

MO-SOD: Micro-Oxidation Small Object Detection Model for Oxygen-Free Copper Surfaces Based on Microscopic Imaging System

Qianqian Li, Taohong Zhang,* Mingyang Yang, Xuxu Guo, Han Chen, and Suli Fan



Cite This: *ACS Omega* 2023, 8, 6608–6620



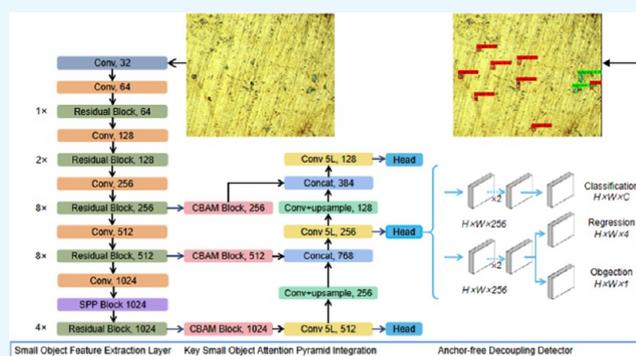
Read Online

ACCESS |

Metrics & More

Article Recommendations

ABSTRACT: Micro-oxidation is a fatal problem for some precision oxygen-free copper materials, and it is difficult to detect with the naked eyes. However, manual inspection using microscope equipment is expensive, subjective, and time-consuming. The automatic high-definition micrograph system equipped with micro-oxidation detection algorithm can detect more quickly, efficiently, and accurately. In this study, a micro-oxidation small object detection model, MO-SOD, is proposed to detect the oxidation degree on oxygen-free copper surface based on microimaging system. This model is developed for rapid detection on the robot platform combined with high-definition microphotography system. The proposed MO-SOD model consists of three modules: small target feature extraction layer, key small object attention pyramid integration layer, and anchor-free decoupling detector. The small object feature extraction layer focuses on the local features of small object to improve the perception of micro-oxidation spots and also takes the global features into account to reduce the impact of noisy background on feature extraction. Key small object attention pyramid integration block couples key small object feature attention and pyramid to detect the micro-oxidation spots in the image. The performance of MO-SOD model is further improved by combining the anchor-free decoupling detector. In addition, the loss function is improved to combine CIOU loss and focal loss to achieve effective micro-oxidation detection. The MO-SOD model is trained and tested from three oxidation levels in an oxygen-free copper surface microscope image data set. The test results show that the average accuracy (mAP) of MO-SOD model is 82.96%, which is superior to other most advanced detectors.



1. INTRODUCTION

Oxygen-free copper is widely used in aerospace and electronic devices because of its high electrical conductivity, high thermal conductivity, good elasticity, corrosion resistance, nonmagnetic, low hydrogen permeability, easy machining, and low cost.¹ However, the reaction of oxygen-free copper with oxygen and water vapor in the environment will produce micro-oxidation, which is fatal to some precision instruments using oxygen-free copper, and the resulting damage or scrap will cause huge economic losses and even cause personnel casualties.^{2,3} Therefore, detecting the degree of micro-oxidation on the surface of oxygen-free copper and predicting it in advance is of great significance for taking effective protective measures in a timely and reasonable manner.⁴

Micro-oxidation is a fatal problem for oxygen-free copper materials, but it is difficult to detect micro-oxidation with the naked eye. Visual inspection using microscope equipment is one of the more widely used inspection techniques. However, the inspector's work is intensive and long-term inspection can affect the inspector's working condition, resulting in low inspection

efficiency, low inspection accuracy, and problems with false and missed inspections.

Before the increase of machine vision, visual inspection could not be adapted to efficient production and nondestructive testing (NDT) is widely used. NDT methods used for oxidation corrosion inspection include eddy current, magnetic leakage, and infrared inspection. Among them, eddy current detection and magnetic leakage detection are prone to detection errors when detecting rough surfaces. Infrared detection due to its more restrictive conditions is generally used for a small range of offline detection. For intelligent manufacturing, the above methods cannot satisfy the automation requirements. Infrared detection is generally only used for offline detection in a small

Received: November 1, 2022

Accepted: February 2, 2023

Published: February 13, 2023



range due to its many limited conditions, which cannot meet the automation requirements of the intelligent manufacturing industry.

Machine vision inspection is a more desirable method of oxidation detection, which processes the images captured by the camera by means of image processing algorithms. Traditional machine vision-based inspection methods consist of two main aspects, one is based on image processing methods and the other is based on manual feature machine learning methods. The image-processing-based approach uses a CCD HD camera to acquire images of copper surface morphology. Shen, Chen, and Chang⁵ use RGB color space to collect oxidation chromatograms and define oxidation color ranges for oxidation identification. Manual feature-based machine learning methods use algorithms such as SIFT,⁶ Canny,⁷ HOG,⁸ LBP,⁹ and GLCM¹⁰ for feature extraction and then train an SVM¹¹ to determine if the input image has oxidation. Son, Hwang, and Kim¹² propose a method for determining the surface area of rust on steel bridges, applying color space transformation, and rust area classification based on the J48 decision tree algorithm. These machine vision-based oxidation inspection methods improve the efficiency of inspection and meet the requirements of automation to a certain extent but are less adaptable and the hand-made feature vectors do not adequately represent the information for complex environments. In addition, classifier inspection cannot accurately locate the oxidation location.

CNN can automatically extract the features of an image, solving the problem of time-consuming and laborious manual feature extraction. Because of their high accuracy, efficiency, and adaptability, deep learning algorithms are gradually becoming an important research direction in oxidation detection. Li, Lin, and Chen¹³ propose an adaptive multithreshold spot area adaptive calibration algorithm to detect small rust spots on ships, and the detection rate and range of this method are better than traditional methods. Liu, Xu, and Xu¹⁴ proposed a convolution neural network-based surface defect detection system with markers for the classification and detection of steel plate surface defects with small samples. Yao, Yang, Wang, and Zhao¹⁵ proposed a convolutional neural network-based method for the detection and identification of corrosion damage on hull structural panels, using a classifier together with an overlapping scan sliding window algorithm to identify and locate the location of corrosion damage. Bastian, Jaspreeth, and Ranjith¹⁶ designed a special convolutional neural network to classify images of water, oil, and gas pipelines based on the degree of corrosion of the pipeline that came to be detected. Papamarkou, Guy, and Kroencke¹⁷ proposed a deep learning-based image classification algorithm for detecting corrosion in stainless steel tanks, including discoloration, pitting, and stress corrosion cracks. Zhang, Deng, and Lu¹⁸ proposed a channel-attention-based metal corrosion detection method (CAMCD) that can automatically and intelligently detect areas with several different levels of corrosion. Forkan, Kang, and Jayaraman¹⁹ developed CorrDetector based on convolutional neural networks for structure identification and corrosion feature extraction. These works achieve good results in their corresponding scenarios, but their focus is on corrosion detection of large areas, such as rust defect identification, and did not address micro-oxidation detection. Micro-oxidation detection shall be taken with a microscopic imaging system. On this basis, dust particles on the surface of oxygen-free copper will generate noise on the surface image. In addition, small indentation and scratch will also cause great impact, resulting in an extremely noisy background on the

surface of oxygen-free copper photographed by microscopy and difficult identification.

Although some existing work has made some achievements in corrosion detection, there are still the following research problems. Oxidation detection is mostly aimed at corrosion detection of large areas, such as rust defect identification, and does not involve micro-oxidation detection. At present, the feature extraction and detection algorithms used in the object detection model for oxidation detection are less robust and accurate.

In this study, a micro-oxidation small object detection model of oxygen-free copper surface based on microscopic imaging system is proposed for rapid detection on a robot platform combined with high-definition microsystem. The main contributions of this paper are as follows:

1. A new type of micro-oxidation small object detection model for oxygen-free copper surface based on microscopic imaging system, MO-SOD model, which can realize high-precision detection on the existing robot computer. MO-SOD model uses three modules to improve the accuracy of micro-oxidation spot detection in noisy oxygen-free copper background.
2. The small object feature extraction layer focuses on the local features of small object to improve the perception of micro-oxidation spots while taking into account the global features to reduce the impact of noisy background on feature extraction. MO-SOD model embeds the key small object attention pyramid integration to effectively focus on small object features, remove redundancy, reduce parameters, and reduce computation.
3. The performance of the detector is further improved by combining the anchor-free decoupling detector. In addition, the loss function is improved to combine CIOU loss and focal loss to achieve effective micro-oxidation detection.
4. The effectiveness of the proposed MO-SOD model in micro-oxidation detection accuracy and performance is evaluated through comprehensive validation experiments, analysis, and comparison.

The organization of this study is as follows. In [Section 2](#), a comprehensive overview of the literature on corrosion testing is presented. In [Section 3](#), details of the proposed MO-SOD model are presented. [Section 4](#) presents the experimental data, experimental environment, and results. And the conclusion of this article is summarized in [Section 5](#).

2. MICRO-OXIDATION SMALL OBJECT DETECTION METHOD

2.1. Motivation. Considering the high similarity in shape, color, and texture of micro-oxidation spots at different levels on the surface of oxygen-free copper, we use yolo series as the micro-oxidation detection tool. Among them, Yolov4²⁰ and Yolov5²¹ may be a little overoptimized for anchor-based pipelines, and we chose Yolov3²² as our starting point. In fact, Yolov3 remains one of the most widely used detectors in the industry due to factors such as computational resource limitations and lack of software support in various practical applications. To improve the accuracy and efficiency of oxygen-free copper surface micro-oxidation detection, we propose a micro-oxidation small object detection model of oxygen-free copper surface based on microscopic imaging system, MO-SOD model, which is used for rapid detection on the robot platform

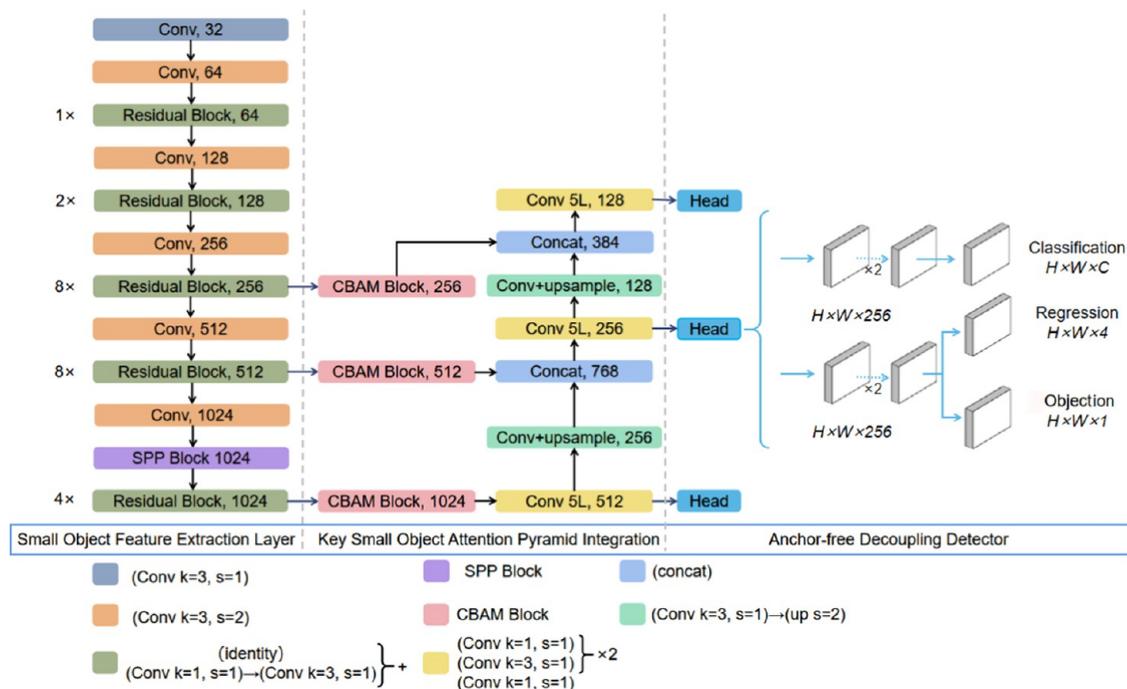


Figure 1. Structure of the proposed MO-SOD model.

combined with a high-definition microphotography system. The overall structure is shown in Figure 1.

2.2. Structure of the Proposed MO-SOD Model. The MO-SOD model proposed in this paper consists of three parts: small object feature extraction layer, key small object attention pyramid integration layer, and anchor-free decoupling detector. The small object feature extraction layer focuses on the local features of small object when extracting image features to improve the perception of micro-oxidation spots. At the same time, the extraction layer takes the global features into account to reduce the impact of noisy background on feature extraction. And we propose a key small object attention pyramid integration block, where the key small object attention is coupled with the pyramid. An anchor-free decoupling detector is introduced to perform micro-oxidation spot detection on the extracted small object features, output the position and category of the detected object, and improve the detection accuracy and efficiency of small objects in a cluttered background.

2.2.1. Small Object Feature Extraction Layer. The size and shape of different micro-oxidation spots vary greatly. Moreover, the micro-oxidation spot background photographed by microscope contains a lot of noise, such as indentation and scratch on the background. The small object feature extraction layer adopts Darknet53 as the basis to perform feature extraction on images, including six traditional convolution layers and five residual convolution layers, as shown in Figure 1. In particular, the small object feature extraction layer focuses on the local features of small object to improve the perception of micro-oxidation spots while taking into account the global features to reduce the impact of noisy background on feature extraction. Small object feature extraction layer adds spatial pyramid pooling at the near end of the feature extraction layer. By maximizing the pool size of different pool cores, feature extraction is performed to improve the receptive field of the network.

The spatial pyramid pooling structure is mixed in the convolution of the last feature layer of the feature extraction network. After three times of convolution of the last feature layer

of the feature extraction network, three different scales of maximum pooling are, respectively, used for processing

$$G(x) = \text{concat}(P^1(x), P^5(x), P^7(x), P^{11}(x)) \quad (1)$$

where x represents the output of the convolution layer, and $P^i(x)$ represents that the pooled check x of size i is used to maximize pooling, and finally the pooled tensor concat is used as the output. Here, the maximum pooling cores i are 11×11 , 7×7 , 5×5 , and 1×1 (1×1 means no processing). The step size of each pooling layer is 1. Finally, the size and depth of the feature map obtained after pooling remain unchanged.

The spatial pyramid pooling module uses pooling of different sizes to achieve the fusion of features at different scales, which can greatly increase the receptive field, separate the most significant context features, extract and fuse local area features. It focuses on the local features of small objects to improve the perception of micro-oxidation spots and considers the global features to reduce the impact of noise background on feature extraction. This method can effectively extract and fuse local area features, so it is more suitable for small target detection in clutter background.

2.2.2. Key Small Object Attention Pyramid Integration. The feature fusion layer embeds the key small object attention pyramid integration, which makes the channel attention and spatial attention for pyramid fusion. Channels focus on global information, while spatial attention is used locally. In this way, comprehensive salient features can be extracted to improve the detection performance of small objects in the clutter background. After attention mechanism processing, multiscale key small object features are fused with the pyramid model for independent detection to improve the detection effect of small objects.

The key small object attention module is computed as

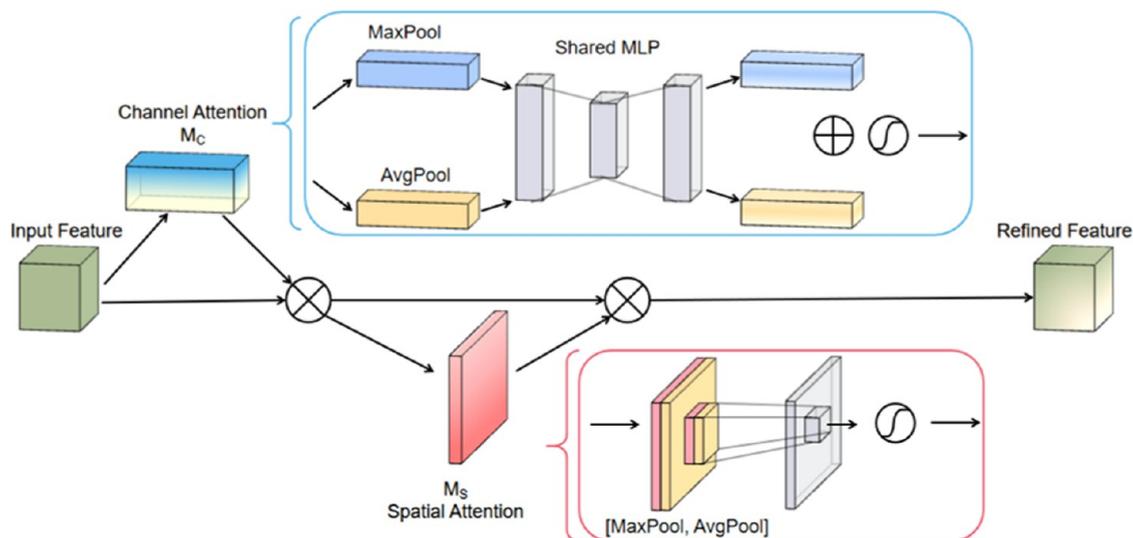


Figure 2. Key small object attention module.

$$\begin{aligned}
 F(f) &= F_s(F_c(f)) \\
 F_c(f) &= M_c(f) \otimes f \\
 F_s(f) &= M_s(f) \otimes f
 \end{aligned} \quad (2)$$

where input feature map $f \in R^{C \times W \times H}$, F_c represents channel attention, and F_s represents spatial attention. As shown in Figure 2, the input feature map passes through the channel attention module and the spatial attention module, in turn. $M_c(f)$ means that the input feature map obtains a one-dimensional attention weight in the channel attention, and $M_s(f)$ means that the input feature map obtains a two-dimensional attention weight in the spatial attention, where \otimes denotes element-wise multiplication

$$\begin{aligned}
 M_c(f) &= \sigma(\text{MLP}(\text{MaxPool}(f)) + \text{MLP}(\text{AvgPool}(f))) \\
 M_s(f) &= \sigma(\text{conv}(\text{concat}(\text{MaxPool}(f), \text{AvgPool}(f))))
 \end{aligned} \quad (3)$$

where σ represents the sigmoid function. The channel attention mechanism performs global average pooling and global maximum pooling for the input feature layer of a single key small object. It uses the shared full connection layer to process and add them together. The weight of each channel in the input feature layer is obtained through the sigmoid function, that is, the channel attention weight of the input feature layer. For the input single key small object feature layer, spatial attention mechanism takes the maximum and average value on the channel of each feature point, stacks them, and then uses the convolution of channel number 1 to adjust the channel, and obtains the weight of each feature point of the input feature layer through the sigmoid function, that is, the spatial attention weight of the input feature layer.

In the feature fusion layer, attention is only introduced to key small object features, instead of concatenating attention mechanisms or serializing attention in the whole feature extraction process. As shown in Figure 3, attention mechanism does not participate in the network structure of feature extraction layer and only strengthens and fuses the features of key small objects, which can remove redundancy, reduce unnecessary parameters, and reduce the amount of computation.

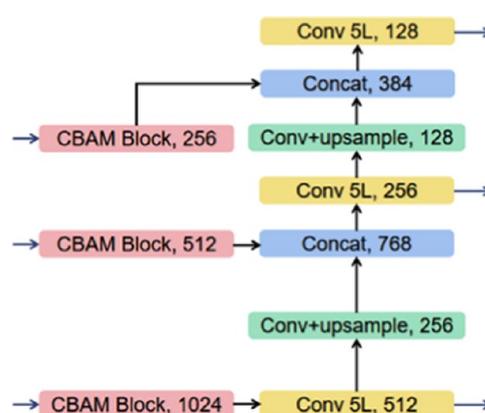


Figure 3. Key small object attention pyramid integration.

2.2.3. Anchor-Free Decoupling Detector. The research shows that the classification and regression can be implemented by using a 1×1 convolution, which will bring adverse effects to the network identification. In the proposed MO-SOD model, an anchor-free decoupling detector is introduced to detect the features from the key small object attention pyramid integration, thus improving the accuracy and efficiency of small target detection in the clutter background. The detector is divided into two parts, which are implemented separately and integrated into the final prediction.

As shown in Figure 4, it contains a 1×1 conv layer to reduce channel dimensions; two parallel branches of 3×3 conv layer

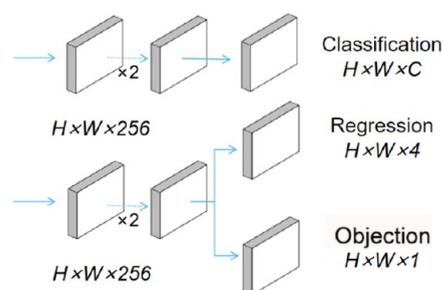


Figure 4. Anchor-free decoupling detector.

are classified and regressed, respectively. Finally, three prediction results are obtained, where regression prediction is used to judge the regression parameters of each feature point, object prediction is used to judge whether each feature point contains objects, and class prediction is used to judge the class of the object contained in each feature point.

2.3. Loss Function. To reduce training time and model hyperparameters, SimOTA dynamically matches positive samples, calculates pairwise matching, i.e., the cost relationship between each ground truth box and each feature point, and selects the top k predictions with the lowest cost in a fixed central region as its positive samples. Finally, the corresponding grids of these positive predictions are classified as positive and the remaining grids are classified as negative. Corresponding to the predictions, the network loss consists of three components, namely, reg-loss, obj-loss, and cls-loss.

Reg-loss is the regression parameter judgment of the feature points. The CIOU takes into account the distance between the ground truth boxes and the predicted boxes, the overlap rate, the scale, and the penalty term, making the predicted box regression more stable and allowing for better convergence speed and accuracy. Reg-loss is calculated as follows

$$L_{\text{reg}} = \sum_{i=1}^n L_{\text{CIOU}} \quad (4)$$

$$L_{\text{CIOU}} = 1 - \text{IOU} + \frac{\rho^2(b, b^{\text{gt}})}{c^2} + \alpha v \quad (5)$$

$$v = \frac{4}{\pi^2} \left(\tan^{-1} \frac{w^{\text{gt}}}{h^{\text{gt}}} - \tan^{-1} \frac{w}{h} \right)^2 \quad (6)$$

$$\alpha_1 = \frac{v}{1 - \text{IOU} + v} \quad (7)$$

where n represents the number of positive samples obtained after dynamic matching of positive samples by SimOTA. $\rho^2(b, b^{\text{gt}})$ represents the Euclidean distance between the centroids of the prediction boxes and the ground truth boxes. c represents the diagonal distance of the smallest closed region that can contain both the prediction boxes and the ground truth boxes. w , w^{gt} , h , and h^{gt} represent the width and height of the predicted and real boxes.

Obj-loss is a judgment of whether a feature point contains an object, as the number of potential bounding boxes containing objects is much smaller than those containing only background. Cross-entropy loss is slow and may not be optimized to optimality during iterations with a large number of simple samples. In this paper, we introduce focal loss in the model, which gives the object a high loss value, which makes the detector sensitive to the object. The focal loss is calculated as

$$F(p, y) = \begin{cases} -\alpha_2(1-p)^\gamma \log(p) & \text{if } y = 1 \\ -(1-\alpha)p^\gamma \log(1-p) & \text{otherwise} \end{cases} \quad (8)$$

In the above, $y \in \{\pm 1\}$ specifies the ground truth class, and $p \in [0, 1]$ is the estimated probability of the model for the class labeled $y = 1$. α_2 is a balancing factor to balance the proportion of positive and negative samples themselves in terms of number. γ represents the exponential scale factor, which reduces the loss of easy-to-classify samples and allows the MO-SOD model to focus more on difficult, misclassified samples. In this paper, α_2 and γ

are 0.25 and 2, respectively. Based on the above formula, Obj-loss can be converted to

$$L_{\text{obj}} = -\alpha_2 \sum_{i=1}^N 1_i^{\text{obj}} (1-p)^\gamma \log(p) \quad (9)$$

$$L_{\text{no obj}} = -(1-\alpha_2) \sum_{i=1}^N 1_i^{\text{no obj}} p^\gamma \log(1-p) \quad (10)$$

where N represents the total number of prediction boxes, the last feature extracted by the backbone network, and each square represents a prediction box. 1_i^{obj} indicates that the i th prediction box contains an object. Also, $1_i^{\text{no obj}}$ means that the prediction box only contains the background.

Cls-loss is a judgment of the kind of objects contained in the feature points, and the cross-entropy loss is calculated based on the kind of the ground truth boxes and the kind of feature points predicted as a component of the loss of the Cls part

$$L_{\text{cls}} = - \sum_{i=1}^n \sum_{c \in \text{classes}} [\hat{p}_i(c) \log(p_i(c)) + (1 - \hat{p}_i(c)) \log(1 - p_i(c))] \quad (11)$$

where n is the number of positive samples obtained after dynamic matching of positive samples by SimOTA. The prediction of classes is represented by $p(c)$. The variables with $\hat{}$ indicate that they are predicted values; otherwise, they indicate ground truth.

The loss function of the proposed MO-SOD model is expressed as

$$\text{LOSS} = w_{\text{reg}} L_{\text{reg}} + L_{\text{obj}} + L_{\text{no obj}} + L_{\text{cls}} \quad (12)$$

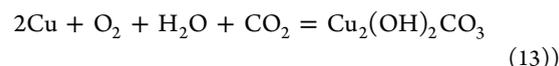
where w_{reg} is the weight of the reg-loss, which has a value of 5 in this paper.

3. EXPERIMENTS

3.1. Data Set. To verify the applicability of the proposed MO-SOD model in different chemical environments, we executed the chemical reactions of oxygen-free copper oxidation in two different environments, namely, the oxidation of oxygen-free copper in a normal atmospheric environment and the oxidation of oxygen-free copper in a high-temperature atmospheric environment. We collected the surface images of oxygen-free copper in different oxidation environments through the microscope shooting system for the experimental analysis of oxygen-free copper oxidation detection:

- (1) Oxidation in a neutral atmospheric environment at room temperature.

We put nine new oxygen-free copper samples in the neutral atmospheric environment at room temperature for 1 month. The oxygen-free copper surface slowly oxidizes with CO_2 , O_2 , and H_2O to produce basic copper carbonate $\text{CuCO}_3 \cdot \text{Cu}(\text{OH})_2$, commonly known as copper green



- (2) Oxidation in a high-temperature neutral atmosphere environment.

Similar to the above, we also take nine new oxygen-free copper samples for oxidation in the high-temperature neutral atmosphere. The nine new oxygen-free copper samples are put

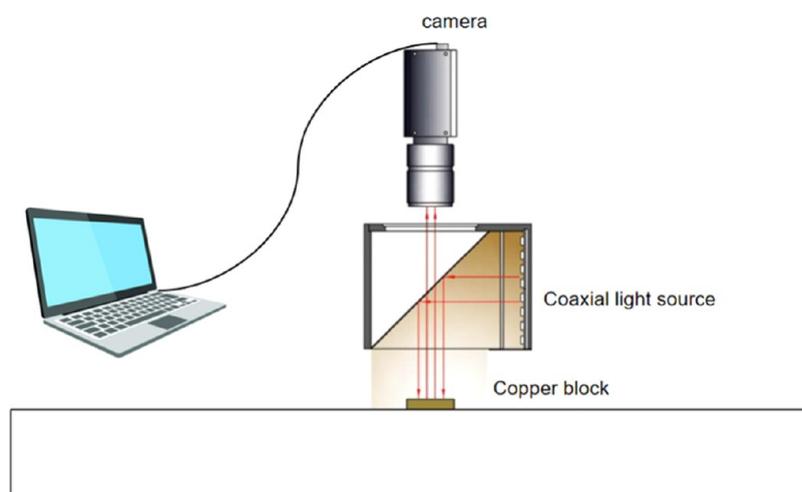


Figure 5. Experimental filming environment.

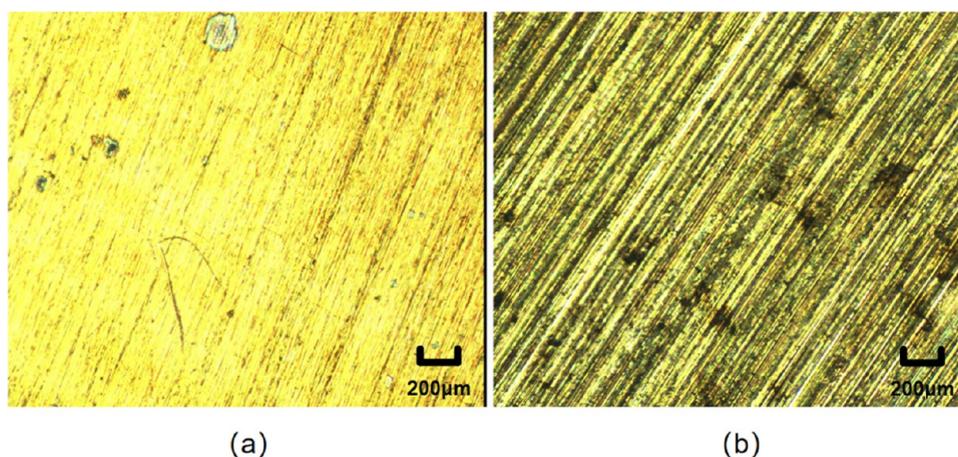


Figure 6. Image of oxygen-free copper surface taken by high-definition microscope equipment. Surface of oxygen-free copper sample oxidized in a neutral atmosphere at room temperature (a). Surface of oxygen-free copper sample oxidized in a neutral atmosphere at high temperature (b).

into a high-temperature oxidation furnace and take them out every half an hour to generate different oxidation gradients. In the high-temperature oxidation furnace, oxygen-free copper reacts with O_2 in the high-temperature neutral atmospheric environment to form black copper oxide



To construct a data set to train and validate the proposed MO-SOD model, we use a microscopic HD filming system with white light as the experimental light source. The filming environment is shown in Figure 5. We collect the surface morphology images of oxygen-free copper sample blocks in different oxidation environments through the high-definition microphotography system. Figure 6a shows the surface image of oxygen-free copper sample oxidized in a neutral atmospheric environment at room temperature. Figure 6b shows the surface image of oxygen-free copper sample oxidized in a high-temperature neutral atmosphere.

The above micro-oxidation images of oxygen-free copper surface collected in different oxidation environments are used to calibrate the data set using the adaptive oxidation calculation method. There are 899 images in the data set from the surface of oxygen-free copper samples oxidized in a neutral atmosphere at room temperature and 1200 images from the surface of oxygen-

free copper samples oxidized in a neutral atmosphere at high temperature:

1. The oxidation spot is labeled with a detecting frame as the smallest outer rectangle of the oxidation spot.
2. Edge segmentation of the oxidation spots and removal of edge noise to calculate the oxidation spot area, where for the i th oxidation spot in a picture, its area is denoted as A_i .
3. Color calibration of the oxidation level of the oxidation spots is carried out to determine the color depth of each oxidation spot. As the degree of oxidation deepens, the shades of oxidation spots segmented from the outside to the inside are green, brown, and black; the gray scale range of the three main colors of oxidation spots is extracted, and the gray scale values of all pixels within the main color range are calculated to determine the color depth of each oxidation spot; for the i th oxidation spot in a picture, its color depth C_i is denoted as

$$C_i = \sum k_1 \mu_{GN} + k_2 \mu_{BN} + k_3 \mu_{BK} \quad (15)$$

where μ_{GN} is the green pixel gray value, μ_{BN} is the brown pixel gray value, and μ_{BK} is the black pixel gray value, and k_1 , k_2 , and k_3 are their corresponding color weight coefficients, respectively, and $k_1 = 0.4$, $k_2 = 0.2$, and $k_3 = 0.7$.

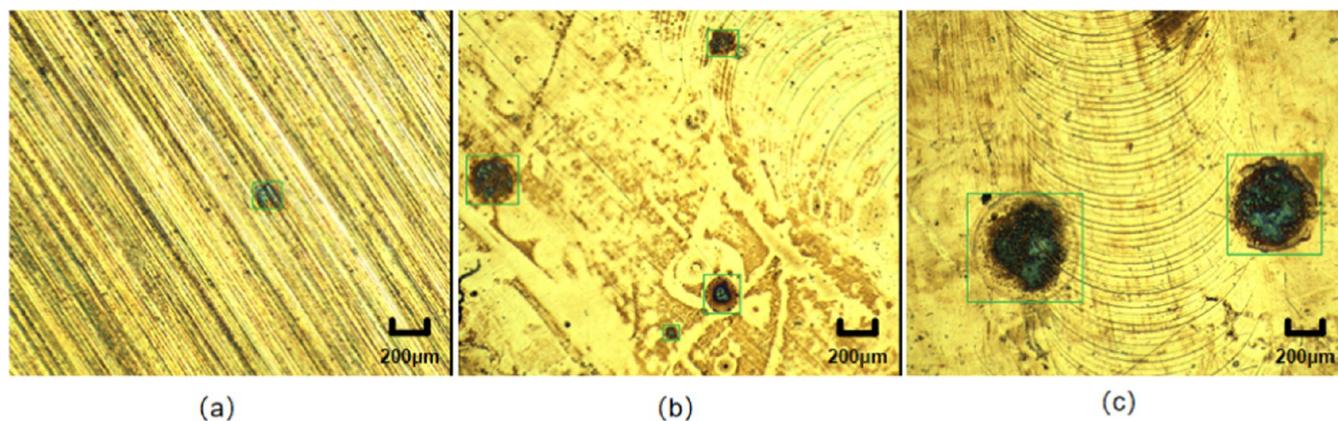


Figure 7. Labeled images. Panel (a) shows a slight oxidation spot, panel (b) shows moderate oxidation, and panel (c) shows a severe oxidation spot. The scale bar in the figures shows the 200 μm length of the sample surface.

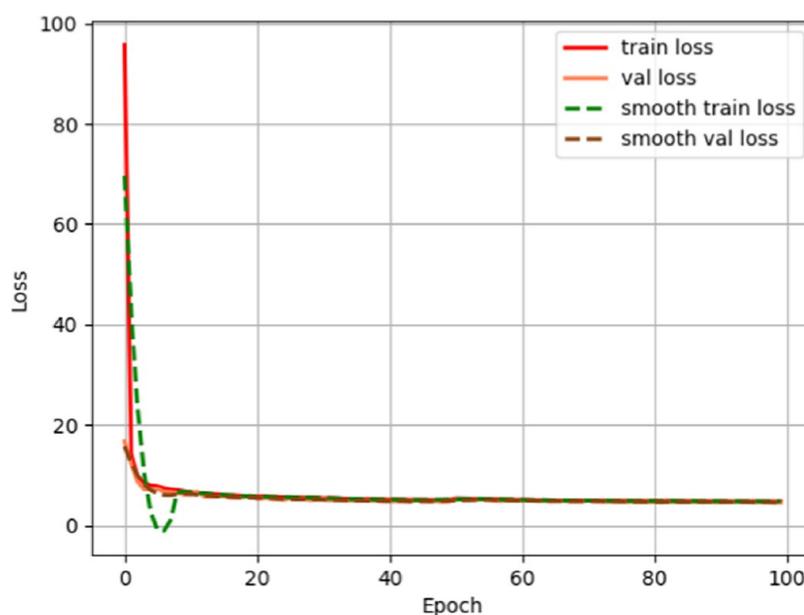


Figure 8. Loss decline curve of the proposed MO-SOD model.

- The sample is placed into an X-ray photoelectron spectroscopy analyzer, and the oxygen-to-copper ratio of the oxidized spot is determined; for the i th oxidized spot in a picture, the oxygen-to-copper ratio is recorded as R_i .
- Based on the oxygen content parameters obtained and the corresponding morphological characteristics, the adaptable oxidation calculation is fitted to the expression as

$$f(A_i^\alpha C_i^\beta) = R_i \quad (16)$$

where α and β are the oxidation coefficient constants, A_i is the oxidation spot area, C_i is the oxidation spot color depth, and R_i is the oxidation spot oxygen-to-copper ratio, i.e., the oxidation volume.

The oxygen-to-copper ratio and corresponding oxidation spot area and color depth of 100 oxidation spot samples are used to fit the oxidation amount calculation method to obtain the adaptable oxidation spot calculation method as

$$K(A_i^\alpha C_i^\beta) + b = R_i \quad (17)$$

where $K = 0.02$, $\alpha = 1$, $\beta = -1$, and $b = 0.2$.

The oxidation amount is calculated for a large number of oxidation spot samples according to the above method, and a total of 10 374 oxidation spot samples are obtained. The oxidation grade is classified according to the oxidation amount, and two thresholds are set through data analysis: $T_1 = 0.7$, $T_2 = 1.0$. These two thresholds divide it into three levels: L1: slight oxidation, L2: moderate oxidation, L3: severe oxidation. Among them, the range of the oxygen-to-copper ratio of L1 is $R_{\min} \leq R < T_1$, $R_{\min} = 0.5546$. The range of oxygen-to-copper ratio of L2 is $T_1 \leq R < T_2$. The range of oxygen-to-copper ratio of L3 is $T_2 \leq R < R_{\max}$, $R_{\max} = 3$. Classes of slight oxidation, moderate oxidation, and severe oxidation contain 7043, 2259, and 1072 samples, respectively.

When the number of input data samples is not enough, it is easy to lead to features insufficient, which is difficult for convergence of the model. The model trained by limited data samples is easy overfitting. The robustness of the model is with inadequate performance. If the new data is applied in this model, the accuracy is unsatisfactory. The oxidation classes are labeled on 2099 images, and several labeled images are shown in Figure 7. To generate the training and test sets, the data set is divided

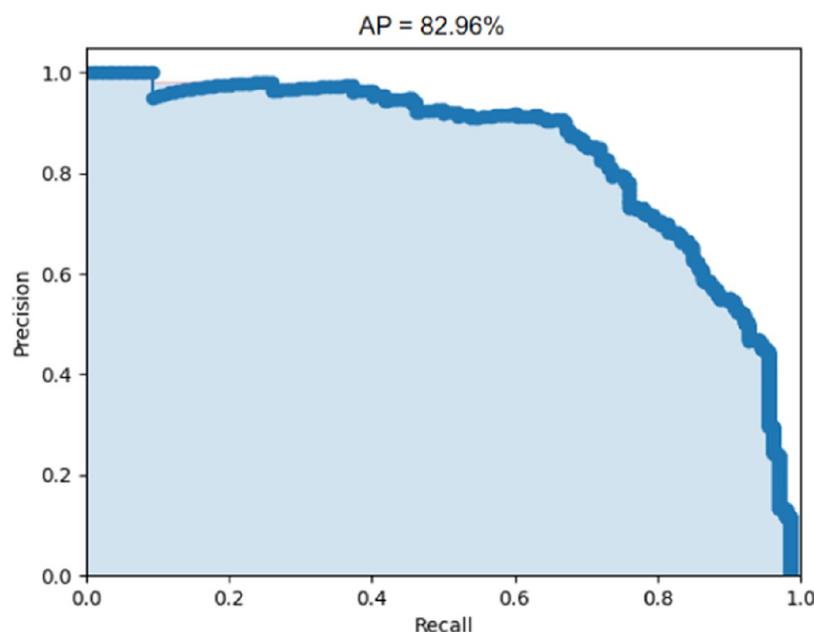


Figure 9. Pr(precision/recall) curve of the proposed MO-SOD model.

randomly and the training and validation data sets contain 1889 images. The other 210 images are used for testing.

3.2. Experimental Environment. The MO-SOD model is implemented with Pytorch 1.7.1 and CUDA 10.2.89. All cnn-based training and testing are performed on a server with an NVIDIA GeForce RTX 2080Ti graphics processing unit (GPU) equipped with an Intel Core i9-9900k @3.60 GHz CPU, 62 GB of installed memory (RAM) and 11 GB of GDDR6 RAM.

The micro-oxidation spot image is taken by a high-definition microcamera with an image resolution of 2448 pixels \times 2048 pixels, and the frame rate is 22 fps. A quadruple objective lens is applied to the front of the camera to enlarge the image, and the corresponding actual size is shown as a scale bar in the sample figures.

The loss function loss and Adam optimizer are used to iteratively update the parameters of the convolution kernel and neurons in the model. The optimizer is parameterized by 100 iterations of training. The first 50 of these are freeze training to speed up training and prevent the weights from being corrupted in the early stages of training. Each training batch size is 8, and the learning rate is set to 1×10^{-3} . The next 50 iterations are unfreeze training. Each training batch size is 4, the learning rate is 1×10^{-4} , and the momentum parameter is 0.92. The loss curve during training can be seen in Figure 8, which shows that the loss function has been optimized and converged to a stable value. The pr(precision/recall) curve obtained by testing the proposed MO-SOD model is shown in Figure 9, which shows the accuracy of the model for micro-oxidation detection.

3.3. Evaluation Metrics. To demonstrate the accuracy of the proposed MO-SOD model for oxidation spot identification, including the accuracy of location information and the accuracy of classification of oxidation class categories, it is evaluated by Precision, Recall, F1Score, AP, and mAP.

In fact, it is unlikely that the predicted results will match the ground truth exactly. Therefore, we use the intersection-over-union (IoU) metric to indicate the overlap between the predicted bounding box and the ground truth box. When the area of overlap between a suspected oxidation spot and ground truth exceeds the IoU threshold, the test result is classified as

positive; otherwise, the test result is classified as negative. In this study, the IoU value is 0.5.

The precision indicates the number of true positive (TP) results divided by all positive assays, which is shown in eq 18. Recall is defined as the percentage of TP among all correct test results, which is shown in eq 19. And F1Score is an overall measurement between the precision and the recall, which is shown in eq 20. True positive (TP) indicates the number of samples for which the prediction is positive and the actual case is also positive. False positive (FP) indicates the number of samples where the prediction is positive and the actual is negative. False negative (FN) indicates the number of samples for which the prediction is negative and the actual case is positive

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (18)$$

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (19)$$

$$\text{F1Score} = \frac{\text{recall} + \text{precision}}{\text{recall} \times \text{precision}} \quad (20)$$

The area under the precision and recall curve is known as the average precision (AP). AP indicates the detector's ability to locate objects and assign them to a single class. In general, the higher the AP of a class of objects, the better the detector's performance in identifying them. Mean accuracy (mAP) indicates the detector's performance across all categories and can be defined as the average of the AP across all categories.

$$\text{AP} = \int_0^1 \text{precision}(r) \, dr \quad (21)$$

where $\text{precision}(r)$ is the curve made with recall as the horizontal coordinate and precision as the vertical coordinate

$$\text{mAP} = \frac{\sum_{i=1}^C \text{AP}_i}{C} \quad (22)$$

where C is the total number of categories.

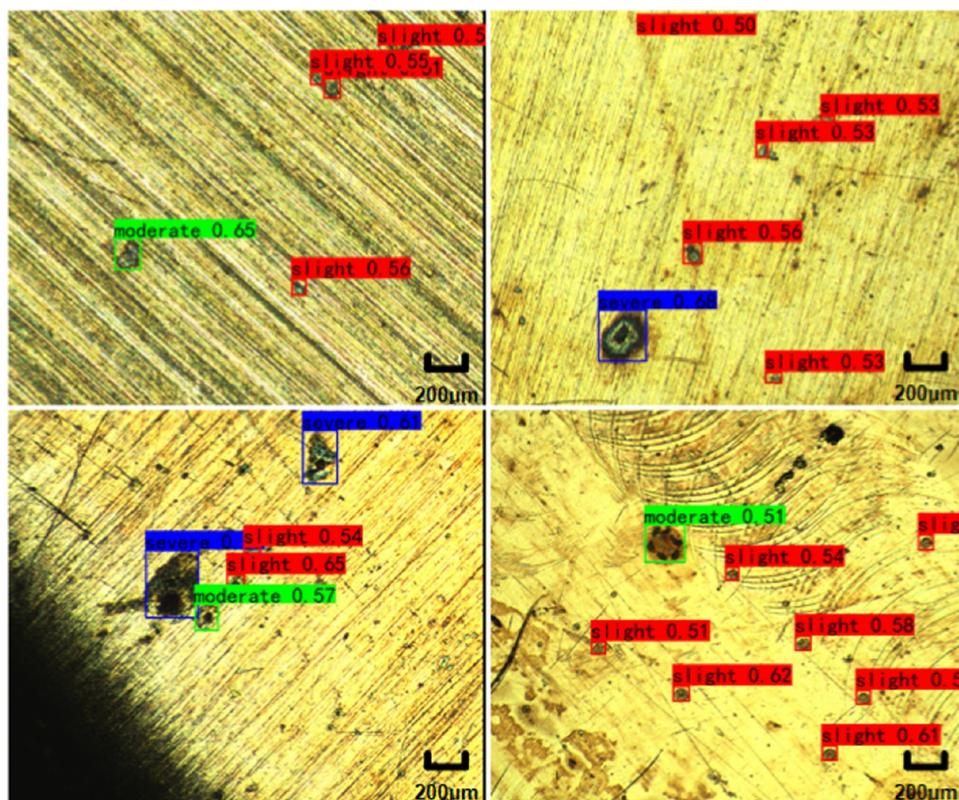


Figure 10. Detection results of the proposed MO-SOD model. The red box indicates the detected slight oxidation spot, the green box indicates the detected moderate oxidation spot, and the blue box indicates the detected severe oxidation spot. The number above the box indicates the discrimination probability of this class. The scale bar in the figure shows the 200 μm length of the sample surface.

Table 1. Detection Results of Three Kinds of Micro-Oxidation

micro-oxidation grade	number of micro-oxidation spot	AP (%)	precision (%)	recall (%)	F1Score
slight oxidation	643	77.17	82.53	60.96	0.70
moderate oxidation	204	85.81	87.13	73.04	0.79
severe oxidation	85	85.89	89.86	82.35	0.80

Table 2. Comparison of Micro-Oxidation Detection Performance

detector	backbone	size	mAP (%)	slight oxidation	moderate oxidation	severe oxidation
Yolov3	Yolov3	416	55.68	56.28	52.84	57.93
Yolov4	Yolov4	416	53.18	58.64	44.06	56.83
YoloX-s	YoloX-s	416	65.42	62.70	61.29	72.27
SSD	VGG16	300	51.04	23.01	56.06	74.04
Centernet	Resnet50	512	55.48	51.13	52.82	62.49
RetinaNet	Resnet50	600	44.38	26.08	54.52	52.54
MO-SOD model	MO-SOD model	416	82.96	77.17	85.81	85.89

4. RESULTS AND DISCUSSION

4.1. Performance of the Proposed MO-SOD Model.

Ten percent of the data set is randomly selected as the test set and input to the trained MO-SOD model to obtain information on the location of oxidation spots on the oxygen-free copper surface images and the results of the oxidation class category assessment; some of the recognition results are shown in Figure 10. Due to the presence of scratches and indentations on the copper surface and the high-definition microscope system used to capture the images, the image background is more cluttered. We can see that all three different levels of micro-oxidation spots can be detected correctly. More importantly, when an image

contains multiple levels of micro-oxidation spots, all of them can be identified.

Table 1 shows the effectiveness of the proposed MO-SOD model for the detection of three levels of micro-oxidation on copper surfaces. It can be seen that the object detection network has high accuracy in the location detection of oxidation spots and the classification of oxidation grade. Although the total number of samples for the class of severe oxidation is relatively small, due to its large oxidation area and deeper color, based on its unique characteristics, the test results show that it has high accuracy for this level of micro-oxidation. The class of slight oxidation is relatively less accurate due to its very small oxidation area and the large number of samples, which are scattered in the background, while the class of moderate oxidation is

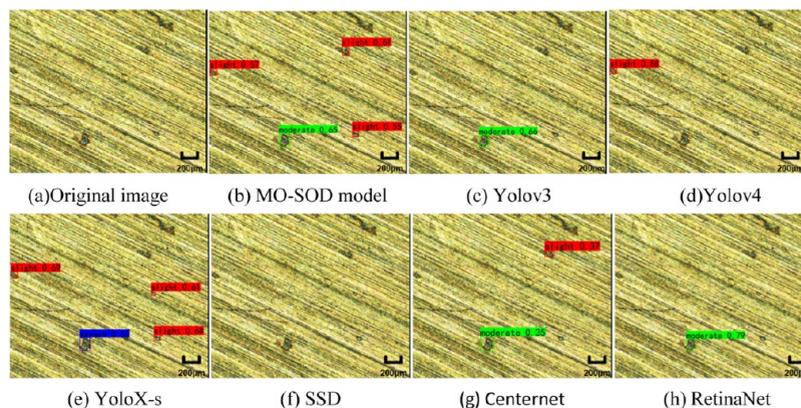


Figure 11. Comparison of the detection results among the proposed MO-SOD model with other models. Original image (a) and detection results produced by the proposed MO-SOD model (b), Yolov3 (c), Yolov4 (d), YoloX-s (e), SSD (f), Centernet (g), and RetinaNet (h). The scale bar in the figure shows the 200 μm length of the sample surface.

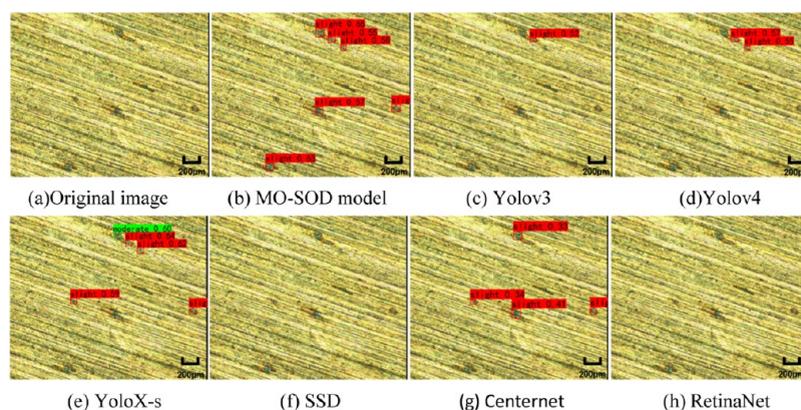


Figure 12. Comparison of the detection results among the proposed MO-SOD model with other models. Original image (a) and detection results produced by the proposed MO-SOD model (b), Yolov3 (c), Yolov4 (d), YoloX-s (e), SSD (f), Centernet (g), and RetinaNet (h). The scale bar in the figure shows the 200 μm length of the sample surface.

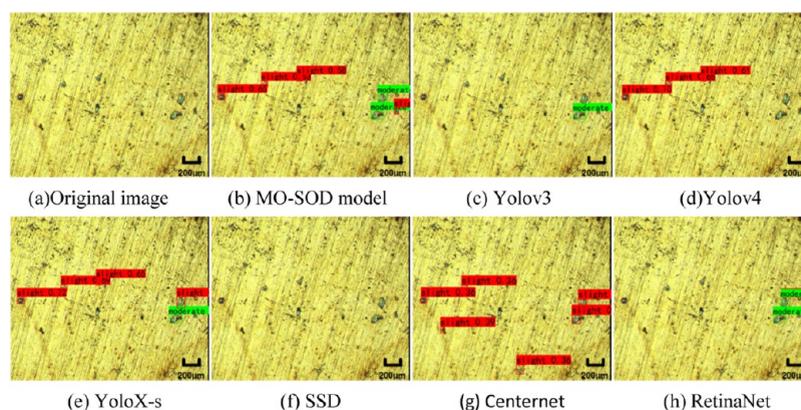


Figure 13. Comparison of the detection results among the proposed MO-SOD model with other models. Original image (a) and detection results produced by the proposed MO-SOD model (b), Yolov3 (c), Yolov4 (d), YoloX-s (e), SSD (f), Centernet (g), and RetinaNet (h). The scale bar in the figure shows the 200 μm length of the sample surface.

intermediate between slight oxidation class and severe oxidation. The model can detect micro-oxidation spots from different oxidation environments, such as oxygen-free copper samples oxidized in a neutral atmosphere at room temperature and oxygen-rich copper samples oxidized in a neutral atmosphere at high temperature.

4.2. Comparison with the Recent Detection Methods.

In this section, the proposed MO-SOD model is compared with

some of the most advanced detectors. Yolov3,²² Yolov4,²³ YoloX,²⁴ single shot multibox detector (SSD),²⁵ Centernet,²⁶ and RetinaNet²⁷ are chosen for comparison. For a fair comparison, the size of the compared detectors is adjusted to a similar scale and all detectors are trained without pretraining. As shown in Table 2, the proposed detector achieves 82.96% mAP, which is the best of these algorithms. The superiority of the proposed MO-SOD model is verified.

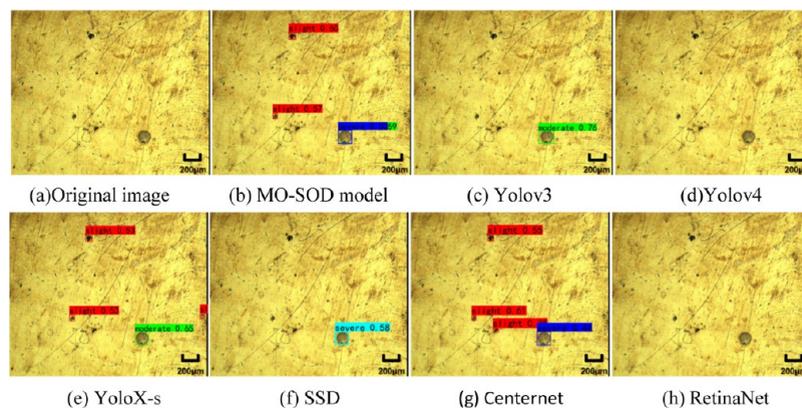


Figure 14. Comparison of the detection results among the proposed MO-SOD model with other models. Original image (a) and detection results produced by the proposed MO-SOD model (b), YOLOv3 (c), YOLOv4 (d), YOLOX-s (e), SSD (f), Centernet (g), and RetinaNet (h). The scale bar in the figure shows the 200 μm length of the sample surface.

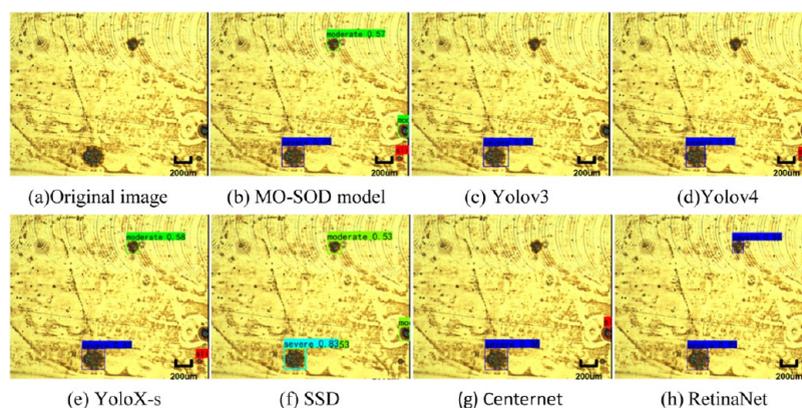


Figure 15. Comparison of the detection results among the proposed MO-SOD model with other models. Original image (a) and detection results produced by the proposed MO-SOD model (b), YOLOv3 (c), YOLOv4 (d), YOLOX-s (e), SSD (f), Centernet (g), and RetinaNet (h). The scale bar in the figure shows the 200 μm length of the sample surface.

Figures 1–5 show that the proposed method achieves the best detection of micro-oxidation spots compared with other recent algorithms. In Figures 1–3, it can be found that the proposed MO-SOD model can accurately locate the combination of slight oxidation spots and moderate oxidation spots in the noisy oxygen-free copper surface image, and the classification is accurate. However, YOLOv3, YOLOv4, SSD, and RetinaNet have poor detection effects and cannot completely detect all oxidation spots. In contrast, YOLOX-s and Centernet cannot accurately classify slight, moderate, and severe oxidation spots. YOLOX-s wrongly identifies moderate oxidation spot as severe oxidation spot in Figure 11, slight oxidation spot as moderate oxidation spot in Figure 12, and moderate oxidation spot as slight oxidation spot in Figure 13. Centernet classifies moderate oxidation spot as slight oxidation spot in Figure 13. Figure 14 shows that Centernet incorrectly detects the surface indentation of oxygen-free copper as oxidation point. In Figure 15, all models except the proposed MO-SOD model cannot detect all oxidation spots or the classification is inaccurate, where the proposed MO-SOD model can detect accurately even if the oxidation spot is truncated in image.

All detectors except the proposed MO-SOD model are limited by the tight layout of micro-oxidation, small differences in size, and features that are difficult to identify. SSD and RetinaNet are greatly affected by the background noise of a high-definition microscopic image of oxygen-free copper surface, and

it is difficult to detect micro-oxidation spots. Among the multiple dense micro-oxidation spots, YOLOv3 and YOLOv4 have very low recall. YOLOX-s is slightly lower than the recall of the proposed MO-SOD model while it has poor classification accuracy. Besides, the accuracy of the generated micro-oxidation spot bounding box for YOLOX-s is also worse than that of the proposed MO-SOD model. Centernet tends to over-recognize scratches on the background, while the classification accuracy is not as good as that of the proposed MO-SOD model. MO-SOD model is superior to all of the above detectors, and multilayer feature fusion enables it to accurately identify slight micro-oxidation spots scattered in the background. MO-SOD model can also accurately identify moderate micro-oxidation spots with unclear characteristics, as well as can accurately locate the severe micro-oxidation spots with obvious characteristics but few samples. This shows that the MO-SOD model can effectively detect surface micro-oxidation spots.

MO-SOD model can achieve such excellent detection results benefiting from its indispensable modules. Small object feature extraction layer adds spatial pyramid pooling at the near end of the feature extraction layer. It focuses on the local features of small object to improve the perception of micro-oxidation spots while taking into account the global features to reduce the impact of noisy background on feature extraction. It can greatly extract and fuse local area features; thus, it is more suitable for the detection of small objects in the clutter background. The

feature fusion layer embeds the key small object attention pyramid integration, which makes the channel attention and spatial attention for pyramid fusion. Channels focus on global information, while spatial attention is used locally. After attention mechanism processing, multiscale key small object features are fused with pyramid model for independent detection, which significantly improves the detection effect of micro-oxidation spots. Accurate classification of micro-oxidation spots at different levels requires features of higher latitude, while the accurate regression of micro-oxidation spots requires features of finer granularity. Therefore, the anchor-free decoupling detector applies two parallel branches for classification and regression to train and learn, respectively, so as to achieve a win–win situation for classification and regression.

5. CONCLUSIONS

In this research, we propose a micro-oxidation small object detection model, MO-SOD, for oxygen-free copper surface oxidation detection based on microscopic imaging system. It can be integrated into a detection platform combined with a high-definition microscopic system for rapid online detection.

The proposed model can be applied for detecting tiny spots, which are with noisy background or many scratches and difficult to distinguish. Based on our small object feature extraction layer, the features of changeable small objects in the noisy background are extracted by focusing on local features and fusing global features. Key small object attention pyramid integration is embedded in the feature fusion layer of the model, where the attention mechanism is applied to pyramid integration of the extracted small object features. Besides, anchor-free decoupling detector and loss function are designed to improve the accuracy of micro-oxidation spot detection.

Compared with other popular detection models, MO-SOD model achieves the highest micro-oxidation spot recognition rate of 82.96%. MO-SOD model provides a new noncontact method to detect the micro-oxidation spot on the oxygen-free copper surface and achieve the oxidation degree. This method can also be used to detect small targets in other noisy backgrounds.

AUTHOR INFORMATION

Corresponding Author

Taohong Zhang – Department of Computer, School of Computer and Communication Engineering, University of Science and Technology Beijing (USTB), Beijing 100083, China; Beijing Key Laboratory of Knowledge Engineering for Materials Science, Beijing 100083, China; orcid.org/0000-0003-3209-5836; Email: zth_ustb@163.com

Authors

Qianqian Li – Department of Computer, School of Computer and Communication Engineering, University of Science and Technology Beijing (USTB), Beijing 100083, China; Beijing Key Laboratory of Knowledge Engineering for Materials Science, Beijing 100083, China

Mingyang Yang – State Key Laboratory of Tribology, Beijing 100084, China

Xuxu Guo – Department of Computer, School of Computer and Communication Engineering, University of Science and Technology Beijing (USTB), Beijing 100083, China; Beijing Key Laboratory of Knowledge Engineering for Materials Science, Beijing 100083, China

Han Chen – Department of Computer, School of Computer and Communication Engineering, University of Science and Technology Beijing (USTB), Beijing 100083, China; Beijing Key Laboratory of Knowledge Engineering for Materials Science, Beijing 100083, China; orcid.org/0000-0003-4224-1366

Suli Fan – Department of Computer, School of Computer and Communication Engineering, University of Science and Technology Beijing (USTB), Beijing 100083, China; Beijing Key Laboratory of Knowledge Engineering for Materials Science, Beijing 100083, China

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acsomega.2c07043>

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the Science and Technology Innovation 2030 Major Project (No. 2020AAA0108703) of the Ministry of Science and Technology. The authors thank all reviewers, editors, and contributors for their contributions and suggestions.

REFERENCES

- (1) Koch, G.; Varney, J.; Thompson, N.; Moghissi, O.; Gould, M.; Payer, J. International measures of prevention, application, and economics of corrosion technologies study. *NACE Int.* **2016**, *216*, 2–3.
- (2) Shen, H. K.; Chen, P. H.; Chang, L. M. Automated steel bridge coating rust defect recognition method based on color and texture feature. *Autom. Constr.* **2013**, *31*, 338–356.
- (3) Koch, G. H.; Brongers, M. P.; Thompson, N. G.; Virmani, Y. P.; Payer, J. H. *Corrosion Cost and Preventive Strategies in the United States, United States. Federal Highway Administration*, 2002, No. FHWA-RD-01-156, R315-01.
- (4) Liu, L.; Tan, E.; Zhen, Y.; Yin, X. J.; Cai, Z. Q. In *AI-Facilitated Coating Corrosion Assessment System for Productivity Enhancement*, 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA); IEEE, 2018; pp 606–610.
- (5) Shen, H. K.; Chen, P. H.; Chang, L. M. Automated steel bridge coating rust defect recognition method based on color and texture feature. *Autom. Constr.* **2013**, *31*, 338–356.
- (6) Lowe, D. G. In *Object Recognition from Local Scale-Invariant Features*, Proceedings of the Seventh IEEE International Conference on Computer Vision; IEEE, 1999; Vol. 2, pp 1150–1157.
- (7) Canny, J. A computational approach to edge detection. *IEEE Trans. Pattern. Anal. Mach. Intell.* **1986**, PAMI-8, 679–698.
- (8) Dalal, N.; Triggs, B. In *Histograms of Oriented Gradients for Human Detection*, 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05); IEEE, 2005; Vol. 1, pp 886–893.
- (9) Ojala, T.; Pietikainen, M.; Maenpaa, T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern. Anal. Mach. Intell.* **2002**, *24*, 971–987.
- (10) Soh, L. K.; Tsatsoulis, C. Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices. *IEEE Trans. Geosci. Remote Sens.* **1999**, *37*, 780–795.
- (11) Joachims, T. 11 Making Large-Scale Support Vector Machine Learning Practical. In *Advances in Kernel Methods: Support Vector Learning*; Academic Press: MIT, 1999; p 169.
- (12) Son, H.; Hwang, N.; Kim, C.; Kim, C. Rapid and automated determination of rusted surface areas of a steel bridge for robotic maintenance systems. *Autom. Constr.* **2014**, *42*, 13–24.
- (13) Li, P.; Lin, L.; Chen, Y. A SROD Algorithm Based Accurate Detection Method for Surface Rust Spots in Ocean Ship. *J. Coastal Res.* **2018**, *83*, 921–926.

- (14) Liu, Y.; Xu, K.; Xu, J. Periodic surface defect detection in steel plates based on deep learning. *Appl. Sci.* **2019**, *9*, No. 3127.
- (15) Yao, Y.; Yang, Y.; Wang, Y.; Zhao, X. Artificial intelligence-based hull structural plate corrosion damage detection and recognition using convolutional neural network. *Appl. Ocean Res.* **2019**, *90*, No. 101823.
- (16) Bastian, B. T.; Jaspreeth, N.; Ranjith, S. K.; Jiji, C. V. Visual inspection and characterization of external corrosion in pipelines using deep neural network. *NDT&E Int.* **2019**, *107*, No. 102134.
- (17) Papamarkou, T.; Guy, H.; Kroencke, B.; Miller, J.; Robinette, P.; Schultz, D.; et al. Automated detection of corrosion in used nuclear fuel dry storage canisters using residual neural networks. *Nucl. Eng. Technol.* **2021**, *53*, 657–665.
- (18) Zhang, S.; Deng, X.; Lu, Y.; Hong, S.; Kong, Z.; Peng, Y.; Luo, Y. A channel attention based deep neural network for automatic metallic corrosion detection. *J. Build. Eng.* **2021**, *42*, No. 103046.
- (19) Forkan, A. R. M.; Kang, Y. B.; Jayaraman, P. P.; Liao, K.; Kaul, R.; Morgan, G.; Ranjan, R.; Sinha, S. CorrDetector: A framework for structural corrosion detection from drone images using ensemble deep learning. *Expert Syst. Appl.* **2022**, *193*, No. 116461.
- (20) Bochkovskiy, A.; Wang, C. Y.; Liao, H. Y. M. *Yolov4: Optimal Speed and Accuracy of Object Detection*, arXiv:2004.10934. arXiv.org e-Print archive. <https://doi.org/10.48550/arXiv.2004.10934>.
- (21) glenn jocher. *yolov5*, <https://github.com/ultralytics/yolov5>2021.
- (22) Redmon, J.; Farhadi, A. *Yolov3: An Incremental Improvement*, 2018. arXiv:1804.02767. arXiv.org e-Print archive. <https://doi.org/10.48550/arXiv.1804.02767>.
- (23) Bochkovskiy, A.; Wang, C. Y.; Liao, H. Y. M. *Yolov4: Optimal Speed and Accuracy of Object Detection*, 2020. arXiv:2004.10934. arXiv.org e-Print archive. <https://doi.org/10.48550/arXiv.2004.10934>.
- (24) Ge, Z.; Liu, S.; Wang, F.; Li, Z.; Sun, J. Yolox: Exceeding Yolo Series in 2021, 2021. arXiv:2107.08430. arXiv.org e-Print archive. <https://doi.org/10.48550/arXiv.2107.08430>.
- (25) Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C. Y.; Berg, A. C. *Ssd: Single Shot Multibox Detector*, European Conference on Computer Vision; Springer: Cham, 2016; pp 21–37.
- (26) Zhou, X.; Wang, D.; Krähenbühl, P. *Objects as Points*, 2019. arXiv:1904.07850. arXiv.org e-Print archive. <https://doi.org/10.48550/arXiv.1904.07850>.
- (27) Lin, T. Y.; Goyal, P.; Girshick, R.; He, K.; Dollár, P. In *Focal Loss for Dense Object Detection*, Proceedings of the IEEE International Conference on Computer Vision, 2017; pp 2980–2988.