

A Generative AI Approach for Image Augmentation and Applications in Industrial Imaging

Munyoung Lee

*Regional ICT Convergence Research Section
Electronics and Telecommunications Research Institute
Daejeon, Republic of Korea
munyounglee@etri.re.kr*

Eun-Hee Kim

*Regional ICT Convergence Research Section
Electronics and Telecommunications Research Institute
Daejeon, Republic of Korea
eunheekim@etri.re.kr*

Hyunwoo Kim

*Regional ICT Convergence Research Section
Electronics and Telecommunications Research Institute
Daejeon, Republic of Korea
kim.hw@etri.re.kr*

Kyu-Sung Lee

*Regional ICT Convergence Research Section
Electronics and Telecommunications Research Institute
Daejeon, Republic of Korea
kyusung.lee@etri.re.kr*

Abstract—The adoption of artificial intelligence (AI) across various industries is rapidly expanding. However, data collection from manufacturing processes is often limited or costly, resulting in insufficient datasets for AI applications. To address this issue, data augmentation methods have been proposed. Recent advances in generative AI, including generative adversarial networks (GANs) and diffusion models, have emerged as promising alternatives. In this study, we investigate the feasibility of generative AI to expand image datasets. Specifically, we demonstrate the application of OpenAI's image generation model to synthesize additional optical microscope images obtained during intermediate stages of semiconductor manufacturing. A comparative analysis and validation of the generated data were conducted to demonstrate the applicability of generative AI-based data augmentation. Our results indicate that generative AI provides an effective and practical approach for supplementing datasets in data-scarce environments.

Index Terms—Generative AI, Data Augmentation, Image Augmentation, openAI API.

I. INTRODUCTION

Recently, the rapid advancement of artificial intelligence (AI) technologies has led to their widespread adoption across various application areas, including smart city management, factory automation, and personalized healthcare services. To build an AI model for specific tasks, it is essential to acquire large volumes of data while preserving its relativeness, quality, and accuracy. However, in some application domains, the amount of available data is limited due to various constraints such as privacy concerns, data acquisition costs, and industrial confidentiality issues.

In such data-scarce environments, developing scalable and high-performing AI models becomes challenging. To overcome these challenges, recent research has increasingly fo-

cused on the use of generative AI technologies to create synthetic data that closely resembles real-world data. As a result, the synthetic data can be integrated with real datasets to improve the performance and reliability of AI models.

For example, in AI services designed for predictive maintenance and anomaly detection in industrial equipment, it is crucial to collect data that implies various early signs of potential failures [1], [2]. Traditionally, experienced operators have recognized subtle indicators, such as abnormal noise or vibrations, to predict equipment malfunctions and proactively identify the underlying causes. However, such anomalies occur infrequently and are often difficult to reproduce systematically. Furthermore, allowing real failures to occur for the sake of data collection is impractical, as it may lead to equipment damage and significant financial loss. In these cases, generative AI can be effectively employed to create realistic abnormal data based on real warning signals, supporting the development of prototype AI systems for predictive maintenance.

Several studies on data augmentation using generative models, such as generative adversarial networks (GANs) and diffusion models, have been actively conducted [3], [4]. Nevertheless, challenges remain, particularly with respect to the fidelity and relevance of the generated data. Inaccuracies or excessive distortions introduced during the generation process may compromise the reliability of the augmented dataset. Therefore, domain-specific expertise is crucial for validating the quality and utility of synthetic data.

In this study, we apply generative AI techniques to augment image data using openAI's API [5], and conduct a comparative analysis to evaluate the validity and applicability of the generated datasets. For this analysis, the original dataset is obtained from optical microscope images during semiconductor manufacturing processes.

This work was supported by the Technology Innovation Program (Development and Practice of an On-device AI Functionality and Performance Testing Framework based on NPU, RS-2025-02307650, 50%) funded by the Ministry of Trade, Industry & Energy (MOTIE, Korea) and ETRI grant funded by the Korea government (25ZT1100, 25YT1100, 50%).

TABLE I
IMAGE AUGMENTATION PROMPTS AND DESCRIPTIONS.

Prompt	Description
Prompt 1: tiny variation	Apply very small, imperceptible random perturbations to RGB pixel values while keeping sharpness, structure, and details identical to the original.
Prompt 2: slight variation	Apply imperceptibly small texture adjustments and near-invisible brightness and color variations, preserving the original tone, sharpness, structure, and details without noticeable change.
Prompt 3: moderate variation	Apply subtle color variations for data augmentation, slightly adjust brightness and hue while keeping overall color tone and structure intact, and preserve sharpness and details.

II. METHODOLOGY

A. OpenAI API for Image Augmentation

The OpenAI API [5] is a cloud-based platform that allows developers to integrate state-of-the-art artificial intelligence capabilities into their products and services. Through this API, users can access a wide range of AI models, including those for image generation, natural language processing, code generation, and speech recognition. For instance, image generation models can transform existing photographs into a different style, such as Studio Ghibli, without requiring extensive computational resources. Leveraging this capability, generative AI can support image augmentation in data-scarce environments. In this study, we demonstrate that OpenAI's gpt-image-1 model can augment optical microscope images obtained from semiconductor manufacturing processes, providing a practical approach to address data scarcity while maintaining high-quality and realistic image representations.

B. Original Image Acquisition

To generate optical microscope image data for this study, patterned wafer samples were fabricated using a standard wet oxidation process on 6-inch silicon (Si) substrates. This process deposited a silicon dioxide (SiO_2) layer with a thickness of 5000 Å. The wafers were subsequently etched to controlled depths to form representative micro-patterns of 500 Å, 3000 Å, 4000 Å, and 4500 Å. Based on the well-known interference characteristics of SiO_2 , where the perceived color originates from thin-film interference and varies with thickness [6]–[8], the etching depth was adjusted to produce clear color differences observable under optical microscopy. Leveraging this property enabled the acquisition of diverse color variation data across different thicknesses, ensuring that the collected images contained visually meaningful information suitable for machine learning-based analysis.

Optical measurements were conducted with an Olympus BX-53M optical microscope. The captured images were stored in sRGB format at a resolution of 3072×2048 pixels with 24-bit color depth. For consistency and relevance to high-resolution inspection tasks, all samples were captured at 10× magnification.

C. Image Augmentation via Generative AI

Using OpenAI's image generation model, gpt-image-1, three prompt levels (Table ??) are applied to the original

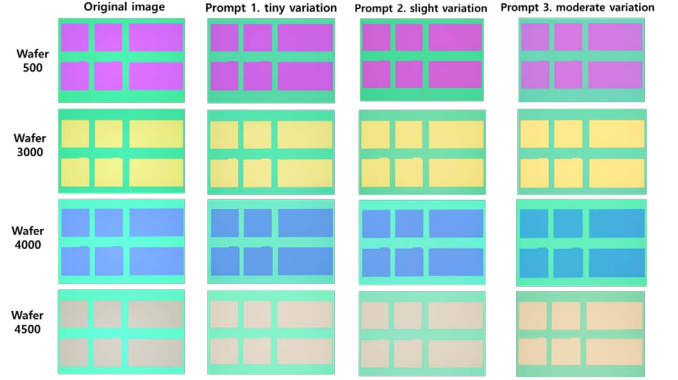


Fig. 1. Images augmented from optical microscope semiconductor wafer data using the three prompt levels.

images, progressively increasing variation from Prompt 1 to Prompt 3. This approach allows the model to generate multiple augmented images from a single input, adhering to the prompt instructions and preserving the essential structure and content.

Figure 1 illustrates images generated from the original samples of four semiconductor wafers. Although the differences are not readily perceptible to the human eye, it can be observed that the degree of variation relative to the original images increases progressively from Prompt 1 to Prompt 3. These results demonstrate that generative AI can effectively produce a series of augmented images that introduce controlled variability while maintaining the core structural and visual characteristics of the original samples, thereby providing a practical approach for image augmentation in data-scarce environments.

III. ANALYSIS OF AUGMENTED IMAGES

Figure 2 shows the histograms of the R, G, and B channels for the Wafer 4500-B image, which corresponds to the bottom region of a wafer etched to 4500 Å, comparing the original image with the three augmented prompts (Prompt 1-3).

In the R channel (left), the original image exhibits two peaks around 100 and 220. For Prompt 1 (tiny variation), the first peak shifts slightly to the right, while the second peak is slightly reduced. Prompt 2 (slight variation) shows a more pronounced rightward shift and a narrowing of the second peak. Prompt 3 (moderate variation) exhibits the largest shift, with the second peak significantly reduced. This indicates that

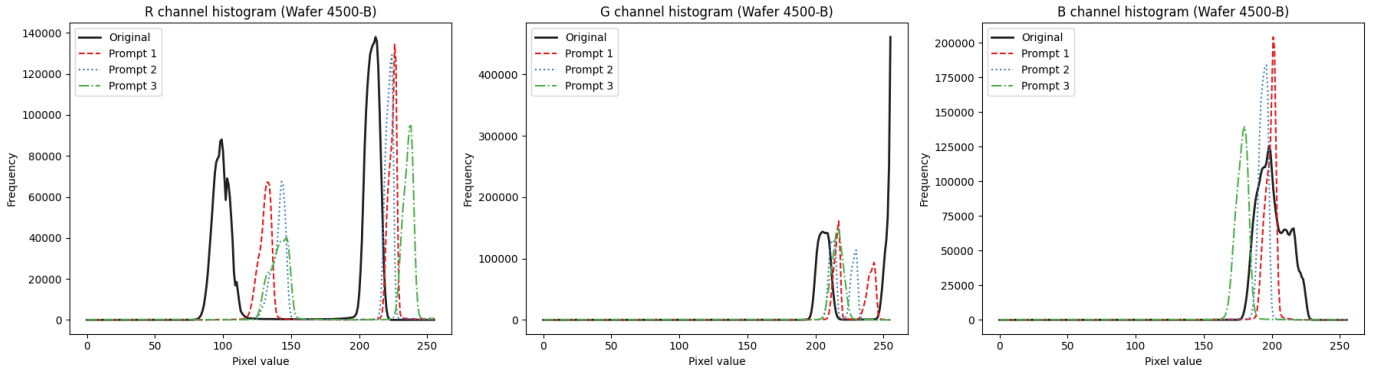


Fig. 2. Histograms of the R, G, and B channels for the wafer 4500-B image.

pixel values in the R channel gradually increase with higher prompt levels.

In the G channel (center), the original image has a major peak near 250 and a smaller peak around 210. Prompt 1 shows a peak around 230–245, indicating a rightward shift. Prompt 2 shifts slightly to the right with a broader spread, and Prompt 3 also shifts rightward, peaking near 220 with the widest distribution.

In the B channel (right), the original image has a peak near 200. Prompt 1 shifts this peak to around 210, Prompt 2 remains nearly aligned with the original or slightly left with modest broadening, and Prompt 3 shifts further leftward toward 180–190 with the widest spread.

Overall, higher prompt levels lead to increased values in the R channel, while the G and B channels exhibit variable shifts, generally tending toward lower values and occasionally showing greater dispersion. Although these differences are subtle to the human eye, the histograms provide quantitative evidence of the controlled modifications introduced by the augmentation process.

Figure 3 illustrates the average RGB color differences between the original images and the generated images using the three prompts for all wafer samples under identical conditions. The analysis shows that Prompt 1 exhibits the lowest color differences across most wafers, with a minimum observed for the Wafer 3000 sample, demonstrating the highest consistency with the original color characteristics. Prompt 2 shows intermediate color differences overall, with a relatively lower difference for the Wafer 4000 sample, indicating improved color fidelity under specific conditions. In contrast, Prompt 3 displays the highest color differences across all wafers, with a sharp increase above 50 for the Wafer 4000 and 4500 samples, indicating pronounced color distortions. These results suggest that Prompt 1 is most suitable for maintaining color fidelity relative to the original images, whereas Prompt 3, due to its greater color variability, is better suited as a supplementary approach for enhancing data diversity rather than preserving the original colors.

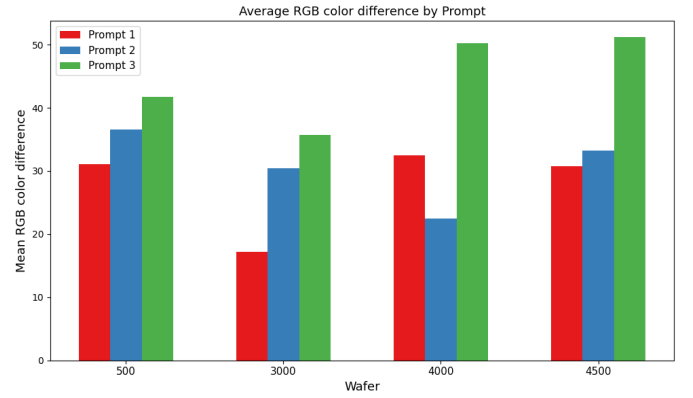


Fig. 3. Average RGB Color Differences by Prompt Level.

IV. APPLICATION SCENARIOS AND DISCUSSION

Optical microscope image data represents a critical data modality, which has been widely adopted in semiconductor manufacturing and inspection processes. The integration of advanced artificial intelligence (AI) techniques has recently driven significant innovation in these domains. Optical microscope images of semiconductor wafers are utilized for Fault Detection and Classification (FDC) to enhance process yield and enable early identification of device defects. Moreover, these images serve as essential inputs for evaluating thin-film performance metrics such as thickness, uniformity, deposition and etching rates, and pattern fidelity during various manufacturing stages.

Despite their importance, acquiring sufficient abnormal or defect data remains challenging, often resulting in highly imbalanced datasets dominated by normal samples. This imbalance complicates the training of robust AI models for accurate fault detection, and variability in surface uniformity and other film properties produces diverse anomaly patterns that are difficult to capture comprehensively.

Building on the methodology and analysis presented in Sections II and III, the generative AI techniques were applied to produce three levels of synthetic optical microscope images that closely resemble real measurements. For each prompt

level, multiple iterations of image generation were performed to assess the characteristics, variability, and practical applicability of the augmented datasets. The controlled variations in the R, G, and B channels, together with the color fidelity trends shown in Figures 2 and 3, demonstrate that these synthetic images can effectively supplement real datasets. This capability is particularly valuable in data-scarce scenarios, enabling more robust AI model training for Fault Detection and Classification (FDC) and other semiconductor inspection tasks, and illustrating the practical application potential of generative AI in industrial imaging workflows.

Despite the promising results, one challenge in using generative AI for image augmentation is maintaining consistent fine-tuning across different samples. Future work will focus on addressing this limitation and systematically evaluating how different combinations of generated and real images influence the performance of AI models.

V. CONCLUSION

We investigated a generative AI-based image augmentation approach for optical microscope images of semiconductor patterned wafers using OpenAI's gpt-image-1 model. Three controlled prompt levels were applied to generate augmented datasets, enabling systematic evaluation of their effects. Analysis of pixel distributions and average RGB color differences shows that Prompt 1 preserves color fidelity, while Prompt 3 introduces greater variability, supporting enhanced dataset diversity.

These results demonstrate that generative AI can produce high-quality augmented images that retain essential structural and visual features while providing controlled variability. This approach offers a practical solution for data-scarce environments and supports the development of AI models, contributing to accelerated AI transformation (AIX) across diverse industrial applications.

REFERENCES

- [1] M. Lee, S. Yoon, and S. Jeon, "Anomaly detection based status diagnosis system of machine tools," in *2022 KICS Conference Fall*, pp. 522–523, KICS, 2022.
- [2] G. Pang, C. Shen, L. Cao, and A. van den Hengel, "Deep learning for anomaly detection: A review," *ACM Computing Surveys (CSUR)*, vol. 54, no. 2, pp. 38:1–38:38, 2021.
- [3] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *J. Big Data*, vol. 6, p. 60, 2019.
- [4] S. Liu, J. Chen, Y. Feng, Z. Xie, T. Pan, and J. Xie, "Generative artificial intelligence and data augmentation for prognostic and health management: Taxonomy, progress, and prospects," *Expert Systems with Applications*, vol. 255, p. 124511, 2024.
- [5] OpenAI, "Openai api," 2023.
- [6] J. Henrie, S. S. Kellis, S. M. Schultz, and A. Hawkins, "Electronic color charts for dielectric films on silicon," *Optics Express*, vol. 12, no. 7, pp. 1464–1469, 2004.
- [7] E. Hecht, *Optics*. Pearson, 5 ed., 2017. Covers Fresnel coefficients, interference, and thin-film optics.
- [8] H. A. Macleod, *Thin-Film Optical Filters*. CRC Press, 4 ed., 2010. Standard text on thin-film interference and reflectance modeling.