FMDB *Transactions on Sustainable Structures and Materials*

Steel Surface Defect Detection Using YOLOv8 with Shuffle Attention

S. Ajay Karthick[1], U. R. Sudarsananarayanan[2], I. I. Shrinaath[3], B. Deepa[4,*], Dilli Kasi Rao Kotha[5]

[1,2,3]Department of Computer Science and Engineering, SRM Institute of Science and Technology, Ramapuram, Chennai, Tamil Nadu, India.
[4]Department of Electronic and Communication Engineering, SRM Institute of Science and Technology, Ramapuram, Chennai, Tamil Nadu, India.
[5]Faculty of Engineering, Environment and Computing, Coventry University, Coventry, England, United Kingdom.
ajaykarthick1207@gmail.com[1], sudarsananarayanan003@gmail.com[2], shrinaath.reddy@gmail.com[3], deepab@srmist.edu.in[4], kothad@coventry.ac.uk[5]

Abstract: Defect detection on steel surfaces presents an important feature of Industry 4.0 and intelligent manufacturing, providing building safety, product quality assurance, and reducing production waste. Original quality control processes involve a lot of manual visual checking. Even the human visual approach requires too many resources, is highly personal, and does not allow one to keep up with the huge amount of data that has been generated by modern continuous steel rolling mills. Deep learning techniques, particularly efficient one-stage detectors such as those in the YOLO family, have been shown to cope with real-time defect detection tasks. At the same time, the practical execution of these algorithms, which are used to detect surface defects under the conditions of a chaotic industrial environment, faces some technical challenges. Particularly, when the problem is characterized by very low visual contrast, elongated and slender defects, very high inter-class similarity, and widespread texture noise, this results in relatively high false negative detection rates and poor localization capabilities of the conventional detectors. Therefore, to overcome the deficiencies of the current detectors, this paper proposes a new attention-enabled defect detection model based on a strong YOLOv8 baseline. The main contribution is that the architecture significantly enhances accurate localization and classification, while strictly maintaining the low interference needed by real-time feeds.

Keywords: Deep Learning; Intelligent Manufacturing; Quality Assurance; Low Visual Contrast; Texture Noise; Industrial Environment; Defect Detection; Steel Surfaces; Building Safety.

Received on: 17/07/2025, **Revised on:** 30/10/2025, **Accepted on:** 09/11/2025, **Published on:** 07/10/2026

Journal Homepage: https://www.fmdbpub.com/user/journals/details/FTSSM

DOI: https://doi.org/10.69888/FTSSM.2026.000815

Cite as: S. A. Karthick, U. R. Sudarsananarayanan, I. I. Shrinaath, B. Deepa, and D. K. R. Kotha, "Steel Surface Defect Detection Using YOLOv8 with Shuffle Attention," *FMDB Transactions on Sustainable Structures and Materials,* vol. 2, no. 2, pp. 88–96, 2026.

Copyright © 2026 S. A. Karthick *et al.,* licensed to Fernando Martins De Bulhão (FMDB) Publishing Company. This is an open access article distributed under CC BY-NC-SA 4.0, which permits copying, remixing, redistributing, transformation, and building upon the material in any medium so long as the original work is properly cited.

1. Introduction

1.1. Background and Industrial Context

*Corresponding author.

According to your write-up on vendor-managed inventory, its model, having interesting implications for productivity, can lead to breakthrough innovations in production provided stakeholders are ready to change, and product quality and local suppliers are available. The supplier-managed inventory model is a more recent model, and they are certainly not free from issues; their heavy dependence on local vendors and quality levels has numerous limitations that they face. The Vendor Managed Inventory Model has been incorporated into manufacturing processes recently. As with many contemporary industry advances, it largely comes from Japan's manufacturing revolution [1]. This type of manufacturing is higher up on the supply chain complexity spectrum. This is due to the vendor-managed inventory model that engages an external vendor, local suppliers, and in-house suppliers as one [20]. The mere mention of zero defects raises eyebrows, as zero defects seem an impossibility. However, the reality is that zero defects are quite possible and do not even involve much fat expense.

1.2. Deep Learning Models Gradually Replaced Traditional Methods

As the limitations of manual inspection became clear. Convolutional neural networks see an image as pixels and not hand-crafted features and capture the pattern [2]. In the domain of object detection. The manufacturing industry prefers single-stage detectors because of their speed and efficiency. They perform so well for one of the reasons. They process detection steps simultaneously, which can effectively shorten time [3]. Despite the many designs available, the industry standard for instant results is the YOLO series, meaning you only look once [4].

1.3. Specific Challenges and Proposed Solution

Using YOLO models to detect defects on steel surfaces is very difficult due to those surfaces' optical and morphological characteristics [5]:

- The defect and healthy metallic background have a very low contrast gradient [6].
- **Scale Variance:** Scale Variance refers to very small strip-like trace defects (micro scratches) with a very small areal occupation fraction of pixel counts [7].
- **Inter-class Similarity:** Similar-looking textures appearing in completely different defect classes are indicative of inter-class similarity [8].
- **Complex Texture Noise:** The complex texture noise is generated by the interactions of factory lighting and the rolling process. These add cluttered and high-frequency background textures over the steel plate [9].

A normal convolution has an effective receptive field that is relatively small, which means that the model is less likely to focus on the discriminative ones [10]. To avoid the massive computational complexity caused by the problem, researchers propose an optimized detector based on YOLOv8, specifically modified with a ShuffleAttention module. This is a plug-and-play module of lightweight nature that integrates channel and spatial attention for improving feature selectivity at a low computational cost [11].

2. Related Work and Literature Survey

Traditional Inspection Methodologies: A distinct approach based on the RGB colour space framework. A method was proposed for detecting defects in a metallic serrated plate. A plate that appears within a realistic context. Accordingly, RGB images of the target will be taken with a camera. The coefficient of similarity in the RGB color space will be calculated. Afterward, the order of the images will be sorted. A chromatic-based algorithm will be used to distinguish backgrounds. In addition, the chromatic saliency map will be visualized. Dispersion shall employ an anomaly detection algorithm. A stable principal component pursuit algorithm will be used. The analysis uses the shape of the feature. The saliency map will be used to determine the scope of area defects [12].

Convolutional Neural Networks and Two-Stage Detectors: CNNs eliminated the need for feature engineering. At first, all the CNN-based adaptations began to use two-stage detectors, specifically the RCNN family (Faster RCNN) [13]. These models utilize a Region Proposal Network (RPN) to produce region proposals. The Faster RCNN model has extremely accurate localization results, but the computational speed of iterating through thousands of region proposals is much too slow for production lines operating at more than 60 FPS [14].

One-Stage Detectors and Attention Evolution: In Single-stage detectors, object detection is considered just a single regression problem and performs a pass of the image through a network and outputs bounding box coordinates and class probabilities [15]. Subsequent iterations of YOLO have undergone slight but significant changes in structure. YOLOv8 is an advanced model featuring an anchor-free decoupled head and multi-scale feature fusion. However, YOLO encounters challenges when trying to detect tiny and low-variation defects. In response to this issue, the scientists presented attention mechanisms [16]. The "squeeze and excitation" (SE) module uses global spatial information to model interdependencies:

- The “squeeze and excitation” (SE) module uses global spatial information to model interdependencies.
- Media attention focuses on the spatial coordinates of the anomaly [17].
- The model utilizes both methods sequentially to refine feature maps.

Nonetheless, memory and inference are expensive with recurrent attention modules. Rephrase More Undo ShuffleAttention resolves this by dividing the feature maps into smaller clusters and performing channel and spatial consideration together [18]. Adding ShuffleAttention to the deeper layers of YOLOv8 increases the ability to detect small and less salient defects without affecting inference speed [19].

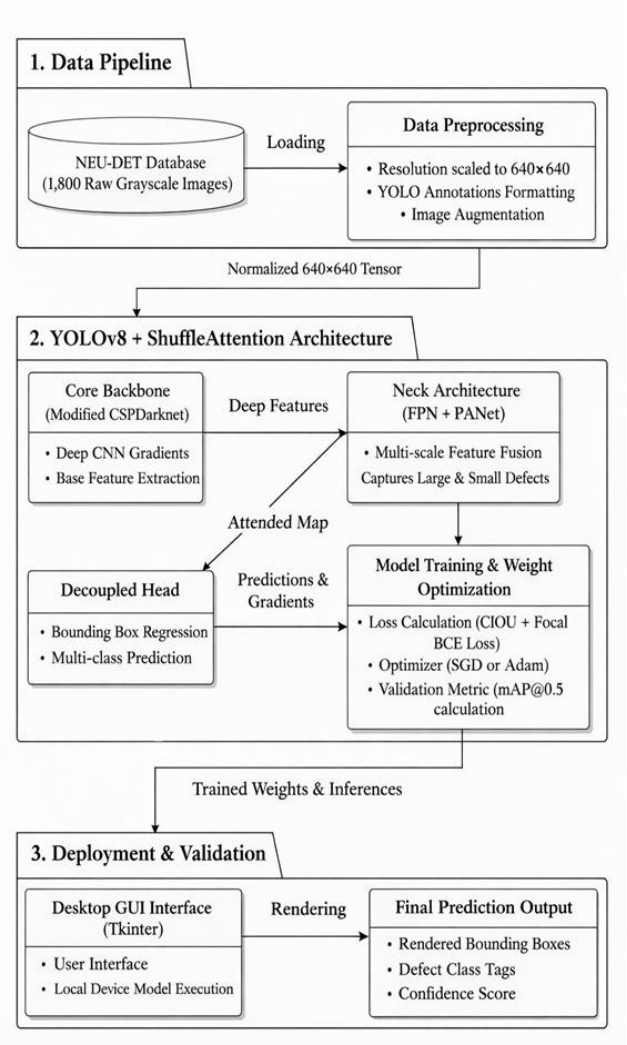


Figure 1: High-level block diagram

3. Proposed Framework and System Architecture

To train the augmented detector, test metrics, and deploy weights stripped down to a low inference mode is the end goal of your entire system pipeline.

3.1. Modification and Augmentation of Data:

The system ingests the public NEU surface defect dataset (NEU-DET). The dataset has six classes, namely crazing, patches, inclusion, pitted_surface, scratches and rolled-in-scale. Annotations are provided for each image in the dataset, represented in (c, x_c, y_c, w, h) form. In this equation, c represents the class index, xc and yc correspond to the normalized center coordinates, and w and h indicate the normalized width and height of the bounding box. The tools take the images from the dataset and convert them to tensors of size 640×640.

3.2. YOLOv8 Baseline Backbone

The First is composed of three functioning stages. The backbone of the cross-stage partial network has the P3, P4, and P5 architecture with the optimal profile. The stages are all such extraction operations, so they allow for high-res operations. First and foremost, the cross-stage partial bottlenecks are important, as they combine these two information flow-optimizing features. The multi-scale bounding box regression and classification with an anchor-free, easy-to-conflict feature in the anchor-free head reduces feature conflict.

3.3. ShuffleAttention Integration

Combining the ShuffleAttention (SA) module into the deep feature extraction pipeline greatly improves sensitivity to fine-grained defects. The SA module divides X into G distinct groups based on the intermediate feature map XRCHW. The two branches of subfeatures groupX{k} are the channel attention branch X{k1} and spatial attention branch X{k2}. Once the outputs are concatenated, researchers perform a channel shuffle operation on the G groups to acquire global context. The proposed architecture starts from the NEU-DET dataset and its data processing pipeline as the first block, then the main block of the architecture, which is YOLOv8 merged with ShuffleAttention here; all training and inference take place, and finally, to make this paper more accessible to a higher percentage of users, a custom graphical user interface is developed in which everything is preloaded. Researchers can do the training and inference with simple button clicks (Figure 1).

3.4. Formal Optimization and Inference Pipeline

The network applies an objective total loss function during training:

$$L_{total} = \lambda_{box} L_{CIoU} + \lambda_{cls} L_{BCE} + \lambda_{dfl} L_{DFL} \quad (1)$$

During inference, researchers use a confidence threshold and NMS to eliminate overlapping bounding boxes with high precision. NMS uses a calculation called Intersection over Union:

$$IoU = \frac{Area(p_{max} \cap p_j)}{Area(p_{max} \cup p_j)} \quad (2)$$

To get better access and make it easier for the end user, a custom graphical user interface is developed to make it more accessible to everyone (Figure 2).

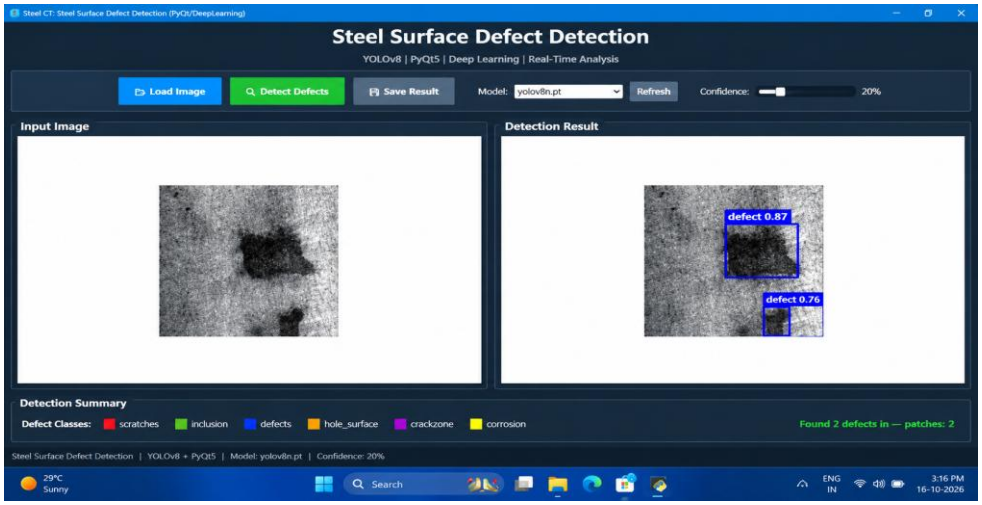


Figure 2: Custom graphical user interface

3.5. Experimental Setup

3.5.1. Dataset Description

Experiments were done on the NEU-DET dataset that comprises grey-scale images of steel strips with extreme intra-class variation and inter-class similarity. To avoid data leakage, the training, validation and testing sets should not overlap with each other (Figure 3).

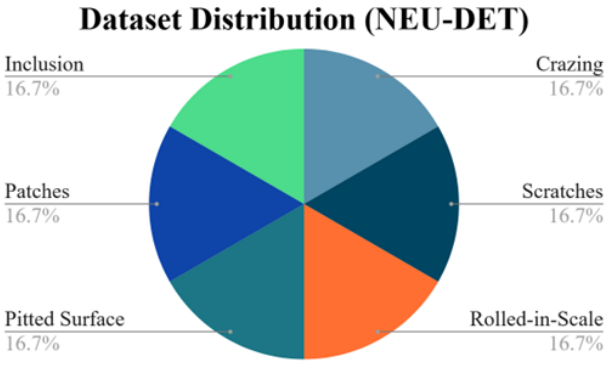


Figure 3: Dataset distribution pie chart

3.5.2. Implementation and Hyperparameters

The defect detector is constructed by adding the ShuffleAttention block after Layer 22 (after the deepest P5 backbone feature extraction layer) using a PyTorch network graph:

- **Input Resolution:** 640x640 pixels
- **Training Duration:** 100 epochs
- **Batch Size:** 8
- **Optimizer:** SGD / AdamW algorithms

3.5.3. Evaluation Metrics Formulation

To proceed with the evaluation of the model's performance, researchers should analyse and measure two things: whether the system correctly detects and identifies defects, and how good and precise the localization of the bounding box is in covering the actual damage. Here are the metrics for evaluation. Precision (P) and Recall (R):

$$P = \frac{TP}{TP+FP} \tag{3}$$

$$R = \frac{TP}{TP+FN} \tag{4}$$

Average Precision (AP) and Mean Average Precision (mAP):

$$AP = \int_0^1 \quad P(R)dR \tag{5}$$

$$mAP = \frac{1}{N}\sum_{i=1}^{N} \quad AP_i \tag{6}$$

4. Results and Analysis

4.1. Quantitative Baseline Comparison

When researchers analyse and go through the raw results from training, they notice that the model hit its best stride when completing epoch 98. It recorded an mAP@50 of 0.7406, a precision of 0.6746, and a recall of 0.6829. The notable takeaway to consider here is the massive jump in absolute precision. Researchers can notice a jump of 10.1% from 64.0% to 74.1% with the addition of the ShuffleAttention module to the architecture (Table 1).

Table 1: Comprehensive performance comparison on NEU-DET

Algorithm	mAP@50 (%)	Precision (%)	mAP@50:95 (%)
Faster R-CNN	68.0	67.8	30.6
YOLOv4n	50.8	49.3	22.5
YOLOv5n	70.8	66.8	36.5

YOLOv7	66.0	65.8	35.6
YOLOv8n	64.0	73.7	40.2
YOLOv8+ShuffleAttention (Ours)	74.1	67.4	40.6

This proves that the addition is doing its job correctly and fulfilling the purpose for which it is designed. This achievement is attained by filtering out uneven metallic background textures of the metal surface to stop the system from throwing false positives.

4.2. Class-Wise Breakdown and Qualitative Analysis

Researchers can understand how the attention mechanism works if they break down the performance by the shape of the defects:

4.2.1. High Contrast Defects (Patches and Inclusions)

The system had almost no issues in this segment, as these marks or shapes have sharp edges and contrast and obvious colour differences when compared to the surface of normal steel with no defects; the detection scores were practically perfect, with an mAP@50 above 0.90.

4.2.2. Granular and Dispersed Defects (Crazing and Pitted Surfaces)

These kinds of defects are harder than the high-contrast ones to catch because they aren't present in a single spot or collectively located together, but they are scattered or spread out like sand all over the surface. There was a need for a special technique to overcome the issues here, and the Spatial Attention mechanism from the ShuffleAttention module stepped up here by having those network lines grouped, which spiked the accuracy for these particular classes by more than 2.5%.

4.2.3. Low Contrast and Elongated Defects (Scratches and Rolled-in Scale)

Notably, the scratches, which are low-contrast and elongated defects, are notoriously hard to catch as they are so thin and narrow, which in turn helps in blending them into the metal surface (Figure 4).

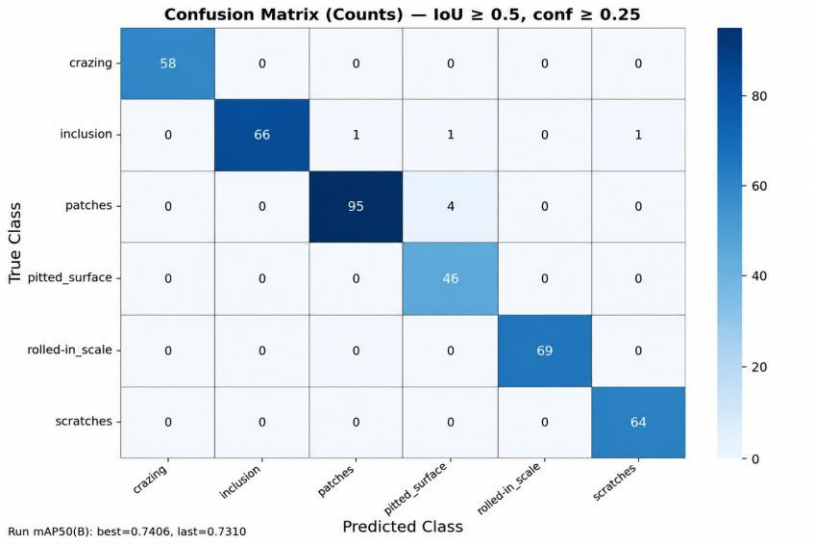


Figure 4: Confusion matrix

Here, the attention module focused on the defective area and left the remaining background out of processing so that the information is shared across different channels, which made sure the faint lines of scratches didn't just disappear as the image was processed with deeper processes in the network.

4.3. Inference Latency and Viability

One of the major concerns that was prevailing when this upgrade was done was whether this would slow down the system. But fortunately, the ShuffleAttention block did not affect or change the overall footprint of the model; only the parameter count

slightly increased from 3.00M to 3.15M. When the inference was run on an edge computing GPU, the system consistently sustained 85 frames per second (FPS). This is more than enough, as the cameras in most industrial assembly lines only do about 30 to 60 FPS; this makes our model more optimal, as it can comfortably run quickly for live factory deployment.

5. Discussions

NEU-DET dataset was tested rigorously, and this led to the conclusion that simply adding global attention mechanisms or generic modules can't always give better results. This was drafted after seriously evaluating the performance results, which are drawn from the original YOLOv8 and Faster R-CNN that struggled heavily with the presence of messy and natural glare of metallic surfaces. These models wrongly predicted lighting shifts that are harmless as actual damage, which in turn caused the spikes in false positives. Here, the new attention module is introduced to eliminate the existing issues of false positives by using the ShuffleAttention mechanism, which rewrites how the network filters noise. The ShuffleAttention module makes use of a different processing technique where spatial and channel attention are both handled in parallel. This new addition helps in preventing washouts of small defects during downsampling and, in turn, helps in pushing our mAP@50 to 74.06%. When evaluating in a real factory scenario, raw accuracy alone doesn't promise a very uplifting result if the underlying software creates a production bottleneck (Figure 5).

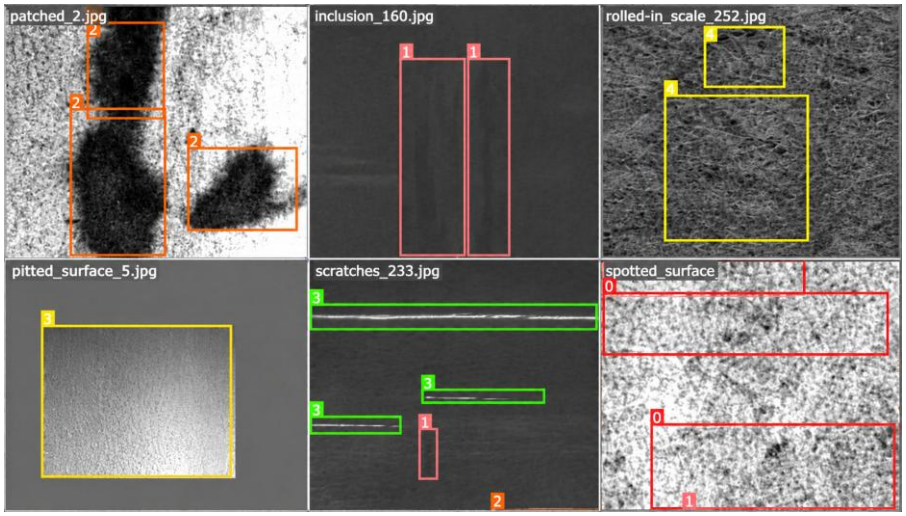


Figure 5: Visual inference results grid

The system or architecture is said to be successful only when the handling of the accuracy-latency trade-off is correctly maintained. As the parameter count researchers attained is incredibly lean (just 3.15M), the model attained a sustained performance on standard edge hardware by having frame processing at over 85 fps. This again means that this architecture can also be upgraded with the live feed, high-speed rolling lays while still giving an absolute precision jump of 10.1%, while also drastically reducing the number of times the human operator has to stop or cut the line for a false alarm. Even after attaining this absolute high accuracy and efficient result, researchers still have to acknowledge a lingering limitation in our current study. The YOLOv8 framework, when merged with ShuffleAttention and trained with the NEU-DET dataset, works best and is calibrated to specific hot-rolled metallurgical surfaces. This doesn't guarantee that our current architecture works well for other types of surfaces and lighting conditions. This architecture should be stress-tested against cold-rolled steel and larger multi-factory datasets to be regarded as a universal fit or plug-and-play solution. It should be ensured that the weights generalize as intended to different metal surfaces and grades with different lighting.

6. Conclusion and Future Work

Industry 4.0 is about integrating sophisticated technologies like AI, machine learning and computer vision into intelligent production processes. In this context, researchers propose a new system for steel surface defect detection that combines the YOLOv8 framework and the ShuffleAttention module to improve the accuracy and reliability of defect detection. The suggested method is tested using the NEU-DET dataset for steel surface defect detection by automated image analysis to reduce the dependency on manual inspection. The proposed architecture achieved the best mAP@0.50 of 0.7406 and an absolute gain of 10.1% in precision, which indicates an improvement in fault detection performance. Incorporation of ShuffleAttention allows the model to pay attention to essential elements of the image and regions connected to defects, leading to more efficient detection of surface abnormalities. Moreover, the authors have designed a strong architecture with multi-factory dataset

portability to evaluate the model in multiple production contexts. This way, it becomes more robust to unknown conditions, such as random lighting and vibrations in the service, which can affect the image quality and the detection reliability. The system shows promise for manufacturing adoption through automated quality inspection and reliable defect identification. But there are still modifications to be made to increase its actual performance. In the future, we will focus on finding the best position of the attention module in the design of YOLOv8 to get better results. In addition, by integrating AI acceleration technologies such as NVIDIA TensorRT, the computation efficiency and detection speed can be improved to enable real-time industrial applications. Another essential line is to develop an unsupervised learning mechanism to alleviate the expensive and time-consuming manual data labeling process. It may ease dataset preparation and boost learning efficiency. The researchers also want to allow the model to adjust on its own to new sorts of surfaces and patterns of defects. The proposed system is a promising solution for intelligent steel surface inspection. Future advances will concentrate on accuracy, speed, adaptability and reliable deployment in varied industrial situations.

Acknowledgment: The authors express their sincere gratitude to SRM Institute of Science and Technology at Ramapuram and Coventry University for their institutional support and encouragement. The authors also acknowledge the academic resources and research environment that contributed to the successful completion of this study.

Data Availability Statement: The data supporting this study are available from the corresponding author upon reasonable request.

Funding Statement: The authors received no external funding for this research.

Conflicts of Interest Statement: The authors declare no conflicts of interest and have appropriately acknowledged all sources.

Ethics and Consent Statement: The study complied with applicable ethical standards, and informed consent was obtained from all participants.

References

1. C. Li, H. Huang, H. Mo, Z. Wen, S. Zhou, and Z. Zhu, "Novel efficient steel surface defect detection model based on ConvNeXt v2 and squeeze aggregated excitation attention," *IEEE Access*, vol. 13, no. 6, pp. 107512–107531, 2025.
2. H. Wang, W. Li, B. Zhang, and Z. Gu, "YOLOv8n-GSE: Efficient steel surface defect detection method," *IEEE Access*, vol. 13, no. 8, pp. 166343–166356, 2025.
3. J. Su, Q. Luo, W. Gui, and C. Yang, "Few-shot parameter efficient fine tuning for SAM in salient steel surface defect detection," *IEEE Transactions on Industrial Informatics*, vol. 21, no. 10, pp. 7742–7753, 2025.
4. Z. Dang and X. Wang, "FD-YOLO11: A feature-enhanced deep learning model for steel surface defect detection," *IEEE Access*, vol. 13, no. 4, pp. 63981–63993, 2025.
5. C. Wang, H. Wang, Y. Jiang, L. Yu, and X. Wang, "CSCP-YOLO: A lightweight and efficient algorithm for real-time steel surface defect detection," *IEEE Access*, vol. 13, no. 6, pp. 113517–113528, 2025.
6. H. Haoyan, T. Jinwu, W. Haibin, and L. Xinyun, "EAD-YOLOv10: Lightweight steel surface defect detection algorithm research based on YOLOv10 improvement," *IEEE Access*, vol. 13, no. 3, pp. 55382–55397, 2025.
7. X. Huang, J. Zhu, and Y. Huo, "SSA-YOLO: An improved YOLO for hot-rolled strip steel surface defect detection," *IEEE Transactions on Instrumentation and Measurement*, vol. 73, no. 1, pp. 1–17, 2024.
8. Y. Du, H. Chen, Y. Fu, J. Zhu, and H. Zeng, "AFF-Net: A strip steel surface defect detection network via adaptive focusing features," *IEEE Transactions on Instrumentation and Measurement*, vol. 73, no. 5, pp. 1–14, 2024.
9. K. Shen, X. Zhou, and Z. Liu, "MINet: Multi-scale interactive network for real-time salient object detection of strip steel surface defects," *IEEE Transactions on Industrial Informatics*, vol. 20, no. 5, pp. 7842–7852, 2024.
10. Q. Luo, B. Li, J. Su, C. Yang, W. Gui, O. Silven, and L. Liu, "CDDNet: Camouflaged defect detection network for steel surface," *IEEE Transactions on Instrumentation and Measurement*, vol. 73, no. 11, pp. 1–13, 2024.
11. L. Liu, X. Feng, F. Li, Q. Xian, Z. Chen, and Z. Jia, "Surface defect detection of industrial components based on improved YOLOv5s," *IEEE Sensors Journal*, vol. 24, no. 15, pp. 23940–23950, 2024.
12. H. Feng, K. Song, W. Cui, Z. Zhou, and Y. Yan, "A suspected defect screening-guided lightweight network for few-shot aviation steel tube surface defect segmentation," *IEEE Sensors Journal*, vol. 24, no. 22, pp. 38113–38124, 2024.
13. W. Yaodong, Y. Hang, G. Baoqing, S. Hongmei, and Y. Zujun, "Research on real-time detection system of rail surface defects based on deep learning," *IEEE Sensors Journal*, vol. 24, no. 13, pp. 21157–21167, 2024.
14. R. Guo, P. Ji, Y. Zhang, J. Hu, W. Liu, X. Li, and M. Li, "Steel surface defect detection based on improved GCHS-YOLO algorithm," *IEEE Access*, vol. 12, no. 12, pp. 190865–190875, 2024.
15. H. Liu, C. Chen, R. Hu, J. Bin, H. Dong, and Z. Liu, "CGTD-Net: Channel-wise global transformer-based dual-branch network for industrial strip steel surface defect detection," *IEEE Sensors Journal*, vol. 24, no. 4, pp. 4863–4873, 2024.

16. C. Song, J. Chen, Z. Lu, F. Li, and Y. Liu, "Steel surface defect detection via deformable convolution and background suppression," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, no. 5, pp. 1–9, 2023.
17. X. Xiang, Z. Wang, J. Zhang, Y. Xia, P. Chen, and B. Wang, "AGCA: An adaptive graph channel attention module for steel surface defect detection," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, no. 2, pp. 1–12, 2023.
18. X. Wang, F. He, Y. Chen, and S. Feng, "Steel Surface Defect Detection Based on Improved YOLOv7," *4th International Conference on Consumer Electronics and Computer Engineering (ICCECE)*, Guangzhou, China, 2024.
19. Q. Wang, M. Wang, J. Sun, D. Chen, and P. Shi, "Review of surface-defect detection methods for industrial products based on machine vision," *IEEE Access*, vol. 13, no. 5, pp. 90668–90697, 2025.
20. B. Hunag, J. Liu, X. Liu, K. Liu, X. Liao, K. Li, and J. Wang, "Improved YOLOv5 Network for Steel Surface Defect Detection," *Metals*, vol. 13, no. 8, pp. 1– 16, 2023.

Publisher's Note: *The publisher remains impartial concerning jurisdictional claims in published maps and institutional affiliations. Responsibility for the content rests entirely with the authors and does not necessarily reflect the publisher's perspectives.*