Dean&Francis
ISSN 2959-6157

Performance Comparison and Analysis of YOLOv8/v9/v10 in Steel Plate Surface Defect Detection

He Chen[1,*]

and Jun Tang[2]

[1]School of Mathematics and Statistics, Beijing Jiaotong University, Beijing, China
[2]International School of Information Science & Engineering, Dalian University of Technology, Dalian, China
*Corresponding author: 23271030@bjtu.edu.cn

Abstract:
During steel plates' continuous casting and rolling, surface defects often arise from rolling equipment/processes, risking production accidents. These defects are small, easily occluded, and require real-time detection for efficiency—demanding high accuracy and speed from detection algorithms. This paper synthesizes relevant research papers on the application of You Only Look Once version 8, version 9, version 10 models (YOLOv8, v9, v10 models), and their improved versions in steel surface quality inspection. It analyzes and summarizes the technological evolution path in the field of industrial quality inspection, with a key focus on the application and performance of the YOLO series models. Studies indicate that single-stage detection models have gained significant attention due to their advantage in balancing speed and accuracy. Among them, YOLO models and their improved variants have become the industry mainstream through a series of enhancements, such as architecture decoupling and feature compression. Specifically, YOLOv10 and its improved models, benefiting from innovations in non-NMS inference, dynamic gate-controlled feature fusion, shape awareness, and loss functions, have achieved a detection accuracy of 85.5% mAP@0.5 on datasets like NEU-DET. This study will further summarize the limitations of current technologies in small-sample dense defect recognition and propose development directions centered on multi-modal fusion and adaptive computing.

Keywords: Industrial Defect Detection; Detection Model; YOLOv10; Detection Accuracy.

1. Introduction

Steel plates, as the core foundational material of the industrial system, are widely used in key fields such as mechanical manufacturing, construction engineering, shipbuilding, and energy equipment. Their surface defects (such as cracks, inclusions, oxide scales, scratches, etc.) directly affect the structural safety and service life of the products - if minor cracks are not detected in time, they may expand during subsequent processing or service, causing equipment failures, engineering collapses, and even safety accidents. Therefore, the detection of surface defects on steel plates is an indispensable key link in the industrial production quality control system. With the transformation of manufacturing towards intelligence and high-speed, modern steel plate production lines have reached a speed of 2-5m/s. At the same time, the steel plate materials have become diverse (such as high-strength alloy steel plates, ultra-thin cold-rolled steel plates) and the defect forms have become more complex (multi-scale, low contrast, sparse distribution), which puts higher requirements on the real-time performance, accuracy, and robustness of the detection technology. Traditional manual detection methods cannot meet the actual production needs, so in-depth research on the steel plate surface defect detection system has become the consensus of today's steel enterprises. The surface defect detection method based on machine vision has the characteristics of high accuracy, fast processing speed, and intelligent processing, and is the main development trend of surface defect detection [1, 2].

Early industrial quality inspection mainly relied on physical sensing methods, such as eddy current detection and laser scanning, although they could achieve micro-level defect recognition, the equipment deployment cost was high (a single set of laser detection system cost over one million yuan), and they were sensitive to environmental temperature, humidity, and the cleanliness of the steel plate surface. In complex industrial sites, they were prone to false detection and missed detection; traditional machine learning methods in the manual feature stage (such as texture analysis combined with SVM, random forest) had the advantage of low computing power requirements, but feature extraction relied on manual experience, and in the face of diverse defect types and dynamic changing industrial environments, their generalization ability was seriously insufficient, with an error rate of over 15%, making it difficult to adapt to modern production lines.

The rise of deep learning technology has driven the transformation of the industrial quality inspection paradigm. Various detection models have gradually been applied to the steel plate defect detection scenarios, but the performance differences among different models are significant. The machine vision-based methods also have limitations. Deep learning, especially convolutional neural networks (CNNs), has shown promise, but it also has limitations in industrial detection. Two-stage detection models (such as Mask R-CNN) can achieve high-precision defect localization, but their dual-stage architecture leads to redundant computation, and the inference delay is generally over 100ms per image, far from meeting the real-time requirements of high-speed production lines (<30ms per image); Transformer series models (such as DETR) have strong global feature modeling capabilities and perform well in complex backgrounds, but they require an extremely large-scale dataset (usually requiring tens of thousands of labeled samples) to converge, and their inference speed is more than 30% lower than the YOLO series, making it difficult to adapt to small sample scenarios and high-speed detection requirements in industrial sites [3].

The single-stage YOLO series models have achieved a balance between speed and accuracy through end-to-end architecture optimization - from YOLOv1 to v8, the models have improved through Anchor-Free design, feature pyramid optimization, etc., with the inference speed increasing by 46% while maintaining stable accuracy; the released YOLOv10 model in 2025 further broke through the efficiency bottleneck, through innovations such as no NMS inference and dynamic feature fusion, achieving dual breakthroughs in lightweighting and detection performance, becoming the current research hotspot in industrial quality inspection. This paper reviews the evolution path of industrial defect detection technology, compares the architectural innovations and performance of the YOLOv8, v9, and v10 series models, analyzes the current technical bottlenecks, and provides theoretical references and practical guidance for the selection of detection models and technical optimization in industrial scenarios.

2. Evolution of Industrial Defect Detection Technology

2.1 Requirements for Industrial Defect Detection

Generally, there are six core challenges to steel surface defects, as shown in table 1 below.

ISSN 2959-6157

Table 1. Three Scheme comparing

Challenges Numble	Description	Cause Analysis	Requirements
1	Complex,multi-scale and sparsely distributed defect morphology	Diverse defect types (scratches, blisters, cracks, folds, scale, rolled-in scale, etc.) with vastly different morphologies; obvious multi-scale issue	Requires the model has strong multi-scale feature extraction capabilities
2	High detection accuracy requirements (low tolerance for missed detections)	Missing tiny cracks may lead to product scrapping or equipment accidents	Aim for near-zero missed detection; low sensitivity to confidence thresholds
3	High real-time requirements	Steel plates move at high speed; inspection must keep pace with production rhythm	High FPS inference capability (>30 FPS, partly >100 FPS)
4	Complex background noise and significant illumination variation	Complex lighting conditions on the industrial site; some defects have low contrast with the background	Model requires strong robustness; feature extraction network must capture fine details
5	Limited dataset scale	Low defect ratio makes collection difficult; high labeling costs	Effective training on small datasets; support for data augmentation
6	Stable and portable industrial deployment requirements	Need deployment on embedded devices or edge servers	Requires the model to be lightweight and easy to deploy; strong hardware adaptability

The YOLO series becomes an optimal solution for industrial scenarios with features such as the multi-scale feature pyramid, dynamic gate control filtering, lightweight deployment and so on. Specifically, by constructing a multi-scale feature pyramid structure, the YOLO series is capable to simultaneously capture the feature information of both large size defects (such as large area oxide scale) and small defects (such as fine cracks) on steel surface, which precisely matches the “multi-scale defect detection” requirement. Its dynamic gating mechanism can adaptively filter key defect features during fusion, effectively filters out interference from complex background noise and lights variations in industrial environments, thereby enhancing model robustness [4, 5]. In terms of lightweight design, the number of parameters and computational overhead are continuously reduced, during the iterations from YOLOv8 to v10, through improvements like architecture decoupling (e.g., Dual-branch decoupling head) and feature compression (e.g., CIB feature compression block). For example, the number of parameters of YOLOv10 can be compressed to 2.67M, enabling stable deployment on edge devices like Jetson Nano to meet the dual requirements of “low computational power adaptation” and “high real-time performance (>30 FPS)” in industrial scenarios. Simultaneously, the YOLO series supports flexible data augmentation strategies (e.g., turning off Mosaic augmentation to suppress noise at the end of training) [6], which can enhance the generalization ability and training effect of the model under the condition of limited steel defect dataset size, further adapting to the practical data conditions of industrial quality inspection.

2.2 Logic of Paradigm Transition

As deep learning becomes the mainstream solution for industrial quality inspection, the performance of various detection models in practical applications varies significantly. Especially in steel defect detection scenarios with high requirements for real-time performance and high robustness, algorithm selection requires a strict trade-off between accuracy, speed, deployment cost, and generalization ability. Industrial detection methods have undergone a three-stage leap.

First, it starts from the physical sensing stage, relying on specialized equipment (e.g., infrared thermal imagers) for micron-level defect identification, but with strong environmental sensitivity and high deployment costs. Then moving to the hand-crafted feature stage, it employs traditional machine learning methods like texture analysis + SVM, relying on experiential knowledge for feature design and limited generalization, while offering interpretability advantages. To the current deep learning stage, single-stage and two-stage detection models that apply deep learning have emerged. Among them, two-stage models are represented by Fast R-CNN and Msk R-CNN, which have high detection and positioning accuracy but are computationally heavy and low. Single-stage models represented by the YOLO series, which achieve spped-accuracy balance through end-to-end training, have become the mainstream solution for real-time quality inspection.

Table 2 compares the practical performance of several current mainstream algorithm types in steel quality inspection, highlighting their core challenges and applicable scenarios in industrial production line environments.

Table 2. Performance of Mainstream Algorithms in Steel Plate Quality Inspection

Algorithm Types	Advantages	Disadvantages	Actual Performance
DETR/Transformer series	Strong label feature modeling ability; More robust performance in complex backgrounds	It requires an extremly largr amount of data to converage well, and small object detection is still not as easy to tune as YOLO. Inference is slower than YOLO	It is suitable for offline high-precision quality inspection or scientific research scenarios, but not for high-spped real-time quality inspection on industrial production
Traditional image processing + SVM/ Random Forest	Easy to implement; Low computing power requirements	Manul feature extraction is difficult to adapt to defect diversity. Light and material variations can lead to high false detection rates.	It was used when the surface texture of th steel plate was monotonous and there were few types of defects, but it has been phased out on modern complex steel plate production lines.
Semantic segmentation calss(U-Net, SegFrmer)	Fine segmentation of defect edges, suitable for subsequent size/shape annalysis	Inference is slower than object detectionand requires denser labeling(which is costing).	It is suitable for defect post-progressing(such as precisly measuring crack length), but is typically used as secondary analysis after YOLO detection.
Faster R-CNN	High precision,especially friendly for small object detection, and good at distinguishing complex backgrounds	Slow inference speed: HIgh latency due to the two-stage strucure Complex deployment: High cost of optimization on embedded devices orindustrial camera ports, high computing power requirements	Faster R-CNN is fine if the detection is offline and the speed requirementare not strict, but it is almost obsolete in real-time scenarios
RetinaNet	Single-stage detection with precision close to Fast R-CNN; The Focal Loss has an advantage over imbalanced samples	Moderate inference speed, less lightweight than the YOLO version, less sensitive to particularly small defects than YOLOv8/v10,fewer industrial deployment cases, and less experince in community optimization and migration	It is suitable for semi-real-time detection, but it can not compete with the optimized YOLO series in steel plate inspection.

2.3 Main Iteration Path of the YOLO Series

The improvement of YOLO models has primarily revolved around the core industrial needs of "accuracy-speed-deployment cost" as illustrated in fig. 1, forming distinct technical thread at different iteration stages.

The early v5-v8 series achieved a 46% increase in inference spped by transitioning from Anchor-based to Anchoe-free design, significantly simplifying the inspection process and reducing Anchor frame redundancy computation. To enhance the ability of complex defect recognition, some improved models (e.g., versions incorporating the Global Attention Mechanism (GAM)) effectively increased the detection rate of occluded defects by deepening the depth of feature fusion, but at the cost of a 70% increase in computation, which highlights the "accuracy vs. computational consumption" dilemma [7]. In the lightweight direction, models like LSKA-YOLOv6 compressed parameters by up to 42% through large kernel decomposition convolution techniques. While it lowered the threshold for edge device deployment, it led to a 1.8% increase in the missed detection rate of large defects due to loss of deep features, exposing the mismatch between lightweighting and detection integrity [8]. It wasn't until YOLOv10, which innovatively used dual-branch decoupled heads to reduce redundant computation and combined CIB blocks to optimize feature processing efficiency, that finally achieved a 40% overall efficiency improvement [4], striking a better balance between accuracy, speed, and lightweight deployment, further aligning with the stringent requirements of industrial production lines for detection models.

ISSN 2959-6157

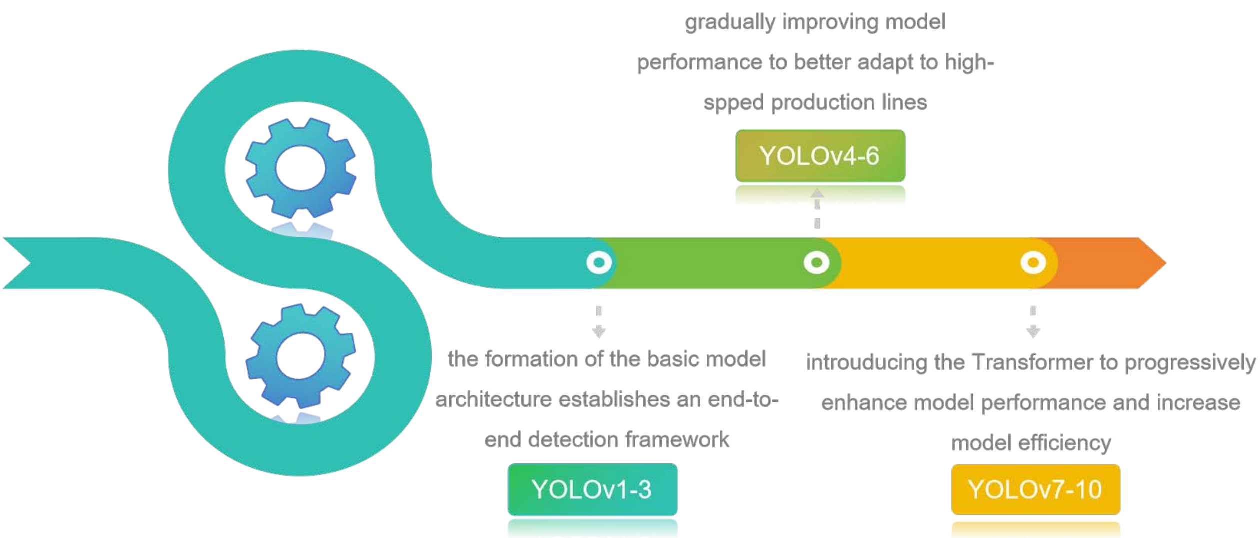

Fig. 1 Evolution process of YOLO models

3. Core Innovations and Paradigm Breakthroughs of YOLOv10

3.1 Architectural Refactoring Strategies

YOLOv10 achieves multi-dimensional innovation at the architectural level, breaking through the limitations of traditional models by refactoring the detection pipeline and feature processing mechanisms. Among them, the NMS-free inference mechanism is a key breakthrough. This mechanism dismissed the non-maximum suppression (NMS) post-processing that traditional detection models rely on and irectly achieves precise screening of bounding boxes through a one-to-many label assignment strategy, reducing inference latency by 40% in the process and significantly improving real-time detection efficiency. In terms of dynamic feature fusion, the innovatively designed SlimFusionCSP module combines the gating mechanism with depthwise separable convolution. During the feature fusion progress, it can adaptively screen the effective features that are critical for defect detection while filtering out 70% of redundant feature information, ensuring feature validity while reducing computational overhead. To address the challenge of identifying low-contrast defects on steel surfaces (e.g., fine scratches, faint rolled-in scale), the DualConv parallel convolution structure is employed to enhance local perception capability and texture detail features are extracted in parallel through multi-scale convolution kernels, significantly improving the recognition rate of such difficult-to-detect defects and further refining the adaptability of the architecture to defects in complex industrial scenarios[4].

3.2 Industrial Adaptability Optimaization

YOLOv10 focuses on the actual needs of industrial quality inspection, conducting targeted optimizations in loss functions, data augmentation, and deployment adaptation to strengthen the industrial application abilities of the model. At the loss function level, Shape-IoU is innovatively introduced and geometric constraint terms are added. For irregularly shaped defects like scale and folding, geometric features are used to assist in optimizing the bounding box regression accuracy, reducing the localization error of such defects by 30% and improving the accuracy of the detection result [4, 5]. Regarding data augmentation strategy, a dynamic adjustment mechanism is adopted. Mosaic augmentation is used in the early training stages to expand sample diversity, but is turned off in the later stages. This effectively suppresses noise interference caused by Mosaic splicing, preventing the model from overfitting to noise features and significantly enhancing the model's generalization in real industrial scenarios [8]. For edge computing adaptation, through architectural simplification and parameter optimization, the model parameter count is compressed to 2.67M, a 13% reduction compared to YOLOv8n. Inference latency on common industrial edge devices like Jetson Nano can be controlled within 20ms, meeting the dual requirements of device deployment cost and real-time performance for high-speed production lines, laying the foundation for the model's widespread application in industrial settings [4].

4. Core Innovations and Paradigm Breakthroughs of YOLOv10

4.1 Industrial Adaptability Optimization

In the collection of the performance of different models from the reviewed papers, this paper compares test data on the NEU-DET dataset, commonly used across these studies, in terms of mAP, parameter count, etc. Here, mAP (mean Average Precision) is a common comprehensive evaluation metric in object detection, reflecting the average detection accuracy of the model under different confidence thresholds. A higher value indicates better model performance. Parameter count is used to measure model complexity. Models with fewer parameters typically have a lightweight advantage and are suitable for scenarios with limited computational resources.

NEU-DET is a public image dataset specifically for surface defect detection, published by Northeastern University. It focuses on six common surface defects in cold-rolled steel strips: cracks, inclusion, patches, pitted surface, rolled oxide scale, and scratches. Each defect category contains 300 images, totaling 1800 grayscale images.

Table 3 summarizes the performance of mainstream models on the NEU-DET dataset.

Table 3. Performance of Mainstream Models on the NEU-DET Dataset

Model	mAP@0.5	Parameter count(M)	Key Advantages
YOLOv10n-SFDC	85.5%	2.67	Inclusions defect mAP increases 10.9%
YOLOv9-CBAM	88.0%	25.5	Scale positioning accuracy increases 15%
ELA-YOLO	81,7%	5.4	Crack recall rate increases 12%
LSKA-YOLOv8	84.9%	2.2	Small target missed detection rate decreases 8%

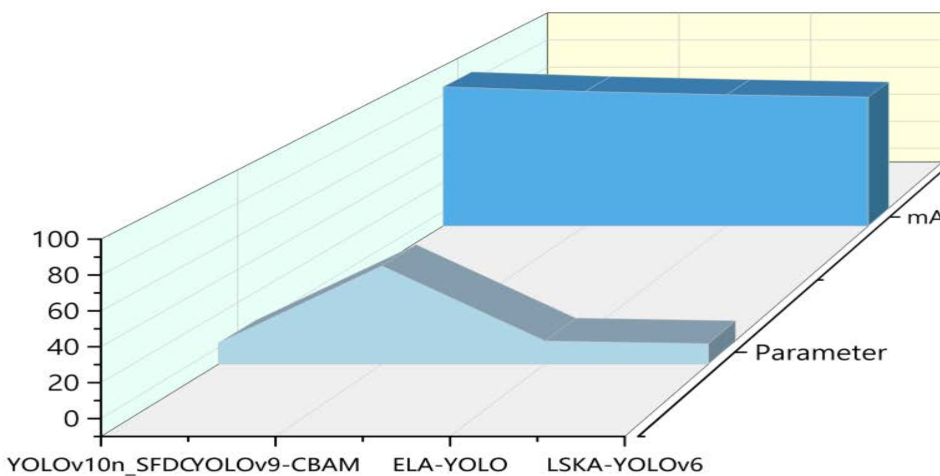

Fig. 2 Performance of Mainstream Models on the NEU-DET Dataset

The core breakthroughs of YOLOv10n-SFDC are mainly reflected in three aspects. First, it excels in balancing lightweight design and accuracy, with only 48% of the parameters of ELA-YOLO's [4, 8]. Second, the model demonstrates good adaptability to multi-defect scenarios by introducing a gating mechanism to dynamically adjust feature flow. As shown in fig. 2, it effectively reduced the missed detection rate by 15% on the NEU-DET dataset [4]. Finally, the robustness for small object detection is significantly enhanced, utilizing the DualConv module increased the recall rate for inclusion defects smaller than 10px by 10.9% [4].

4.2 Industrial Adaptability Optimization

YOLOv10n-SFDC still faces several key bottlenecks. Firstly, it performs poorly in scene migration, particularly on welding defect datasets where mAP drops significantly to 69.3%, indicating insufficient cross-domain generalization ability. Secondly, the model's capability for detecting dense defects is limited. When the target overlap ratio exceeds 60%, the recall rate drops by up to 12%, exposing issues of feature confusion and missed detection. Furthermore, there is a noticeable boundary in real-time performance. Affected by the computational overhead of the

ISSN 2959-6157

SlimFusionCSP module, FPS decreases by approximately 10%, indicating that the balance between computational power and accuracy still requires to be further optimized.

5. Future Directions and Development Challenges

5.1 Deployment Challenges in Industrial Applications

In the future development of industrial quality inspection applications, detection models such as YOLOv10n-SFDC still face several severe deployment challenges. Currently, the computational power limitation of edge devices is a significant bottleneck. For instance, larger models like YOLOv10-X have an inference delay of over 50ms on embedded platforms such as Jetson Nano, which is insufficient for high real-time requirements and requires hardware collaboration design and model compression for deeper optimization [9]. Additionally, the existing detection systems lack an efficient real-time interaction mechanism with the manufacturing execution system (MES), preventing the rapid feedback of detection results to the production control process, thereby affecting the response efficiency of the quality inspection loop [10]. Moreover, in the face of newly emerging defect types in the production line, existing models rely on re-training, and their transfer learning and zero-sample adaptation capabilities are not yet mature, severely restricting the application of models in actual dynamic environments.

5.2 Key Technology Breakthrough Points

Introducing multimodal perception fusion technology, combined with the semantic prior of large language models (LLM), enables defect detection based on text descriptions (such as „locating metal fatigue cracks"), thus breaking through the limitation of fixed categories and enhancing the understanding and recognition capabilities for unknown defects. For the edge computing bottleneck, reconfigurable dynamic computing architectures are being developed, which can adaptively allocate computing resources based on the complexity of defect, significantly improving inference efficiency while maintaining accuracy [6]. At the same time, a synthetic data engine based on generative adversarial networks (GAN) can generate a large number of high-quality samples for industrial defects, effectively alleviating the problems of data scarcity and distribution bias, and providing a foundation for model transfer in zero-sample and few-sample scenarios [8]. These technological advancements are expected to systematically improve the real-time performance, adaptability, and integration of existing detection system, driving visual quality inspection towards a more intelligent and flexible new stage.

6. Conclusion

Industrial surface defect detection technology has undergone a fundamental transformation from physical sensing, manual features to the deep learning paradigm. This study systematically compared and analyzed the performance evolution and innovative breakthroughs of the YOLOv8, v9, and the latest v10 series models in the task of detecting surface defects on steel plates. The YOLO series models have become an important component in the field of industrial defect detection, and their continuous evolution has significantly promoted the development of real-time quality inspection technology. In the future, by integrating cognitive intelligence methods and adaptive computing technologies, it is expected to achieve a more intelligent, flexible, and robust next-generation industrial quality inspection system.

The YOLO series has achieved breakthroughs in lightweighting and multi-defect adaptability through architecture decoupling and feature compression mechanisms, establishing a new standard for real-time quality inspection. YOLO and its improved models continue to make breakthroughs in various aspects, such as YOLOv9-E/C for high-precision scenarios, YOLOv8n improved version for lightweight edge deployment, and YOLOv10 specializing in small target defects. Through multi-scale feature fusion mechanisms (such as combining global and local features to deal with weak defects), architecture optimization (such as ADown operator, dynamic gated fusion), and targeted loss function design, significant breakthroughs have been achieved in lightweighting, multi-defect adaptability, and small target detection capabilities, laying a new standard for real-time online quality inspection. However, cross-scenario generalization and dense defect detection remain key bottlenecks. In the future, it is necessary to promote the evolution of industrial quality inspection systems towards cognitive intelligence through the integration of visual-language models, adaptive computing architectures, and synthetic data engines.

Authors Contribution

All the authors contributed equally and their names were listed in alphabetical order.

References

[1] Yang L, Huang X, Ren Y, Han Q, Huang Y. Steel plate surface defect classification technology based on image enhancement and combination feature extraction. Engineering

Computations, 40(6), 1305–1329. 2023. https://doi.org/10.1108/ec-12-2022-0720
[2] Xue B, Wu Z. Key Technologies of steel plate surface defect detection system based on artificial intelligence machine vision. Wireless Communications and Mobile Computing, 2021 (1). https://doi.org/10.1155/2021/5553470
[3] Hu X, Lin S. DFFNet: a lightweight approach for efficient feature-optimized fusion in steel strip surface defect detection. Complex & Intelligent Systems, 10(5), 6705–6723. 2024. https://doi.org/10.1007/s40747-024-01512-1
[4] Liao, L F, et al. A novel YOLOv10-based algorithm for accurate steel surface defect detection.» Sensors 25.3 2025: 769.
[5] Ma Ruichen, et al. ELA-YOLO: An efficient method with linear attention for steel surface defect detection during manufacturing. Advanced Engineering Informatics 65 2025: 103377.
[6] Li Hanlin, et al. Steel surface defect detection based on multi-layer fusion networks. Scientific Reports 15.1 2025: 10371.
[7] Huang Fan, et al. «A comparative study of YOLOv9 and YOLOv10 for steel surface defect detection.» 2024 30th International Conference on Mechatronics and Machine Vision in Practice (M2VIP). IEEE, 2024.
[8] Tie Jun, et al. LSKA-YOLOv8: A lightweight steel surface defect detection algorithm based on YOLOv8 improvement. Alexandria Engineering Journal 109 2024: 201-212.
[9] Chen Cong, Hoileong Lee, and Ming Chen. Steel surface defect detection method based on improved YOLOv9. Scientific Reports 15.1 2025: 25098.
[10] Wang, Yuanyuan, et al. A steel defect detection method based on edge feature extraction via the Sobel operator. Scientific Reports 14.1 2024: 27694.