Shahrood University of Technology

Journal of Artificial Intelligence and Data Mining (JAIDM)
Journal homepage: http://jad.shahroodut.ac.ir

Research paper

Automated Surface Defect Detection in Copper Blanks Using YOLOv8 Segmentation and EfficientNetV2-S Classification

Hossein Ghayoumi Zadeh[1]*, Ali Fayazi[1], khosro Rezaee[2], Afsaneh Aminaee[3], Hadi Halavati[4], Mehdi Tahernejad[5], Hadi Memarzadeh[6], Ali Masoumi[6] and Mohammad Sadegh Jafari[5]

1. Department of Electrical Engineering, Vali-e-Asr University of Rafsanjan, Rafsanjan, Iran.
2. Department of Biomedical Engineering, Meybod University, Meybod, Iran.
3. Director of Training and Competency Development, Sarcheshmeh Copper Complex, Rafsanjan, Iran.
4. Research and Development Division, Sarcheshmeh Copper Complex, Rafsanjan, Iran.
5. Technical and Engineering Research, Research and Development Department, Sarcheshmeh Copper Complex, Rafsanjan, Iran.
6. Refinery and Casting Division, Sarcheshmeh Copper Complex, Rafsanjan, Iran.

Article Info

Article History:
Received 20 September 2025
Revised 19 December 2025
Accepted 03 January 2025

DOI:10.22044/jadm.2026.16853.2818

Keywords:
Copper Electrorefining, Surface Defect Detection, Industrial Image Classification, EfficientNetV2-S, Deep Learning in Quality Control, Computer Vision in Metallurgy.

**Corresponding author: h.ghayoumizadeh@vru.ac.ir (H. Ghayoumizadeh)*

Abstract

In this study, an intelligent deep learning–based system is proposed for the automated detection of surface defects in copper cathode blanks used in the electrorefining process. The proposed pipeline combines a YOLOv8-based segmentation model with an EfficientNetV2-S classifier to localize and analyze defect-relevant regions of each blank. The segmentation module identifies the main copper regions, edge strips, and defect-prone areas associated with surface anomalies such as scratches, dents, misalignment, and discoloration, effectively reducing background interference and improving classification reliability. The dataset includes 5,266 labeled images with a significant class imbalance, addressed using focal loss and class weighting during training. Experimental results on the test set demonstrate strong performance, achieving 98.32% accuracy, 96.71% precision, 95.67% recall, an F1-score of 96.19%, and an AUC of 0.9953. Grad-CAM visualizations and error analysis further confirm that the model consistently focuses on meaningful defect regions while remaining robust to background and illumination variations. These results highlight the effectiveness of the proposed approach for reliable quality control in industrial copper electrorefining lines.

1. Introduction

Copper electrorefining is a well-established metallurgical process that enables the production of copper with very high purity levels, often exceeding 99.99% [1, 2]. In this process, impure copper from smelting serves as the anode, while stainless steel cathodes act as reusable substrates onto which copper ions are deposited. Despite continuous improvements in cell design and electrolyte chemistry, maintaining uniform and defect-free deposition remains a persistent challenge. Even small anomalies in the deposited layer can complicate the stripping stage, causing excessive adhesion or incomplete separation of the copper from the cathode [3, 4]. Such problems not only lower process efficiency but also increase operational costs through equipment damage and production downtime. Figure 1 provides a schematic overview of the electrorefining setup [5].

The durability of stainless steel cathodes, which are designed to undergo numerous production cycles, is another critical factor in sustaining efficient electrorefining operations [6]. Mechanical deformation or surface damage to these blanks accelerates replacement requirements and disrupts production continuity. In large-scale industrial

facilities, such as the Sarcheshmeh Copper Complex in Rafsanjan, Iran, surface defect identification has traditionally relied on visual inspection by human operators.

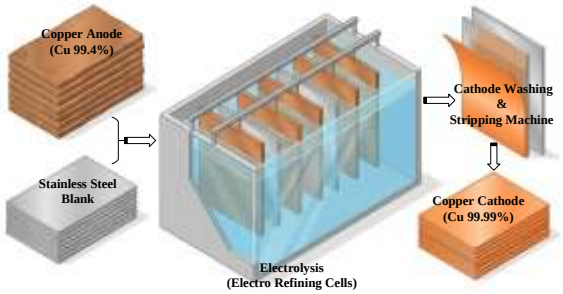

Figure 1. Schematic overview of the copper electrorefining process. The figure is an original redraw by the authors and does not reproduce any previously published illustration; it is based on the general process description reported in [5].

While simple in principle, manual inspection suffers from significant limitations, including operator fatigue, subjective judgment, and reduced reliability under harsh industrial conditions. Exposure to electrolyte fumes, high temperatures, and chemically aggressive environments further degrade inspection accuracy and pose health and safety risks to operators.

Delayed or inaccurate detection of surface defects allows defective blanks to progress to subsequent production stages, potentially causing equipment damage, increased scrap rates, and deterioration of final product quality. Over time, the physical and mental demands placed on human inspectors also elevate the risk of workplace accidents and operational inefficiencies. These challenges highlight the need for automated inspection solutions that can operate consistently and reliably under real-world electrorefining conditions.

In this context, advanced image processing and deep learning techniques offer a promising alternative to conventional inspection methods. By enabling continuous, rapid, and objective defect detection, such approaches can significantly reduce human error, minimize operational downtime, and lower overall production costs. Accordingly, this study develops an intelligent inspection system that combines a YOLOv8-based model for detecting and segmenting critical regions of copper blanks with an EfficientNetV2-S network for the final classification of surface defects.

By focusing analysis on defect-relevant regions rather than entire images, the proposed approach aims to enhance detection accuracy while maintaining robustness to environmental variability commonly encountered in industrial electrorefining lines.

2. Related work

Ensuring reliable detection of surface defects is essential for maintaining product quality and production-line efficiency [7]. Early industrial inspection systems largely depended on manual visual checks or classical image-processing techniques such as thresholding, Gabor filtering, and wavelet-based feature extraction [8]. While these approaches can perform adequately in controlled settings, they often degrade under real factory conditions where lighting fluctuations, specular reflections, and background texture variations are unavoidable—particularly when defects are subtle, small, or embedded in noisy surface patterns [8].

With the shift toward data-driven solutions, deep learning—especially convolutional neural networks (CNNs)—has become the dominant paradigm for automated surface inspection. Following major breakthroughs in large-scale visual recognition, CNN backbones such as AlexNet, VGG, and ResNet have been widely adopted for defect classification and recognition in industrial applications [9–13]. However, pure image-level classification only indicates whether a defect exists and does not provide spatial localization. This limitation motivated the adoption of object detection frameworks that simultaneously localize and recognize defect regions [14].

Two-stage detectors (e.g., Faster R-CNN) have been successfully adapted for inspection tasks, often by strengthening multi-scale representation and proposal quality. Representative enhancements include feature pyramid integration for improved detection of small defects [15], guided-anchor strategies and backbone optimizations for PCB and metal inspection [16], and deformable convolution or refined suppression techniques for steel-related applications [17]. Despite strong accuracy, two-stage pipelines typically involve heavier computation and slower inference, which can be restrictive for real-time deployment on production lines.

To meet real-time constraints, many recent studies have moved toward one-stage detectors such as SSD and YOLO [14]. Numerous YOLO variants have been tailored for industrial defect detection by improving feature fusion, reducing model complexity, and strengthening small-target sensitivity. For example, YOT-Net extended YOLOv3 using triplet-loss learning for copper elbow inspection [8], and other works combined YOLO backbones with lightweight networks to achieve favorable speed–accuracy trade-offs on metallic defect datasets [19]. Variants of YOLOv4 and YOLOv5 have incorporated lightweight

designs, shallow–deep fusion, pruning strategies, and attention mechanisms to enhance detection under complex textures [20–23]. More recent anchor-free developments and improved label assignment schemes (e.g., YOLOv6) further improved robustness and localization quality in challenging scenarios [17].

Within metal-surface inspection, copper-related studies increasingly adopt YOLO due to its balance of speed and accuracy. For instance, Zhang et al. integrated a lightweight backbone and attention mechanisms to detect nodules on cathode plates while maintaining real-time capability [24]. CSC-YOLO, derived from a YOLOv4-tiny design with cross-region fusion and channel attention, reported strong accuracy on copper strip/plate defect datasets with high throughput [25]. Similar trends are observed in steel inspection, where modified YOLO pipelines (e.g., Res2Net backbones, double-FPN designs, and transformer/attention augmentation) consistently report measurable gains over baseline models [26–29]. Recent 2025 studies continue this direction: Zhou et al. proposed AEB-YOLOv8n for copper strips, improving mAP@0.5 while reducing parameters and computation, with explicit emphasis on small-defect representation and efficient multi-scale fusion [33]. Likewise, Zhang et al. introduced an improved YOLOv10-based model (LAM-YOLOv10n) for steel surface defects, reporting a precision gain over the YOLOv10n baseline while targeting practical real-time constraints [34].

Beyond YOLO-centric detectors, researchers have explored alternative backbones and hybrid designs to improve feature quality and deployment efficiency. EfficientNet-based modifications have been used to balance accuracy and inference cost, including attempts to replace standard YOLO backbones with EfficientNet variants combined with attention modules [30]. ResNet-based designs remain common for multi-scale defect recognition and PCB inspection, especially when paired with attention and feature pyramid strategies [16, 31]. In parallel, unsupervised and generative approaches (e.g., autoencoders and GANs) have been investigated for anomaly detection, particularly when defect labels are scarce or class distributions are highly imbalanced. Lightweight deployment has also been addressed through pruning and edge-oriented optimization, enabling faster inference on embedded GPUs without a major loss in accuracy [32].

Although prior studies demonstrate that YOLO-based detectors can localize defects efficiently, most existing pipelines emphasize bounding-box detection on generic metal datasets (often copper strips or steel plates) and do not explicitly address inspection requirements where (i) background suppression is crucial, (ii) defect regions are small and visually ambiguous, and (iii) downstream decision-making depends on fine-grained region analysis rather than coarse localization. In copper electrorefining environments, reflections, surface texture variability, and operational artifacts can further complicate detection. Moreover, many defect inspection workflows treat localization and final classification as a single-stage objective, whereas practical quality-control settings may benefit from a structured pipeline that first isolates critical regions and then performs a dedicated classification with loss functions tailored to class imbalance. These gaps motivate the proposed two-stage system in this work, where YOLOv8-based segmentation is used to focus on key regions and suppress irrelevant background before EfficientNetV2-S performs defect classification under focal loss and class-weighting constraints.

3. Method

3.1. Dataset

In this study, to train and evaluate the proposed system for automated surface defect detection in copper blanks, a dedicated dataset comprising 5,266 real images was collected from the electrorefining process at the Sarcheshmeh Copper Refinery in Rafsanjan, Iran. The data were gathered over several weeks through continuous sampling in collaboration with the plant's quality control team. Image acquisition was carried out using a Basler acA1300-200uc industrial camera, known for its high resolution and suitability for harsh environments. The camera was mounted in a fixed position to ensure a consistent angle and distance relative to the blank in all samples.

To ensure high image quality and consistency during acquisition, a high-precision Time-of-Flight (ToF) laser distance sensor (wavelength: 940 nm, Class 1, measurement range: 1 cm to 5 m, accuracy: ±1.5 cm) was employed to measure the distance between the camera and the blank surface in real time. Images were only captured when the measured distance was within a predefined range and the reflected signal from the blank surface met the required intensity threshold. This automated validation mechanism ensured that only high-quality images under optimal visual and geometric conditions were stored in the dataset, thereby minimizing noise and inconsistencies.

Maintaining standardization in image acquisition is crucial in computer vision and deep learning applications, particularly in industrial environments where minor deviations in lighting,

angle, or object placement can significantly impact model performance. To enhance dataset diversity and simulate real-world production conditions, image capture was conducted at different times of day and under varying workshop conditions, including the presence of electrolyte vapor, natural and artificial light fluctuations, and changes in humidity. This approach resulted in a dataset that better represents real operating conditions and improves the generalization capability of the trained model.

The dataset was divided into two primary categories: Normal, representing blanks with consistent copper deposition and no critical flaws, and Reject, which includes blanks with significant issues that can interfere with the stripping process. As shown in Figure 2.a, normal blanks exhibit a smooth and uniform copper layer. In contrast, Figures 2.b–2.f present typical defective cases. In Figure 2.b, the copper layer bridges across the two upper windows of the blank, preventing proper separation. Figure 2.c highlights the absence of the edge strip, allowing copper to spread over the edges and increasing the likelihood of damage. Figure 2.d shows incomplete deposition in the lower part of the blank, signaling a failure in the process. In Figure 2.e, the edge strip is present but misaligned, leading to abnormal copper overflow around it. Lastly, Figure 2.f depicts a bent and damaged upper window, which can cause mechanical interference in the stripping equipment.

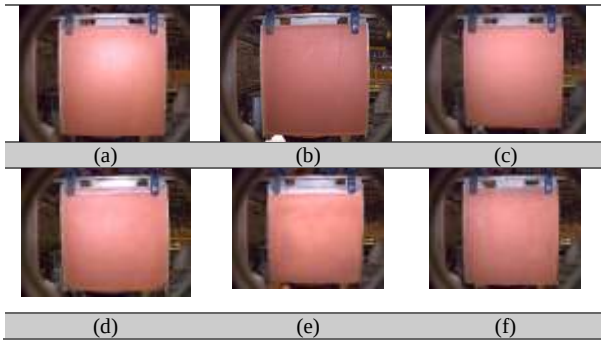

Figure 2. Sample images of copper blanks. (a) Normal blank with uniform copper deposition. (b–f) Various defect types including: copper bridging upper windows (b), missing edge strip (c), incomplete lower deposition (d), misaligned edge strip (e), and deformed upper window (f).

Together, these examples illustrate the range and complexity of defects in the Reject class and reinforce the necessity for reliable, automated inspection systems.

3.2. Segmentation

Semantic segmentation plays a crucial role in computer vision tasks that require precise localization and classification of each pixel in an image. In contrast to traditional classification models that assign a single label to the entire image, segmentation models analyze the spatial structure of the image to determine which pixels belong to which object or region [35]. This is particularly important in industrial inspection problems where the shape, position, and boundary of an object directly affect its functional assessment. In surface defect detection scenarios, segmentation can isolate regions of interest—such as the surface of the product—while ignoring irrelevant background information, thereby improving the accuracy and robustness of subsequent classification stages. By leveraging segmentation, the system is empowered to focus on specific zones of the object under inspection, which is essential when defects are subtle, localized, and possibly embedded in noisy surroundings [36].

In our study, early experiments showed that using classification alone on full-frame images often led to misclassification—especially in borderline cases where the defect was visually subtle or narrowly localized. In several instances, the classifier sometimes confused background textures, lighting variations, or equipment markings with actual defects. This confusion led to both false positive and false negative results [37]. This issue became more prominent due to the nature of the copper blanks used in electrorefining: although the position of each blank was fixed, variations in surface oxidation, lighting reflection, and environmental conditions introduced significant background noise. Furthermore, some reject conditions, such as incomplete copper deposition at the bottom or a thin bridge of copper between the top windows, were small in size and located at the periphery. Without spatial filtering, these features could be easily overlooked by a deep classifier trained on the entire image. Therefore, we decided to incorporate a segmentation stage as a preprocessing step to extract only the relevant regions of the blank from the original image. This not only reduced input complexity but also allowed the classifier to focus exclusively on the critical areas where defects are likely to occur.

YOLOv8-seg was adopted as the segmentation backbone due to its practical balance between mask quality and inference efficiency for industrial deployment. Compared with earlier YOLO generations commonly used in defect inspection (e.g., YOLOv5/YOLOv7), YOLOv8 employs a more modern head design and training strategy that reduces sensitivity to anchor settings and improves localization of small, low-contrast regions—an important requirement in our application where

defects may appear as thin copper bridges, edge-strip anomalies, or incomplete deposition near boundaries. In addition, as a one-stage instance segmentation model, YOLOv8-seg directly outputs pixel-level masks with a compact architecture, enabling standardized cropping and background suppression prior to classification. This provides a favorable speed–accuracy trade-off compared with heavier two-stage alternatives such as Mask R-CNN or semantic segmentation pipelines like DeepLab, particularly under the practical constraints of memory usage and real-time inference. In our preliminary trials, the segmentation-guided input reduced misclassifications caused by background textures and illumination artifacts while maintaining the throughput needed for online quality control in copper refineries [38].

Mathematically, the YOLOv8 model is designed as a single-stage detector that maps an input image $x \in \mathbb{R}^{H\times W\times 3}$ to a set of predictions including bounding boxes $b=(x,y,w,h)$, class probabilities $p \in [0,1]^C$, and binary segmentation masks $m \in \{0,1\}^{H\times W}$. The model optimizes a multi-part loss function:

$$\text{Loss}_{\text{YOLOv8}} = \lambda_{\text{mask}}\text{Loss}_{\text{mask}} + \lambda_{box}\,\text{Loss}_{\text{box}} + \lambda_{\text{cls}}\text{Loss}_{\text{cls}} \quad (1)$$

In this formulation, the total loss is expressed as a weighted sum of three components. The classification term (λ_{cls}) measures how well the model distinguishes between classes and is typically implemented using Binary Cross-Entropy. The regression term (λ_{box}) evaluates the accuracy of bounding box localization, often with a loss such as CIoU. The mask term (λ_{mask}) applies binary cross-entropy to the predicted segmentation masks to ensure precise pixel-level delineation. The λ coefficients act as balancing factors, regulating the relative contribution of each component to the overall optimization process.

This unified design allows YOLOv8 to jointly capture object existence, spatial location, and precise contours, leading to substantial improvements in identifying small, fine-grained, or low-contrast defects—characteristics commonly observed in the reject samples of our dataset.

In training the YOLOv8 segmentation model, labeling was carried out at the blank-surface level, including the copper deposition zone, edge strips, and potential defect regions. This choice stemmed from the observation that distinguishing normal from reject blanks often depends on very fine details—such as slight miscoverage of copper or small bridging—that are only meaningful within the geometry of the blank itself. Conversely, the surrounding scene introduced irrelevant variations in lighting, textures, and background structures that could confuse the model. Focusing annotations solely on the blank region eliminated this noise, resulting in clearer masks and more consistent preprocessing. Such structured labeling not only enabled precise cropping but also enhanced the overall robustness and interpretability of the defect detection pipeline.

Figure 3 demonstrates how preprocessing improves blank-region segmentation. In the left panel (a), the raw camera image contains background artifacts and environmental noise that can interfere with reliable feature extraction. In contrast, the right panel (b) shows the outcome after YOLO-based segmentation and cropping, where a clear and standardized view of the blank surface is obtained, providing a more suitable input for downstream analysis and classification.

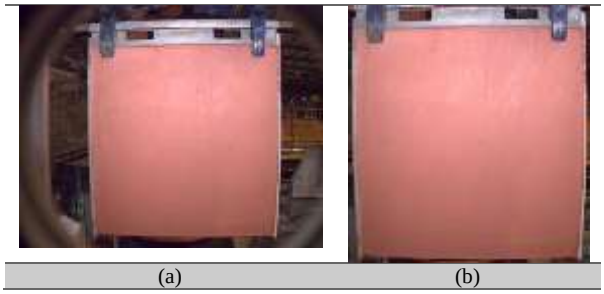

Figure 3: (a) Example of a raw input image captured from the production line, including background clutter; (b) Preprocessed and segmented blank region using YOLOv8, isolating the area of interest for reliable defect detection.

3.3. Classification Using EfficientNetV2-S

Accurate classification of copper blank surfaces—particularly in distinguishing between normal and reject categories—requires a model capable of detecting fine-grained visual differences across localized regions. In our dataset, many of the reject cases were characterized by subtle anomalies such as incomplete copper deposition, narrow bridging between zones, or slight deformations in edge areas, all of which were difficult to capture without a high-capacity yet detail-sensitive network. These challenges ruled out the use of shallow models or standard CNNs, which often fail to detect nuanced surface-level inconsistencies, especially when trained on small or imbalanced datasets. To address these demands, we selected EfficientNetV2-S, a state-of-the-art convolutional neural network that combines efficient scaling, faster convergence, and enhanced representational capacity. Its compound scaling method allows the model to maintain a good trade-off between accuracy and computational cost—making it suitable for deployment in industrial environments while ensuring robustness in detecting delicate surface

defects. The network's proven performance in various fine-grained classification benchmarks further supported its adoption in our pipeline [39, 40].

For the classification task, we selected EfficientNetV2-S, a modern convolutional neural network architecture that offers a strong trade-off between accuracy and computational efficiency. This model is particularly well-suited for our application, as distinguishing between normal and reject copper blanks often depends on fine-grained local features such as incomplete copper deposition, irregular connections near the windows, or subtle anomalies around the edge strip area. These features are challenging to capture using conventional CNNs, especially under varying illumination and industrial background noise. All images were resized to a resolution of 384×384 and normalized using standard ImageNet statistics:

$$x_{norm} = \frac{x - \mu}{\sigma} \tag{2}$$

EfficientNetV2 is based on compound scaling, which uniformly scales the depth d, width w, and resolution r of the network using a set of fixed coefficients (α, β, γ) subject to the constraint:

$$\alpha\beta^2\gamma^2 \approx 2, \text{ for a } 2x \text{ increase in FLOPs} \tag{3}$$

The goal is to maintain optimal accuracy while increasing the model size within a bounded computational budget. Each scaling factor contributes as follows:

- d: increases network depth (number of layers),
- w: increases the number of channels per layer,
- r: increases the input image resolution.

The model is composed of multiple stages of Fused-MBConv and MBConv blocks. Each MBConv block applies a depthwise separable convolution with an expansion phase followed by a squeeze-and-excitation (SE) mechanism. The output of an MBConv block can be expressed as:

$$y = SE(BN(DWConv(x, W\_d))) + x \tag{4}$$

Where x is the input feature map, DWConv is depthwise convolution, BN is batch normalization, and SE denotes the channel attention module. In early layers, EfficientNetV2-S uses Fused-MBConv blocks that replace the separate expansion and depthwise convolution with a single 3×3 convolution:

$$y = BN\left(Conv_{3\times3}(x, W_{conv})\right) \tag{5}$$

This design improves GPU parallelism and reduces training time. After the final convolutional stage, the feature map $F \in \mathbb{R}^{H\times W\times C}$ is passed through a Global Average Pooling (GAP) layer and a fully connected Dense layer with sigmoid activation to perform binary classification:

$$\hat{y} = \sigma(W \cdot GAP(F) + b) \tag{6}$$

Where:

$$GAP(F) = \frac{1}{HW}\sum_{i=1}^{H}\sum_{j=1}^{W} F_{i,j} \tag{7}$$

$$\sigma(z) = \frac{1}{1+e^{-z}} \tag{8}$$

This formulation allows the network to combine spatial features effectively and make its final classification decision within a broader contextual view. Compared to the original EfficientNet, the EfficientNetV2-S variant incorporates several refinements that enhance both accuracy and efficiency. Early layers employ fused operations to streamline computations, while a progressive training strategy gradually increases input resolution and augmentation difficulty to improve learning stability. In addition, the architecture has been carefully tuned for faster GPU inference and stronger convergence on small- to medium-scale datasets, making it well-suited for practical industrial applications.

These improvements reduce training time by up to 30% while maintaining or improving accuracy on tasks involving subtle visual distinctions.

The architecture depicted in Figure 4 illustrates the sequential stacking of fused-MBConv and MBConv blocks, incorporating squeeze-and-excitation modules and compound scaling to optimize the model's capacity for depth, width, and resolution.

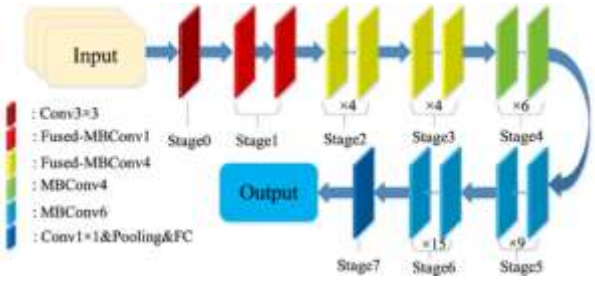

Figure 4. The architecture of EfficientNetV2-S, consisting of progressive convolutional stages with fused-MBConv and MBConv blocks optimized for both speed and accuracy.

3.4. Addressing Class Imbalance: Class Weighting and Focal Loss

One of the primary challenges encountered in training deep learning models for industrial defect classification is class imbalance, a condition where the number of samples in one class (typically the "normal" or non-defective cases) significantly outweighs the number of samples in the other class (i.e., "reject" or defective cases). In our dataset, this imbalance was particularly pronounced, as reject copper blanks—those exhibiting surface defects, misaligned edge strips, or incomplete copper deposition—occur infrequently in the production line of the Sarcheshmeh copper refinery. As a result, the training set contained several times more normal samples than reject ones, leading to a natural bias in favor of the majority class.

This imbalance can severely hinder the model's ability to detect and generalize rare but critical anomalies. A classifier trained on such a skewed dataset may achieve deceptively high accuracy simply by predicting the dominant class for most inputs, while entirely missing subtle reject cases—particularly dangerous in real-world inspection scenarios where failure to identify a defect can lead to process interruptions or equipment damage.

To mitigate this problem, we first adopted a class weighting strategy that modifies the standard binary cross-entropy loss to emphasize the minority class during optimization. Let y∈ {0, 1} denote the true label, and $\hat{y} \in (0, 1)$ the predicted probability of the positive class (i.e., "reject"). The weighted binary cross-entropy loss is defined as:

$$L_{\text{weighted}} = -w_1 y \log(\hat{y}) - w_0 (1-y) \log(1-\hat{y}) \quad (9)$$

In this formulation, w1 represents the weight assigned to the positive class (minority), while w_0 corresponds to the negative class (majority). Typically, w_1 is set higher than w_0 to compensate for class imbalance, and the values are calculated based on the inverse frequency of each class in the dataset. This weighting strategy ensures that the model pays greater attention to the underrepresented reject samples during training.

$$w_c = \frac{N}{C \cdot N_c} \quad (10)$$

Here, N is the total number of samples, N_c is the number of samples in class c, and C is the number of classes (2 in our binary case). This approach encourages the model to pay more attention to underrepresented reject cases, improving sensitivity and reducing bias.

While class weighting helps reduce the impact of imbalance, it is often insufficient in scenarios where the classifier continues to overfit "easy" samples and fails to learn from hard or ambiguous examples. To enhance the learning process further, we incorporated focal loss, a modified cross-entropy loss function that dynamically scales the contribution of each sample based on its classification difficulty. The focal loss function is defined as:

$$L_{\text{focal}} = -\alpha_t (1-p_t)^{\gamma} \log(p_t) \quad (11)$$

Where:

- $p_t = \hat{y}$ if y=1 , and $p_t = 1-\hat{y}$ if y=0
- $\alpha_t \in [0,1]$ is a class-balancing parameter similar to class weighting
- $\gamma \geq 0$ is the focusing parameter that reduces the contribution of well-classified samples

The intuition behind focal loss is that it down-weights the loss for easy examples (where $p_t \approx 1$) and focuses more on hard or misclassified examples (where $p_t \ll 1$). This makes it particularly effective in imbalanced datasets with fine-grained classification needs. The Focal Loss parameters, $\alpha=0.6$ and $\gamma=1.5$, were selected after extensive empirical experiments with various values on the validation set to achieve optimal performance. The final loss function used for training the classifier was:

$$L_{\text{final}} = L_{\text{focal}}(\hat{y}, y; \alpha = 0.6, \gamma = 1.5) \quad (12)$$

This formulation improved the network's ability to detect borderline reject cases, which are often visually subtle and underrepresented. Moreover, it helped reduce false negatives—instances where defective blanks might be misclassified as normal—thereby increasing the reliability and robustness of the inspection system for industrial deployment.

4. Result

To evaluate the effectiveness of the proposed defect detection system, the dataset was divided into three distinct subsets: training, validation, and test sets. The training set included a total of 2,814 images, comprising 2,300 samples labeled as normal blanks and 514 samples as reject blanks. The validation set, used for model tuning and early stopping, consisted of 1,202 images, with 919 belonging to the normal class and 283 to the reject class. Finally, the test set, which remained unseen during training, included 1,250 images—973 from the normal category and 277 from the reject category. This distribution reflects a class imbalance, particularly in the reject class, and

necessitated the use of specialized training strategies such as focal loss and class weighting to mitigate potential performance degradation. In the following subsections, we present and analyze the classification results using metrics including accuracy, precision, recall, F1-score, AUC, and confusion matrix.

Figure 5 illustrates the accuracy progression of the proposed model during the training process over 100 epochs. As shown in the plot, both training and validation accuracy improve significantly during the initial epochs, with the training accuracy starting from approximately 75% and steadily approaching 100%. The validation accuracy follows a similar trend, rapidly increasing to around 97% and maintaining a stable trajectory throughout the later epochs.

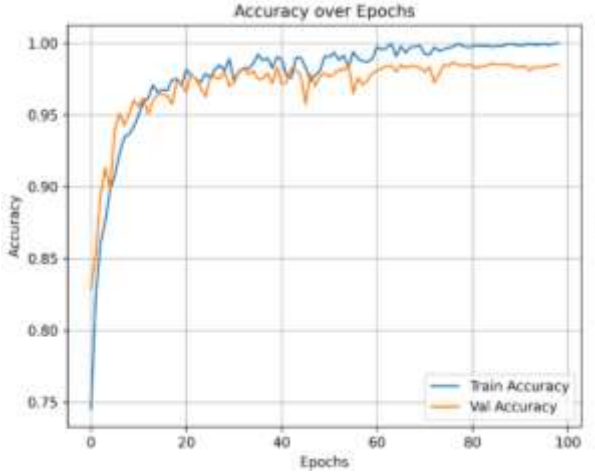

Figure 5. Accuracy trends during training and validation over 100 epochs using EfficientNetV2-S.

The close alignment between training and validation curves suggests that the model generalizes well to unseen data and does not exhibit overfitting. This consistent accuracy is indicative of the model's ability to correctly distinguish between normal and defective copper blanks based on subtle surface patterns and localized defects. The use of a strong backbone (EfficientNetV2-S), along with appropriate regularization techniques such as dropout and batch normalization, has likely contributed to this high and stable performance.

Figure 6 presents the training and validation loss curves over 100 epochs. As illustrated, both loss functions exhibit a steep and consistent decline during the early training stages, reflecting rapid convergence and effective learning. The training loss begins around 0.9 and decreases smoothly to near-zero values, indicating successful error minimization on the training set. The validation loss closely follows the same trend, which suggests that the model generalizes well to unseen data and does not overfit.

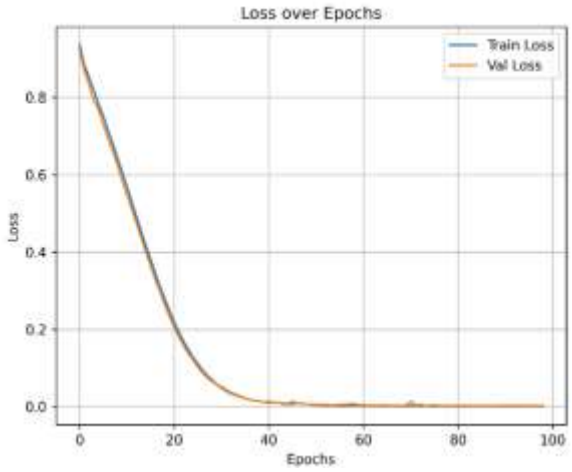

Figure 6. Loss reduction trends during training and validation, indicating effective model convergence.

The close alignment between training and validation loss curves demonstrates that the regularization techniques—such as dropout, batch normalization, and class weighting—were effective in stabilizing the learning process. Moreover, the overall smoothness and convergence of the loss functions validate the suitability of the chosen model architecture (EfficientNetV2-S) for classifying surface defects in copper blanks. The low final loss on both datasets confirms that the model has learned discriminative features that distinguish normal blanks from reject cases, even under subtle or noisy variations.

Figure 7 depicts the evolution of the Area under the Curve (AUC) metric for both the training and validation datasets over the course of 100 epochs. AUC is a robust measure that reflects the model's ability to distinguish between classes across various decision thresholds. A value closer to 1.0 indicates excellent separability, while a value of 0.5 suggests performance equivalent to random guessing. As shown in the plot, the training AUC increases rapidly during the early stages and quickly approaches 1.0, maintaining a nearly perfect score in the later epochs. The validation AUC follows a similar trajectory, reaching over 0.95 within the first 10 epochs and remaining stable with minimal fluctuation thereafter. The proximity of the validation AUC to the training curve is a strong indicator of the model's generalization capability and suggests that the learned representations are robust against overfitting. The consistently high AUC scores confirm that the model performs well not only at a single classification threshold (like 0.5), but across the entire range of possible thresholds—an essential requirement in industrial inspection tasks where the cost of misclassification may vary. This reliability across thresholds demonstrates that the classifier is

highly effective in identifying both subtle and obvious defects on the copper blank surfaces.

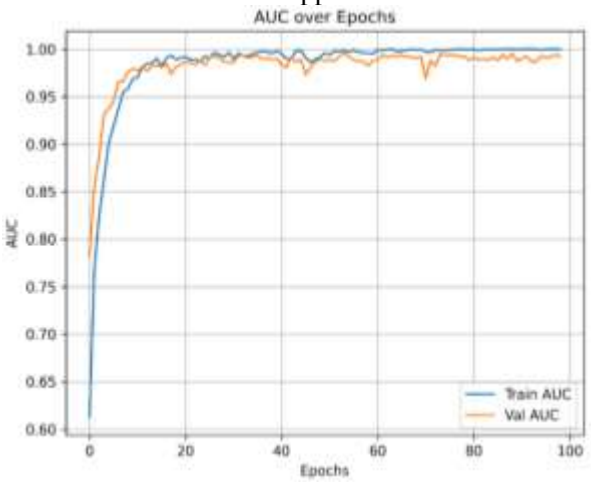

Figure 7. AUC progression over epochs, reflecting improved class separability during training and validation.

Figure 8 shows the Receiver Operating Characteristic (ROC) curve for the proposed classification model, evaluated on the test dataset. The ROC curve represents the trade-off between the true positive rate (TPR) and false positive rate (FPR) across various classification thresholds. The closer the curve hugs the top-left corner, the better the model is at distinguishing between the two classes—here, normal and reject blanks.

The computed area under the curve (AUC = 0.9953) indicates near-perfect classification performance. AUC is a threshold-independent metric and is especially useful in imbalanced classification tasks such as this, where the reject class is underrepresented. A value above 0.99 signifies that the model is highly capable of ranking positive samples higher than negative ones, regardless of the chosen decision threshold.

From the shape of the curve, it is evident that the classifier achieves a high TPR with minimal FPR, confirming that the majority of reject samples are correctly identified without misclassifying many normal samples. This characteristic is critical in real-world copper production environments, where false negatives (i.e., missed defects) can lead to operational failures, and false positives (i.e., unnecessary rejections) can increase costs.

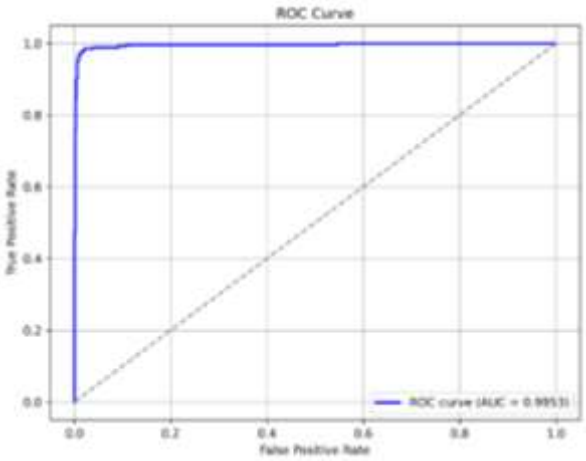

Figure 8. ROC curve for the proposed model, demonstrating excellent separability with an AUC of 0.9953.

Figure 9 illustrates the Precision–Recall (PR) curve for the proposed classification model. The PR curve is especially informative in imbalanced classification scenarios, such as the current task where the number of reject samples is significantly lower than the normal class. Unlike the ROC curve, which considers true negatives, the PR curve focuses solely on the positive class (here, reject) and offers a better understanding of the trade-off between precision and recall. As shown, the model maintains a high level of precision across a wide range of recall values, with minimal drop-offs until very high recall thresholds. This behavior confirms that the model is able to detect most of the defective blanks (high recall) without producing many false positives (high precision). The curve remains close to the upper-right corner, which indicates superior performance. The reported Average Precision (AP) of 0.9871 further quantifies this performance. AP is computed as the area under the precision–recall curve and summarizes the model's ability to balance both objectives across all thresholds. An AP near 1.0 suggests that the classifier can reliably detect defects even when they are rare and subtle—an essential requirement in industrial defect detection systems.

Figure 10 displays the confusion matrix obtained from evaluating the proposed model on the test set. The matrix provides a detailed breakdown of correct and incorrect predictions across the two classes: Normal and Reject. Out of 973 normal samples, the model correctly classified 964 and misclassified only 9 as rejects. Similarly, among 277 reject samples, 265 were correctly identified while 12 were incorrectly predicted as normal.

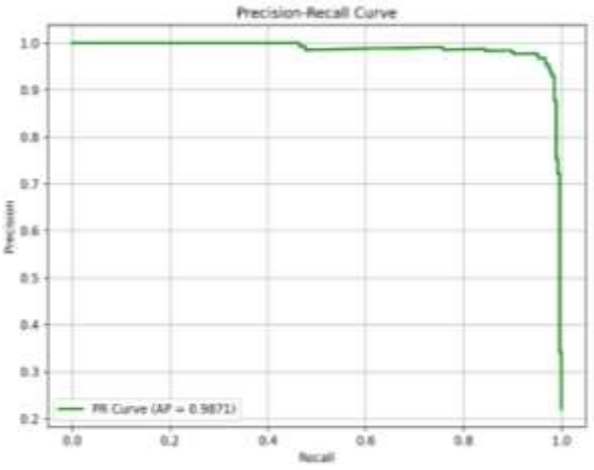

Figure 9. Precision–Recall curve with an AP of 0.9871, highlighting the model's effectiveness in distinguishing classes under imbalanced data conditions.

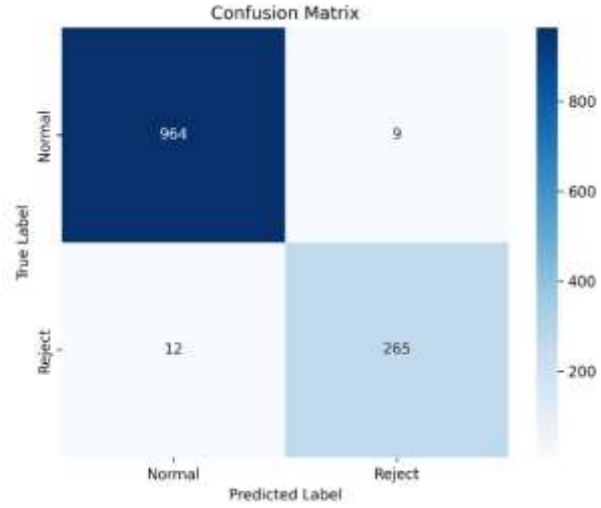

Figure 10. The confusion matrix shows strong performance on both classes, with minimal false positives and false negatives—especially important for reliable reject detection.

This confusion matrix serves as the foundation for computing various evaluation metrics:

- Accuracy:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \approx 0.9832 \quad (13)$$

- Precision (for Reject class):

$$Precision = \frac{TP}{TP + FP} \approx 0.9671 \quad (14)$$

- Recall (Sensitivity):

$$Recall = \frac{TP}{TP + FN} \approx 0.9567 \quad (15)$$

- F1-Score:

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \approx 0.9619 \quad (16)$$

These results clearly indicate that the model maintains a high true positive rate for both classes while minimizing false alarms and missed detections. Particularly, the low false negative count (12 misclassified rejects) is critical in industrial contexts where failing to detect a defective blank could lead to downstream process failures or quality issues.

5. Discussion

The experimental results demonstrate that the proposed deep learning framework, combining segmentation with classification, is highly effective in identifying surface defects on copper blanks in an industrial electrorefining setting. The model achieved strong performance across multiple evaluation metrics, including accuracy, AUC, precision, recall, and F1-score, confirming its robustness in handling class imbalance and subtle visual anomalies. In particular, the integration of YOLOv8-based segmentation significantly improved classification reliability by eliminating background noise and focusing the classifier on relevant regions. However, to better understand the strengths and limitations of the model, it is essential to explore specific case examples, especially those involving misclassifications. In the following sections, we analyze a series of representative correct and incorrect predictions to gain insight into the model's behavior in real-world scenarios.

Figure 11 provides visual examples of misclassified samples, offering critical insight into the model's limitations. The top row shows instances of False Reject, where visually acceptable blanks were mistakenly predicted as defective. In these cases, factors such as surface texture irregularities, minor lighting artifacts, or subtle background shadows may have contributed to false positive decisions. For example, the third sample exhibits superficial blotches and oxidation stains that might have been interpreted as defect patterns by the classifier. The bottom row presents False Normal predictions, where genuinely defective blanks were incorrectly classified as normal. These samples often include edge-related anomalies or thin copper bridging near the top windows—features that may be either too fine or too localized for the model to detect reliably. In some instances, the uniformity of copper deposition in central regions may have outweighed the peripheral defects in the model's decision-making. These examples highlight the inherent challenge of classifying industrial surface anomalies, particularly in the presence of diverse background patterns, lighting variations, and fine-

grained defects. They also emphasize the need for continued refinement of segmentation accuracy and consideration of spatial context in future model iterations.

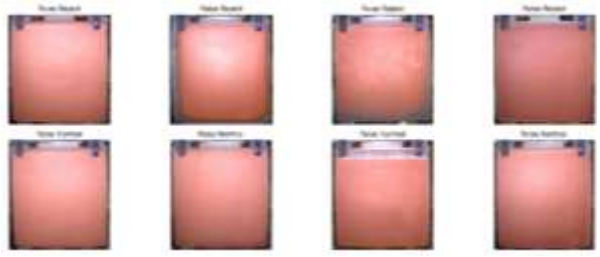

Figure 11. Examples of Misclassified Blanks: False Rejects (Top) and False Normals (Bottom).

Figure 12 illustrates the Grad-CAM visualization for a test sample, highlighting the spatial attention of the EfficientNetV2-S classifier during prediction. The yellow regions denote areas of high activation, indicating zones that strongly influenced the model's decision, while blue and purple zones had minimal impact.

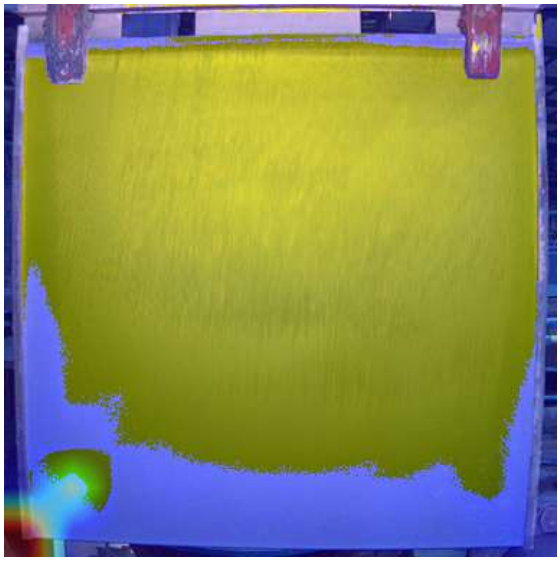

Figure 12. Grad-CAM visualization showing the model's focus on the central copper deposition area and lower edges, confirming attention to critical defect regions during classification.

As shown, the model focuses primarily on the central and upper regions of the copper deposition area, where structural uniformity and deposition texture are most informative for classifying a blank as normal or reject. Interestingly, the activation map avoids background clutter and clamp regions at the top corners, suggesting that the segmentation pre-processing step successfully guided the model's attention toward the relevant copper surface. However, some dispersed activation at the lower corners, particularly in the left area with faint discoloration, indicates potential sensitivity to peripheral noise or artifact. This type of visualization not only supports trust in the model's interpretability but also confirms that the network has learned to prioritize the correct physical zones of interest—especially useful in applications where visual cues may be subtle and context-dependent.

Table 1 provides a comparative analysis of various deep learning architectures evaluated on the copper blank surface defect dataset. The results highlight the performance of five baseline models (MobileNetV2, MobileNetV3 Small, ResNet101, Swin-Tiny, and DenseNet121) against the proposed model based on EfficientNetV2-S. Given the industrial context of this study, where real-time deployment and accurate identification of rare reject cases are both critical, multiple evaluation metrics were considered. While some architectures such as ResNet101 and DenseNet121 showed strong recall or F1-score values, they typically required higher computational resources. In contrast, MobileNetV3 Small, though lightweight, failed to capture fine-grained defect features effectively, resulting in relatively poor recall (0.7040). The proposed EfficientNetV2-S model achieved the highest overall performance, with an accuracy of 98.32%, precision of 96.71%, recall of 95.67%, F1-score of 96.19%, and an AUC of 0.9953. These results confirm its superior capability in detecting subtle reject patterns, while maintaining a lightweight architecture suitable for industrial environments with limited computational capacity.

Table 1. EfficientNetV2-S delivers the best F1-score and AUC with superior efficiency.

Model	Acc	Pre	Recall	F1	AUC
MobileNetV2	0.9712	0.9617	0.9061	0.9331	0.9926
MobileNetV3Small	0.8960	0.8025	0.7040	0.7500	0.8905
ResNet101	0.9800	0.9468	0.9639	0.9553	0.9932
Swin-Tiny	0.9328	0.9214	0.7617	0.8340	0.8716
DenseNet121	0.9768	0.9559	0.9386	0.9472	0.9916
Proposed Model	0.9832	0.9671	0.9567	0.9619	0.9953

In particular, high recall was a priority in this study due to the critical need to minimize false negatives—i.e., mistakenly classifying a defective cathode as normal, which could lead to downstream processing failures. EfficientNetV2-S effectively balances model complexity with performance, making it a reliable choice for deployment on smart inspection systems within copper refineries.

6. Conclusion

This study proposes a deep learning pipeline combining YOLOv8 segmentation and EfficientNetV2-S classification for automated defect detection on copper cathode blanks at the

Sarcheshmeh Copper Refinery. By effectively isolating critical regions, the system ensures robustness against industrial background noise and surface variability. The results demonstrate strong classification performance, achieving 98.32% accuracy, 96.71% precision, 95.67% recall, and an AUC of 0.9953 on the test set. These metrics confirm the system's effectiveness in distinguishing subtle surface anomalies that could otherwise compromise the stripping process or damage production equipment. Techniques such as class weighting and focal loss were essential in addressing the inherent class imbalance, ensuring high sensitivity to the underrepresented "reject" category. Visual analyses using confusion matrices, ROC and PR curves, and Grad-CAM heatmaps further validate the system's interpretability and generalization capability. In particular, the Grad-CAM results highlight that the classifier correctly focuses on defect-prone regions while ignoring irrelevant structures. Overall, the system offers a scalable, accurate, and interpretable solution for automated quality inspection in metal refining lines. Future work may focus on integrating multi-modal inputs (e.g., thermal imaging or depth sensing) and utilizing multi-camera setups to achieve a more comprehensive view of the blank edges. Additionally, real-time edge deployment or unsupervised anomaly detection could be explored to further improve adaptability and minimize manual intervention.

References

[1] J. Hait, R. Jana, and S. Sanyal, "Processing of copper electrorefining anode slime: A review," *Mineral Processing and Extractive Metallurgy*, vol. 118, no. 4, pp. 240–252, 2009.

[2] R. Moskalyk and A. Alfantazi, "Review of copper pyrometallurgical practice: today and tomorrow," *Minerals Engineering*, vol. 16, no. 10, pp. 893-919, 2003.

[3] A. Artzer, M. Moats, and J. Bender, "Removal of antimony and bismuth from copper electrorefining electrolyte: Part I—A review," *JOM*, vol. 70, no. 10, pp. 2033–2040, Oct. 2018.

[4] J. Djokić, A. M. Alfantazi, R. Moskalyk, and M. Moats, "Influence of electrolyte impurities from e-waste electrorefining on copper extraction recovery," *Metals*, vol. 11, no. 9, Art. No. 1383, Sep. 2021.

[5] P. T. Smelting, "Electrolytic refining (ISA process)," May 2025. [Online]. Available: http://www.ptsmelting.com/e-re-isa.html [Accessed: May 29, 2025].

[6] J. Lu, Y. Wang, Z. Li, X. Zhang, and H. Liu, "Effect of rapid hollow cathode plasma nitriding treatment on corrosion resistance and friction performance of AISI 304 stainless steel," *Materials*, vol. 16, no. 24, Art. No. 7616, Dec. 2023.

[7] K. R. Ahmed, "DSteelNet: A real-time parallel dilated convolutional neural network with atrous spatial pyramid pooling for detecting and classifying defects in surface steel strips," *Sensors*, vol. 23, no. 1, Art. No. 544, Jan. 2023.

[8] Y. Xian, H. Zhang, Z. Liu, J. Wang, and L. Li, "YOT-Net: YOLOv3 combined triplet loss network for copper elbow surface defect detection," *Sensors*, vol. 21, no. 21, Art. No. 7260, Nov. 2021.

[9] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, Jun. 2017.

[10] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv: 1409.1556*, 2014.

[11] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA, pp. 1–9, 2015.

[12] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, pp. 770–778, 2016.

[13] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, pp. 4700–4708, 2017.

[14] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "SSD: Single Shot MultiBox Detector," in *Computer Vision – ECCV 2016* (Lecture Notes in Computer Science, vol. 9905), pp. 21–37, 2016.

[15] R. Wei and Y. Bi, "Research on recognition technology of aluminum profile surface defects based on deep learning," *Materials*, vol. 12, no. 10, Art. no. 1681, 2019.

[16] B. Hu and J. Wang, "Detection of PCB surface defects with improved Faster-RCNN and feature pyramid network," *IEEE Access*, vol. 8, pp. 108335–108345, 2020.

[17] W. Zhao, F. Chen, H. Huang, D. Li, and W. Cheng, "A new steel defect detection algorithm based on deep learning," *Computational Intelligence and Neuroscience*, vol. 2021, Art. No. 5592878, 2021.

[18] F. Huang, B.-w. Wang, Q.-p. Li, and J. Zou, "Texture surface defect detection of plastic relays with an enhanced feature pyramid network," *Journal of*

Intelligent Manufacturing, vol. 34, no. 3, pp. 1409–1425, 2023.

[19] X. Chen, J. Lv, Y. Fang, and S. Du, "Online detection of surface defects based on improved YOLOV3," *Sensors*, vol. 22, no. 3, Art. No. 817, 2022.

[20] Y. Xie, W. Hu, S. Xie, and L. He, "Surface defect detection algorithm based on feature-enhanced YOLO," *Cognitive Computation*, vol. 15, no. 2, pp. 565–579, 2023.

[21] H. Huang, X. Tang, F. Wen, and X. Jin, "Small object detection method with shallow feature fusion network for chip surface defect detection," *Scientific Reports*, vol. 12, no. 1, Art. no. 3914, 2022.

[22]B. Fan and W. Li, "Application of GCB-net based on defect detection algorithm for steel plates ",*Research square* ,2022.

[23] Y. Ma, J. Yin, F. Huang, and Q. Li, "Surface defect inspection of industrial products with object detection deep networks: a systematic review," *Artificial Intelligence Review*, vol. 57, no. 12, 2024.

[24] Z. Zhang, X. Huang, D. Wei, Q. Chang, J. Liu, and Q. Jing, "Copper nodule defect detection in industrial processes using deep learning," *Information*, vol. 15, no. 12, Art. No. 802, 2024.

[25] G. Zhang, T. Chen, and J. Wang, "CSC-YOLO: An image recognition model for surface defect detection of copper strip and plates," *Journal of Shanghai Jiaotong University (Science)*, vol. 30, pp. 1037–1049, 2025.

[26] H. Zhao, J. Liu, X. Liu, Y. Shi, and Y. Qiao, "LSD-YOLOv5: A steel strip surface defect detection algorithm based on lightweight network and enhanced feature fusion mode," *Sensors*, vol. 23, no. 14, Art. no. 6558, 2023.

[27] C. Zhao, Y. Liu, H. Zhang, J. Wang, and X. Li, "RDD-YOLO: A modified YOLO for detection of steel surface defects," *Measurement*, vol. 214, Art. No. 112776, 2023.

[28] J. Shi, J. Yang, and Y. Zhang, "Research on steel surface defect detection based on YOLOv5 with attention mechanism," *Electronics*, vol. 11, no. 22, Art. no. 3735, 2022.

[29] Z. Guo, Y. Liu, Y. Zhang, X. Wang, and J. Sun, "MSFT-YOLO: Improved YOLOv5 based on transformer for detecting defects of steel surface," *Sensors*, vol. 22, no. 9, Art. no. 3467, 2022.

[30] L. Li, Y. Wang, X. Zhao, Z. Liu, and H. Chen, "The bearing surface defect detection method combining magnetic particle testing and deep learning," *Applied Sciences*, vol. 14, no. 5, Art. no. 1747, 2024.

[31] Y. Xia, J. Xiao, and Y. Weng, "Surface defect detection of polarizer based on improved Faster R-CNN," *Optical Techniques*, vol. 47, no. 6, pp. 695–702, 2021.

[32] Y. Xian, H. Zhang, Z. Liu, J. Wang, and X. Li, "An EA-based pruning on improved YOLOv3 for rapid copper elbow surface defect detection," *Engineering Applications of Artificial Intelligence*, vol. 123, Art. No. 106412, 2023.

[33] B. Zhou, H. Chen, J. Luo, P. Li, B. Xiang, and K. Li, "AEB-YOLO: An efficient multi-scale defect detection algorithm for copper strips," *AIP Advances*, vol. 15, no. 9, Art. No. 095310, 2025.

[34] L. Zhang, Z. Wang, Y. Ma, and G. Li, "Steel surface defect detection algorithm based on improved YOLOv10," *Scientific Reports*, vol. 15, Art. No. 32827, 2025.

[35] J. Cao, Z. Wang, Y. Wang, X. Ma, and H. Li, "A graph-based approach for module library development in industrialized construction," *Computers in Industry*, vol. 139, Art. No. 103659, 2022.

[36] Y. Fang, L. Sun, Z. Wang, and J. Chen, "Modulation of porphyrin photoluminescence by nanoscale spacers on silicon substrates," *Applied Surface Science*, vol. 285, pp. 572–576, 2013.

[37] D. Ma, Y. Liu, X. Chen, J. Zhang, and Z. Wang, "Multi-sensing signals diagnosis and CNN-based detection of porosity defect during Al alloys laser welding," *Journal of Manufacturing Systems*, vol. 62, pp. 334–346, 2022.

[38] M. Bellaoui, K. Bouchouicha, and I. Oulimar, "Estimation of daily global solar radiation based on MODIS satellite measurements: The case study of Adrar region (Algeria)," *Measurement*, vol. 183, Art. No. 109802, 2021.

[39] M. Tan and Q. V. Le, "EfficientNetV2: Smaller models and faster training," in *Proceedings of the International Conference on Machine Learning (ICML)*, pp. 10096–10106, 2021.

[40] S. A. Amiri and Z. Davoudi, "Enhanced Deep Learning Approaches for Wildfire Detection Using Satellite Imagery," *Journal of AI and Data Mining*, vol. 13, no. 4, pp. 491–500, Oct. 2025.

تشخیص خودکار عیوب سطحی در صفحات مسی کاتدی با استفاده از قطعه‌بندی YOLOv8 و طبقه‌بندی EfficientNetV2-S

حسین قیومی‌زاده[۱،*]، علی فیاضی[۱]، خسرو رضایی[۲]، افسانه امینایی[۳]، هادی حلاوتی[۴]، مهدی طاهرنژاد[۵]، هادی معمارزاده[۶]، علی معصومی[۶]، محمدصادق جعفری[۵]

۱ گروه مهندسی برق، دانشگاه ولی‌عصر (عج) رفسنجان، رفسنجان، ایران.

۲ گروه مهندسی پزشکی، دانشگاه میبد، میبد، ایران.

۳ مدیر امور آموزش و توسعه شایستگی، مجتمع مس سرچشمه، رفسنجان، ایران.

۴ واحد تحقیق و توسعه، مجتمع مس سرچشمه، رفسنجان، ایران.

۵ تحقیقات فنی و مهندسی، امور تحقیق و توسعه، مجتمع مس سرچشمه، رفسنجان، ایران.

۶ امور پالایشگاه و ریخته‌گری، مجتمع مس سرچشمه، رفسنجان، ایران.

ارسال ۲۰۲۵/۰۹/۲۰؛ بازنگری ۲۰۲۵/۱۲/۱۹؛ پذیرش ۲۰۲۵/۰۱/۰۳

چکیده:

در این پژوهش، یک سامانه هوشمند مبتنی بر یادگیری عمیق برای تشخیص خودکار عیوب سطحی در صفحات مسی کاتدی مورد استفاده در فرایند الکتروریفاینینگ ارائه شده است. چارچوب پیشنهادی شامل یک مدل قطعه‌بندی مبتنی بر YOLOv8 برای جداسازی نواحی مهم صفحه و یک شبکه طبقه‌بندی EfficientNetV2-S برای تحلیل و تشخیص نهایی عیوب می‌باشد. مرحله قطعه‌بندی با استخراج نواحی مسی، نوارهای لبه و بخش‌های مستعد عیب، اثر پس‌زمینه و نویزهای محیطی را کاهش داده و دقت طبقه‌بندی را افزایش می‌دهد. مجموعه‌داده مورد استفاده شامل ۵۲۶۶ تصویر برچسب‌گذاری شده از محیط صنعتی واقعی بوده که دارای عدم‌توازن کلاس‌ها است. برای رفع این چالش از روش‌های وزن‌دهی کلاس‌ها و تابع هزینه فوکال استفاده شده است. نتایج تجربی روی مجموعه آزمون نشان می‌دهد که مدل پیشنهادی به دقت ۹۸/۳۲ درصد، دقت مثبت ۹۶/۷۱ درصد، بازخوانی ۹۵/۶۷ در صد، امتیاز F1 برابر ۹۶/۱۹ در صد و سطح زیر منحنی ۰/۹۹۵۳ دست یافته است. تحلیل‌های بصری مبتنی بر Grad-CAM نیز بیانگر تمرکز مدل بر نواحی معنادار عیب و پایداری آن در برابر تغییرات نور و شرایط محیطی است. نتایج حاصل، کارایی روش پیشنهادی را برای کنترل کیفیت خودکار در خطوط صنعتی پالایش مس تأیید می‌کند.

کلمات کلیدی: الکتروریفاینینگ مس، تشخیص عیوب سطحی، طبقه‌بندی تصاویر صنعتی، EfficientNetV2-S، یادگیری عمیق در کنترل کیفیت، بینایی ماشین در متالورژی.