Frontiers in Computing and Intelligent Systems
ISSN: 2832-6024 | Vol. 16, No. 2, 2026

Deep Learning Model Fused with Attention Mechanism for Defect Detection of Electroplated NdFeB Products

Jingbo Hao, Junfeng Li, Yang Du, Qian Chen *, Zhen Liang

Equipment Engineering Department, Advanced Technology & Materials Co., Ltd., Beijing, China

*** Corresponding author:** Qian Chen (Email: sanchen97@163.com)

Abstract: Electroplated neodymium-iron-boron (NdFeB) products are widely used in electronics, automobiles, and new energy industries. Surface defects, such as cracks, pinholes, scratches, and edge chipping, can directly affect product performance and service life. To address the problems of insufficient detection accuracy, weak anti-interference capability, and difficulty in identifying tiny defects in traditional deep learning models for NdFeB surface defect detection, an improved defect detection model integrating an attention mechanism is proposed. Based on YOLOv8n, the proposed model introduces an improved Recurrent Channel Attention (RCA) mechanism into the backbone network to enhance the extraction of critical defect features and suppress complex background interference. Meanwhile, Switchable Atrous Convolution (SAConv) is incorporated to enlarge the receptive field and improve the model's perception and representation ability for multi-scale defects. Experimental results show that the proposed model achieves an mAP@0.5 of 85.3% on a self-built electroplated NdFeB defect dataset, representing an improvement of 6.7% over the original YOLOv8n model. The detection speed reaches 62 FPS, meeting the real-time detection requirements of industrial production lines. The proposed model can effectively identify various common surface defects in electroplated NdFeB products and demonstrates strong detection capability for tiny and subtle defects under complex lighting conditions, thereby providing reliable technical support for intelligent quality inspection of electroplated NdFeB products.

Keywords: Electroplated NdFeB; Surface Defect Detection; YOLOv8n; Attention Mechanism; Switchable Atrous Convolution.

1. Introduction

Although deep learning methods have demonstrated strong feature representation capability and detection performance in industrial surface defect detection, they still face numerous challenges for the specific task of surface defect detection of electroplated NdFeB products. First, the surface of electroplated NdFeB products exhibits strong metallic reflection characteristics, and high-light regions and shadow regions tend to coexist under different lighting conditions. This results in indistinct grayscale and texture differences between defects and the background, increasing the difficulty of model discrimination. Second, there are various types of surface defects on electroplated NdFeB, which differ greatly in shape, size, distribution location and edge features. Some defects also suffer from mutual occlusion and blurred boundaries, which imposes higher requirements on the multi-scale feature extraction ability and robustness of the model. Third, tiny defects such as pinholes, fine cracks and local edge chippings account for a small proportion of pixels in images and are often accompanied by noise interference. Traditional detection models are prone to missed detection and false detection, making it difficult to meet the requirements of high precision and high stability for industrial online inspection [1,2].

According to existing research, object detection algorithms have gradually become an important technical route for industrial defect detection. Among them, two-stage object detection algorithms represented by Faster R-CNN have high positioning accuracy and classification performance. However, due to their complex model structure and slow inference speed, their application in industrial production lines with high real-time requirements is limited to a certain extent. In contrast, one-stage object detection algorithms represented by the YOLO series achieve a good balance between detection speed and accuracy, and are especially suitable for real-time online inspection tasks. As one of the latest representatives of the YOLO series, YOLOv8 shows strong advantages in network structure design, feature fusion capability and inference efficiency. Among them, the YOLOv8n model has fewer parameters and low computational cost, making it more suitable for deployment on industrial edge devices or embedded platforms. Nevertheless, while lightweight models achieve high-speed detection performance, they may also suffer from insufficient feature extraction capability. Especially when facing complex backgrounds and tiny targets, the original model tends to ignore fine-grained defect information, thereby affecting the overall detection effect [3-6].

To further improve the detection performance of models in complex industrial scenarios, researchers have attempted to introduce various attention mechanisms into object detection networks to enhance feature representation by strengthening the network's attention to key regions and important channel information. For example, attention mechanisms such as SE, CBAM and ECA have achieved certain effects in defect detection, object recognition and other tasks. These methods can improve the model's perception of target regions to a certain extent, but still have problems such as insufficient modeling of local details and limited adaptability to complex scale variations in the task of electroplated NdFeB defect detection [7,8]. In addition, conventional convolution operations alone can hardly balance the capture of local details and the modeling of global context information. Especially for defect categories with large size differences, insufficient feature response is prone to occur. Therefore, how to improve the detection ability of multi-scale and tiny defects under complex backgrounds on the premise of ensuring model lightweight and real-time performance has become a key issue in the current research on intelligent defect detection of electroplated NdFeB.

To address the above problems, this paper takes YOLOv8n as the baseline model and proposes a surface defect detection method for electroplated NdFeB that fuses attention mechanism and switchable atrous convolution. An improved Recurrent Channel Attention (RCA) mechanism is introduced into the backbone network to model the interdependencies between feature channels, strengthen the model's representation of key defect features such as cracks, pinholes, scratches and edge chippings, and reduce the interference of complex backgrounds and metallic reflections on detection results. Meanwhile, to enhance the model's perception of defects at different scales, a Switchable Atrous Convolution (SAConv) module is introduced. By dynamically adjusting the convolutional receptive field, the model's ability to handle both large-scale and tiny defects is improved, thereby further enhancing detection accuracy and robustness. While maintaining high detection speed, the proposed method can more effectively extract discriminative features of defect regions, providing technical support for real-time online inspection of industrial production lines.

Accordingly, this paper proposes a deep learning model for NdFeB defect detection fused with an attention mechanism. Based on YOLOv8n, the model optimizes its feature extraction capability and multi-scale defect perception by introducing an improved attention mechanism and switchable atrous convolution, so as to achieve high-precision and real-time detection of defects in electroplated NdFeB products. The main research contents of this paper are as follows:

(1) Construct an electroplated NdFeB defect dataset and annotate common defect types to provide data support for model training and validation;

(2) Design a backbone network fused with the improved RCA attention mechanism to enhance the model's ability to capture defect features;

(3) Introduce SAConv to expand the receptive field of the model and improve the performance of multi-scale defect detection;

(4) Verify the effectiveness and superiority of the proposed model through comparative experiments, providing theoretical and experimental basis for industrial practical applications.

2. Model Structure

2.1. Overall Process

The proposed NdFeB defect detection model fused with attention mechanism takes YOLOv8n as the baseline model. Aiming at its shortcomings in NdFeB defect detection, two improvements are made.First, the improved RCA attention mechanism is integrated into the backbone network to enhance the extraction ability of defect features.Second, SAConv is introduced to replace some ordinary convolutions to expand the receptive field of the model and improve the performance of multi-scale defect detection.

The overall structure of the model mainly consists of four parts: input layer, backbone feature extraction network, neck feature fusion network, and head detection network. The specific structure is shown in Figure 1 (the schematic diagram of the model structure can be supplemented here, and the corresponding chart can be inserted in practical applications).

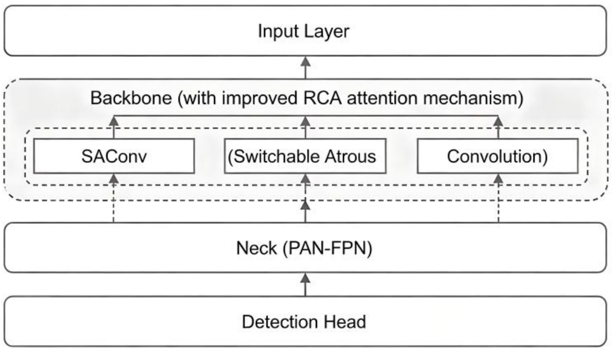

Figure 1. Overall Structure Diagram

The model follows the classic pipeline of object detection with a clear and logically rigorous data flow from top to bottom:The Input Layer first receives image data and performs standardized preprocessing such as normalization and resizing to lay a foundation for subsequent feature extraction. As the core feature extraction module, the Backbone is equipped with an improved RCA attention mechanism and integrates three core operators: SAConv, Switchable Atrous Convolution, and standard convolution. It not only enhances the aggregation capability of global contextual information but also dynamically adapts to the feature extraction requirements of objects at different scales. The Neck adopts the PAN-FPN structure, which effectively bridges the semantic gap between shallow and deep features through top-down and bottom-up bidirectional feature fusion paths, achieving multi-scale feature enhancement and fusion and providing highly robust feature representations for detection tasks. Finally, the Detection Head performs the final prediction of object category classification and bounding box localization based on the fused multi-scale features, outputting detection results containing category probabilities and coordinate information.

2.2. SAConv

Figure 2 illustrates the architecture of the core Switchable Atrous Convolution (SAC) module in the DetectoRS framework. This module adopts a dual-branch weight-sharing design, where multi-scale features are extracted by two 3×3 atrous convolutions with atrous rates of 1 and 3, respectively. A spatially adaptive switching function S generates position-wise fusion weights to achieve a dynamic balance between local details and global context.

Global context injection mechanisms are embedded at both ends of the module: first, image-level semantic information is projected via global average pooling and a 1×1 convolution, then fused with the original features before being fed into the core convolution branch. After multi-scale feature fusion, a symmetric post-global context module reinforces feature consistency and feeds the enhanced features back into the main stream in a residual form.This design equips the network with the ability to adapt to varying object scales without significantly increasing the number of parameters, effectively alleviating the limitation of fixed receptive fields in traditional atrous convolutions.

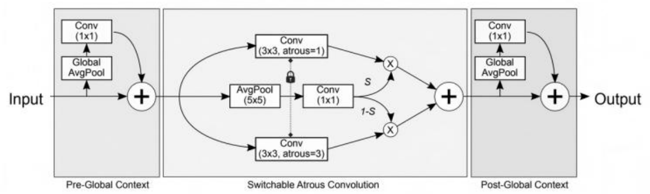

Figure 2. Switchable Atrous Convolution

The defect sizes of NdFeB products vary greatly, including tiny pinholes and fine cracks (size less than 1mm), as well as larger edge chippings and scratches (size greater than 5mm). The receptive field of traditional convolution is fixed, making it difficult to simultaneously meet the detection needs of defects of different sizes. Dilated convolution can expand the receptive field without increasing the number of parameters by inserting dilations into the convolution kernel, thereby improving the model's ability to detect large-size defects. Its core calculation process can be expressed as Equation (1), where the input feature map is $X \in \mathbb{R}^{C_{in} \times H \times W}$, the convolution kernel is $K \in \mathbb{R}^{C_{out} \times C_{in} \times k \times k}$, r is the dilation rate, and each element of the output feature map Y after dilated convolution is calculated as follows:

$$Y(i,j,k)=\sum_{m=1}^{C_{in}}\sum_{n=1}^{k}\sum_{p=1}^{k}X(i+r\cdot n,j+r\cdot p,m)\cdot K(k,m,n,p) \qquad (1)$$

However, dilated convolution with a fixed dilation rate tends to cause sparsity in feature maps, which affects the detection accuracy of tiny defects. Switchable Atrous Convolution (SAConv) can adaptively adjust the dilation rate according to the content of the input feature map: it adopts a smaller dilation rate for defect regions to ensure that the feature details of tiny defects are not lost, and a larger dilation rate for background regions to expand the receptive field and capture more background context information, thereby achieving accurate detection of multi-scale defects.

The core advantage of SAConv lies in its adaptive dilation rate generation mechanism. First, a lightweight generation network $g(\cdot)$ is used to predict the initial dilation rate, and then the Sigmoid function is used to map it to the preset dilation rate range $[r_{min}, r_{max}]$, which is specifically shown in Equation (2). Here, r_{min} and r_{max} are the lower and upper bounds of the dilation rate, respectively, and $\sigma(.)$ is the Sigmoid activation function:

$$R(x)=r_{min}+(r_{max}-r_{min})\cdot\sigma(g(X)) \qquad (2)$$

2.3. Design of the Improved RCA Attention Mechanism

Typical channel attention modules represented by SE-Net only rely on global average pooling to generate channel-wise weight descriptors. Such a single compression method tends to discard detailed local feature responses and shows poor robustness against noise, making it less effective when dealing with complex industrial backgrounds and tiny defects in NdFeB surface detection. In view of these limitations, this paper presents an enhanced version of the Recurrent Channel Attention (RCA) mechanism, aiming to strengthen the model’s capability in capturing fine-grained defect features and suppressing irrelevant background interference.

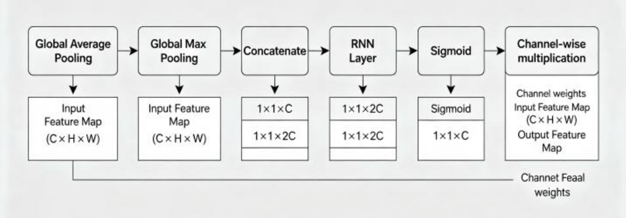

Figure 3. Switchable Atrous Convolution

The improved RCA structure can be divided into three key stages: feature compression, recurrent attention learning, and feature recalibration. First, both global average pooling and global max pooling are applied to the feature maps from the backbone network. By fusing the outputs of these two pooling operations, the two-dimensional spatial features are compressed into one-dimensional channel descriptors, preserving both global statistical information and locally salient details simultaneously. Next, a recurrent neural network is employed to perform iterative learning on the compressed channel features, which helps model long-range semantic dependencies among different channels. In this way, the network adaptively emphasizes channels related to defect regions while weakening those dominated by background clutter. Finally, the learned channel weights are normalized through the Sigmoid function and multiplied channel-wisely with the original feature maps to obtain reweighted and enhanced feature representations.

This improved RCA module is integrated into the backbone of YOLOv8n by inserting it after each C2f block. Such a design allows the model to dynamically focus on defect areas during feature extraction, improves the distinguishability of defect characteristics, and provides more representative and discriminative features for the subsequent feature fusion and object detection stages.

2.4. Moving Average

The overall workflow of the model is described as follows.First, the preprocessed images of electroplated NdFeB products are fed into the model, with their input size uniformly resized to 640×640.Next, multi-scale features of the input images are extracted via the backbone network, which integrates the improved RCA attention mechanism and SAConv, thereby generating feature maps at different hierarchical levels.After that, the PAN-FPN neck network is adopted to fuse these multi-level feature maps, so as to strengthen the representative capability of the features.Finally, the detection head performs defect classification and location regression on the fused feature maps, and outputs the category, coordinates and confidence score of each defect, by which the whole defect detection task is completed.

3. Results and Analysis

In this study, the experimental hardware platform consists of an Intel Core i7-12700H processor, an NVIDIA RTX 4090 graphics card, and 16 GB of RAM. The software environment is built on the Windows 10 operating system, using PyTorch 1.13.0 as the deep learning framework and Python 3.9 as the programming language, with LabelImg employed for dataset annotation. At present, publicly available defect datasets for electroplated neodymium-iron-boron are scarce and cannot meet the requirements of practical industrial detection scenarios. Accordingly, a dedicated defect dataset is constructed in this paper, collected from the electroplating production line of a neodymium-iron-boron manufacturing enterprise. The dataset contains 500 images with a resolution of 1280 × 960, covering four typical post-electroplating defects: cracks, scratches, color differences, and edge chipping. The dataset is randomly divided into training, validation, and test sets at a ratio of 7:2:1. To further enhance the generalization ability and anti-interference performance of the model, data augmentation operations including random flipping, random cropping, brightness adjustment, contrast adjustment, and Gaussian noise injection are applied to the training set to expand the sample size.

To fully verify the effectiveness and superiority of the proposed model in the task of defect detection for electroplated neodymium-iron-boron, comparative experiments were carried out to compare it with three different models. Specifically, the comparison models include the original YOLOv8n baseline model, the YOLOv8n improved model integrated with the SE attention mechanism (denoted as YOLOv8n+SE), and the YOLOv8n improved model embedded with the CA attention mechanism (denoted as YOLOv8n+CA). To ensure the fairness and rigor of the comparative experiments, all tests were conducted under the same experimental hardware and software environment, and the same set of evaluation metrics (mAP@0.5, Precision, Recall, FPS) was adopted for comprehensive comparison. The detailed experimental data and specific comparison results of each model are presented in Table 1.

Table 1. Comparison of Experimental Results of Different YOLOv8n-Based Models

Model	Precision	Recall	mAP@0.5	FPS
YOLOv8n	**78.6**	**79.2**	**78.6**	**68**
+SE	**81.3**	**82.5**	**81.8**	**65**
+CA	**82.1**	**83.7**	**82.9**	**64**
Our	**84.5**	**85.8**	**85.3**	**62**

As can be seen from Table 1, the proposed model in this paper outperforms all the comparison models in all evaluation metrics. Compared with the original YOLOv8n baseline model, the proposed model improves precision by 5.9%, recall by 6.6%, and mAP@0.5 by 6.7%, respectively. Although the detection speed is reduced by 6 FPS, it still maintains 62 FPS, which can fully meet the real-time detection requirements in industrial scenarios. Compared with the YOLOv8n+SE and YOLOv8n+CA models, the mAP@0.5 of our model is increased by 3.5% and 2.4%, respectively. This indicates that the improved RCA attention mechanism is more effective than the traditional SE and CA attention mechanisms in improving the defect detection accuracy. It can better capture the subtle defect features of neodymium-iron-boron and suppress the interference of complex background in industrial scenes.

To further validate the effectiveness of the proposed model in detecting different categories of defects, the average precision (AP) values corresponding to each defect type are statistically analyzed for all comparison models. The detailed results are listed in Table 2.

Table 2. AP Values of Different Models for Various Defect Types

Model	Crack	Scratch	Color Difference	Edge Chipping
YOLOv8n	**76.2**	**79.5**	**77.8**	**75.4**
YOLOv8n+SE	**79.8**	**82.1**	**80.9**	**78.7**
YOLOv8n+CA	**81.5**	**83.6**	**82.3**	**81.2**
Our	**84.1**	**85.3**	**84.7**	**87.1**

As can be observed from Table 2, the proposed model achieves the highest AP values across all four defect categories, exhibiting superior detection performance compared with the baseline YOLOv8n and its attention-enhanced variants.In comparison with the original YOLOv8n, our model obtains obvious improvements in AP for cracks, scratches, color differences and edge chipping, respectively. Especially for tiny and inconspicuous defects such as cracks, the proposed model shows a more significant performance gain, which demonstrates its strong ability in locating subtle defect features. When compared with YOLOv8n+SE and YOLOv8n+CA, our model still achieves higher AP values on all defect types. This reveals that the improved attention mechanism adopted in this work can more effectively extract discriminative feature information of different defects, suppress background interference in industrial scenes, and enhance the model's adaptability to various defect morphologies.

Overall, the proposed model maintains stable and outstanding detection performance for multiple types of electroplated NdFeB defects, verifying its effectiveness and robustness in practical defect detection tasks.

4. Summary

To address insufficient public datasets, difficult subtle defect detection and obvious background interference in electroplated NdFeB defect detection, this paper proposes an improved YOLOv8n model. A dedicated dataset is constructed with data augmentation to enhance generalization. Comparative experiments with original YOLOv8n, YOLOv8n+SE and YOLOv8n+CA (using precision, recall, mAP@0.5 and FPS as indicators) show the proposed model achieves 84.5% precision, 85.8% recall, 85.3% mAP@0.5 and 62 FPS—outperforming all comparison models, with mAP@0.5 6.7% higher than original YOLOv8n and satisfying real-time industrial needs. It also gains the highest AP for cracks, scratches, color differences and edge chipping, especially excelling in tiny/weak defect identification. The model effectively extracts defect features, suppresses background interference, balances accuracy and speed, providing a reliable technical scheme for practical scenarios.

References

[1] LI Junfeng, CHEN Qian, HAO Jingbo, et al. Surface defect detection of NdFeB based on FFA-YOLO[J]. Metallic Functional Materials, 2025, 32(3): 91-96 https://doi.org/ 10. 13228/j.boyuan.issn1005-8192.20240164.

[2] Geng J ,Li M ,Tang C .RFG-YOLO: Accurate and Lightweight Detection of Multi-Scale Defects on Metallic Surfaces[J]. Concurrency and Computation: Practice and Experience, 2026, 38(2).

[3] Lu J ,Yu R ,Zhang H , et al.Highly reflective metal industrial parts surface detection with polarization vision and modified YOLO[J].Digital Signal Processing,2025,163105238-105238. DOI: 10.1016/J.DSP.2025.105238.

[4] Yiming X ,Kai Z ,Li W .Metal Surface Defect Detection Using Modified YOLO[J].Algorithms,2021,14(9):257-257.DOI: 10. 3390/ A14090257.

[5] Jiang K ,Wang Y ,He H , et al.Application of YOLO-v7 and YOLO-v8 Transfer Learning Models in Breast Lesion Classification and Diagnosis.[J].Current medical imaging, 2026, DOI:10.2174/0115734056422406251128144702.

[6] SHAN Bin, XIAN Ziqi, WEN Yanwei, et al. Applications of machine learning in atomic manufacturing: status, challenges, and outlook[J]. Metallic Functional Materials, 2025, 32(4): 1-14 https://doi.org/10.13228/j.boyuan.issn1005-8192.20250130.

[7] ZHANG Youliang, ZHANG Xin, WANG Qiyan, et al. Electrolytic ash removal process for NdFeB permanent magnetic materials[J]. Metallic Functional Materials, 2025, 32(4): 130-134 https://doi.org/10.13228/j.boyuan.issn1005-8192.20240215.

[8] Zhang H, Li Y, Wang Z, et al. Surface Defect Recognition of NdFeB Raw Materials Based on Improved YOLOX-Tiny Model[C]//2023 3rd International Conference on Electronic Technology and Information Science (ICETIS). 2023: 1024-1028.