

Small Target Defect Detection Method on Aluminum Ingot Surface based on Improved YOLOv8n-SimAM

Guangxu Liu¹, Batu Nasheng², Wei Zheng², Liangliang Lv², Tianhua Zhang², Guodong Sun²

¹ College of Control Science and Engineering, Zhejiang University, Hangzhou, Zhejiang, 310000, China

² Inner Mongolia Huomei-Hongjun Aluminum and Electricity Co., Ltd., Hologola, Inner Mongolia, 029200, China

Abstract: In this paper, an intelligent detection solution based on the improved YOLOv8n architecture is proposed to address the key challenges in the detection of surface defects in aluminum ingots. As an important raw material in modern industry, the surface defects of aluminum ingots increase significantly in high-speed production environments. Traditional detection methods have difficulty dealing with the problems of small targets and morphological diversity. This study effectively improves the model's ability to recognize multi-morphological small defects by introducing the parameter-free SimAM three-dimensional attention mechanism. The innovations include: 1) constructing a high-definition aluminum ingot defect dataset based on an actual production line and systematically analyzing the scale distribution characteristics of the defects; 2) integrating the three-dimensional attention mechanism in the network backbone layer to enhance the feature extraction capability of small targets and solve the morphological diversity problem of burrs and slag inclusions; 3) through comparative experimental verification, the improved YOLOv8n-SimAM model achieved a mAP of 64.5% in aluminum ingot surface defect detection, an increase of 9 percentage points over the baseline model, and improved the detection of burrs and slag inclusions by 6.3% and 9.9% in the F1 score, respectively. Experimental results show that this method achieves a good balance between detection accuracy and computational efficiency, providing a reliable technical solution for the practical application of intelligent manufacturing quality monitoring systems for aluminum ingots.

Keywords: Surface Defects of Aluminum Ingots; YOLOv8n; SimAM Attention; Small Target Detection.

1. Introduction

As a key basic raw material for modern industry, the quality and stability of aluminum alloy ingots directly impacts the performance of core components in high-end equipment such as aircraft and electric vehicles. Amid the rapid development of intelligent manufacturing technology, the operating speed of continuous casting production lines for aluminum ingots has now reached over 100 meters per minute. However, the high-speed production environment and complex process conditions have led to a significant increase in the incidence of surface defects such as burrs on the edges of ingots and oxide inclusions. These process defects not only result in a decrease in product qualification rates, but are also likely to cause secondary damage during subsequent deep processing. Therefore, establishing an efficient aluminum ingot surface quality inspection system is of great significance for ensuring quality control in the industrial chain.

Existing research on aluminum alloy defect detection primarily focuses on two scenarios: geometric dimension inspection of finished workpieces[1], and composition analysis of alloy materials[3]. However, at the initial stage of the casting process, the detection of surface defects in aluminum ingots still faces the dual challenges of basic theory and technical application. Current detection technologies face the following core challenges: First, surface defects in aluminum ingots appear as distinct, tiny target features in industrial camera imaging (the pixel ratio of the defect area is typically less than 5%), and they often appear densely clustered. Second, defect morphology is highly diverse, with significant visual differences between different types of defects, and even within the same defect, there is significant variation in shape, texture, and spatial distribution.

To address the above problems, this study developed an

intelligent detection solution for aluminum ingot surface defects based on an improved YOLOv8n architecture. By introducing the parameter-free SimAM attention mechanism, the model's ability to recognize tiny multi-morphological defects was effectively improved. The innovative work of this study is mainly reflected in the following aspects: in response to the scarcity of aluminum ingot defect data, a high-definition image dataset based on actual production lines was constructed, and the scale distribution characteristics of defects were systematically analyzed; a three-dimensional attention mechanism was integrated into the network backbone layer to enhance the feature extraction capability of tiny targets and successfully solve the morphological diversity problem of burrs and slag inclusions; through systematic comparative experiments, it was proved that the proposed method achieved a good balance between detection accuracy and computational efficiency, providing a reliable technical solution for the practical application of intelligent manufacturing quality monitoring systems for aluminum ingots.

2. Related Work

The development of industrial surface defect inspection technology has consistently kept pace with the evolution of inspection targets and the advancement of automation. Throughout the historical evolution of inspection methods, we observe three distinct phases: the early reliance on manual visual inspection and the current intelligent phase based on machine vision. This study will focus on the application of machine vision technology in surface defect inspection in complex industrial scenarios.

In recent years, deep learning[4] as an efficient machine learning method, has achieved remarkable results in target detection tasks. Ma et al.[5] proposed a lightweight aluminum

strip surface defect detection method YOLO-DCSAM based on the improved YOLOv4. By optimizing the network structure through deep separable convolution and dual-channel attention modules, the detection speed was increased by 3 times while maintaining high detection accuracy, significantly improving the applicability of real-time detection in industrial sites. Zhu et al.[6] proposed an improved YOLOv5s algorithm, which optimized the same-scale feature fusion capability by introducing cross-convolution feature enhanced connection, enhanced feature concat module and channel attention mechanism, and achieved a 3.8% improvement in the accuracy of infrared laser sensor in the detection of small defects on metal-based surfaces. Zhang et al.[7] proposed a wind blade defect detection model SOD-YOLO based on an improved YOLOv5. By adding a microscale detection layer, optimizing anchor box clustering, introducing the CBAM attention mechanism and channel pruning, they achieved a 95.1% mAP (a 7.82% increase over the original model) and a 28.3% FPS improvement in drone image analysis. The KPE-YOLOv5 algorithm proposed by Yang et al[8]. achieved a 5.3% mAP improvement and a 7% accuracy improvement on the VisDrone-2020 dataset through three improvements: K-means++ optimization of anchor boxes, integration of scSE attention modules, and addition of small target detection layers, significantly enhancing the small target detection capability. In the field of aluminum workpiece detection, Nascimento R et al.[9] proposed an automatic detection system for aluminum casting quality based on the YOLOv8s-seg instance segmentation model and traditional image processing technology, which realizes flash detection and hole detection through a robot system. In response to the challenges of multi-scale changes, category imbalance and high background similarity in the detection of aluminum ingot surface defects, the YOLO-PDC model proposed by Li[10] optimizes feature extraction by combining PConv and DCNv2 and introduces the SimAM attention mechanism to enhance detection accuracy. In the field of aluminum alloy ingot detection, Han et al.[11] developed a real-time detection

algorithm for aluminum ingot alloy casting lines and verified the applicability of single-stage detectors in aluminum material inspection. Yan et al.[12] used a mask gradient response threshold segmentation method to solve the problem of detecting surface defects of milled aluminum ingots. Existing research focuses on aluminum products after processing, while research on the detection of native defects on the surface of aluminum ingots at the source of casting is relatively scarce.

3. Method

3.1. Improved Model Based on YOLOv8

YOLOv8 is an advanced object detection algorithm that inherits the strengths of its predecessors, the YOLO family of algorithms, while also incorporating several significant improvements and optimizations. YOLOv8 optimizes adaptive anchor box adjustment, enabling more accurate prediction of object positions and sizes, which helps improve detection precision and recall. It also introduces improved multi-scale prediction technology for better detection of objects of varying sizes. The YOLOv8 model consists of four components: Input, Backbone, Neck, and Head.

Input: This is the input of the YOLOv8 model. It is responsible for preprocessing the input image and converting it into a format that the model can handle.

Backbone: This is the backbone network of YOLOv8, responsible for extracting features from the input image and converting the image into a feature representation rich in semantic information. YOLOv8 uses CSPDarknet53 as its backbone network, which offers significant improvements in performance and speed compared to the previous Darknet53.

Neck: YOLOv8 utilizes the concept of the Feature Pyramid Network (FPN), removing the convolutions in the FPN upsampling stage from the original YOLOv5 and replacing the C3 module with a C2f module. This structure facilitates the integration of features from the Backbone architecture, improving model performance.

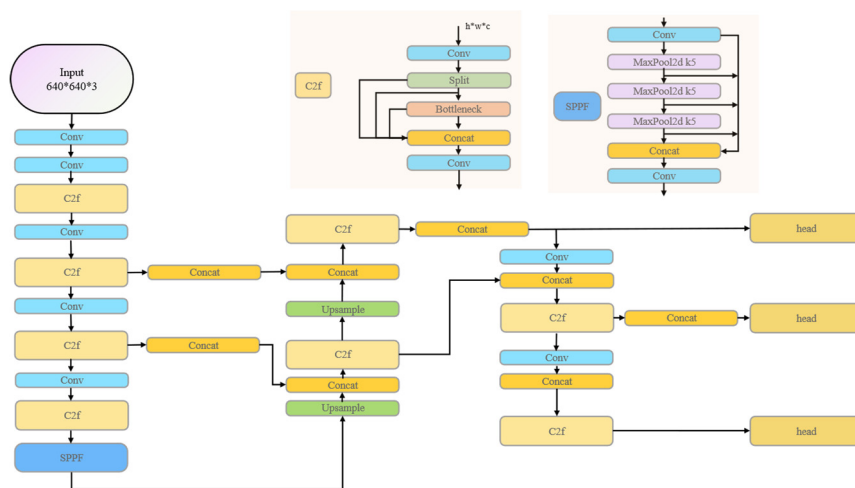


Figure 1. Improved YOLOv8 network architecture diagram

Head: This part is the last layer of the model. Unlike the coupled head of YOLOv5, YOLOv8 uses Decoupled Head. In the detection head part, YOLOv8 adopts the Anchor-free idea[13], which accelerates the post-processing process of non-maximum suppression (NMS) and further improves the accuracy and efficiency of the model.

In addition, to meet the needs of different application scenarios, YOLOv8 provides five models of different scales (N, S, M, L, X) to adapt to various resource conditions[14]. After comparing the detection accuracy and efficiency of YOLOv8 models of different sizes, it was found that YOLOv8n achieved the best balance between efficiency and

accuracy, and subsequent improvements were based on this model. The improved aluminum ingot surface defect detection network structure is shown in Figure 1. The SimAM attention module is introduced at the backbone of the network to guide the model to focus on the detailed structures at different scales on the aluminum ingot surface.

Channel attention uses a 1D mechanism to distinguish channels but ignores spatial differences, while spatial attention focuses on position information through a 2D mechanism but ignores channel differences, which restricts the ability to distinguish features. Existing methods such as BAM and CBAM[15] simply connect the two types of attention in parallel or in series, failing to simulate the collaborative mechanism of human brain attention. Therefore, it is necessary to construct an attention module with unified weights to achieve a better attention mechanism by evaluating the importance of neurons.

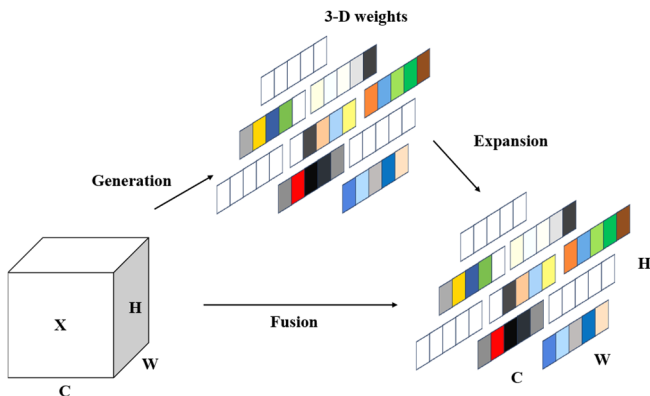


Figure 2. ECA module structure

Based on mature neuroscience theories, Yang et al.[16] proposed the SimAM attention mechanism in 2021. This is a fully 3-D, weighted, and parameter-free attention mechanism. Its principle is shown in Figure 2, where H and W are the height and width of the feature map, and X is the input feature map. Compared with other existing attention mechanisms, SimAM considers the correlation between spatial and channel factors and can efficiently generate realistic 3-D weights for feature mapping without the need for additional parameters.

4. Experiment

4.1. Dataset And Experimental Sett

To meet the industrial needs of aluminum ingot surface defect detection, this study used data actually collected from an aluminum production line in Inner Mongolia, using an MV-CS016-10GM industrial camera (1440×1080 resolution, 10 μ m pixel size) to acquire more than 2,000 images. After process analysis and quality assessment, flash and slag inclusions were identified as the main defect types: flash is a thin linear protrusion (5-20 μ m wide) formed by the overflow and solidification of rolled aluminum liquid; slag inclusions are irregular pits (50-200 μ m) caused by residual Al_2O_3 impurities. Through stratified sampling, 1,889 high-quality images without blur and uniform distribution were selected as samples. The results are shown in Figure 3.

To address the challenges of on-site data collection in aluminum ingot production (high temperature limitations, mechanical vibration, high reflective interference, and insufficient defect morphology diversity), this study employed data augmentation techniques to expand the sample size. Specific methods include: 1) Geometric transformation:

$\pm 15^\circ$ random rotation to simulate pose changes, 0.7-1.3 $^\circ$ scaling to adapt to shooting distance, and shear correction to compensate for installation errors; 2) Photometric adjustment: High-Speed Vision (HSV) spatial brightness perturbation to simulate illumination fluctuations; and 3) Semantic enhancement: Multi-image stitching to promote multi-scale feature learning, and image blending to enhance inter-class boundary recognition.

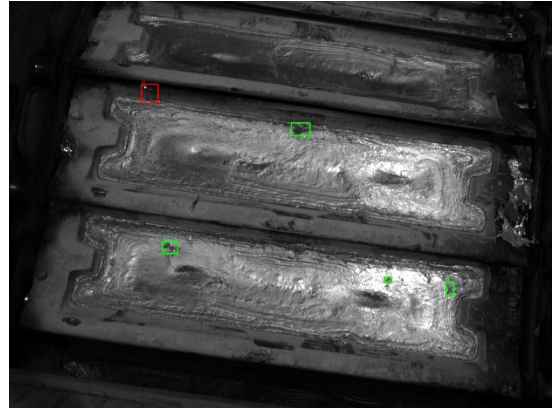


Figure 3. Schematic diagram of aluminum ingot surface defects

This study constructed a high-performance computing experimental platform with the following hardware and software configuration: It uses the Ubuntu 22.04 LTS operating system, is equipped with a parallel computing array consisting of 16 x 32GB DDR4-3200MHz ECC REG memory modules, and four NVIDIA GeForce RTX 4090 GPUs. The software environment is based on the PyTorch 1.13.1 deep learning framework, and is equipped with a 1TB Samsung 980 Pro NVMe SSD as system storage. The platform provides 16.3 TFLOPS of single-precision floating-point computing power, fully meeting the computational requirements of large-scale image processing.

4.2. Algorithm Comparison Experiment

In order to systematically evaluate the optimization effect of each improved module, this study designed a rigorous ablation experiment based on a self-built aluminum ingot surface defect dataset.

As shown in Table 1, the YOLOv8n-SimAM combined model demonstrates significant advantages in both defect detection categories: the overall F1-score reaches 0.731, with flash detection F1-score improving by 6.3% and slag inclusion detection by 9.9%. This demonstrates the synergistic effect of SimAM's 3D attention mechanism.

Table 1. Comparison of F1 values of various models in ablation experiments

Model	Fin	Slag inclusion	Overall category
YOLOv8n	0.712	0.686	0.699
YOLOv8n-SimAM	0.731	0.762	0.746

Table 2 shows the detection performance results for the AP and mAP values for each defect. The YOLOv8n-SimAM model achieved a mAP of 64.5% for aluminum ingot surface defect detection, a 9 percentage point improvement over the baseline YOLOv8n. It also achieved an 8.1% improvement in AP for flash detection and a 10% improvement in AP for slag

inclusion detection. The model incorporating SimAM alone achieved the best performance for slag inclusion detection.

Table 2. Comparison of AP and mAP values among different models in ablation experiments

Model	AP (%)		mAP@
	Fin	Slag inclusion	
YOLOv8n	53.8	57.1	55.5
YOLOv8n- SimAM	61.9	67.1	64.5

To investigate the performance differences between the improved YOLOv8n model and other models for aluminum ingot surface defect detection, we conducted performance comparison tests using the YOLOv5, YOLOv8n, YOLOv8m, YOLOv8x, and YOLOv10n detection models. Table 3 shows the performance results for both AP and mAP. The experimental results show that the YOLOv8n-SimAM model achieves optimal performance in both flash and slag inclusion detection tasks, achieving APs of 61.9% and 67.1%, respectively, and mAP@0.5 of 64.5%, significantly outperforming the other comparison models. In contrast, the basic version of YOLOv8n achieves an mAP@0.5 of 55.5%, while the larger-scale YOLOv8m and YOLOv8x models show decreased performance, reaching 45.1% and 44.7%, respectively, indicating that model performance does not improve with an increase in the number of parameters. In addition, the performance of YOLOv10n (50.2%) and YOLOv5 (47.0%) is lower than that of YOLOv8n, which further verifies the superiority of YOLOv8n-SimAM in small object detection tasks.

Table 3. Comparative experiment on AP and mAP values of different models

Model	AP(%)		mAP@ 0.5(%)
	Fin	Slag inclusion	
YOLOv8n-SimAM	61.9	67.1	64.5
YOLOv8n	53.8	57.1	55.5
YOLOv8m	42.5	47.7	45.1
YOLOv8x	42.7	46.7	44.7
YOLOv10n	43.4	57.0	50.2
YOLOv5	44.2	49.8	47.0

The F1 indicators of each model are shown in Table 4. The experimental results indicate that the YOLOv8n SimAM model achieves the best performance in both edge and slag detection tasks, with AP values of 0.768 and 0.759, respectively, and a comprehensive mAP of 0.763, significantly better than other compared models. In contrast, the mAP of the basic version YOLOv8n is 0.699, while the performance of larger YOLOv8m and YOLOv8x decreases, only 0.602 and 0.566, respectively, indicating that the model performance does not improve with an increase in the number of parameters. In addition, although YOLOv10n (0.628) and YOLOv5 (0.639) perform better than YOLOv8m and YOLOv8x, they are still inferior to YOLOv8n and its SimAM improved version, further verifying the effectiveness of attention mechanism in improving the accuracy of small object detection.

Table 4. Comparison of F1 values of various models in ablation experiments

Model	Fin	Slag inclusion	Overall category
YOLOv8n- SimAM	0.731	0.762	0.746
YOLOv8n	0.712	0.686	0.699
YOLOv8m	0.582	0.623	0.602
YOLOv8x	0.547	0.586	0.566
YOLOv10n	0.591	0.665	0.628
YOLOv5	0.626	0.653	0.639

5. Experiment

This article proposes an intelligent detection method for surface defects of aluminum ingots based on an improved YOLOv8n architecture. By introducing a parameter free SimAM three-dimensional attention mechanism, it effectively solves the detection problem of small targets and diverse defects. A high-definition aluminum ingot defect dataset based on actual production lines was constructed, and SimAM module was integrated into the network backbone layer, significantly improving the feature extraction ability of the model. The experimental results showed that the improved YOLOv8n SimAM model achieved a mAP of 64.5% in aluminum ingot surface defect detection, which was 9 percentage points higher than the baseline model. The F1 scores of burrs and slag inclusion defects increased by 6.3% and 9.9%, respectively. Compared with mainstream models such as YOLOv5 and YOLOv8m, this method shows significant advantages in detection accuracy and computational efficiency, providing a reliable technical solution for the intelligent manufacturing quality monitoring system of aluminum ingots.

References

- [1] Mery D. Aluminum casting inspection using deep object detection methods and simulated ellipsoidal defects. *Machine Vision and Applications*. 2021 May;32(3):72.
- [2] Chen T, Cai C, Zhang J, Dong Y, Yang M, Wang D, Yang J, Liang C. RER-YOLO: improved method for surface defect detection of aluminum ingot alloy based on YOLOv5. *Optics Express*. 2024 Feb 27;32(6):8763-77.
- [3] Chen H, Yan F, Yang J, Yan J, Qin T, Zhang J. Real-time surface defect detection algorithm on aluminum ingot alloy casting lines. *Physica Scripta*. 2024 Dec 31;100(1):016018.
- [4] Liu Gx, Liao P, Yang N X. Active vision method for pressure vessel weld quality parameter detection based on deep learning [J]. *Chinese Journal of Scientific Instrument*, 2023, 44(5): 1-9.
- [5] Zhuxi MA, Li Y, Huang M, Huang Q, Cheng J, Tang S. A lightweight detector based on attention mechanism for aluminum strip surface defect detection. *Computers in Industry*. 2022 Apr 1;136:103585.
- [6] Zhu X, Liu J, Zhou X, Qian S, Yu J. Enhanced feature Fusion structure of YOLO v5 for detecting small defects on metal surfaces. *International Journal of Machine Learning and Cybernetics*. 2023 Jun;14(6):2041-51.
- [7] Zhang R, Wen C. SOD-YOLO: A Small Target Defect Detection Algorithm for Wind Turbine Blades Based on Improved YOLOv5. *Advanced Theory and Simulations*, 2022, 5.
- [8] Yang R, Li W, Shang X, Zhu D, Man X. KPE-YOLOv5: An improved small target detection algorithm based on YOLOv5. *Electronics*. 2023 Feb 6;12(4):817.

- [9] Nascimento R, Ferreira T, Rocha CD, Filipe V, Silva MF, Veiga G, Rocha L. Quality Inspection in Casting Aluminum Parts: A Machine Vision System for Filings Detection and Hole Inspection. *Journal of Intelligent & Robotic Systems*. 2025 Apr 26;111(2):53.
- [10] Li N, Wang Z, Zhao R, Yang K, Ouyang R. YOLO-PDC: algorithm for aluminum surface defect detection based on multiscale enhanced model of YOLOv7. *Journal of Real-Time Image Processing*. 2025 Apr;22(2):86.
- [11] Han Y, Li X, Cui G, Song J, Zhou F, Wang Y. Multi-defect detection and classification for aluminum alloys with enhanced YOLOv8. *PloS one*. 2025 Mar 20;20(3):e0316817.
- [12] Yan S, Guo H, Liu S. A Partition Stacking Classification Framework With Oversampling for Quality Prediction of Aluminum Alloy Ingots. *International Journal of Artificial Intelligence and Green Manufacturing*. 2025 Apr 19;1(1).
- [13] Li Y, Yu W, Guan X. 3D localization for multiple AUVs in anchor-free environments by exploring the use of depth information. *IEEE/CAA Journal of Automatica Sinica*. 2023 Aug 1;11(4):1051-3.
- [14] Vijayakumar A, Vairavasundaram S. Yolo-based object detection models: A review and its applications. *Multimedia Tools and Applications*. 2024 Oct;83(35):83535-74.
- [15] Magacho G, Espagne E, Godin A. Impacts of the CBAM on EU trade partners: consequences for develop** countries. *Climate Policy*. 2024 Feb 7;24(2):243-59.
- [16] Liang L, Zhang Y, Zhang S, Li J, Plaza A, Kang X. Fast hyperspectral image classification combining transformers and SimAM-based CNNs. *IEEE Transactions on Geoscience and Remote Sensing*. 2023 Aug 28;61:1-9.