Sensors and Materials, Vol. 38, No. 6 (2026) 3543–3562
MYU Tokyo

S & M 4519

AI-assisted Vision-based Sensing System for Printed Circuit Board Defect Detection Using You Only Look Once

Yi-Heng Wu,[1] Chung-Wen Hung,[2] Yi-An Chen,[2] and Lung-Fa Wu[3*]

[1]Department of Computer Information and Network Engineering, Lunghwa University of Science and Technology,
No. 300, Sec. 1, Wanshou Rd., Guishan District, Taoyuan City 333326, Taiwan
[2]Department of Electrical Engineering, National Yunlin University of Science and Technology,
No. 123, Section 3, University Road, Douliu City, Yunlin County 640301, Taiwan
[3]Department of Electrical Engineering, National Ilan University
No. 1, Sec. 1, Shenlong Rd., Yilan City 26047, Taiwan

(Received August 21, 2025; accepted June 5, 2026)

Keywords: YOLO, vision-based sensing, automated optical inspection, machine vision, printed circuit board inspection, defect detection

The integration of AI into automated optical inspection enables the enhancement of printed circuit board manufacturing efficiency, accuracy, and cost-effectiveness. We developed an AI-assisted, vision-based defect sensing system that conducts multiclass electronic component recognition and surface defect detection by coupling an industrial optical sensing module with an enhanced You Only Look Once version 9-e (YOLOv9-e) deep learning architecture. The developed system enables the integration of controlled ring lighting and a high-resolution CMOS visual sensor to overcome image degradation in raw sensory signals, establishing a highly accurate, noncontact optical inspection concept. Printed circuit board samples containing diverse soldering defects from a national technical examination framework were utilized to compile a dataset of 220 images. Comparative YOLOv9-e outperformed YOLOv7, achieving a component recognition mean average precision at the intersection over union threshold of 0.50 (mAP@0.5) of 93.6%, an F1-score of 90.0%, and a defect detection accuracy of 89.8%. Although the noncontact sensing configuration of the developed system provides robust, real-time diagnostic capabilities, limitations exist, including dataset diversity and susceptibility to ambient illumination variations during sub-millimeter solder wetting inspection. To address the limitations, computational structural re-parameterization in the network layers is required to preserve critical geometric reflections.

1. Introduction

Various industrial sensing methodologies have been developed for accurate defect detection. However, translating these methodologies into the optical sensing domain presents various challenges. To address these challenges, advanced signal processing and deep architectures have

---

[*] Corresponding author: e-mail: lfwu@niu.edu.tw

https://doi.org/10.18494/SAM5912

been introduced across diverse engineering sectors. For instance, a variational autoencoder integrated with a deep neural network classifier has been utilized to convert complex sensory data into Gaussian-distributed latent variables for precise defect detection and classification.(1) Similarly, probabilistic confidence convolutional neural networks (CNNs) have been employed to enhance adaptability when isolating known and unknown defect categories from dynamic streams.(2) To relate structural signal analysis and visual spatial awareness, multichannel image conversion methods, such as transforming time- and frequency-domain signals into spatial matrices for CNN-based diagnosis, have been widely used owing to their significant performance enhancements in automated defect isolation.(3)

On the basis of such cross-domain image conversion techniques, deep learning architectures such as CNNs and You Only Look Once (YOLO) models have become essential components in industrial visual inspection owing to their capacity for real-time object detection and feature classification.(4) In high-precision surface quality monitoring, a system combining the YOLOv4 architecture with CCD sensors presents highly precise, real-time defect extraction on thin-film transistor liquid crystal display glass substrates.(5) Enhanced YOLOv9 models integrated with the ByteTrack algorithm overcome multiscale tracking challenges and dense occlusions in heavy industrial transport scenarios,(6,7) whereas a compressed YOLOv7-Tiny network paired with U-Net (a specialized CNN architecture used for semantic image segmentation) ensures superior localization accuracy for complex topological patterns in agricultural disease sensing.(8) In smart manufacturing, these noncontact machine vision architectures are integrated with automated optical inspection (AOI) hardware for printed circuit board (PCB) defect detection, offering high-speed, nondestructive, and exceptionally accurate surface verification.(9–11) AOI technology operates as a noncontact, optomechanical inspection framework that replaces manual visual scrutiny with synchronized data acquisition and digital signal processing. The technology relies on a three-tier architecture: an engineered illumination array (typically multi-angle LED rings or coaxial lights) to highlight surface contours, a high-resolution optical sensor configuration, such as CCD or CMOS cameras, to capture the reflected light fields, and a computational processing unit. By orchestrating these components, AOI systems rapidly evaluate surface characteristics, such as checking component alignment, verifying part presence, and evaluating the wetting boundaries of solder joints, offering high-speed, nondestructive, and exceptionally accurate surface verification.

Despite the rapid speed and noncontact operational advantages of integrating deep architectures with AOI technologies, high-accuracy defect detection and classification largely depend on the signal quality of the raw optical inputs captured by visual sensors. Microsoldering surfaces often exhibit high specular reflection and variable gloss profiles, making traditional machine vision methods prone to false positives triggered by even slight fluctuations in the ambient illumination angle.(12) A primary problem in vision-based quality sensing is the accurate modeling of multiscale, sub-millimeter geometric surface structural features under variable production light fields. Conventional methods are used to resolve light field variations using polarized illumination, dedicated sensor arrays, or multisensor data fusion networks. However, such physical interventions considerably increase the hardware footprint, procurement cost, and systemic complexity of the optical apparatus.(13) Therefore, there is a demand to develop alternative computational methods that utilize programmable gradient information with high-speed CMOS sensors. By optimizing the gradient flow, this method effectively resolves the information loss common in traditional deep convolutional layers, ensuring the strict preservation of fine-grained geometric reflections down to the final output layer.

In this study, we developed an AI-assisted vision-based sensing system for PCB defect detection, which balances real-time edge processing constraints with high detection accuracy for surface microdefects. By implementing a localized spatial patching strategy to maximize target resolution feasibility, the developed system successfully mitigates sensor performance deterioration caused by ambient illumination variations and overcomes the inherent difficulty of resolving sub-millimeter-sized component profiles under standard industrial pixel densities. The developed system enhances dataset representation across complex industrial layouts, establishing a robust computational foundation for the deployment of real-time inline AOI hardware setups in modern manufacturing lines.

2. System Architecture

On the basis of industrial benchmarks of the Electric Power Electronics Category B Technician Skill Certification Practical Examination guidelines,(14) an automated defect detection system that interfaces vision-based sensor data with a deep learning framework was developed. The developed system's workflow begins with the optimization of an automated imaging system, which functions as an integrated optomechanical staging rig to systematically capture and establish a high-resolution surface defect dataset. In this phase, multilayer PCBs are the target objects, with structural defects introduced to capture diverse surface anomalies. Following raw data acquisition, the captured images are annotated, and targeted data augmentation is applied to low-exposure or under-represented defect classes to ensure data balance across the computational network. YOLOv9 and YOLOv7 architectures are deployed for model optimization and comparison. In the collaborative sensing environment, the YOLOv9-e model and the YOLOv7-d6 model are trained on identical dataset divisions to evaluate sensor-to-model compatibility, spatial feature extraction accuracy, and practical inference throughput during real-time multiclass defect detection tasks.

2.1 Architecture design

The developed system integrates a high-resolution industrial CMOS sensor with adjustable multi-angle LED ring illumination to stabilize light field conditions. The architecture consists of four modules (Fig. 1).

(1) Image acquisition module

A CMOS camera with regulated LED ring illumination serves as the optical sensor for automated PCB imaging. The camera employs a 0.5-inch progressive scan CMOS sensor with 3072 × 2048 resolution (6.3 megapixels), 2.4 µm pixel pitch, and a global electronic shutter to eliminate motion blur. A 16 mm low-distortion industrial lens with a C-mount ensures precise capture. The sensing principle relies on directional reflection. Appropriate

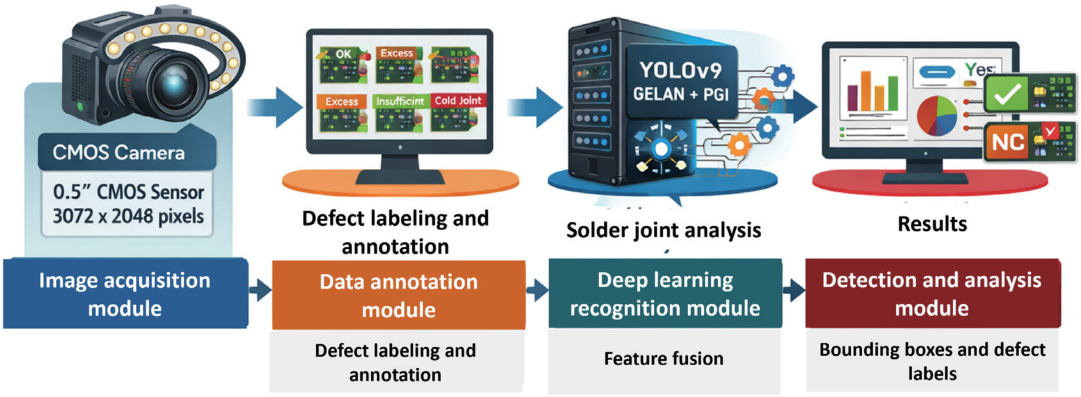


Fig. 1. (Color online) PCB defect detection system architecture (created in this study using AI).

solder volumes form concave meniscus profiles reflecting light uniformly, whereas deviations distort scattering patterns. This process translates three-dimensional surface topology into two-dimensional spatial feature maps.

(2) Data annotation module

Captured features are mapped to defect categories such as proper wetting, excessive volume, insufficient volume, and cold solder joints. Raw pixel coordinates are converted into standardized annotated matrices, forming training labels for deep learning optimization.

(3) Deep learning recognition module

Digitized feature maps are processed using the YOLOv9 architecture for real-time solder joint anomaly classification. The network combines a generalized efficient layer aggregation network (GELAN) with programmable gradient information (PGI) to mitigate information loss across convolutional layers. GELAN with a PGI framework is used to maximize inference speed on edge devices and incorporate lightweight, highly efficient convolutional modules without sacrificing the depth needed to isolate delicate surface features while maintaining sub-millimeter localization accuracy. This ensures that fine-grained geometric reflections from sub-millimeter solder joints are preserved through to the final prediction layer.(6,15)

(4) Detection and analysis module

Model validation is performed using YOLOv9 and YOLOv7 outputs, evaluated with precision, recall, and F1-score metrics. A feedback loop logs defect distributions and issues real-time control flags to inline inspection alerts, enabling rapid monitoring and automated quality control.(16)

2.2 Image processing

Input data is verified to ensure proper image format. Gaussian-distributed random noise of identical dimensions is added pixel by pixel to the original image to improve robustness against illumination variation. To prevent distortion in brightness, pixel values are clipped within the valid range [0, 255]. This preprocessing step enhances generalization by simulating real-world noise conditions encountered during automated optical inspection.

2.2.1 Model selection

To select an appropriate architecture, we compared diverse YOLO models (Table 1). YOLO is a family of computer vision models designed to detect and classify objects in images with high speed and accuracy. Each version introduces architectural modifications that affect performance. YOLOv4 employs a backbone known as cross-stage partial darknet with 53 layers (CSPDarknet53) combined with a path aggregation network (PANet), achieving high accuracy but requiring significant computational resources. YOLOv7-Tiny is a lightweight variant based on the extended efficient layer aggregation network (E-ELAN), optimized for devices with limited computing power. It is fast and memory-efficient, but its shallow depth reduces accuracy, limiting its ability to detect subtle solder joint defects. YOLOv7-d6 is a variant of YOLOv7 that enhances gradient flow, improving accuracy, but its extended architecture increases inference time and memory usage, which constrains its suitability for real-time industrial inspection. YOLOv9-e introduces GELAN in combination with the re-parameterized cross-stage partial network (RepCSPNet) and ELAN. It also integrates PGI, a mechanism that ensures fine details such as subtle optical reflections from solder joints are preserved as data propagates through deep layers.

The YOLOv4 model provides a strong baseline accuracy but lacks efficiency for real-time PCB inspection, whereas the YOLOv7-Tiny model is efficient but sacrifices precision, making it unsuitable for fine solder joint defect detection. The YOLOv7-d6 model improves accuracy but increases inference time and memory usage. In contrast, the YOLOv9-e model, which incorporates GELAN and PGI, maintains the optimal balance by preserving fine-grained reflections, enabling multiscale fusion, and sustaining computational efficiency.

YOLOv9-e shows the highest accuracy and efficiency. By preserving fine-grained optical information while maintaining computational speed, YOLOv9-e provides a robust solution for real-time PCB defect detection. This balance of precision and efficiency underscores its

Table 1
Comparative analysis results of YOLO models.

Feature	YOLOv4(17)	YOLOv7-Tiny(18)	YOLOv7-d6(19)	YOLOv9-e(20)
Core architecture	CSPDarknet53 backbone + PANet	Lightweight E-ELAN	Deep E-ELAN with extended layers	GELAN with RepCSPNet + ELAN
Information loss control	Standard gradient flow	Reduced due to shallow depth	Improved gradient flow with deeper stacking	PGI auxiliary branch preserves fine-grained reflections
Feature extraction module	CSP + PANet	RepConv layers (lightweight)	RepConv + extended ELAN	RepCSPNet + hierarchical ELAN
Computational efficiency	High accuracy but slower inference	Very fast, low memory footprint	Higher accuracy, slower inference	Balanced speed and accuracy via re-parameterization
Multiscale fusion precision	Standard FPN + PANet	Limited due to shallow depth	Enhanced multiscale fusion	Superior multiscale fusion with PGI
Application suitability	General object detection	Edge devices, low power	High-accuracy industrial inspection	Real-time PCB defect detection

suitability for intelligent manufacturing applications, where reliable inspection under variable optical conditions is essential. Therefore, YOLOv9 is well suited for PCB defect detection under variable optical conditions.

The model comprises an input module, backbone, neck, and output unit (Fig. 2). GELAN integrates RepCSPNet and ELAN modules to achieve high-performance feature extraction. RepCSPNet was adopted to employ structural re-parameterization, converting multibranch training structures into streamlined single-path inference structures. This reduces memory consumption and accelerates computation. ELAN stacks convolutional groups hierarchically, preserving gradient flow and enabling deeper feature processing. These modules enhance multiscale feature fusion and detection precision while maintaining computational efficiency, making YOLOv9 well suited for real-time PCB defect detection where accuracy and speed are critical.(6,7,12)

2.2.2 Defect detection

The criteria for defects are defined for two interrelated constraints: component pin geometry and solder material volume boundaries. For pin treatment, either the solder first, then trim or trim first, then solder sequence is acceptable. However, the post-trimming lead extension must not exceed 0.5 mm to prevent capacitive coupling or short-circuiting. For specific component groups, including IC sockets, small variable resistors, transformers, and high-current terminal blocks, lead trimming is omitted to preserve mechanical anchoring integrity.

In terms of solder volume control and joint morphology, the cooled metallic material must exhibit a smooth, specularly reflective surface. From a materials sensing perspective, phase-change defects such as cold solder joints, localized bubbling, burnt flux residue, and dull, micrograngular surface textures must be avoided. The volume of the multi-alloy solder must be carefully regulated; both excessive accumulation and insufficient wetting introduce boundary conditions that degrade electrical conductivity and reduce long-term mechanical shear strength.

The qualitative and quantitative assessment of solder integrity depends on characterizing the volumetric distribution and light-scattering properties of the joint meniscus. As shown in Fig. 3(a), excessive solder produces a thick, convex profile that overflows the copper pad mask, obscuring component leads and indicating an unacceptable condition. In Fig. 3(b), an optimal solder volume forms a bright, uniform concave fillet, yielding a structurally sound electrical connection classified as acceptable. Conversely, Fig. 3(c) shows insufficient solder volume,

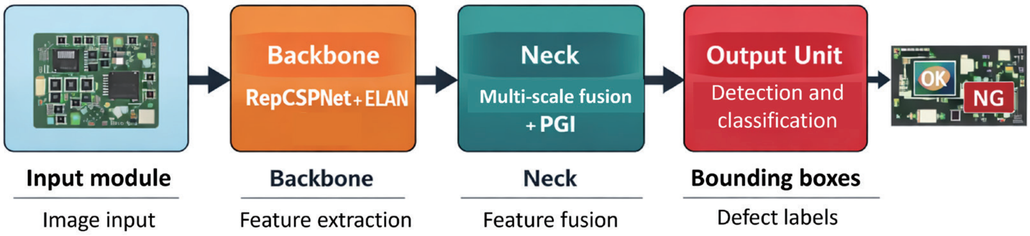
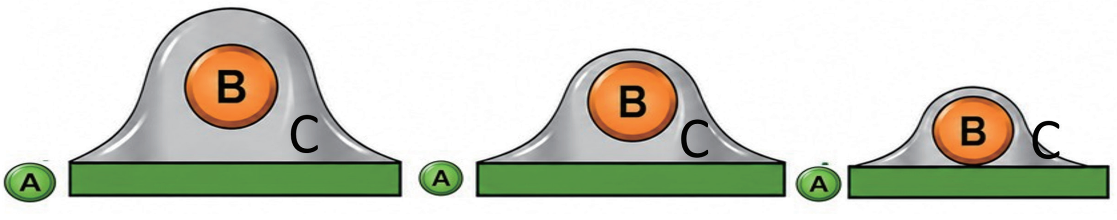


Fig. 2. (Color online) Simplified structure of YOLOv9-e (NG: no good, modified using AI).(20)



(a) (b) (c)

Fig. 3. (Color online) Pin treatment and solder joint appearance requirements: (a) excessive solder volume (not acceptable), (b) appropriate solder volume (acceptable), and (c) insufficient solder volume (not acceptable) (A: PCB substrate, B: copper wire, and C: protective solder mask layer).

where the alloy fails to encapsulate the lead or climb the component terminal, resulting in an incomplete metallurgical bond.

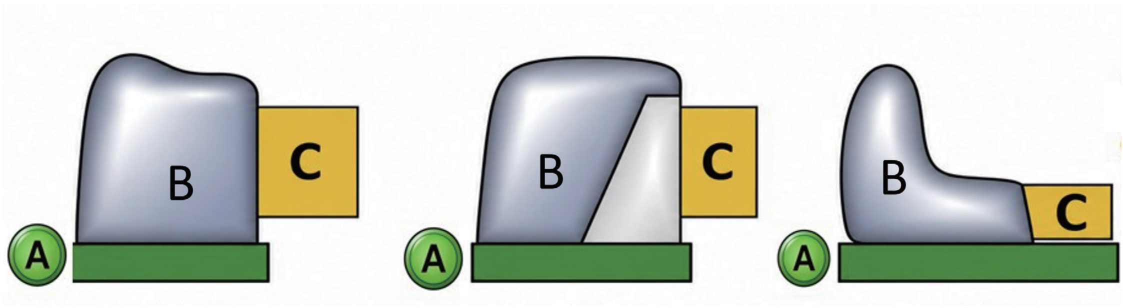
For surface-mount devices, the solder joint morphology is evaluated by analyzing boundary contours and optical reflection patterns. As shown in Fig. 4(a), a joint with good tin content exhibits complete wetting, forming a smooth concave boundary layer across the contact pads and ensuring a reliable electrical connection. In Fig. 4(b), insufficient tin content results in incomplete wetting, characterized by steep wetting angles and partial boundary contact, which fails to establish a stable electrical pathway. In contrast, Fig. 4(c) illustrates excessive tin content, where the solder accumulates into a large, uneven mound or droplet. This condition increases the risk of bridging adjacent traces and is classified as structurally defective.

3. Model Training and Performance Evaluation

3.1 Model training

After completing image preprocessing, accurate annotation is essential for the effective training of deep learning models. In particular, bounding box coordinates for each target object and their corresponding class labels provide the model with spatial distribution and visual characteristics, thereby enhancing the accuracy of both detection and classification. To establish a standardized training dataset, we used LabelImg, an open-source image annotation tool widely used in object detection research. LabelImg is a lightweight graphical interface developed in Python that offers a user-friendly environment for labeling tasks. It supports common image formats, such as portable network graphics, joint photographic experts, and scalable vector graphics, and accommodates multiple annotation standards, including Pascal visual object classes, YOLOv9, and common objects in context. These features are used for the seamless integration of images into deep learning frameworks and image processing pipelines, whereas the intuitive design improves the consistency and efficiency of manual labeling.

By using LabelImg, each solder joint image was manually annotated by specifying its category (good solder joint, excessive solder, insufficient solder, and poor wetting) and drawing the corresponding bounding boxes. Then, a standard YOLO-format .txt file was created, including class identification, bounding box center coordinates, width, and height. An example



(a) (b) (c)

Fig. 4. (Color online) Surface mount device (SMD) solder joints in wetting condition: (a) good tin content, (b) insufficient tin content, and (c) too much tin content (A: PCB substrate, B: protective solder mask layer, and C: SMD material).

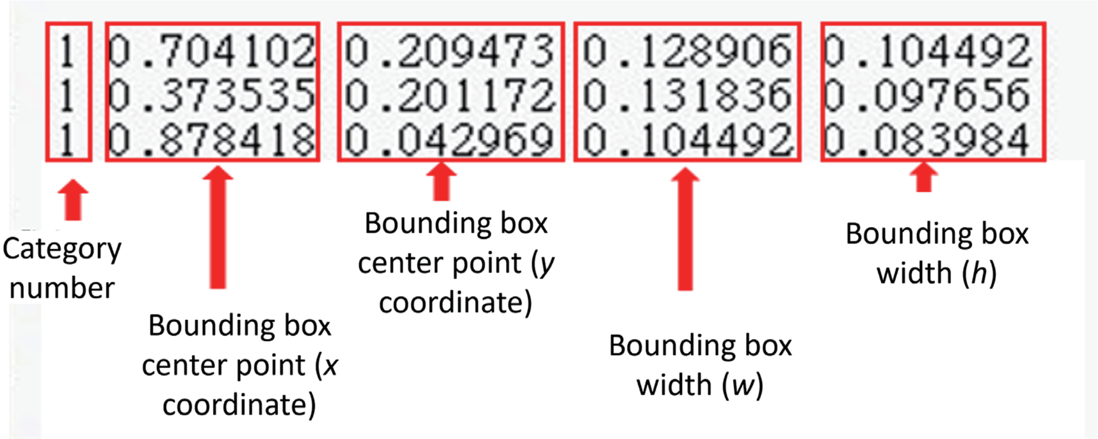
of this annotation file format is shown in Fig. 5. These annotation files served as the input data for subsequent model training, ensuring consistency and reproducibility throughout the learning process.

In this study, images were acquired through an automated imaging system and manually annotated for training a YOLOv9-based deep learning model. The dataset was split into training, validation, and test sets. The training set was used for model learning, while the validation set supported parameter tuning and helped prevent overfitting. The test set, composed of independent images, was reserved for final performance evaluation to assess generalization capability. Training was conducted in two stages: component recognition and defect detection. In the defect detection stage, each component type was further labeled as either normal or defective. Multiple YOLOv7 model variants with different depths were also trained and compared to evaluate detection accuracy and computational efficiency. The objective was to achieve an optimal balance between accuracy and speed, thereby ensuring the model's reliability and applicability in real-world inspection scenarios. Through systematic comparison and parameter adjustment, the most suitable model configuration for the research objectives was identified.

3.2 Evaluation metrics

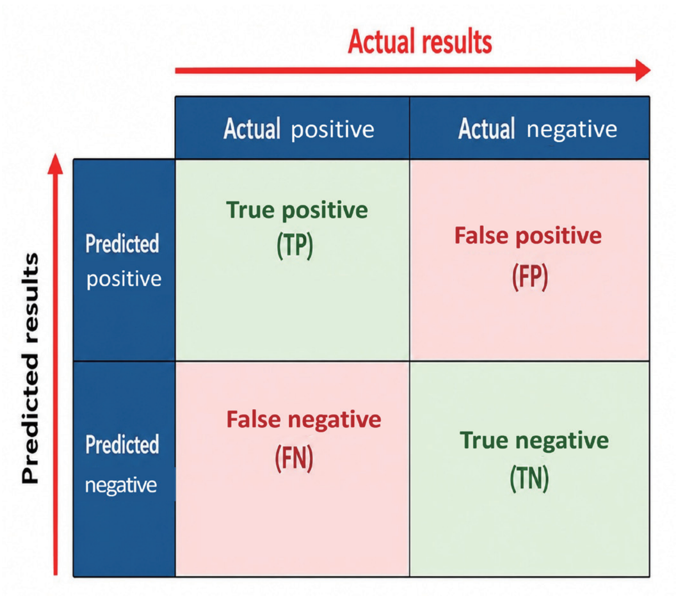
The performance of the YOLO model is evaluated using standardized metrics, including the confusion matrix, accuracy, precision, recall, F1-score, average precision, mean average precision (mAP), and intersection over union (IoU). A confusion matrix is constructed for an intuitive representation of prediction outcomes by distinguishing true positive (TP), false positive (FP), true negative (TN), and false negative (FN), thereby offering valuable insight into both correctness and error distribution (Fig. 6).(5,21)

The following performance metrics are calculated to evaluate the performance of the model. *Precision* is calculated to quantify the proportion of correctly predicted positive cases relative to all predicted positives.



```
1 0.704102 0.209473 0.128906 0.104492
1 0.373535 0.201172 0.131836 0.097656
1 0.878418 0.042969 0.104492 0.083984
```

Fig. 5. (Color online) Standardized image naming rules for PCB samples.



	Actual positive	Actual negative
Predicted positive	True positive (TP)	False positive (FP)
Predicted negative	False negative (FN)	True negative (TN)

Fig. 6. (Color online) Confusion matrix for classification accuracy evaluation.

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

Recall is used to measure the model's ability to correctly identify relevant targets.

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

Mean average precision (mAP) is calculated as an overall measure of detection accuracy by averaging the precision values across all object classes.

$$mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i \tag{3}$$

Here, AP_i is the average precision for class i, and N is the total number of classes.

F1-score is calculated by computing their harmonic mean, providing a single measure of performance that accounts for correctness and completeness.

$$F1-score = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \tag{4}$$

These metrics, combined with energy efficiency indicators such as inference time, memory usage, storage requirements, and peak power consumption, are used for the evaluation of both detection accuracy and computational efficiency.(22–24)

4. Results and Discussion

In this study, PCB samples produced by students during the Class B technical certification examination (a standardized vocational testing system managed by the Ministry of Labor in Taiwan) were used, among which ten defective samples were identified. For subsequent image analysis and processing, the captured PCB images were split into 16 or 25 equal parts, depending on their size, to obtain more detailed image samples. Two hundred twenty images were compiled for training and testing in this study. When using the automated imaging system, each image was coded in accordance with a standardized format comprising the sample number, area number, and shooting number (Table 2, Fig. 7).

To enhance the clarity of blurred regions, the original image was first converted to grayscale, followed by Gaussian blur processing. A weighted sharpening technique was then applied, and the sharpened pixel values were substituted for the corresponding values in the blurred regions of the original image. This procedure improved the overall visual clarity and ensured that fine structural details were preserved for subsequent analysis.

Table 2
PCB image coding.

Code	Description	Classification / Example
Sample number	Identification of PCB samples from which image was taken	20240531_205354
A	Segmented area of PCB image	Area designation
t	Tile index within segmented PCB image	Tile number (0, 1, 2…)
L/R	Orientation of image capture	L = normal (left/standard orientation) R = rotated (right/rotated orientation)
Normal/rotated	Shooting orientation.	Normal view (L) or rotated view (R)

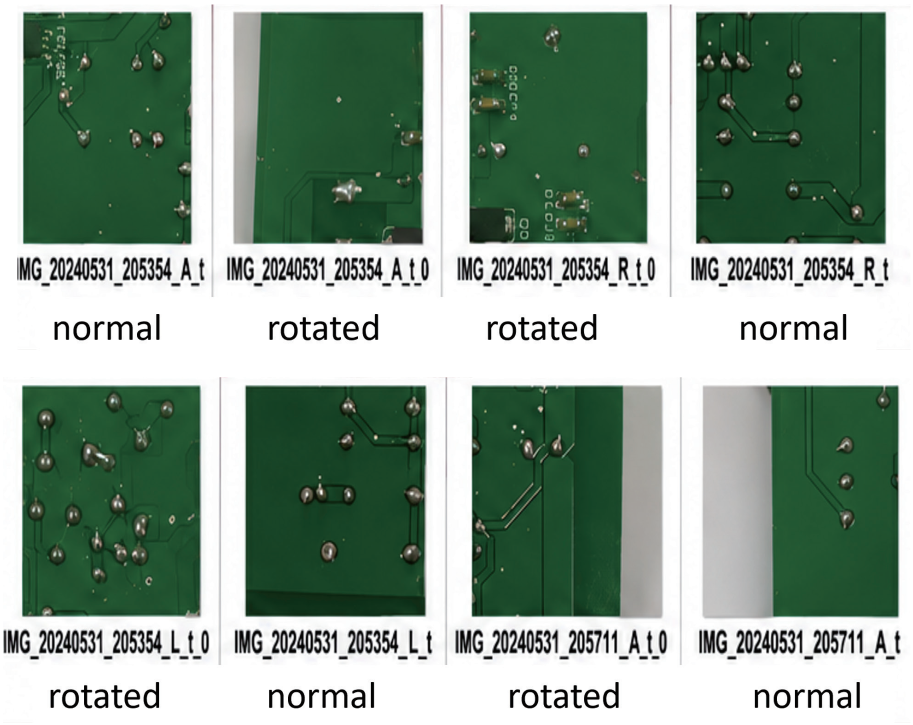


Fig. 7. (Color online) Examples of images taken in this study.

Figure 8 shows a collection of defect sample images, with red bounding boxes indicating the positions of defective components. Annotation was performed in the YOLOv9 format, with class ID 0 representing the defect category. Each bounding box specifies the spatial location and extent of a solder defect on the PCB. The dataset encompasses a variety of defect types, including excessive solder, insufficient solder, foreign object attachment, and solder bridging, under diverse visual and background conditions. This variability enhances the model's generalization capability and improves detection accuracy for real-world solder joint defects.

The confusion matrix in Fig. 9 was constructed to evaluate the classification performance of the model across three representative categories: defects, nondefective solder joints, and nonrelevant background. The diagonal elements represent correctly classified images, whereas the off-diagonal cells show misclassifications. The model showed strong recognition of defect regions and nondefective solder joints, achieving accuracies of 0.90 and 0.89, respectively. These values indicate that the model can reliably distinguish between defective and properly soldered components, which is consistent with findings in other defect detection studies where clear structural and morphological differences aided classification. However, the model presents a difficulty in correctly identifying background areas, exhibiting a classification accuracy of 0.00. Instead, background samples are frequently misclassified as either defect regions or nondefective solder joints. For example, 10% of background samples were labeled as defects, whereas 2% were labeled as nondefective joints. Similarly, nearly half of the defect samples (45%) and more than half of the nondefective samples (56%) were misclassified as background. This indicates

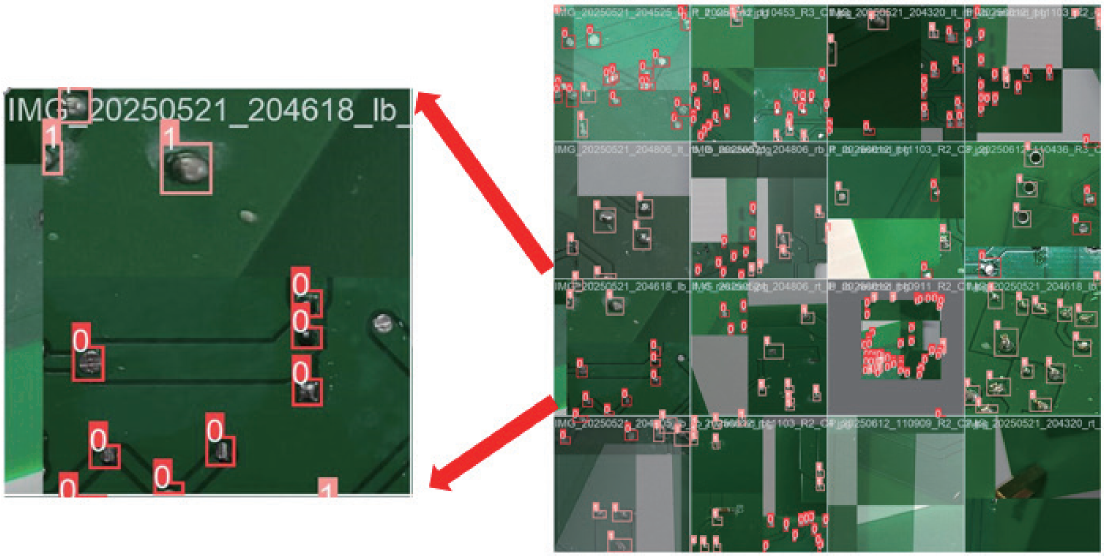


Fig. 8. (Color online) Annotated defect samples in YOLOv9 format.

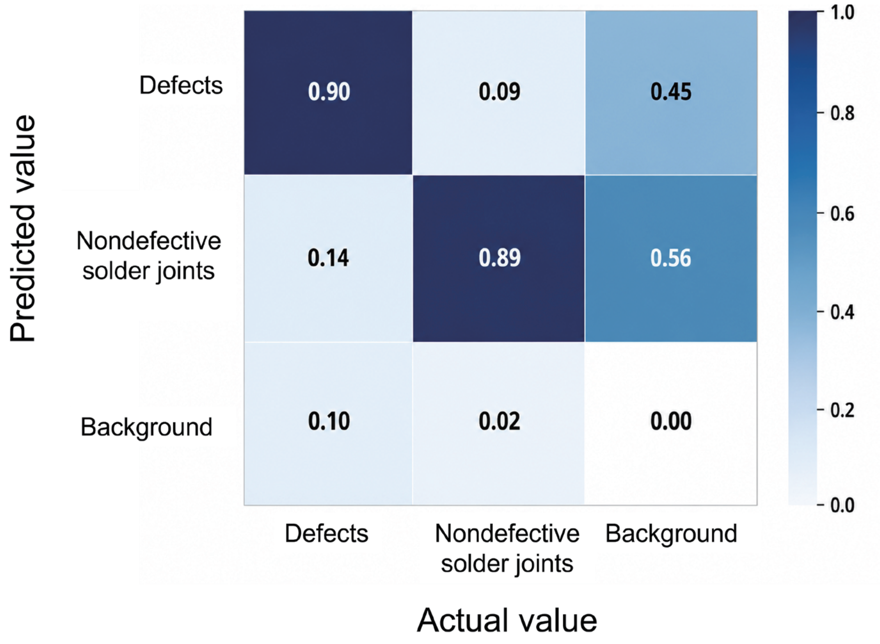


Fig. 9. (Color online) Confusion matrix for PCB defect detection using YOLOv9-e.

that the model struggles to separate background noise from meaningful solder joint features, a challenge commonly reported in visual inspection tasks where background textures overlap with low-contrast or shadowed regions.

Overall, the results show that although the model performs well in distinguishing defective and nondefective solder joints, it struggles to separate background from foreground categories.

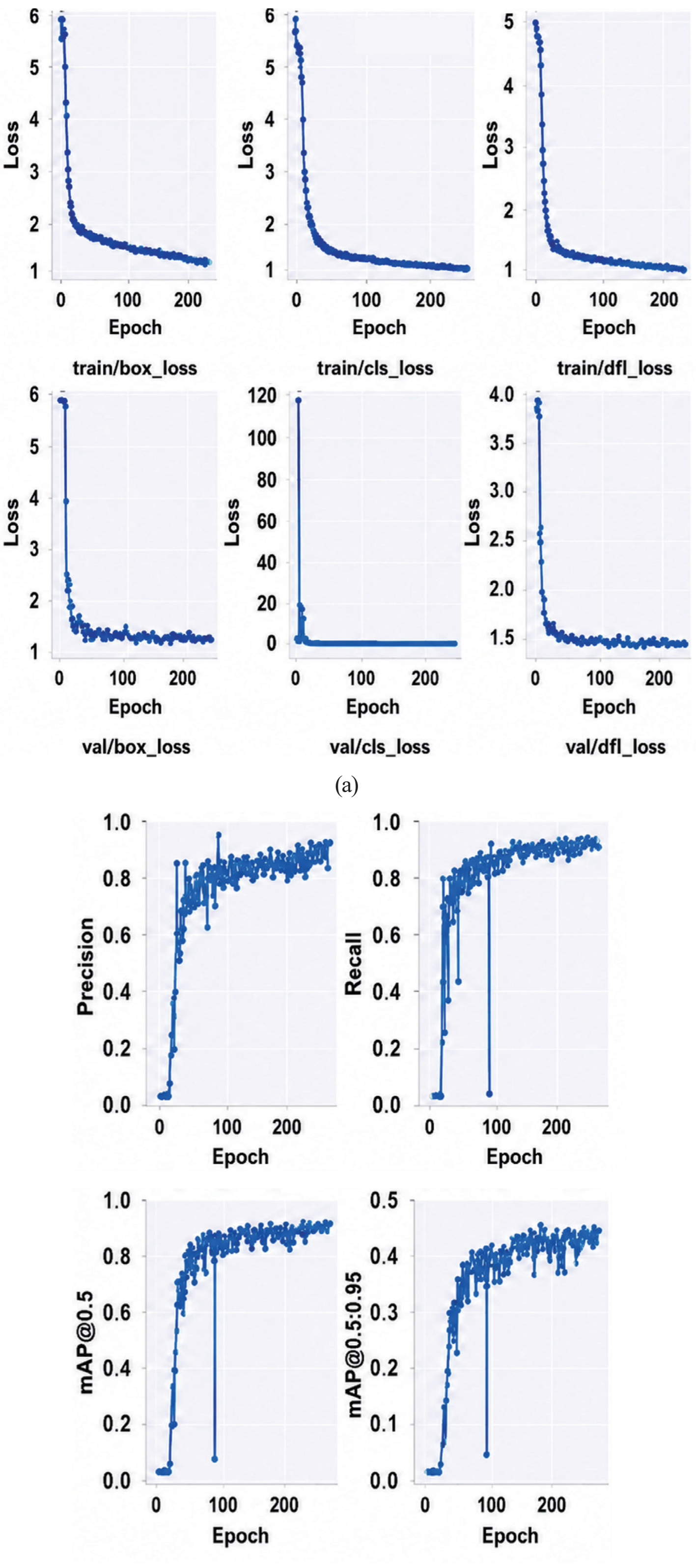
This imbalance suggests a bias toward detecting solder-related features, which reduces sensitivity to background regions. To address this limitation, more data with diverse backgrounds must be integrated, and illumination normalization techniques or attention-based deep learning mechanisms are applied to improve discrimination between subtle background noise and true solder joint features.

The training and validation metrics of the YOLOv9-e model are shown in Fig. 10. The top row shows training losses, including box loss (train/box_loss), classification loss (train/cls_loss), and distribution focal loss (train/dfl_loss), along with precision and recall metrics. The bottom row displays validation losses (val/box_loss, val/cls_loss, val/dfl_loss) and mAP at the IoU thresholds of 0.50 (mAP@50) of 0.5 and 0.95. All loss values steadily decreased with increasing epochs, indicating successful convergence. Meanwhile, precision, recall, and mAP consistently improved and stabilized, demonstrating that the model effectively learned spatial and categorical patterns with minimal signs of overfitting.

Figure 11 shows the annotated data distribution and characteristics in the training set. The top-left bar chart shows that the sample quantities of defect regions and nondefective solder joints are nearly equal, indicating a class balance. The top-right heatmap shows that most annotated objects are located near the image center, reflecting consistency in capture angles and framing. The bottom-left scatter plot indicates that object centers are broadly distributed across the image, while the bottom-right scatter plot demonstrates that most bounding boxes fall within a width range of 0.05–0.25 and a height range of 0.02–0.18. This near-normal distribution provides stable scale references for model training.

The inter-relationships among annotation attributes in the YOLOv9-e format, including center coordinates (x,y), width, and height, are shown in Fig. 12. The diagonal plots represent the univariate histograms of each attribute, whereas the lower triangular plots display 2D heatmaps capturing the pairwise distribution patterns. The x and y coordinates are nearly uniformly distributed, indicating that the labeled objects are spatially well-dispersed across the image area. A strong positive correlation is observed between width and height, with most bounding boxes clustered around a width of approximately 0.10 and a height of approximately 0.08, suggesting consistent object sizes that are conducive to stable model convergence. Additionally, the weak correlation between spatial position (x/y) and bounding box size implies that object location and scale are largely independent, an attribute that enhances the model's generalization capability.

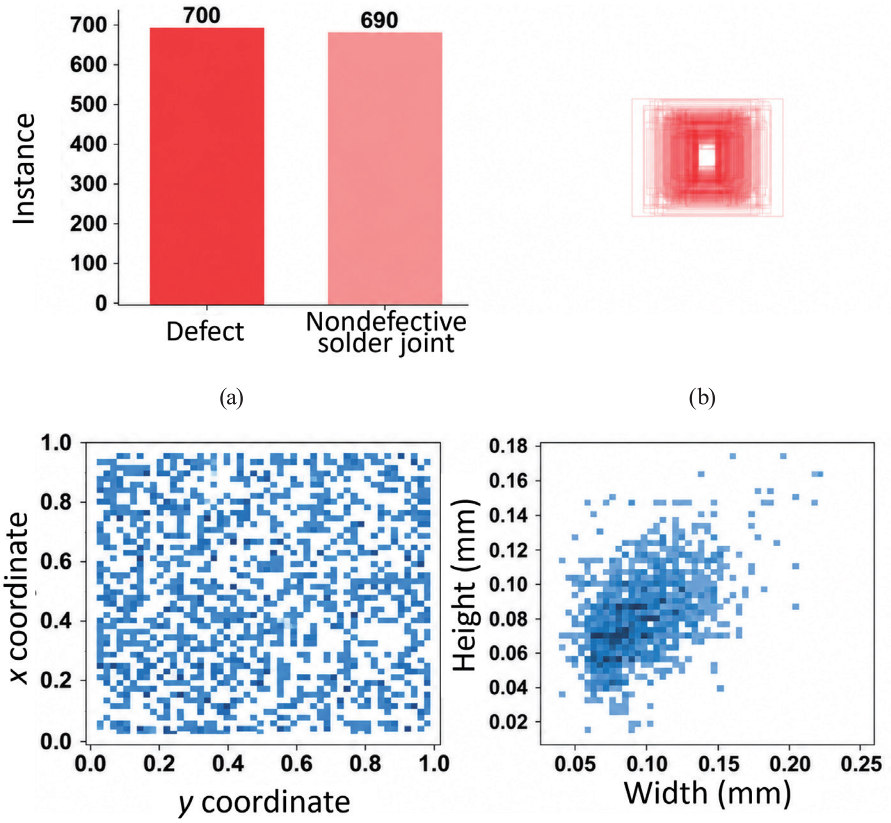
To evaluate the performance of the YOLO model trained in this study for defect recognition tasks, four evaluation curves are presented: the precision–confidence curve, recall–confidence curve, precision–recall curve, and F1–confidence curve (Fig. 13). Figure 13(a) shows that the model achieves a maximum precision of 1.00 across all classes at a confidence threshold of 0.639, which indicates extremely low false positive rates under high-confidence conditions. However, such high thresholds may reduce recall, requiring trade-off adjustments. The recall–confidence curve in Fig. 13(b) shows that recall reaches a maximum of 1.00 when confidence is near zero, meaning nearly all defect regions are detected. However, recall decreases sharply when the threshold exceeds 0.6, revealing a tendency for the model to miss defects under stricter confidence criteria. In the precision–recall curve shown in Fig. 13(c), the model achieves the mAP@0.5 of 0.917 for defect regions and 0.955 for nondefective solder joints, with an overall



(a)

(b)

Fig. 10. (Color online) Training and validation metrics of YOLOv9-e model: (a) training and validation losses and (b) precision, recall, and mAP.



(a) (b)

(c) (d)

Fig. 11. (Color online) Annotation data distribution and statistics: (a) instances per class, (b) bounding box heatmap, (c) object center distribution, and (d) width and height distribution.

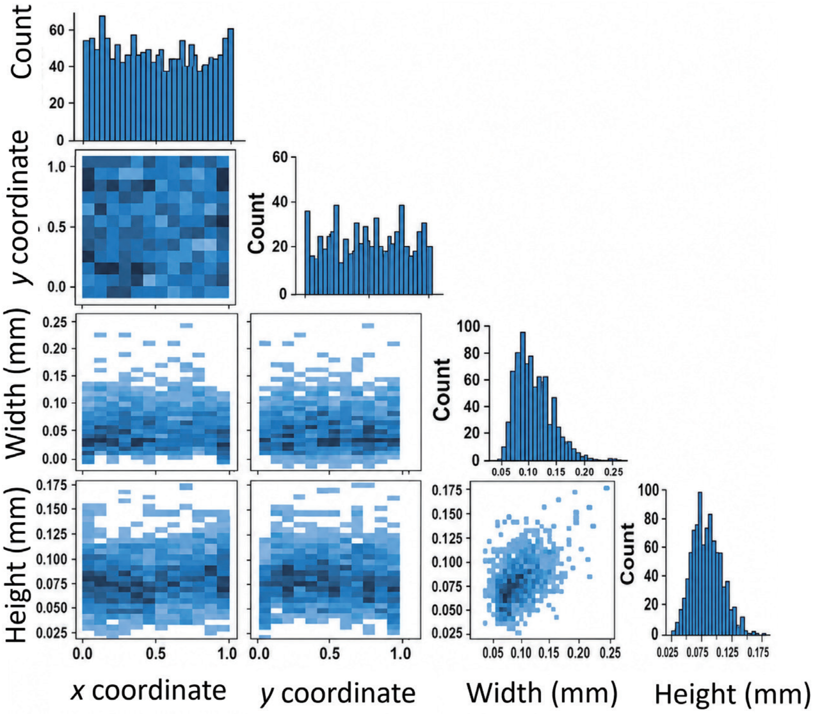
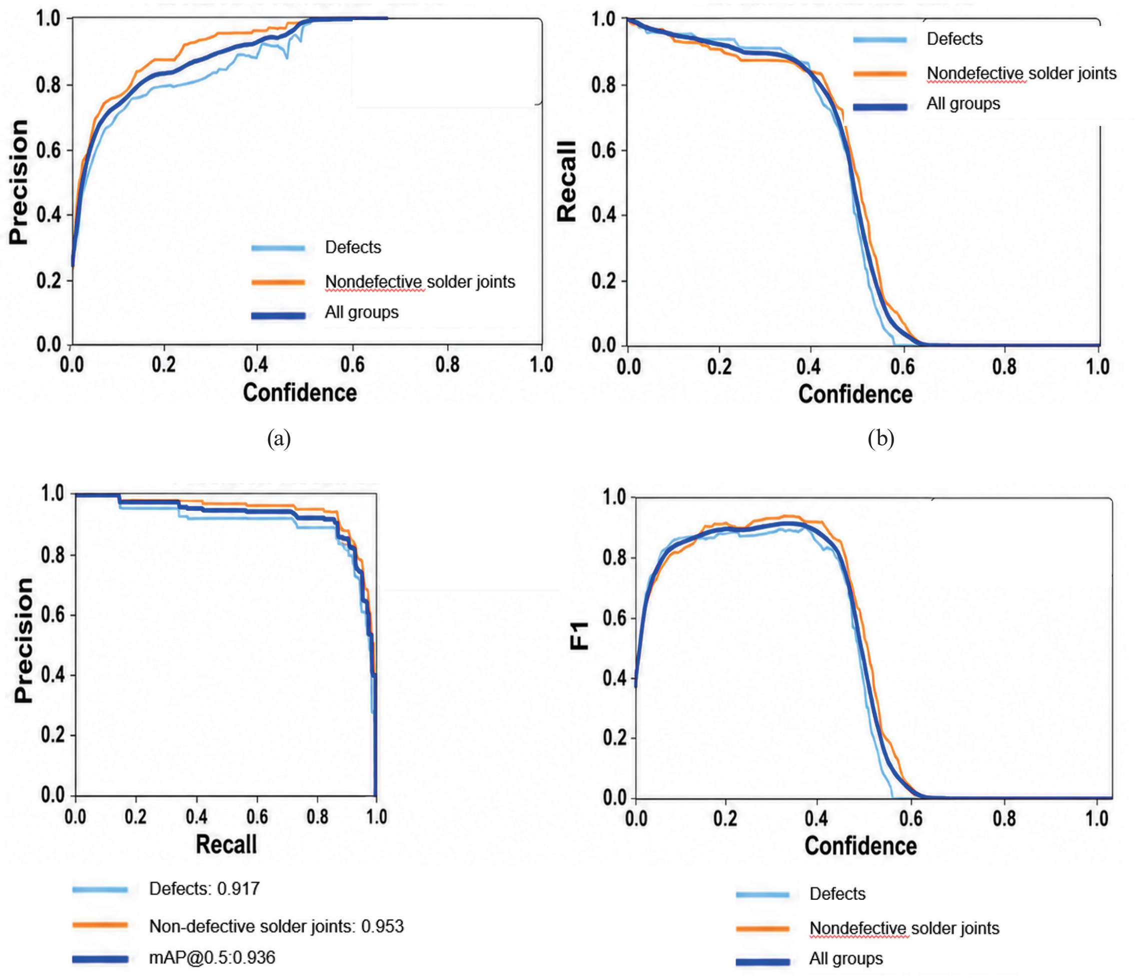


Fig. 12. (Color online) Pairwise distribution and correlation of annotation features.



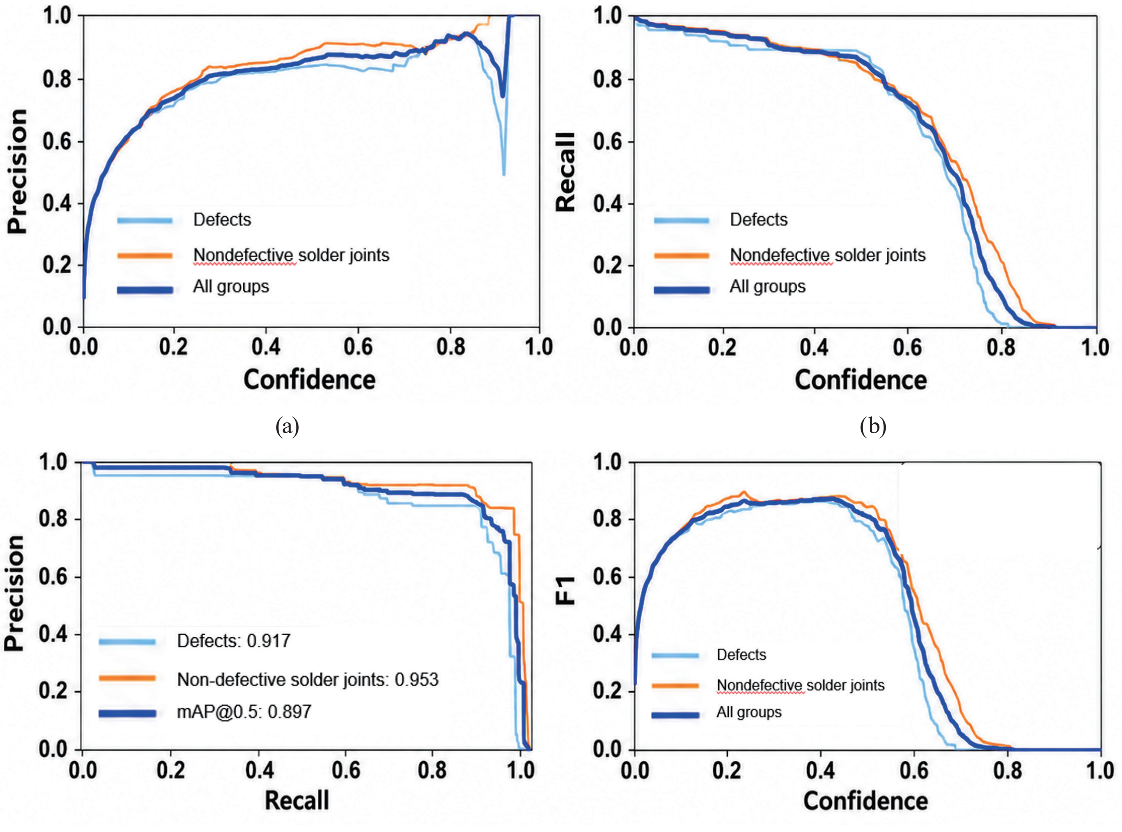
(a) (b)

(c) (d)

Fig. 13. (Color online) Evaluation curves for performance of YOLOv9-e model: (a) precision–confidence curve, (b) recall–confidence curve, (c) precision–recall curve, and (d) F1–confidence curve.

average of 0.936. These results indicate stable and high-quality detection across both defect and nondefect categories, while maintaining robustness against background noise. Finally, the F1–confidence curve shows that the model achieves its optimal F1-score of 0.90 at a confidence threshold of 0.4 [Fig. 13(d)]. This threshold is recommended for practical applications because it balances recall and precision, thereby improving overall detection performance in real-world inspection scenarios.

Performance evaluation curves were drawn to compare the inference performance of the YOLOv7 and YOLOv9 models. In this study, multiple versions of YOLOv7 were trained and analyzed, and their performance metrics were compared against those of YOLOv9-e. The precision–confidence curve shows that YOLOv9-e maintains a smooth and steadily increasing precision as the confidence threshold rises, indicating consistent prediction quality across different levels of confidence. By contrast, YOLOv7 shows notable fluctuations around a confidence level of 0.8, suggesting less stable predictions under stricter confidence requirements.



(a) (b)

(c) (d)

Fig. 14. (Color online) Evaluation curves for inference performance of YOLOv7 model: (a) precision–confidence curve, (b) recall–confidence curve, (c) precision–recall curve, and (d) F1–confidence curve.

Overall, YOLOv9-e produces a smoother curve but also achieves a more substantial precision improvement, highlighting its superior stability. The recall–confidence curve shows that YOLOv9 consistently achieves higher recall values across most confidence thresholds compared with YOLOv7, reflecting stronger defect detection capability [Fig. 14(b)]. The precision–recall and F1–confidence curves show that YOLOv9-e delivers more stable overall performance and achieves a better balance between precision and recall. These results collectively demonstrate that YOLOv9-e is more reliable and better suited for PCB defect inspection applications than YOLOv7.

5. Conclusions

By deploying an advanced automated imaging system, multi-angle optical data was captured efficiently and consistently, ensuring comprehensive spatial coverage of PCB surfaces. In parallel, deep-learning-based object detection algorithms were integrated to manage multiclass component recognition and solder joint defect identification. Through a multiphase evaluation strategy comparing sequential generations of the YOLO architecture, we proposed an accessible, high-performance AOI solution to lower deployment barriers for small and medium-sized electronics manufacturing enterprises, enhancing line throughput and final product quality.

The development of an integrated hardware–software workflow enabled real-time edge processing constraints to be balanced with high detection sensitivity for sub-millimeter solder joint anomalies. The developed system showed excellent performance in the classification and detection of PCB solder joint defects. The optimized YOLOv9 network achieved a maximum precision of 1.00 at a confidence threshold of 0.639 and a maximum recall of 1.00 at a confidence threshold of 0.000, validating its ability to minimize false positives while maintaining comprehensive defect coverage. A peak F1-score of 0.90 was recorded at a confidence threshold of 0.400, offering an operational balance recommended for practical factory-floor deployment. The system showed the mAP@0.5 of 0.936 across all categories, with individual benchmarks of 0.917 for defect regions and 0.955 for nondefective solder joints. Comparative analysis results showed that YOLOv9-e provides superior advantages over YOLOv7, exhibiting faster convergence rates across all training loss components, more stable validation curves, and an expanded mAP@0.5:0.95 boundary, which confirms the enhanced generalization capability.

Despite these advancements, challenges remain. Hardware resolution limitations and dataset scarcity were encountered during system implementation. When treating the entire captured substrate image as a single input tensor, the native pixel resolution proved insufficient for deep convolutional layers to accurately segment sub-millimeter anomalies. This limitation was mitigated by designing a localized spatial patching strategy, dividing larger high-resolution frames into 16 and 25 discrete subregions to amplify the relative pixel footprint of localized structural features. However, severe specular glare and limited sample variations for specialized component joints prevented reliable resolution of microscale lead-level wetting characteristics within the current scope. Dataset diversity is also required to accommodate complex, high-density industrial layouts, managing ambient illumination variations, and deploying the system into live, real-time inline AOI hardware setups. Through the integration of hardware-level polarization filters and the incorporation of attention-based deep neural network extensions, surface glare can be isolated from material boundaries, and the robustness and adaptability of the visual inspection system can be enhanced in dynamic industrial environments.

References

1 A. He and X. Jin: IEEE Trans. Reliab. **70** (2021) 1581. https://doi.org/10.1109/TR.2021.3090310
2 B. Ma, W. Cai, Y. Han, and G. Yu: IEEE Trans. Instrum. Meas. **70** (2021) 1. https://doi.org/10.1109/TIM.2021.3077965
3 M. J. Hasan, M. M. M. Islam, and J.-M. Kim: Measurement **138** (2019) 620. https://doi.org/10.1016/j.measurement.2019.02.075
4 M. Alom, T. M. Taha, C. Yakopcic, S. Westberg, P. Sidike, M. S. Nasrin, B. C. Van Essen, A. A. S. Awwal, and V. K. Asari: Electronics **8** (2019) 292. https://doi.org/10.3390/electronics8030292
5 L. Xu, E. Xu, and L. Li: Int. J. Prod. Res. **56** (2018) 2941. https://doi.org/10.1080/00207543.2018.1444806
6 S. Aldirmaz-Colak, M. Namdar, A. Basgumus, S. Özyurt, S. Kulac, and N. Calik: IEEE Access **13** (2025) 97152. https://doi.org/10.1109/ACCESS.2025.3573371
7 G. Zhang, Z. Li, C. Tang, J. Li, and X. Hu: SSRN Electron. J. (2024) 4857945. https://doi.org/10.2139/ssrn.4857945
8 J.-S. Park, M.-S. Lee, C.-Y. Park, Y.-S. Kim, H.-J. Song, and J.-D. Kim: Sens. Mater. **32** (2020) 2345. https://doi.org/10.18494/sam.2020.2806
9 W.-H. Lee, C.-C. Chen, and J. Jiang: Sens. Mater. **38** (2026) 935. https://doi.org/10.18494/sam5789
10 I.-J. Ding, M.-C. Hsieh, and Z.-X. Wu: Sens. Mater. **35** (2023) 1099. https://doi.org/10.18494/sam4236
11 R. Tan, Y. Liu, H. Zhang, and X. Chen: Sens. Mater. **37** (2025) 2385. https://doi.org/10.18494/sam5574

12 Z. He, Y. Lian, Y. Wang, and Z. Lu: Results Eng. **27** (2025) 106437. https://doi.org/10.1016/j.rineng.2025.106437

13 Q. Lin, Z. Xie, K. Takamasu, and M. Chen: Infrared Phys. Technol. **154** (2026) 106357. https://doi.org/10.1016/j.infrared.2025.106357

14 Workforce Development Agency, Ministry of Labor, Taiwan: https://www.wdasec.gov.tw (accessed May 2026).

15 C.Y. Wang, I.H. Yeh, and H.Y.M. Liao: Lect. Notes Comput. Sci. **15089** (2025) 1. https://doi.org/10.1007/978-3-031-72751-1_1

16 G. Ikmel and E. A. E. I. Najiba: Proc. 2024 International Conference on Ubiquitous Networking (IEEE, 2024) 1. https://doi.org/10.1109/UNET62310.2024.10794724

17 A. Bochkovskiy, C.Y. Wang, and H.Y.M. Liao: arXiv:2004.10934 (2020). https://doi.org/10.48550/arXiv.2004.10934

18 P. Cheng, X. Tang, W. Liang, Y. Li, W. Cong, and C. Zang: Lect. Notes Comput. Sci. **14357** (2023) 5. https://doi.org/10.1007/978-3-031-46311-2_5

19 C.-Y. Wang, A. Bochkovskiy, and H.-Y.M. Liao: Proc. IEEE/CVF Conference on Computer Vision Pattern Recognition (IEEE, 2023) 7464. https://doi.org/10.1109/CVPR52729.2023.00721

20 C.-Y. Wang, I.-H. Yeh, and H.-Y.M. Liao: arXiv:2402.13616 (2024). https://doi.org/10.48550/arXiv.2402.13616

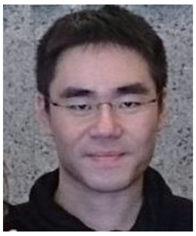
21 C.-H. Shih and Y.-Y. Xu: Sens. Mater. **37** (2025) 117. https://doi.org/10.18494/sam5325

22 Y. Wang, X. Liu, H. Zhang, J. Li, and T. Chen: Comput. Mater. Continua **80** (2024) 4973. https://doi.org/10.32604/cmc.2024.055839

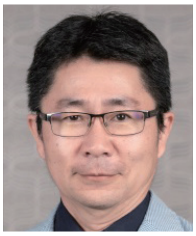
23 Z. Zheng, P. Wang, W. Liu, J. Li, R. Ye, and D. Ren: IEEE Trans. Cybern. **52** (2021) 8574. https://doi.org/10.1109/TCYB.2021.3095305

24 C.-C. Hsieh, Y.-W. Lin, L.-H. Tsai, W.-H. Huang, S.-L. Hsieh, and W.-H. Hung: Sens. Mater. **32** (2020) 3429. https://doi.org/10.18494/sam.2020.2921

About the Authors



Yi-Heng Wu received his B.S. degree in computer science and engineering from Tungnan University, Taipei, Taiwan, in 2006, and his M.S. degree from the Graduate School of Mathematical Sciences at Aletheia University, Taipei, Taiwan, in 2008. He obtained his Ph.D. degree in computer science and information engineering from Tatung University, Taipei, Taiwan, in 2011. He is currently an assistant professor at the Department of Computer Information and Network Engineering at Lunghwa University of Science and Technology, Taoyuan, Taiwan. His research interests include virtual reality, augmented reality, multimedia information networks, and information security. (zzz7136@mail.lhu.edu.tw)



Chung-Wen Hung received his Ph.D. degree in electrical engineering from National Taiwan University, Taipei City, Taiwan, in 2006. He is currently a professor and the director of the Advanced Instrumentation Center at National Yunlin University of Science and Technology, Yunlin, Taiwan. His research interests include AI landing, IoT, and Industrial IoT. (wenhung@yuntech.edu.tw)

Yi-An Chen is currently an undergraduate student in the Department of Electrical Engineering at National Yunlin University of Science and Technology, Taiwan. His research interests include AI-based recognition and its applications. (b11112151@yuntech.edu.tw)

Lung-Fa Wu received his Ph.D. degree in automation and control from National Taiwan University of Science and Technology, Taipei, Taiwan, in 2022. He is currently an assistant professor with the Department of Electrical Engineering of National Ilan University, Yilan, Taiwan. His research interests include AI, deep learning, and optomechanical system integration. (lfwu@niu.edu.tw)