International Journal of Science and Engineering Applications
Volume 14-Issue 07, 34 – 39, 2025, ISSN:- 2319 - 7560
DOI: 10.7753/IJSEA1407.1007

Research on Welding Quality Inspection and Visual Anti-attack Protection for Industrial Scenarios

Jingyang Zhou
Yangtze University
Jingzhou, China,434100 China

Gengpei Zhang*
Yangtze University
Jingzhou, China,434100 China

Xiaohan Dou
Yangtze University
Jingzhou, China,434100

Abstract: With the advancement of intelligent manufacturing, automated inspection of welding quality has become a vital direction in industrial vision research. To address the limitations of traditional methods in complex working conditions, this paper proposes a deep learning-based detection framework tailored to industrial welding scenarios, incorporating systematic image enhancement and adversarial defense mechanisms to improve model robustness and security. In the image preprocessing stage, the SCUNet denoising model is adopted to significantly improve image clarity. To address uneven illumination, a DarkIR module based on Retinex theory is introduced, effectively enhancing local texture and contrast in low-light images. For target detection, we construct a YOLOv11 model, which achieves efficient detection of small welding defects through multi-scale feature fusion and attention mechanisms. To tackle the vulnerability of deep models to adversarial attacks, this study designs an Alpha channel attack method and proposes defense strategies including channel purification, simulated perturbation enhancement, and channel attention mechanisms, significantly enhancing system robustness against adversarial samples. Experimental results validate the effectiveness and practicality of the proposed method in improving the accuracy and safety of welding quality inspection, providing technical support for the engineering deployment of industrial vision systems.

Keywords: computer vision, image processing, object detection, security

1. INTRODUCTION

Against the backdrop of rapid advances in intelligent manufacturing, welding—being a fundamental joining process in industrial production—plays a pivotal role in determining the structural integrity and service life of products. Traditional weld quality inspection methods primarily rely on manual visual checks or offline evaluations, which are not only inefficient but also subjective, and thus fall short of the modern industry's demands for high precision, efficiency, and real-time responsiveness. With the rapid development of computer vision and deep learning technologies, automatic visual inspection of weld quality has emerged as a research hotspot and is increasingly applied in industrial scenarios such as automotive, shipbuilding, rail transit, and construction machinery.

In the quality control process of industrial components, surface defect detection remains a critical task. In recent years, the application of computer vision and deep learning has significantly expanded, aiming to improve the accuracy and efficiency of defect detection. High-precision detection models based on the YOLOv5 algorithm [1] have been proposed for surface defect detection in industrial parts, demonstrating the effectiveness of deep learning architectures in this domain. Meanwhile, the application of convolutional neural networks (CNNs) [2] has been extensively studied, highlighting their crucial role in addressing the challenges of defect detection across various industrial applications.

Traditional defect detection approaches rely largely on classical vision-based techniques [3], which are primarily designed to locate and classify defects to facilitate repair and quality assurance. However, such methods often encounter limitations under complex industrial environments, prompting the adoption of more advanced machine learning solutions. Deep learning frameworks that combine regression and classification techniques have also been applied to surface defect detection, offering a comprehensive approach for identifying various types of flaws [4].

To address the challenge of limited annotated data, procedural synthetic data generation has been employed to expand training datasets and improve model robustness and generalization [5]. This approach is particularly valuable in scenarios where labeled images of industrial defects are scarce. Furthermore, deep learning strategies often face obstacles due to insufficient training data, driving the development of systematic frameworks based on synthetic data, especially in the context of additive manufacturing [6]. These frameworks contribute to the detection of geometric defects and ensure the integrity of manufactured components. In the semiconductor industry, defect detection and inspection systems aim to cover a wide range of yield-related applications, underscoring the importance of comprehensive defect detection solutions in high-precision industries. In addition, surveys of surface defect detection methods emphasize the challenges posed by small datasets, and the potential for specialized algorithms to operate effectively under such constraints [7].

The integration of advanced deep learning models, synthetic data generation, and automated machine learning (AutoML) services has become a prevailing trend to improve the accuracy of industrial defect detection. These methods collectively aim to enhance detection precision, reduce dependence on large annotated datasets, and streamline quality control processes across various manufacturing sectors.

However, vision-based inspection systems still face multiple challenges in practical deployment. Under complex working conditions, image noise interference is severe, weld seam appearances vary widely, and deep learning models are vulnerable to adversarial sample attacks during deployment—potentially leading to misclassification and hidden quality risks. Therefore, in addition to improving weld detection accuracy, enhancing the adversarial robustness of vision systems has become a key issue to ensure the safe and intelligent operation of industrial systems.

This study aims to develop an efficient and robust vision-based weld quality inspection system for industrial scenarios, incorporating adversarial defense mechanisms to effectively cope with harsh environments and potential attacks. The proposed approach is expected to improve the practicality and reliability of industrial visual systems, providing theoretical support and technical pathways for quality control in intelligent welding processes and driving the advancement of industrial visual inspection technologies.

2. METHODOLOGY

2.1 Image Illumination Enhancements

In real-world industrial inspection environments, the surfaces of metallic components often exhibit strong shadows, specular highlights, and non-uniform exposure due to the periodicity of geometrical structures, high reflectivity of materials, and dramatic variations in ambient lighting. These factors pose significant interference to vision-based defect detection models. Saeed et al. [8] investigated the impact of lighting variation and makeup on recognition performance, proposing a normalization preprocessing strategy to enhance the model's adaptability to challenging inputs. Hu et al. [9] embedded Retinex theory into the FIN-GAN framework, achieving robust face recognition under complex lighting conditions. Dash et al. [10] combined homomorphic filtering with CLAHE to develop an enhancement strategy for retinal images, thereby improving contrast and standardizing illumination. Meanwhile, Iqbal et al. [11] proposed a face normalization method based on hierarchical feature extraction and histogram processing, and demonstrated its generalizability and robustness across multiple public datasets.

To improve the consistency of input images and robustness to lighting variations, we introduce an image-level illumination normalization module, DarkIR, during the preprocessing stage.

In this study, we adopt an illumination model grounded in classical Retinex theory as the theoretical foundation for lighting correction. The standard Retinex model decomposes an image I(x,y) into a reflectance component R(x,y)and an illumination component L(x,y), such that:

$$I(x,y) = R(x,y) \cdot L(x,y)$$

In the logarithmic domain, this multiplicative model becomes additive, which facilitates subsequent processing:

$$\log I(x,y) = \log R(x,y) + \log L(x,y)$$

To extract reflectance information that is more sensitive to structural defects, we propose a local brightness-constrained strategy, incorporating a joint normalization approach that combines local contrast enhancement with gamma compression. Based on the image's luminance channel, spatial mean filtering is performed to suppress low-frequency components and recover the reflectance term. Simultaneously, the contrast range is dynamically adjusted to enhance the discernibility of fine textures and edge responses.

In implementation, we adopt the Retinex by Adaptive Filtering (RAF) method as the backbone due to its fast deployability, and integrate it with a gamma compression function as follows:

$$I_{norm}(x,y) = \left(\frac{I(x,y)}{G_\sigma(I)}\right)^\gamma$$

Here, G(I) denotes the Gaussian-blurred output of the luminance channel, and γ controls the intensity of nonlinear brightness compression. This method effectively suppresses overexposed highlights in local regions while enhancing fine details in low-light areas. It is particularly effective in reducing pseudo-defect artifacts caused by structural light reflections, such as those encountered in the internal threads of metal components.

2.2 Image Denoising

In industrial weld quality inspection, visual images serve as the primary input for deep learning models, and their clarity and stability directly determine the accuracy of the detection system. However, in practical industrial environments, image quality often suffers from various noise sources, such as welding fumes, specular reflections, metal splatter, and sensor interference. These factors degrade image quality and impair the model's decision-making performance.

Significant progress has been made in the field of image denoising, particularly due to the application of deep learning techniques. Tian et al. [12] conducted a comprehensive review, highlighting the effectiveness of discriminative deep learning methods in dealing with Gaussian noise. Chang et al. [13] proposed a novel spatial-adaptive network (SADNet) for efficient blind denoising, demonstrating that convolutional neural networks (CNNs) can adapt to spatially variant noise characteristics. Gurrola-Ramos et al. [14] introduced a residual dense U-Net architecture, while Xu et al. [15] developed a "Noisy-As-Clean" (NAC) strategy that enables denoising model training without clean reference images—especially effective in signal-dependent real noise conditions. Huang et al. [16] further proposed a "Neighbor2Neighbor" strategy, simplifying the training of denoising models using only noisy images, thereby addressing limitations in previous self-supervised techniques. Quan et al. [17] investigated complex-valued CNNs and their potential in enhancing denoising and image restoration tasks. Tian et al. [18] developed a multi-stage CNN framework integrated with wavelet transforms, leveraging cascaded enhancement blocks to mine structural information. Mansour et al. [19] proposed a lightweight "zero-shot Noise2Noise" network capable of high-quality denoising without training data or prior knowledge of the noise distribution.

In this work, we incorporate a denoising module based on SCUNet (Self-Calibrated U-Net), which effectively captures various types of noise features by combining local modeling (residual convolution) with non-local modeling (Swin transformer). This design enables the model to handle diverse noise patterns, including those commonly encountered in low-light and high-dynamic-range industrial settings. The network architecture is built upon a multi-scale U-Net backbone, which is particularly important for industrial images where both fine details and overall structures must be addressed simultaneously.

By introducing the novel Swin-Conv block, SCUNet balances local detail preservation and global structural integrity, enhancing its denoising performance—especially for industrial images featuring periodic textures or repetitive structures. The model's non-local modeling capacity further improves its ability to suppress noise and recover important visual details.

For blind image denoising tasks, SCUNet follows the maximum a posteriori (MAP) estimation formulation to simulate the denoising process:

$$\hat{x} = \underset{x}{\operatorname{argmin}} D(x, y) + \lambda P(x)$$

To address the diversity of industrial noise, SCUNet adopts a synthetic training data pipeline that simulates combinations of multiple noise types. The dual degradation strategy and random shuffling process improve the model's adaptability to different noise scenarios, thereby enhancing its real-world performance and generalization capabilities.

2.3 Target Detection

Accurate localization and identification of weld defects are central tasks in weld quality inspection. With the evolution of deep learning, convolutional neural network (CNN)-based object detection algorithms have become mainstream. Among them, the YOLO (You Only Look Once) series has stood out. YOLOv3 introduced anchor boxes and multi-scale prediction to enhance detection performance across objects of different sizes, particularly under complex conditions [20]. Pedoeem et al. [21] developed YOLO-LITE to achieve real-time detection at approximately 21 FPS. Gai et al. [22] proposed an improved YOLOv4 model for cherry fruit detection, integrating DenseNet connections and circular bounding boxes to better fit the shape of the targets and improve accuracy. Lawal [23] adapted YOLOv3 for tomato detection, achieving an average precision (AP) as high as 99.5%. Liu et al. [24] introduced Image-Adaptive YOLO (IA-YOLO) to improve detection in adverse weather, addressing the trade-off between image enhancement and detection. Wu et al. [25] applied YOLOv5 with a multi-scale anchor mechanism to enhance small-object detection in remote sensing images. The Spiking-YOLO model [26] employed spiking neural networks to reduce power consumption while maintaining competitive performance on PASCAL VOC and MS COCO datasets, comparable to Tiny-YOLO.

These developments reflect ongoing efforts to balance accuracy and computational efficiency. The rapid evolution from YOLOv1 to YOLOv8 and beyond [27][28][29] highlights its central role in real-time object detection across diverse fields, including robotics, autonomous driving, and industrial manufacturing.

In this study, we adopt the YOLOv11 model for weld defect detection in industrial images. YOLOv11 builds upon the high-speed architecture of previous YOLO models and incorporates multi-scale feature fusion and attention mechanisms to enhance the detection of minute defects such as fine cracks and pores. In addition, the model integrates an anchor adaptation mechanism and lightweight residual modules, significantly reducing parameter counts while maintaining high precision—making it suitable for deployment on embedded devices.

During training, a dataset of real-world industrial weld images containing various typical defect types was used. Data augmentation techniques such as Mosaic augmentation and color perturbation were applied to improve robustness against environmental interference (e.g., smoke, glare). The loss function design incorporates a GIoU/Loss balancing strategy to optimize localization accuracy, and Focal Loss is introduced to suppress the impact of hard-to-classify samples.

Experimental results demonstrate that the YOLOv11 model achieves high recall and low false detection rates in weld defect detection tasks. Its strong adaptability supports real-world deployment in industrial welding inspection systems and provides a reliable foundation for subsequent defect classification and quality assessment.

2.4 Adversarial Defense

In industrial vision-based inspection systems, the security of deep learning models has become an increasingly prominent concern. While these models typically perform well under ideal conditions, they remain vulnerable to adversarial attacks in complex industrial environments, which can compromise the stability and reliability of defect detection. Among these threats, alpha channel-based attacks are particularly insidious and have begun to attract considerable attention. Such attacks manipulate the alpha channel (transparency information) of an image, subtly introducing perturbations without visibly altering the image's appearance. These perturbations can effectively disrupt feature extraction and classification, leading to misclassification, missed detections, or even systemic misjudgment—ultimately posing significant risks to weld quality assessment.

The stealthiness of alpha channel attacks stems from their exploitation of flaws in the model's preprocessing pipeline. In certain image acquisition or processing chains, images may be stored in a four-channel RGBA format. If the deep learning model fails to explicitly discard the alpha channel, it may inadvertently include the malicious perturbation in its input. To mitigate this risk, we propose a multi-faceted defense strategy against alpha channel attacks.

First, at the image preprocessing stage, all input images are forcefully converted to standard three-channel RGB format, thereby removing abnormal signals in the alpha channel and eliminating the attack vector at the source. Second, a simulated perturbation augmentation mechanism is introduced during model training, incorporating images with forged alpha channels to enhance the model's perceptual robustness against such attacks. Third, we integrate a channel attention mechanism into the network architecture, which dynamically monitors activation patterns across channels. This enables the model to identify and suppress potential abnormal responses, thereby enhancing its adaptive defensive capability against adversarial interference.

3. EXPERIMENTS

3.1 Image Preprocessing

In the original weld images, substantial salt-and-pepper noise can be observed around the image borders and on the weld surface. This type of high-frequency noise often arises in industrial image acquisition due to sensor interference, low illumination, or signal compression. Such noise severely hampers feature extraction around the weld edges and molten pool areas, consequently degrading the detection accuracy of the model.

After denoising with the Self-Calibrated U-Net (SCUNet) model, the comparative results (see Figure 1) show a significant improvement in overall visual quality. Details of the weld seam, color gradation, and metallic texture are preserved, while most of the background and structural noise is effectively removed. The edge clarity of the molten pool is visually enhanced, and the transition of the heat-affected zone (HAZ) appears more natural, presenting a more realistic industrial image.

Figure 1. Image Denoising Comparison

We also incorporate DarkIR (Dark Illumination Recovery) technology to enhance low-light weld images during preprocessing. Objective evaluation metrics are used to assess its performance. As shown in Figure 2, the original image suffers from insufficient overall brightness in the weld area, with significant occlusion and information loss in the background and edges. If such images are directly fed into a deep learning model, key weld features are likely to be missed or misinterpreted, leading to false or missed detections and compromising system stability.

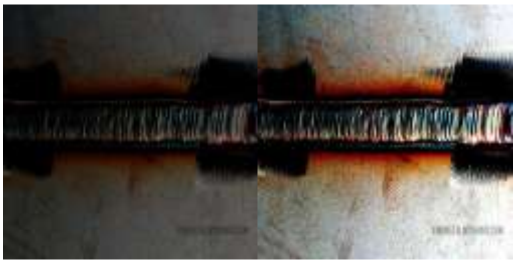

Figure 2. Low-light Image Enhancement Comparison

After enhancement via DarkIR, overall image brightness is significantly improved, and the boundaries of the molten pool and weld seam become clearer. Color gradation and metallic textures are effectively restored. Enhanced images retain the periodic stripe patterns of the weld beads and the boundary transitions of the heat-affected zone, resulting in higher visual contrast and clarity. This aids segmentation algorithms in accurately delineating weld regions and localizing defects.

Quantitative evaluations, shown in Table 1, further confirm the effectiveness of the enhancement. The enhanced images achieve a Structural Similarity Index (SSIM) of 0.842, indicating good preservation of texture and local structural features. The Peak Signal-to-Noise Ratio (PSNR) reaches 20.06 dB, showing higher pixel-level restoration accuracy. These metrics collectively highlight the practical value and balanced performance of DarkIR.

For denoising, the PSNR reaches 29.24 dB, demonstrating excellent noise suppression, while the SSIM is 0.880, indicating effective retention of structural and textural information with minimal distortion. These results confirm SCUNet's practicality for preprocessing industrial weld images under noisy conditions, providing cleaner, high-quality inputs for subsequent defect detection and classification.

Table 1. Evaluation Metrics for Denoising and Low-Light Enhancement

Task	SSIM	PSNR
Low-light enhancement	0.842	20.06
denoising	0.880	29.24

3.2 Target Detection

Figure 3 illustrates the target detection results of industrial weld images based on the deep learning model, covering two quality categories: "Good Weld" and "Bad Weld." Overall, the model is capable of accurately localizing weld regions in most images and assigning high confidence scores. For some "Good Weld" labels, the confidence reaches up to 0.9, indicating that the model has learned and can effectively recognize typical features of high-quality welds.

Moreover, in certain images, the model successfully identifies multiple weld regions simultaneously, demonstrating its capability for multi-target detection and its applicability to complex industrial scenarios involving multiple weld seams. In addition, the model is able to detect weld defects, with labels such as "Bad Weld" or "Defect" achieving confidence scores between 0.5 and 0.8. This suggests the model has a preliminary but functional ability to identify welding anomalies.

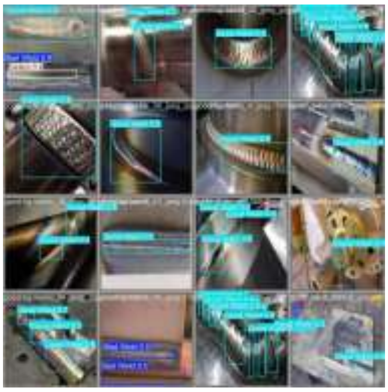

Figure 3. Weld Detection Results

The training curve shown in Figure 4 presents the entire training process of YOLOv11 on the industrial weld dataset, including both training and validation phases. It also displays the convergence behavior of the model and the trend of key performance indicators.

From the loss curves, the initial training loss is relatively high . As training progresses, all loss terms drop sharply and stabilize after around 150 epochs. The validation loss follows a similar pattern, indicating good convergence and no signs of overfitting. Specifically, the final box_loss and cls_loss values converge to around 0.2–0.3, reflecting substantial improvements in localization and classification accuracy.

From the metric curves, the model achieves a precision of approximately 0.85 and a recall of around 0.7, suggesting it maintains a good balance between accurate detection and reduced missed detections. Most notably, the mean Average Precision (mAP) at an IoU threshold of 0.5 reaches approximately 0.87, while mAP@[0.5:0.95] reaches around 0.65. These metrics demonstrate the YOLOv11 model's strong capability for multi-scale target recognition in weld imagery.

As shown in Figure 4, YOLOv11 successfully identifies multiple weld categories. In particular, "Good Weld" targets are assigned high confidence values (e.g., 0.9), highlighting the model's strong learning ability for high-quality weld features. "Bad Weld" and "Defect" targets are also detected with mid-to-high confidence scores (e.g., 0.5–0.8), confirming its preliminary ability to detect weld anomalies. Such capabilities hold considerable value for practical industrial weld quality control.

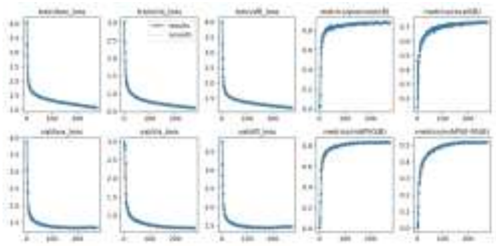

Figure 4. YOLOv11 Training Curve and Detection Performance

3.3 Adversarial Defense Evaluation

Figure 5 illustrates the visual outcomes of an alpha channel attack, comprising three key images: the original weld image (IEye, left), the adversarial perturbation (IAI, center), and the final adversarial sample (IAttack, right). The original image IEye is a typical input to an industrial weld inspection system, containing visible weld defects. The arrow highlights a critical area that should be accurately identified by the deep learning model.

Figure 5 Result of Alpha Channel Attack

The center image IAI represents a perturbation intentionally crafted by an attacker. Although the image content is unrelated to welding (in this case, a white dog), it is normalized and transformed into a noise matrix used for adversarial manipulation. By embedding this perturbation into the alpha channel of the target image and modulating its transparency via a blending formula, an adversarial image (IAttack) is generated. Visually, IAttack is nearly indistinguishable from IEye, and the defect regions remain clearly visible. However, due to the alpha channel perturbation, the deep model's inference is misled, potentially classifying defective regions as "Good Weld" or other incorrect categories—severely compromising detection reliability.

This example demonstrates the stealth and deception potential of alpha channel-based attacks. Even if the visual content remains unaltered, hidden perturbations in auxiliary data channels can introduce security threats that corrupt the model's judgment. Therefore, robust preprocessing strategies are essential for industrial weld inspection systems operating in real-world environments.

To counteract this threat, we propose input-layer interventions to prevent alpha attacks from propagating through the system. Specifically, all input images are screened to ensure they are in standard three-channel RGB format. Any four-channel (RGBA) inputs are either stripped of their alpha channel or rejected entirely. Furthermore, compliance checks based on image headers or channel distributions can help detect anomalous alpha content—such as nonzero transparency or irregular patterns—which may signal a potential adversarial attempt.

Figure 5. Alpha Channel Attack Visualization

4. CONCLUSION

This study addresses the complex conditions and security challenges commonly encountered in industrial welding environments by developing a comprehensive weld quality inspection system that integrates image denoising, illumination enhancement, object detection, and adversarial defense.

To mitigate noise interference in weld imagery, a SCUNet-based denoising module—built upon a hybrid architecture of Swin Transformer and U-Net—was employed. This approach effectively suppresses high-frequency noise and significantly improves image quality. In terms of illumination enhancement, the DarkIR algorithm was introduced to correct uneven brightness and improve contrast, yielding clearer and more informative visual inputs.

For defect detection, the YOLOv11 model was utilized, incorporating multi-scale feature fusion and attention mechanisms. This design enables the model to achieve high precision and recall in weld defect recognition tasks. Experimental results demonstrate its deployability and practical utility in real-world industrial applications.

Furthermore, the study addresses the growing concern of adversarial threats, specifically alpha channel-based attacks. A series of defensive strategies were proposed and evaluated, including input channel filtering, synthetic perturbation-based training augmentation, and attention-based anomaly suppression. These enhancements collectively improve the robustness and safety of industrial visual inspection systems.

The proposed system offers a technically sound and application-ready solution for intelligent weld quality assessment, providing both theoretical guidance and engineering support for industrial deployment.

5. REFERENCES

[1] Chen Z, Feng X, Liu L, et al. Surface defect detection of industrial components based on vision[J]. Scientific Reports, 2023, 13(1): 22136.

[2] Khanam R, Hussain M, Hill R, et al. A comprehensive review of convolutional neural networks for defect detection in industrial applications[J]. IEEE Access, 2024.

[3] Czimmermann T, Ciuti G, Milazzo M, et al. Visual-based defect detection and classification approaches for industrial applications—a survey[J]. Sensors, 2020, 20(5): 1459.

[4] He Z, Liu Q. Deep regression neural network for industrial surface defect detection[J]. IEEE Access, 2020, 8: 35583-35591.

[5] Schmedemann O, Baaß M, Schoepflin D, et al. Procedural synthetic training data generation for AI-based defect detection in industrial surface inspection[J]. Procedia CIRP, 2022, 107: 1101-1106.

[6] Li R, Jin M, Paquit V C. Geometrical defect detection for additive manufacturing with machine learning models[J]. Materials & Design, 2021, 206: 109726.

[7] [7]Jin Q, Chen L. A survey of surface defect detection of industrial products based on a small number of labeled data[J]. arXiv preprint arXiv:2203.05733, 2022.

[8] Saeed U, Masood K, Dawood H. Illumination normalization techniques for makeup-invariant face recognition[J]. Computers & Electrical Engineering, 2021, 89: 106921.

[9] Hu Y, Lu M, Xie C, et al. FIN-GAN: Face illumination normalization via retinex-based self-supervised learning and conditional generative adversarial network[J]. Neurocomputing, 2021, 456: 109-125.

[10] Dash S, Senapati M R, Sahu P K, et al. Illumination normalized based technique for retinal blood vessel segmentation[J]. International Journal of Imaging Systems and Technology, 2021, 31(1): 351-363.

[11] Iqbal M, Riaz M M, Ghafoor A, et al. Illumination normalization of face images using layers extraction and histogram processing[J]. Arabian Journal for Science and Engineering, 2021, 46(4): 3319-3328.

[12] Tian C, Fei L, Zheng W, et al. Deep learning on image denoising: An overview[J]. Neural Networks, 2020, 131: 251-275.

[13] Chang M, Li Q, Feng H, et al. Spatial-adaptive network for single image denoising[C]//Proceedings of the European Conference on Computer Vision (ECCV). 2020: 171-187.

[14] Gurrola-Ramos J, Dalmau O, Alarcón T E. A residual dense U-Net neural network for image denoising[J]. IEEE Access, 2021, 9: 117648-117658.

[15] Xu J, Huang Y, Cheng M M, et al. Noisy-As-Clean: Learning self-supervised denoising from corrupted image[J]. IEEE Transactions on Image Processing, 2020, 29: 9316-9329.

[16] Huang T, Li S, Jia X, et al. Neighbor2Neighbor: Self-supervised denoising from single noisy images[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2021: 14761-14770.

[17] Quan Y, Chen Y, Shao Y, et al. Image denoising using complex-valued deep CNN[J]. Pattern Recognition, 2021, 111: 107634.

[18] Tian C, Zheng M, Zuo W, et al. Multi-stage image denoising with the wavelet transform[J/OL]. arXiv:2202.00790, 2022. https://arxiv.org/abs/2202.00790.

[19] Mansour Y, Heckel R. Zero-shot Noise2Noise: Efficient image denoising without any data[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2023: 4228-4238.

[20] Chen W, Huang H, Peng S, et al. YOLO-face: A real-time face detector[J]. The Visual Computer, 2020, 36(6): 1421-1431. (IF: 5)

[21] Pedoeem J, Huang R. YOLO-LITE: A real-time object detection algorithm optimized for non-GPU computers[J/OL]. arXiv:1805.06361, 2018. https://arxiv.org/abs/1805.06361.

[22] Gai R, Chen N, Yuan H. A detection algorithm for cherry fruits based on the improved YOLO-v4 model[J]. Neural Computing and Applications, 2021, 33(22): 15481-15491.

[23] Lawal M O. Tomato detection based on modified YOLOv3 framework[J]. Scientific Reports, 2021, 11(1): 1447. (IF: 5)

[24] Liu W, Ren G, Yu R, et al. Image-adaptive YOLO for object detection in adverse weather conditions[J/OL]. arXiv:2105.06048, 2021. https://arxiv.org/abs/2105.06048.

[25] Wu W, Liu H, Li L, et al. Application of local fully convolutional neural network combined with YOLOv5 algorithm in small target detection of remote sensing image[J]. PLOS ONE, 2021, 16(5): e0251792.

[26] Kim S, Park S, Na B, et al. Spiking-YOLO: Spiking neural network for energy-efficient object detection[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2020, 34(07): 11262-11270.

[27] Terven J R, Córdova Esparza D M. A comprehensive review of YOLO: From YOLOv1 to YOLOv8 and beyond[J/OL]. arXiv:2304.00501, 2023. https://arxiv.org/abs/2304.00501.

[28] Terven J, Cordova-Esparza D. A comprehensive review of YOLO architectures in computer vision: From YOLOv1 to YOLOv8 and YOLO-NAS[J/OL].arXiv:2307.08582,2023. D

[29] Hussain M. YOLO-v1 to YOLO-v8, the rise of YOLO and its complementary nature toward digital manufacturing and industrial defect detection[J]. Machines, 2023, 11(3): 285.