A Universal AI-Powered Segmentation Model for PCBA and Semi-Conductor

Charlie Zhu, Stephan Pirner, Srinivas Subramanian, Bahir Usanmaz, Robert Gray
Menghua Jiang, Brad Perkins, Robert Jung
Advanced Technology Solutions (ATS), Nordson Corp
Westlake, OH, USA

Abstract

The integration of artificial intelligence (AI) in automated optical inspection (AOI) systems for printed circuit board assemblies (PCBAs) and semiconductors has garnered significant attention, driven by the need for improved accuracy, speed, and consistency in defect detection. Traditionally, specialized models are trained to address specific application tasks, such as defect detection or classification. However, inspecting a complete product often requires the application of multiple models, presenting challenges in training, maintenance, and scalability.

This paper introduces a novel universal deep learning model designed to segment AOI images for both PCBA and semiconductor components, offering a more robust and adaptable solution for defect detection. Experimental results demonstrate that the proposed model performs effectively across both PCBA and semiconductor AOI tasks, highlighting its versatility and potential for streamlining inspection processes. In addition, this segmentation model can also facilitate auto-programming in a high-mix low-volume production environment.

Introduction

Automated Optical Inspection (AOI) systems are pivotal in ensuring the quality of Printed Circuit Board Assemblies (PCBAs) and semiconductor components. Traditional AOI systems primarily utilized rule-based algorithms that depended on handcrafted features for defect detection. While these methods were effective for straightforward tasks, they often struggled with the increasing complexity and variability inherent in modern electronic components. This limitation has driven the adoption of machine learning (ML) and deep learning (DL) techniques, which offer enhanced adaptability and accuracy in defect detection. For instance, convolutional neural networks (CNNs) have been employed to identify defects in PCBAs, demonstrating superior performance over traditional methods[1]

Task-Specific Segmentation Models

In the realm of defect detection, task-specific segmentation models have been extensively researched. The U-Net architecture, for example, has been widely adopted for biomedical image segmentation due to its encoder-decoder structure, which effectively captures spatial features. However, its application in industrial inspection tasks, such as PCBAs and semiconductors, has been limited. Other studies have proposed CNN-based models tailored for specific inspection tasks, achieving high accuracy but lacking the flexibility to generalize across different domains[1].

Challenges in High-Mix, Low-Volume Production Environments

The shift towards high-mix, low-volume production presents significant challenges for AOI systems. In such environments, the diversity of products necessitates frequent reconfiguration of inspection models, leading to increased maintenance efforts and reduced scalability. Traditional task-specific models are often inadequate, as they require extensive retraining to adapt to new product variations, resulting in inefficiencies and higher operational costs. This underscores the need for more adaptable inspection models capable of handling diverse product types without extensive reconfiguration.

Universal Segmentation Models

The concept of universal segmentation models has gained traction, aiming to develop models capable of generalizing across various tasks and domains. For instance, the UniverSeg model has been proposed for medical image segmentation, demonstrating the ability to handle unseen segmentation tasks without additional training[1]. Similarly, the UniSeg model employs a prompt-driven approach for multi-task medical image segmentation, outperforming other universal models on multiple tasks[1]. These advancements highlight the potential of universal models to streamline segmentation processes across different applications.

In summary, while significant progress has been made in developing advanced AOI systems and task-specific segmentation models, challenges persist in adapting these systems to high-mix, low-volume production environments. The emergence of universal segmentation models offers a promising avenue for creating adaptable and scalable inspection solutions applicable to both PCBAs and semiconductor components.

Proposed Universal Segmentation Models for PCBA and Semi-Conductor

The proposed universal segmentation model leverages the robust capabilities of convolutional neural networks (CNNs) to address the diverse inspection needs of both PCBAs and semiconductor components. Unlike traditional task-specific models, this universal model is designed to perform segmentation tasks across a wide range of AOI applications, eliminating the need for multiple specialized models. The architecture of the proposed model incorporates advanced feature extraction layers to effectively capture spatial patterns and variations inherent in AOI images. Additionally, it is optimized for adaptability in high-mix, low-volume production environments, enabling efficient auto-programming and reducing the complexity of inspection processes. By unifying segmentation tasks under a single model, this approach simplifies training, maintenance, and scalability, ensuring consistent and reliable defect detection across both PCBAs and semiconductors. Experimental results validate the versatility and robustness of this model, showcasing its potential to streamline inspection workflows while maintaining high accuracy and operational efficiency.

Experimental Methodology

Model Architecture

The proposed universal segmentation model is designed to process a single field of view from the AOI sensor, utilizing RGB 8-bit images as input. Its output is a segmentation mask with four distinct classes: Substrate, Component, Solder Joint, and Non-Board-Area. The architecture is built to balance accuracy and computational efficiency, making it well-suited for a variety of inspection tasks in PCBAs and semiconductor manufacturing.

To address the diverse requirements of different applications, multiple versions of the model are available. For tasks like auto-programming, where precision is critical, a larger model version is used to ensure high segmentation accuracy. On the other hand, for applications which is more time-critical, such as inspections, a smaller version prioritizing inference speed is deployed. This adaptability ensures the model can cater to a range of use cases, providing both flexibility and efficiency in high-mix, low-volume production environments.

Test Setup

A 3D AOI system equipped with a 7 µm 3D sensor, was utilized to acquire high-resolution 2D and 3D images for all experimental samples. Uniformly illuminating white light was used to capture the 2D images This setup ensured the acquisition of detailed and consistent imaging data, critical for accurate training and evaluation of the segmentation model. Three distinct datasets were prepared for the experiments, each containing annotated images where component bodies and pins were highlighted for segmentation tasks:

1. *Training Set:* Images from Nordson R&D PCBs, as in Figure 1, were used to train the model. The images were annotated/labeled manually. Three types of features are being annotated in each image, as in Figure 2, i.e. substrate as green, component body as red, and lead including solder Joint as purple. This set included a diverse range of components and assembly configurations to ensure robust feature learning.
2. *Validation Set:* Similar to the training set, the validation set comprised images from Nordson R&D PCBs. It was used during training to monitor the model’s performance and prevent overfitting.
3. *Test Set:* The test set contained images from both Nordson R&D PCBs and iNEMI test vehicles[1], as shown in Figure 2, which are intentionally excluded from the training and validation sets. The specially designed iNEMI test vehicles consist of a standardized set of samples widely used in the industry for benchmarking AOI systems. This dataset allowed for an unbiased evaluation of the model's generalization capabilities and adaptability to unseen board configurations.

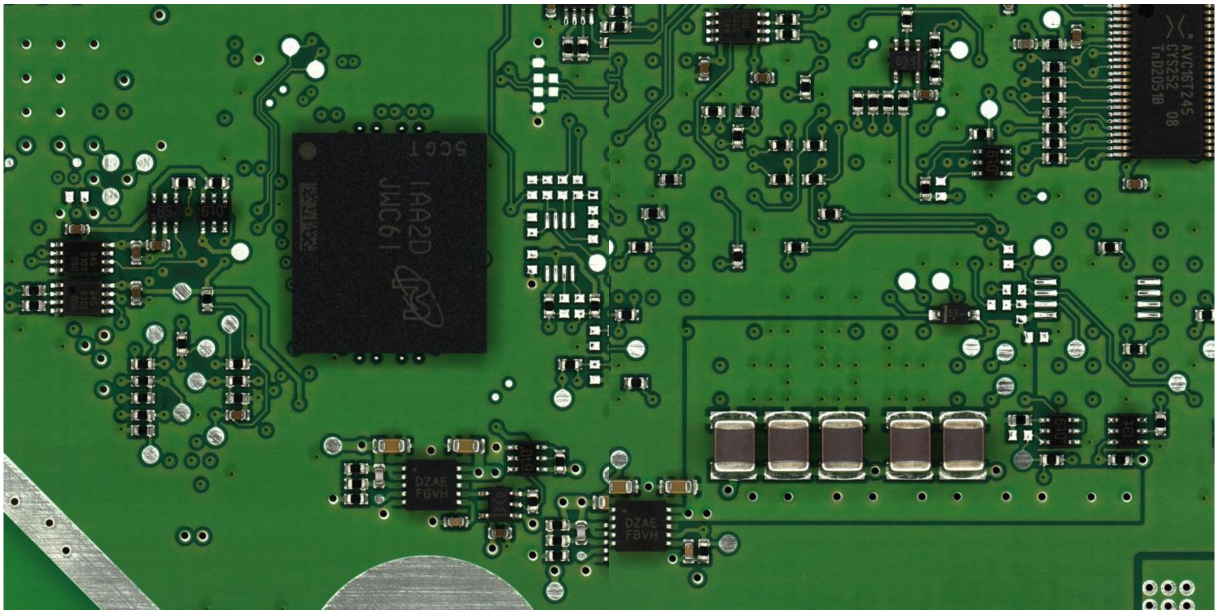

Figure 1. Sample image from Nordson R&D PCB

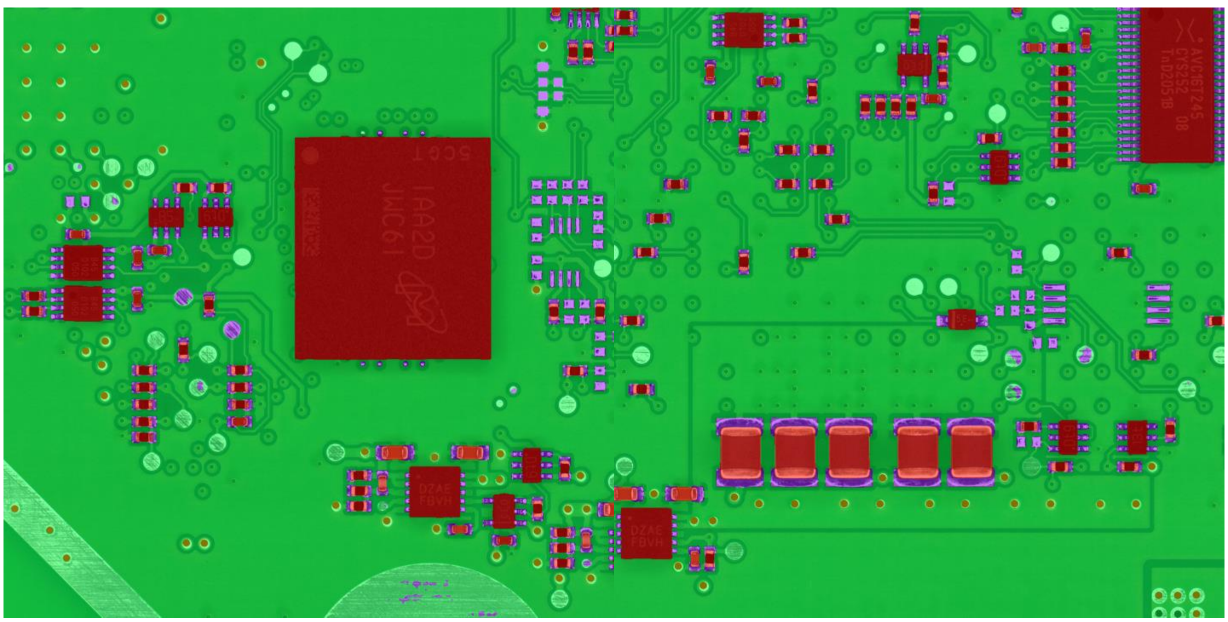

Figure 2. Labeled data

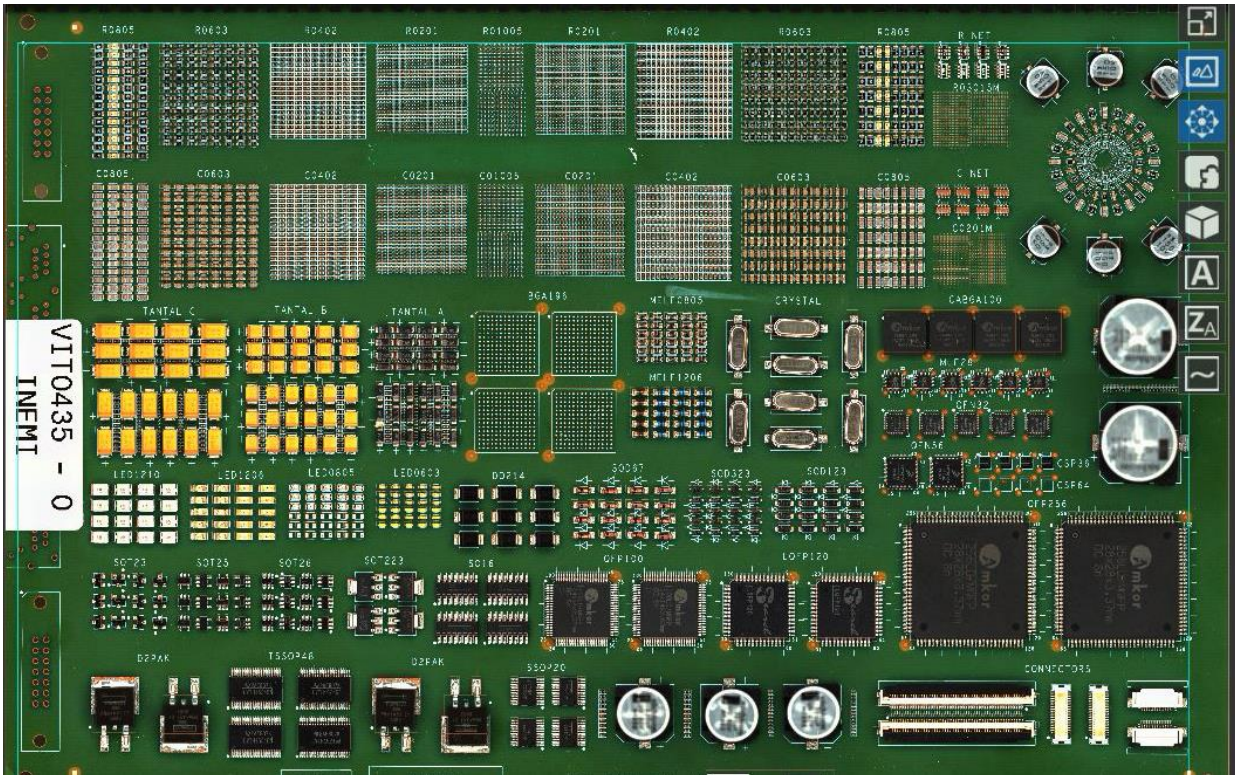

Figure 3. Image of iNEMI AOI test vehicle

Training

The model was trained on the training set using supervised learning with annotated data. Augmentation techniques such as rotation, scaling, and brightness adjustment were applied to increase the diversity of the training data. The model's performance was validated during training using the validation set, with early stopping criteria applied based on validation loss.

Evaluation

The model's performance was rigorously evaluated using the test set, consisting exclusively of images from iNEMI test vehicles. The primary metric for accuracy was the percentage of pixels in the predicted segmentation masks that matched the ground truth annotations. This pixel-level accuracy provided a detailed assessment of the model’s capability to correctly classify each region in the AOI images. Additionally, the evaluation focused on the model’s ability to generalize across unseen board configurations and its robustness in identifying the four target classes: Substrate, Component body, and leads including solder Joint.

Results

The evaluation of the proposed universal segmentation model yielded strong results, demonstrating its robustness and adaptability across different datasets. A critical challenge during training was managing the inherent inaccuracies in human annotations, as achieving pixel-perfect precision in ground truth masks is impractical. To address this, the training process emphasized strategies to avoid overfitting, ensuring the model could generalize effectively to new data while maintaining high accuracy. Visual inspection of the segmentation results revealed that the model consistently identified the four target classes—Substrate, Component, Solder Joint, and Non-Board-Area—with remarkable precision.

Performance on Validation Set

Quantitative evaluation further validated the model’s performance. On the validation set, which consisted exclusively of Nordson R&D PCB samples, the model achieved an impressive accuracy of approximately 96%. This high accuracy highlights the model’s ability to adapt to the intricate patterns and features present in these training-like datasets without over-fitting. Moreover, the consistent performance on the validation set suggests that the model has successfully learned to distinguish subtle details while avoiding the pitfalls of overfitting.

Performance on Test Set

Table 1 shows an overview of the results of the test dataset. The test set provided a more rigorous assessment of the model’s generalization capabilities, as it included images from both Nordson R&D PCBs and iNEMI test vehicles, which are industry-standard samples. On this mixed dataset, the model achieved a solid accuracy of about 90%. While slightly lower than the validation set, this result demonstrates the model's ability to handle unseen data and adapt to variations in component designs and assembly configurations, as shown in Figure 3. These findings underscore the model’s potential to deliver reliable and accurate segmentation results across diverse AOI tasks in PCBA and semiconductor manufacturing.

Table 1. Test Results Overview

	Validation Accuracy	Testing Accuracy
Nordson R&D PCB	96%	93%
iNEMI Test Vechicle	-	86%
Overall	96%	90%

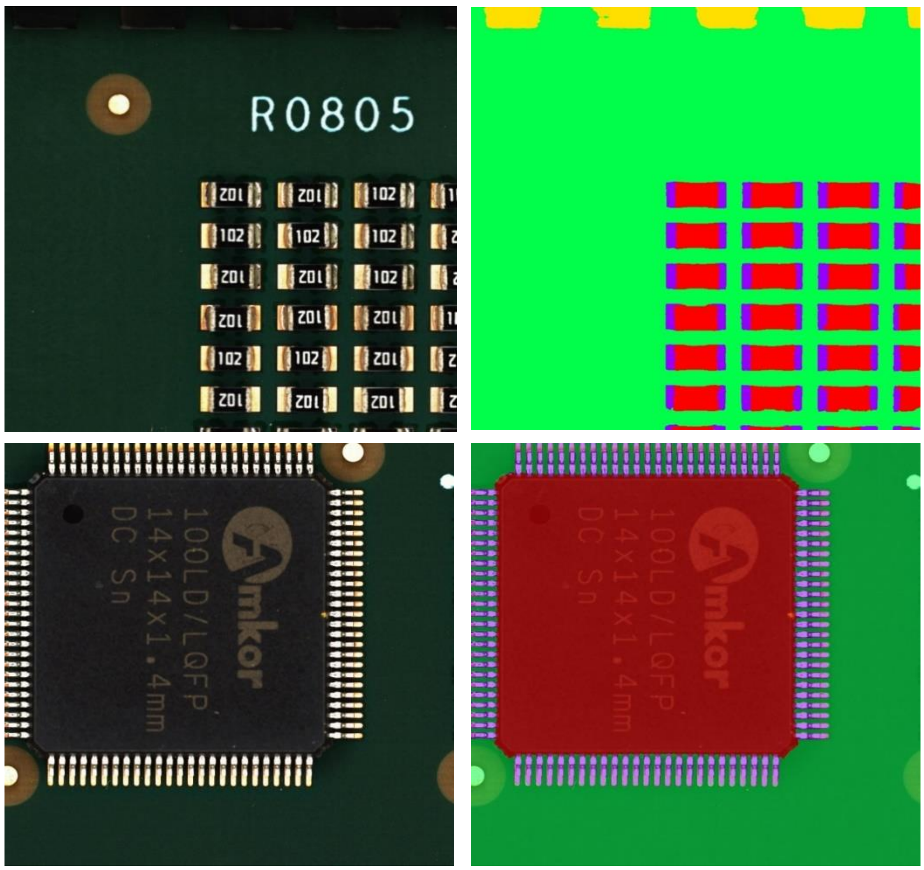

Figure 3. Output Mask Images on the iNEMI Test Vehicle

Discussion

Semi-Conductor Applications

The proposed universal AI-powered segmentation model has demonstrated significant efficacy in segmenting components within printed circuit board assemblies (PCBAs) images. Building upon this success, extending the model's application to wafer-level tasks, such as segmenting wafer bumps from the substrate, presents a logical progression. Accurate segmentation of wafer bumps is crucial for quality control in semiconductor manufacturing, as it directly impacts the reliability of subsequent packaging processes. Recent studies have explored deep learning approaches for wafer defect detection[1], indicating the potential for such models in wafer applications. Integrating wafer bump segmentation into the current model could enhance its versatility, providing a unified solution across multiple stages of electronic component manufacturing.

Transformer Models

Incorporating transformer-based architectures into the segmentation model offers another avenue for improvement. Transformers have shown remarkable success in capturing long-range dependencies in image data, which is beneficial for complex segmentation tasks[1]. However, their computational complexity poses challenges for real-time applications, especially in manufacturing environments where swift processing is essential. To address this, hybrid models that combine the strengths of transformers and convolutional neural networks (CNNs) have been proposed. For instance, a hybrid convolutional-transformer architecture has been developed for real-time image segmentation, achieving a balance between accuracy and efficiency [1]. Adopting such hybrid architectures could enable the deployment of transformer-based models in scenarios with stringent real-time requirements, thereby enhancing the model's performance without compromising operational speed.

In summary, expanding the model's capabilities to include wafer bump segmentation and integrating transformer-based architectures can significantly enhance its applicability and performance. These advancements would provide a more comprehensive and efficient tool for defect detection and quality control in the electronics manufacturing industry.

Conclusions

The results of this study demonstrate that a generalized artificial intelligence (GAI) model is both feasible and effective for inspecting images from any Automated Optical Inspection (AOI) system, regardless of the specific product being analyzed. By achieving high levels of segmentation accuracy across diverse datasets, the proposed model highlights the potential to unify the inspection process, making it more efficient and scalable. This capability represents a significant advancement over traditional task-specific models, which often require extensive retraining and fine-tuning for each unique application.

The ability to deploy a generic model across various AOI systems introduces substantial operational benefits. It reduces the time and resources required to train new models for different products, thereby streamlining the quality control workflow. This not only accelerates the inspection process but also ensures consistent and reliable performance across a wide range of manufacturing scenarios. The flexibility and adaptability of the proposed model provide a clear path toward improving the scalability of AOI systems in modern production environments.

In addition to its role in defect inspection, the segmented results generated by the model open up new possibilities for process optimization. For inspection tasks, the segmentation masks enable the application of statistical techniques to identify defects such as missing or incorrect components and insufficient or excessive solder. In addition, the mask image can be combine with other data such as 3D data to detect 3D defects such as tombstone and measure coplanarity. These insights can enhance the accuracy and reliability of quality control, ensuring that only defect-free products progress through the manufacturing pipeline. Furthermore, the model’s outputs are ideally suited for auto-programming tasks. By using masked images as inputs for recipe programming, the need for CAD files and manual tuning is eliminated. This not only saves significant programming time but also reduces the reliance on specialized process knowledge, making the workflow more accessible and efficient.

Overall, the proposed universal AI-powered segmentation model demonstrates the potential to revolutionize AOI systems by delivering both technical and operational advantages. By combining high accuracy, scalability, and versatile applications, this model offers a robust solution for addressing the challenges of modern manufacturing and quality control.

Acknowledgements

We would like to extend our gratitude to the teams at Nordson ATS for their invaluable support in this project, including the provision of essential resources and equipment. We also sincerely appreciate the contributions of the iNEMI Investigation of AI Enhancement to AOI for PCBA Project team for supplying the test vehicles used in our study.

References

[1] A. Legon, et al., (2022 May 23-26). Detection and Classification of PCB Defects using Deep Learning Methods. *2022 IEEE 31st Microelectronics Design & Test Symposium (MDTS)*, Albany, NY, USA.

[2] Y. Ye, et al., (2023 Oct 8-12). UniSeg: A Prompt-driven Universal Segmentation Model as well as A Strong Representation Learner. *26th International Conference on Medical Computing and Computer Assisted Intervention)*, Vancouver, Canada.

[3] W. Zhang, et al., (2024 Oct 17-18). Enhancing AOI with AI for Printed Circuit Board Assembly– A Comprehensive Industry Study. *40th International Electronics Manufacturing Technology (IEMT) Conference*, Penang, Malaysia.

[4] TY. Kim, et al., (2024 Feb). Deep Learning-Based Detection of Defects in Wafer Buffer Zone During Semiconductor Packaging Process. *Multiscale Sci. Eng.,* Vol. 6, P.25-32.

[5] X. Li, et al., (2024 Dec). Transformer-Based Visual Segmentation: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol. 46, P. 1586-1601.

[6] J. Fan, et al., (2024 Feb). SegTransConv: Transformer and CNN Hybrid Method for Real-Time Semantic Segmentation of Autonomous Vehicles. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 25, No. 2, P.10138-10163.