CDSEM AUTOMATED INSPECTION AND DEFECTS CLASSIFICATION USING A MULTICLASS DEEP LEARNING APPROACH

Jerald M. Constantino
James C. Gonzales

Operations Excellence
Western Digital, 109 Technology Ave., SEPZ, Laguna Technopark, Biñan, Laguna, Philippines 4024
jerald.constantino@wdc.com, james.gonzales@wdc.com

ABSTRACT

Incorporating machine learning techniques, particularly convolutional neural networks (CNNs), has shown remarkable potential in transforming manufacturing inspection processes. This project presents a solution for enhancing inspection efficiency in the manufacturing industry by developing an automated inspection system tailored for Critical Dimension Scanning Electron Microscopy (CDSEM) images retrieved from a Big Data Platform (BDP) server.

The proposed system utilizes custom multiclass image classifier models constructed on CNN architectures to classify CDSEM images, thereby streamlining the inspection process accurately. Additionally, a dashboard is integrated into the system, offering summary and detailed inspection results. This dashboard also provides valuable insights for the quality assurance (QA) team, facilitating informed decision-making and enabling timely interventions when necessary.

By merging state-of-the-art machine learning techniques with intuitive visualization tools, this automated inspection system marks a substantial advancement in manufacturing quality control. It leads to workforce optimization by 50 percent and pledges to enhance the manufacturing industry's efficiency, quality, and reliability of inspection processes.

1.0 INTRODUCTION

In Western Digital Philippine Head Office (PHO), Critical Dimension Scanning Electron Microscopy (CDSEM) plays a crucial role in ensuring the quality and integrity of incoming semiconductor wafers. CDSEM images from the Send Ahead (SA) monitoring process samples are used to determine whether a wafer can proceed to the Main Build (MB) input stage.

Traditionally, the CDSEM inspection process involves capturing CDSEM images of SA samples, followed by a manual inspection by the Supplier Quality Engineering (SQE) team. They evaluate the images based on established criteria via a monitor. Subsequently, the SQE team utilizes their inspection judgment results of the samples representing the wafer. Pass wafers are released into the system, while wafers with samples requiring more attention undergo further risk assessment to mitigate potential downstream impacts. See Fig. 1 for the traditional CDSEM Manual Inspection Flow.

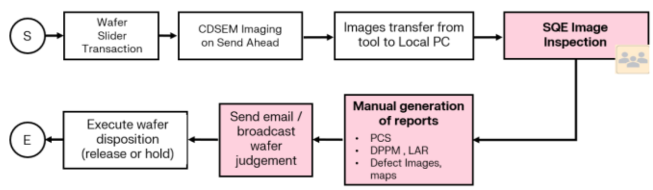


Fig. 1. Traditional CDSEM Manual Inspection Flow. Manually generated reports contain the following: Process Control Sheet (PCS), Defects Parts Per Million (DPPM), Lot Acceptance Rate (LAR), Defect Images, and Wafer Map.

While the traditional CDSEM inspection process serves its purpose, it also presents several challenges. Firstly, it demands round-the-clock support from operators and technicians, increasing operational costs and potential human errors. Secondly, the manual gathering of inspection results and generation of multiple reports add to the administrative burden, hindering efficiency and timely decision-making. In response to these challenges, there is a pressing need for innovative solutions that harness the power of advanced technologies to streamline the CDSEM inspection process, improve accuracy, and reduce dependency on manual labor.

This project presents a novel approach to addressing these challenges by proposing the development of an automated inspection system leveraging machine learning techniques, specifically convolutional neural networks (CNNs), to classify CDSEM images. Additionally, the system incorporates a dashboard to provide comprehensive insights into inspection results, empowering the SQE team with actionable information for effective decision-making. By integrating cutting-edge technology and intuitive visualization tools, this automated inspection system can revolutionize CDSEM inspection processes, enhancing efficiency, accuracy, and quality control in semiconductor manufacturing.

2.0 REVIEW OF RELATED WORK

The utilization and enhancement of CDSEM has recently gained significant attention in the semiconductor industry and research field. For instance, several studies have focused on improving the cycle time of CDSEM measurement by proposing a new wafer alignment methodology and data storage optimization strategies[1]. Another study explored the opportunities of employing a generative adversarial network to improve low-quality images[2]. Furthermore, previous research had extended the usage of CDSEM tools by implementing on-device, target-free overlay measurements[3]. It involves extracting sub-pixel contours from CDSEM images and using design data to calculate overlay differences. Meanwhile, further study has developed an unsupervised machine learning model for process window monitoring[4]. The algorithm extracts critical features from a dataset of CDSEM images, encodes them into latent feature vectors, and ranks images based on similarity indices. Lastly, another study employed a deep learning model to enhance the 3D profiling accuracy of high aspect ratio features using high-voltage CDSEM[5]. These advancements pave the way for further exploration of CDSEM's potential, particularly in real-time process monitoring and defect detection during production.

While deep learning finds applications in various fields such as defect detection[6], predictive maintenance[7], financial forecasting[8], and recommendation systems[9], this project focuses on utilizing deep learning for multiclass image classification of defects in CDSEM-generated images.

This project utilized a deep learning technique called multiclass defect classification based on the Visual Geometry Group (VGG) architecture. The VGG network has gained popularity in recent years for image classification due to its balance of performance and simplicity. Its design prioritizes high accuracy by using deep convolutional layers with small kernels. It effectively captures complex image features, leading to impressive results, especially for large datasets[10].

3.0 METHODOLOGY

Although the traditional CDSEM inspection process fulfills its purpose, it also presents several challenges. Manually transferring and storing images is time-consuming, and the potential to make mistakes may lead to scattered data and difficulty tracking trends. Furthermore, manually classifying defects is slow and subjective, hindering consistent analysis. Static reports offer limited insights, making proactive decision-making challenging.

3.1 CDSEM AI Process Flow

This project addresses the pain points directly (refer to Fig. 2 for the project's flow). By leveraging a BDP server, CDSEM images are consolidated and securely archived. Subsequently, an automated CDSEM (Automated Inspection) AI model classifies defects within the images, saving valuable time and ensuring consistent results. These classifications are then translated into interactive visualizations on a user-friendly dashboard, providing the most up-to-date production insights. An auto-email alert system also notifies relevant personnel of critical findings, facilitating faster response times. This project streamlines the CDSEM inspection process by automating tedious tasks and offering real-time data visibility. It leads to increased efficiency, improved accuracy in defect classification, enhanced production visibility, quicker problem-solving through timely alerts, and data-driven decision-making for optimized production quality and yield.

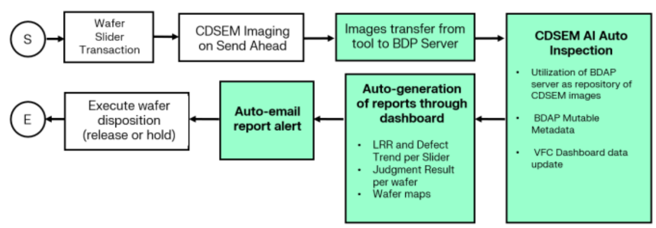


Fig. 2. The updated project flow with automation on CDSEM inspection and report generation through dashboards and email reports.

3.2 Data Architecture

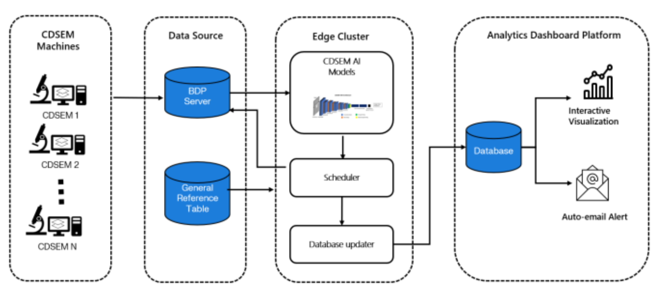


Fig. 3. High-level data architecture of the project.

Illustrated in Fig. 3 is the high-level data architecture of the project. The BDP Server acts as a repository for CDSEM images of wafer samples. Complementing this is the General Reference Table, a data source containing details about class information and standard reference data for the products.

The deployed model is containerized using Docker within an edge cluster. It ingests the images for classification. The scheduler module processes the classification results in metadata and uploads this information to the CDSEM image. The Database Update module updates the dashboard's database by processing information from the classification results and the general reference table.

The Analytics Dashboard Platform houses a database updated by a dedicated module. The platform boasts an interactive dashboard and an auto-email alert module, rounding out the functionality.

3.3 CDSEM Imaging Slider Images

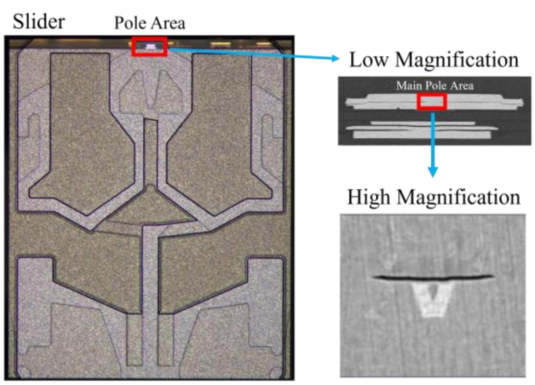


Fig. 4. Sample CDSEM images for low and high magnification inspection process.

Before the wafers proceed to their subsequent processing destination, the wafer slider samples undergo CDSEM imaging. This process involves two magnification modes: high magnification (highmag) and low magnification (lowmag). Each magnification has its own set of classes. Fig. 4 shows a sample of CDSEM images for highmag and lowmag.

Image samples manually classified for highmag and lowmag processes are used to train and develop this project's two multiclass image classification models.

3.4 CDSEM AI CNN Architecture

The CDSEM CNN architecture is illustrated in Fig. 5. It is based on the VGG network for its simplicity and effectiveness. The network takes a constant input image size of 512 x512 grayscale pixels.

Convolutional layers, followed by max-pooling layers, are used throughout the network. These layers extract features from the images and reduce their spatial dimensions. Each convolutional layer uses a ReLU activation function to introduce non-linearity. After several convolutional and max-pooling layers, global average pooling is applied to reduce spatial dimensions further and extract global features.

The output from the global average pooling layer then travels through fully connected dense layers with ReLU activation for classification. A dropout layer with a rate of "D" is included to prevent the model from overfitting during training. The final layer has a different number of units depending on the model type: "H" units for the high-magnification model (corresponding to the number of class types) and "L" units for the low-magnification model.

Finally, a sigmoid activation function is applied to the final layer for multiclass classification tasks. Sigmoid is chosen because the model is expected to classify an image that belongs to multiple classes independent of each other. Each output neuron in the final layer corresponds to a binary decision ("1" for presence and "0" for absence) for a specific class. Sigmoid activation is preferred over the common multiclass activation SoftMax function for this case because understanding the probability of each class is considered more important than having a strict probability distribution across all classes. The CNN's output, a list of probabilities for each class, is then further processed by applying a threshold to each class based on user-defined criticality levels.

CDSEM images with no predicted classification (zero probability) from the models will be assigned an "unknown class." These images are collected for future training, ensuring continuous improvement and adaptation of the models to evolving inspection requirements.

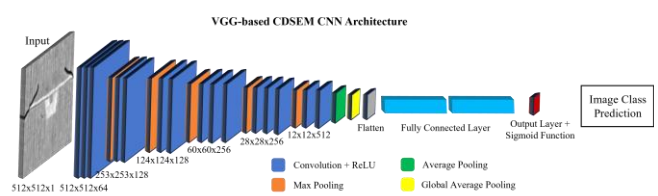


Fig. 5. VGG-based CDSEM CNN architecture.

3.5 Machine Learning Development Life Cycle

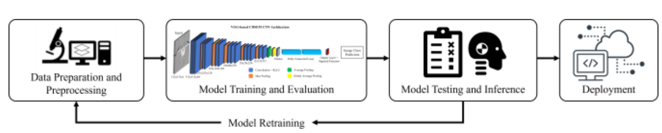


Fig. 6. Machine learning development and deployment workflow.

The machine learning development life cycle for multiclass image classification encompasses several vital stages: Image dataset preparation and preprocessing, model training, testing, and deployment (see Fig. 6). This iterative process is crucial for building and deploying effective defect classification systems. The following sections elaborate on each stage of this development life cycle:

3.5.1 Data Preparation and Preprocessing

The CDSEM images containing good and defective samples are gathered from the BDP at this stage. Folders are created for each category, including major, minor, and monitoring defects. Each image is then categorized and stored in the folder associated with its respective defect type. Subsequently, the dataset is partitioned into training, testing, and validation subsets. The training dataset is utilized to train the model, while the testing and validation datasets are employed to evaluate its performance and generalize its ability to unseen data.

3.5.2 Classification Model Training and Evaluation

During this stage, the prepared dataset is used to train a deep learning model to recognize and classify defects within the images. As illustrated in Fig. 5, The model's architecture is configured with layers based on the VGG network. Training also involves optimizing the model's parameters by configuring different optimizers, learning rates, batch size, and other essential hyperparameters. Confusion matrix is also utilized to provide a comprehensive breakdown of a model's performance, highlighting areas of over or under-rejection, which is crucial for understanding where the model might be making errors.

3.5.3 Classification Model Testing and Inference

Following training and evaluation, the trained model is tested using the validation dataset to assess its generalization performance on unseen defect instances. This step ensures that the model can effectively detect defects in real-world scenarios. The model is deployed during inference to predict new images and identify defects within them.

3.5.4 Model Deployment

Once the trained model has provided a reasonable prediction output during the testing and inference stage, the model will then integrate with the CDSEM AI system and deploy on the edge cluster. The system runs every nth hour to fetch data from the BDP. It performs the classification task to produce a prediction output displayed on the dashboard for monitoring purposes. The CDSEM AI system will also generate a metadata tag containing information about the predicted category name and its associated probability. This metadata tag will be uploaded to the BDP, as shown in Fig. 7.

Fig. 7. The metadata tag for prediction result is uploaded in BDP.

3.6 Dashboard

Using the Analytics dashboard platform, a dashboard is created to visualize and interact with the status tables and trend charts. This dashboard is a vital tool for defect analysis, providing stakeholders with actionable insights to improve production efficiency. The dashboards feature the overall distribution of defect categories per day, daily lot rejection rate, and defect trend per wafer. It also includes information about the final disposition and whether the wafer can be released or held. Lastly, it also contains a wafer map where the user can track the location of the slider with its specific defects in the wafer. See the Appendix for some of the features of the dashboard.

4.0 RESULTS AND DISCUSSION

Implementing the CDSEM AI system has brought about notable improvements in the efficiency and accuracy of inspection processes. These systems can accurately analyze images, significantly reducing the need for manual intervention using the created CNN architecture based on the VGG network. Consequently, this has substantially optimized the workforce required for the inspection process by as much as 50 percent. By automating repetitive tasks and streamlining workflows, CDSEM personnel can allocate their time more effectively to strategic and value-added activities.

Moreover, integrating an automated dashboard and Key Performance Indicator charts has provided real-time insights into the performance of the inspection processes. These dashboards offer a comprehensive overview of key metrics such as defect detection rates, inspection throughput, and quality assurance metrics, enabling timely decision-making and process optimization. Additionally, automated notification systems have been implemented to streamline the communication process for holding and releasing wafers based on inspection results. Automatically generating alerts and notifications ensures prompt action in response to identified defects or anomalies, minimizing production delays, and optimizing yield.

Furthermore, automated inspection systems have standardized the inspection process, eliminating variation introduced by human judgment. Moreover, the custom CNN architecture enables the system to enhance sensitivity to unusual defects or anomalies that might be overlooked in manual inspection. It improves defect detection rates and enhances overall product quality. Overall, implementing automated image inspection systems has yielded significant benefits in efficiency, accuracy, and process optimization, ultimately leading to enhanced productivity and product quality.

5.0 CONCLUSION

In conclusion, adopting automated image inspection and judgment systems has revolutionized inspection processes, significantly improving efficiency, accuracy, and process optimization. Through deep learning algorithms, manual intervention has seen a remarkable reduction, leading to workforce optimization by 50 percent. Integrating automated interactive dashboards and KPI charts has provided invaluable real-time insights, enabling timely decision-making and effective process optimization. Additionally, implementing automated notification systems has streamlined communication and minimized production delays, ultimately optimizing yield.

Furthermore, standardizing the inspection process and enhancing sensitivity to unusual defects achieved through automation has improved defect detection rates and overall product quality.

As the company embraces technological innovation, its commitment to automation ensures competitiveness in the global market.

6.0 RECOMMENDATIONS

Based on the findings of the project, the following recommendations are proposed to improve the processes further:

Explore Alternative Object Detection and Classification Algorithms. It is recommended that alternative object detection and classification algorithms be explored to enhance the classification model's performance. By experimenting with different algorithms, approaches can be identified that offer superior accuracy and efficiency, improving the overall effectiveness of the image classification system.

Fully Automate Final Wafer Disposition. Integrating CDSEM AI with the company's auto-holding system can further streamline the processes by fully automating the final wafer disposition. By seamlessly integrating AI-driven CDSEM capabilities with the existing systems, greater efficiency, and accuracy in determining the disposition of wafers might be achieved.

Enhance Dashboard Functionality. The dashboard's functionality should be enhanced to meet the evolving needs of the end users.

By implementing these recommendations and pursuing future work, we can continue to enhance our image inspection and classification systems' effectiveness, efficiency, and reliability, driving continuous improvement and innovation within our organization.

7.0 ACKNOWLEDGMENT

The authors would like to extend their heartfelt appreciation to the QA Engineering department for providing the image dataset and for their dedication in evaluating the classification model's performance from user acceptance testing to deployment. They also thank the IT department for assisting with the application software necessary for the project's development. Finally, the authors acknowledge the Process Excellence department's valuable knowledge and support in this project.

8.0 REFERENCES

1. R. Brown, C. Power, E. Pike, and M. Lenahan, "Improved Cycle Time of CDSEM Measurements Through Novel Alignment Methodologies & Data Storage Optimization," *2023 34th Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC),* Saratoga Springs, NY, USA, 1-6, 2023, doi: 10.1109/ASMC57536.2023.10121084.

2. Y. Midoh, and K. Nakamae, "Image Quality Enhancement of a CD-SEM Image using Conditional Generative Adversarial Networks," *Proc. SPIE 10959, Metrology, Inspection, and Process Control for Microlithography XXXIII, 109590B,* 2019, doi: 10.1117/12.2515152.

3. T. Bourguignon *et al*., "Enabling On-device and Target-free Overlay Measurement from CD-SEM Contours," *Proc. SPIE 12053, Metrology, Inspection, and Process Control XXXVI, 120530C*, 2022, doi: 10.1117/12.2613327.

4. B. Dey *et al*., "Unsupervised Machine Learning based CD-SEM Image Segregator for OPC and Process Window

Estimation," *Proc. SPIE 11328, Design-Process-Technology Co-optimization for Manufacturability XIV, 113281G*, 2020, doi: 10.1117/12.2552055.

5. W. Sun, Y. Goto, T. Yamamoto, and K. Hitomi, "Evaluation of Deep Learning Model for 3D Profiling of HAR Features using High-voltage CD-SEM," *Proc. SPIE 11611, Metrology, Inspection, and Process Control for Semiconductor Manufacturing XXXV, 116110X*, 2021, doi: 10.1117/12.2592052.

6. I. Chen, R. Hwang, and H. Huang, "PCB Defect Detection Based on Deep Learning Algorithm," *Processes 2023*, 11, 775, 2023, doi: 10.3390/pr11030775.

7. R. Piyush *et al*., "Predictive Maintenance for Industrial Equipments Using ML & DL," *2023 International Conference on Advanced Computing & Communication Technologies (ICACCTech)*, Banur, India, 391-396, 2023, doi: 10.1109/ICACCTech61146.2023.00071.

8. Y. Ma, "Integrated Forecast of Financial Time Series Data Based on Deep Learning Model," *2022 IEEE Conference on Telecommunications, Optics and Computer Science (TOCS)*, Dalian, China, 1140-1144, 2022, doi: 10.1109/TOCS56154.2022.10016211.

9. Z. Xie, N. Somani, Y. J. S. Tan, and J. C. Y. Seng, "Automatic Toolpath Pattern Recommendation for Various Industrial Applications based on Deep Learning," *2021 IEEE/SICE International Symposium on System Integration (SII)*, Iwaki, Fukushima, Japan, 60-65, 2021, doi: 10.1109/IEEECONF49454.2021.9382625.

10. N. Zakaria, and Y. M. Mohmad Hassim, "A Review Study of the Visual Geometry Group Approaches for Image Classification," *Journal of Applied Science, Technology and Computing*, 1(1), 14-28, 2024, https://publisher.uthm.edu.my/ojs/index.php/jastec/article/view/16010.

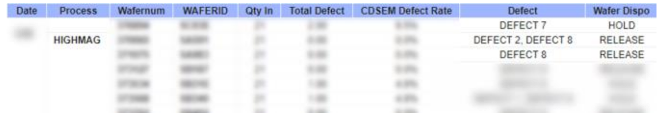
9.0 ABOUT THE AUTHORS

Jerald Constantino is an Associate Data Scientist at Western Digital, specializing in machine learning and generative AI projects. He holds a Bachelor of Science in Computer Engineering from Batangas State University - TNEU Alangilan.

James Gonzales is a Data Scientist at Western Digital, Philippine Head Office. He holds a master's degree in data science from Asian Institute of Management and bachelor's degree in Electronics and Communications Engineering from Polytechnic University of the Philippines – Taguig. He has over 15 years combined experience working in magnetic head industry as a Test Engineer and Data Scientist.

10.0 APPENDIX

Appendix A – The Final Wafer Disposition per Wafer and per Inspection Process. The table contains information about which wafers should be held or released. It also includes the predicted number of sliders, the total defect, and the defect rate detected within the wafer.



Date	Process	Wafernum	WAFERID	Qty In	Total Defect	CDSEM Defect Rate	Defect	Wafer Dispo
[illegible]	HIGHMAG	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	DEFECT 7	HOLD
		[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	DEFECT 2, DEFECT 8	RELEASE
		[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	DEFECT 8	RELEASE
		[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
		[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
		[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Appendix B – The Wafer Map and the Location of the Defected Slider for the Selected Wafer. The illustration provides a representation of Sliders and their associated defects within the wafer. A summary table for each detected defect is also provided on the right side of the dashboard.

Appendix C – The Auto-Email Alerts. Two emails are sent to the users. The first is a snapshot alert of the current trend regularly sent to users. The second is an immediate auto-email alert sent when the final wafer dispositions are held.