[Public]

The Application of Machine Learning in the Next Frontier of Failure Analysis Fault Isolation

Bernice Zee*, Angeline Phoa, Syahirah Zulkifli, Qiu Wen, Oh Ziying
Advanced Micro Devices (Singapore) Pte Ltd, Device Analysis Lab, 508 Chai Chee Lane, Singapore 469032
Phone: +65 6796 9888 *Email: bernice.zee@amd.com

Abstract

The continued scaling of transistor technology in tandem with the rapid development of next generation packaging technologies has presented interesting challenges to current failure analysis techniques. The ability to localize and visualize defects prior and during physical failure analysis (PFA) is essential for successful root cause analysis. Failure analysis (FA) typically follows a workflow to effectively isolate the failure and determine the root cause. In recent years, machine learning methodologies have been applied to both fault isolation and non-destructive imaging steps of the FA workflow to help aid defect detection as defects have become more subtle and challenging to differentiate in denser and more complex semiconductor packages [2]. Depending on the technique, different machine learning methodologies are applied. This paper provides an overview of recent applications of machine learning in the semiconductor FA workflow. Some of the areas covered include the use of machine learning based methodologies for computer vision (CV) based defect detection through images comparison between good and reject device [8], unsupervised and supervised learning techniques using independent component analysis (ICA) to analyze acoustic microscopy data [11, 12], and deep learning based high resolution reconstruction technique 3D X-ray imaging [13, 14].

1. Introduction

Transistor scaling continues to defy predictions that Moore's Law is dead. New transistor architectures and patterning processes are driving transistor densities of next generation semiconductor chips required for high performance computing and data centers. However, pursuing transistor scaling alone to achieve performance and more functionalities comes at a high price and the economics are becoming unrealistic for most to adopt. Furthermore, the semiconductor industry is no longer driven purely by performance; miniaturization, low latency, and high bandwidth requirements are becoming just as important. This has led to innovations in next generation packaging technologies for heterogeneous integration [1]. Examples of such packaging technologies include Wafer-Level Fan-Out (WLFO), Through-Silicon-Vias (TSV), 3D interconnects and embedded bridges. Heterogeneous integration enables communication between different functional dies at a lower cost on a package substrate as compared to a fully integrated monolithic silicon die. Moreover, heterogeneous integration allows faster time to market as designers could potentially adopt intellectual property (IP) reuse strategies.

2. Failure Analysis (FA) Workflow

Failure analysis (FA) is needed at every step of the product time to market cycle, from identifying design issues at early product development stages, process issues during technology qualification and high-volume manufacturing, and field failures. FA is needed at both the silicon transistor and package interconnect levels as the types of defects observed vary vastly.

Failure analysis typically follows a workflow shown in Figure 1 to effectively isolate the failure and determine the root cause [2].

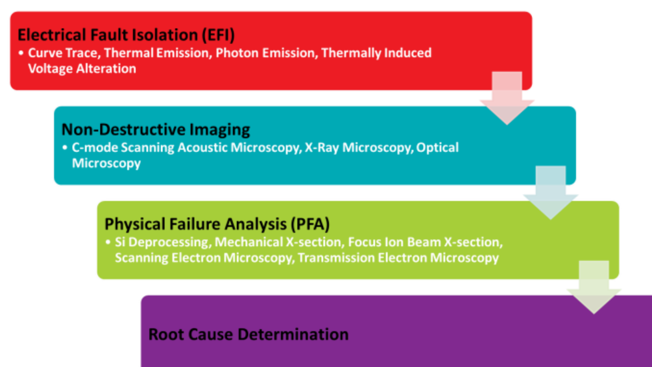

Figure 1. Schematic of FA flow.

First, electrical fault isolation (EFI) techniques like lock-in thermography (LIT), photon emission microscopy (PEM), and time domain reflectometry (TDR) are used to characterize short, leakage, high resistance, and open failures [2, 3, 4] in a device under test (DUT). Thermal or photon emission sites observed during static or dynamic testing help to localize an area of inspection. Second, non-destructive imaging techniques such as Confocal-mode (C-mode) scanning acoustic microscopy (SAM), and X-ray inspection are used to look for and image defects. CSAM is used to inspect semiconductor package interfacial integrity [5] and X-ray is used to study the internal structural integrity of interconnects and metal traces [6]. Third, destructive physical failure analysis (PFA) is conducted based on the position identified by EFI and non-destructive imaging. PFA of the DUT is done by mechanical cross-section or Focused Ion beam (FIB). The prepared sample is then imaged using scanning electron microscopy (SEM) or transmission electron microscopy (TEM) and elemental analysis can be done using energy dispersive X-ray (EDX). The consolidated results obtained from steps one to three are finally used to determine the root cause of failure.

3. Machine learning to aid FA

The continued scaling of transistor technology in tandem with the rapid development of next generation packaging technologies has presented interesting challenges to current FA techniques. Higher transistor densities equate to longer times required for emission site verifications across units and a higher chance for erroneous human judgement in silicon level EFI. Additional silicon sample preparation steps are also required to improve signal to noise ratio (SNR) and sensitivity for defect detection. Complex die stacking and numerous die-to-die interconnect with no direct electrical contact to electrically stimulate specific functions in a die, make package level EFI equally challenging. The resolution of non-destructive imaging techniques has also been pushed to theoretical limits with highly miniaturized interconnects. Trade-offs in terms of field of view (FOV), imaging depth and scan times must be balanced with little wriggle room. Nanometer sized three-dimensional (3D) transistor architectures have led to more occurrences of Non-Visual Defects (NVD) and complicated sample preparation workflows of defect imaging [7].

This paper aims to provide an overview of recent applications of machine learning in the semiconductor FA workflow to aid defect prediction, localization, and visualization. Some of the areas covered include the use of machine learning based methodologies for computer vision (CV) based defect detection through images comparison between good and reject device [8], unsupervised and supervised learning techniques using independent component analysis (ICA) to analyze acoustic microscopy data [11, 12], and deep learning based high resolution reconstruction technique for 3D X-ray imaging [13, 14].

4. Case Studies

4.1 AI Based PEM

Photon Emission Microscopy (PEM) is a commonly used semiconductor fault isolation technique. Photons in infra-red wavelength emitted by transistors and other semiconductor devices, such as p-n junction diodes, are collected with the help of highly sensitive detectors while the device under test is electrically exercised through a static or dynamic setup. [2] The photons would be displayed as hot spots when overlaid with the image of the die as shown in Figure 2.

PEM is useful to detect failures such as shorts, damage at electrostatic discharge (ESD) protection structures when there are excessive photons being emitted at the site of the failure. However, photons are also being emitted for good transistors during switching. Multiple emission sites are expected during an analysis, so it is critical to compare to a reference unit to find the anomalous location. This method, however, is manual and error prone.

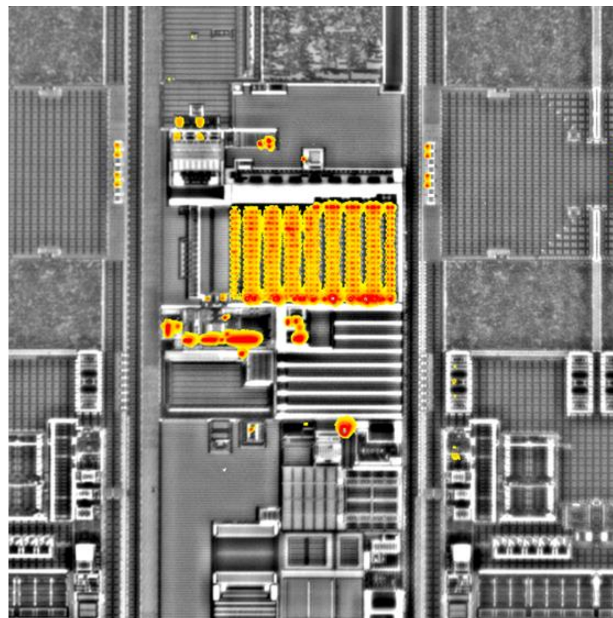

Figure 2. Photon Emission image with hot spots overlaid on die image as yellow/red spots.

Single Shot Detection (SSD) model, one of Computer Vision (CV) object detection algorithm based on the VGG16 neural network [8], is implemented for this work. SSD is chosen as it is the first deep network-based object detector that does not resample pixel or features for bounding box hypotheses and is as accurate as approaches that do [9]. This is made possible through the elimination of bounding box proposals and using default boxes instead, thereby improving the speed of real time applications.

Machine learning or deep learning takes training data then learns from its features to be able to execute user's purpose, e.g., object detection on the testing data (a totally different set of images from training data set) correctly. The training data set should be as relevant as possible to the testing data but covers its variation as exhaustive enough. In general, the more training data set the more exhaustive it is.

Augmentation [9] is commonly used to multiply training data set to reduce overfitting problem. The advantages are no additional image is required and it still helps to reduce overfitting problem by adding data set variation. A common augmentation technique is image mirroring and rotation, as shown in Figure 3. For each image in data set, every object present in it is noted down along with its bounding box coordinate (e.g., in the format of Xmin, Ymin, Xmax, Ymax). This annotation list will be used by the SSD network during learning process.

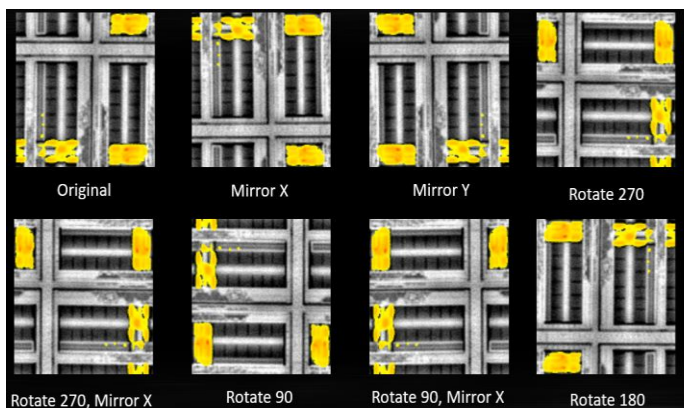

Figure 3. Augmentation to multiply dataset

4.1.1 Applications of AI Based PEM

In a typical PEM image, the hot spots with highest intensity have highest probability of being the defective sites. The model has been trained to highlight the top five highest intensity hot spots (in purple boxes) in a single image, as shown in Figure 4. This reduces the need for manual screening to pick out these sites. In addition, the rectangular features in the die are also detected in this mode, as highlighted in green boxes. The hot spot and rectangular circuitry are detected with the accuracy of 70% and 90% respectively.

Figure 4. Results from a single image detection

A more definitive defect detection is done on two PEM images. The hot spots from a failing unit and a good unit, aligned to each other using the green boxes, are compared. By implementing Jaccard Similarity Index [10] method, hot spots found only on either image will be highlighted, as shown in Figure 5 below. The blue box indicates hot spots found on a reject's PEM but not on a good PEM image. White boxes indicate hot spots found on a good PEM but not on a bad reject image. This highlights the differences between the images and helps to decide on the location for further in-depth analysis.

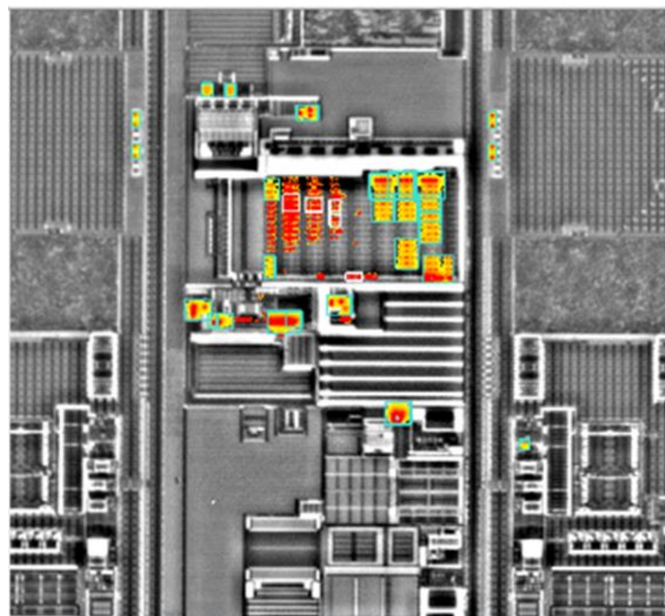

Figure 5. Results of two image comparison

More training is in progress to improve the accuracy of hot spot detection.

4.2 AI in 3D X-ray Microscopy

3D X-ray Microscopy (XRM) is a well-known FA technique for visualizing defects in semiconductor packages without destructive PFA. X-ray absorption varies between different materials and geometries, and this provides the basic contrast mechanism for X-ray imaging. In 3D XRM, a DUT is placed between an X-ray source and detector and rotated axially. A series of 2D X-ray projection images taken at different angles is initially enlarged through geometric magnification. Further magnification of these images is achieved when a high-resolution scintillator converts X-rays into photons, thus enabling optical magnification. Subsequently, a 3D volume model is reconstructed from the collected 2D X-ray images using mathematic models and algorithms.

The application of 3D XRM in next generation packaging technologies FA has been demonstrated and reported [6, 16-17]. It provides improved spatial resolution and contrast which is necessary when imaging subtle internal defects in semiconductor packages with dense build up, miniaturized interconnects, and stacked components. However, the reconstructed 3D volume model may suffer from noise and under- sampling artifacts carried over from the 2D projection images when using traditional 3D reconstruction methods such as Feldkamp-Davis-Kress (FDK) method [13,14]. Noise can be reduced by increasing the exposure time per projection, and/or collecting more projections but this comes at the cost of scan throughput.

To improve 3D XRM scan efficiency without compromising image quality and resolution, in references [13, 14], the

authors propose using a new convolutional neural network algorithm to reconstruct 3D volume model. The method is termed “Deep Learning High-Resolution Reconstruction (DLHRR)” and it is based on “noise2noise” model with their proprietary cost function. Improved scan throughput is achieved through the implementation of pre-trained neural networks which extracts signals from low-dose data more efficiently than the FDK method. The training input comprises a dataset with a low number of 2D projection images with high pixel noise and the training target is a 3D volume model created from a high number of projection images with low pixel noise which serves as the “ground truth” data. The neural network training is constrained to a strict sample class and X-ray acquisition condition; thus, retraining is needed if the type of sample and/or the X-ray scan parameters changes. Conversely, the retraining can be done relatively quickly using one 3D volume data set which contains thousands of projection images which are augmented to account for possible variations in sample and data acquisition conditions.

4.2.1 Applications of AI Based 3D XRM

Figure 6 shows a comparison in image quality of an AMD product with μbump interconnect defect when using traditional FDK reconstruction method versus DLHRR reconstruction method. FDK reconstructed slices shown in Figure 6 (a-b) typically required 1600 projections and a 9.6-hour tomographic scan to image an approximately 1μm thick bump crack. On the other hand, DLHRR reconstructed slices in Figure 6 (c-d) showed very similar image quality with less noise on the corresponding crack with only 400 projections and a 2.4-hour tomographic scan. The results clearly demonstrate the benefit of deep learning-based reconstruction algorithms to improve scan efficiency without compromising image quality.

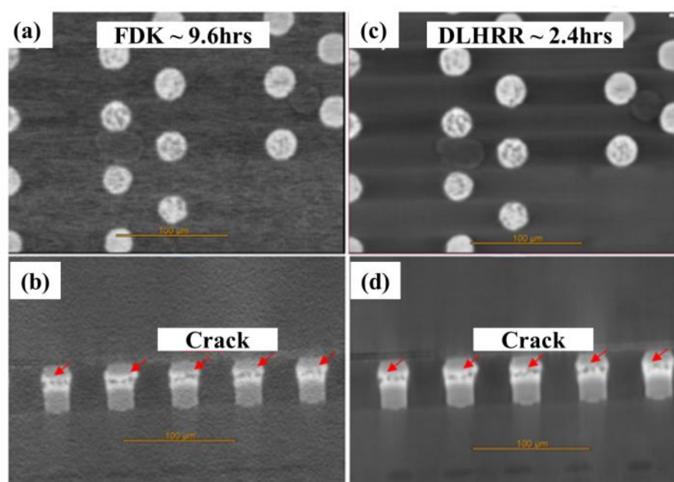

Figure 6. Comparison of results from DLHRR to FDK methods on a 50x50 mm AMD HBM μbump package. (a-b) virtual cross section slices from the FDK 9.6-hour scan, and (c-d) virtual slices from the DLHRR 2.4-hour scan. (a) and (c) are the top-down views, and (b) and (d) are the cross-sectional views.

4.3 AI Based SAM

Scanning acoustic microscopy (SAM) is also another frequently used FA tool for visualizing defects in semiconductor packages without destructive PFA [18]. Acoustic waves in the range of 5 to 500 MHz are generated by a piezoelectric transducer and these are focused and transmitted to a DUT through distilled water. When the acoustic waves interact with a device, part of the waves is reflected to the transducer and the other part is transmitted through the device. Either reflected or transmitted acoustic waves are processed to analyze the internal features. C-mode SAM (C-SAM) images are most frequently used for inspecting interfacial integrity (e.g., interfacial delamination, TIM inspection), and detecting defects (e.g., voids, white spots, black spots, die crack) in semiconductor packages. C-SAM images are created using reflected echoes from a fixed depth within a selected thickness of the device. This is done by “gating” the echoes from A-scan signals which display a reflected acoustic pulse at a given location on the device.

The application of C-SAM in next generation packaging technologies FA has been discussed in reference [5]. The appropriate hardware such as high-speed analogue-to-digital converter (ADC) for nano-seconds gating, transducers of suitable focal length and frequency as well as active z-axis positioning of the transducer during C-scan are optimized for high-resolution acoustic imaging of defect detection in 2.5D packages. However, deciphering acoustic signals and images is becoming more challenging, especially as defects are shrinking with interconnect size and are getting masked by features generated from more complex package architectures. Data interpretation relying on an analyst’s experience is usually subjective and this makes the results vulnerable to human error.

Automated defect detection in semiconductor interconnect bumps using acoustic imaging was explored more than a decade ago [19]. It made used of backscatter amplitude integral (BAI) signal analysis to digitally filter and pre-process raw echo signals offline to suppress noise and amplify signals of interest such as those associated with interconnect bumps. These signals obtained were then further inspected using wavelet, and pulse separation analysis. Once all the appropriate signal parameters related to the interconnect bumps were derived, they were combined in a classification model for deriving a binary void-intact decision to classify defective and non-defective interconnect bumps. Such a methodology required substantial human intervention which is not ideal.

Image-based machine learning algorithms have advanced significantly and their broad application for C-SAM defect detection in semiconductor packages has been considered. However, the challenge of developing such machine learning models is the scarcity of training data, especially in a FA lab setting where data is typically high-mix and low-volume [12]. In fact, C-SAM images of defective devices are particularly challenging to obtain and image augmentation is necessary to generate sufficient data for training. Another challenge faced when using image-based machine learning models is the

quality of training data; data consistency in the definition of defects is crucial and the mantra "Garbage in, Garbage out" aptly describes what can go wrong. Defects have different grey values in different time domain and marginal differences in the echo gating will affect the C-SAM image intensity and the resultant image consistency. Furthermore, manual labeling of large image data sets is time consuming, and even if the model could be successfully trained, poor image contrast and/or the lack of pixel definition can cause a defect to go undetected in the C-SAM image.

To address the above challenges, the authors of references [11] and [12] have proposed methods focused on classifying A-scan signals with a small percentage of labeled sample data to train a deep learning based neural network. A-scan signal-based networks are not affected by shape or intensity of defects, assuming the same transducer and scanning settings are used for collecting the data. Reference [11] used independent component analysis (ICA) to extract independent signal components and their weighting matrices to characterize and label interconnect bumps into two categories: (1) intact and (2) defective. Subsequently, a labeled dataset is generated and used to train their one-dimensional convolutional neural networks (1D-CNN). Reference [12] used wavelet filtering to denoise the A-scan signals prior to labeling. Labeling is done for a group of signals from a region of interest defined in a C-SAM image and four classes were defined in their model: (1) defect type I, (2) defect type II, (3) intact, and (4) structure. Their labeled data set was used to train four different deep neural networks, namely recurrent neural network-long short-term memory (RNN-LSTM), recurrent neural network-gated recurrent units (RNN-GRU), RNN-BiLSTM and RNN-1D CNN. The highest accuracy was obtained with RNN-1D CNN. Both references demonstrated uses for their model with good accuracy.

4.4 Other AI Based FA Techniques

A combination of modelling and machine learning to detect FinFET transistor level defects is discussed in reference [7]. Technology Computer-Aided Design (TCAD) modelling is used to generate electrical responses for transistor defects at different locations and sizes. These electrical responses serve as the attribute in the datasets for training, testing, and validating the random forest based predictive model as actual transistor failure data is scarce due to the sporadic occurrence and the many variants of such defects. The predictive model is expected to improve FA success rate by predicting real failures in device. SEM image denoising is another key area where there is much opportunity to apply machine learning for enhancing image resolution and defect detection especially as the dimensions of die circuitry shrink in the more advanced Si nodes [20].

5. Conclusion

Shrinking transistor and interconnect dimensions are pushing the limits of current FA techniques in terms of time to results, resolution, and sensitivity. In addition, manually conducted activities such as inspection are becoming very challenging. The implementation of machine learning in the FA flow to aid defect prediction, localization, and visualization helps to alleviate these challenges. There is no “one-size-fits-all” machine learning approach for all the FA techniques. In fact, the method used for each FA technique depends on the type of data that can be collected from the tool to teach the model successfully and achieve good performance that will translate to better FA success rate. With the vast number of machine learning approaches available, there is a lot of research potential for machine learning in the FA flow.

6. Acknowledgements

The authors would like to thank their collaborators and colleagues, especially Nathan Linarto, and Teo Chea Wei, who have contributed to this work.

© 2022 Advanced Micro Devices, Inc. All rights reserved. AMD, the AMD Arrow logo, and combinations thereof are trademarks of Advanced Micro Devices, Inc. Other product names used in this publication are for identification purposes only and may be trademarks of their respective companies.

References

1. Li, T.; Hou, J.; Yan, J.; Liu, R.; Yang, H.; Sun, Z. Chiplet Heterogeneous Integration Technology — Status and Challenges. *Electronics* 2020, *9*, 670. https://doi.org/10.3390/electronics9040670.
2. J.M Chin, *et al.*, “Fault isolation in semiconductor product, process, physical and package failure analysis: Importance and overview”, 22nd European Symposium on Reliability of Electron Devices, Failure Physics and Analysis, Oct 2011, pp. 1–10.
3. Cao, L., Venkata, M., Huynh, J., Tan, J., Tay, M. Y., Qiu, W., “Lock-in Thermography for Flip-chip Package Failure Analysis”, *Proc 38th Int’l Symp for Testing and Failure Analysis*, Phoenix, AZ, November 2012, pp. 316–324.
4. Qiu, W., Zee, B., Deslandes, H., Lai, B., Tien, D., “Defect Z-depth Determination in Flip-chip Using Lock-in Thermography”, *Proc* 24th *Int’l Symp on the Physical and Failure Analysis of Integrated Circuits*, Chengdu, China, July 2017, pp. 1946-1550.
5. Z. Y. Oh, F. J. Foo and B. Zee, "Optimization and Application of Acoustic Imaging for Defect Detection in Stack Die Packages," 2018 IEEE International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA), 2018, pp. 1-5, doi: 10.1109/IPFA.2018.8452515.
6. S. Md Zulkifli *et al.*, “High-Res 3D X-ray Microscopy for Non-Destructive Failure Analysis of Chip-to-Chip Micro-bump Interconnects in Stacked Die Packages”, Proceedings of 24th International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA), 2017.
7. C. -W. Teo, K. L. Low, V. Narang and A. V. -Y. Thean, "TCAD-Enabled Machine Learning Defect Prediction to Accelerate Advanced Semiconductor Device Failure

Analysis," 2019 International Conference on Simulation of Semiconductor Processes and Devices (SISPAD), 2019, pp. 1-4, doi: 10.1109/SISPAD.2019.8870440.

8. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.Y., Berg, A.C.: SSD: Single Shot MultiBox Detector. arXiv e-prints arXiv:1512.02325 (dec 2015).
9. Perez, Luis, and Jason Wang. 'The Effectiveness of Data Augmentation in Image Classification Using Deep Learning'. ArXiv:1712.04621 [Cs], Dec. 2017. arXiv.org, http://arxiv.org/abs/1712.04621.
10. Moulton, Ryan; Jiang, Yunjiang (2018), "Maximally Consistent Sampling and the Jaccard Index of Probability Distributions", International Conference on Data Mining, Workshop on High Dimensional Data Mining.
11. M. Kögel, et. al., "Machine learning assisted signal analysis in Acoustic Microscopy for nondestructive defect identification", Proc 45th Int'l Symposium for Testing and Failure Analysis, Portland, OR, Nov. 2019, pp. 35-42.
12. A. S. Nair, et al., "Automated defect classification in semiconductor devices using deep learning networks", Proceedings of 29th International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA), 2022.
13. A. Gu, et. al., "Accelerate Your 3D X-Ray Failure Analysis by Deep Learning High Resolution Reconstruction", Proc 48th Int'l Symposium for Testing and Failure Analysis, Phoenix, AZ, Oct. 2021, pp. 291-295.
14. A. Gu, et. al., "From System to Package to Interconnect: An Artificial Intelligence Powered 3D X-ray Imaging Solution for Semiconductor Package Structural Analysis and Correlative Microscopic Failure Analysis", 2022 IEEE 29th International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA), 2022.
15. S. Göke, *et al.*, " Scaling AI in the sector that enables it: Lessons for semiconductor-device makers," [Online]. Available: https://www.mckinsey.com/industries/semiconductors/our-insights/scaling-ai-in-the-sector-that-enables-it-lessons-for-semiconductor-device-makers. [Accessed 8 8 2022].
16. Y. Sylvester et al., "3D X-ray microscopy: A non-destructive high-resolution imaging technology that replaces physical cross-sectioning for 3DIC packaging," ASMC 2013 SEMI Advanced Semiconductor Manufacturing Conference, Saratoga Springs, NY, 2013, pp. 249-255.
17. C. Y. Liu, P. S. Kuo, C. H. Chu, A. Gu and J. Yoon, "High resolution 3D X-ray microscopy for streamlined failure analysis workflow," 2016 IEEE 23rd International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA), Singapore, 2016, pp. 216-219.
18. Aryan P, Sampath S, Sohn H., "An Overview of Non-Destructive Testing Methods for Integrated Circuit Packaging Inspection.", Sensors (Basel). 2018 Jun 21;18(7):1981. doi: 10.3390/s18071981. PMID: 29933589; PMCID: PMC6068802.
19. S. Brand, P. Czurratis, P. Hoffrogge, M. Petzold, "Automated inspection and classification of flip-chip-contacts using scanning acoustic microscopy", Microelectronics Reliability, Volume 50, Issues 9–11, 2010, Pages 1469-1473.
20. Bappaditya Dey, Sandip Halder, Kasem Khalil, Gian Lorusso, Joren Severi, Philippe Leray, Magdy A. Bayoumi, "SEM image denoising with unsupervised machine learning for better defect inspection and metrology," Proc. SPIE 11611, Metrology, Inspection, and Process Control for Semiconductor Manufacturing XXXV, 1161115 (22 February 2021); https://doi.org/10.1117/12.2584803