Applications: Proceedings of the 2020 Intelligent Systems Conference (IntelliSys) Volume 1 (pp. 505-515). Springer International Publishing.

[15] Zhao, C., Shu, X., Yan, X., Zuo, X., & Zhu, F. (2023). RDD-YOLO: A modified YOLO for detection of steel surface defects. Measurement, 214, 112776.

[16] Yuan, M., Zhou, Y., Ren, X., Zhi, H., Zhang, J., & Chen, H. (2024). YOLO-HMC: An improved method for PCB surface defect detection. IEEE Transactions on Instrumentation and Measurement.

[17] Gao, A., Fan, Z., Li, A., Le, Q., Wu, D., & Du, F. (2023). YOLO-Weld: A Modified YOLOv5-Based Weld Feature Detection Network for Extreme Weld Noise. Sensors, 23(12), 5640.

[18] Song, L., Kang, J., Zhang, Q., & Wang, S. (2023). A weld feature points detection method based on improved YOLO for welding robots in strong noise environment. Signal, Image and Video Processing, 17(5), 1801-1809.

[19] Shang, L., Chang, Y., Chen, Y., Yu, L., Zeng, Z., & Chen, J. (2024, October). Research on YOLOv8 Detection Algorithm Based on SKAttention Fusion. In 2024 5th International Conference on Machine Learning and Computer Application (ICMLCA) (pp. 571-574). IEEE.

[20] Liu, L. J., Zhang, Y., & Karimi, H. R. (2024). Resilient machine learning for steel surface defect detection based on lightweight convolution. The International Journal of Advanced Manufacturing Technology, 134(9), 4639-4650.

[21] Trinh, D. C., Mac, A. T., Dang, K. G., Nguyen, H. T., Nguyen, H. T., & Bui, T. D. (2024). Alpha-EIOU-YOLOv8: an improved algorithm for rice leaf disease detection. AgriEngineering, 6(1), 302-317.

[22] Ren, J., Zhang, H., & Yue, M. (2025). YOLOv8-WD: Deep Learning-Based Detection of Defects in Automotive Brake Joint Laser Welds. Applied Sciences, 15(3), 1184.

[23] Hua, G., Zhang, H., Huang, C., Pan, M., Yan, J., & Zhao, H. (2024). An enhanced YOLOv8-based bolt detection algorithm for transmission line. IET Generation, Transmission & Distribution, 18(24), 4065-4077.

[24] Lu, J., Lee, S. H., Kim, I. W., Kim, W. J., & Lee, M. S. (2023). Small foreign object detection in automated sugar dispensing processes based on lightweight deep learning networks. Electronics, 12(22), 4621.

[25] Li, Z., Luo, S., Xiang, J., Chen, Y., & Luo, Q. (2024). Improved Chinese Giant Salamander Parental Care Behavior Detection Based on YOLOv8. Animals, 14(14), 2089.

[26] Jiang, T., Du, X., Zhang, N., Sun, X., Li, X., Tian, S., & Liang, Q. (2024). YOLOv8-GO: A Lightweight Model for Prompt Detection of Foliar Maize Diseases. Applied Sciences, 14(21), 10004.

[27] Yao, L., Zhao, B., Wang, X., Mei, S., & Chi, Y. (2024). A Detection Algorithm for Surface Defects of Printed Circuit Board Based on Improved YOLOv8. IEEE Access.

[28] Wang, Z., Chen, J., Yu, P., Feng, B., & Feng, D. (2024). SC-YOLOv8 Network with Soft-Pooling and Attention for Elevator Passenger Detection. Applied Sciences, 14(8), 3321.

[29] Ma, J., Tong, J., Xue, M., & Yao, J. (2024). A lightweight model for echo trace detection in echograms based on improved YOLOv8. Scientific Reports, 14(1), 1-17.

[30] Xiao, L., & Zhang, H. (2025, January). Improved drainage pipe defect detection algorithm based on YOLOv8. In The International Conference Optoelectronic Information and Optical Engineering (OIOE2024) (Vol. 13513, pp. 1025-1030). SPIE.

[31] Shi, Y., Qing, S., Zhao, L., Wang, F., Yuwen, X., & Qu, M. (2024). Yolo-peach: a high-performance lightweight yolov8s-based model for accurate recognition and enumeration of peach seedling fruits. Agronomy, 14(8), 1628.

[32] Lei, L., Duan, R., Yang, F., & Xu, L. (2024). Low Complexity Forest Fire Detection Based on Improved YOLOv8 Network. Forests, 15(9), 1652.