Proceedings of the Singapore Scientific Conference – Forging a Sustainable Future Through Science & Technology

doi: 10.3850/081125_SSC2025_OR-04-0269

SSC2025

SSC2025 Category/Theme:: Manufacturing — Next-Gen Manufacturing

AOI Advisory System for PCB Inspection using Retrieval Augmented Generation

Min Sik Byun,[1] Malcom Yoke Hean Low,[1] and David Weidong Lin[2]
[1] Infocomm Cluster, Singapore Institute of Technology, Singapore
[2] Engineering Cluster, Singapore Institute of Technology, Singapore

Keywords: Vision Language Model, RAG, AOI, PCB Inspection, Defect Detection

BACKGROUND

In high-volume electronics manufacturing, effective inspection of Printed Circuit Boards (PCBs) is critical. While Automated Optical Inspection (AOI) is the industry standard, it frequently suffers from high false call rates, creating production bottlenecks [1]. Previous AI solutions using Convolutional Neural Networks (CNNs) showed promise [2] but are often impractical due to their reliance on large, specialized datasets unavailable in dynamic manufacturing environments [3].

To address this gap, our previous work introduced a novel AOI advisory system to act as a "second opinion" tool, built on a Vision Language Model (VLM) enhanced with Retrieval-Augmented Generation (RAG) [4]. In its initial evaluation on full-board images from the public MIXED PCB DEFECT DATASET [5], this system achieved 66.67% accuracy. However, a key limitation remained: the presence of vast, non-defective areas acting as background noise. This study is directly motivated by the hypothesis that model accuracy can be significantly improved by eliminating this extraneous information through a focused, cropped-image approach.

OBJECTIVES

- To introduce and evaluate methodological refinement using input images tightly cropped around specific defect locations.
- To quantify the performance improvement of this focused approach compared to the previous full-board methodology.
- To validate that this cropped-image technique provides a more direct and rigorous benchmark of a VLM's core defect recognition capabilities.

METHODOLOGY

A rigorous quantitative analysis was conducted to validate our hypothesis. Using the cropped-image dataset, defect types were programmatically extracted from the VLM's structured text responses. These predictions were then systematically compared against ground-truth labels derived from corresponding YOLO annotation files. This method enabled a direct, pair-wise evaluation of classification accuracy for both the RAG and non-RAG model configurations under identical conditions.

RESULTS AND DISCUSSION

The results demonstrate the new methodology's clear superiority. The RAG-enabled model consistently outperformed the non-RAG baseline across all levels of prompt detail. For prompts with 5-6 descriptions, the RAG model achieved an accuracy of approximately 79.5%, a substantial improvement over the non-RAG baseline's 44.35%. This outcome not only highlights the value of the RAG framework but also validates our core hypothesis that a focused, cropped-image approach significantly boosts VLM accuracy by eliminating distracting background information.

CONCLUSIONS

This research confirms that a focused, cropped-image methodology offers a more precise and reliable benchmark for VLMs in industrial inspection. The findings show that combining this targeted visual input with a RAG framework is essential for building robust and highly accurate quality control systems. This work paves the way for practical, two-stage AOI systems where initial rapid defect localization is followed by precise VLM-based analysis on cropped regions, promising more trustworthy and efficient automated inspection.

REFERENCES

[1] Wu, Xing & Ge, Yuxi & Zhang, Qingfeng & Zhang, Dali. (2021). PCB Defect Detection Using Deep Learning Methods. 873-876. 10.1109/CSCWD49262.2021.9437846.

[2] Bhattacharya, A., Cloutier, S.G. End-to-end deep learning framework for printed circuit board manufacturing defect classification. Sci Rep 12, 12559 (2022). https://doi.org/10.1038/s41598-022-16302-3.

[3] Fonseca, L.A.L.O., Iano, Y., Oliveira, G.G.d. et al. Automatic printed circuit board inspection: a comprehensible survey. Discov Artif Intell 4, 10 (2024). https://doi.org/10.1007/s44163-023-00081-5.

[4] M. S. Byun, M. Y. H. Low, D. W. Lin, C. C. Hong, Y. C. Tan and D. P. N. Ong, "Vision-Language RAG for Post-AOI PCB Inspection," in Proc. IEEE Int. Conf. on Service Operations and Logistics, and Informatics (SOLI), Singapore, 2025, to appear.

[5] V. Kumar Ancha, F. N. Sibai, V. Gonuguntla and R. Vaddi, "Utilizing YOLO Models for Real-World Scenarios: Assessing Novel Mixed Defect Detection Dataset in PCBs," in IEEE Access, vol. 12, pp.100983-100990, 2024, doi:10.1109/ACCESS.2024.3430329.