

A Hardware–Software Interactive Inspection Framework for Automated Copper Cathode Blank Defect Detection and Robotic Sorting

Hossein Ghayoumi Zadeh^{1✉}, Ali Fayazi¹, Khosro Rezaee², Mehdi Tahernejad³, Hadi Memarzadeh⁴, Ali Masoumi⁵, Saeed Khodadadi⁶, Mohammad Sadegh Jafari³

1. Corresponding author, Associate Professor, Department of Electrical Engineering, Vali-e-Asr University of Rafsanjan, Rafsanjan, Iran. Email: h.ghayoumizadeh@vru.ac.ir
2. Associate Professor, Department of Biomedical Engineering, Meybod University, Meybod, Iran. Email: a.fayazi@vru.ac.ir
3. Associate Professor, Technical & Engineering Unit, R&D Department, Sarcheshmeh Copper Complex, Rafsanjan, Iran. Email: khosrorezaee@meybod.ac.ir
4. Head of Operations, Refinery and Casting Division, Sarcheshmeh Copper Complex, Rafsanjan, Iran. Email: mehditahernejad@nicico.com
5. Expert, Refinery and Casting Division, Sarcheshmeh Copper Complex, Rafsanjan, Iran. Email: alimasoumi@nicico.com
6. Technical Safety Supervisor of Mines, Health, Safety and Environment (HSE), Sarcheshmeh Copper Complex, Rafsanjan, Iran. Email: saeedkhodadadi@nicico.com

Article Info	ABSTRACT
Article type: Research Article Article history: Received 9 May 2026 Received in revised form 19 June 2026 Accepted 16 September 2026 Published online 1 October 2026 Keywords: copper electrorefining, copper cathode blank, automated visual inspection, robotic sorting, MobileNetV4, local-global attention	Automated inspection and robotic sorting of copper cathode blanks are essential for improving quality control in electrorefining production lines, where surface defects can disturb stripping operations, damage reusable stainless-steel blanks, and increase downtime. This study proposes a hardware–software interactive inspection framework for automated copper cathode blank defect detection and robotic sorting in a real production environment at the Sarcheshmeh Copper Complex in Iran. The system integrates an industrial camera, a Time-of-Flight distance sensor for cathode-to-camera position verification, a real-time monitoring interface, a deep learning-based inspection module, and a PLC-based output mechanism for sending reject commands to the robotic sorting unit. In the proposed pipeline, YOLOv8 segmentation first isolates defect-relevant regions, and the segmented image is then classified using a lightweight MobileNetV4 network, enhanced with Local–Global Attention, which combines CBAM-based local refinement and MHSA-based global contextual modeling. To address class imbalance, an alpha-balanced focal loss was used during training. The dataset included 5,489 labeled images collected from the real electrorefining production line. On the independent test set, the proposed framework achieved 98.5% accuracy, 97.6% precision, 96.2% recall, 96.9% F1-score, 0.992 AUC, and 0.982 average precision. The results demonstrate that the proposed framework provides an accurate, deployment-oriented solution for automated inspection and robotic sorting of copper cathode blanks.

Cite this article: Zadeh, H. & et al, (2026), Risk Assessment Model for Software Development Projects (A Case Study for Iranian Messaging Applications). *Journal of Engineering Management and Soft Computing*, 12 (4). 83-107.

DOI: <https://doi.org/10.22091/jemsc.2026.15723.1367>



© The Author(s) retains the copyright.

Publisher: University of Qom

DOI: <https://doi.org/10.22091/jemsc.2026.15913.1378>

1) Introduction

Copper electrorefining is a key industrial process for producing high-purity copper, in which impure copper anodes are dissolved in an electrolytic medium and copper ions are selectively deposited onto reusable stainless-steel cathode blanks (Liu et al., 2024). The quality and uniformity of this deposition directly affect the efficiency of subsequent production stages, particularly the stripping operation, where the deposited copper layer must be separated cleanly from the blank surface. Although modern electrorefining plants benefit from optimized electrolyte composition, improved cell design, and controlled operating parameters, surface irregularities on cathode blanks remain a persistent operational challenge. Localized over-deposition, copper bridging near structural openings, abnormal growth around edge regions, and incomplete deposition can interfere with stripping, increase adhesion between the copper layer and the blank, and eventually, lead to mechanical damage or production delays (Joukar et al., 2025; Sahlman et al., 2021).

In copper refineries, stainless-steel blanks are expected to tolerate repeated electrorefining and stripping cycles. Therefore, maintaining their structural integrity is essential for stable production and cost control (Monfaredi et al., 2022). Even limited damage to sensitive regions of the blank, such as edge strips or upper window areas, can reduce reusability, increase replacement costs, and compromise the continuity of the production line (Lu et al., 2023). Figure 1 illustrates the general electrorefining mechanism, including anode dissolution and cathode deposition.

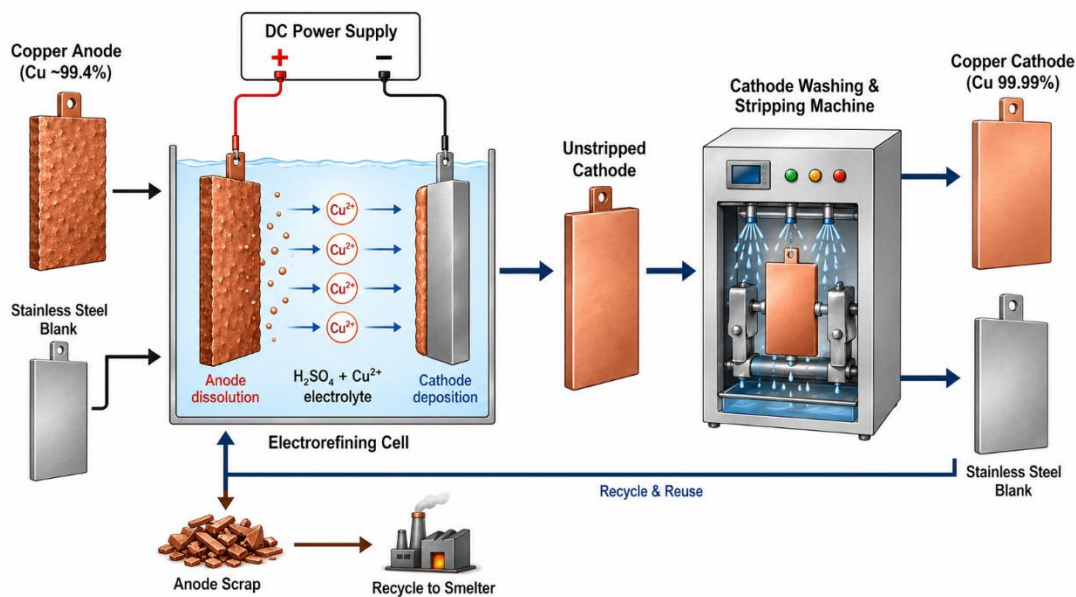


Figure 1. Simplified Schematic of the Copper Electrorefining Process, Showing Anode Dissolution and Cathode Deposition

The present study is motivated by quality-control challenges observed in the copper electrorefining line of the Sarcheshmeh Copper Complex in Rafsanjan, Iran. In this industrial environment, the inspection of copper cathode blanks has traditionally relied on operator-based visual assessment before defective blanks enter downstream operations. Although manual visual inspection is simple and widely used, it is inherently affected by operator fatigue, subjective judgment, variable illumination, electrolyte vapor, heat, and chemically aggressive production-line conditions. These factors can reduce inspection consistency and increase the probability of both missed defects and the

unnecessary rejection of acceptable blanks. More importantly, undetected defective blanks may continue through the production process, where they can cause stripping problems, equipment malfunction, reduced product quality, and increased downtime (Lins et al., 2023).

To address these limitations, an automated inspection system must not only classify images accurately but also operate reliably as part of a real manufacturing workflow. In practice, this requires a coordinated hardware-software structure, including stable image acquisition, Time-of-Flight (ToF)-based position verification, real-time monitoring (Goftari et al., 2026; Hosseini et al., 2026), decision generation, and communication with production-line actuators through a programmable logic controller (PLC) (Bagheri, 2025). In the present work, a hardware-software inspection setup was developed for the refinery unit, consisting of an industrial camera for image acquisition, a ToF distance sensor for verifying the cathode-to-camera position, an interactive software interface for real-time monitoring, and a PLC-based output mechanism for sending reject commands to the robotic sorting unit. This integration enables the proposed system to move beyond offline defect classification and function as an interactive inspection framework for automated quality control and robotic sorting.

Recent advances in computer vision and deep learning have created new opportunities for automated inspection in harsh industrial environments (Eskandari et al., 2025). Deep models can learn discriminative visual patterns from production data and provide rapid, repeatable, and objective decisions. In our previous work, we developed an automated copper blank defect detection framework based on YOLOv8 segmentation and EfficientNetV2-S classification. That study demonstrated that isolating defect-relevant regions before classification can substantially improve robustness against background clutter and illumination variations, achieving strong classification performance on a real industrial dataset (Ghayoumi Zadeh et al., 2026). However, the previous framework mainly focused on validating the segmentation–classification pipeline and used EfficientNetV2-S as the final classifier. Although effective, further investigation is required to determine whether a more lightweight and attention-guided architecture can improve the balance between accuracy, sensitivity to subtle reject patterns, computational efficiency, and suitability for integration into a hardware–software production-line inspection system (Hütten et al., 2024; Zahra & Rezaeenour, 2025).

To address this need, the present work proposes a hardware-software interactive inspection framework for automated copper cathode blank defect detection and robotic sorting. The framework retains YOLOv8-based segmentation for defect-relevant region isolation and replaces the classification stage with a MobileNetV4-based architecture equipped with a local-global attention mechanism.

The proposed classifier is designed to capture fine local abnormalities, such as thin copper bridges, edge-strip misalignment, and small incomplete deposition regions, while simultaneously modeling broader structural relationships across the blank surface. By integrating local feature refinement and global contextual reasoning, the proposed method aims to improve reject detection reliability while maintaining computational efficiency for practical deployment in copper electrorefining quality-control systems.

The main contributions of this study are summarized as follows. First, a hardware–software interactive inspection framework is developed and evaluated in a real copper electrorefining production environment. Second, an industrial image acquisition setup is implemented using an

industrial camera and a Time-of-Flight distance sensor to improve the reliability of cathode blank positioning before image processing. Third, a YOLOv8 segmentation stage is used to isolate defect-relevant regions and reduce background interference under harsh refinery conditions. Fourth, a lightweight MobileNetV4 classifier, enhanced with Local–Global Attention, is proposed to combine local defect-sensitive features with global structural dependencies. Finally, the framework is connected to a PLC-based output mechanism for sending reject commands to the robotic sorting unit, supporting automated decision-making in the production line.

2) Related work

Automated visual inspection has become an important tool in modern industrial quality-control systems, especially in production lines where surface defects can reduce product quality, interrupt downstream operations, and increase maintenance costs. Early inspection approaches were mainly based on manual assessment or handcrafted image-processing techniques, including thresholding, edge detection, texture descriptors, Gabor filters, and wavelet-based analysis (Ahmed et al., 2023; Xian et al., 2021). Although these methods are computationally simple and can perform well under controlled imaging conditions, their reliability often decreases in practical production environments. Changes in illumination, surface reflection, background texture, vibration, dust, and environmental contamination can make defects difficult to distinguish from normal surface variations, particularly when they are small, low-contrast, or embedded in noisy regions (Aghelizadeh et al., 2026).

The limitations of handcrafted methods have encouraged the use of machine learning and deep learning for automated defect inspection. Convolutional neural networks (CNNs) are widely used because they can learn hierarchical and discriminative visual features directly from image data. Early deep learning-based inspection systems often employed classification backbones, such as AlexNet, VGG, ResNet, and DenseNet, to separate defective samples from non-defective ones (Huang et al., 2017; Krizhevsky et al., 2017; Simonyan & Zisserman, 2014). These models improved recognition performance compared with manually designed features; however, image-level classification alone provides limited spatial information. In many industrial applications, a simple normal/defective label is not sufficient because the position, size, and relation of the defect to functional regions of the product may determine its practical impact. This limitation has motivated the use of object detection and segmentation-based approaches for industrial defect analysis.

Two-stage object detectors, particularly Faster R-CNN and its variants, have been investigated extensively for surface defect inspection. These models usually provide strong localization accuracy by first generating candidate regions and then refining their classification. Feature pyramid structures have been used to improve multi-scale defect representation in metal-surface inspection tasks (Wei & Bi, 2019). Similarly, guided region proposal strategies, deformable convolutions, and non-local attention mechanisms have been adopted in PCB, steel strip, and plastic component inspection to improve the detection of small and visually complex defects (Hu et al., 2020). Despite their accuracy, two-stage detectors are generally more computationally demanding and may be less suitable for time-sensitive production-line inspection, especially when continuous monitoring and limited hardware resources are involved.

To satisfy real-time inspection requirements, one-stage detectors, such as SSD and the YOLO family, have received significant attention in industrial defect detection (Liu et al., 2016). These

models predict object classes and locations in a single inference pass, offering a practical balance between localization accuracy and processing speed. Several YOLO-based methods have been adapted for metal and manufacturing inspection. YOT-Net combined YOLOv3 with triplet-loss learning for copper elbow defect detection, improving feature discrimination between defective and normal regions (Krizhevsky et al., 2017; Xian et al., 2021). Other studies have modified YOLOv3 and YOLOv4 using lightweight backbones, multi-scale feature fusion, dense connections, pruning strategies, and attention modules to improve real-time performance while preserving detection accuracy (Xie et al., 2023). These studies show that YOLO-based models are well suited for industrial applications that require both fast inference and the robust localization of defect regions (Pakzad & Pakzad, 2026).

In metal-surface inspection, several studies have specifically focused on copper and steel products. For cathode copper plate inspection, customized YOLO-based models have been proposed to detect small nodules and irregular defects using lightweight backbones and attention mechanisms (Zhang et al., 2024). In copper strip and plate inspection, CSC-YOLO introduced cross-region feature fusion and efficient channel attention to improve small-defect detection while maintaining high inference speed (Zhang et al., 2025). Similar developments have been reported in steel inspection, where improved YOLO variants have incorporated Res2Net backbones, double-FPN structures, CBAM modules, and transformer-based components to enhance fine-grained feature representation (Zhang et al., 2025). These studies demonstrate that domain-specific modifications, including lightweight design, multi-scale fusion, and attention-based refinement, can significantly improve defect detection performance in metallurgical inspection tasks.

Beyond YOLO-based detectors, modern CNN and hybrid architectures have also been used for industrial defect recognition and classification. EfficientNet and EfficientNetV2 have attracted interest because of their strong trade-off between accuracy and computational efficiency and their ability to act as effective feature extractors in industrial vision systems (Li et al., 2024). ResNet-based architectures are still widely used in two-stage inspection frameworks, especially when combined with feature pyramid networks or attention mechanisms for multi-scale representation (Zhao et al., 2023). In addition, deployment-oriented techniques such as pruning, quantization, and compact model design have been explored to support real-time inference on embedded and edge devices (Eslami et al., 2025; Xian et al., 2023). These developments indicate that industrial defect inspection is not only concerned with maximizing accuracy but also with achieving low latency, robustness to environmental variation, and stable performance on imbalanced datasets.

Copper blank inspection in electrorefining environments presents additional challenges compared with many standard surface-defect datasets. Defects are not always isolated texture marks; instead, they may appear as abnormal copper growth near edge strips, thin bridges between upper windows, incomplete deposition in lower regions, or the deformation of the blank structure. The visual appearance of the blank can also be affected by electrolyte vapor, surface oxidation, reflections, and lighting changes. Therefore, the direct classification of full-frame images may confuse true defects with irrelevant background or process-related visual artifacts (Indukuri et al., 2026). Segmentation-guided inspection is particularly useful in this context, as it can isolate defect-relevant regions before final classification.

In our previous work, an automated copper blank defect detection framework was developed using YOLOv8 segmentation followed by EfficientNetV2-S classification (Ghayoumi Zadeh et al., 2026). That study showed that isolating the blank region before classification improves robustness to background interference and enables reliable detection of defective copper blanks under real industrial conditions. However, its classification stage mainly relied on EfficientNetV2-S and did not explicitly combine local defect-sensitive attention with global structural reasoning. Since copper blank defects may be both highly localized, such as edge-strip misalignment and thin copper bridging, and structurally global, such as deformation around the upper window area, a classifier capable of integrating local and long-range contextual information is needed.

Attention mechanisms provide a suitable direction for addressing this requirement. Local attention modules, such as the Convolutional Block Attention Module (CBAM), can improve feature representation by emphasizing informative channels and spatial regions. This is useful for highlighting subtle local patterns such as small deposition irregularities or edge-related defects. In contrast, global attention mechanisms, including Multi-Head Self-Attention (MHSA), can model long-range relationships between distant regions of the feature map and capture broader structural dependencies. Combining local and global attention is therefore well matched to copper blank inspection, where classification decisions often depend on both fine surface cues and the overall geometric consistency of the blank.

Based on these observations, the present study extends the segmentation-guided inspection concept by introducing a MobileNetV4-based classifier enhanced with a Local–Global Attention mechanism. Unlike heavier CNN backbones, MobileNetV4 provides efficient feature extraction and is more suitable for practical deployment in real-time industrial systems. By integrating CBAM for local feature refinement and MHSA for global dependency modeling, the proposed architecture aims to improve sensitivity to subtle reject patterns while maintaining computational efficiency. This design addresses the remaining gap in copper blank defect inspection: the need for a lightweight, attention-guided, and deployment-oriented classifier that can use both localized defect evidence and global structural context under real-world electrorefining conditions.

3) Methodology

3.1) Overall Hardware–Software Inspection Framework

The proposed system is designed as a hardware–software interactive inspection framework for automated defect detection and robotic sorting of copper cathode blanks. As shown in Figure 2, the framework integrates front-view Time-of-Flight (ToF)-based position verification, front-view image acquisition, software-based preprocessing, deep learning-based defect analysis, real-time monitoring, programmable logic controller (PLC)-based decision output, and robotic sorting in a unified production-line workflow. The main objective of this integrated structure is to provide reliable and repeatable inspection under real electrorefining plant conditions while reducing the limitations of operator-based visual assessment.

The workflow begins when a copper cathode blank arrives at the inspection station on the conveyor. A Time-of-Flight (ToF) distance sensor is positioned in front of the blank and directed toward the broad surface of the cathode blank. This sensor verifies whether the blank is located within an acceptable distance range from the imaging system before image acquisition. In the implemented

system, image capture is allowed only when the measured distance remains within a predefined tolerance around the target distance and the reflected signal strength satisfies the acceptance criterion. This step helps prevent image acquisition under unstable positioning, improper cathode placement, or weak sensor feedback.

After position verification, a single-lens industrial camera, also installed in a front-view configuration, captures the surface image of the copper cathode blank. The front-view arrangement of both the ToF sensor and the camera helps maintain a consistent inspection geometry and reduces variations caused by perspective changes. The acquired image is then transferred to the software layer, where basic signal checking, image transfer, and preprocessing operations are performed before deep learning analysis.

In the first stage of the deep learning inspection module, YOLOv8-based segmentation is applied to isolate defect-relevant regions of the copper cathode blank. These regions include copper deposition areas, edge-strip zones, and structurally important surface regions. This segmentation-guided step suppresses irrelevant background information and provides a more standardized representation of the blank surface for the classification stage. In the second stage, the segmented image is analyzed by a lightweight MobileNetV4 classifier enhanced with a Local–Global Attention mechanism. This classifier performs the final Normal/Reject decision by combining local defect-sensitive features with global structural dependencies across the blank surface.

The classification result is displayed through the monitoring software to support real-time visualization and operator supervision. Based on the final inspection decision, the software sends a binary Normal/Reject decision to the PLC. The PLC then generates the corresponding reject command for the downstream robotic sorting unit. If the blank is classified as Normal, it continues through the normal production path; otherwise, the PLC-based reject command activates the robotic sorting unit for separation. This terminology is used consistently throughout the paper as: classification decision -> PLC-based reject command -> robotic sorting unit.

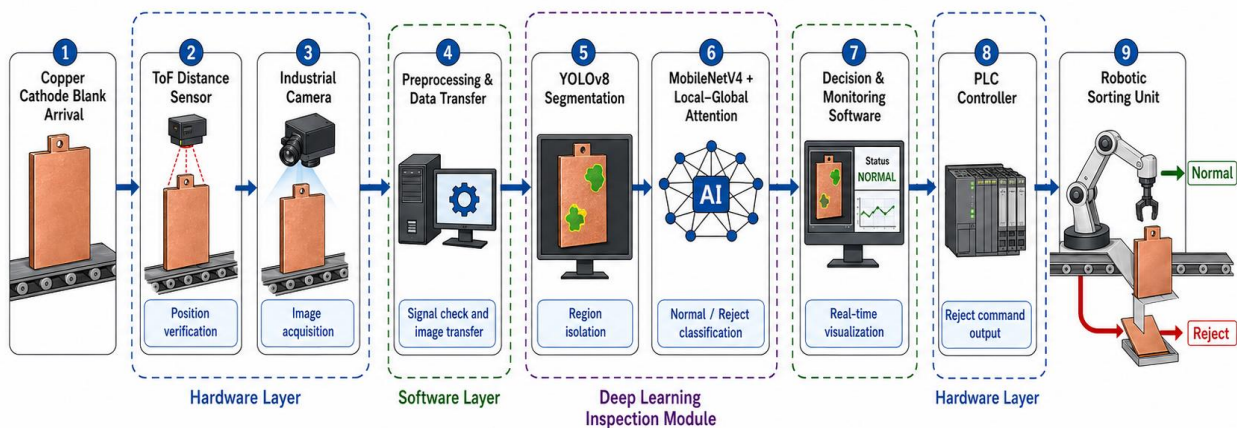


Figure 2. Overall Workflow of the Proposed Hardware-Software Inspection Framework

3.2) Dataset

The experimental evaluation was conducted using an expanded industrial image dataset collected from the copper electrorefining production line at the Sarcheshmeh Copper Refinery, Rafsanjan, Iran. The dataset extends the image collection used in a previous YOLOv8–EfficientNetV2-S study

(Ghayoumi Zadeh et al., 2026) and contains 5,489 real images of copper cathode blanks captured under operational refinery conditions. This expanded dataset was used to evaluate the proposed MobileNetV4–LGA classifier under the same practical imaging environment while providing a larger and a more representative set of Normal and Reject samples.

Unlike laboratory datasets acquired under controlled conditions, the images in this dataset reflect real production variability. The samples include changes in surface appearance, illumination, humidity, electrolyte vapor, copper-surface reflections, and normal fluctuations in the working environment. These conditions make the dataset suitable for assessing the robustness of automated defect detection models intended for deployment in continuous industrial inspection systems.

Image acquisition was performed using a Basler acA1300-200uc industrial camera installed at a fixed inspection point along the production line. The camera position was kept constant during data collection to reduce perspective variation and maintain a consistent front-view of the blank surface. To improve acquisition reliability, the imaging setup was coupled with a Time-of-Flight (ToF) laser distance sensor. The sensor operated at a wavelength of 940 nm, with a Class 1 safety rating, a measurement range of 1 cm to 5 m, and an accuracy of approximately ± 1.5 cm. During acquisition, images were recorded only when the measured distance and reflected signal quality satisfied predefined acceptance criteria. This procedure helped exclude samples affected by unsuitable positioning, unstable geometry, or poor acquisition conditions (Ghayoumi Zadeh et al., 2026; Unal et al., 2026a; Unal et al., 2026b).

Although the camera geometry was fixed, the refinery environment introduced natural visual variations that are important for model generalization. Therefore, images were collected over multiple working periods rather than during a single acquisition session. This strategy allowed the dataset to include different illumination conditions, electrolyte mist, surface oxidation, humidity changes, and process-related visual variations. Such variability is important in copper blank inspection because subtle defect patterns may be visually similar to normal surface changes caused by the production environment.

Each image was assigned to one of two classes: Normal or Reject. The Normal class includes blanks with acceptable copper deposition and no visible defect in the critical regions affecting the stripping process. The Reject class includes blanks with abnormal deposition patterns or structural defects that may disturb downstream handling, reduce stripping quality, or damage the reusable stainless-steel blank. Typical Reject cases include copper growth outside the intended deposition area, copper bridging between upper openings, missing or displaced edge strips, incomplete deposition in lower regions, and mechanical deformation of the blank structure.

To improve reproducibility, image-level Normal/Reject labels were assigned according to production-line quality criteria and were verified by personnel familiar with copper blank stripping problems. Samples were labeled as Normal when the copper deposition pattern and the blank geometry were acceptable for downstream handling. Samples were labeled as Reject when visible abnormalities could disturb stripping or damage the reusable stainless-steel blank, including copper bridging, missing or displaced edge strips, incomplete lower deposition, abnormal copper spread, and mechanical deformation. For the YOLOv8 segmentation stage, pixel-level masks were prepared for defect-relevant blank regions, including the main copper deposition area, edge-strip zones, and

structurally important upper-window regions. Ambiguous cases were reviewed again before inclusion in the final dataset, and the train, validation, and test subsets were kept separate to avoid leakage during model selection and final evaluation.

The final dataset was divided into training, validation, and independent test subsets. The training set contained 2,863 images, including 2,303 Normal and 560 Reject samples. The validation set included 1,241 images, with 919 Normal and 322 Reject samples. The independent test set consisted of 1,385 images, including 1,042 Normal and 343 Reject samples. This split preserves the natural class imbalance observed in the production line, where defective blanks occur less frequently than acceptable ones, and provides a realistic benchmark for evaluating Reject-class detection performance.

Figure 3 illustrates representative examples of the dataset. Figure 3(a) presents a normal sample with a relatively uniform copper layer and acceptable surface condition. Figures 3(b)–3(f) show typical reject cases observed in the production line. Figure 3(b) illustrates copper bridging between the two upper windows, which can interfere with proper separation during stripping. Figure 3(c) shows a missing edge strip, allowing copper to spread toward the side regions and increasing the possibility of adhesion or damage. Figure 3(d) demonstrates incomplete copper deposition in the lower part of the blank. Figure 3(e) presents a case in which the edge strip exists but is misaligned, leading to abnormal copper flow around the protected region. Figure 3(f) shows a mechanically deformed upper window, which may cause physical interference in the stripping machine. These examples highlight the diversity of reject patterns and confirm that the inspection task requires sensitivity to both local surface defects and larger structural abnormalities.

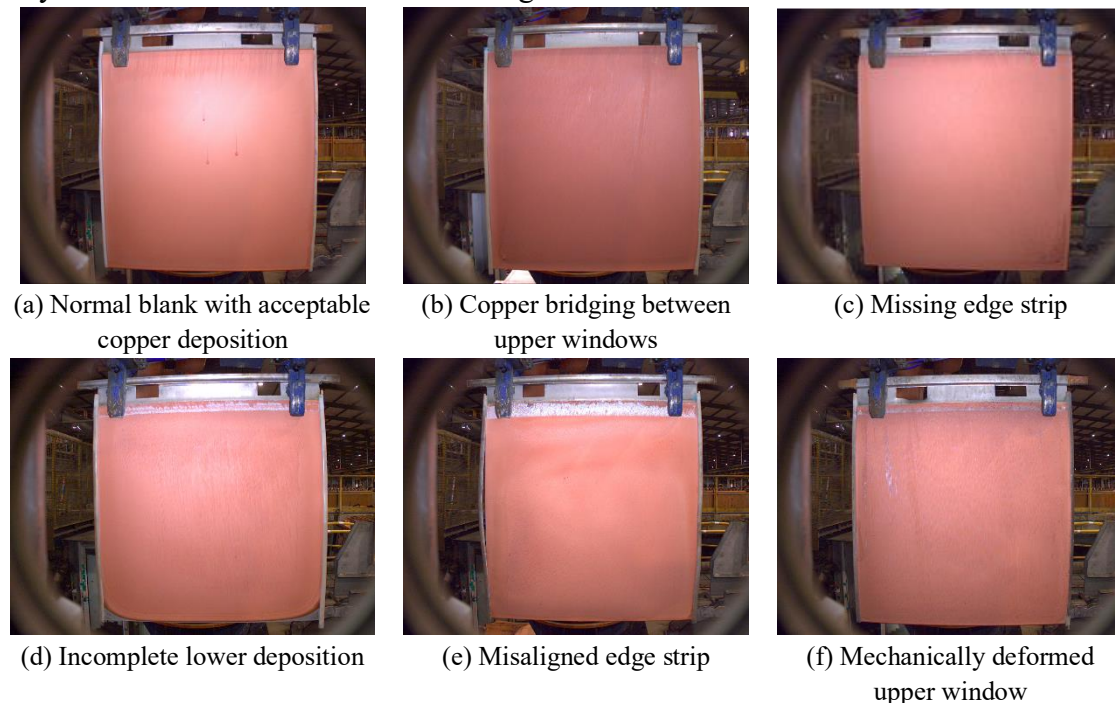


Figure 3. Representative Samples from the Copper Cathode Blank Dataset

3.3) Segmentation-Guided Region Extraction

In the proposed framework, segmentation is used as a region-extraction stage before final defect classification. The purpose of this stage is not to perform the final normal/reject decision, but to

produce a clean and standardized representation of the copper blank by separating the relevant surface areas from the surrounding industrial background. This step is particularly important in copper electrorefining environments, where visual information outside the blank region—such as fixtures, shadows, reflections, stains, and workshop structures—may introduce misleading patterns for the classifier.

Copper blank defects are often spatially dependent. Some reject cases appear near the edge strips, some occur around the upper window regions, and others are related to incomplete copper deposition near the lower part of the blank. Therefore, analyzing the whole camera frame without spatial filtering may reduce classification reliability, especially when the defect is small or located close to the boundary of the deposition area. By extracting only the blank-related regions, the subsequent classifier receives an input image that is more consistent in geometry and less affected by irrelevant background variations.

For this purpose, the YOLOv8-seg model was employed as the segmentation module. YOLOv8-seg is a one-stage instance segmentation architecture capable of predicting object classes, bounding boxes, and pixel-level masks in a single inference process. This makes it suitable for industrial inspection tasks where both localization accuracy and processing speed are important. In the present study, the segmentation model was used to identify the main copper deposition area, edge-strip regions, and other defect-relevant blank components. The generated masks were then used to suppress background regions and construct a refined input for the classification network.

The use of YOLOv8-seg is motivated by three practical considerations. First, it provides mask-level localization rather than only bounding-box detection, allowing the system to preserve the actual shape of the blank and its critical regions. Second, its one-stage design supports fast inference, which is important for online inspection in production lines. Third, it has sufficient robustness for images captured under variable industrial conditions, including surface reflections, electrolyte vapor, and non-uniform illumination. These characteristics make YOLOv8-seg an appropriate preprocessing module for copper blank inspection, where the classifier must focus on subtle deposition abnormalities rather than background artifacts.

In this study, the segmentation stage follows the region-isolation strategy validated in our previous YOLOv8–EfficientNetV2-S framework (Ghayoumi Zadeh et al., 2026). However, in the present work, this stage is used only as a fixed preprocessing component, while the main contribution lies in the proposed MobileNetV4-based Local–Global Attention classifier. The role of YOLOv8-seg is therefore to generate a defect-focused representation of each copper blank by suppressing irrelevant background regions and preserving the critical surface areas required for classification. Given an input image $x \in \mathbb{R}^{H \times W \times 3}$, the YOLOv8-seg module predicts bounding boxes, class probabilities, and pixel-level masks for the target blank regions. Its optimization objective can be expressed as:

$$L_{\text{seg}} = \lambda_{\text{cls}} L_{\text{cls}} + \lambda_{\text{box}} L_{\text{box}} + \lambda_{\text{mask}} L_{\text{mask}} \quad (1)$$

Where L_{cls} is the classification loss, L_{box} indicates the bounding-box localization loss, and L_{mask} represents the mask prediction loss. The coefficients λ_{cls} , λ_{box} , and λ_{mask} balance the contribution of each term during training. In this work, the output masks were not used as the final defect decision;

instead, they were applied to remove background regions and generate standardized inputs for the subsequent classifier.

The segmentation annotations were prepared to cover the defect-relevant blank regions, including the main copper deposition area, edge-strip zones, and structurally important regions around the upper windows. The annotation process was performed at the region level rather than only at the image level because many reject patterns are meaningful only when evaluated in relation to the geometry of the blank. For example, copper bridging, edge-strip misalignment, and incomplete lower deposition are not merely texture changes but spatial abnormalities occurring in specific functional areas. The masks were checked visually after annotation, and incorrectly aligned masks were corrected before training the segmentation module. By isolating these regions, the preprocessing stage reduces the influence of reflections, workshop background, and irrelevant surface artifacts, thereby improving the consistency of the classifier input. Figure 4 shows the effect of segmentation-guided preprocessing. Figure 4(a) presents a raw production-line image containing background clutter and environmental noise, while Figure 4(b) shows the processed output after YOLOv8-based segmentation and cropping. The resulting image provides a cleaner and more standardized view of the copper blank, allowing the proposed MobileNetV4–LGA classifier to focus on defect-relevant visual patterns.

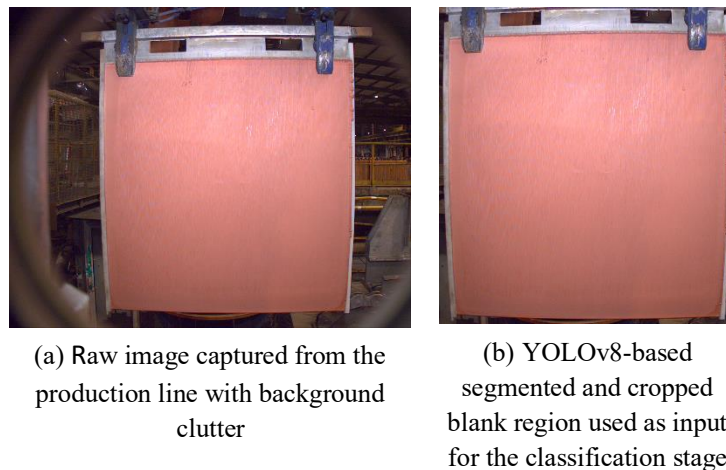


Figure 4. Effect of Segmentation-Guided Preprocessing

3.4) Classification with the Proposed Model

After segmentation-guided preprocessing, the refined blank images are passed to the proposed classification network for the final Normal/Reject decision. Designing an effective classifier for copper blank inspection requires a balance between accuracy, computational efficiency, and robustness to real industrial variations. Although conventional CNNs have achieved strong performance in image classification, copper blank defect detection remains challenging because reject patterns may be highly localized, visually subtle, and strongly dependent on the spatial structure of the blank. Examples include thin copper bridges near the upper windows, edge-strip abnormalities, incomplete lower deposition, and small surface irregularities that may be confused with reflections, oxidation marks, or process-related texture variations.

To address these challenges, this study proposes a MobileNetV4-based classifier enhanced with a Local–Global Attention (LGA) module. In the implementation, MobileNetV4-Hybrid-Large

pretrained on ImageNet-1K was used as the backbone, with an input resolution of 384×384 . The backbone first extracts high-level convolutional feature maps from the segmented blank image. These feature maps are then passed to the LGA module, which combines two complementary attention branches. The local branch uses the Convolutional Block Attention Module (CBAM) to refine channel and spatial features, enabling the network to emphasize small defect-sensitive regions such as edge strips and copper bridging areas. The global branch applies Multi-Head Self-Attention (MHSA) on a downsampled token representation of the feature map to capture long-range structural dependencies across the blank surface. The outputs of the local and global branches are concatenated and fused using a 1×1 convolution, followed by batch normalization and SiLU activation. The refined feature map is then reduced using adaptive average pooling and passed through a multi-layer fully connected classifier. The classifier consists of linear layers with batch normalization, ReLU activation, and dropout regularization, producing a single output logit for binary defect classification. This design enables the model to combine fine local evidence with broader geometric context, which is important for distinguishing true reject cases from visually similar normal variations. To address the natural imbalance between Normal and Reject samples, the model was trained using an improved focal loss with an α -balancing factor. This loss function reduces the contribution of easy samples and increases the relative importance of hard or misclassified examples, helping the classifier focus on subtle reject cases without being dominated by the majority Normal class. In the experiments, the focal loss parameters were set to $\alpha = 0.55$ and $\gamma = 1.65$. Overall, the proposed MobileNetV4–LGA classifier is designed to provide a lightweight and attention-guided solution for copper blank defect detection under real-world electrorefining conditions. As illustrated in Figure 5, the classification pipeline begins with the segmented, cropped blank image as input. MobileNetV4 generates the backbone feature map, refined by the LGA block through local CBAM-based attention and global MHSA-based contextual modeling. The fused representation is then processed by both adaptive pooling and fully connected layers to produce the final defect prediction.

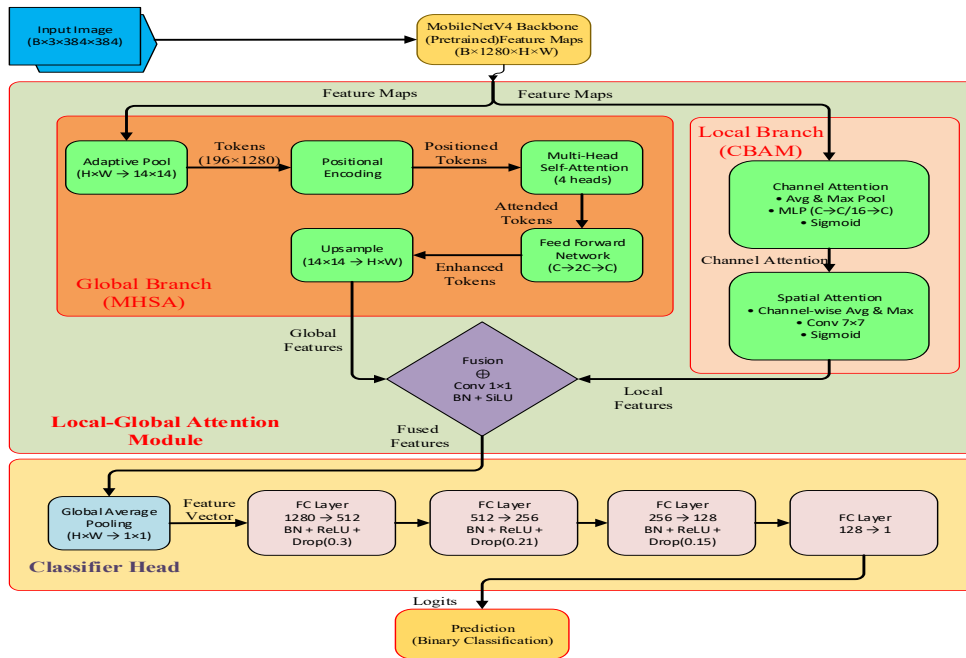


Figure 5. Architecture of the Proposed MobileNetV4–LGA Classifier

3.4.1) Lightweight Backbone (MobileNetV4)

In the proposed classifier, we adopt MobileNetV4 as a lightweight backbone to balance accuracy and latency under industrial constraints. Its efficiency comes primarily from depthwise-separable convolutions and inverted residual (MBConv/fused-MBConv) blocks. According to Qin et al. (2024), for a standard convolution with kernel size $k \times k$, M input channels, N output channels, and feature map of size $F \times F$, the multiply–accumulate cost is:

$$C_{\text{standard}} = k^2 M N F^2 \quad (2)$$

A depthwise-separable convolution decomposes this into a depthwise convolution followed by a pointwise 1×1 projection,

$$C_{\text{sep}} = \underbrace{k^2 M F^2}_{\text{depthwise}} + \underbrace{M N F^2}_{\text{pointwise}} \quad (3)$$

Yielding the reduction ratio

$$\frac{C_{\text{sep}}}{C_{\text{standard}}} = \frac{1}{N} + \frac{1}{k^2} \quad (4)$$

Which, for the common case, $k = 3$ and $N \gg 1$ is approximately $\frac{1}{9}$, i.e., about 8–9× fewer MACs with minimal accuracy loss. Each inverted residual block expands channels by a factor t , filters with a depthwise $k \times k$, then projects back to a linear bottleneck; algebraically,

$$y = \underbrace{\text{PW}_{1 \times 1}^{\text{proj}}}_{\mathbb{R}^M \rightarrow \mathbb{R}^N} \left(\underbrace{\text{DWConv}_{k \times k}}_{\text{per-channel}} \left(\underbrace{\sigma \left(\text{PW}_{1 \times 1}^{\text{expand}} \right)}_{\mathbb{R}^M \rightarrow \mathbb{R}^M} (x) \right) \right) \quad (5)$$

With a residual shortcut $y \leftarrow y + x$ when stride is 1 and the channel dimensions match, improving gradient flow in thin bottlenecks. Early stages may use the fused variant (expansion and spatial convolution combined) to reduce launch overhead on GPUs/NPUs without changing the computational intent. To enhance representational power for fine defects, MobileNetV4 integrates squeeze-and-excitation (SE) channel attention. Given a feature tensor $U \in \mathbb{R}^{H \times W \times C}$, the channel descriptor is computed as $Z_c = \frac{1}{HW} \sum_{i=1}^H \sum_{j=1}^W U_c(i, j)$, followed by a gating vector $s = \sigma(W_2 \delta(W_1 z))$, and the output is rescaled as $\tilde{U}_c = s_c \cdot U_c$. This selectively amplifies informative channels (e.g., thin copper bridges or edge-strip artifacts) while suppressing background. Finally, MobileNetV4 follows the common compound-scaling principle to balance depth/width/resolution across deployment budgets, often parameterized as $d = \alpha \phi$, $w = \beta \phi$, and $r = \gamma \phi$ with a resource constraint (e.g., $\alpha \beta^2 \gamma^2 \approx \text{const}$), allowing us to select an input resolution and model size that meet real-time throughput without sacrificing accuracy. In practice, this backbone provides rich global–local features at a fraction of the cost of heavy CNNs, leaving headroom for our subsequent local–global attention and imbalance-aware training while keeping inference suitable for online copper-blank inspection (Qin et al., 2024).

3.4.2) Local–Global Attention Module

While the lightweight backbone efficiently extracts hierarchical features, copper blank defect detection requires a mechanism that can simultaneously capture localized irregularities and global structural context. Many defects are subtle and spatially limited, such as thin copper bridges or slight

misalignments of the edge strip, whereas others are large-scale or affect the geometry of the entire blank. Conventional CNNs, constrained by fixed receptive fields, may focus disproportionately on either local details or broad patterns, missing critical cross-scale dependencies. To overcome this limitation, the proposed model incorporates a Local–Global Attention module designed to fuse fine-grained spatial information with holistic structural cues.

The intuition behind this module is rooted in the complementary roles of convolution and attention. Convolutions are excellent at capturing local textures and patterns; however, they lack the ability to dynamically relate distant regions of the image. In contrast, attention mechanisms can model long-range dependencies, allowing the network to relate, for example, a defect on the lower edge of the blank with corresponding structural context in the upper windows (Hosseini & Shanechi, 2025). By combining these two paradigms, the Local–Global Attention module ensures that subtle surface variations are not misclassified due to missing global awareness, and that large-scale structural inconsistencies are reinforced by local evidence (Zhao et al., 2021).

Mathematically, let the backbone output be a feature map $F \in \mathbb{R}^{H \times W \times C}$. The local branch applies CBAM-based channel and spatial refinement to emphasize defect-sensitive local regions, such as edge-strip abnormalities and thin copper bridges. This can be expressed as:

$$F_{local} = CBAM(F) \quad (6)$$

In parallel, the global branch reshapes an adaptively pooled feature map into tokens and applies multi-head self-attention to capture long-range dependencies across the blank surface. A typical formulation of scaled dot-product attention is:

$$Attention(Q, K, V) = Softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (7)$$

where Q, K, and V are, respectively, the query, key, and value projections obtained from the tokenized feature representation. The attended tokens are then reshaped and upsampled to the original feature-map resolution to produce F_{global} . Finally, the outputs of the local and global branches are fused using channel-wise concatenation followed by a 1×1 convolution, batch normalization, and SiLU activation:

$$F_{fused} = SILU(BN(Conv_{1 \times 1}([F_{local}, F_{global}]))) \quad (8)$$

where $[.]$ denotes channel-wise concatenation. This fusion strategy allows the model to combine CBAM-refined local features with MHSA-based global contextual features before final classification.

In practice, the Local–Global Attention module is particularly useful for challenging borderline case, as it enables the classifier to consider both localized defect cues and the overall blank geometry. For instance, incomplete copper deposition near the bottom edge may be small in size but critical for classification. Conversely, when copper deposition spreads across multiple windows, the resulting broader structural abnormality can be captured more effectively through the global attention branch.

By integrating this module into the pipeline, the proposed model achieves a more balanced representation that improves classification accuracy under real-world operating conditions. It

effectively mitigates the weaknesses of purely convolutional or purely attention-based designs, providing a robust mechanism for defect detection in copper electrorefining applications.

3.4.3) Imbalance-Aware Learning Strategy

In industrial defect detection tasks, defective samples are usually much less frequent than normal samples. This imbalance is also present in the copper blank dataset used in this study, where the Normal class contains more samples than the Reject class. If this issue is not considered during training, the model may become biased toward the majority class and produce high overall accuracy while still missing important defective cases. In copper electrorefining, such missed reject samples are critical because they may continue to the stripping stage and cause mechanical problems, equipment damage, or production delays.

To reduce this bias, the proposed classifier was trained using an α -balanced focal loss. Compared with standard binary cross-entropy, focal loss gives less importance to easy and correctly classified samples and places more emphasis on hard or misclassified cases. This property is useful in the present application because many reject patterns are subtle and localized, such as thin copper bridging, slight edge-strip displacement, incomplete lower deposition, or abnormal copper growth near critical regions. By emphasizing these difficult samples during training, the model becomes more sensitive to the minority Reject class without simply over-predicting defects. For binary classification, let $y \in \{0,1\}$ denote the ground-truth label, where $y = 1$ corresponds to the Reject class and $y = 0$ corresponds to the Normal class. Let $p \in (0,1)$ be the predicted probability of the Reject class. The focal loss is expressed as:

$$FL(p_t) = -\alpha_t(1 - p_t)^\gamma \log(p_t) \quad (9)$$

Where p_t is defined as:

$$p_t = \begin{cases} p & y = 1 \\ 1 - p & y = 0 \end{cases} \quad (10)$$

and α_t is defined as:

$$\alpha_t = \begin{cases} \alpha & y = 1 \\ 1 - \alpha & y = 0 \end{cases} \quad (11)$$

In this formulation, alpha controls the relative importance of the positive class, whereas gamma is the focusing parameter. For binary classification, $y = 1$ denotes the Reject class, while $y = 0$ denotes the Normal class. If p is the predicted probability of the Reject class, then $p_t = p$ for $y = 1$, and $p_t = 1 - p$ for $y = 0$. Similarly, $\alpha_t = \alpha$ for $y = 1$, and $\alpha_t = 1 - \alpha$ for $y = 0$. The α -balanced focal loss is therefore written as $FL(p_t) = -\alpha_t(1 - p_t)^\gamma \log(p_t)$. When a sample is classified correctly with high confidence, p_t becomes close to 1 and the modulating factor $(1 - p_t)^\gamma$ reduces its contribution to the total loss. In contrast, misclassified or low-confidence samples receive a larger loss contribution.

In this study, the focal loss parameters were set to $\alpha = 0.55$ and $\gamma = 1.65$ based on validation performance. The value of α gives slightly greater importance to the Reject class, while γ helps the model focus on visually challenging samples. This imbalance-aware learning strategy improves the

model's ability to detect rare defect patterns while maintaining stable performance on the majority Normal class.

3.5) Training Configuration

All experiments were implemented in Python using the PyTorch deep learning framework and the timm model library. The proposed classifier was trained on segmented and cropped copper blank images generated by the YOLOv8-based preprocessing stage. The MobileNetV4-Hybrid-Large backbone pretrained on ImageNet-1K was used, and all input images were resized to 384×384 pixels before being passed to the network.

The dataset was divided into training, validation, and test subsets. The training set was used to optimize the model parameters, the validation set was used for model selection and early stopping, and the test set was kept unseen until the final evaluation. Images were normalized using the standard ImageNet mean and standard deviation values. Data augmentation was applied only to the training set and included small random rotations, slight translation and scaling, brightness variation, and contrast adjustment. These transformations were selected to reflect realistic production-line variations without introducing unrealistic geometric changes.

The model was trained with a batch size of 16 for a maximum of 40 epochs. AdamW was used as the optimizer with an initial learning rate of 5×10^{-5} and a weight decay of 1×10^{-4} . A cosine annealing learning-rate scheduler was applied to improve convergence stability. Dropout regularization was used in the classifier head with a dropout rate of 0.30. For the Local-Global Attention module, the global attention branch used a 14×14 token grid, four attention heads, and a feed-forward expansion factor of 2. The CBAM reduction ratio was set to 16. To address class imbalance, the model was trained using an alpha-balanced focal loss, with $\alpha = 0.55$ and $\gamma = 1.65$. Weighted sampling was not used in the final experiments, so imbalance handling was performed through the focal loss formulation. Automatic mixed precision was enabled when a CUDA-capable GPU was available. Gradient clipping with a maximum norm of 1.0 was applied to improve training stability. The best model was selected based on validation performance using a combined score derived from the area under the receiver operating characteristic curve (AUC) and the optimal F1-score (Maydanchi et al., 2023). Early stopping was applied with a patience of 20 epochs.

For final testing, the best saved model was loaded and evaluated on the independent test set. The binary decision threshold was selected from the validation set to avoid test-set leakage. Final performance was reported using accuracy, precision, recall, F1-score, AUC, average precision, specificity, sensitivity, and the confusion matrix.

4) Results

The proposed MobileNetV4-LGA classifier was evaluated on an independent test set to assess its ability to distinguish Normal and Reject copper blanks under real industrial conditions. The full dataset contained 5,489 images and was divided into training, validation, and test subsets. The training set included 2,863 images, consisting of 2,303 Normal and 560 Reject samples. The validation set contained 1,241 images, including 919 Normal and 322 Reject samples. The independent test set consisted of 1,385 images, with 1,042 Normal and 343 Reject samples. This distribution reflects the natural imbalance of industrial inspection data, where defective samples occur

less frequently than normal samples. To reduce the effect of this imbalance during training, an α -balanced focal loss was used.

Figure 6 shows the training and validation accuracy curves of the proposed model. The training accuracy starts at approximately 72–75% and increases rapidly during the early epochs, reaching about 98–99% after nearly 10–15 epochs. The validation accuracy follows a similar trend and quickly stabilizes around 97%, indicating that the model generalizes well to unseen validation samples. The close behavior of the training and validation curves suggests that the model does not suffer from severe overfitting. This stable convergence can be attributed to the MobileNetV4–LGA architecture, dropout and batch normalization in the classifier head, data augmentation, and the use of α -balanced focal loss. Overall, the accuracy curves demonstrate that the proposed classifier can effectively learn discriminative features for distinguishing Normal and Reject copper blanks.

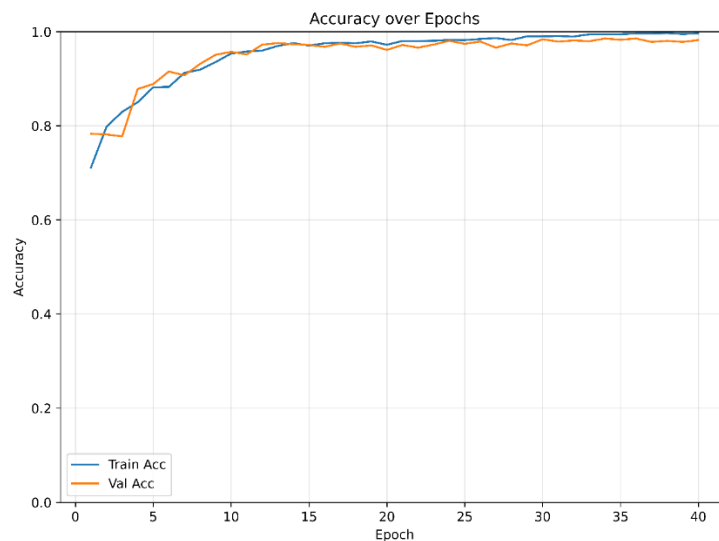


Figure 6. Training and Validation Accuracy Over the First ~40 Epochs Using MobileNetV4 with Local-Global Attention

Figure 7 presents the training and validation loss curves during the training process. The training loss decreases steadily from approximately 0.09 to nearly 0.01, indicating the effective optimization of the proposed MobileNetV4–LGA classifier. The validation loss remains within a low numerical range while exhibiting several fluctuations during training, particularly around the middle epochs. Since the vertical scale of the loss curve is relatively small, these variations mainly reflect changes in prediction confidence rather than large classification errors.

The observed fluctuations in validation loss can be attributed to several factors. First, the dataset is naturally imbalanced, with the Normal class being more frequent than the Reject class. As a result, the composition of validation batches may contain different proportions of difficult or minority-class samples, causing temporary variations in the loss value. Second, the use of α -balanced focal loss increases the contribution of hard or low-confidence samples, which can make the validation loss more sensitive to visually ambiguous reject cases. Third, regularization techniques, such as dropout and weight decay, may also introduce minor variations during training while helping to reduce overfitting.

Despite these fluctuations, the validation loss remains low, and the final test-set metrics confirm stable generalization performance. Therefore, the loss behavior does not indicate overfitting or training instability. Instead, it reflects the sensitivity of the model to challenging samples in an imbalanced industrial dataset. Together with the accuracy curve, the loss trend supports the effectiveness of the proposed training strategy for robust copper blank defect classification. Figure 7 illustrates the training and validation loss curves of the proposed MobileNetV4–LGA classifier, showing steady training convergence and low validation loss with minor fluctuations.

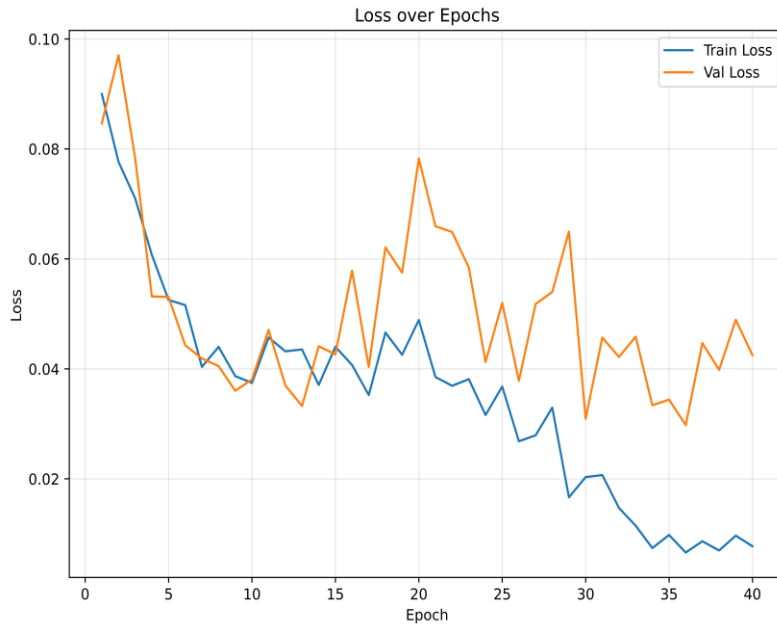


Figure 7. Training and Validation Loss Curves of the Proposed MobileNetV4–LGA Classifier

Figure 8 presents the Receiver Operating Characteristic (ROC) curve and the corresponding Area Under the Curve (AUC) for the proposed MobileNetV4–LGA classifier on the independent test set. The ROC curve illustrates the trade-off between the true positive rate and the false positive rate across different decision thresholds. As shown in the figure, the curve remains close to the upper-left region of the plot, indicating strong discriminative capability between Normal and Reject copper blanks.

The proposed model achieved an AUC of 0.992, confirming its ability to separate the two classes with high reliability. This result is particularly important for the present task because the dataset is imbalanced and the Reject class is less frequent than the Normal class. A high AUC indicates that the classifier maintains robust ranking performance across thresholds rather than depending only on a single fixed decision threshold. In practical industrial inspection, this behavior is valuable because the operating threshold can be adjusted according to production requirements, such as prioritizing defect detection sensitivity or reducing unnecessary false rejections.

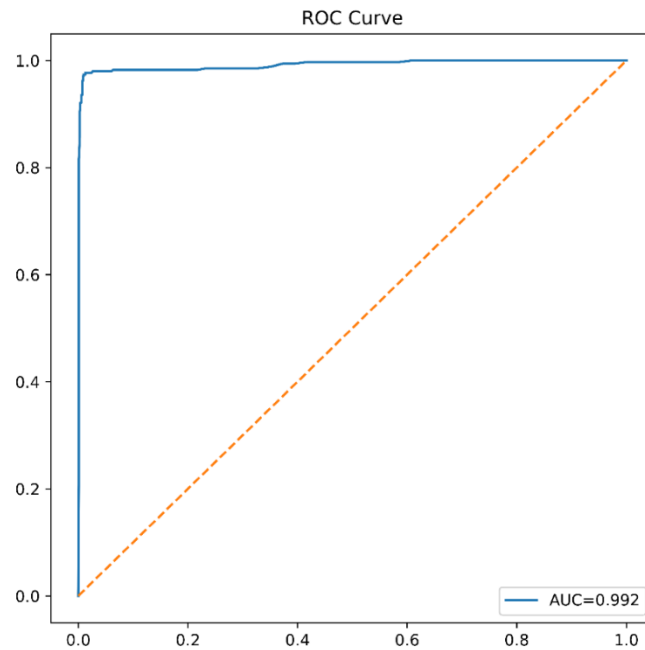


Figure 8. ROC Curve of the Proposed MobileNetV4–LGA Classifier on the Independent Test Set, Achieving an AUC of 0.992

Figure 9 presents the Precision–Recall (PR) curve of the proposed MobileNetV4–LGA classifier on the independent test set. Since the Reject class is less frequent than the Normal class, the PR curve provides an informative evaluation of the model’s ability to detect defective blanks under class imbalance. The curve indicates that the classifier maintains high precision across a wide range of recall values, showing that most defective blanks can be detected while keeping false alarms relatively low.

The Average Precision (AP) value of 0.982 further confirms the strong balance between precision and recall across different decision thresholds. This result is important for industrial deployment because missed reject samples may lead to downstream stripping problems, while excessive false rejections can increase production costs. Therefore, the PR curve demonstrates that the proposed model can provide reliable Reject-class detection while maintaining practical control over unnecessary rejection of normal blanks.

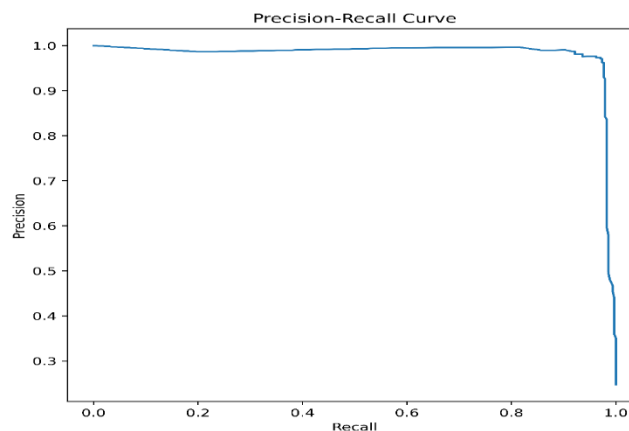


Figure 9. Precision–Recall Curve of the Proposed MobileNetV4–LGA Classifier on the Independent Test Set, Achieving an Average Precision (AP) of 0.982.

Figure 10 shows the confusion matrix of the proposed MobileNetV4–LGA classifier on the independent test set. Class 0 represents Normal blanks, while Class 1 represents Reject blanks. Out of 1,042 Normal samples, 1,034 were correctly classified as Normal, and only 8 were incorrectly classified as Reject, corresponding to a low false-positive rate. In the Reject class, 330 out of 343 defective blanks were correctly detected, while 13 samples were misclassified as Normal. These results demonstrate that the proposed model achieves strong performance for both classes, with high specificity for Normal blanks and high sensitivity for Reject detection. The low number of missed Reject samples is particularly important in industrial inspection, where false negatives may allow defective blanks to proceed to the stripping stage and cause downstream operational problems.

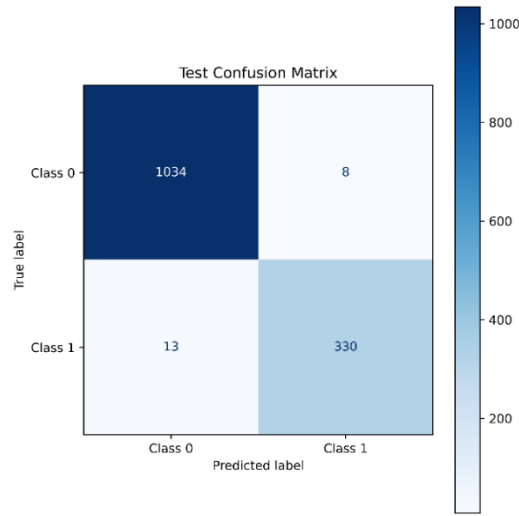


Figure 10. Confusion Matrix of the Proposed MobileNetV4–LGA Classifier on the Independent Test Set

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

- Accuracy:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \approx 0.985 \quad (12)$$

- Precision (for Reject class):

$$Precision = \frac{TP}{TP+FP} \approx 0.976 \quad (13)$$

- Recall (Sensitivity):

$$Recall = \frac{TP}{TP+FN} \approx 0.962 \quad (14)$$

- F1-Score:

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \approx 0.969 \quad (15)$$

These results indicate that the proposed model achieves reliable classification performance for both Normal and Reject samples while maintaining a low number of false alarms and missed detections. In particular, only 13 Reject samples were misclassified as Normal, which is important in industrial inspection because undetected defective blanks may proceed to the stripping stage and cause downstream quality problems, equipment interference, or production delays.

To demonstrate the practical applicability of the proposed system, a dedicated software application was developed and deployed on the copper production line, as shown in Figure 11. The application provides real-time visualization of the captured blank, segmentation-guided inspection output, and final Normal/Reject decision. The final decision is transmitted to the PLC, which issues the reject command to the robotic sorting unit. This implementation confirms that the proposed framework is not limited to offline image classification but can be integrated into an industrial decision-and-actuation workflow.

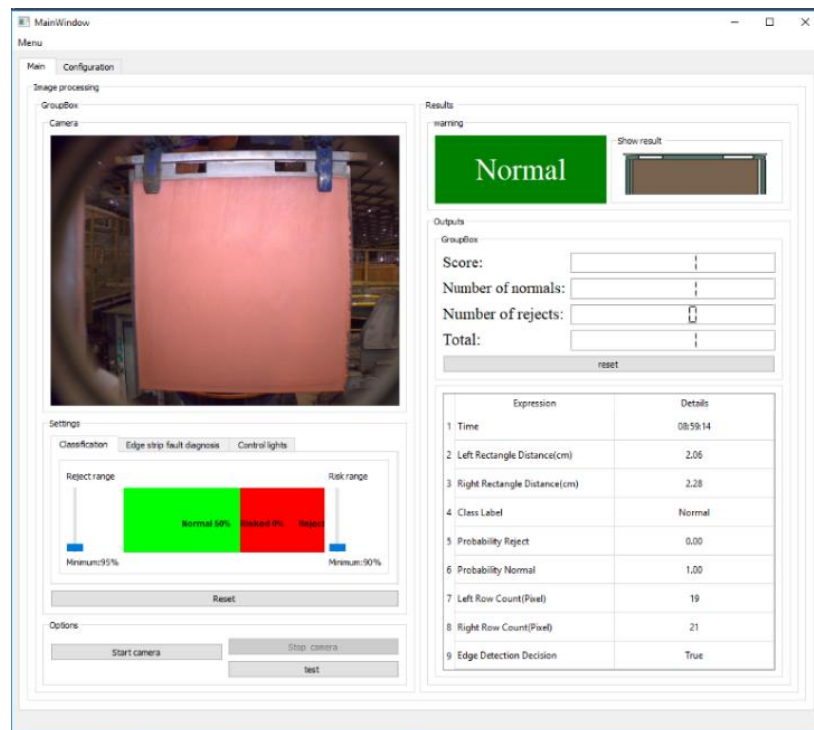


Figure 11. Interface of the Developed Defect Detection Software Deployed on the Copper Production Line

5) Discussion

The experimental results demonstrate that the proposed MobileNetV4–LGA framework is effective for automated surface defect detection in copper cathode blanks under real industrial electrorefining conditions. The combination of YOLOv8-based segmentation and attention-guided classification provides a structured inspection pipeline in which irrelevant background regions are first removed and the classifier is then applied to a cleaner, defect-focused representation of the blank. This design is particularly important for copper blank inspection because many reject patterns are not distributed uniformly across the image; rather, they occur in specific functional regions such as edge strips, upper windows, and lower deposition boundaries. By narrowing the classifier’s attention to these critical regions, the proposed pipeline reduces the influence of background clutter, illumination changes, surface reflections, and workshop artifacts.

The strong test-set performance confirms the suitability of the proposed architecture for this task. The model achieved 98.5% accuracy, 97.6% precision, 96.2% recall, 96.9% F1-score, and 0.992 AUC on the independent test set. These results indicate that the model can reliably distinguish Normal and Reject blanks despite the natural class imbalance in the dataset. In industrial inspection, recall and F1-score are particularly important because false-negative predictions may allow defective blanks to continue to the stripping stage, where they can cause adhesion problems, mechanical interference, or equipment damage. The confusion matrix further supports this observation: among 343 Reject samples in the test set, 330 were correctly detected and only 13 were missed. At the same time, only 8 out of 1,042 Normal samples were incorrectly classified as Reject, indicating that the model also maintains a low false-alarm rate.

A key factor contributing to this performance is the Local–Global Attention mechanism. Copper blank defects may appear either as small localized abnormalities or as broader structural inconsistencies. For example, thin copper bridging near the upper windows or slight edge-strip displacement requires sensitivity to local spatial patterns, while deformation of the blank structure or abnormal deposition distribution requires a wider contextual understanding of the entire blank. The CBAM branch of the proposed LGA module helps emphasize informative local channels and spatial regions, while the MHSA branch captures long-range dependencies across the feature map. By combining these two forms of attention, the classifier can use both fine-grained defect evidence and global geometric context when making the final Normal/Reject decision.

Table 2 compares the proposed model with several baseline architectures, including MobileNetV2, MobileNetV3Small, ResNet101, Swin-Tiny, DenseNet121, and EfficientNetV2-S. The results show that lightweight MobileNet-based models do not necessarily guarantee strong defect recognition unless supported by an appropriate attention mechanism. For example, MobileNetV3Small achieved relatively low recall, suggesting that its compact feature representation was not sufficient for detecting subtle reject patterns in this dataset. Swin-Tiny also showed limited performance, particularly in recall and F1-score, which may be related to the dataset size and the need for the careful adaptation of transformer-based models to this industrial task.

Table 2. Performance Comparison of Baseline Models and the Proposed MobileNetV4–LGA Architecture on the Copper Blank Defect Dataset

Model	Accuracy	Precision	Recall	F1-Score	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
EfficientNetV2-S	0.9832	0.9671	0.9567	0.9619	0.9953
Proposed Model	0.985	0.976	0.962	0.969	0.992

Deeper convolutional models, such as ResNet101 and DenseNet121, performed better than the lightweight baselines, especially in recall and F1-score. However, these models are generally more computationally demanding, which may limit their suitability for deployment in production-line inspection systems where low latency and resource efficiency are important. EfficientNetV2-S provided the strongest baseline performance, achieving 98.32% accuracy, 96.71% precision, 95.67% recall, 96.19% F1-score, and the highest AUC of 0.9953. This confirms that EfficientNetV2-S is a robust architecture for copper blank defect classification and provides a strong reference point for

evaluating the proposed method. Compared with EfficientNetV2-S, the proposed MobileNetV4-LGA model achieved slightly higher accuracy, precision, recall, and F1-score, while its AUC was marginally lower. Specifically, the proposed model improved accuracy from 98.32% to 98.5%, precision from 96.71% to 97.6%, recall from 95.67% to 96.2%, and F1-score from 96.19% to 96.9%, while AUC decreased slightly from 0.9953 to 0.992. This trade-off suggests that the proposed model provides a more favorable operating-point performance for the Reject detection task, particularly in terms of recall and F1-score, which are critical in industrial quality control. The slightly lower AUC indicates that EfficientNetV2-S retains a marginal advantage in threshold-independent ranking performance; however, the proposed MobileNetV4-LGA classifier provides stronger classification performance at the selected operating threshold. Overall, the comparison indicates that the proposed model offers a balanced solution for copper blank defect detection by combining lightweight feature extraction with local and global attention. Rather than relying solely on a heavier backbone, the architecture improves defect sensitivity through targeted feature refinement and contextual modeling. This is especially useful in the present application, where defects may be visually subtle, spatially localized, and embedded in industrial noise. The results suggest that the proposed MobileNetV4-LGA framework is suitable for practical deployment in copper electrorefining inspection systems, particularly when both detection reliability and computational efficiency are required. Nevertheless, several limitations should be acknowledged. First, although the dataset was collected under real production conditions, it was obtained from a specific industrial site and imaging setup. Therefore, further validation on additional copper refineries, different camera configurations, and broader environmental conditions would be valuable for assessing generalization. Second, the current system relies on RGB image data; incorporating complementary sensing modalities such as depth, thermal imaging, or multi-view acquisition may improve detection of defects that are difficult to distinguish from surface texture alone. In addition, longer-term line deployment should be evaluated to quantify maintenance requirements, false-reject cost, missed-defect risk, and the effect of changing illumination or process conditions over time. This discussion highlights that the proposed framework improves the practical balance between classification performance and deployment suitability. The integration of segmentation-guided preprocessing, MobileNetV4-based feature extraction, Local-Global Attention, and α -balanced focal loss enables robust detection of rare and subtle reject cases while maintaining strong performance on Normal samples. Therefore, the proposed approach can serve as a promising foundation for automated quality-control systems in copper electrorefining lines.

6) Conclusion

This study presented a hardware-software interactive inspection framework for automated copper cathode blank defect detection and robotic sorting in an electrorefining production line. The proposed system was developed with emphasis on real industrial deployment at the Sarcheshmeh Copper Refinery in Iran, where blank inspection had traditionally relied on operator-based visual assessment and where surface defects can interfere with stripping operations, damage reusable stainless-steel blanks, and reduce production efficiency. The framework integrates front-view cathode positioning verification using a Time-of-Flight sensor, front-view image acquisition using an industrial camera, real-time monitoring software, deep learning-based defect detection, PLC-based decision output, and robotic sorting. The inspection module combines YOLOv8-based segmentation-guided

preprocessing with a MobileNetV4 classification network enhanced by a Local–Global Attention mechanism. The segmentation stage isolates defect-relevant regions, including copper deposition areas, edge-strip zones, and structurally important blank regions, thereby reducing background interference and providing a cleaner input for classification. The proposed MobileNetV4–LGA classifier was designed to capture both localized defect patterns and broader structural irregularities. The CBAM-based local attention branch improves sensitivity to fine surface abnormalities such as copper bridging, edge-strip misalignment, and incomplete deposition, while the MHSA-based global attention branch captures long-range contextual relationships across the blank surface. To address the natural imbalance between Normal and Reject samples, the model was trained using an α -balanced focal loss. Experimental results on the independent test set demonstrated strong classification performance, achieving 98.5% accuracy, 97.6% precision, 96.2% recall, 96.9% F1-score, 0.992 AUC, and 0.982 AP. These results confirm that the proposed framework can reliably distinguish Normal and Reject copper cathode blanks under real industrial operating conditions while maintaining a low number of missed defects and false rejections. Overall, the proposed framework provides an accurate, deployment-oriented, and interactive solution for automated copper cathode blank inspection and robotic sorting. By combining hardware-based position verification, industrial image acquisition, segmentation-guided deep learning, real-time visualization, and PLC-based reject command generation, the system moves beyond offline image classification and supports practical manufacturing decision-making. Future work may focus on validating the framework across additional refinery sites, improving real-time inference and hardware performance analysis, integrating multi-camera or multi-modal sensing such as depth or thermal imaging, and exploring adaptive or unsupervised anomaly detection methods to improve robustness against unseen defect patterns.

Statements and Declarations

Ethics Statement: The author affirms that this research was carried out in compliance with recognized ethical standards and principles. All information sources have been appropriately cited and credited. Where required, ethical clearance was granted, and informed consent was obtained from all participants. No instances of plagiarism, data fabrication, or falsification were involved in this study.

Use of Artificial Intelligence: The author confirms that AI-based tools were employed exclusively for language polishing and grammatical correction. All AI-generated assistance was thoroughly examined, validated, and revised by the author, who bears full responsibility for the precision, authenticity, and integrity of the final manuscript. AI tools were not utilized to produce or alter research data or to substitute essential scholarly reasoning.

Conflicts of Interest: Any possible conflicts of interest, including financial or personal relationships that might affect the research, are disclosed. The author states that there are no competing interests.

Compliance with Regulations: The author verifies that this study adhered to all relevant institutional, national, and international rules, and that all required permissions for data usage and reproduction were secured.

References

- Aghelizadeh, P., Tang, Y., Sychterz, A., & Sneed, L. H. (2026). Review on vibration-based damage detection of vehicular bridges considering traffic effects. *Journal of Civil Structural Health Monitoring*, 16(5), 53. <https://doi.org/10.1007/s13349-026-01102-4>
- Ahmed, K. R. (2023). DSTEELNet: A real-time parallel dilated CNN with Atrous spatial pyramid pooling for detecting and classifying defects in surface steel strips. *Sensors*, 23(1), 544. <https://doi.org/10.3390/s23010544>

- Bagheri, M. (2025). *Advancing architecture through shape memory alloy actuators* (Doctoral dissertation, Virginia Tech). <https://vtechworks.lib.vt.edu/items/c13d5432-b85f-490e-b194-c2a4f6cf3ca8>
- Eskandari, M., Indukuri, M., Lukin, S. M., & Matuszek, C. (2025). LLM-supported safety annotation in high-risk environments. In *HRI 2025 Workshop VAM-HRI*. <https://openreview.net/forum?noteId=Ewg3WsMBRv>
- Eslami, N., Mohajjel, M., Rasouli Kenari, A., & Abbasi, S. (2025). Energy-Aware multi-exit vision transformer with early exits for efficient inference on energy-harvesting edge Devices. *Engineering Management and Soft Computing*. <https://doi.org/10.22091/jemsc.2025.13413.1289>
- Ghayoumi Zadeh, H., Fayazi, A., Rezaee, K., Aminaee, A., Halavati, H., Tahernejad, M., Memarzadeh, H., Masoumi, A., & Jafari, M. S. (2026). Automated Surface Defect Detection in Copper Blanks Using YOLOv8 Segmentation and EfficientNetV2-S Classification. *Journal of AI and Data Mining*, 14(2), 169-181. <https://doi.org/10.22044/jadm.2026.16853.2818>
- Goftari, E., London, J. K., Carias, J., Gittings, E., Fulford, L., Capps, S., Louie, B., Regalado, S., Ruiz, E., Torres, C., & Li, H. (2026). Building Capacity for Advanced Community Air Monitoring in California's Overburdened Communities. *Environmental Justice*, 0(0). <https://doi.org/10.1177/19394071261474817>
- Hosseini, M., & Shanechi, M. M. (2025). Dynamical modeling of behaviorally relevant spatiotemporal patterns in neural imaging data. *Proceedings of machine learning research*, 267, 23846.
- Hosseini, M., Khatounabadi, H., & Fakharzadeh, M. (2026). A Low-Power Wearable Respiratory Sensor for Non-Invasive Stress Monitoring. *arXiv preprint arXiv:2608.05697*. <https://doi.org/10.48550/arXiv.2608.05697>
- Hu, B., & Wang, J. (2020). Detection of PCB surface defects with improved faster-RCNN and feature pyramid network. *IEEE Access*, 8, 108335–108345. <https://doi.org/10.1109/ACCESS.2020.3001349>
- Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017, July). Densely connected convolutional networks. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 2261–2269). Honolulu. <https://doi.org/10.1109/CVPR.2017.243>
- Hütten, N., Alves Gomes, M., Hölken, F., Andricevic, K., Meyes, R., & Meisen, T. (2024). Deep learning for automated visual inspection in manufacturing and maintenance: A survey of open-access papers. *Applied System Innovation*, 7(1), 11. <https://doi.org/10.3390/asi7010011>
- Indukuri, M., Eskandari, M., Kollu, S. N., Lukin, S., & Matuszek, C. (2026). Hazard or Anomaly? Evaluating VLMs for Understanding Dangers and Discrepancies. *arXiv preprint arXiv:2607.18325*. <https://doi.org/10.48550/arXiv.2607.18325>
- Joukar, A., Karnik, S., Noori-Dokht, H., Younesi, S., Trippel, S. B., & Wagner, D. R. (2025). Mechanical Wear of Degraded Articular Cartilage: A. Joukar et al. *Annals of Biomedical Engineering*, 53(4), 956-965. <https://doi.org/10.1007/s10439-025-03680-8>
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84–90. <https://doi.org/10.1145/3065386>
- Li, L., Liu, Z., Zhao, H., Xue, L., & Wu, J. (2024). The bearing surface defect detection method combining magnetic particle testing and deep learning. *Applied Sciences*, 14(5), 1747. <https://doi.org/10.3390/app14051747>
- Lins, R. G., Santos, R. E. D., & Gaspar, R. (2023). Vision-based measurement for quality control inspection in the context of Industry 4.0: A comprehensive review and design challenges. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, 45(4). <https://doi.org/10.1007/s40430-023-04050-y>
- Liu, J., Meng, Y., Li, C., & Tie, J. (2024, September). Research on purity control on high-purity copper by electrorefining. In *Tenth International Conference on Mechanical Engineering, Materials, and Automation Technology (MMEAT 2024)* (Vol. 13261, pp. 627-631). SPIE. <https://doi.org/10.1117/12.3046539>
- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C., & Berg, A. C. (2016). SSD: Single shot MultiBox detector. In *Lecture Notes in Computer Science* (pp. 21–37). https://doi.org/10.1007/978-3-319-46448-0_2
- Lu, J., Dou, H., Zhou, Z., Li, H., Wang, Z., Jiang, M., Li, F., Gao, Y., Song, C., Fang, D., He, Y., & Li, Y. (2023). Effect of rapid hollow cathode plasma nitriding treatment on corrosion resistance and friction performance of AISI 304 stainless steel. *Materials*, 16(24), 7616. <https://doi.org/10.3390/ma16247616>

- Maydanchi, M., Ziaei, A., Basiri, M., Azad, A.N., Pouya, S., Ziaei, M., Haji, F. and Sargolzaei, S. (2023, April). Comparative Study of Decision Tree, AdaBoost, Random Forest, Naïve Bayes, KNN, and Perceptron for Heart Disease Prediction. In *SoutheastCon 2023* (pp. 204-208). IEEE. <https://doi.org/10.1109/SoutheastCon51012.2023.10115189>
- Monfaredi, P., Emami, S. M. M., & Moghadam, A. S. (2022). Seismic behavior of hollow-core infilled steel frames; an experimental and numerical study. *Journal of Constructional Steel Research*, 192, 107244.
- Pakzad, A., & Pakzad, A. (2026). Application of deep learning algorithms in clustering production units: Enhancing regulatory processes using soft computing approaches (Case Study: North Khorasan Province). *Engineering Management and Soft Computing*, 12(1), 82-101. <https://doi.org/10.22091/jemsc.2026.13325.1286>
- Qin, D., Leichner, C., Delakis, M., Fornoni, M., Luo, S., Yang, F., Wang, W., Banbury, C., Ye, C., Akin, B., Aggarwal, V., Zhu, T., Moro, D., & Howard, A. (2024). MobileNetV4: Universal Models for the Mobile Ecosystem. In *Lecture Notes in Computer Science* (pp. 78–96). https://doi.org/10.1007/978-3-031-73661-2_5
- Sahlman, M., Aromaa, J., & Lundström, M. (2021). Copper cathode contamination by nickel in copper electrorefining. *Metals*, 11(11), 1758. <https://doi.org/10.3390/met11111758>
- Simonyan, K., & Zisserman, A. (2015, April). Very deep convolutional networks for large-scale image recognition. In *3rd international conference on learning representations (ICLR 2015)*. Computational and Biological Learning Society. <https://doi.org/10.48550/arXiv.1409.1556>
- Unal, H. B., Zeynali, S., Anttila, E., Dharmakumar, R., Gross, D. C., & Sharif, B. (2026b). Enhanced T1 sensitivity and reduced artifacts in ferumoxytol-enhanced CMRI at 0.55 T compared to 3 T: a preclinical in vivo study. *Journal of Cardiovascular Magnetic Resonance*, 28. <https://doi.org/10.1016/j.jocmr.2025.102252>
- Unal, H.B., Zeynali, S., Ahmed, A., Anttila, E., Mastouri, R., Kreutz, R.P., Dharmakumar, R., Gross, D.C. and Sharif, B. (2026a). Ferumoxytol-Enhanced Myocardial T1 Tracking Using a Hybrid 2D/3D Steady-State MRI Sequence Captures Cyclic Intramyocardial Blood Volume Dynamics. *NMR in Biomedicine*, 39(6), e70308. <https://doi.org/10.1002/nbm.70308>
- Wei, R., & Bi, Y. (2019). Research on recognition technology of aluminum profile surface defects based on deep learning. *Materials*, 12(10), 1681. <https://doi.org/10.3390/ma12101681>
- Xian, Y., Liu, G., Fan, J., Yu, Y., & Wang, Z. (2021). YOT-Net: YOLOv3 combined triplet loss network for copper elbow surface defect detection. *Sensors*, 21(21), 7260. <https://doi.org/10.3390/s21217260>
- Xian, Y., Yu, Y., Lian, Y., Fan, J., & Wang, Z. (2023). An EA-based pruning on improved YOLOv3 for rapid copper elbow surface defect detection. *Engineering Applications of Artificial Intelligence*, 123, 106412. <https://doi.org/10.1016/j.engappai.2023.106412>
- Xie, Y., Hu, W., Xie, S., & He, L. (2023). Surface defect detection algorithm based on feature-enhanced YOLO. *Cognitive Computation*, 15(2), 565–579. <https://doi.org/10.1007/s12559-022-10061-z>
- Zahra, R., & Rezaeenour, J. (2025). Review community detection algorithms in multilayer networks; Traditional methods and deep learning. *Engineering Management and Soft Computing*, 10(2), 1-25. <https://doi.org/10.22091/jemsc.2025.11154.1196>
- Zhang, G., Chen, T., & Wang, J. (2025). CSC-YOLO: An image recognition model for surface defect detection of copper strip and plates. *Journal of Shanghai Jiaotong University (Science)*, 30(5), 1037–1049. <https://doi.org/10.1007/s12204-024-2723-2>
- Zhang, Z., Huang, X., Wei, D., Chang, Q., Liu, J., & Jing, Q. (2024). Copper nodule defect detection in industrial processes using deep learning. *Information*, 15(12), 802. <https://doi.org/10.3390/info15120802>
- Zhao, C., Shu, X., Yan, X., Zuo, X., & Zhu, F. (2023). RDD-YOLO: A modified YOLO for detection of steel surface defects. *Measurement*, 214, 112776. <https://doi.org/10.1016/j.measurement.2023.112776>
- Zhao, W., Chen, F., Huang, H., Li, D., & Cheng, W. (2021). A new steel defect detection algorithm based on deep learning. *Computational Intelligence and Neuroscience*, 2021(1). <https://doi.org/10.1155/2021/5592878>