

Wire bonding object detection through deep learning architecture of oriented bounding boxes with YOLOv8

Yuen Chin Wong¹, Hea Choon Ngo¹, Umami Raba'ah Hashim², Mohd Hilmi Hasan³

¹Department of Intelligent Computing and Analytics, Faculty of Artificial Intelligence and Cyber Security, Universiti Teknikal Malaysia Melaka, Melaka, Malaysia

²Department of Software Engineering, Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka, Melaka, Malaysia

³Department of Computing, Centre for Research in Data Science, Universiti Teknologi Petronas, Seri Iskandar, Malaysia

Article Info

Article history:

Received Jul 18, 2025

Revised Aug 3, 2026

Accepted Aug 20, 2026

Keywords:

Efficient multi-scale attention
Feature pyramid network
Oriented bounding box
Semiconductor manufacturing
Wire bond detection
You only look once v8

ABSTRACT

Wire bonding inspection is a critical process in semiconductor packaging to ensure product quality and reliability. This study compares you only look once version 5 (YOLOv5) axis-aligned bounding box (AABB) and YOLOv8 oriented bounding box (OBB) for wire bond defect detection. The dataset comprises 1,063 images with 14,778 annotated bonding instances in both AABB and OBB formats. Experimental results are evaluated using precision, recall, F1-score, mAP@0.5, mAP@0.5:0.95, and inference time. The findings show that YOLOv8 OBB achieved superior recall (0.9904) and mAP@0.5:0.95 (0.6787) compared to YOLOv5 AABB, demonstrating improved localization of rotated and dense wire bonds. Customized YOLOv8 variants with feature pyramid network (FPN)+efficient multi-scale attention (EMCA) and spatial squeeze and excitation (SpatialSE)+EMCA were also tested, with FPN+EMCA achieving the best balance of accuracy and robustness. This work highlights the potential of OBB-based object detection in semiconductor manufacturing and establishes a foundation for future real-time inspection systems.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Hea Choon Ngo

Department of Intelligent Computing and Analytics, Faculty of Artificial Intelligence and Cyber Security
Universiti Teknikal Malaysia Melaka

Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia

Email: heachoon@utem.edu.my

1. INTRODUCTION

Semiconductor backend manufacturing, which comprises assembly and final test processes as shown in Figure 1, particularly wire bonding, is a critical stage in integrated circuit (IC) manufacturing where chips are interconnected with substrates. As packaging density increases, the detection of defects in wire bonds has become increasingly challenging due to higher packaging density, smaller pad sizes, and rotated or overlapping layouts [1], [2]. Bonding defects can cause electrical failure, reduced yield, and costly rework, necessitating robust automated inspection methods. Accurate and efficient inspection is therefore essential. Conventional inspection methods, including manual visual inspection and rule-based computer vision, are labor-intensive and often fail under high-density or rotated bonding conditions, leading to low accuracy and scalability issues [3], [4].

Deep learning has revolutionized industrial defect detection, with convolutional neural network (CNN)-based object detectors delivering significant improvements in robustness and scalability [5]. Recent advances in deep learning, particularly the you only look once (YOLO) family of detectors, have

significantly improved object detection accuracy and efficiency [6]. However, existing YOLO models up to YOLOv5 rely on axis-aligned bounding boxes (AABB), which struggle to accurately capture rotated or overlapping bonds [7]. YOLOv8 introduces oriented bounding boxes (OBBs), enabling orientation-aware detection more suitable for wire bond geometry. While OBB detection has been successfully applied in aerial imagery [8], PCB inspection [9], and scene text recognition [10], its potential for semiconductor wire bond inspection has not been fully explored.

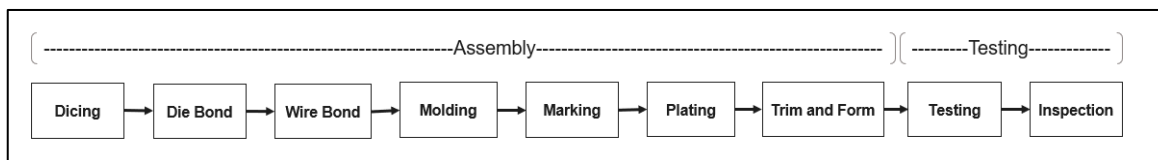


Figure 1. Semiconductor backend manufacturing process flow

Recent studies have also introduced architectural enhancements, such as feature pyramid networks (FPN) for multi-scale feature fusion [11], efficient multi-scale attention (EMCA) for context-aware feature refinement [12], and squeeze-and-excitation (SE) modules for channel recalibration [13]. These mechanisms have improved detection accuracy in general domains, yet their effectiveness within OBB-based architectures for semiconductor defect detection has not been systematically studied. This research addresses these gaps by proposing an enhanced YOLOv8 OBB-based framework for wire bond detection with the following contribution: i) development of a wire bond dataset annotated in both AABB and OBB formats, ii) comparative evaluation of YOLOv5 AABB and YOLOv8 OBB models for wire bond defect detection, and iii) integration of customized YOLOv8 models with FPN+EMCA and spatial squeeze and excitation (SpatialSE)+EMCA modules to improve detection performance.

2. METHOD

This research presents a new and easily customized approach which eases the potential of switching to improved versions of model architecture in shorter time to address the current issue. Three main phases of the research methodology are demonstrated in Figure 2, can be divided into three distinct stages. These stages form a systematic problem-solving pathway as the future guidance for future researcher.

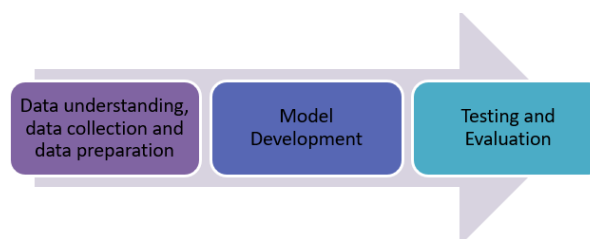


Figure 2. Three main phases of the research methodology

In the earlier stage, a clear understanding of the challenges associated is developed with rotational variations and bond structures in current chips, followed by prioritizing the collection of a representative dataset that can be utilized for subsequent testing and evaluation. First, the inputs images data required to be annotated, using labelling tool for bounding box drawing. After images is labeled, it will required to be transformed into the required structure for model training for both YOLOv5 and YOLOv8. The same input images dataset will be prepared to load into both model for comparison. However, the annotation are different, as YOLOv5 required an AABB and YOLOv8 that we intent to trained required OBB. The annotation served as the most time taxing process to product sufficient data for model training. Then, the annotated data will be submitted to data normalization and data splitting to train and validate the model performance. The preprocessed dataset will be submitted for a model using a pre-trained model of YOLOv8-obb.pt on DOTA dataset with customize adaptation on classification. An AABB YOLOv5 will also be trained with similar parameters for comparison. In this phase, the switching algorithm design will be developed in order to support the hybrid approach and the explanation on how the design is selected.

In testing and evaluation, the performance of the new proposed model with extensive experiments is verified. The result of two customized YOLOv8 architectures: FPN+EMCA and SpatialSE+EMCA models' architecture were evaluated based on performance over the YOLOv8 base-model architecture. The evolution of object detection is marked by two major phases. The traditional era before 2014, relying on handcrafted computer vision techniques like edge detection, template matching, and feature extraction methods often limited by their sensitivity and resources. While the deep learning era post-2014, which revolutionized performance through models like CNNs and R-CNNs [14]. The advent of deep neural networks, especially with the introduction of YOLO in 2016, enabled real-time, high-precision object detection suitable for practical applications [15]. Deep architectures are broadly categorized into discriminative models, like CNNs and RNNs [16], which excel in classification tasks, and generative models, such as autoencoders and deep belief networks, used for data generation and augmentation.

While R-CNN and its variants, Faster R-CNN [17] deliver high accuracy, their two-stage architecture demands significant computational power, limiting real-time application. The emergence of YOLO in 2015 [6], [18] revolutionized object detection by using a single neural network to simultaneously predict class probabilities and bounding boxes, achieving much higher speeds and making real-time deployment feasible. Particularly at YOLOv5 addressed some limitations by introducing anchor box learning and customized spatial attention mechanisms [19]. The latest, YOLOv8, further enhances detection through anchor-free architectures, multiscale prediction, and efficient backbone modules. Demonstration on adding an angle prediction regressor to YOLOv5 on the DOTA dataset does show improvement with mAP but only reached about 75.6% [20], indicating YOLOv5 may not be ideal for applications requiring precise rotational object detection despite its computational speed benefits. In 2023, YOLOv8 became the first version of YOLO introduces significant advancements over its predecessors by incorporating OBB [21] to enhance object detection accuracy, especially for tasks involving rotated or non-axis-aligned objects.

2.1. Dataset and annotation

A dataset of 1,063 high-resolution wire bond images was compiled, comprising 14,778 annotated bonding instances. The dataset includes variations in bond orientation, density, and lighting to ensure robustness. Annotation was conducted in two phases: i) manual annotation of 100 seed images by domain experts and ii) model-assisted annotation, where YOLOv8-generated predictions were refined to produce high-quality labels. Two annotation formats were prepared: AABB, used for YOLOv5, and OBB, used for YOLOv8. The dataset was split into 70% training, 30% testing. The purpose of splitting data into training and test dataset is to evaluate the performance of new developed model on unseen data. OBB is preferable in the two-stage detector as well and observed achieving better in overall object detection performance and achieving lower computational consumption [22] compared to AABB. AABB represents axis-aligned rectangles defined by four normalized values for center coordinates, width, and height. Conversely, OBB features a bounding box that can be rotated and not necessarily aligned with the axes, characterized by parameters including width, height, and rotation angle, with three types used in YOLOv8: Θ -based, point-based, and h-based. Figure 3 presents examples of the AABB and OBB annotation formats applied to wire bond images, including the corresponding object classes. Figure 3(a) shows the AABB and OBB annotation on wire bond. Figure 3(b) shows the sample of AABB-annotated image and classes, while Figure 3(c) shows the sample of OBB-annotated image and classes.

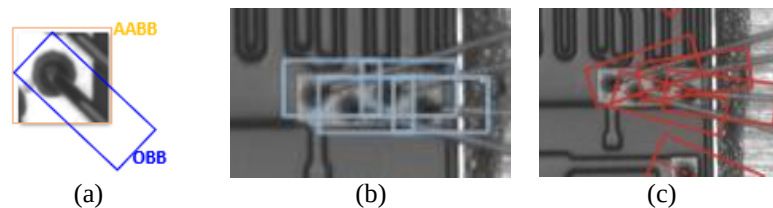


Figure 3. AABB and OBB annotations on wire bonds; (a) AABB and OBB annotations on wire bonds, (b) sample of AABB-annotated image and classes, and (c) sample of OBB-annotated image and classes

2.2. Model architectures

In this study, four models were evaluated: i) YOLOv5 AABB baseline, ii) YOLOv8 OBB baseline, iii) YOLOv8 OBB (FPN+EMCA), and iv) YOLOv8 OBB (SpatialSE+EMCA). YOLOv8, a lightweight backbone for efficient feature extraction, a neck for multi-scale feature aggregation [23], and a detection head

that predicts bounding boxes, class probabilities, and objectness scores. Compared to YOLOv5, YOLOv8 introduces enhanced modules such as improved convolutional blocks, better feature fusion techniques, and more refined anchor and bounding box prediction methods. These advancements allow YOLOv8 to achieve higher accuracy, better generalization across diverse datasets, and faster inference times, making it more effective in real-time object detection tasks. Figure 4 shows the customized YOLOv8 architectures. Figure 4(a) shows the architecture of FPN+EMCA YOLOv8 OBB model, while Figure 4(b) shows the architecture of SpatialSE+EMCA YOLOv8 OBB model. The performance of these two models will then be compared based on these metrics to evaluate their relative strengths.

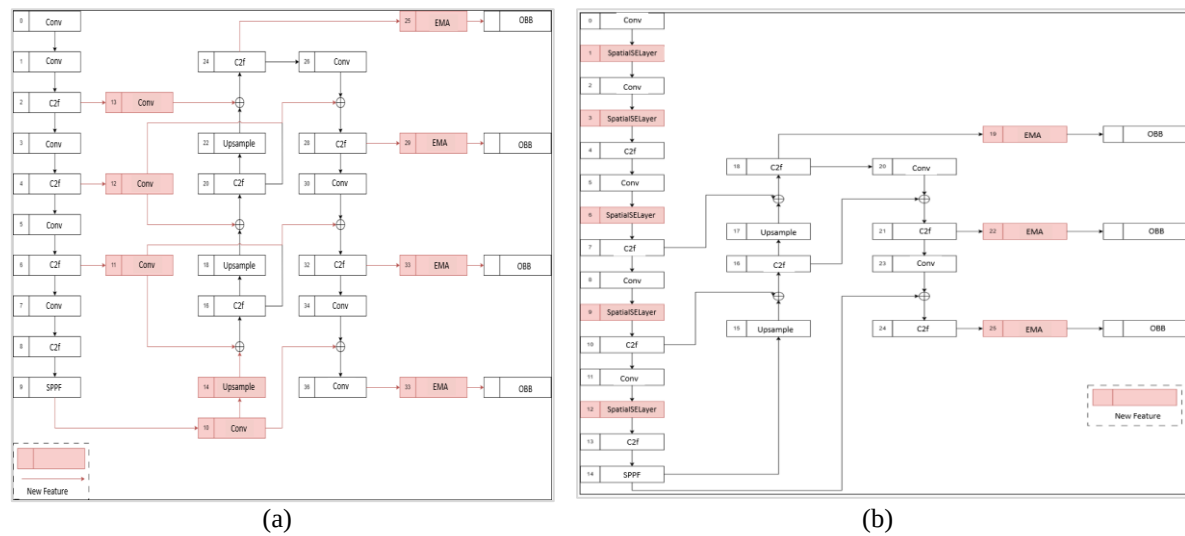


Figure 4. Customized YOLOv8 architectures of; (a) FPN+EMCA and (b) SpatialSE+EMCA

2.2.1. Efficient multi-scale attention

Traditional object detection methods often struggle with varying object sizes, leading to inaccuracies, which EMCA aims to address by enhancing multi-scale feature learning. Integrating EMCA into object detection frameworks enhances detection accuracy and computational efficiency, as demonstrated in a 2021 study showing improved mAP and processor efficiency when compared to standard models [24]. Its lightweight design ensures effective real-time operation and adaptability to diverse environmental conditions, such as lighting changes or occlusions, advancing the robustness of object detection applications.

2.2.2. Feature pyramid network

YOLOv8 features an advanced architecture comprising a backbone for feature extraction, a neck for feature aggregation, and a head for generating object detections, which collectively improve both speed and accuracy in object detection tasks. However, it struggles with detecting small and tightly packed objects due to scale issues and insufficient pixel data, often resulting in misclassifications or undetected items within cluttered scenes. These limitations, particularly pronounced in instances where objects are smaller than 8×8 pixels [25], lower the mean average precision (mAP) and overall efficacy, underscoring the need for architectural optimizations. Proposed enhancements, such as integrating FPN [26] could utilize a top-down approach with lateral connections to merge upsampled high-level features with lower-level features, enhancing feature representation and improving the detection of small objects. This design is adopted in ProFPN [27], resulting in significant reductions in FLOPs while maintaining effective feature utilization.

2.2.3. Spatial squeeze and excitation

The SE block [13], enhances CNNs by adaptively recalibrating feature map channels to focus on the most informative features, thus improving detection of small and variable-sized objects in complex environments. SE blocks have been successfully applied not only in object detection but also in segmentation tasks like U-Net, yielding significant accuracy improvements with minimal complexity increase [28]. For instance, integrating SE blocks into YOLO (e.g., ISE-YOLO) has shown positive gains in detection accuracy and inference speed.

2.3. Training setup and evaluation metrics

All models were trained on an NVIDIA RTX 3090 GPU (24 GB VRAM) using PyTorch 2.0. The training configuration was optimizer: SGD with momentum 0.937; initial learning rate: 0.01 (cosine decay); batch size: 16; epochs: 100; input resolution: 1280×1280; data augmentation: random flips, ±15° rotations, brightness and contrast jitter. Performance was evaluated using precision, recall, F1-score, mAP@0.5, mAP@0.5:0.95, and inference time (FPS). Robustness was ensured through 5-fold cross-validation, reporting 95% confidence intervals. Paired t-tests were conducted to assess statistical significance.

3. RESULTS AND DISCUSSION

This study clearly outlines two primary objectives. The first involves training a YOLOv8 OBB model using a newly developed OBB image dataset. The model's performance will be compared to that of the existing YOLOv5 ABB model to validate the hypothesis that YOLOv8 OBB addresses the limitations of YOLOv5 ABB, particularly in detecting rotated and densely packed objects. An additional notable contribution is the implementation of a model-assisted labelling algorithm, which enhances the efficiency of data annotation. As electronic components continue to miniaturize and increase in complexity, there is a growing need for an optimized YOLOv8 OBB model capable of accurately detecting smaller objects. Motivated by this, the second objective focuses on evaluating the performance of two improved YOLOv8 OBB architectures—enhanced FPN+EMCA and SpatialSE+EMCA—in terms of detection accuracy and efficiency for small objects.

A diverse and well-annotated dataset enables the trained model to generalize effectively across various conditions, such as different lighting, angles, and occlusions, while a limited dataset risks overfitting and poor performance on new images. The annotation process was significantly streamlined through a new developed model-assisted labelling algorithm on YOLOv8 OBB model, inspired by tools used for YOLOv5 ABB [29]. This algorithm utilizes pretrained models to predict object locations, reducing manual effort. Initially, manual annotation was performed on 100 images, which trained the first version of the YOLOv8 OBB model. This model then predicted bond locations for an additional 50 images, achieving over 70% accuracy. This accuracy level reduced manual validation to 30% of the annotations, greatly improving efficiency. The iterative process of retraining and model-assisted annotation further enhanced detection performance and consistency, ultimately leading to a total of 1,063 annotated images with 14,778 annotated first and second bonds. With an increase of annotated set close to 265% more than initial planning, highly generalized YOLOv8 OBB models are trained.

The upcoming sections evaluate the performance of YOLOv8 and YOLOv5 models, focusing on their ability in object detection especially on dense rotational bonds. For both objectives, thorough assessment between them using key metrics, including:

- Precision represents the ratio of true positive predictions over the total number of positive observations:

$$Precision = \frac{TP}{TP+FP} \quad (1)$$

- Recall, also known as sensitivity, represents the ratio of true positives to the number of actual positive observations:

$$Recall = \frac{TP}{TP+FN} \quad (2)$$

- F1-score combining precision and recall into single value for a balanced measure, and normally use when both false positives and false negatives are important to the consideration:

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision+Recall} \quad (3)$$

- mAP@0.5 represents the metric to evaluate objection with measurement of accuracy of the model correctly detecting and locating the objects within images at IoU [30], [31], the overlapping between 2 boundaries, with a threshold 0.5.
- mAP_0.5:0.95 computing the average precision for each class at each IoU threshold in the interval of 0.5, up to 0.95. These metrics provide a comprehensive assessment of each model's detection capabilities, especially when balancing the importance of false positives and false negatives.

The experimental approach used involves training a YOLOv5 model using a similar dataset with ABB annotations for 100 epochs at a standardized high resolution [22] input size of 1280 pixels.

3.1. Baseline comparison

The experiment result shown in Table 1, YOLOv8 OBB model consistently outperformed YOLOv5 ABB across all key metrics, except precision. YOLOv5 achieved higher precision (0.99413) but slightly lower recall (0.98978) compared to YOLOv8 (precision 0.99374, recall 0.99041). YOLOv8 outperformed YOLOv5 in mAP@0.5:0.95 (0.67873 vs. 0.66317), indicating better localization of rotated and dense wire bonds. These improvements were statistically significant ($p < 0.05$).

Table 1. Performance comparison between YOLOv5 ABB and YOLOv8 OBB

Models	Metrics/precision	Metrics/recall	Metrics/F1-score	Metrics/mAP_0.5	Metrics/mAP_0.5:0.95
YOLOv5 ABB	0.99413	0.98978	0.99195	0.99074	0.66317
YOLOv8 OBB	0.99374	0.99041	0.99207	0.99194	0.67873

Based on the mAP results for the YOLOv5 ABB and YOLOv8 OBB models, it's observed that YOLOv8 OBB outperforms YOLOv5 ABB in mAP@0.5 with a value of 0.99194 compared to 0.99074. This indicating that YOLOv8 achieves higher detection accuracy at a moderate IoU threshold. In another view from the mAP@0.5:0.95 metric, which evaluates performance across a range of IoU thresholds, further emphasizes the robustness of YOLOv8 OBB, yielding a score of 0.67873 versus YOLOv5 ABB's 0.66317.

In addition to accuracy, another significant factor used to determine the suitability of a object detection models viability for real-time inspection process, called operational efficiency. This efficiency best represented by the measurement of inference time and frames per second (FPS) to determine model efficiency performance. Figure 5 shows two histograms of inference time distribution over 344 images with some basic statistics of two evaluated models. Figure 5(a) show the inference time statistics for YOLOv5 ABB model, while Figure 5(b) shows the inference time statistics for YOLOv8 OBB model. As the assessment of inference time of both models in the same platform, the statistic could still be a reference. YOLOv5 showcased an average inference time of 0.603s, translating to approximately 1.658 FPS. In contrast, YOLOv8 displayed a longer average inference time of 1.063s and an FPS of 0.941, indicating the shortcoming of YOLOv8 OBB model.

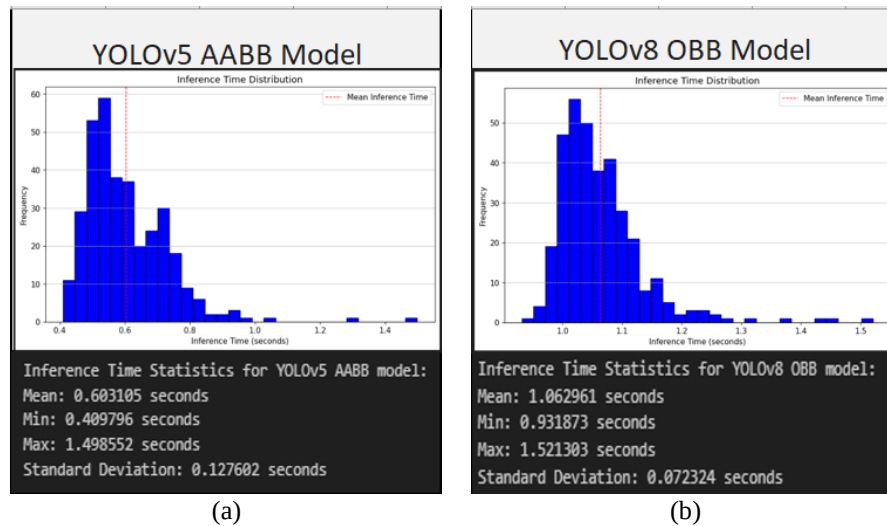


Figure 5. Comparison of inference time between; (a) YOLOv5 ABB and (b) YOLOv8 OBB model

There is two important statistic figures sound interesting. First is the maximum value. The maximum inference time taken between 2 models are close, with just a 0.03s gap. This could indicates in the condition of complex bond layout, YOLOv8 likely could perform more efficiently compared to YOLOv5 ABB. Secondly, the standard deviation YOLOv8 obtained a inference time o f 0.072s, which is lower than 0.127s in YOLOv5 ABB. This longer standard deviation observed in YOLOv5 ABB model is a clear insight that the model faces inconsistencies in performance or that the model may struggle occasionally with certain inputs.

3.2. Customized architectures

In the previous section, the result confirmed that YOLOv8 OBB model outperforms YOLOv5 ABB model in accuracy. However, YOLOv8 base model does struggles with detecting small objects due to scale issues and feature representation challenges in cluttered scenes. To address these limitations, 2 different optimizations suggested over the YOLOv8 OBB base model, first named as FPN+EMCA, which expects FPN to improve the model's sensitivity to small objects through multi-scale feature processing [32], while EMCA aims to achieve a better focus on relevant features, in helping to detect varying object sizes. The second optimized YOLOv8 also making use of the EMCA on feature attention improvement, but integrated PartialSE into the head. The SE block aims to enhance channel relationships by using global features which involved 2 main steps: "squeeze" and "excitation". This SpatialSE+EMCA setup allows networks to emphasize the right channels without being overly simplistic.

Table 2 presents results of customized YOLOv8 models. It is observed that YOLOv8 (FPN+EMCA) hold the highest precision at 99.5% and highest recall rate of 99.12%. Additionally, YOLOv8 with FPN+EMCA achieves the highest performance indicating that this optimized architecture setup is effective in handling object detection tasks. The combination of FPN+EMCA enhanced precision and recall, resulting a superior F1-score and mAP values. In another optimized version of YOLOv8, SpatialSE+EMCA generally achieved slightly higher F1-score comparing to ordinary YOLOv8, existing YOLOv5 ABB. The strength of SpatialSE+EMCA seem not significant improvement, but more to comparable to ordinary YOLOv8 as expected may be due to data characteristics or the misalignment with the feature hierarchies.

Table 2. Results of customized YOLOv8

Models	Metrics/precision	Metrics/recall	Metrics/F1-score	Metrics/mAP_0.5	Metrics/mAP_0.5:0.95
YOLOv8 OBB (FPN+EMCA)	0.99523	0.99122	0.993221	0.9921	0.68102
YOLOv8 OBB (SpatialSE+EMCA)	0.99428	0.99054	0.992406	0.9923	0.67312

In summary, the enhancements in YOLOv8, particularly with the integration of FPN+EMCA technique, contribute significantly on better precision and recall rate on detecting small and multi-scale objects. This new optimized YOLOv8 FPN+EMCA would be able to handle increasing complexity of bonds detection from various chip layout. Key contributions include demonstrating how customized architecture modifications can enhance performance, supported by recent studies [33], [34], emphasizing the importance of a deep understanding of model design and domain expertise. As the future trend of in semiconductor design on downscaling chip size, and more complex packaging technologies on chip on chip, chip by chip, and even modular layout, object detection will become even more importance to be automated and supported by AI. Thus the robustness of object detection model required to be always monitored and evaluated. The trade-off between inference speed and detection accuracy is critical for deployment. YOLOv5 is better suited for applications requiring real-time, high-throughput inspection, while YOLOv8 particularly with FPN+EMCA is ideal for high-accuracy defect detection in critical production stages.

4. CONCLUSION

This study introduced an OBB-based framework for wire bond object detection in semiconductor manufacturing. Comparative results showed that YOLOv8 OBB significantly outperforms YOLOv5 ABB in detecting rotated and dense bonding layouts, though with slower inference. Among the customized architectures, YOLOv8 with FPN+EMCA achieved the best performance, confirming the value of multi-scale feature attention. The contributions of this work include: i) development of a specialized wire bond dataset annotated in both ABB and OBB formats, ii) demonstration that YOLOv8 OBB achieves superior detection accuracy and robustness compared to YOLOv5 ABB, and iii) validation of FPN+EMCA as an effective architecture for enhancing wire bond inspection. Future work should focus on expanding the dataset to capture broader variability, developing lightweight architectures for real-time inspection, and deploying the system in industrial production environments.

ACKNOWLEDGEMENTS

The authors would like to thank the Universiti Teknikal Malaysia Melaka (UTeM) for their invaluable support.

FUNDING INFORMATION

The authors express their gratitude to the Universiti Teknikal Malaysia Melaka (UTeM) for its financial support of this research. This funding has contributed to covering publication costs and providing research facilities.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Yuen Chin Wong	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	
Hea Choon Ngo		✓				✓		✓	✓	✓	✓	✓		✓
Umami Raba'ah Hashim	✓		✓				✓			✓		✓		✓
Mohd Hilmi Hasan				✓	✓		✓			✓	✓		✓	

C : **C**onceptualization

M : **M**ethodology

So : **S**oftware

Va : **V**alidation

Fo : **F**ormal analysis

I : **I**nterpretation

R : **R**esources

D : **D**ata Curation

O : **O**riginal Draft

E : **E**diting

Vi : **V**isualization

Su : **S**upervision

P : **P**roject administration

Fu : **F**unding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

Data will be made available on request.

REFERENCES

- [1] Y. S. Zhou, C. L. Gan, M. H. Chung, and H. Takiar, "A review of interconnect materials used in emerging memory device packaging first-and second-level interconnect materials," *Journal of Materials Science in Electronics*, vol. 32, no. 23, pp. 27133-27147, Oct. 2021, doi: 10.1007/s10854-021-07105-9
- [2] J. Pan and P. Fraud, "Wire bonding challenges in optoelectronics packaging," in *Proceedings of the 1st SME Annual Manufacturing Technology Summit: Dearborn*, 2004, pp. 1-8.
- [3] N. G. Shankar and Z. W. Zhong, "A rule-based computing approach for the segmentation of semiconductor defects," *Microelectronics Journal*, vol. 37, no. 6, pp. 500-509, Jun. 2006, doi: 10.1016/j.mejo.2005.07.018.
- [4] L. Alam and N. Kehtarnavaz, "A survey of detection methods for die attachment and wire bonding defects in integrated circuit manufacturing," *IEEE Access*, vol. 10, pp. 83826-83840, 2022, doi: 10.1109/access.2022.3197624.
- [5] C. C. Lim, H. C. Ngo, U. R. A. Hashim, and M. H. Hasan, "Surface-mount device design cycle time reduction using hybrid predictive modeling and optimization algorithm," *Bulletin of Electrical Engineering and Informatics*, vol. 14, no. 4, pp. 2889-2898, Aug. 2025, doi: 10.11591/eei.v14i4.9058.
- [6] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, real-time object detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 779-788, Jun. 2016.
- [7] Z. Zheng, T. Zhang, Z. Liu, B. Liu, Y. Li, and C. Sun, "Arbitrarily oriented object detection in remote sensing images based on improved YOLOv4-CSP," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 9355-9368, Apr. 2022, doi: 10.1109/JSTARS.2022.3214541.
- [8] J. Leng *et al.*, "Recent Advances for aerial object detection: a survey," *ACM Computing Surveys*, vol. 56, no.12, pp.1-36, Jul. 2024, doi: 10.1145/3664598.
- [9] K. Fu, Z. Chang, Y. Zhang, G. Xu, K. Zhang, and X. Sun "Rotation-aware and multi-scale convolutional neural network for object detection in remote sensing iamges," *ISPRS Journal of Photogrammetry and Remote Sensing*, 161, pp. 294-308, Feb. 2020, doi: 10.1016/j.isprsjprs.2020.01.025.
- [10] Q. Ling, N. A. M. Isa, and M. S. M. Asaari, "Precise detection for dense PCB components based on modified YOLOv8," *IEEE Access*, vol. 11, pp. 116545-116560, Oct. 2023, doi: 10.1109/access.2023.3325885.
- [11] T. Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 2117-2125, doi: 10.1109/cvpr.2017.106.
- [12] B. Khalili and A. W. Smyth, "SOD-YOLOv8—Enhancing YOLOv8 for Small Object Detection in Traffic Scenes," *arXiv*, Aug. 2024, doi: 10.48550/arXiv.2408.04786.
- [13] J. Hu, L. Shen, and G. Sun, "Squeeze-and-Excitation Networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018, pp. 7132-7141, doi: 10.1109/cvpr.2018.00745.
- [14] Z. Zou, K. Chen, Z. Shi, Y. Guo, and J. Ye, "Object Detection in 20 Years: A Survey," in *Proceedings of the IEEE*, vol. 111, no. 3, pp. 1-20, 2023, doi: 10.1109/JPROC.2023.3238524.

- [15] U. Sirisha, S. Praveen, P. N. Srinivasu, P. Barsocchi, and A. K. Bhoi, "Statistical Analysis of Design Aspects of Various YOLO-Based Deep Learning Models for Object Detection," *International Journal of Computational Intelligence Systems*, vol. 16, no. 1, Aug. 2023, doi: 10.1007/s44196-023-00302-w.
- [16] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Region-Based Convolutional Networks for Accurate Object Detection and Segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 1, pp. 142–158, Jan. 2016, doi: 10.1109/tpami.2015.2437384.
- [17] M. Safaldin, N. Zaghdien, and M. Mejdoub, "An Improved YOLOv8 to Detect Moving Objects," *IEEE Access*, pp. 1–1, Jan. 2024, doi: 10.1109/access.2024.3393835.
- [18] M. Hussain, "YOLO-v1 to YOLO-v8, the Rise of YOLO and Its Complementary Nature toward Digital Manufacturing and Industrial Defect Detection," *Machines*, vol. 11, no. 7, p. 677, Jul. 2023, doi: 10.3390/machines11070677.
- [19] Z. Li, "Road Aerial Object Detection Based on Improved YOLOv5," in *International Conference on Computer, Big Data and Artificial Intelligence (ICCBDAI 2021)*, vol. 2171, no. 1, p. 012039, Jan. 2022, doi: 10.1088/1742-6596/2171/1/012039.
- [20] Q. Hu, R. Li, C. Pan, and O. Gao, "Remote Sensing Image Object Detection Based on Oriented Bounding Box and YOLOv5," in *2022 IEEE 10th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)*, pp. 657–661, Jun. 2022, doi: 10.1109/itaic54216.2022.9836953.
- [21] G. Wang, Y. Chen, P. An, H. Hong, J. Hu, and T. Huang "UAV-YOLOv8: A small-object-detection model based on improved YOLOv8 for UAV aerial photography scenarios," *Sensors*, vol. 23, no. 16, 7190, doi: 10.3390/s23167190.
- [22] Z. Fang, J. Ren, H. Sun, S. Marshall, J. Han, and H. Zhao, "SAFDet: A Semi-Anchor-Free Detector for Effective Detection of Oriented Objects in Aerial Images," *Remote Sensing*, vol. 12, no. 19, pp. 3225–3225, Oct. 2020, doi: 10.3390/rs12193225.
- [23] H. Yi, B. Liu, B. Zhao, and E. Liu, "Small Object Detection Algorithm Based on Improved YOLOv8 for Remote Sensing," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 17, pp. 1734–1747, Jan. 2024, doi: 10.1109/jstars.2023.3339235.
- [24] K. Blachut and T. Kryjak, "Real-Time Efficient FPGA Implementation of the Multi-Scale Lucas-Kanade and Horn-Schunck Optical Flow Algorithms for a 4K Video Stream," *Sensors*, vol. 22, no. 13, p. 5017, Jul. 2022, doi: 10.3390/s22135017.
- [25] Y. Hu, S. Wang, Y. Hong, and R. Li, "A YOLOv8s Model Integrating GELAN Module and HAT Attention for PCB Defect Detection," in *2024 5th International Conference on Electronic Communication and Artificial Intelligence (ICECAI)*, pp. 297–300, May 2024, doi: 10.1109/icecai62591.2024.10675021.
- [26] A. Li, R. Hamzah, A. Rahim, and Y. Gao, "YOLO Algorithm with Hybrid Attention Feature Pyramid Network for Solder Joint Defect Detection," *IEEE Transactions on Components Packaging and Manufacturing Technology*, vol. 14, no. 8, pp. 1493–1500, Jun. 2024, doi: 10.1109/tcpmt.2024.3409773.
- [27] J. Ke, L. He, B. Han, J. Li, and X. Gao, "ProFPN: Progressive feature pyramid network with soft proposal assignment for object detection," *Knowledge-Based Systems*, vol. 299, pp. 112078–112078, Jun. 2024, doi: 10.1016/j.knosys.2024.112078.
- [28] Z. Huangfu and S. Liu, "Lightweight You Only Look Once v8: An Upgraded You Only Look Once v8 Algorithm for Small Object Identification in Unmanned Aerial Vehicle Images," *Applied Sciences*, vol. 13, no. 22, pp. 12369–12369, Nov. 2023, doi: 10.3390/app132212369.
- [29] J. Kim, and M. F. Kabir "Automated Data Labelling for Object-Detection via Iterative Instance Segmentation," in *2023 International Conference on Machine Learning and Applications (ICMLA)*, pp. 845–850, Dec. 2023, doi: 10.1109/icmla58977.2023.00124.
- [30] E. Dehaerne, B. Dey, and S. Halder, "A Comparative Study of Deep-Learning Object Detectors for Semiconductor Defect Detection," in *2021 28th IEEE International Conference on Electronics, Circuits, and Systems (ICECS)*, pp. 1–2, Oct. 2022, doi: 10.1109/icecs202256217.2022.9971022.
- [31] N. Katyayani, K. Bhardwaj, and T. Poongodi, "Deep Learning Approach for Multi-Object Detection Using Yolo Algorithm," in *2023 6th International Conference on Contemporary Computing and Informatics (IC3I)*, Gautam Buddha Nagar, India, 2023, pp. 689–693, doi: 10.1109/IC3I59117.2023.10398124.
- [32] Y. Gong, X. Yu, Y. Ding, X. Peng, J. Zhao, and Z. Han, "Effective Fusion Factor in FPN for Tiny Object Detection," *arXiv*, Jan. 2020, doi: 10.48550/arxiv.2011.02298.
- [33] H. Wang, X. Fu, Z. Yu, and Z. Zeng, "DSS-YOLO: an improved lightweight real-time fire detection model based on YOLOv8," *Scientific Reports*, vol. 15, no. 1, Mar. 2025, doi: 10.1038/s41598-025-93278-w.
- [34] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017, doi: 10.1109/tpami.2016.2577031.

BIOGRAPHIES OF AUTHORS



Yuen Chin Wong    received his bachelor of Electrical and Electronics Engineering degree specializing in Computer Engineering, from Universiti Multimedia Malaysia, in 2003. Currently, he serves as a Director at Infineon Technology, located in Melaka, where he is a part of the Product Test Engineering Department. His professional focus includes the transition data analytics and management ideas into practical production solutions. He can be contacted at email: cyberelax@gmail.com.



Hea Choon Ngo    received his bachelor's degree in computer science from the Universiti Teknikal Malaysia Melaka (UTeM), Master of Information Technology from the University of New South Wales (UNSW, Sydney) and Ph.D. in Computer Science from the Universiti Sains Malaysia. His research interests involve soft computing, data science and analytics, dynamic planning, computational modelling and intelligent systems. He is currently a senior lecturer in the Department of Intelligent Computing and Analytics, Faculty of Artificial Intelligence and Cyber Security, UTeM. He is also a member of the Center for Advanced Computing Technology, UTeM. He can be contacted at email: heachoon@utem.edu.my.



Umami Raba'ah Hashim    received her bachelor's degree in computer science from the Universiti Teknologi Malaysia in 2000, Master of Multimedia from Multimedia University in 2008, and Ph.D. in Computer Science from the Universiti Teknologi Malaysia in 2015. She is currently a senior lecturer in the Department of Software Engineering, Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka (UTeM). Her research interests include pattern recognition, machine vision, and automated defect inspection. She is also a member of the Center for Advanced Computing Technology, UTeM. She can be contacted at email: ummi@utem.edu.my.



Mohd Hilmi Hasan    received his bachelor's degree in information technology from the Universiti Teknologi Petronas (UTP), the master's degree in information technology from The Australian National University, Australia and Ph.D. in Information Technology from UTP. He currently serves as an associate professor and chair of the computing department, UTP. Driven by a strong commitment to innovation, his expertise in data science and data analytics has led to his attachment to the Centre for Research in Data Science. His work in industry-based projects involves the whole spectrum of the data science cycle and managing big data, including the study and development of appropriate system architecture and platforms. He also serves as a panel assessor for the Malaysia accreditation body and contributes to the Malaysia Board of Technologists' Seoul Accord initiative. He can be contacted at email: mhilmi_hasan@utp.edu.my.