information

Article

Copper Nodule Defect Detection in Industrial Processes Using Deep Learning

Zhicong Zhang [1], Xiaodong Huang [2], Dandan Wei [1], Qiqi Chang [1], Jinping Liu [1] and Qingxiu Jing [1,*]

[1] Faculty of Materials Metallurgy and Chemistry, Jiangxi University of Science and Technology, Ganzhou 341000, China; zhang_work@mail.jxust.edu.cn (Z.Z.)
[2] School of Economics and Management, Jiangxi University of Science and Technology, Ganzhou 341000, China
* Correspondence: 9120030009@jxust.edu.cn

Abstract: Copper electrolysis is a crucial process in copper smelting. The surface of cathodic copper plates is often affected by various electrolytic process factors, resulting in the formation of nodule defects that significantly impact surface quality and disrupt the downstream production process, making the prompt detection of these defects essential. At present, the detection of cathode copper plate nodules is performed by manual identification. In order to address the issues with manual convex nodule identification on the surface of industrial cathode copper plates in terms of low accuracy, high effort, and low efficiency in the manufacturing process, a lightweight YOLOv5 model combined with the BiFormer attention mechanism is proposed in this paper. The model employs MobileNetV3, a lightweight feature extraction network, as its backbone, reducing the parameter count and computational complexity. Additionally, an attention mechanism is introduced to capture multi-scale information, thereby enhancing the accuracy of nodule recognition. Meanwhile, the F-EIOU loss function is employed to strengthen the model's robustness and generalization ability, effectively addressing noise and imbalance issues in the data. Experimental results demonstrate that the improved YOLOv5 model achieves a precision of 92.71%, a recall of 91.24%, and a mean average precision (mAP) of 92.69%. Moreover, a single-frame detection time of 4.61 ms is achieved by the model, which has a size of 2.91 MB. These metrics meet the requirements of practical production and provide valuable insights for the detection of cathodic copper plate surface quality issues in the copper electrolysis production process.

Keywords: electrolytic cathodic copper plate; YOLOv5; attention mechanism; deep learning; mobilenetv3; object detection

Citation: Zhang, Z.; Huang, X.; Wei, D.; Chang, Q.; Liu, J.; Jing, Q. Copper Nodule Defect Detection in Industrial Processes Using Deep Learning. *Information* **2024**, *15*, 802. https://doi.org/10.3390/info15120802

Received: 8 November 2024
Revised: 3 December 2024
Accepted: 6 December 2024
Published: 11 December 2024

1. Introduction

In industrial production, high-purity copper cathodes are commonly produced through electrolytic refining. However, during the process of purification, convex nodules and particles appear on the surface of the cathode copper plate. These nodules grow continuously and eventually make contact with the anode, resulting in a short circuit between the cathode and anode, which ultimately leads to a significant decrease in current efficiency and deterioration in the quality of the copper plate [1,2]. Current research indicates that the formation of nodules on copper cathode plates is caused by a combination of multiple factors. For instance, during the electrolysis process, the dissolution of the anode leads to the production of solid particles, such as anode sludge, which adhere to the surface of the cathode copper plates and induce nodule formation [3]. Improper use of additive components and their ratios can cause roughening of the cathode copper plate surface, thereby increasing the risk of nodule formation [4]. Additionally, uneven current distribution often leads to nodule formation on the cathode copper plates. If the cathode is tilted, the current increases significantly in the area closer to the anode, thereby raising the likelihood of nodule formation [5]. However, in actual industrial production environments, the factors

contributing to nodule formation on copper plates are highly complex and difficult to avoid completely. The occurrence of nodulation on the surface of a copper cathode plate not only deteriorates the quality of the copper plate but also impacts downstream production processes, making the detection of nodulation defects essential. Currently, the detection of nodule formation on the surface of cathode copper plates primarily relies on manual visual inspection methods [6,7], as illustrated in Figure 1. This method is often influenced by subjective human judgment, making it challenging to ensure detection accuracy. Additionally, the manual identification method faces several challenges, including high labor costs, low identification efficiency, and a lack of capability for continuous identification tasks. Therefore, it is of great significance to develop an efficient and accurate algorithm for detecting nodules on copper cathode plates. This algorithm aims to reduce labor costs in production, improve industrial efficiency, and promote the advancement of intelligent manufacturing.

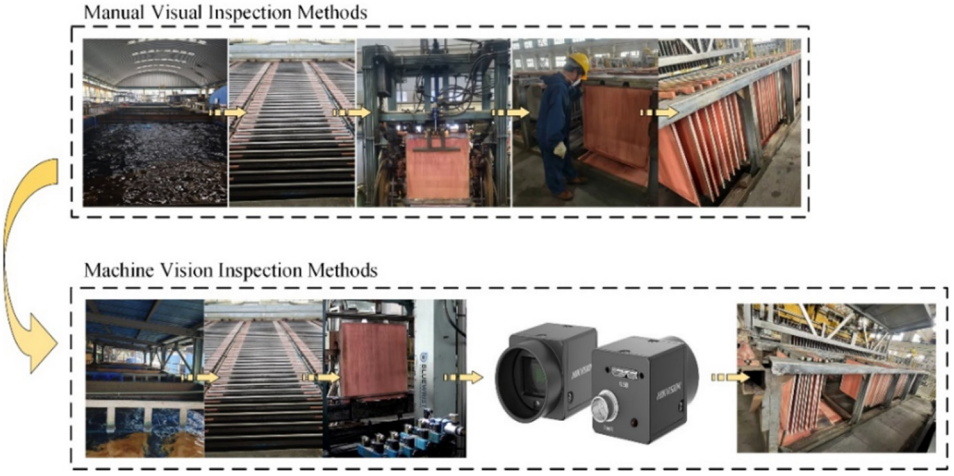


Figure 1. Schematic diagram of the copper electrolytic process.

Prior to the widespread use of deep learning technology, defect identification of metal materials relied on conventional image processing methods and machine learning approaches. Traditional image processing procedures consist of image acquisition and preprocessing, feature extraction and selection, classification, and post-processing. Lee et al. [8] used a combined surface normal Gabor filter (SNGF) to identify surface imperfections. Firstly, the topological structure of the surface of the object was standardized, then the surface normal vector was converted into a complex form, and the geometric features of the defect were extracted, confirming the effectiveness and applicability of the SNGF method. Cui et al. [9] proposed a detection method for strip steel surface defects based on wavelet denoising and an improved Canny algorithm. This method enhances image contrast and addresses uneven brightness issues through improved wavelet denoising and homomorphic filtering. Additionally, the optimized Canny algorithm is utilized to accurately detect defect edges on the strip steel, thereby significantly improving the detection accuracy of surface defects. Chen et al. [10] proposed an unsupervised edge detection algorithm based on k-means (PK) principal component analysis for identifying surface cracks on industrial metal components. The algorithm selects temperature features using a particle swarm optimization method and optimizes the hyperparameters of support vector machines and k-nearest neighbors using a Bayesian search approach. The method achieved a crack detection rate of 82.9% and an accuracy rate of 86.2%. The results demonstrate that traditional image processing and machine learning methods can effectively achieve defect detection on metal surfaces.

With the rapid advancements in convolutional neural networks in the realm of object recognition and the widespread adoption of deep learning in computer vision applications, more and more researchers have begun to implement deep learning-driven surface defect detection technologies across various industrial applications [11–17]. Jiang et al. [18] in-

vestigated a surface defect detection method for metal mesh based on an improved Faster R-CNN algorithm. The method replaces the original Visual Geometry Group network (VGG16) with a deep residual network (ResNet50). Additionally, k-means and genetic algorithms were introduced to design anchor box sizes better suited to the metal mesh dataset, effectively addressing the issue of inaccurate defect localization. Li et al. [19] proposed the YOLOv5s-STCD algorithm. This approach integrates the last C3 module of the backbone extraction network with the Swin Transformer structure, introduces the coordinate attention (CA) mechanism, replaces the original detection head with a decoupled head, and substitutes the CIoU loss function with the SIoU loss function. These improvements enhance the model's detection performance, making it suitable for defect detection in steel.

To detect and locate the nodule defects on a cathode copper plate, enhance detection efficiency, and facilitate intelligent production, a lightweight YOLOv5 model combined with the BiFormer attention mechanism is proposed. The main contributions of this article can be summarized as follows:

1. As the primary feature extraction network, the lightweight MobileNetV3 network is presented, significantly decreasing the model's computational demands and parameter count while expediting the training process.
2. In the neck network, the model's performance in detecting small targets is enhanced by incorporating the BiFormer attention mechanism.
3. By combining the advantages of Efficient IoU loss and Focal loss, the initial CIoU loss function is replaced with the Focal-EIoU loss function. This change addresses the disadvantages of low resolution, significant noise interference, and low contrast between nodules and the cathode copper plate images, ultimately improving detection accuracy.
4. In order to apply the detection model to industrial production, a detection system for cathode copper plate surface nodules was built.

The subsequent parts of this paper are organized as follows: In Section 2, we review current work on defect detection. Section 3 details the specific framework of YOLOv5 and the improvements made. In Section 4, we present the copper plate dataset and experimental environment, and provide an analysis of the results. Finally, Section 5 summarizes the paper.

2. Related Work

With the advancement of convolutional neural networks, the performance of deep learning-based object detection algorithms has significantly improved, leading to important applications in product identification and defect detection. Currently, there are two primary types of target detection algorithms. The first type is the two-stage object detection algorithm, which relies on candidate regions and detects targets by extracting a series of candidate regions from the image for classification. The second type is the one-stage object detection algorithm based on regression, which identifies all locations in the image and utilizes a regression algorithm to determine the position and dimensions of the target.

2.1. Two-Stage Object Detection Algorithm

Girshick et al. [20] proposed the object detection algorithm R-CNN, which combines region proposals with CNNs based on the standard metrics of the PASCAL VOC dataset. This algorithm performs region proposals from top to bottom through extensive convolutions and applies supervised pre-training on limited data. The findings demonstrate that the mean average precision (mAP) of the algorithm improves by over 30%. However, the operational speed decreases due to the redundant data generated during the algorithm's execution, making it unable to meet the high-efficiency demands of factory application scenarios.

Wu et al. [21] improved the Faster R-CNN algorithm for steel strip defect detection, achieving an mAP of 79.5%. Their approach integrates deformable convolutions and a multi-scale detection module into the FPN backbone, and incorporates the CBAM attention

mechanism in the RPN to optimize feature representation. However, despite its high accuracy, the method requires considerable computational and memory resources, which may limit its applicability on resource-constrained devices.

Yang et al. [22] improved the Mask R-CNN framework and proposed a new algorithm for metal surface defect detection. The method adds bottom-up feature fusion paths to the original FPN, and integrates the CBAM attention mechanism with deformable convolutions to enhance the contextual awareness of ResNet. Additionally, the IOU and DIOU loss functions are used to optimize the localization accuracy of the predicted bounding boxes. Experimental results show that the algorithm achieved an mAP of 78.6% on a specialized dataset, validating its effectiveness. However, due to the complex hyperparameter configuration, the optimization process requires significant time and computational resources.

The two-stage algorithm offers significant insights into the processes and improvement directions in object detection. It includes two fundamental phases: identifying potential target areas and subsequently classifying and localizing the targets. Although the RPN (Region Proposal Network) has been refined to enhance performance using various techniques, like Libra R-CNN [23] and Cascade R-CNN [24], these improvements increase the algorithm's complexity and training difficulty. In light of the need to enhance algorithm efficiency, experts are increasingly turning to one-stage object detection algorithms to achieve a better balance between training difficulty, detection accuracy, and speed.

2.2. One-Stage Object Detection Algorithm

Wang et al. [25] developed an enhanced algorithm based on YOLOv3 and YOLOv4 to improve the precision and efficiency of metal surface defect detection by introducing the concepts of focus structure and anchorless frames. Wang et al. [26] proposed the YOLO-Gear method, which is based on the YOLOv5s network and achieves efficient end-face defect detection of metal gears through enhancements in the adaptive convolutional attention module and feature pyramid network. Yang et al. [27] introduced a method that utilizes the coordinate attention mechanism alongside improvements at the input end, additive module, and detection end of YOLOv5 to effectively enhance the model's generalization capability and detection precision. Wang et al. [28] proposed the YOLOv5-GSE model, which integrates Ghost convolution by replacing the traditional convolution operation in the backbone network, significantly improving the efficiency and accuracy of metal surface defect detection. Song et al. [29] proposed an improved YOLOv8 algorithm that replaces the convolution in the Bottleneck module with DCNv2, enhances feature fusion using BiFPN, introduces the BiFormer attention mechanism, and employs the WIoUv3 loss function to improve the performance of steel surface defect detection. Xie et al. [30] developed the LMS-YOLO model, which incorporates the lightweight C2f_LMSMC module into the backbone network, introduces an efficient global attention mechanism, utilizes three independent decoupled heads for regression and classification, and replaces CIoU with NWD as the regression loss to enhance the detection of small-scale defects.

In this article, we selected YOLOv5-S as the base model for several reasons: (1) YOLOv5 has been validated by extensive academic research and practical application, which demonstrates its effectiveness and stability in target detection tasks, ensuring reliable performance in various environments. Additionally, YOLOv5 has a smaller parameter size, lower computational complexity, and relatively lower resource requirements, making it more adaptable in resource-limited scenarios. (2) YOLOv5 offers comprehensive development documentation and numerous examples, enabling users to adopt and implement the model quickly. Additionally, the active community support offers users access to a wealth of resources, discussions, and solutions, greatly facilitating learning and technical exchanges during the development process.

3. Method

3.1. YOLOv5

YOLO (You Only Look Once) is an end-to-end object detection algorithm designed to meet the demands of real-time detection [31]. YOLOv5 is the fifth generation of the YOLO algorithm, subdivided into four models of varying scales: YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x, with YOLOv5s being the most lightweight model in the series. In this section, the YOLOv5s model will be used as an example to provide a detailed explanation of the composition of the YOLOv5 architecture, which is primarily composed of four components: input, backbone, neck, and prediction module. The main structure is illustrated in Figure 2 below.

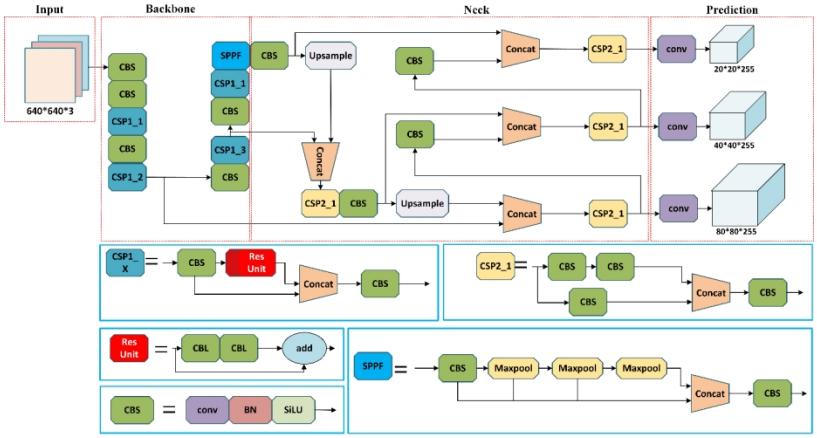


Figure 2. Network structure of the YOLOv5s algorithm.

Input: Mosaic data augmentation, adaptive anchor frame calculation, and adaptive image zoom techniques are employed in the input stage of YOLOv5s.

Backbone: The backbone is designed to extract relevant features from the input image and convert them into a multi-layer feature map for subsequent object detection tasks. It consists of the CBS (Conv BatchNorm SiLU) structure, Focus structure, CSP (Cross Stage Partial) structure, and SPP (Spatial Pyramid Pooling) structure.

Neck: The neck network consists of a feature pyramid network (FPN) [32] with upsampling operations and a pyramid attention network (PAN) [33] with downsampling operations, effectively integrating multiple levels of features. The design allows the network to combine high-level features with low-level features, facilitating strong semantic information transfer through the FPN and precise positioning information extraction through the PAN, effectively capturing target information across different scales.

Prediction Module: The prediction module employs a convolutional layer to convert feature maps into object detection outputs. It performs category prediction and annotation, determines the position and bounding box, and resizes the feature map to match the original image.

3.2. YOLOv5 Algorithm Improvements

3.2.1. MobileNetV3 Backbone Network

To minimize the number of parameters and computations in the model, the feature extraction network utilizes the lightweight MobileNetV3 [34], employing depthwise separable convolution instead of traditional convolution, adding the SE attention mechanism module, and introducing the nonlinear activation function, h-swish, which further enhances the model's performance and generalization capability. Figure 3 illustrates the structure of MobileNetV3.

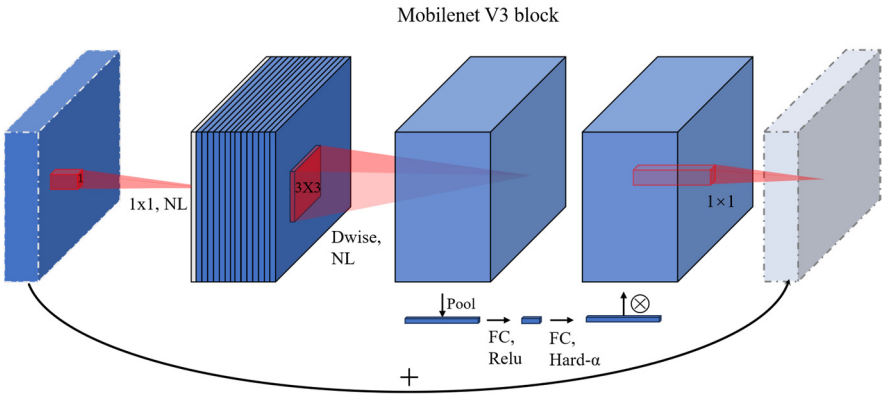


Figure 3. Basic network unit of MobileNetV3.

Depthwise separable convolution comprises two essential components: depthwise convolution and pointwise convolution. In depthwise convolution, the unique characteristics of each channel in the input feature matrix are effectively extracted by applying the convolution kernel to each channel, resulting in an equivalent number of feature maps. This approach reduces both the floating-point computations and the parameter count compared to standard convolution. In pointwise convolution, the input features are convolved with a 1×1 convolution kernel. This process is repeated n times to increase the number of channels to n layers. The processes of traditional convolution and depthwise separable convolution are illustrated in Figure 4.

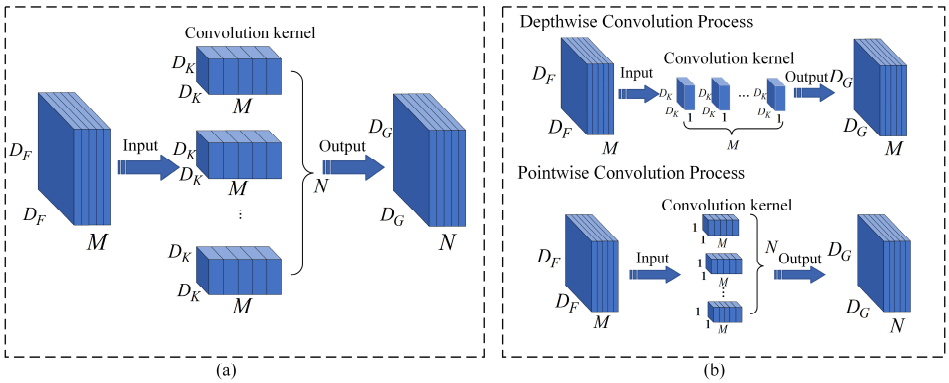


Figure 4. (**a**) Traditional convolution process diagram; (**b**) Depthwise separable convolution process diagram.

The total amount of operations during the convolution process can be decreased by using depthwise separable convolution procedures, given that the size of the input data is $D_F \times D_F \times M$, and the number of output channels is N. If the convolution is performed with N convolution kernels of size $D_K \times D_K \times M$ employing the traditional convolution method, the total number of parameters required is $D_K \times D_K \times M \times N \times D_F \times D_F$.

If the depthwise separable convolution operation is employed, the convolution is performed using M convolution kernels with a size of $D_K \times D_K \times 1$ during the depthwise convolution process. In pointwise convolution, the convolution operation is conducted with N convolution kernels of size $1 \times 1 \times M$ to achieve the same convolution effect as traditional convolution. The total number of required parameters is $D_K \times D_K \times M \times D_F \times D_F + M \times N \times D_F \times D_F$. Equation (1) illustrates the ratio of the number of parameters for separable convolution to those for conventional convolution.

$$\frac{D_K \times D_K \times D_X \times D_X \times M + D_K \times D_K \times M \times D_X \times D_X \times N}{D_K \times D_K \times M \times D_X \times D_X \times N} = \frac{1}{N} + \frac{1}{D^2} \quad (1)$$

The cost of computing resources saved is proportional to the dimensions of the output channels and the dimensions of the convolution kernel.

The inverted residual structure from MobileNetV2 [35] is incorporated into MobileNetV3, as shown in Figure 5a. The inverted residual algorithm proceeds as follows: First, a 1×1 convolution is employed to map low-dimensional compressed features into high-dimensional space. Then, depthwise separable convolution is utilized for feature extraction. Finally, another 1×1 convolution is utilized to reduce dimensionality and map the features back to a low-dimensional space.

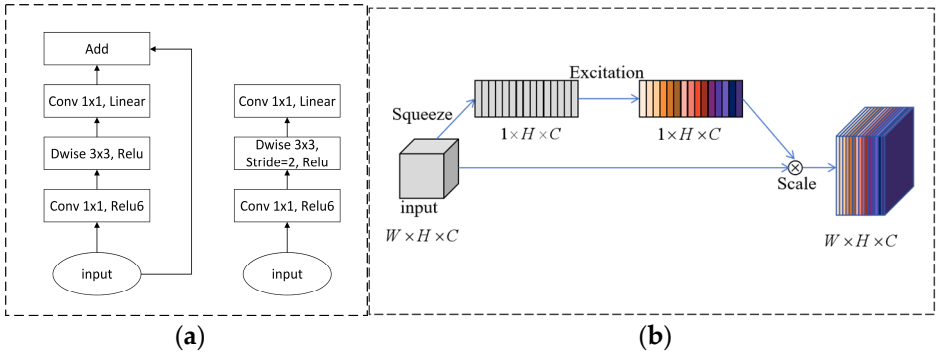


Figure 5. (**a**) Inverted residual structure of MobileNetV3; (**b**) Schematic diagram of SE network structure.

The SE (Squeeze-Excitation) [36] module, which consists of two parts—squeeze and excitation—is incorporated into MobileNetV3, as illustrated in Figure 5b. In the figure, W, H, and C represent the dimensions of the feature map—width, height, and number of channels—while the size of the input feature map is $W \times H \times C$. The calculation process is outlined as follows: First, the input feature map undergoes squeeze, which performs global average pooling to compress it into a vector of size $1 \times 1 \times C$. Second, the excitation operation processes this vector through two fully connected layers. The final step involves scaling, where the calculated weights for each channel are multiplied by the corresponding two-dimensional matrix of the original feature map to yield the output. The SE module is capable of adaptively learning the significance of each channel and modifying its contributions in the feature map based on task requirements. This enables the network to focus more on important feature channels, thereby enhancing the model's performance.

3.2.2. BiFormer Attention Mechanism

Due to the presence of various irrelevant interferences on the copper cathode plate's surface, the extraction of features is affected. To address this, the attention mechanism [37] guides the algorithm to concentrate on the critical areas by applying weights to the input features. Thus, the BiFormer attention mechanism is incorporated to assist the model in focusing on critical areas with potential defects on the copper plate's surface, thereby enhancing detection accuracy.

BiFormer [38] utilizes the Bi-Level Routing Attention (BRA) mechanism to enable efficient computation distribution in a flexible and query-aware manner. Figure 5 illustrates that the BRA structure diagram mainly comprises three steps. The first stage involves segmenting the region into input mappings to efficiently process the data and enhance information relevance by constructing a region-level affinity map. For a given two-dimensional input feature map $X \in R^{H \times W \times C}$, segment it into $S \times S$ non-overlapping regions to ensure that each region contains $\frac{HW}{S^2}$ collections of feature vectors, transform X into X^r via a reshape operation, and finally obtain Q, K, and V using a linear mapping function. Equations (2)–(4) indicate the process.

$$Q = X^r W^q \tag{2}$$

$$K = X^r W^k \tag{3}$$

$$V = X^r W^v \tag{4}$$

The symbols W^q, W^k, and W^v represent the mapping matrix of query, key, and value, respectively.

The second stage is to implement inter-region routing using a directed graph. To do this, calculate the average value of Q and K for each region, and obtain the region-level query Q^r and key K^r, and then perform matrix multiplication and softmax operations on Q^r and K^r to obtain an inter-region adjacency matrix. This matrix is then pruned to obtain I^r by retaining only the first k connection, as illustrated in Equations (5) and (6).

$$A^r = Q^r (K^r)^T \tag{5}$$

$$I^r = topkIndex(A^r) \tag{6}$$

where Q^r represents the query matrix, K^r represents the key matrix, A^r represents the attention matrix, and I^r represents the computational index matrix.

The third stage focuses on fine-grained computation. The index matrix I^r is utilized to perform self-attention computation on fine-grained tokens in each region. The gather function collects the key values, Q and K, followed by self-attention computation on Q^g and K^g within fine-grained regions. Equations (7)–(9) illustrate the process.

$$K^g = gather(K, I^r) \tag{7}$$

$$V^g = gather(V, I^r) \tag{8}$$

$$O = attention(Q, K^g, V^g) + LCE(V) \tag{9}$$

where the *attention* function is used to calculate the attention weights, K^g and V^g represent the mapping of keys and values on coarse-grained granularity, and Q represents the mapping of the query matrix. $LCE(V)$ represents the enhancement of local context within the specified region.

The network architecture of the BiFormer attention mechanism, which employs BRA as a fundamental building block for constructing a four-stage pyramid structure, is illustrated in Figure 6. The initial phase involves embedding overlapping patches. During the second to fourth stages, patch merging modules are employed to reduce the spatial resolution of the input while increasing the number of channels. Subsequently, N_i consecutive BiFormer blocks are applied to modify the features. Each BiFormer block initially encodes relative position information by utilizing a 3 × 3 deep convolution. Subsequently, a two-layer routing attention mechanism module and a two-layer MLP module, with an extension ratio of e, are employed to model the cross-position relationships and the embedding positions.

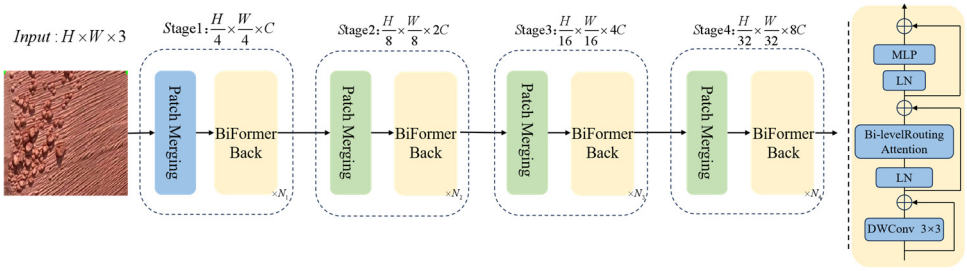


Figure 6. Network structure diagram of the BiFormer's attention mechanism.

By introducing the two-layer routing mechanism of BRA, the BiFormer balances the relationship between addressing long-distance dependencies and optimizing computational efficiency while maintaining model flexibility, thereby effectively enhancing computational efficiency.

3.2.3. F-EIOU Loss Function

The effectiveness of the loss function in deep learning algorithms directly influences the training results and the final performance of the model. In this model, the original CIOU [39] loss function is replaced with the EIOU [40] loss function, which penalizes the discrepancy between length and breadth and substitutes the portion of the CIOU loss that tracks aspect ratio consistency. This modification more accurately reflects the true differences between width and height, thus facilitating faster network convergence. The calculation formula for EIOU is detailed below:

$$L_{EIOU} = L_{IOU} + L_{dis} + L_{asp} = 1 - IOU + \frac{\rho^2(b, b^{gt})}{C^2} + \frac{\rho^2(W, W^{gt})}{C_w^2} + \frac{\rho^2(h, h^{gt})}{C_h^2} \quad (10)$$

where C_h and C_w represent the height and width of the minimum circumscribed rectangle for both the predicted box and the actual box; W and W^{gt} are the widths of the prediction box and the real box; h and h^{gt} are the lengths of the predicted box and the actual box; b is the center point of the predicted box, and b^{gt} represents the central coordinates of the actual box; $\rho(W, W^{gt}) = |W - W^{gt}|$, $\rho(h, h^{gt}) = |h - h^{gt}|$.

The EIOU divides the loss function into three components: the direction loss function $L_{asp} \equiv \frac{\rho^2(W,W^{gt})}{C_w^2} + \frac{\rho^2(h,h^{gt})}{C_h^2}$; the overlap loss between the predicted box and the actual box $L_{IOU} = 1 - IOU$; and the distance loss function $L_{dis} = \frac{\rho^2(b,b^{gt})}{C^2}$. Figure 7 shows the schematic diagram of the EIOU loss.

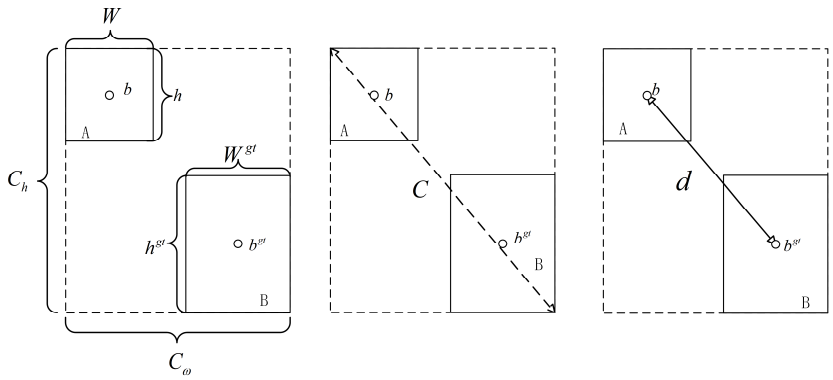


Figure 7. Schematic diagram of EIOU loss.

Poorly regressed samples significantly impact the regression loss. To enhance the model's performance and precision, the Focal concept is integrated into the EIOU loss function to balance samples with varying regression quality. The final F-EIOU loss function is derived by combining the EIOU loss function and the Focal loss [41] function, and its expression is shown in Equation (11).

$$L_{F-EIOU} = IOU^{\gamma} L_{EIOU} \quad (11)$$

Among them, γ is applied to fine-tune the loss function's sensitivity to various IOU regions, and IOU^{γ} is used to adjust the effect of the intersection union ratio loss.

The F-EIOU bounding box loss function is designed to address the issue of sample imbalance and to introduce consistency penalty terms and adaptability parameters. This enhances the model's robustness and contributes to improved training efficiency for the network. The neural network's predictions and expected values can be evaluated using three types of loss functions. The closer the predictions align with the expected values, the lower the value of the loss function. Figure 8 illustrates the variations in box_loss and obj_loss for the F-EIOU loss function compared to the CIOU, DIOU, GIOU, and EIOU loss functions after 500 training epochs. The results show that the obj_loss of the F-EIOU loss function converges faster than that of the other loss functions. Additionally, for the same

training epochs, the F-EIOU loss function demonstrates lower values for both obj_loss and box_loss. This indicates that the enhanced model demonstrates superior convergence speed and reliability, facilitating more accurate image sample recognition.

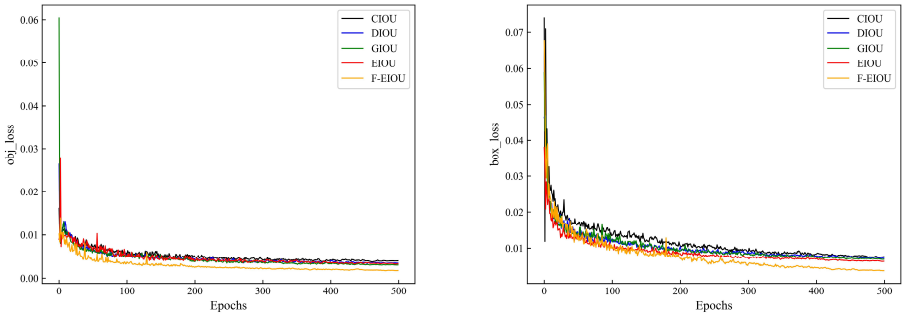


Figure 8. Comparison of loss functions: (**a**) obj_loss comparison; (**b**) box_loss comparison.

4. Experiments

4.1. Dataset Preparation

The experimental dataset used in this study was collected from an electrolytic copper cathode production line at a factory in Jiangxi Province, China. Images of the cathode copper plates were captured using a Dahua color industrial camera (model: A3A10CG8), ensuring precise capture of lump formation details. The data collection environment was a standardized industrial setting designed to minimize external interference and maintain consistent lighting conditions.

The dataset contains 2117 images, which are divided into training (1695 images), validation (211 images), and test (211 images) sets. Among these, 1906 images were annotated using Labelme software (Version: 5.5.0), with a total of approximately 34,992 nodule formation labels. To prevent excessive label occlusion due to the large number of targets, labels are represented by "0". The annotations underwent multiple rounds of by a team of three members to ensure quality and consistency. This dataset is highly representative, effectively reflecting the challenges encountered in industrial lump detection, such as small-sized targets and diverse shapes. An example of nodulation images is shown in Figure 9.

Figure 9. Legend for nodule defects on cathode copper plates.

4.2. Experimental Environment

During the training process of the experiments, we carefully configured the training parameters to ensure optimal model performance. The batch size was set to 16, and the training duration was 500 epochs. The initial learning rate was set to 0.01, the momentum was set to 0.937, and the weight_decay was set to 0.0005. The SGD optimization function was used, and cosine annealing was applied to dynamically modify the learning rate.

In this paper, Windows 10 Professional is used as the operating system version on the experimental computer, the CPU model is 16 vCPU Intel(R) Xeon(R) Platinum 8352V CPU @ 2.10 GHz, the GPU used is RTX 4090 for computing, the video memory size is 24 GB, the Pytorch 2.0 deep learning framework serves as the model's foundation, and the Python version is 3.10.13, and Cuda 11.8 is used for GPU acceleration.

4.3. Evaluation Metrics

In this paper, precision, recall, mean average precision (mAP), params, GFLOPS, SpeedGPU, and model size are selected as evaluation indicators. Precision is calculated by dividing the number of samples predicted as positive by the number of correctly predicted positive samples. Recall is the ratio of correctly predicted positive samples to the total number of actual positive samples. SpeedGPU refers to the time required for the model to detect an image, and the model size is the weight file size attained at the conclusion of training. The specific calculation formula is as follows:

$$R_P = \frac{TP}{TP + FP} \tag{12}$$

$$R_R = \frac{TP}{TP + FN} \tag{13}$$

$$AP = \int_0^1 R_P(t)dt \tag{14}$$

$$mAP = \frac{\sum_{t=1}^{N} AP(t)}{N} \tag{15}$$

where TP measures how many samples the model correctly classifies as positive, FP represents the number of samples where the model incorrectly classifies negative examples as positives, and FN represents the number of samples where the model incorrectly classifies positive examples as negatives.

4.4. Experimental Results and Analysis

4.4.1. Ablation Experiments

The improved YOLOv5 model was trained, and the graph generated during the training is shown in Figure 10.

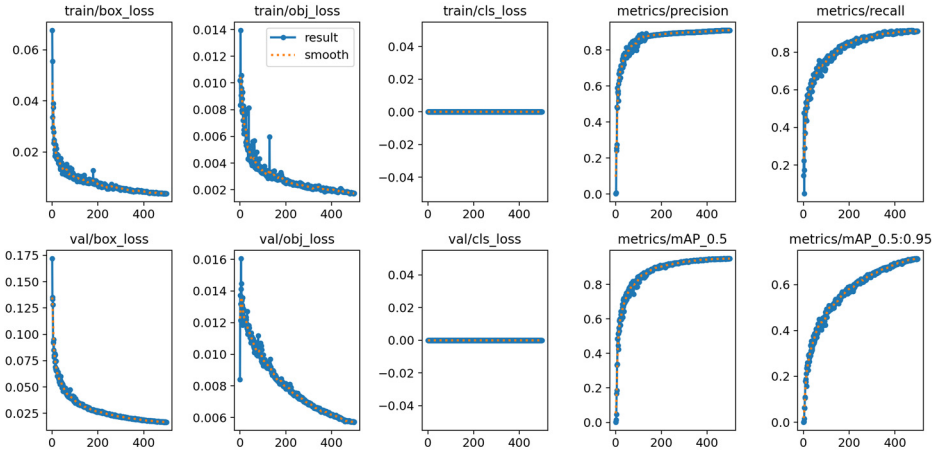


Figure 10. Graphs produced during training of the improved YOLOv5.

Considering the improvements made during the lightweighting of the network structure, this research conducts ablation experiments to assess the actual effects of each component of YOLOv5 on both model performance and efficiency, thus verifying the specific impact of each improvement. The experimental results are shown in Table 1.

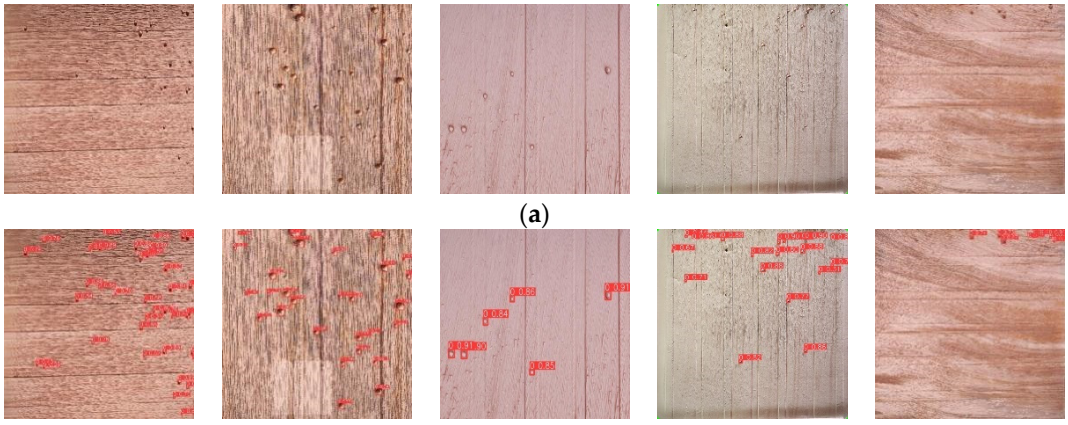
Table 1. Ablation experiments result (√ means we used this module).

Module			Precision (%)	Recall (%)	mAP (%)	Params (B)	SpeedGPU (ms)	Weight (M)
MobileNetV3	BiFormer	EIOU						
			93.27	90.40	92.38	7,020,913	6.99	13.62
√			87.34	85.93	86.23	1,509,487	4.57	2.87
	√		94.71	92.06	93.01	7,020,913	7.14	13.75
		√	93.45	91.03	92.74	7,088,241	6.87	13.62
√	√		92.13	89.42	91.67	1,531,951	4.69	2.91
√		√	87.83	86.48	86.77	1,509,487	4.64	2.87
√	√	√	92.71	91.24	92.69	1,531,951	4.61	2.91

With Mobilenetv3 as the primary network for feature extraction, the number of parameters was reduced by 85% compared to the original model, resulting in a significant compression of the model size and a 34.6% increase in detection speed under GPU acceleration. This improvement is largely attributed to the use of depthwise separable convolution (DSC), which is especially beneficial given the limited computing power of mobile platforms. However, while the model is lightweight, its performance is slightly reduced because the internal parameter calculation of Mobilenetv3 is not as effective as that of CSPDarknet-53, resulting in a loss of key information during feature extraction and a subsequent decrease in certain classification accuracies, thereby reducing overall defect detection accuracy. Following the addition of the lightweight BiFormer attention mechanism, precision, recall, and mean average precision increased by 5.4%, 4%, and 6.3%, respectively, while the number of parameters only increased by 1.4%. This improvement is due to BiFormer paying attention to a small part of the relevant markers in a query-adaptive way, without diverting attention to unrelated markers, thereby enhancing the model's feature extraction capability and enabling accurate target detection. The lightweight model's precision increased by 0.5%, and the mean average precision increased by 0.6% with the addition of the F-EIOU loss function.

Although the improved YOLOv5 model exhibits slightly lower precision than the original version, the recall and mean average precision values have increased by 1% and 0.4%, respectively. Additionally, the size of the weight file has been reduced by 78.6%, the number of parameters has decreased by 78.1%, and the detection speed has increased by 34%. These enhancements fulfill the goal of achieving a lightweight model.

The visualization results of this model for identifying nodules on a copper cathode plate surface are presented in Figure 11.



(**a**)

(**b**)

Figure 11. Model detection effect diagram: (**a**) The original image; (**b**) The image after model detection.

4.4.2. Comparison of the Improved Model with Different Models

To evaluate the performance of different YOLO models in cathodic copper plate surface nodule detection, a comparative analysis of various YOLO models was conducted to select the most suitable model for this task. The results presented in Table 2 show the performance comparison of these models in detecting nodules on cathodic copper plate surfaces.

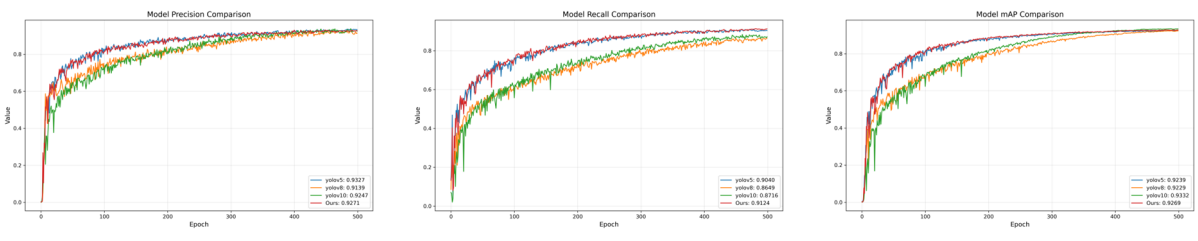
Table 2. Comparison of different YOLO models.

Module	Precision (%)	Recall (%)	mAP (%)	Params (B)	Weight (M)
YOLOv5s	93.27	90.40	92.38	7,020,913	13.62
YOLOv8s	91.39	86.49	92.29	11,166,544	22.6
YOLOv10s	92.47	87.16	93.32	8,067,126	16.6
Ours	92.71	91.24	92.69	1,531,951	2.91

As shown in Table 2, YOLOv5 demonstrates outstanding performance in terms of precision (93.27%) and recall (90.40%), particularly excelling in achieving a balance between these two metrics. Although the mAP of YOLOv8 and YOLOv10 is comparable to or slightly higher than that of YOLOv5, their recall rates are 86.49% and 87.16%, respectively, which are significantly lower than YOLOv5′s recall rate of 90.40%. A lower recall rate could lead to the omission of critical targets, such as missing nodule regions in cathodic copper plate surface nodule detection, thereby compromising the comprehensiveness and accuracy of detection. Therefore, in applications such as nodule detection, which require a high recall rate, YOLOv5 demonstrates a clear advantage.

In addition, the parameters of YOLOv8 and YOLOv10 are 11,166,544 B and 8,067,126 B, respectively, exceeding that of YOLOv5 (7,020,913 B) by approximately 59.2% and 14.9%. This increase may impose a higher computational burden in practical applications, particularly in hardware-constrained scenarios, potentially leading to slower processing speeds or memory overflows. In contrast, YOLOv5 offers higher computational efficiency, making it more suitable for resource-limited environments. For nodule detection on copper plates, considering the hardware limitations, YOLOv5 strikes a better balance between performance and computational requirements, ensuring efficient and accurate detection even under constrained resource conditions.

The improved YOLOv5 model demonstrates higher mAP (92.69%) and recall (91.24%) compared to the original YOLOv5, while significantly reducing the number of parameters (1,531,951 B). The weight file size is only 2.91 MB, greatly enhancing computational efficiency. This allows the model to maintain high detection accuracy while reducing the computational burden. Therefore, the improved YOLOv5 model further optimizes both performance and computational efficiency based on the advantages of YOLOv5, making it particularly suitable for resource-limited application scenarios. Figure 12 presents a comparison of precision, recall, and mAP for each model.



(**a**) (**b**) (**c**)

Figure 12. Performance Comparison of YOLO Models: (**a**) Precision Comparison of YOLO Models; (**b**) Recall Comparison of YOLO Models; (**c**) mAP Comparison of YOLO Models.

To further validate the enhanced lightweight YOLOv5 model's detection performance, experiments were conducted to compare it with the single-stage lightweight object detection

algorithms YOLOv3-tiny and SSD, as well as the two-stage object detection algorithm Faster RCNN. Evaluations were based on multiple indicators, such as mean average precision and accuracy. Table 3 presents the results of the comparative experiments.

Table 3. Comparison of different models.

Module	Precision (%)	Recall (%)	mAP (%)	Weight (M)
Yolov3-tiny	79.72	75.13	73.72	17.5
SSD	86.25	75.72	89.75	94.34
Faster RCNN	79.31	70.68	68.51	108.29
Ours	92.71	91.24	92.69	2.91

The table clearly illustrates that the model proposed in this research performs noticeably better than alternative target recognition methods. With a size of just 2.91 M, the trained model provides significant advantages for deployment on portable computers or embedded systems while maintaining remarkable performance.

5. Industrial Application

To successfully apply the improved YOLOv5 model in real-world industrial production, a real-time detection system based on the PyQt5 framework has been developed specifically for identifying nodule defects on the surface of electrolytic cathode copper plates. The system is capable of performing real-time online object detection, fully leveraging the advantages of deep learning models to accurately identify nodules on the copper plate surface. By adjusting the IOU (intersection over union) and confidence thresholds within the system, the detection precision and recall can be flexibly tuned to meet the requirements of different production environments.

Specifically, higher IOU and confidence thresholds ensure more accurate object localization and classification, reducing false positives and false negatives. Conversely, lower thresholds increase the number of detected nodules, thereby enhancing recall, which is suitable for scenarios requiring high detection rates, such as identifying as many defects as possible. This adaptive adjustment capability allows the system to optimize its performance according to specific task requirements and application scenarios, ensuring accurate detection results in diverse environments. This flexibility makes the system applicable not only to defect detection in electrolytic copper plates but also to similar industrial inspection tasks.

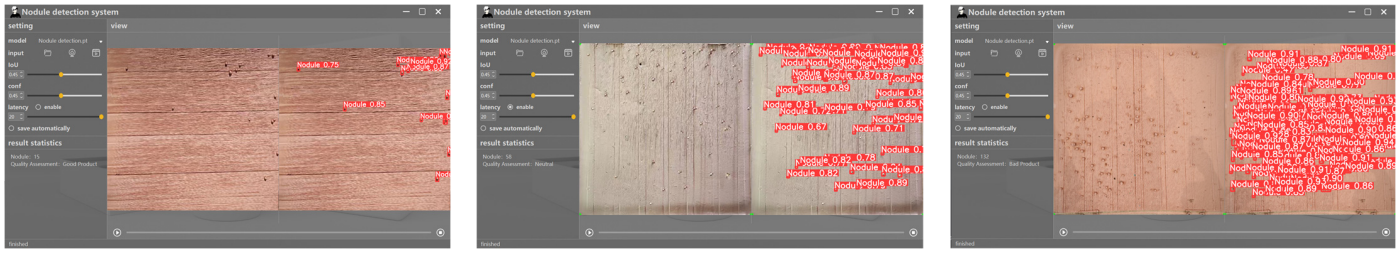
In addition to detecting the number of nodules, this study incorporates a quality assessment of nodulations to provide a more comprehensive analysis of copper plate quality. The specific evaluation criteria are as follows:

(1) When the number of nodules is less than or equal to 30, the quality is assessed as "Good", indicating that there are few surface defects and the overall quality is high.
(2) When the number of nodules is between 30 and 100, the quality is assessed as "Neutral", suggesting that there are a certain number of defects on the copper plate surface, but they remain within an acceptable range.
(3) When the number of nodules exceeds 100, the quality is assessed as "Poor", meaning that the copper plate surface has significant defects that may impact the product's quality and subsequent use.

By combining both the quantity and quality assessment of nodules, the system provides more detailed defect information for the production line, helping production personnel take timely actions for quality control. The application of this real-time detection system not only enhances production efficiency but also reduces the workload of manual inspection.

Figure 13 demonstrates the application scenario of the real-time detection system in the detection of nodules on electrolytic cathode copper plates. The system interface is clear and user-friendly, providing a real-time display of detection results and quality

assessments, allowing production personnel to promptly understand the quality status of the copper plates.



(**a**) Good Product (**b**) Neutral Product (**c**) Bad Product

Figure 13. Nodule detection system.

6. Limitation and Future Work

Although this study has yielded relatively promising results, we believe there is still room for further improvement and optimization. In real production environments, the complexity of the data is far greater than in controlled laboratory settings, and may involve more noise, occlusions, and other non-ideal conditions. Therefore, future research should place greater emphasis on how to effectively deploy this model in actual production lines to address the unforeseen challenges that arise during manufacturing. Moreover, long-term operational validation is crucial. Only through a comprehensive evaluation of the model's performance over extended periods and large-scale operations can we better ensure its stability and reliability, thereby achieving its intended effects in real-world production environments.

In addition, with the continuous advancement of object detection technologies, we plan to explore other lightweight object detection models in future work, alongside further optimization of the existing models. These lightweight models typically require fewer computational resources and offer higher operational efficiency, making them more suitable for resource-constrained industrial environments. Through comparative studies, we aim to evaluate the performance of these lightweight models versus the approach proposed in this study across different industrial scenarios, considering factors such as detection accuracy, computational efficiency, and hardware resource requirements. This will not only help enhance the applicability and practicality of copper nodular detection technologies but could also provide valuable insights for optimizing other deep learning applications in industrial settings.

7. Conclusions

To address the issues related to the manual identification of nodulation on the surface of copper cathode plates, characterized by low efficiency, high workload, and inaccuracy, this study presents a lightweight YOLOv5 model integrated with the BiFormer attention mechanism. The MobileNetV3 module significantly reduces the network parameters, achieving a lightweight design. The BiFormer attention mechanism introduced in the neck network significantly enhances the model's feature extraction capability and target localization accuracy. Additionally, the F-EIOU loss function has replaced the CIOU loss function, significantly improving the bounding box regression performance through its unique focal mechanism and the integration of various geometric information. The experimental results indicate that, although the improved YOLOv5 model exhibits slightly lower precision than the original YOLOv5 model, the recall and mean average precision values have improved by 1% and 0.4%, respectively. In addition, compared to the original YOLOv5 model, the improved model's weight file size has been reduced by 78.6%, the number of parameters has decreased by 78.1%, and the detection speed has improved by 34%, demonstrating outstanding detection performance. The overall performance of this model exceeds that of the original YOLOv5 model, achieving a balance between lightweight

design and accuracy, thereby providing valuable references for model deployment. To verify the performance of the proposed model in actual industrial production, a cathode copper plate surface nodule defect detection system was developed using PyQt5. This system features real-time detection capabilities, flexibly meeting the needs of nodule defect detection in production environments and confirming the industrial application value of the proposed model.

Author Contributions: Conceptualization, Z.Z. and X.H.; resources, X.H., D.W., Q.C. and J.L.; data curation, D.W. and Q.C.; writing—original draft preparation, Z.Z.; writing—review and editing, Z.Z., X.H. and Q.J.; supervision, Q.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Due to the nature of this research, participants of this study did not agree for their data to be shared publicly, so supporting data is not available.

Acknowledgments: Thank you to our researchers for their collaboration and technical assistance.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Wu, Q. Present Situation and Trend of Copper Industry in China. *Nonferrous Met. Process.* **2018**, *47*, 1–6. [CrossRef]
2. Schlesinger, M.E.; Sole, K.C.; Davenport, W.G. *"Electrolytic Refining" in Extractive Metallurgy of Copper*, 6th ed.; Elsevier Ltd.: Oxford, UK, 2021; pp. 331–359.
3. Chen, S. The Reason for Nodules Forming on the Bottom of Cathode Copper and Countermeasures. *Copper Eng.* **2022**, *39*, 42–45. [CrossRef]
4. Fang, Y.; Pan, M.; Huang, H.; Shao, Y.; He, Y.; Chen, B.; Guo, Z. Current situation and prospect of additives in copper electrolysis deposition process. *Min. Metall.* **2021**, *30*, 61–69. [CrossRef]
5. Meng, Y.; Liu, H.; Li, C. Cathodic current change and nodulation morphology during short circuit of copper electrolysis. *Chin. J. Nonferrous Met.* **2022**, *32*, 262–270. [CrossRef]
6. Xu, Y.; Wang, D.; Liu, H.; Yu, H. Intelligent recognition method for surface defects of cold rolling copper strip. *J. Nonferrous Met.* **2022**, *32*, 2950–2964. [CrossRef]
7. Yang, X.; Tang, D.; Cai, B.; Yuan, H.; Pan, X.; Chang, Y.; Zhang, C. Image threshold segmentation method for surface nodulation detection of cathode copper plate based on IGWO. *J. Shaanxi Univ. Technol. Nat. Sci. Ed.* **2023**, *39*, 31–39. [CrossRef]
8. Lee, E.T.; Fan, Z.; Burak, S. A new approach to detect surface defects from 3D point cloud data with surface normal Gabor filter (SNGF). *J. Manuf. Process.* **2023**, *92* (Suppl. C), 196–205. [CrossRef]
9. Cui, Y.; Zhao, L.; Li, H.; Liu, H. Strip surface defect detection based on wavelet denoising and improved Canny algorithm. *Mod. Electron. Tech.* **2024**, *47*, 148–152. [CrossRef]
10. Chen, H.; Zhang, Z.; Yin, W.; Zhou, G.; Wang, L.; Li, Y.; Zhao, C.; Wang, C. Shape characterization and depth recognition of metal cracks based on laser infrared thermography and machine learning. *Expert Syst. Appl.* **2024**, *238*, 122083. [CrossRef]
11. Xie, W.; Ma, W.; Sun, X. An efficient re-parameterization feature pyramid network on YOLOv8 to the detection of steel surface defect. *Neurocomputing* **2024**, *614*, 128775. [CrossRef]
12. Zhang, Y.; Feng, T.; Song, Y.; Shi, Y.; Cai, G. An Improved Target Network Model for Rail Surface Defect Detection. *Appl. Sci.* **2024**, *14*, 6467. [CrossRef]
13. Cao, Y.; Zhu, W.; Yang, J.; Fu, G.; Lin, D.; Cao, Y. An effective industrial defect classification method under the few-shot setting via two-stream training. *Opt. Lasers Eng.* **2023**, *161* (Suppl. C), 107294. [CrossRef]
14. Li, W.; Zhang, S.; Zhang, T.; Shen, Y.; Han, L.; Peng, Z.; Xie, Z.; Zhong, C.; Jia, S. Surface Defect Detection and Evaluation Method of Large Wind Turbine Blades Based on an Improved Deeplabv3+ Deep Learning Model. *Struct. Durab. Health Monit.* **2024**, *18*, 553–575. [CrossRef]
15. Zhang, H.; Fu, W.; Wang, X.; Li, D.; Zhu, D.; Su, X. An efficient model for metal surface defect detection based on attention mechanism and multi-scale feature. *J. Supercomput.* **2024**, *81*, 40. [CrossRef]
16. Cardellicchio, A.; Nitti, M.; Patruno, C.; Mosca, N.; di Summa, M.; Stella, E.; Renò, V. Automatic quality control of aluminium parts welds based on 3D data and artificial intelligence. *J. Intell. Manuf.* **2023**, *35*, 1629–1648. [CrossRef]
17. Liso, A.; Cardellicchio, A.; Patruno, C.; Nitti, M.; Ardino, P.; Stella, E.; Renò, V. A Review of Deep Learning-Based Anomaly Detection Strategies in Industry 4.0 Focused on Application Fields, Sensing Equipment, and Algorithms. *IEEE Access* **2024**, *12*, 93911–93923. [CrossRef]
18. Jiang, F.; Li, N.; Qiu, C.; Liu, D. Defect detection method of wire mesh based on improved Faster RCNN. *China Sciencepap.* **2024**, *19*, 153–159. [CrossRef]

19. Li, H.; Chen, Z.; Xiu, H.; Huang, X. Improved YOLO steel surface defect detection algorithm. *Mod. Electron. Tech.* **2024**, *47*, 7–14. [CrossRef]
20. Girshick, R.; Donahue, J.; Darrell, T.; Jitendra, M. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. In Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Columbus, OH, USA, 23–28 June 2014; pp. 580–587. [CrossRef]
21. Wu, J.; Wang, J.; Fu, M.; Wang, Z.; Lu, Y. Determination of defect-testing for steel strips based on improved faster-RCNN algorithm. *Angang Technol.* **2022**, *23*, 23–28. [CrossRef]
22. Yang, F.; Huo, J.; Cheng, Z.; Chen, H.; Shi, Y. An improved mask R-CNN micro-crack detection model for the surface of metal structural parts. *Sensors* **2024**, *24*, 62. [CrossRef]
23. Pang, J.; Chen, K.; Shi, J.; Feng, H.; Ouyang, W.; Lin, D. Libra r-cnn: Towards balanced learning for object detection. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019.
24. Cai, Z.; Vasconcelos, N. Cascade r-cnn: Delving into high quality object detection. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake, UT, USA, 18–23 June 2018; pp. 6154–6162.
25. Wang, H.; Han, W.; Wang, K. Metal surface defect detection algorithm based on YOLOX. *Comput. Eng. Des.* **2023**, *44*, 2803–2810. [CrossRef]
26. Wang, C.; Yamg, S.; Zhou, L.; Hua, B.; Wang, S.; Lv, J. Research on metal gear end-face defect detection method based on adaptive multi-scale feature fusion network. *J. Electron. Meas. Instrum.* **2023**, *37*, 153–163. [CrossRef]
27. Yang, S.; Ding, F.; Wen, H.; Li, P.; Hu, S. Hot-rolled steel strip surface defects detection based on CAYOLOv5. *J. Optoelectron. Laser* **2024**, *35*, 21–28. [CrossRef]
28. Wang, H.; Liu, H.; Shao, Y. An Improved YOLOv5 Algorithm for Metal Surface Defect Detection. *Mech. Sci. Technol. Aerosp. Eng.* **2023**, *43*, 1–6. [CrossRef]
29. Song, X.; Cao, S.; Zhang, J.; Hou, Z. Steel Surface Defect Detection Algorithm Based on YOLOv8. *Electronics* **2024**, *13*, 988. [CrossRef]
30. Xie, W.; Sun, X.; Ma, W. A light weight multi-scale feature fusion steel surface defect detection model based onYOLOv8. *Meas. Sci. Technol.* **2024**, *35*, 055017. [CrossRef]
31. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788.
32. Lin, T.Y.; Dollár, P.; Girshick, R.; He, K.; Hariharan, B.; Belongie, S. Feature Pyramid Networks for Object Detection. *arXiv* **2017**, arXiv:1612.03144.
33. Li, H.; Xiong, P.; An, J.; Wang, L. Pyramid Attention Network for Semantic Segmentation. *arXiv* **2018**, arXiv:1805.10180.
34. Howard, A.; Sandler, M.; Chu, G.; Chen, L.-C.; Chen, B.; Tan, M.; Wang, W.; Zhu, Y.; Pang, R.; Vasudevan, V. Searching for MobileNetV3. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019.
35. Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; Chen, L.C. Mobile-NetV2: Inverted residuals and linear bottlenecks. *arXiv* **2018**, arXiv:1801.04381.
36. Hu, J.; Shen, L.; Sun, G. Squeeze-and-excitation networks. In Proceedings of the 2018 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPR), Salt Lake City, UT, USA, 18–22 June 2018; pp. 7132–7144.
37. Qi, X.; Zhi, M. Review of Attention Mechanisms in image Processing. *J. Front. Comput. Sci. Technol.* **2024**, *18*, 345–362. [CrossRef]
38. Zhu, L.; Wang, X.; Ke, Z.; Zhang, W.; Lau, R. BiFormer: Vision Transformer with Bi-Level Routing Attention. *arXiv* **2023**, arXiv:2303.08810.
39. Liu, X.; Yang, X.; Chen, Y.; Zhao, S. Object detection method based on CloU improved bounding box loss function. *Chin. J. Liq. Cryst. Dis.* **2023**, *38*, 656–665. [CrossRef]
40. Naik, T.B.; Hashmi, M.F.; Gupta, A. EIoU-distance loss: An automated team-wise player detection and tracking with jersey colour recognition in soccer. *Connect. Sci.* **2024**, *36*, 2291991. [CrossRef]
41. Zhang, Y.-F.; Ren, W.; Zhang, Z.; Jia, Z.; Wang, L.; Tan, T. Focal and efficient IOU loss for accurate bounding box regression. *Neurocomputing* **2022**, *506*, 146–157. [CrossRef]