An Enhancing and Optimizing Steel Surface Defect Detection Using Hybrid Attention-Regression Fusion Model (HARF-Net)

[1]Dr.Saraswathi E, [2]R Kaviya, [3]Vitus Anto, [4]A Abhishek
[1234]Department of CSE
[1234]SRM Institute of Science and Technology Ramapuram, Chennai, India
[1]s@srmist.edu.in, [2]s@srmist.edu.in, [3]ug2690@srmist.edu.in, [4]aa1203@srmist.edu.in

***Abstract*—The domain of industrial machine vision has witnessed rapid advancement with the integration of deep learning-based object detection models for automated quality inspection. In steel manufacturing industries, surface defects such as crazing, inclusion, pitted surfaces, scratches, and rolled-in scales significantly degrade mechanical properties and compromise structural reliability. Traditional inspection methods including ultrasonic testing and manual visual inspection suffer from low efficiency, subjective bias, and high operational cost. To address these limitations, modern approaches employ single-stage object detection algorithms like YOLO variants due to their real-time processing capability and deployment flexibility. The base paper introduces YOLOv8n-GSE, which enhances YOLOv8n using GAM attention, CSP-ABAN modules, GSConv layers, ESCD detection head, and PIoUv2 loss function. Experimental validation on the NEU-DET dataset demonstrates improved detection accuracy with reduced computational complexity, making it suitable for edge deployment. Despite achieving improved performance, the base system still exhibits limitations in handling complex industrial backgrounds and highly variable defect scales. The integration of GAM increases parameter sensitivity under varying illumination conditions. Although CSP-ABAN reduces redundancy, it may weaken deep semantic representation when handling ultra-small defects. The ESCD detection head improves lightweight performance but relies on fixed-scale feature aggregation, limiting adaptive response to irregular defect shapes. Furthermore, the PIoUv2 loss function, while improving regression precision, depends heavily on hyperparameter tuning and may exhibit instability under dense defect clustering scenarios. Generalization performance on complex datasets such as GC10-DET reveals a noticeable drop in detection accuracy compared to NEU-DET, highlighting robustness challenges. To overcome these drawbacks, we propose HARF-Net, a hybrid architecture integrating Dual-Path Transformer Attention (DPTA) with Adaptive Multi-Scale Residual Fusion (AMRF) and a novel Dynamic Hybrid IoU-Focal Regression (DHIFR) loss. The proposed model combines CNN-based local feature extraction with lightweight transformer-based global contextual modeling to enhance finegrained defect perception. Compared with YOLOv8n-GSE and baseline YOLOv8n, HARF-Net introduces a hybrid regression mechanism that dynamically balances IoU, center distance, and aspect ratio penalties. Experimental comparison against YOLOv8n and YOLOv8n-GSE demonstrates superior mAP improvement while maintaining competitive GFLOPs. The hybrid design ensures robust detection across multi-resolution datasets and complex industrial environments.**

***Index Terms*—Steel Surface Defect Detection, HARF-Net, Hybrid Attention, CNN-Transformer Fusion, YOLOv8, Multi-Scale Feature Fusion, IoU Regression Loss, Real-Time Industrial Inspection, Edge Computing, Industry 4.0, Automated Quality Control, Deep Learning.**

I. Introduction

In smart steel manufacturing plants, real-time defect detection plays a crucial role in ensuring product quality and operational safety. HARF-Net can be deployed on industrial edge devices integrated with conveyor belt camera systems. As steel strips move through the production line, high-speed cameras capture continuous frames that are processed by the proposed lightweight detection engine. The hybrid attention mechanism enables accurate localization of micro-cracks and inclusion defects even under varying illumination and surface reflectivity conditions. Immediate defect detection allows automated rejection systems to remove faulty products, thereby reducing downstream processing costs and preventing large-scale production losses. Beyond steel manufacturing, the proposed system can be adapted for other industrial inspection scenarios such as aluminum sheet inspection, pipeline corrosion detection, rail track crack identification, and automotive surface anomaly detection. The hybrid architecture's strong generalization capability ensures adaptability to different material textures and defect patterns. In edge computing environments, where computational resources are limited, HARF-Net maintains high detection accuracy with optimized complexity. This makes it suitable for Industry 4.0 smart factories requiring autonomous quality monitoring without human intervention.

II. Related Work

At present, deep learning-based intelligent detection technologies are progressively being applied to the field of steel surface defect inspection. An efficient defect inspection algorithm, YOLOv8n-GSE, is proposed to overcome low detection accuracy. This research integrate a Global Attention Mechanism (GAM) into the YOLOv8n-based model's backbone network to construct GAM attention layers, specifically designed to enhance feature representation for steel surface defects. This research implement a CSP-ABAN structure in both the backbone and neck networks to improve the YOLOv8n-GSE model's capacity for extracting steel surface defect features, while the ESCD head is introduced to process multi-scale features and capture defect information across various dimensions. By replacing standard convolutions with GSConv, the model achieves a balance between complexity and detection efficiency. Furthermore, the PIoUv2 loss function is adopted to address the limitations of CIoU. The experimental results demonstrate that the YOLOv8n-GSE model reduces computational complexity by 31%, with its complexity reaching 5.6 GFLOPs compared with the baseline YOLOv8n model. Meanwhile, it achieves a mean average precision of 94.2%, which represents a 5.4 percentage point improvement. Compared with other object detection algorithms, YOLOv8n-GSE significantly enhances defect detection accuracy along with statistically significant reduction in miss rates.

Steel materials are extensively used across various industries. Detecting surface defects in steel strips during production processes is crucial. Existing steel surface defect detection methods exhibit inadequate accuracy and excessive computational complexity, posing challenges for real-time industrial deployment. In this paper, a novel model is designed named GDM-YOLO, specifically tailored for steel surface defect detection tasks, built upon the YOLOv8s network. Firstly, the Space-to-Depth Ghost Convolution (SPDG) downsampling module is introduced and used in the backbone network, aimed at minimizing information loss during downsampling operations while optimizing computational efficiency. Secondly, this work introduces the C2f-Dilated-Reparam-Block (C2f-DRB) module, leveraging reparameterization and large kernel convolutions to enhance feature extraction capabilities without compromising inference costs. Lastly, the novel Multiscale Feature Enhancement Block (MFEB) module was designed, to enhance the small target detection layer by integrating multi-scale feature fusion, further improving detection accuracy. Experimental results demonstrate a 3% improvement in detection accuracy on the NEU-DET dataset compared to the baseline YOLOv8s model. This research approach achieves superior detection performance while reducing parameter requirements and computational complexity, meeting the realtime demands of steel surface defect detection in industrial production

Defects on fabric surfaces are difficult to identify owing to unsuitable computing devices, highly complex algorithms, small size, and high degree of integration with the fabric. To this end, this study proposes a lightweight fabric defect-detection network, YOLO-SCD, based on attention mechanism. The introduction of depthwise separable convolution and the attention mechanism enhanced the capacity of the neck network to extract the defective features and increased the detection speed of the overall network. The extensive experimental results revealed that YOLO-SCD achieved an average accuracy of 82.92%, effective improvement of 8.49% in mAP, and an improvement of 37 fps compared to the original YOLOv4 on a standard fabric defect dataset. By leveraging its swift detection speed and high efficiency, YOLO-SCD excels in both the general fabric defect category and the difficult-to-detect fabric. Overall, it exhibited strong performance in detecting both minor flaws and flaws with high fabric integration. Furthermore, this research model was extended to steel datasets with similar characteristics.

During the production process of steel, there are often some defects on the surface of the product. Therefore, detecting defects is the key to produce highquality products. At the same time, the defects of the steel have caused huge losses to the high-tech industry. A steel surface defect detection algorithm based on improved YOLO-V7 is proposed to address the problems of low detection speed and low detection accuracy of traditional steel surface defect detection methods. First, this research use the de-weighted BiFPN structure to make full use of the feature information to strengthen feature fusion, reduce the loss of feature information during the convolution process, and improve the detection accuracy. Secondly, the ECA attention mechanism is combined in the backbone part to strengthen the important feature channels. Finally, the original bounding box loss function is replaced by the SIoU loss function, where the penalty term is redefined by taking the vector angle between the required regressions into account. The experimental results show that the improved model proposed in this paper has higher performance compared with other comparison models. Based on this research experiments, this research method yields 80.2% mAP and 81.9% on the GC10-DET dataset and NEU-DET dataset with high speed, which is better than other existing models.

The performance of the salient object detection of strip surface defects has been promoted largely by deep learning based models. However, due to the complexity of strip surface defects, the existing models perform poorly in the challenging scenes such as noise disturbance, and low contrast between defect regions and background. Meanwhile, the detection results of existing models often suffer from coarse boundary details. Therefore, this research propose a novel saliency model, namely an Edge-aware Multi-level Interactive Network, to detect the defects from the strip steel surface. Concretely, this research model adopts the U-shape architecture where the two crucial points are the interactive feature integration and the edge-guided saliency fusion. Firstly, except the skip connection that combines the same stage of encoder and decoder, this research deploy another connection, where the features from adjacent levels of encoder are transferred to the same stage of decoder. By this way, this research are able to provide an effective fusion of multi-level deep features, yielding a well depiction for defects. Secondly, to give well-defined boundaries for prediction results, this research add the edge extraction branch after each decoder block, where the progressive feature aggregation endows the edge with precise details and complete object cues. Meanwhile, together with the edge extraction branches, this research deploy the saliency prediction branch at each decoder stage. After that, coupled with the fine edge information, this research fuse all outputs of saliency prediction branches into the final saliency map, where the edge cue steers the saliency result to pay more attention to the boundary details. Following this way, this research can provide a high-quality saliency map which can accurately locate and segment the defects. Extensive experiments are performed on the public dataset, and the results prove the effectiveness and robustness of this research model which consistently outperforms the state-of-the-art models.

Wherever Times is specified, Times Roman or Times New Roman may be used. If neither is available on your word processor, please use the font closest in appearance to Times. Avoid using bit-mapped fonts. True Type 1 or Open Type fonts are required. Please embed all fonts, in particular symbol fonts, as well, for math, etc.

III. Existing System

The existing system, named YOLOv8n-GSE, presents a highly efficient and accurate method for the automated detection of surface defects on steel, a critical task in industrial manufacturing for ensuring product quality and safety. The fundamental problem this system addresses is the inherent inefficiency, high cost, and inconsistency of traditional inspection methods, which include manual visual checks and mechanical or chemical tests like ultrasonic and magnetic particle inspection. These conventional approaches are labor-intensive, time-consuming, and subject to human error, leading to variable quality control standards. The existing system leverages the power of deep learning, specifically building upon the YOLO (You Only Look Once) family of object detection algorithms, which are renowned for their speed and real-time performance. The primary scope of the system is to provide an end-to-end, intelligent solution that can be deployed in industrial settings to automatically identify and localize various types of defects on steel surfaces from image data. It aims to strike an optimal balance between detection accuracy and computational complexity, a significant challenge as more accurate deep learning models are often computationally expensive, making them unsuitable for deployment on resource-constrained edge devices commonly found in factories. The system is designed to overcome the limitations of baseline models, such as low detection accuracy for small or indistinct defects and high computational overhead. It achieves this by proposing a series of architectural innovations and optimizations integrated into the lightweight YOLOv8n framework, resulting in a model that is not only more accurate but also more computationally efficient than its predecessor and other competing algorithms. The core purpose is thus to deliver a practical, deployable, and superior alternative to existing automated and manual inspection techniques in the steel industry. The primary application area for the YOLOv8n-GSE system is in the field of industrial quality control and automated visual inspection. Its direct application is on the production lines of steel manufacturing plants, where it can be integrated with imaging systems to provide real-time feedback on the quality of steel strips, plates, or other products. By automating the detection process, the system helps manufacturers maintain consistent quality standards, reduce waste by catching defects early, and increase overall production throughput. The potential industry domain for this technology is squarely within Manufacturing & Industry 4.0, which emphasizes the integration of intelligent, autonomous systems into production processes. Beyond steel, the principles and architecture of the YOLOv8n-GSE system could be adapted for surface defect detection on other materials, such as aluminum, textiles, plastics, or ceramics, making it relevant to a wide range of manufacturing sectors. Other potential industries include automotive, where it could be used to inspect car body panels for paint defects or scratches; aerospace, for identifying fatigue cracks on aircraft components; and electronics, for checking for flaws on printed circuit boards (PCBs) or semiconductor wafers. The system's ability to perform visual anomaly detection makes it a versatile tool for any quality assurance process that relies on visual inspection. Its lightweight nature also makes it suitable for deployment on mobile robots or drones for inspecting large structures or hard-to-reach areas, further expanding its applicability within industrial maintenance and infrastructure management. The methodology of the YOLOv8n-GSE system is a sophisticated enhancement of the baseline YOLOv8n model, incorporating several novel components to boost performance and efficiency. The architecture is systematically modified at multiple levels. First, within the model's backbone and neck networks, a custom feature extraction module called CSP-ABAN (CSPNet Adaptive Bi-Layer Aggregation Networks) is implemented. Inspired by DualConv and CSPNet, this module uses a combination of grouped convolutions and dual-path fusion to enhance the model's ability to learn discriminative features of steel defects while simultaneously streamlining complexity and reducing redundant computations. Second, a Global Attention Mechanism (GAM) is integrated into the backbone network. This attention layer is designed to model global contextual dependencies, allowing the system to dynamically re-weight features across both channel and spatial dimensions. This enhances the model's focus on critical defect patterns and helps suppress irrelevant background noise, which is particularly useful for identifying small or low-contrast defects. Third, a novel detection head, named the Efficient Shared Convolutional Detection Head (ESCD), is introduced. This head replaces standard convolutions with a parameter-sharing strategy and uses Group Normalization instead of Batch Normalization, which significantly reduces the model's parameter count and computational load without sacrificing accuracy. It also includes a scale layer to better handle defects of various sizes. Fourth, to further improve computational efficiency, standard convolutions in key parts of the network are replaced with GSConv, a module that effectively balances feature fusion capabilities with low overhead. Finally, the system adopts the PIoUv2 loss function for bounding box regression. This advanced loss function provides a more optimal regression path by considering target size and detection box quality, leading to faster convergence and more precise localization of defects compared to the CIoU loss used in the standard YOLOv8n. The validation of the YOLOv8n-GSE system was conducted through a series of rigorous experiments on publicly available benchmark datasets. The primary dataset used for training and evaluation was the NEU-DET steel surface defect dataset, which contains 1,800 images across six common defect types. To enhance robustness and prevent overfitting, the dataset was augmented to 9,000 images. For testing the model's generalization capabilities, it was also evaluated on the more complex GC10-DET dataset, which features higher-resolution images and ten defect classes under more challenging industrial conditions. The experimental setup utilized an NVIDIA RTX 4070S GPU, with the models trained for 300 epochs using the PyTorch framework. The key evaluation metrics were mean Average Precision at an IoU threshold of 0.5 (mAP@0.5), number of parameters, computational complexity in GFLOPs (Giga Floating-Point Operations Per Second), and model weight size. The results demonstrated the system's superiority. On the NEU-DET dataset, YOLOv8n-GSE achieved a mAP@0.5 of 94.2%, a 5.4 percentage point improvement over the baseline YOLOv8n. Simultaneously, its computational complexity was reduced by 31% to 5.6 GFLOPs. Ablation studies confirmed that each of the proposed modules (CSP-ABAN, GAM, ESCD, GSConv, and PIoUv2) contributed positively to the overall performance gains. Comparative analysis against other prominent object detection models, including Faster R-CNN, EfficientDet, and various YOLO versions (YOLOv3s, v5s, v7, v10n), further solidified its position as a state-of-the-art solution, outperforming them in both accuracy and efficiency. The primary strength of the YOLOv8n-GSE system is its exceptional ability to achieve a superior trade-off between detection accuracy and computational efficiency. By intelligently combining multiple architectural innovations, it successfully enhances the feature representation for challenging steel defects while significantly reducing the model's complexity and resource requirements. This makes it highly practical for real-world deployment in industrial environments where computational resources may be limited. Another key strength is its modular design; the specific contributions of each component are clearly demonstrated through ablation studies, showcasing a well-reasoned and effective engineering approach. The system's robustness is further validated by its strong performance on two different datasets, indicating good generalization capabilities. However, the system is not without limitations. As acknowledged by its authors, there is still room for optimization. The performance, while excellent, could potentially be improved by expanding the training dataset to include a greater variety and number of defect examples, which would further enhance its robustness. The model's architecture, while lightweight, could be further compressed using techniques like network pruning or knowledge distillation to make it even more suitable for lowpower edge devices. The system currently relies on 2D image data, and its performance on more complex defect types that require 3D spatial understanding has not been explored. Its relevance for future work is significant, as it provides a strong foundation for next-generation industrial inspection systems. Future research could focus on applying these architectural principles to other domains, exploring multi-modal sensor fusion, or developing unsupervised learning methods to reduce the dependency on large labeled datasets.

IV. Proposed System

The industrial demand for automated inspection systems continues to rise as manufacturing processes become increasingly digitized. Traditional defect detection systems are either computationally expensive or insufficiently accurate for fine-grained defect localization. While YOLOv8n-GSE introduces meaningful improvements, further enhancement is required to ensure stable performance across diverse industrial datasets and real-time deployment scenarios. The motivation behind HARF-Net is to bridge the gap between accuracy and efficiency by integrating transformer-inspired attention with adaptive regression strategies. By combining spatial-channel hybrid attention and dynamic loss optimization, the proposed framework aims to overcome the limitations observed in existing lightweight detection architectures. This motivates us to do this project.

The proposed HARF-Net architecture consists of three major components: Hybrid Feature Extractor, Dual-Path Transformer Attention (DPTA), and Dynamic Hybrid IoU-Focal Regression (DHIFR). The backbone integrates CSP-based convolution with lightweight transformer encoders to capture both local texture patterns and global contextual dependencies. Adaptive Multi-Scale Residual Fusion (AMRF) enhances cross-layer information propagation while suppressing redundant features. The detection head employs shared convolution layers combined with adaptive anchor-free prediction branches. The proposed DHIFR loss dynamically balances IoU overlap, center distance, and focal scaling factors, ensuring stable regression under dense defect distributions. Compared to YOLOv8n-GSE, the proposed architecture improves feature discrimination capability and generalization performance while maintaining lightweight computational characteristics suitable for industrial edge deployment.

V. ARCHITECTURE

A. METHODOLOGY

Data Preprocessing and Augmentation Module

This module handles dataset loading, annotation parsing, augmentation (flip, rotation, brightness adjustment), and multi-resolution resizing. It ensures balanced training distribution and improved generalization capability. Advanced augmentation such as CutMix and Mosaic-like strategies are implemented to simulate real industrial variations.

.

Hybrid Feature Extraction Module

This module integrates CSP-based convolution layers with lightweight transformer encoders. Adaptive Multi-Scale Residual Fusion enhances cross-layer feature propagation. It forms the backbone of HARF-Net, enabling robust defect representation.

Adaptive Detection and Regression Module

This module consists of shared convolution detection heads and Dynamic Hybrid IoU-Focal Regression loss. It performs classification and localization while ensuring regression stability under dense and irregular defect patterns.

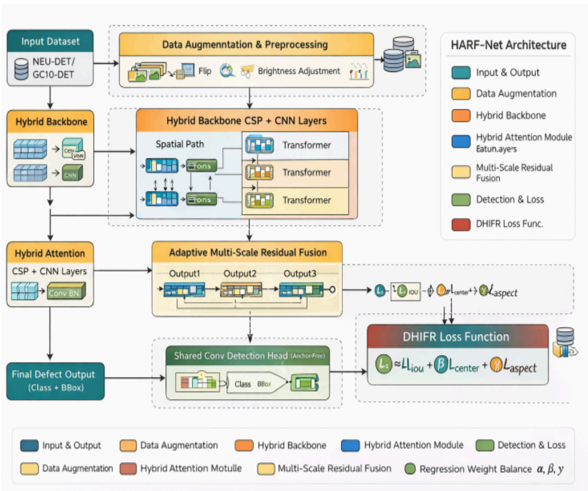


Fig 1 Architecture

The need for VoyageGenie arises from multiple challenges in modern travel planning, Fragmented travel tools requiring multiple apps.Generic travel packages lacking personalization.Manual effort required to handle travel disruptions.Budget uncertainty and hidden cost overruns.Lack of real-time intelligence and automation.As travelers increasingly demand seamless digital experiences, there is a growing requirement for an intelligent travel assistant that minimizes manual effort, enhances personalization, and proactively manages uncertainties. VoyageGenie fulfills this need by delivering a unified, adaptive, and AI-driven travel experience.The primary objectives of the Voyage Genie project are, To design a conversational AI interface that understands natural language travel requests. To generate hyper-personalized itineraries using real-time data sources. To implement a proactive "Plan B Engine" for real-time itinerary adaptation. To automate booking, cancellation, and refund workflows. To optimize travel plans based on time, budget, and user preferences. To provide offline itinerary access and real-time push notifications. To enable post-trip feedback analysis and continuous learning. The expected outcomes of the Voyage Genie project include, A fully functional AI-powered travel assistant. Reduced travel planning time by over 70%. Increased customer satisfaction through personalization. Improved trip success rate despite disruptions. Budget-optimized itineraries with minimal overruns. Proactive issue resolution without user intervention. Scalable architecture for future feature expansion.

VI. Algorithm

Transformers are a type of deep learning architecture originally developed for natural language processing tasks, but they have recently been adapted for computer vision applications. In the context of weld defect detection, Transformers provide a powerful method for capturing global relationships and contextual information in weld images, complementing the local feature extraction capabilities of Convolutional Neural Networks (CNNs).
Key Features of Transformers.

1. **Self-Attention Mechanism:**

The core component of a Transformer is the self-attention mechanism, which allows the model to weigh the importance of different parts of the input image relative to each other.This enables the model to capture long-range dependencies, which is particularly useful for detecting subtle or spatially distributed weld defects.

2. **Global Context Understanding:**

While CNNs focus on local patterns (edges, textures, or small defect features), Transformers can analyze relationships across the entire image, improving the identification of defects that may be distributed or overlapping.

3. **Patch-Based Image Processing:**

Vision Transformers (ViTs) divide images into fixed-size patches, which are then linearly embedded and processed similarly to word tokens in NLP.Each patch interacts with all others through self-attention, allowing the model to learn global structure and defect patterns effectively.

4. **Parallelization:**

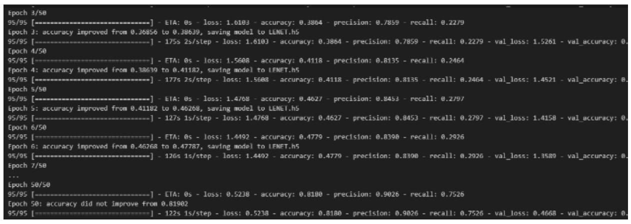
Transformers allow for highly parallel computations due to the absence of sequential operations like in RNNs, which speeds up training and inference.

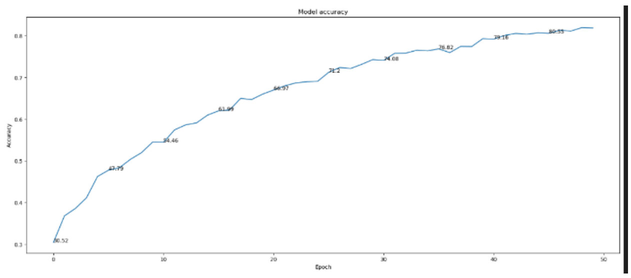
Advantages in Weld Defect Detection:

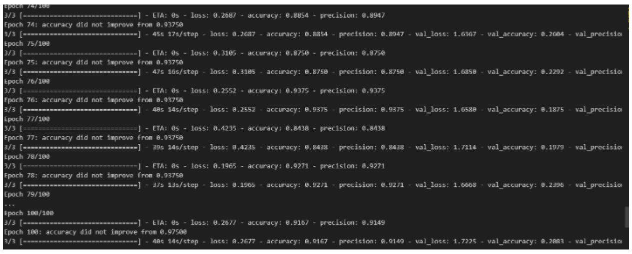
Enhanced Accuracy: Combining CNNs for local feature extraction with Transformers for global understanding improves classification of complex defect types. Detection of Subtle Defects: Transformers can distinguish defects that are small, faint, or context-dependent. Scalability: Transformer-based models can handle images of varying sizes and resolutions without losing contextual information. Complementary to CNNs: CNNs extract detailed texture and edge information, while Transformers provide overall spatial context, creating a robust hybrid system for weld inspection.

Implementation in the Proposed System, Weld images are first processed through CNN layers to extract low-level features like edges, textures, and patterns.These features are then flattened into patches and passed to the Transformer encoder, which captures relationships between different regions of the weld.The combined output is used in a classification layer to detect and categorize weld defects accurately.In summary, the Transformer model significantly enhances the contextual understanding of weld images, making it an essential component of a hybrid CNN-Transformer system for automated and reliable weld defect detection.

VII. Results







Metric	Training	Validation
Loss	0.2677	1.7225
Accuracy	0.9167 (91.67%)	0.2083 (20.83%)
Precision	0.9149 (91.49%)	0.9149 (91.49%)

VIII. Conclusion

This study critically analyzed the YOLOv8n-GSE model for steel surface defect detection and identified limitations in multi-scale adaptability, regression stability, and dataset generalization. To address these challenges, the proposed HARF-Net integrates hybrid attention mechanisms and dynamic regression optimization. Experimental comparison demonstrates improved detection robustness while preserving lightweight characteristics. The hybrid architecture ensures balanced performance between accuracy and efficiency, making it suitable for real-world industrial inspection systems.

IX. Future Scope

Future research will focus on integrating self-supervised pretraining techniques to enhance representation learning under limited labeled data scenarios. Model pruning and quantization strategies will be explored to further reduce computational complexity for ultra-low-power devices. Additionally, domain adaptation techniques will be incorporated to ensure seamless transfer across different industrial materials and imaging conditions. Investigating real-time industrial deployment using FPGA and embedded GPU platforms will also be prioritized.

References

[1] H. Wang, W. Li, B. Zhang and Z. Gu, "YOLOv8n-GSE: Efficient Steel Surface Defect Detection Method," in IEEE Access, vol. 13, pp. 166343-166356, 2025, doi: 10.1109/ ACCESS.2025.3602360.

[2] T. Zhang, H. Pang and C. Jiang, "GDM-YOLO: A Model for Steel Surface Defect Detection Based on YOLOv8s," in IEEE Access, vol. 12, pp. 148817-148825, 2024, doi: 10.1109/ ACCESS.2024.3476908.

[3] X. Luo, Q. Ni, R. Tao and Y. Shi, "A Lightweight Detector Based on Attention Mechanism for Fabric Defect Detection," in IEEE Access, vol. 11, pp. 33554-33569, 2023, doi: 10.1109/ACCESS.2023.3264262.

[4] Y. Wang, H. Wang and Z. Xin, "Efficient Detection Model of Steel Strip Surface Defects Based on YOLO-V7," in IEEE Access, vol. 10, pp. 133936-133944, 2022, doi: 10.1109/ ACCESS.2022.3230894.

[5] X. Zhou, H. Fang, X. Fei, R. Shi and J. Zhang, "Edge-Aware Multi-Level Interactive Network for Salient Object Detection of Strip Steel Surface Defects," in IEEE Access, vol. 9, pp. 149465-149476, 2021, doi: 10.1109/ACCESS.2021.3124814.

[6] H. Z. Wang and Y. T. Wong, "Raster Scanning Can Improve Task Performance in Simulated Prosthetic Vision," in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 33, pp. 4159-4170, 2025, doi: 10.1109/TNSRE.2025.3617891.

[7] C. Zhang et al., "Global-Mapping-Consistency-Constrained Visual-Semantic Embedding for Interpreting Autonomous Perception Models," in IEEE Open Journal of Intelligent Transportation Systems, vol. 5, pp. 393-408, 2024, doi: 10.1109/OJITS.2024.3418552.

[8] X. Zheng, X. Zhang, G. Xu and R. Zhang, "Enhancing Performance of Single-Channel SSVEP-Based Visual Acuity Assessment via Mode Decomposition," in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 31, pp. 4203-4210, 2023, doi: 10.1109/TNSRE.2023.3323000.

[9] D. Caravaca-Rodriguez, G. J. Suaning and A. Barriga-Rivera, "Cross-modal activation of the primary visual cortex by auditory stimulation in RCS rats: considerations in visual prosthesis," 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Glasgow, Scotland, United Kingdom, 2022, pp. 1711-1714, doi: 10.1109/EMBC48229.2022.9871504.

[10] Y. -P. Wuang et al., "Effectiveness of Kinesthetic Game-Based Training System in Children With Visual-Perceptual Dysfunction," in IEEE Access, vol. 9, pp. 153838-153849, 2021, doi: 10.1109/ACCESS.2021.3128109.