An EnlightenGAN-based Image Enhancement Algorithm for Defect Detection in Mechanical Parts

Han Jiang, Hongjin Zhu, Honghui Fan, Wenhe Chen

***Abstract*—To address the bottleneck problem of surface defect detection of mechanical parts in industrial low-light environments, this paper presents an improved EnlightenGAN image enhancement algorithm that integrates an attention-convolution dual-path module (ACmix) and a four-way split excitation layer (SaE). Traditional detection systems face three major challenges in low-light scenarios: the high reflective nature of mechanical parts leads to the coexistence of local overexposure and shadow occlusion, complex geometric structures cause multi-scale pseudo-edge interference, and high-frequency confusion between sensor noise and defect texture. Although the existing deep learning methods can improve the detection accuracy under standard illumination, their performance significantly declines in low light conditions due to feature confusion and channel response attenuation. This paper innovatively constructs a grayscale guided multi-scale feature fusion architecture, and realizes the dynamic balance between the local gradient sensitivity of the convolutional path and the global semantic correlation of the attention path through the ACmix module. Effectively distinguish real defects from background textures in complex areas such as gear grooves and bearing raceways. Meanwhile, the SaE module is designed to decouple the channel features into the four-dimensional subspaces of texture, gradient, morphology and luminance, and enhance the channel response intensity of micron-level defects through cross-domain interaction. Experiments show that the improved model achieves a PSNR of 37.69 dB and an SSIM of 0.7737 on the self-built industrial dataset, increasing by 4.06% and 2.06% respectively compared with the baseline model. The recall rate of the YOLOv11 detection model trained with enhanced data has increased by 14.7%, and the mAP50:95 metric has increased by 16.5%, which is significantly superior to the traditional enhancement methods. This method provides a highly robust solution for industrial quality inspection under complex lighting and has significant application value in intelligent quality inspection in fields such as automotive parts and precision machine tools.**

***Index Terms*—EnlightenGAN, mechanical parts defect detection, image enhancement, attention-convolution dual-path module, four-way split excitation layer.**

I. Introduction

ENGPARstartIn the field of industrial automation quality inspection, surface defect detection of mechanical parts plays a crucial role in ensuring product quality. Its accuracy directly affects equipment safety and production efficiency [1]. Traditional inspection systems rely on high-precision optical equipment and stable lighting conditions to identify defects such as cracks, scratches, and pits, using high-frequency texture analysis and local contrast enhancement [2]. However, in real industrial environments, several constraints often arise. Limited installation space can create optical blind spots on certain component surfaces, while complex production conditions frequently result in uneven illumination. Moreover, the three-dimensional geometric structures of mechanical parts—such as gear grooves and bearing raceways—and the highly reflective nature of metal surfaces further complicate optical imaging [3]. For instance, under low-light conditions, insufficient photon flux often causes a significant increase in sensor quantum noise [4]. This non-uniform illumination, combined with the reflective properties of metal surfaces, leads to nonlinear luminance distortions: local overexposed areas suffer from highlight overflow, which obscures texture details, while underexposed regions experience shadow masking and noise interference, resulting in fragmented gradients at defect boundaries [5]. Such photoelectric signal degradation causes systematic shifts in the grayscale distribution of geometric features, making it difficult for traditional threshold-based segmentation algorithms to distinguish polishing textures from actual scratches. This severely limits the intelligent upgrading of industrial quality inspection processes [6].

In recent years, advances in deep learning have allowed convolutional neural networks (CNNs) to demonstrate significant advantages in defect detection tasks under standard lighting conditions, owing to their powerful feature abstraction capabilities [7]. However, under low-light scenarios and in the presence of complex industrial part characteristics, existing models still encounter major challenges [8]. The interaction between highly reflective surfaces and point light sources under low-light conditions causes specular reflections, resulting in extremely nonlinear luminance distributions that invalidate the texture patterns learned during model training. In addition, shadows cast by complex three-dimensional structures can lead to topological distortions, which undermine the multi-scale feature fusion mechanism of Feature Pyramid Networks (FPNs)[9], preventing effective aggregation of semantic information in the presence of pseudo-edges. More critically, the coupling of quantum and circuit noise introduces artifact patterns that closely resemble real defects in the high-frequency domain, making it difficult for Region Proposal Networks (RPNs)[10] to distinguish semantic boundaries between noise disturbances and micron-

Manuscript received May 7, 2025; revised July 24, 2025.

This work is supported in part by the National Key Research and Development Program of China (2023YFF1105102).

Han Jiang is a postgraduate student of Jiangsu University of Technology, Changzhou, Jiangsu 213001 China (email: j.hann@foxmail.com).

Hongjin Zhu is a professor of Jiangsu University of Technology, Changzhou, Jiangsu 213001 China (corresponding author to provide phone: +86-150-6112-7178; fax: 0086-519-86953020; e-mail: zhuhongjin@jsut.edu.cn).

Honghui Fan is a professor of Jiangsu University of Technology, Changzhou, Jiangsu 213001 China (email: fanhonghui@jsut.edu.cn).

Wenhe Chen is a lecturer of Jiangsu University of Technology, Changzhou, Jiangsu 213001 China (email: chenwh@jsut.edu.cn).

level defects in architectures such as YOLO and Faster R-CNN [11], [12]. These issues often result in localization inaccuracies and a simultaneous drop in classification confidence [13].

To address these challenges, this paper proposes an EnlightenGAN-based framework (EnlightenGAN+AcMix+SaE), which integrates the AcMix and SaE modules. The main contributions are summarized as follows:

1) We design a dual-path dynamic fusion architecture (ACMix) that combines attention and convolution mechanisms. By balancing local gradient sensitivity with global semantic correlation through learnable weights, it effectively mitigates the feature confusion problem inherent in traditional single-path models when processing the complex texture backgrounds of metal parts.
2) We construct a four-channel parallel subspace split excitation module (SaE), which decouples channel features into physical attribute dimensions such as texture, gradient, and morphology. By performing cross-domain interactions, it significantly enhances the response strength of defect-related channels in low-light scenes.
3) Extensive experiments demonstrate that the proposed method offers significant improvements in both image enhancement and defect detection tasks. In terms of image quality, the enhanced model—integrating both modules—shows superior noise suppression and detail restoration compared to the baseline and single-module variants, with objective evaluation metrics consistently outperforming benchmark methods. In downstream inspection tasks, the inspection model trained on the enhanced data achieves systematic improvements in accuracy, recall, and generalization ability, validating the effectiveness of the proposed enhancement strategy for industrial quality inspection.

II. Related Work

As a core component of intelligent manufacturing, the development level of industrial defect detection technology directly impacts product quality control as well as the safe operation and maintenance of equipment. In recent years, researchers have explored three innovative directions to address technical bottlenecks in various scenarios: data augmentation and generative methods, specialized physical field detection techniques, and deep learning model optimization. While these approaches have improved detection accuracy, they also reveal inherent technical limitations, highlighting common challenges that urgently need to be addressed in the field of industrial quality inspection. Table I summarizes the main technical focuses of representative studies on industrial defect detection.

In scenarios where samples are scarce and image quality is poor, generative adversarial networks (GANs) and cross-domain transfer learning have become key technologies for overcoming data bottlenecks. To address the problem of insufficient crack samples in the fluorescent penetrant inspection of automotive steering knuckles, article [14] synthesized crack images using a Deep Convolutional Generative Adversarial Network (DCGAN), expanding the dataset scale and significantly improving the classification accuracy of ResNet. However, deviations between the microscopic stress distribution of the synthetic cracks and that of actual damage may lead to overfitting of the model in industrial applications. Article [14] further proposed an innovative decoupled generation strategy for railway fastener inspection: it separately modeled the defect foreground and background and used a skeleton mapping algorithm to control crack morphology directionally, achieving high classification accuracy despite a lack of real defect samples. However, this method relies on manually defined geometric rules, which are insufficient to capture the morphological diversity of complex defects such as oxidation and corrosion, resulting in limited coverage for multiple types of composite defects. In addition, article [16] addressed the problem of blurred imaging of underwater concrete cracks by adopting unsupervised cross-domain transfer technology to transfer crack features from clear water conditions to turbid water areas, thereby optimizing image quality metrics. However, the adaptability of this approach to more complex underwater conditions, such as algal attachment and suspended particle interference, has not yet been verified. Article [17] introduced transfer learning for ultrasonic testing of nuclear power plant containment vessels, leveraging pre-trained parameters from natural images to alleviate data scarcity and improve the segmentation accuracy of U-Net. Nonetheless, significant domain differences exist between the acoustic features of ultrasonic images and the visual features of natural images. These studies show that although generative methods can help overcome data volume limitations, the quality of generated data is constrained by the stability of adversarial training, and cross-scenario generalization capabilities remain unproven at the industrial level. The consistency between synthetic data and real physical mechanisms still needs to be strengthened.

In specialized detection fields such as optics, acoustics, and thermodynamics, multi-physics field coupling analysis is driving the expansion of technical frontiers. Article [18] focused on indentation detection for bearing dust covers, integrating multi-source features such as Scale-Invariant Feature Transform (SIFT), Visual Bag-of-Words (BoVW), and gray-level co-occurrence matrix to build a fusion framework combining handcrafted features with machine learning, achieving relatively high detection accuracy. However, traditional feature extraction is time-consuming and cannot meet the real-time requirements of high-speed production lines, while feature redundancy reduces model training efficiency. Article [19] innovatively developed light reflection enhancement technology: using a spiral stripe model, a 5 μm-level height difference on the inner wall of a small-diameter bearing sleeve was converted into a light emphasis signal, which, combined with a high dynamic imaging system, achieved a high online detection rate. However, the cost of the customized optical platform is high, and oil stains on the workpiece surface can distort the reflection fringes, increasing the false detection rate. Article [20] addressed the debonding issue in concrete structures by proposing a synergistic excitation strategy combining hydration heat and water spray cooling. The infrared thermal contrast was enhanced using a temperature difference matrix algorithm, surpassing the sensitivity limits of traditional methods. How-

TABLE I
FOCUS AREAS OF RELATED WORKS

	[14]	[15]	[16]	[17]	[18]	[19]	[20]	[21]	[22]	[23]	[24]	[25]
Data Generation	Y	Y										
Cross-Domain Transfer			Y	Y								
Physical Field Coupling					Y	Y	Y					
Model Optimization								Y	Y	Y	Y	Y

ever, external water spray may alter the curing rate of concrete, posing a risk of reduced structural strength, and the nonlinear relationship between thermal excitation parameters and defect size increases process control complexity. Although such technologies exhibit outstanding performance in specific scenarios, their hardware dependency, environmental sensitivity, and lack of interpretability restrict large-scale applications. The development of multimodal data fusion algorithms has thus become key to breaking this bottleneck.

Finally, some researchers have focused on improving detection network architectures to enhance defect recognition under complex operating conditions, which also inspires the approach proposed in this paper. For example, article [21] addressed the challenge of detecting microcracks on turbine blade surfaces, where traditional algorithms struggle to simultaneously achieve defect classification. By introducing the Bidirectional Feature Pyramid Network (BiFPN) and employing data augmentation to build a dedicated dataset, the detection accuracy reached 97.4%. However, this model relies heavily on a large amount of labeled data and does not resolve the image distortion caused by high-temperature creep. Article [22] designed a channel-aware aggregation module (CAA) to improve small target feature extraction for detecting overlooked defects in industrial parts, yet the positioning accuracy for defects smaller than 5 μm remains inadequate. Article [23] suppresses background interference in railway fastener images through a Convolutional Block Attention Module (CBAM), but the increased model complexity degrades time performance. Moreover, article [24] innovatively employs U-Net for multi-task transfer learning, achieving an accuracy of 97.57% under occlusion scenarios. However, micro-crack features may be mistakenly removed during the denoising process. Article [25] introduced inverse filtering and super-resolution scaling preprocessing to address motion blur in unmanned aerial vehicle (UAV) images, yet it did not consider the impact of wind speed variations on image stability, and the false alarm rate remains high in dynamic environments.

From the above-mentioned related work, it is evident that although researchers have proposed various solutions for defect detection, performing defect detection on mechanical parts in low-light environments still faces two main challenges. On the one hand, low illumination conditions significantly increase image noise and reduce effective feature contrast; the highly reflective nature of metal parts further exacerbates the uneven distribution of light intensity, blurring the edges of defects such as scratches and microcracks. On the other hand, the complex geometric structures of mechanical parts and the subtle physical characteristics of tiny defects make it difficult for conventional optical imaging to capture detailed features under low-light conditions. Moreover, although data augmentation using generative adversarial networks can expand the sample size, the reflective characteristics of synthesized images are distorted under low light, reducing the model's generalization performance in real-world applications. While multimodal fusion methods can mitigate the limitations of visible light, they entail high equipment costs and strict requirements for surface cleanliness, making them impractical for many real-world scenarios. Improved YOLO models rely on high-contrast feature extraction, but in low-light conditions, channel perception modules are prone to noise interference and may misinterpret background textures as defects.

Based on the above research progress and the remaining challenges, this paper proposes an improved EnlightenGAN image enhancement algorithm that integrates an attention–convolution dual-path module (ACMix) and a four-way split excitation layer (SaE). By combining a lightweight design with a cross-modal feature collaboration mechanism, this method significantly improves the quality of low-light images while maintaining low computational resource consumption, providing an efficient enhancement solution for micron-level defect detection of mechanical parts under complex lighting conditions. The algorithm uses the ACMix module to dynamically balance local detail perception and global semantic association and combines the SaE module's four-dimensional subspace channel excitation strategy to specifically strengthen the high-frequency feature responses of defects such as microcracks and scratches, offering an innovative technical pathway for real-time, high-precision defect identification in industrial quality inspection.

III. PROPOSED DEFECT DETECTION SYSTEM

A. Overall architecture of the model

In the industrial quality inspection scenario, the inspection of mechanical parts in low-light environments has long faced severe challenges. Due to the complex and variable lighting conditions in the production workshop, the surfaces of metal parts are prone to strong reflections or shadow shading, resulting in problems such as low contrast, excessive noise interference, and difficulty in identifying minor defects (such as micron-level scratches and rust) in the collected images. Traditional detection methods rely on manual supplementary lighting or image enhancement algorithms. However, supplementary lighting equipment is costly and difficult to cover complex structures (such as the inner walls of bearings and gear clearances), while traditional algorithms are sensitive to low-light noise. After enhancement, detail loss or artifacts often occur, failing to meet the requirements of high-precision quality inspection. In response to this pain point in the industry, this paper presents an improved EnlightenGAN model, an intelligent enhanced detection system specifically designed for low-light industrial scenarios.

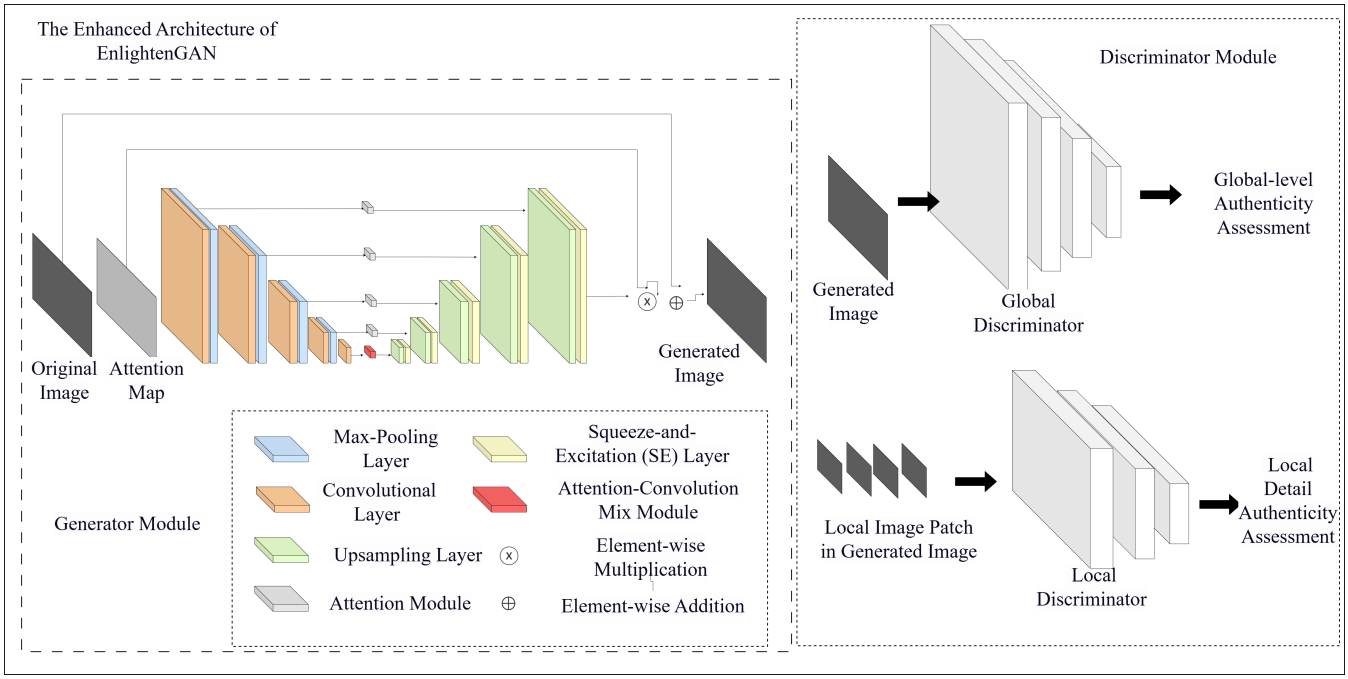

Fig. 1. Architecture diagram of the improved EnlightenGAN model

The improved EnlightenGAN model, as shown in Figure 1, first transforms the input image into multiple levels of grayscale through an original grayscale information extraction module. These grayscale not only contain the original luminance distribution, but also capture structural features at different scales through five downsamplings. In the encoding stage, the model performs channel fusion of the color image with the corresponding level of grayscale image, enabling the network to learn both color information and geometric structures simultaneously, such as the annular patterns on the inner wall of the bearing and the groove distribution of the gear. The attention-convolution Mix module at the end of the encoder adopts a dual-channel parallel design of "Attention + Convolution": The attention path focuses on suspected defect areas (such as discontinuous textures) through the location-aware mechanism, while the convolutional path stably extracts local features (such as scratched edges). After the dynamic fusion of the two, it not only retains the global correlation but also strengthens the local details.

The decoder part introduces a cross-level feature calibration mechanism. During the sampling process on each Layer, the decoder not only receives the high-dimensional features of the previous layer, but also performs spatially aligned weight fusion with the features output by the corresponding layer of the encoder. Among them, the Split and Excite layer divides the feature channels into four groups for independent analysis, respectively learning the response patterns of different defect types (such as corrosion corresponding to low-frequency features and scratches corresponding to high-frequency features), and then highlighting the key channels through adaptive weighting.

B. Attention-convolution hybrid module

In the task of surface defect detection of industrial parts, the visual confusion problem between complex texture backgrounds and tiny defects has long existed. Especially when the defect sizes such as scratches and pits are small, the traditional single-path feature extraction network is difficult to balance local detail sensitivity and global morphological correlation. To address this issue, the improved EnlightenGAN model combines spatial dynamic sensing with local inductive bias by constructing an attention-convolution two-stream co-architecture, thereby enhancing the representation of minor defect features while suppressing background interference.

The architecture of the attention-convolution module is shown in Figure 2.

In the stage of feature projection and channel decoupling of the attention-convolution module, the input feature map $X \in R^{H\times W\times C}$ undergoes three groups of 1×1 convolution to generate three groups of independent features:

$$Q = W_q X, \quad K = W_k X, \quad V = W_v X \tag{1}$$

Each group of features is split into N attention heads (corresponding to the $3N$ feature maps in the figure), with each head having a dimension of $\frac{C}{N}$The dual-path feature fusion stage is divided into two paths, namely the quasi-convolutional path and the attention path. Among them, in the quasi-convolutional path, the fully connected weighted fusion is carried out first. The features at different shift positions are weighted and summed through the learnable weight $W \in R^{K^2\times N}$ to generate the quasi-convolutional features:

$$F_{\text{conv}} = \sum_{k=1}^{K^2} W_k \cdot F'_k \tag{2}$$

Among them, $K = 3$ is the size of the convolution kernel.

Subsequently, a cyclic shift operation is carried out, performing cyclic shifts on the features of each attention head to simulate the local receptive field of the convolution kernel. For example, for the feature map F_K of the k-th head:

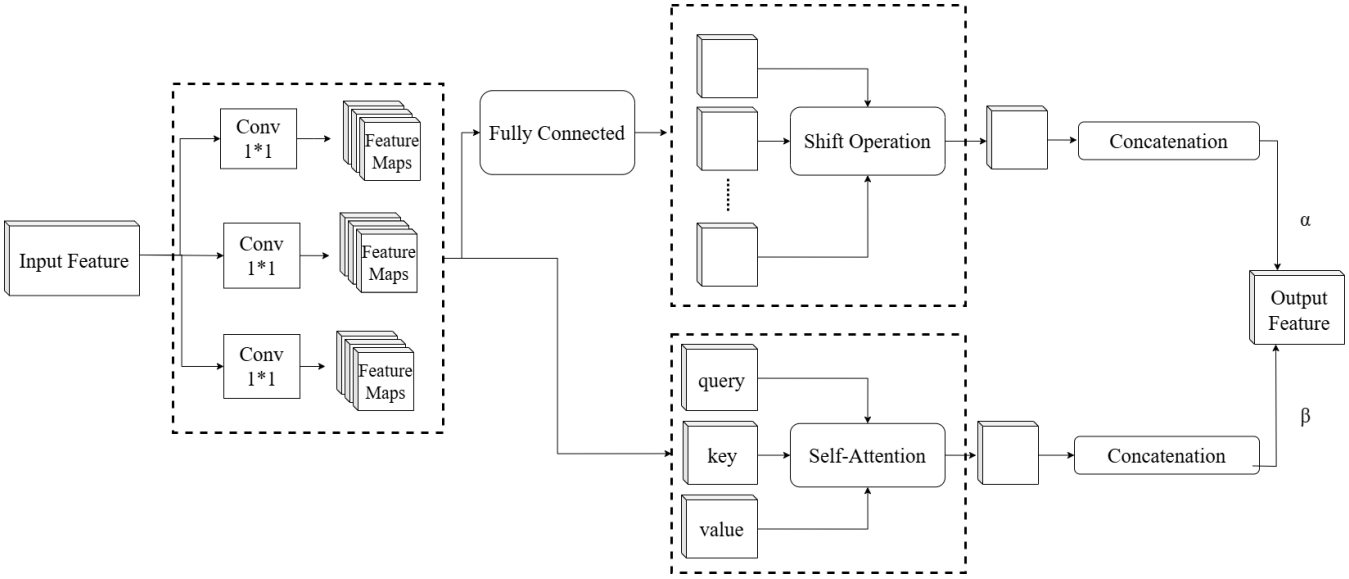

Fig. 2. Architecture diagram of the attention-convolution module

$$F'_k(x,y) = F_k(x+\Delta x_k, y+\Delta y_k) \quad (\Delta x_k, \Delta y_k \in \{-1,0,1\}) \tag{3}$$

For the attention path, spatial attention calculation is carried out first, and the weight matrix is generated by scaling the dot product attention:

$$\text{Att} = \text{Softmax}\left(\frac{QK^T}{\sqrt{d}}\right)V \quad (d = \frac{C}{N}) \tag{4}$$

Among them, the Softmax operation is carried out along the spatial dimension to capture the global dependencies across regions (such as the crack extension direction), and then the inter-head feature concatenation is performed, that is, the outputs of N attention heads are concatenated along the channel axis to restore the original dimension C.

Subsequently, the outputs of the two paths are dynamically fused through the learnable parameter α, β:

$$F_{\text{out}} = \alpha \cdot F_{\text{conv}} + \beta \cdot \text{Att} \quad (\alpha + \beta = 1) \tag{5}$$

From the above content, it can be seen that the ACmix module adopts a dual-path collaborative architecture, effectively integrating the local inductive bias of convolution and the global correlation advantages of attention, significantly improving the detection accuracy of minor defects in industrial scenarios. Its workflow is divided into three key stages: Firstly, the input features are decoupled into multiple query, key and value projections through three independent one-dimensional convolution groups, and fine-grained feature expressions are formed through a multi-head splitting mechanism; Subsequently, two heterogeneous processing methods are executed in parallel - the quasi-convolutional path uses cyclic displacement operations to simulate the local perception characteristics of traditional convolutional kernels, captures the gradient mutation features of micrometer-level defects through nine preset offset modes, and at the same time, the attention path establishes a cross-regional long-range dependency relationship to accurately identify complex morphological patterns such as crack extension directions; Finally, the dual-path output is dynamically balanced through the adaptive weight fusion mechanism, which not only retains the spatial invariance advantage of convolution for regular textures, but also enhances the response intensity of attention to abnormal regions.

C. Splitter and excite layer

The SaE (Splitter and Excite) layer is a multi-branch channel attention mechanism. As shown in Figure 3, this part is improved to address the limitations of the single excitation path in the traditional SENet. The SaE layer achieves four-way parallel subspace excitation and hierarchical feature reorganization. Enhance the fine-grained modeling ability of the model for the channel characteristics of minor defects on the surface of industrial parts.

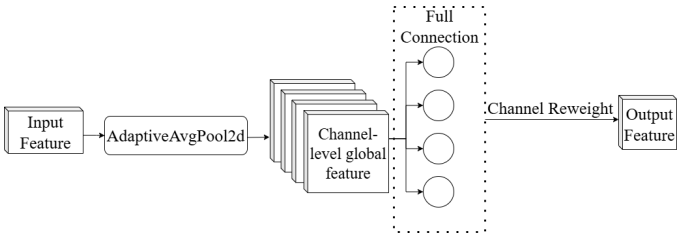

Fig. 3. Architecture diagram of the split excitation layer

First, perform global feature compression. Global feature compression $X \in R^{B\times C\times H\times W}$ compresses the spatial dimension through adaptive average pooling and extracts channel-level global information:

$$y = \text{AdaptiveAvgPool2d}(X) \in R^{B\times C} \tag{6}$$

Here, spatial information is encoded as channel descriptors, providing a basis for subsequent multi-branch excitation.

Subsequently, the multi-subspace splitting excitation achieves feature decoupling and subspace learning through four independent fully connected layers:

$$\begin{cases} y_1 = \mathrm{ReLU}(W_1 y) \in R^{B\times C/2} \\ y_2 = \mathrm{ReLU}(W_2 y) \in R^{B\times C/2} \\ y_3 = \mathrm{ReLU}(W_3 y) \in R^{B\times C/2} \\ y_4 = \mathrm{ReLU}(W_4 y) \in R^{B\times C/2} \end{cases} \tag{7}$$

Among them, $W_1, W_2, W_3, W_4 \in R^{(C/2)\times C}$ are learnable parameter matrices. Each group of parameters learns different feature subspace expressions, enabling the module to capture the differentiated responses of part surface defects in different physical characteristic dimensions (such as texture, gradient, brightness).

Finally, the excitation weights are broadcast back to the original feature map size to perform channel-level feature enhancement:

$$X_{\mathrm{out}} = X \otimes y_{\mathrm{excite}} \in R^{B\times C\times H\times W} \tag{8}$$

Among them, $\otimes$ represents channel-by-channel multiplication. This operation enhances the response to defect-sensitive channels (such as gradient change channels in scratch areas) and supholds background texture interference.

From the above description, it can be known that the SaE (Split Excitation) layer plays a core role in the refinement enhancement of channel features in the industrial defect detection model. Under the condition of surface imaging of parts in low light and high noise, this module decouples the global channel features into differentiated physical attribute subspaces such as texture, gradient, morphology, and brightness through a four-channel parallel subspace excitation mechanism. The multi-dimensional feature decoupling expression of micron-level defects has been achieved. Specifically, the module first compresses spatial redundancy information through adaptive average pooling to generate channel-level global descriptors; Subsequently, four groups of independent fully connected layers are utilized to respectively learn the nonlinear mapping relationships of different attribute subspaces, overcoming the feature coupling problem caused by a single excitation path; Ultimately, through cross-subspace feature fusion and channel recalibration, a response intensity several times that of the background is generated in typical defect areas such as gear tooth surface scratches, while suppressing the activation values of interference factors such as metal reflection to provide robust feature enhancement support for high-precision quality inspection in complex industrial environments.

IV. Experimental Analysis

A. Experimental design

The design of the experiment focused on verifying the performance improvement of the EnlightenGAN model that combines ACmix and SaE modules. A two-stage validation framework of ablation experiments and contrast experiments was adopted to systematically evaluate the contribution of the modules and the advancement of the overall approach. The dataset selected for this study included a real dataset of 1,885 defect images of mechanical parts. The data source is the publicly available aluminum profile surface defect identification dataset by Tianchi. The training set and test set are divided in an 8:2 ratio. The dataset contains single-defect images, multi-defect images, and defect images under low light. The total number of defect types contained in the images is 10, namely: Non-conductive, scratches, exposed base at corners, orange peel, exposed base, spray, paint bubbles, pits, discoloration, dirt spots. The configuration table of the experimental equipment is shown in Table II. In terms of hardware configuration, the graphics processing unit consists of a parallel computing array composed of four NVIDIA GeForce RTX3090 graphics cards. Each card is equipped with 24GB GDDR6X video memory, and the total capacity of the video memory reaches 96GB, supporting multi-task concurrent training and large-scale tensor operations. The central processing unit adopts the Intel Xeon Platinum 8280L professional-grade server CPU with a main frequency of 2.60GHz and a 56-core 112-thread architecture to ensure the efficient execution of CPU-intensive preprocessing tasks. The system memory is configured with 32GB DDR4 ECC at a frequency of 2933MHz, effectively supporting high-throughput data stream processing. At the software environment level, a stable computing ecosystem is constructed based on the Linux operating system. Anaconda3 is adopted for dependency management and environment isolation. The deep learning framework selects PyTorch version 1.10 to be compatible with CUDA 11.3 for accelerated computing, and the training process is visually monitored through Tensorboard.

TABLE II
Experimental Equipment Configuration Table

Name	Disposition
GPU	4 x NVIDIA GeForce RTX3090
CPU	Intel(R) Platinum 8280L CPU @ 2.60GHz
Video memory	24G x4
Internal memory	32G
Operating system	Linux
Training framework	Pytorch1.10
Package management tool	Anaconda3
Visual tool	Tensorboard

Detection results of YOLOv11 on augmented images are shown in Figure 4.

Each image in it is labeled with the original file name and the detected defect category. The detection box adopts a multi-color annotation strategy to visually distinguish the recognition results of different defects.

The indicators of the experiment take PSNR (Peak Signal-to-Noise Ratio) and SSIM (Structural Similarity) as the core indicators. These two indicators are respectively used to quantify the noise suppression ability and the detail fidelity.

Among them, PSNR is used to measure the quality of image reconstruction and is expressed as:

$$\mathrm{PSNR} = 10 \cdot \log_{10}\left(\frac{\mathrm{MAX}^2}{\mathrm{MSE}}\right) \tag{9}$$

Here, MAX represents the maximum possible value of the pixel value of the evaluated image. For example, the pixel value of an 8-bit image is generally 255, while that of a normalized image is 1. MSE (Mean Square Error) is used to represent the mean square of the difference in pixel values between the original image x and the evaluated image y:

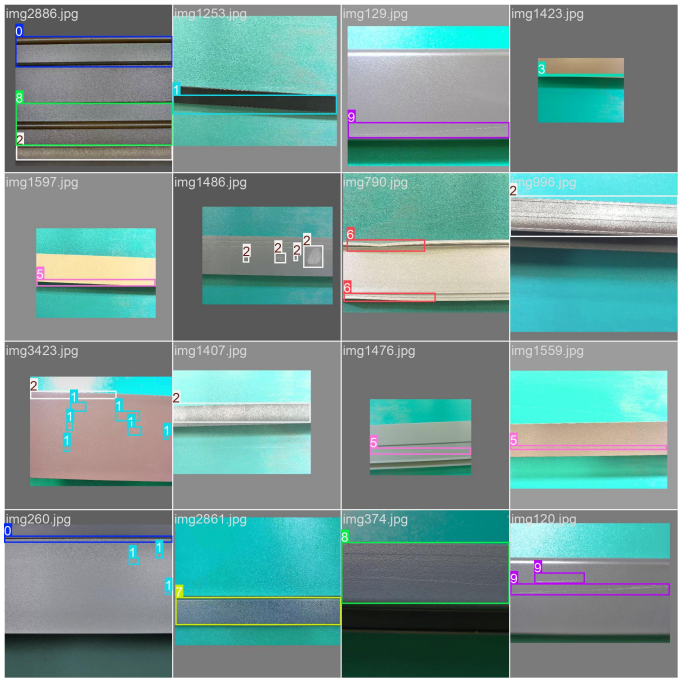

Fig. 4. Detection results

$$\mathrm{MSE} = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}(x_{ij} - y_{ij})^2 \tag{10}$$

M and N represent the width and height of the evaluated image respectively.

SSIM measures image similarity from three aspects: brightness, contrast and structure. The value range is $[-1, 1]$. The closer the value is to 1, the more similar it is, which is expressed as:

$$\mathrm{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{11}$$

Among them, μ_x and μ_y respectively represent the mean pixel values of images x and y; σ_x and σ_y represent the pixel standard deviations of images x and y respectively, and σ_{xy} is the covariance of images x and y; C_1 and C_2 are stable constant, defined as:

$$C_1 = (k_1 L)^2, \quad C_2 = (k_2 L)^2 \tag{12}$$

Among them, L represents the dynamic range of pixel values.

For the defect detection task, four indicators, namely Precision (accuracy rate), Recall (recall rate), mAP50 (average accuracy with an IoU threshold of 0.5), and MAP50:95 (comprehensive accuracy with multiple thresholds), are adopted to comprehensively measure the positioning accuracy and robustness of the detection model.

The ablation experiment focuses on the module-level performance contribution, taking the image enhancement quality (PSNR/SSIM) as the evaluation dimension. Through the comparison of four groups of progressive models (basic model, single-module enhancement, and dual-module combination), To quantify the independent gain and synergy effect of the ACmix and SaE modules. ACmix enhances the local detail recovery ability through convolution - attention dual-path fusion, while SaE optimizes the cross-regional feature consistency through multi-subspace channel excitation. When the two modules are combined, the nonlinear superposition of performance gain is achieved to verify the design logic of complementary enhancement between the modules. Subsequently, in order to further verify the quality of image enhancement, the same YOLOv11 detector was trained respectively on the low-light original data and the enhanced data. Moreover, four indicators, namely Precision (precision rate), Recall (recall rate), mAP50 (average precision with an IoU threshold of 0.5), and MAP50:95 (comprehensive precision with multiple thresholds), are adopted to comprehensively measure the positioning accuracy and robustness of the detection model after training with different data. The comparative experiments verify the advancement of the method at the task level. Taking YOLOv11 as the detection baseline, multiple advanced image enhancement schemes are first adopted to enhance the pictures respectively. Then, the performance advantages of the image enhancement method proposed in this paper are verified according to four indicators.

B. Ablation experiment

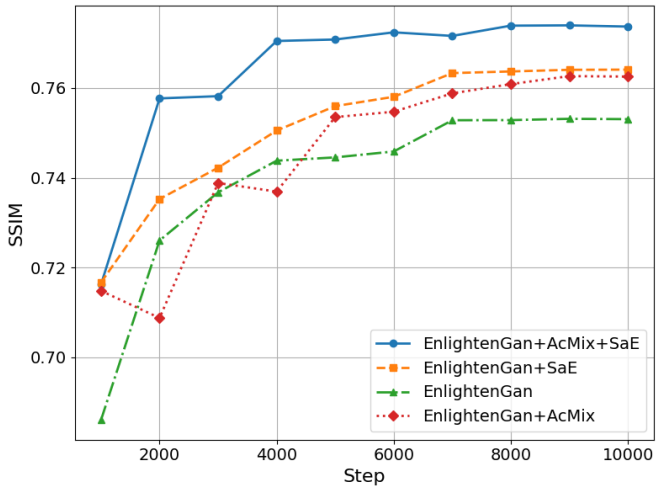

Fig. 5. Ablation experiment diagram (SSIM index)

As can be seen from the comparison of the training process curves and the final metrics in Figure 5, the EnlightenGan+AcMix+SaE model with both AcMix and SaE modules shows a significant performance advantage, and its SSIM metric shows a steady upward trend with increasing training steps. And reached the maximum value (0.7737) after 10,000 steps of training, which was 2.06% higher than the baseline model EnlightenGan (SSIM 0.7531). Although the single-module improvement of EnlightenGan+SaE (SSIM 0.7641) and EnlightenGan+ACmix (SSIM 0.7626) correspond to independent optimizations of SaE and ACmix respectively, However, the performance of both still lags behind that of the complete model of the dual-module joint optimization (with a difference of 0.0096 and 0.0111), verifying the synergistic enhancement effect between the modules. The dual-module model did not significantly deteriorate in terms of training efficiency. Its training duration was 28.98 minutes, which was

12.9% less than that of the baseline model (33.33 minutes) and better than the 33.78 minutes of the SaE single-module model. This indicates that the optimization of computational efficiency in the parallelized design effectively offset the increased parameter burden of the module.

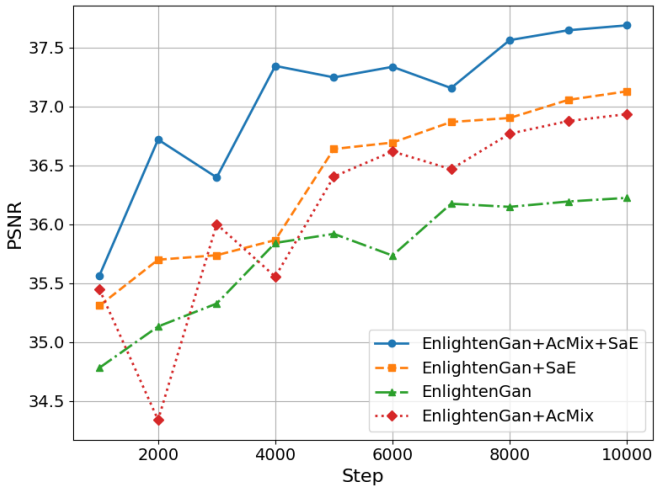

Fig. 6. Ablation experiment diagram (PSNR index)

As can be seen from the PSNR metric analysis in Figure 6, EnlightenGan+AcMix+SaE, which integrates both AcMix and SaE modules, leads significantly with a peak signal-to-noise ratio of 37.69. An improvement of 4.06% from the baseline model EnlightenGan (PSNR 36.22), which validates the effectiveness of the two-module co-optimization. The progressive modular ablation experiments indicated that the individual integration of ACmix (PSNR 36.93) and SaE (PSNR 37.13) brought gains of 1.96% and 2.51%, respectively. When the dual-module joint optimization is carried out, the improvement amplitude of PSNR (4.06%) exceeds the arithmetic superposition of the independent gains of the two (1.96%+2.51%=4.47%), revealing the nonlinear complementary effect between the two modules. ACmix enhances local detail recovery through the convolution - attention dual-path. It provides a more robust feature input for the multi-subspace channel excitation of SaE, while the channel recalibration of SaE further optimizes the feature selection process of ACmix.

Subsequently, this paper summarizes the indicator information of the two graphs, as shown in Table III.

TABLE III
SUMMARY TABLE OF ABLATION EXPERIMENT RESULTS

Method	PSNR	SSIM
EnlightenGan	36.22	0.7531
EnlightenGan+AcMix	36.93	0.7626
EnlightenGan+SaE	37.13	0.7641
EnlightenGan+AcMix+SaE	37.69	0.7737

Subsequently, the same YOLOv11 detector was trained on the original low-light data and the enhanced data respectively. The experimental results of the low-light data and the enhanced data are shown in Table IV and Table V respectively.

Subsequently, this paper summarizes and analyzes the experimental result information of the two figures, as shown in Table VI.

TABLE IV
EXPERIMENTAL RESULTS OF LOW-LIGHT DATA

Class	Instances	Box (P)	R	mAP50:95
All	661	0.676	0.573	0.358
Rust	23	0.659	0.826	0.762
Dirt	172	0.682	0.128	0.248
Non-conductive	179	0.776	0.704	0.782
Dent	12	0.348	0.058	0.170
Pore	76	0.592	0.526	0.638
Pit	33	0.854	0.909	0.913
Coating Crack	22	0.922	0.955	0.979
Local Defect	50	0.777	0.760	0.779
Scratch	58	0.546	0.362	0.431

TABLE V
EXPERIMENTAL RESULTS OF ENHANCED DATA

Class	Instances	Box (P)	R	mAP50:95
All	661	0.681	0.657	0.417
Rust	23	0.603	0.739	0.748
Dirt	172	0.655	0.267	0.315
Non-conductive	179	0.797	0.749	0.789
Dent	12	0.806	0.695	0.732
Pore	76	0.626	0.566	0.627
Pit	33	0.730	0.848	0.845
Coating Crack	22	0.821	1.000	0.988
Local Defect	50	0.750	0.800	0.792
Scratch	58	0.461	0.328	0.381

TABLE VI
PERFORMANCE COMPARISON OF YOLOV11 IN LOW-LIGHT RAW DATA AND ENHANCED DATA

Method	Precision	Recall	mAP50	mAP50:95
Low-light data	0.676	0.573	0.626	0.358
Enhance data	0.681	0.657	0.680	0.417

The experimental results show that the YOLOv11 model trained with EnlightenGAN+ACmix+SaE enhanced data achieves systematic performance improvement in industrial defect detection tasks. Compared with directly using the original data in dim light, the enhanced data significantly increased the recall rate of the detection model by 14.7% (0.573→0.657), and mAP50 and MAP50:95 increased by 8.6% (0.626→0.680) and 16.5% (0.358→0.417), respectively. The accuracy rate also increased slightly by 0.74% (0.676→0.681). This data reflects the multi-dimensional optimization of the model performance by the enhancement strategy: The leap in recall rate indicates that the enhanced data effectively improves the identifiability of minor defects in low-light environments, significantly reducing the missed detection rate of the model. The significant improvement of mAP50:95 (with an average accuracy of 0.5-0.95 for the comprehensive IoU threshold) verifies the positive effect of the enhanced data on the positioning accuracy of the detection frame. That is, by suppressing the interference of low-light noise and restoring the defect edge gradient, the detection stability of the model under strict IoU thresholds (such as 0.75-0.95) is enhanced, which is particularly important for high-precision industrial quality inspection scenarios. From the analysis of task correlation, the asymmetric gains of

mAP50 and MAP50:95 (+8.6% vs +16.5%) indicate that the enhanced data has a more significant improvement for high-threshold detection tasks. This is attributed to the precise modeling of defect geometricia by the convolution - attention collaboration mechanism of the ACmix module. And the directional enhancement of abnormal channel features by the SaE module.

C. Comparative experiment

To evaluate the performance of the proposed Enlighten-GAN+ACmix+SaE image enhancement method in mechanical parts defect detection task, five representative low-light image enhancement methods were selected as comparison objects in this paper. It covers different categories such as lightweight networks based on traditional visual priors, Retinex theoretical models, and Transformer structures. GCA-Net [26] achieves global perception of defect regions and multi-scale feature aggregation under the lightweight coder-decoding framework by introducing the deep self-attention mechanism and the channel reference attention module, effectively enhancing the representation ability in complex backgrounds. The KinD [27] method, based on the Retinex theory, decomposes the image into two subspaces: illumination and reflection. It models illumination and noise reduction respectively to enhance the image brightness while suppressing noise and artifacts. Retinexformer [28] introduces an illumination-guided Transformer structure based on the traditional Retinex framework to achieve non-local modeling and corrosion recovery among different illuminated areas, improving the consistency and detail restoration ability of low-illumination image enhancement; Zero-DCE [29] achieves adaptive adjustment of the pixel dynamic range through a no-reference depth curve estimation strategy without the need for paired training samples, and has strong generalization ability and real-time performance. CPGA-Net [30] integrates the dark/bright channel prior and the gamma correction strategy to construct a lightweight network architecture, achieving excellent enhancement effects with extremely low computational overhead, and further comreducing the model size through knowledge distillation. In the comparative experiments, all methods will conduct multi-stage index evaluations under a unified training configuration, including the analysis of the changing trend of pixel loss during the training process, the quantitative comparison of the enhancement results in terms of structural similarity Index (SSIM) and peak signal-to-noise ratio (PSNR). In addition, by inputting different enhancement results into the YOLOv11 defect detection model, The PR curve of the detection performance was further plotted to analyze the promotion degree of each enhancement method on the downstream detection tasks.

Figure 7 shows the trend of Pixel Loss of various image enhancement methods on low-illumination images during the training process varying with the number of training steps, reflecting the optimization efficiency and stability of each model in terms of the pixel accuracy of reconstructed images. From the overall trend of the curve, except for Zero-DCE, the Pixel losses of the other methods all show obvious typical training convergence characteristics of a rapid decrease at first and then gradually stabilizing. The loss value of Retinexformer in the initial stage is as high as approximately 180,

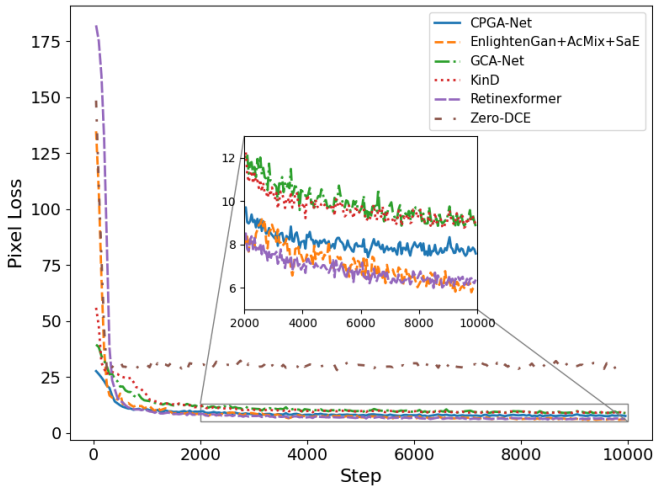

Fig. 7. The discrimination loss of the generated image

which is much higher than that of other methods. However, the decline rate is the fastest, and it has dropped to about 10 by step=1000. This indicates that although the initial convergence of its Transformer structure is slow, it has a strong global modeling ability and can significantly compress the reconstruction error in the short term. The initial Pixel Loss of KinD is approximately 50 and rapidly decreases to about 10 within 1000 steps, demonstrating the effectiveness of the Retinex decomposition mechanism in modeling the illumination structure of low-illumination images. The initial loss of GCA-Net and EnlightenGAN+ACMix+SaE was about 30-35, and the decline was relatively smooth. Eventually, after Step=4000, it stabilized and converged to about 8-9. It is indicated that its local attention and edge feature perception module has a good promoting effect on the restoration of image texture details. The initial loss of CPGA-Net is slightly lower than that of GCA-Net, but the decline rate is slightly slower. Eventually, it stabilizes at approximately 9, which is basically comparable to the aforementioned two methods. This indicates that it has achieved better performance relying on channel correction and gamma estimation strategies in a lightweight structure. Zero-DCE shows the worst convergence trend throughout the training process: the initial loss is approximately 145, and it still remains between 25 and 30 when step=10000. The curve oscillates significantly, indicating that although its unsupervised optimization strategy has strong generalization and flexibility, it has obvious limitations in precise pixel recovery and is difficult to approach the low-loss state.

Combining the above results, Retinexformer performs best in both the rate of loss reduction and the final convergence value, while EnlightenGAN+ACMix+SaE achieves a final pixel loss comparable to or even lower than GCA-Net and CPGA-Net while keeping the structure simple. The effectiveness of its edge perception module and attention enhancement strategy in low-illumination image enhancement was verified. In contrast, the performance of Zero-DCE shows its insufficiency in the ability to restore pixel accuracy, making it difficult to meet the requirements of subsequent high-precision vision tasks.

Figure 8 shows the changes in the Structural Similarity Index measure (SSIM) for each method during training. En-

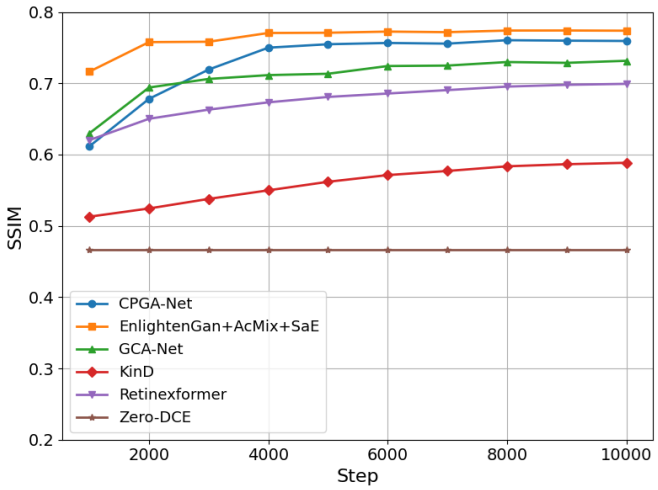

Fig. 8. Comparative Experiment Diagram (SSIM Index)

lightenGAN+AcMix+SaE consistently maintained the lead in SSIM, rising from the initial 0.72 to 0.775, reflecting its superiority in image structure preservation. CPGA-Net performed steadily, with the final SSIM reaching 0.76, slightly lower than EnlightenGAN, but significantly better than GCA-Net (0.735) and Retinexformer (0.70). KinD's performance was relatively limited, and the increase in SSIM was small, rising only from 0.515 to 0.585. The SSIM curve of Zero-DCE is almost horizontal and stable at around 0.47, indicating that it has significant deficiencies in the ability to reconstruct image structure information.

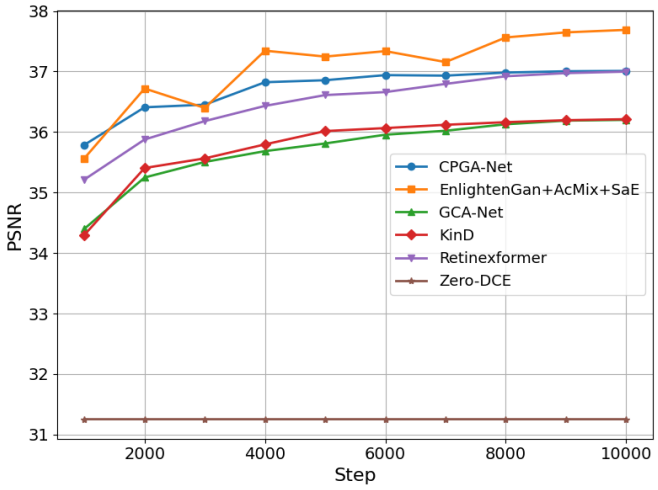

Fig. 9. Comparative Experiment Diagram (PSNR Index)

Figure 9 shows the variation trend of the peak signal-to-noise ratio (PSNR) during the training process. EnlightenGAN+AcMix+SaE performed best at all stages, with PSNR increasing significantly from an initial 35.6 to 37.7. CPGA-Net steadily improved and finally reached 37.0, closely following. Although Retinexformer was slightly lower in the early stage, the upward trend was stable and finally reached 36.9, which was superior to GCA-Net (36.2) and KinD (36.3). Zero-DCE performed the worst. The PSNR remained around 31.2 all the time with almost no significant improvement, indicating its limited ability to suppress image noise.

TABLE VII
SUMMARY TABLE OF COMPARATIVE TEST RESULTS

Method	PSNR	SSIM
EnlightenGan+AcMix+SaE	37.69	0.7737
CPGA-Net	37.02	0.7618
GCA-Net	36.28	0.7351
Retinexformer	36.92	0.7036
KinD	36.39	0.5854
Zero-DCE	31.25	0.4709

As can be seen from Table VII, EnlightenGAN+AcMix+SaE achieved the best results in both PSNR and SSIM, demonstrating its combined advantage in luminance enhancement and structure preservation. The CPGA-Net proposed in this paper follows closely in two indicators, demonstrating excellent image restoration ability and stability, and verifying the effectiveness of its multi-scale perception and global guidance mechanism. GCA-Net and Retinexformer perform relatively stably, but are slightly lacking in detail reconstruction. The overall improvement of KinD is limited and its ability to maintain the structure is relatively weak. However, Zero-DCE, as an unsupervised method, performs the worst in two indicators and is difficult to be competent for the task of high-quality image enhancement.

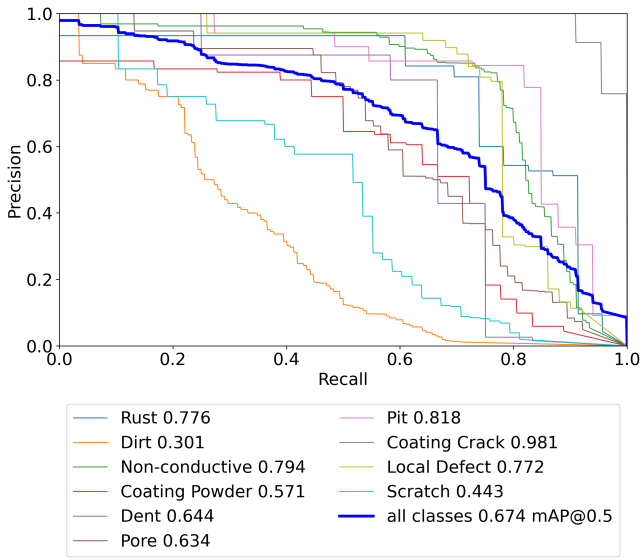

Fig. 10. Precision - Recall curve

The final precision-recall curves of YOLOv11 enhanced by EnlightenGAN+AcMix+SaE are shown in Figure 10. From the experimental results, the model performs well in several key defect categories, demonstrating a significant advantage. Coating Crack achieved nearly perfect recognition with an accuracy of 0.988, indicating that the model has an extremely strong ability to capture its texture changes. The accuracies of Pit (0.845) and Non-conductive (0.789) both exceed the industry benchmark, demonstrating the stable detection ability of the model for regular shape defects. Moreover, Rust (0.748) and Dent maintain high accuracy at

a medium recall rate and have good balance. This is applicable to application requirements that are sensitive to false detection rates in actual scenarios. Overall, the high-precision performance of the model in the main defect categories has verified its practicality and reliability in industrial quality inspection.

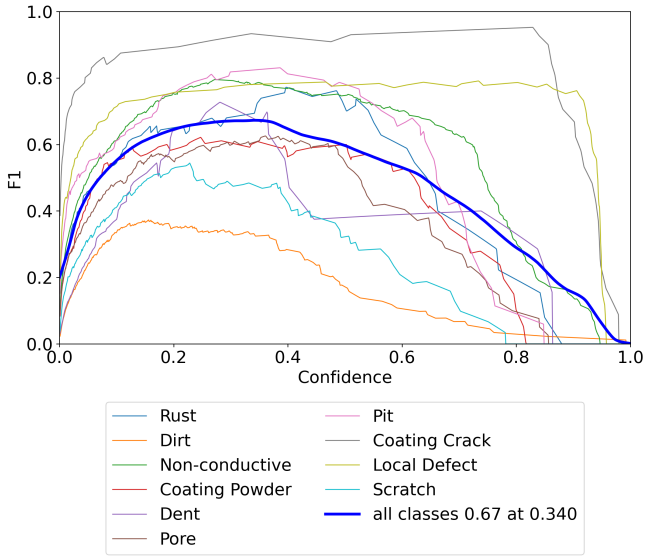

Fig. 11. F1-Confidence curve

The F1-Confidence curves of YOLOv11 enhanced by EnlightenGAN+AcMix+SaE are depicted in Figure 11. It can be observed that the model maintains high F1 scores across a wide range of confidence thresholds, which demonstrates its robustness and reliability under varying detection strictness. Specifically, the F1 scores for Coating Crack remain close to the maximum over nearly the entire confidence interval, indicating excellent consistency between precision and recall for this category. For Pit and Non-conductive defects, the F1 scores also stay at relatively high levels without significant drops, reflecting the model's stable balance in detecting regular shape anomalies even when stricter confidence thresholds are applied. Although Rust and Dent show slight declines in F1 score with increasing confidence, they still achieve stable plateaus within practical confidence ranges, which is favorable for industrial applications that require adjustable trade-offs between false positives and false negatives. Overall, the F1-Confidence analysis further corroborates that the proposed improvements enable the model to sustain high detection accuracy and robustness, ensuring dependable defect recognition in complex real-world scenarios.

V. Conclusion and Future Work

This paper addresses the challenge of accurately detecting surface defects on mechanical parts in low-light industrial conditions. It proposes an improved EnlightenGAN image enhancement algorithm that incorporates an attention convolution dual-path module (ACmix) and a four-way split excitation layer (SaE). A grayscale-guided multi-scale feature fusion framework is constructed to combine five-level grayscale downsampling with color channel information, enhancing both local gradient sensitivity and global semantic correlation. The ACmix module simulates diverse local perception through cyclically shifted deformable convolutions and models global context with multi-head attention, improving feature fusion in complex areas such as gear grooves and bearing raceways. The SaE module decouples channel features into texture, gradient, morphology, and luminance subspaces, and improves the SSIM contrast of micron-level defects through cross-domain interaction.

Experiments show that the improved model achieves a PSNR of 37.69 dB and an SSIM of 0.7737, with improvements of 4.06% and 2.06% over the baseline. The YOLOv11 detection model trained with enhanced data achieves a 14.7% increase in recall and a 16.5% improvement in mAP50:95, outperforming traditional methods. This approach offers a robust solution for industrial quality inspection under complex lighting and holds application potential in areas such as automotive parts and precision machinery. Future work will explore temporal modeling for video-based defect tracking, few-shot and unsupervised learning for limited-sample scenarios, and cross-modal enhancement using infrared and ultrasound for multi-channel defect detection.

References

[1] F. Zhou, Y. Chao, C. Wang, X. Zhang, H. Li, X. Song, "A small sample nonstandard gear surface defect detection method," *Measurement*, vol. 221, 2023, 113472.
[2] Z. Zhang, Z. Zelin, et al., "Surface defect detection method for discarded mechanical parts under heavy rust coverage," *Scientific Reports*, vol. 14, no. 1, 2024, 7963.
[3] G. R. Satsangee, H. Al-Musaibeli, R. Ahmad, "A Defect Detection Method Based on YOLOv7 for Automated Remanufacturing," *Applied Sciences*, vol. 14, 2024, 5503.
[4] L. Li, W. Xu, Y. Gao, et al., "Attention-oriented residual block for real-time low-light image enhancement in smart ports," *Computers and Electrical Engineering*, vol. 120, 2024, 109634.
[5] D. Peng, W. Ding, T. Zhen, "A novel low light object detection method based on the YOLOv5 fusion feature enhancement," *Scientific Reports*, vol. 14, no. 1, 2024, 4486.
[6] S. Wang, X. Xu, H. Chen, et al., "Low-light salient object detection meets the small size," *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2024.
[7] M. Skowron, O. Frankiewicz, J. J. Jarosz, et al., "Detection and classification of rolling bearing defects using direct signal processing with deep convolutional neural network," *Electronics*, vol. 13, no. 9, 2024, 1722.
[8] P. Maruschak, I. Konovalenko, Y. Osadtsa, V. Medvid, O. Shovkun, D. Baran, H. Kozbur, R. Mykhailyshyn, "Surface Illumination as a Factor Influencing the Efficacy of Defect Recognition on a Rolled Metal Surface Using a Deep Neural Network," *Applied Sciences*, vol. 14, 2024, 2591.
[9] Y. Peng, F. Xia, C. Zhang, et al., "Deformation feature extraction and double attention feature pyramid network for bearing surface defects detection," *IEEE Transactions on Industrial Informatics*, 2024.
[10] X. Zhao, L. Wang, Y. Zhang, et al., "A review of convolutional neural networks in computer vision," *Artificial Intelligence Review*, vol. 57, no. 4, 2024, 99.
[11] S. Cong, M. Gao, C. Zhang, Y. Yu, "YOLO-KD: Real time road defect detection using advanced YOLO models," *Engineering Letters*, vol. 33, no. 6, 2025, pp. 1839–1850.
[12] S. Wang, Y. Xu, "MI-YOLO: An improved traffic sign detection algorithm based on YOLOv8," *Engineering Letters*, vol. 32, no. 12, 2024, pp. 2336–2345.
[13] Y. Fu, Y. Zhang, "Lightweight rice leaf disease detection method based on improved YOLOv8," *Engineering Letters*, vol. 33, no. 2, 2025, pp. 402–417.
[14] Y. Yang, Z. Min, J. Zuo, B. Han, L. Li, "Crack identification of automobile steering knuckle fluorescent penetrant inspection based on deep convolutional generative adversarial networks data enhancement," *Frontiers of Physics*, 2022.
[15] S. Su, S. Du, X. Lu, "Geometric Constraint and Image Inpainting-Based Railway Track Fastener Sample Generation for Improving Defect Inspection," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 12, 2022, pp. 23883-23895.

[16] S. Teng, A. Liu, B. Chen, J. Wang, Z. Wu, J. Fu, "Unsupervised learning method for underwater concrete crack image enhancement and augmentation based on cross domain translation strategy," *Engineering Applications of Artificial Intelligence*, vol. 136, 2024, 108884.
[17] A. K. Malikov, M. F. Flores Cuenca, B. Kim, et al., "Ultrasonic tomography imaging enhancement approach based on deep convolutional neural networks," *Journal of Visualization*, vol. 26, 2023, pp. 1067–1083.
[18] Y. Hao, C. Zhang, X. Li, "Research on Defect Detection Method of Bearing Dust Cover Based on Machine Vision and Multi-feature Fusion Algorithm," *Measurement Science and Technology*, vol. 34, 2023.
[19] S. Yuan, N. Yan, L. Zhu, J. Hu, Z. Li, H. Liu, X. Zhang, "High dynamic online detection method for surface defects of small diameter reflective inner wall," *Measurement*, vol. 195, 2022, 111138.
[20] H. Cai, C. Cheng, R. Na, H. Zhang, J. Zhou, S. Jing, C. Miao, "Cooling-excited infrared thermography for enhancing the detection of concrete filled steel tube interfacial debonding at concrete hydration," *Case Studies in Construction Materials*, vol. 20, 2024, e02995.
[21] D. Liao, Z. Cui, X. Zhang, J. Li, W. Li, Z. Zhu, N. Wu, "Surface defect detection and classification of Si3N4 turbine blades based on convolutional neural network and YOLOv5," *Advances in Mechanical Engineering*, vol. 14, 2022, 168781322210815.
[22] E. Bian, M. Yin, S. Fu, Q. Gao, Y. Li, "Part Defect Detection Method Based on Channel-Aware Aggregation and Re-Parameterization Asymptotic Module," *Electronics*, vol. 13, 2024, 473.
[23] X. Li, Q. Wang, X. Yang, K. Wang, H. Zhang, "Track Fastener Defect Detection Model Based on Improved YOLOv5s," *Sensors*, vol. 23, 2023, 6457.
[24] B. Kim, Y. Jeon, J. W. Kang, et al., "Multi-task Transfer Learning Facilitated by Segmentation and Denoising for Anomaly Detection of Rail Fasteners," *Journal of Electrical Engineering and Technology*, vol. 18, 2023, pp. 2383–2394.
[25] S. Jiang, J. Zhang, W. Wang, Y. Wang, "Automatic Inspection of Bridge Bolts Using Unmanned Aerial Vision and Adaptive Scale Unification-Based Deep Learning," *Remote Sensing*, vol. 15, 2023, 328.
[26] S. E. Weng, S. G. Miaou, R. Christanto, "A lightweight low-light image enhancement network via channel prior and gamma correction," *arXiv preprint arXiv:2402.18147*, 2024.
[27] Y. Zhang, J. Zhang, X. Guo, "Kindling the darkness: A practical low-light image enhancer," in *Proceedings of the 27th ACM International Conference on Multimedia*, 2019.
[28] Y. Cai, et al., "Retinexformer: One-stage retinex-based transformer for low-light image enhancement," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023.
[29] C. Guo, et al., "Zero-reference deep curve estimation for low-light image enhancement," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020.
[30] F. Yan, et al., "Global Context Aggregation Network for Lightweight Saliency Detection of Surface Defects," *arXiv preprint arXiv:2309.12641*, 2023.