Studies on Circuit Layout Analysis for the Reliability of LSI Devices

presented by

Yoshikazu Nagamura

The Graduate School of System Design

Tokyo Metropolitan University

Abstract

Physical failure analysis (PFA) specifies layout designs that affect large-scale integration (LSI) failure. Because of their capability and cost effectiveness, convolutional neural networks (CNNs) have been applied to LSI layout analysis. However, the information on failure for root cause analyses is generally limited. Moreover, information over a large area, which includes many geometries, is required to understand the effects of a layout on failure. These problems pose challenges in training the CNN models to achieve high accuracy for layout segment classification. In this study, a scheme for layout segment classification is proposed, which uses CNN to analyze the effects of layout on failure. Thus, multiple segment images of LSI layouts are used as inputs, and the outputs of CNN models are used to classify the input images as either risk or non-risk segments. First, a simple 3-layer model was evaluated, and then a 15-layer model fine-tuned using the transfer learning of the VGG16 model was evaluated. The 15-layer model outperformed the 3-layer model for classifying the LSI layout segments. From a cross-validation of the 15-layer model, a true positive rate of >80% and a false positive rate of <10% are obtained for extracting layout regions related to actual defects. The outputs of the CNN models from the input layout segment demonstrate similarity with the defective layouts. In the regional layouts across the LSI chips, the contour plot of model outputs is visualized as a hazard map of failures. This information is necessary for additional failure cause analyses.

Root cause analysis (RCA) of failures is mandatory to obtain the reliability and productivity of LSIs. Although analyzing layout-induced defects

is crucial to optimize design rules and to predict unknown defects, it is a challenging task due to the difficulty in explaining the relationship between defects and circuit layouts. CNNs were applied to classify LSI layout images to perform the RCA of layout-induced defects in a former study. However, due to the low resolution of images, actual defect positions were not clearly distinguished. In the next study, image segments of different sizes and resolutions were used for the CNN classification. Experimental results indicate that the validity of the extracted layout features depends on the resolution of image segments. Using the visual explanation technique GradCAM++, the features of defective layouts can be accurately captured in local areas including a group of patterns with their surroundings when CNN models are trained on smaller image segments with higher resolution. Conversely, utilizing smaller-size segments deteriorates the classification accuracy due to the incorporation of less information from the images. In the conducted experiments, even in the case of using smaller segmentI s, acceptable performance (the detection rate of defect positions DTR ≅90%, and the risk-image classification rate RCR ≅10%) can be obtained by increasing the size of training datasets. Partial layouts extracted as features of defective layouts can then be used in RCA and in designing future products.

Test quality is critical to eliminate test escapes and to achieve high-reliability LSI devices. A new concept called " physical test coverage " is proposed to verify test coverage based on the physical layout of LSI circuits. The physical test coverage is calculated as the ratio between the critical area of all wires in a device and that of wires undetected by LSI tests. From the critical area of undetected wires and the defect density of a manufacturing

line, the risk of test escapes can be predicted. To effectively develop LSI tests that can minimize the number of test patterns, undetected wires are prioritized by the critical area related to each wire. Even when the conventional "logical" test coverage is high enough to satisfy the coverage criterion, some LSI devices investigated in this study showed low physical test coverage depending on the physical layout of the LSI circuit. The concept of physical coverage was applied in the test development of some LSI products, and the test quality was substantially improved, such that 90% of test escapes of a device were eliminated.

Contents

1 Introduction **1**
1.1 Background 1
1.2 Objectives of this work 5
1.3 Structure of this dissertation 7

2 Preliminaries **11**
2.1 LSI failure analysis 11
2.2 Prediction of LSI failure rate and yields 15
2.3 Design for manufacturing(DFM) and layout analysis 19
2.4 Machine learning and Artificial intelligence(AI) 22

3 CNN-based layout segment classification for analysis of LSI layout-induced failures **27**
3.1 Introduction 27
3.2 CNN models 31
3.3 Proposed scheme 34
3.3.1 Overview of layout segment classification 34
3.3.2 Input data for classification 36
3.3.3 Learning process (model training) 36
3.3.4 Outputs of CNN models 37
3.4 Experiments and results 38
3.4.1 Datasets for model training 38
3.4.2 Definition to judge " Risk " in defect positions 38
3.4.3 Comparison of 3- and 15-layer CNN classification 40
3.4.4 Cross-validation of 15-layer CNN classification 43
3.4.5 Visualization of impact on failures in the LSI layout 47
3.5 Conclusions 51

4 Layout feature extraction using CNN classification in root cause analysis of LSI defects **53**
4.1 Introduction 53
4.2 Overview of experiments 56

4.2.1 Proposed scheme of the layout analysis based on CNN classification 56
4.2.2 Experimental conditions 59
4.3 Feature extraction using high-resolution images 61
4.3.1 Visualization of the circuit layout features extracted by the CNN models 61
4.3.2 Extracting layout features using a virtual layout image 63
4.3.3 Effect of image resolution on the results of feature extraction 65
4.4 Classification of layout images 67
4.4.1 Performance of CNN models in classifying layout images 67
4.4.2 Improvement of model classification results by increasing the size of training datasets 71
4.4.3 Hazard maps of defect generation 72
4.4.4 Using model outputs in root cause analysis 75
4.5 Conclusions 76

5 Layout-based test coverage verification for high-reliability devices 79
5.1 Introduction 79
5.2 Physical test coverage 81
5.2.1 Critical area (CA) 81
5.2.2 Calculation of physical test coverage to estimate the risk of test escapes 83
5.2.3 Flow of physical test coverage calculation 84
5.3 Comparison of logical and physical test coverages 85
5.4 Discussion 87
5.4.1 CA and length of undetected wires 87
5.4.2 Relation between physical test coverage and lengths of undetected wires 88
5.4.3 Relation between length and critical area of undetected wires 89
5.4.4 Case studies of physical test coverage 92

5.5 Conclusions 94

6 Conclusions **95**

References **99**

Acknowledgements **119**

List of Figures

1 Major elements in LSI manufacturing related to failures. 27
2 Layouts of an LSI circuit and defects in a Si substrate. . 29
3 Conceptual diagram of CNN model for classifying the LSI layout segments. 32
4 3- and 15-layer CNN models used in the experiments. . 34
5 Proposed scheme of CNN-based layout segment classification for root cause analyses of LSI failures. 35
6 Training dataset preparation in the learning process of the CNN models. 37
7 Illustration of four layout segments nearest a defect position. 39
8 Confution matrices of 3- and 15-layer models trained with Gate-layer images. 41
9 Locations of segments judged as risk by 3- and 15-layer CNN models 42
10 Histograms of layout segments classified by 15-layer model trained with M2-layer images. 43
11 Confusion matrices of 15-layer models trained with M2-layer images in different epochs. 44
12 Confusion matrix of a 15-layer model trained with M2-layer images for 100 epochs using cross-validation. . . . 46
13 Locations of top 100 layout segments judged as risk segments in the LSI chip. 47
14 Contour plots of a 15-layer model outputs in an LSI chip layout. 48
15 Examples of layout segments used in the classification and a criterion to judge the extraction of defect positions. 49
16 Limited hit rates of top 50 segments extracted by 15-layer models trained with M2, Gate and AA-layer images. 51
17 Proposed scheme of CNN classification for failure analysis. 57

18 Layout segment and the saliency maps of the CNN models trained on the layout images of different size. 62
19 Saliency map obtained based on the input of a layout image of 20 μm. 64
20 Saliency maps of the CNN models trained on 20 μm images. 65
21 Visualization of the layout feature extracted by the CNN model in a virtual layout. 66
22 Dimensions of layout patterns of L- and T- shape, and spaces in the layout of STD cells. 67
23 Classification performance of the CNN models trained on images with different sizes. 70
24 Classification performance of the CNN models trained on datasets of different volumes. 73
25 Contour plots of the outputs of CNN models trained on images of different sizes. 74
26 Contour plots of the outputs of the CNN models trained on datasets of different volumns. 75
27 Critical areas of short and open defects in a wiring layer of a logic circuit. 81
28 Flowchart of physical test coverage calculation using critical area analysis 86
29 Physical and logical test coverages of LSI devices (A to M). 87
30 Wire-length distribution in an LSI logic circuit. 88
31 Comparison of average lengths of undetected wires and physical test coverages. 89
32 Relation between lengths and critical areas of undetected wires. 90
33 Comparison of θs between two device products. 91
34 Examples of CA generation for undetected wires in a logic circuit. 92
35 Test quality improvement using physical test coverage. . 93

36 Test quality improvement using physical test coverage. . 94

List of Tables

1 Datasets used for CNN Model Training 39

2 Summary of Cross-validation of Layout Segment Classification using the 15-layer CNN model 45

3 Numbers of layout images used in the training and testing datasets for CNN models. 60

4 Pixel sizes and pattern dimensions. 68

5 Results of judging the defect position detection and test-segment classifications using the CNN models. 69

6 Results of analyzing defect positions' detection and test segment classification. 72

1 Introduction

1.1 Background

LSI is important in electronic devices. With the rapid advancements in IT digital technology, LSI is being increasingly employed in several electronic applications. Here, we briefly review the various types of LSIs presently used worldwide. In 2018, the ownership of mobile devices in Japanese households exceeded 95%. A smartphone, which is the most common mobile device and is possessed by 80% of the households, comprises approximately 20 LSI circuits. Because of the advancements in LSI integration, there are only a few LSIs on the motherboard of a PC, apart from memory LSIs in the CPU. Although the consolidation of LSIs in LCD TVs has also been advancing, new image processing LSIs for 4 K and 8 K high-definition images have been introduced to the market. Such devices usually employ a system-on-a-chip (SoC), which is a high-performance LSI that integrates several different functional blocks in a single package to perform complex computational processes. In addition, microcontrollers are LSIs with functions to achieve specific operations and are often installed in most home appliances. SoCs and microcontrollers are used in many parts of automobiles. Automotive LSIs are being developed to realize new functions such as automatic driving, and their reliability must be ensured to safeguard passengers and people around the vehicle. Meanwhile, the use of information and communication systems via the Internet is expanding every year, and the society is gradually transforming into a digital society. Digital authentication, such as in the form of self-identification cards and cashless decision making, has become inevitable

in our daily life. The authentication of these IC cards and mobile terminals uses LSI. Thus, LSI plays a leading role in the social infrastructure of a smart society. The fundamental role of LSIs in protecting authentication, personal information, and security in these new lifestyles is crucial.

As described above, the current LSIs widely used in society must have a high reliability [1]. In general, the reliability of LSIs is quantitatively indicated by the occurrence rate of failures during long endurance tests at the start of mass production. However, there is a non-zero risk of failure after the product has been distributed because of defects in the manufacturing process or unobserved issues in design. LSI manufacturers have need to demonstrate the reliability of their products; thus, they may have to maintain records without any defective products to adhere to a stringent requirement called “ zero-defect requests ” from users (customers and society).

Multiple tests are conducted to examine the operation and performance of the LSI at the time of manufacture and upon acceptance of the product and thus guarantee the reliability of LSIs. These tests include functional tests to verify whether the logic circuit operates according to the designed specifications, scan tests to check for open and short failures in the wiring, and quiescent power supply current (IDDQ) tests to check for leakage. These tests are performed in a normal environment as well as a harsh environment with a high temperature and humidity. Reliability is also examined through burn-in testing, in which durability (degradation) is verified through temperature cycling under high-temperature and high-pressure conditions, and destructive testing, in which excessive voltage is applied. In the burn-in process, initial defects (defects detected in the pre-shipment stage) are screened

to prevent the shipping of defective products.

The information on defects detected in the test process is used to identify good products as well as to determine the normality and abnormality of the process. If abnormal processes are determined, the defective products are fed back to the manufacturing process to ensure stable production. Defective LSI products can be selected in the test process because of various reasons, but most of these causes can be traced back to the manufacturing process. Among them, an important aspect to consider is whether the process conditions set at the time of development are optimal. If the margin of variation for process conditions (process window) can be set wide and the actual manufacturing conditions are in the center of that window, the process can be considered optimal. However, if the window is narrow and the actual conditions applied are closer to the edge of that window, the specifications of the product chips may most likely deviate from the required specifications (acceptance threshold). Owing to environmental issues during the process, foreign particles and defects that do not normally occur on a silicon wafer may occur. Owing to these problems, the structure (dimensions and shape) of the device formed on the wafer surface changes, and normal functions cannot be obtained, which results in defects.

There may be issues in the design of the circuit, rather than in the manufacturing process. Even if the design rules are followed, the process window may become narrower owing to the shape and dimensions of the circuit layout. With advancements in LSIs, design dimensions (wiring and transistor gate widths and spaces) have decreased and process margins have become narrower for integration. Meanwhile, lithography and other manufacturing

processes have been developed to achieve process margins that can withstand manufacturing costs. The areas of the layout that have narrow windows in the lithography process are called hot spots. The hot spots should be removed at the design stage because they may cause defects on silicon. Hot spots can be identified with an optical simulation. Another issue that is relevant to layouts is related to the chemical mechanical polishing (CMP) process. In CMP, the surface of stepped interconnections and insulating layers after forming a film is polished using a special chemical solution and a polishing slurry such as silica. The polishing speed is affected by the density, size, and shape of the layout pattern. Simulation studies have been conducted to predict the polishing results, but accurate results have not yet established owing to several factors that must be taken into account. It can also be assumed that the size, shape, and configuration of the layout pattern affect the mechanical stresses caused by the pattern at different locations on the chip. If different materials are stacked on top of each other, a stress is generated owing to different thermal expansion coefficients of the materials. Localized stress may accumulate and lead to the destruction of the device structure and materials.

Design for manufacturing (DFM) refers to the steps taken to simplify the manufacturing process by considering the impact of the abovementioned layout designs on the process. DFM considers hot spots, CMP, and critical area analysis (CAA). The DFM is described in detail in Chapter 2.

Finally, the most important factor in LSI development is the yield, which can be calculated from the results obtained in the test process. The yield is simply expressed as the ratio of the number of LSI chips that passed the

test to the total number of LSI chips on a silicon wafer in the manufacturing process. The complement of this yield is the defect rate, which represents the number of defective chips detected through the test. It is not feasible to achieve 100% yield in the case of microcontrollers and SoCs used in consumer and automotive products because there are thousands of steps involved in the manufacturing process. A few defective products may not be detected during testing. Predicting the occurrence of such defects, that is, the yield rate, is important for examining the production plan of LSI products and to further enhance the reliability of LSI products. Furthermore, predicting the failure factors at an early stage, such as the development stage, is important to prevent the defects from passing the screening test (which are called test-escapes).

1.2 Objectives of this work

This study focuses on failure caused by the LSI layout, proposes a method for analyzing the layout of LSI, and evaluates a method for quantifying the risk of test escape due to layout-related failure. In this thesis, the research results are summarized based on the following: (1) techniques to determine the LSI layout induced defects, (2) techniques to identify the cause of the defects, and (3) techniques to quantify the risk of test escapes based on layout analysis.

In certain cases, such as logic function tests, it is feasible to obtain information that identifies the related wire nets for phenomena that are defective in the test. In some cases, the information obtained from the test is used

for identifying the location of the net of defects or even the location of the circuit responsible for the defect, and in many cases, the location of the defect cannot be identified. The information from this test is combined with other related information to determine the causes of the failure. Then, the estimated points on the LSI chip that are determined to be defective are observed and analyzed to investigate the situation in detail. The actual observation and analysis of the estimated defective locations is called physical analysis or physical failure analysis (PFA). The primary objective of this work is to propose a scheme to analyze the circuit layout of a defective area as an information to estimate the location of the defective area on the chip, which will help improve the physical analysis.

Physical analysis is used to estimate the cause and mechanism of defect formation in defective products by observing the position of the occurrence of defective products. Furthermore, the ultimate goal of this research is to address the causes of defects in cooperation with the manufacturing and design departments to eliminate defects. However, it is difficult to identify the defective part itself, and unknown (unexpected) defects are rarely detected; hence, it is sometimes difficult to analyze and estimate the cause of defects. However, it is desirable to be able to identify and address the causes of defects at an early stage when they are detected. Therefore, the second objective of this work is to propose a method of extracting information to estimate the cause of defects, especially relevant to LSI layout, based on the information of a small number of defects.

The third objective of this work is to propose a method to quantitatively understand the risk of test escape by analyzing the layout. As mentioned

above, the test escape is a serious problem that reduces the reliability of LSI products (and the manufacturer itself). The critical area (CA) is a concept used to describe the relationship between the LSI layout and the risk of failure. This concept simplifies the cause of defects and assumes short-circuit defects between wirings and between open defects that cause wire breakage. The critical area analysis (CAA) is explained in detail in Chapter 2. The CAs in the layout can be estimated and the risk for each shape, configuration, and size of the layout pattern is quantified. To ensure the reliability of LSIs, all faults must be detected by testing. Nevertheless, the wiring cannot be completely tested owing to cost and efficiency. In addition, the degree of reliability ensured by testing is traditionally assessed by test coverage. However, traditional test coverage is not reliable to discuss the magnitude of the actual risk of failure. In this work, a method is proposed for calculating test coverage using CA to express the probability that a defect will slip through the test by linking the risk of defect occurrence to the test. Using this new index (physical coverage), we can efficiently generate test patterns and reduce the risk of test escape by increasing the effective test coverage, thus ensuring the reliability of the LSI products.

1.3 Structure of this dissertation

The remainder of this dissertation is organized as follows:

In Chapter 2, the technologies related to the analysis of LSI defects are reviewed. First, in section 2-1, the analysis methods for LSI defects are summarized. The purpose of physical analysis is to identify, observe, and analyze

the points that cause defects, and each technique is discussed in detail. Section 2-2 describes the yield models and calculation algorithms for predicting the failure rate and yield. The CAA calculation method for analyzing the layout and predicting the yield is explained. Section 2-3 summarizes the issues on DFM from the design stage to improve yields and reliability. Circuit redundancy, optical proximity correction (OPC), hot spot countermeasures for lithography process-margin expansion, CMP, and CAA analysis are discussed. In Section 24, an overview of machine learning and AI technologies is provided.

In Chapter 3, we discuss a method to estimate the defective locations by analyzing the layout of a certain LSI product. The defects were identified by PFA, focusing on the defects that were presumed to be caused by the layout of the product. Here, the image of the circuit layout is cut into squares of a specific size (called “segment”) and used as the data for analysis. We propose a method to classify each segment with the computational model generated by learning the image data with a convolutional neural network (CNN), which is a type of neural network (NN), and to identify the layout that has a risk of defect occurrence. Models with 3 and 15 CNN computational layers were generated to compare the classification performance of the layout images. This study verifies whether the computational model learned by cross-validation can classify layout images near actual defects as having a risk of failure. A prototype of a hazard map to identify the defective locations in the layout of an entire LSI product chip is created using the output values of the calculation model.

In Chapter 4, the feature extraction of the layout images, which is the

basis of the classification performance of the circuit layout images of LSI using the CNN model in Chapter 3, is mainly explored. It is shown that the extracted features of layout images are affected by the resolution of the image data. The layout of feature points extracted from high-resolution images is investigated in detail, and the features of the layouts of defective points are shown. In addition, the amount of data to train the models is increased to compensate for the degradation of classification performance of high-resolution images owing to the limitation of the input size of the model.

Chapter 5 presents a method to quantify the risk of LSI test escape by applying CAA analysis using a layout. Test coverage using CAA is referred to as physical coverage. The details of CAA and physical coverage calculation are described in this section. The physical coverage analysis focuses on wires that are undetected in tests. The risk of test escape is estimated using the CA of undetected wires. Investigating the CA of each net and applying the test pattern to particularly long lines reduces the CA of undetected wires, and test escape is effectively reduced. The conventional test coverage and physical coverage for multiple LSI products were compared. Examples of test pattern development schemes using physical coverage are estimated to effectively reduce the rate of test escapes.

Finally, Chapter 6 concludes the thesis and summarized the work presented.

2 Preliminaries

2.1 LSI failure analysis

An LSI failure is a condition in which an LSI product cannot be operated as designed. Most failures are caused by defects in manufacturing. The failure to form the expected normal circuit pattern on the silicon substrate due to foreign particles in the manufacturing equipment, abnormalities in the process, or fluctuations in manufacturing conditions can cause current leakage in the wiring or gates, as well as signal interruptions due to wire breakage. In certain cases, the expected performance cannot be achieved because the manufacturing process window is extremely narrow. The design (layout) of the circuits formed on the silicon substrate as well as the manufacturing process is related to the margins. For example, the layout of hot spots in the lithography process is complicated for LSI developers who follow the design rules and have a small margin for error during manufacturing. It is extremely important to identify the location of the defect in the chip, observe the state of the location, and determine the reason for the failure; thus, measures must be taken to prevent similar failures from occurring in the future. The method is described in the following section.

An LSI has a logic circuit that is uniquely designed according to the specifications of each product. Logic circuits are made up of wires and tested for defects through functional tests based on the functions required for the product. First, the relationship between the test pattern and the defective behavior (e.g., in a functional test) is analyzed using special software prior to physical analysis, and then the defective locations are narrowed down. This

is called software analysis [2]. Soft analysis can narrow down the wiring (net) that causes the defective operation. However, subsequent physical analysis is used to narrow down the candidate wires, and then the point of defect (anomaly), which is the true cause of the defect, is identified in the wires.

The main analysis methods used for the identification of defects are emission analysis, optical beam induced resistance change method (OBIRCH), and electron beam irradiation analysis (EBAC).

Luminescence analysis is mainly used to monitor the current abnormality and estimate the location of the failure [3, 4, 5, 6]. A high-sensitivity camera detects light generated when hot carriers recombine, which is caused by gate leakage and localized electric field concentration in the insulator film when a voltage is applied to an LSI circuit. Defective areas were identified by superimposing the luminescent areas and the layout pattern. However, light emission may occur in locations other than the abnormality. When a wiring short occurs, a large amount of light may be emitted from multiple transistors connected to the shorted wires. OBIRCH is a technique for identifying defective locations by monitoring the change in current due to the change in resistance of the wiring heated by laser beam irradiation [7, 8, 9]. If the defective part is composed of a material with a different thermal conductivity than that of the normal part, the laser light is scanned along the wiring while a constant voltage is applied to the circuit, and the current change is continuously measured. When the laser light scans the defective area, a different change in the current value from the previously measured value can be observed, and anomalies can be detected. The near-infrared light ($\lambda = 1.3$ μm) used in OBIRCH penetrates the silicon substrate and can be observed

nondestructively from the backside of the substrate. Because many wires are located on the surface of the silicon, it is sometimes difficult to observe them; moreover, since the backside is closer to the substrate surface, analysis from the backside is often more effective. By connecting the OBIRCH to the LSI tester, the area of interest can be set to a potential state suitable for observation and analysis. For example, by switching the potential of the gate with a test pattern (such as the IDDQ pattern), it is easy to find the current abnormality in the wiring to be observed. EBAC [10, 11] and resistive contrast imaging (RCI) are technologies that lead the defective positions to a narrower range than in OBIRCH, which results in semi-destructive analysis (in OBIRCH, the analysis is essentially non-destructive). When the prober terminal is placed in contact with the suspected defective wiring and the electron beam is irradiated onto the device surface, the contrast caused by the difference in the current absorbed is used to form an SEM image. The area where the probe terminals are electrically connected to the wiring in contact with the probe terminals appears brighter than the surrounding area, making it possible to identify open and short defects.

These techniques, which use electron beams, can also monitor the potential status of the circuit. An EB tester applies a test signal to the LSI circuit in the SEM system and monitors the wiring potential chronologically in the SEM image. Abnormal areas can be identified by comparing potential contrast images with normal areas. The wiring to be observed is identified beforehand from the net information obtained by software analysis and the results of non-destructive OBIRCH analysis. By connecting a prober with the wiring in the circuit and applying an electric potential to it,

the wiring disconnection and gate operation characteristics can be directly checked. Similarly, a nanoprober is a device that can monitor electrical characteristics by connecting probe terminals to contacts around the wires and gates in an SEM system.

Another analysis method for locating defects is the detection of abnormal IDDQ [12, 13, 14], which monitors the current leakage (through current) between the power supply and GND in the non-operating state. Because IDDQ testing can detect failures in modes that cannot be detected by logic testing, the IDDQ test is used in conjunction with logic testing to improve the failure detection rate, remove initial defects before shipping, and prevent defective products from test escaping. Using this technique, the heat generation and the current in the defective area can be observed, and the defective location can be identified. It is also possible to identify the defective nets without physical analysis by analyzing the combination of detected patterns using the IDDQ test pattern.

A method for finding the locations of common causes of defects (abnormalities) is described here. As mentioned above, it is important to investigate the location of the cause of failure and thus improve the reliability. However, in many cases, the location of failure cannot be identified by these techniques (e.g., when the wiring that is narrowed down by software calculation is long; or when the analysis is difficult; or when the position determined by emission, OBIRCH, or EB analysis is not a direct defective location). The present study aims to obtain information for identifying defective locations and apply it to these techniques so that defective locations can be accurately identified.

2.2 Prediction of LSI failure rate and yields

LSIs are classified as defective or normal through various tests at the time of manufacture, before shipment, and at the time of customer acceptance. The percentage of LSIs passing the normal tests corresponds to the yield, and its complement is the defect rate. The yield (defect rate) depends mainly on the quality and accuracy of the manufacturing process, but it varies from product to product, even if the same manufacturing process is used. This is because the design (layout) as well as the manufacturing process affect the yield. Thus, it is important to estimate (predict) the yield of each product in advance at the LSI development stage.

The yield prediction model employs a formula to calculate the predicted value of yield based on the factors affecting the yield loss and their influence (weight) on the yield. The following two types of defects are assumed to be the causes of lower yields. The first is margin failure, in which the voltage and resistance values vary depending on the dimensions and shape of the components (e.g., gate pattern) formed on the silicon, resulting in deviations from the required specifications for the LSI's operating speed and noise margin. Margin failure, also called parametric failure, affects both the circuit design and the manufacturing process; therefore, it is necessary to take fundamental measures to identify the causes of failure and correct the design and process. The other defect type is a random failure that occurs randomly and affects logic operations, such as open wiring defects and short-circuit defects. The factors that generally affect the risk of random defects (probability of occurrence) are chip area, pattern size, and defect occurrence level (defect density)

in each process. When the number of defects is relatively large and the chip area is small, the defect occurrence can be approximated to a Poisson distribution. Poisson distribution is widely used for yield prediction because it requires few parameters, and the formula is simple and easy to understand intuitively. As an addendum, it has been confirmed that this Poisson distribution is different from the actual distribution at large chip areas; thus, the negative binominal distribution is closer to reality. This is because the actual distribution of defects is not completely random but rather occurs in a pattern in which defects are in close proximity to each other (clustering). The yield model is described in detail in the following references [15, 16, 17]. Yield prediction is also important in manufacturing. In many cases, yield enhancement activities are carried out simultaneously in multiple manufacturing processes, and the concept of limited yield [18], which considers the yield impact of each process, is useful for yield prediction. The yield of the final product is expressed as the product of the yields of the individual processes. Since yield improvement activities in manufacturing are performed at each process level, estimating the yield of each process is effective. By focusing on the process with high yield impact first and repeating the yield improvement cycle within a short period of time, the yield of the product can be efficiently improved. The simplest yield model assumes that electrical defects occur randomly (Monte Carlo simulation), and the yield is determined from the area and defect density as parameters for calculation. However, in reality, the yield is related to the complexity of the process, congestion (density) of the layout pattern, and the dimensions (from the following references). The concept of critical area analysis (CAA), which quantifies the

congestion of this layout pattern, is described as follows. Two models of defects are assumed for the occurrence of defects in the layout pattern of the LSI wiring layer: one is an open defect that causes an open failure, in which a defect occurs on the wiring in such a way that the formation of the wiring is inhibited; the other is a short defect that exists between the wirings and causes a short-circuit failure between the wires. The locations where these defects occur are not completely randomly selected but are related to the shape and dimensions of the layout pattern and defect size. Open defects are more likely to occur at narrower interconnect widths. Short-circuit defects are more likely to occur between narrower spaces between wires. The larger the size of the defects, the easier it is for both open and short-circuit defects to occur. To express these relationships, the CA is defined as the place or area where open and short defects occur when a defect of a set size is placed on the layout (e.g., wiring layer layout). The CA is used to quantitatively indicate the risk of defects in the product and the wiring layer and is therefore used to predict the yield rate. Using the CA in the Poisson model, the following yield prediction equation can be formulated.

$$Y = \exp(-D_0 \cdot A_c) \tag{1}$$

where Ac is the critical area, and D0 is the dedicated defect density value for the critical area. The differences in yield due to differences in product design, process, and interconnection structure for various uses of CMOS can be efficiently estimated using CA [19]. However, the yield estimated by CAA cannot ignore the differences in processes [20]. When Al is replaced with Cu

in the wiring process, it is not possible to predict the yield of one product from the yield of the other product based on the CA alone. Regarding the modeling of defects for CAA, the modeling of the shape of defects when assuming defect generation in CAA is not yet established; this is because the defects are not circular or square-shaped, but their shapes are indefinite [21, 22]. It is reasonable to calculate CA according to the shape of the defect; however, owing to the problem of computation load, a rectangular shape is often considered. In addition, defects have a size distribution. The following studies were conducted to estimate the size distribution of defects [23, 24, 25]. It is possible to estimate the incidence of open and short defects (failure rate) with a pattern for evaluation called TEG, which uses two long wires placed in close proximity. By preparing test patterns with different wiring widths and spaces, the size of the defective part can be estimated from the pattern dimensions. In practice, it is effective to check the defective area with SEM to obtain information such as size, shape, and the material used (estimated) to accurately estimate the size distribution of defects and to estimate the causes of defects that occur in manufacturing. CAA can also be applied to the prediction of the risk of degradative defects, and the yield estimation method for soft faults (a defect that is not completely defective but may become defective) is reported in the following literature [26]. The method involves re-reading the yield, assuming that the product of the defect size distribution and the CA relative to the defect size (called the probability of failures (POF) curve) decreases compared to the product obtained using the actual size. This method is also expected to be used to obtain a reliability evaluation index, and trials to quantify the risk of gate leakage degradation

defects have been reported [27]. Chapter 5 of this thesis proposes a method to quantify the risk of defect escaping in LSI testing by a yield prediction model with CAA.

2.3 Design for manufacturing(DFM) and layout analysis

Innovative techniques must be employed to simplify the manufacturing process in the design stage and mass-production of products with fewer defects and higher yields, lower manufacturing costs, and higher quality and reliability. This approach at the design stage is called DFM, and it has been applied in LSI development and manufacturing. The gap between the progress of manufacturing process technology development and the specifications required for LSI product devices must be addressed to realize new functions (changes in product generations). Future design nodes will further increase the amount and scope of DFM applications for products that require high reliability with a narrow margin of variation allowed in manufacturing.

The first DFM (still in implementation) was implemented to optimize the wiring, the design dimensions of the holes, the location of the layout patterns, and the shape of the patterns. Then, the wiring width was reduced to increase the integration of LSIs. For this, an improved pattern resolution was required, and the auxiliary pattern sub-resolutional assist features (SRAF) and OPC related to super-resolution optics were applied to expand the process margin. In OPC, the shapes and dimensions of the design layout patterns were derived to achieve the expected final shape based on optical

simulations[28, 29, 30, 31].

Redundant circuit patterns are often used to revive a product by switching the wiring to the redundant circuit in case of an emergency that renders the product inoperable because of a defect [32]. The presence or absence of redundant circuits has a significant impact on yield, and methods for predicting yield have been studied for such a redundancy system.

Recently, machine learning is expected to be applied to the problem of the computational cost of optical simulation to optimize OPC. Layout patterns are classified based on the similarity of the layouts, and the same OPC is applied to similar layouts. The following literature on algorithms for classification of layouts is available [33, 34, 35, 36, 37, 38].

The layout area with a small process margin in the lithography process is called a hotspot. Hotspots can also be detected by optical simulation, but it is not realistic to search for hotspots in the entire LSI product (or even just the logic circuit part) by simulation when considering the cost (time) of the calculation. Extensive research has been conducted with the expectation of applying machine learning in this field, as described below [39, 40, 41, 42, 43, 44, 45].

In recent years, CMP has been playing an important role in the formation of the stacked structure of LSI [46, 47]. CMP combines chemical and mechanical (physical) actions to flatten the substrate; for example, polishing the wiring layer on which tungsten or copper are deposited, which are used as the wiring materials for LSIs. The surface of the insulating film that fills the space between the wiring layers and between the wiring layer and the substrate surface is polished and flattened. However, the surface to be

polished contains a mixture of different materials formed along the circuit pattern, and it is extremely difficult to obtain a flat surface by polishing these materials at different speeds at the same time. The optimal polishing conditions can be set by considering the polishing speed of each material simultaneously. However, the situation becomes further complicated by the influence of the circuit layout. Because the polishing speed is affected by the pattern's density (occupancy), pattern size, and shape, it is difficult to predict the shape of the pattern after polishing if these variables are taken into account. Machine learning is also expected to optimize the CMP process conditions. CAA is a layout analysis method used to quantify the risk of failure of a layout. CAA is commonly used in the semiconductor industry. Defects in LSI products do not occur in a completely uniform distribution of chips, but the risk of defect occurrence varies depending on the degree of congestion in the layout pattern. The narrower the space between the wires, the higher the risk of short-circuit defects. In addition, the narrower the wiring width, the higher the risk of open defects, and the higher the risk in areas where there are many thin wires. CAA analyzes the wiring width, space, pattern shape, and configuration of the layout pattern and quantifies the risk of short-circuit and open defects as an area called the CA.

The application of AI technology to the field of DFM has been discussed for a long time [48], and it was expected that it could counter the unrealistic cost of obtaining exact solutions through CAD and simulation.

2.4 Machine learning and Artificial intelligence(AI)

Machine learning is a technique that uses machines (computers) to analyze the relationship between input information and the outputs. Unlike methods that use formulas based on physical theories to predict phenomena such as manufacturing processes, machine learning uses variable parameters, which are presumed to affect the occurrence of a phenomenon (output), to construct models that derive the most easily explainable laws according to the input-output relationships based on the actual phenomena. Therefore, events that represent the relationship between input and output (data) are important for prediction accuracy; the greater the number of such data, the more accurate the prediction of events that actually occur. The usefulness of machine learning has expanded in recent years as advances in information technology have made it possible to handle large amounts of data quickly and efficiently.

Machine learning can be divided into two types depending on the method of handling training data, which is information that we possess in advance to train a model. They are unsupervised learning, which does not use training data, and supervised learning, which uses training data. K-means clustering is the main unsupervised learning method. This sets up k random hypothetical cluster centers and assigns the data to be classified to each of them appropriately. The hypothetical cluster center and the center of gravity of the assigned data are determined. The data assignment is modified so that the distance between the data and the center of gravity is the smallest. By repeating this process, the data are finally classified into k clusters according

to the similarity between them. This technique can be applied without using training data, which makes it effective for predicting events for which it is difficult to prepare information in advance.

The training data used in supervised learning are the input data for which the resultant event (output) is known in advance, that is, if the input data are used, the expected resultant event will occur. In supervised learning, a model is trained to predict outcome events using pre-labeled training data. New methods for training models are being studied extensively to achieve a higher prediction accuracy. The basic representative supervised learning methods are the k-neighborhood method and support vector machine (SVM). The k-neighbor method prepares several input data for which the resulting event (output) is known, and k events are selected from the data that are close to the new input data (the Euclidean distance in feature space) to be predicted. The most common events among them are predicted as the output. Although the output of the new input data is not known, this method allows us to determine the group to which the input data are classified in the training data. This type of calculation is called classification. SVM is a calculation method that is mainly used for this classification, where a group of data with distinct categories (classes) is used as training data, and the boundaries (lines) are set such that they have the largest margin when they are divided in the feature space. In other words, a boundary must be determined such that the distance between the boundary plane (line) and the data elements in each category that is closest to it is the largest. The data that are closest to these boundaries are called support vectors. According to the set boundaries, the category to which the data belong is determined.

The k-neighborhood method, SVM, and neural networks can be used not only for classification but also for regression analysis using input and output data of continuous values. One of the methods for calculating classification is neural network (NN). A NN is a computational architecture with a network structure that mimics the function of the neurons and synapses that make up the human brain. It has been widely used in recent years owing to enhancements in the performance of general-purpose computers, lower computation costs, and the performance of distributed processing machines called GPUs. NNs can be used for both classification and regression analysis purposes, both with and without training data, depending on the training method. In particular, its image classification performance has been greatly improved compared to conventional algorithms, and it is expected to be used for face recognition, object detection, character recognition, image classification, and a wider range of applications in the future. Many of these are commonly referred to as artificial intelligence (AI) technologies. The NN computational method is also called deep learning because of the depth of its computational hierarchy (repetition).

Machine learning and AI technologies are also widely used in LSI manufacturing. Recently, machine learning has been used to classify the wafer map, which shows the position of a defective chip in LSI tests (logic operation test and analog value measurement test), and a number of research results have been reported [49, 50, 51, 52, 53, 54, 55, 56]. It is also being studied for the purpose of automatically classifying images (optical and SEM photographs) of defects on LSI chips detected in the inspection process that is carried out during the manufacturing process [57, 58, 59, 60]. It is

also realized by using sensor measurements to monitor the state of manufacturing equipment and processes to detect abnormalities (fault detection and correction (FDC)) that have been overlooked by conventional threshold judgments [61]. The controlled coordination of production planning with AI is also being investigated [62, 63]. These types of production sites using AI technology are called "smart factories". Machine learning is also used to identify the optimal conditions for the manufacturing process [64]. A study reported the optimization of CMP conditions using a technique called reinforcement learning, which uses NN [52]. This method has been used for predicting hotspots in the lithography process where the resolution margin is low [65]. In addition, methods to generate OPCs and SRAFs to expand the lithomargin by utilizing NNs called generative adversarial networks (GANs) are being studied [66, 67]. OPC is the mask pattern deformation based on optical simulations. These methods are expected to provide a solution to the problem of forecasting calculations due to the expansion of the level of integration of LSIs, the increasing complexity of micro-processes applied to advanced products, and the development of multiple processes. In Chapters 3 and 4 of this thesis, the circuit layout images of LSI using CNNs, which have been reported to have exhibit a performance in image classification, and their application to the analysis of defective factors are examined.

3 CNN-based layout segment classification for analysis of LSI layout-induced failures

3.1 Introduction

Defective chips, identified in LSI tests by manufacturers or customers, are used to investigate the cause of failure using PFA to improve production quality[68, 69]. Process conditions, equipment, materials, and the LSI layout are the major factors affecting LSI failure (Figure 1). They are closely related to each other and affect the quality, reliability, and yield of LSI products. For example, random particles from either equipment[70, 71] or human operation and the perturbation of process conditions[72, 73] may change the shape or dimensions of device structures (such as poly-Si gate or metal lines) from their optimal ranges or specifications, leading to defects.

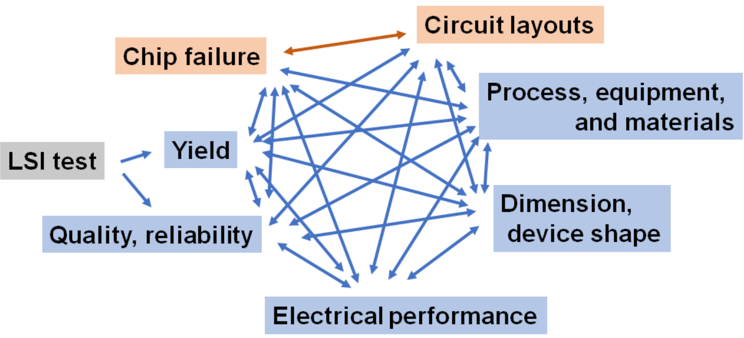


Figure 1: Major elements in LSI manufacturing related to failures.

The layout of the LSI circuits affects failure. In the lithography process, hotspots are a systematic failure related to the LSI layout[74, 75]. In the LSI chips, mechanical stresses induced by the layout can cause defects in Si substrates[76, 77]. To prevent failures, DFM rules are adopted in the lay-

out design to consider manufacturability[78, 79]. To estimate the impact of the LSI layout on failures, optical lithography simulation[80, 81] and stress simulation[82, 83, 84] have been used; however, these techniques have high calculation costs and operational difficulties, particularly for processes in high-end device manufacturing (e.g., multi-patterning). Therefore, simulations are conducted for a local region of an LSI chip layout; however, it is difficult to determine specific layout segments that may cause failure. Figure 2(a) shows the layout of the LSI chip used in this study. Using PFA, some defects were identified in Si substrates under specific standard (STD) cells with the same layout design at different coordinates in different chips. Figures2(b) and 2(c) show the diagrams of the layout of the STD cell and a defect in the Si substrate, respectively. Generally, defects in the Si substrate are formed by defect seeds that grow because of mechanical stresses[85]. Because the defects were identified in cells with the same layout, the failures are attributed to the layout design. The STD cells designed with a layout identical to that of the defective cells were loaded at over 1,000 locations in the LSI chip. Although all these STD cells may cause failure, the risk level is different for each cell in different regions because of differences in the surrounding layout design. The STD cells in which actual defects exist can be considered to have been subjected to relatively high mechanical stresses from the neighboring blocks in the layout. To clarify the reason why the defects were generated in the specific STD cells, it is crucial to distinguish the difference in the risk of failure for each layout design. Thus, layout analysis methods for quantifying the impact of a layout on failures are required.

Machine learning (ML), a statistical technique, is extensively used in ap-

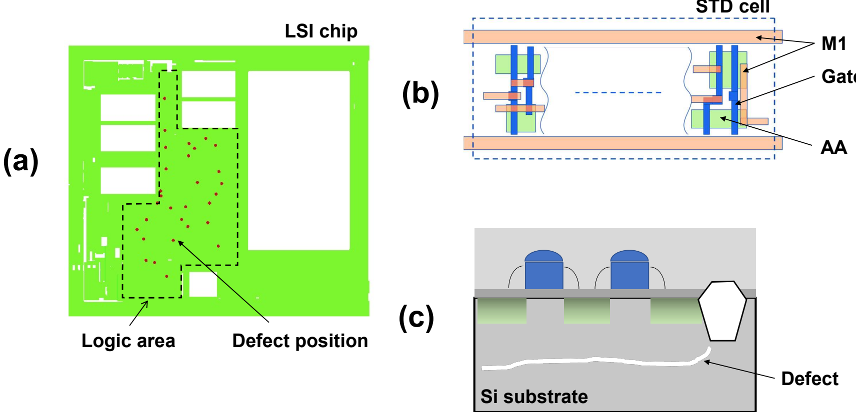


Figure 2: Layouts of an LSI circuit and defects in a Si substrate. (a) LSI chip layout used in this study. The 30 defects identified by PFA in the logic circuit area on different chips were all for the same chip layout. All the defects were in a specific STD cell with the same layout design. (b) Sketch of the layout of STD cell with defects. (c) Cross-sectional view of a defect in a Si substrate below STD cell.

plications such as image recognition[86, 87, 88, 89, 90] and manufacturing process optimization[91, 92]. In the semiconductor sector, various computational techniques that use Big Data have been developed, which can be used for smart factories[93, 94, 95], process optimization[96, 97, 98, 99], testing cost reduction[100, 101, 102, 103], and other applications[104, 105, 106, 107]. Moreover, convolutional neural networks (CNNs) belong to a supervised learning method in which the models are trained with training data, which indicate the target values or features to optimize the data for prediction[108, 109, 110]. CNNs can extract image features without additional data having to be prepared in advance. They can classify images at a

reasonable cost and are thus strong candidates for root cause analyses suitable for the LSI failures[111, 112]. However, in most cases, information on failures in the LSI chips is limited; thus, there is a small amount of training data for a CNN classifier. Furthermore, at the position of each defect detected by PFA, the contribution of a layout segment to failure is different, which reduces the efficiency of model learning when the layout segments are used as the training data. In [113, 114, 115, 116, 117], the CNN has been applied for hotspot detection and excellent classification accuracy has been reported. Because the hotspot detection requires exact recognition of the location of the layout patterns, especially pattern edges, the layout images used in the classification are small and contain a few rectangles. However, the images used in proposed method are larger, containing more rectangles, than those used in hotspot detection. I intend to analyze the impact of the layout in a wide area surrounding the defect position, but large images complicate the classification. To address these challenges, I propose a method for the LSI layout analysis based on a CNN and evaluate its feasibility. The chip layout is divided into rectangular segments and the impact of the layout segments on failures is quantified using the CNN classifier. The final aim of this study, in general, is to clarify the cause of defects generated in the specific STD cells and take countermeasures to avoid failures. Specifically, the target of this study is to properly detect the candidate layouts of unknown defects, which are classified as false positives using CNN with small training datasets. However, training the model with small datasets is a constraint because information of defects is always limited during the early stages of the PFAs. For the layout analysis, I propose a scheme to distinguish the risk

level of each layout segment and determine unsuitable layouts. Using cross-validation, the capability of the CNN models to classify the layout segments is confirmed.

The contributions of this study are as follows.

- Framework for the analysis of LSI layouts using a CNN is proposed., which classifies the layout segments and quantifies the layout's impact on failures.

- Index to quantify the ability of CNN models to classify LSI layout segments is introduced from the viewpoint of failure analysis. Using this index, the performance of the models can be quantitatively compared to optimize the model training method.

The remainder of this chapter is organized as follows. In Section 3.2, the concept of the CNN model is explained. In Section 3.3, the proposed scheme for image classification for analyzing the impact of the LSI layouts on failures is described. In Section 3.4, the experiments and results of the LSI chip image classification are described, and issues related to layout segment classification are discussed from the viewpoint of failure analysis. Finally, in Section 3.5, the conclusion of this work is presented.

3.2 CNN models

Figure 3 shows the conceptual diagram of a CNN model for the classification of images of the LSI layout segments. The intensity of each pixel of the input image $\{x_i\}$ is numerically processed in the convolution (conv) and pooling layers, which are used for the conversion and compaction of image

data, respectively. The filters during image processing are slid over the entire image by a specified number of pixels called a stride. The outputs of image processing are computed using sum-product arithmetic in each node of the fully connected (FC) layers.

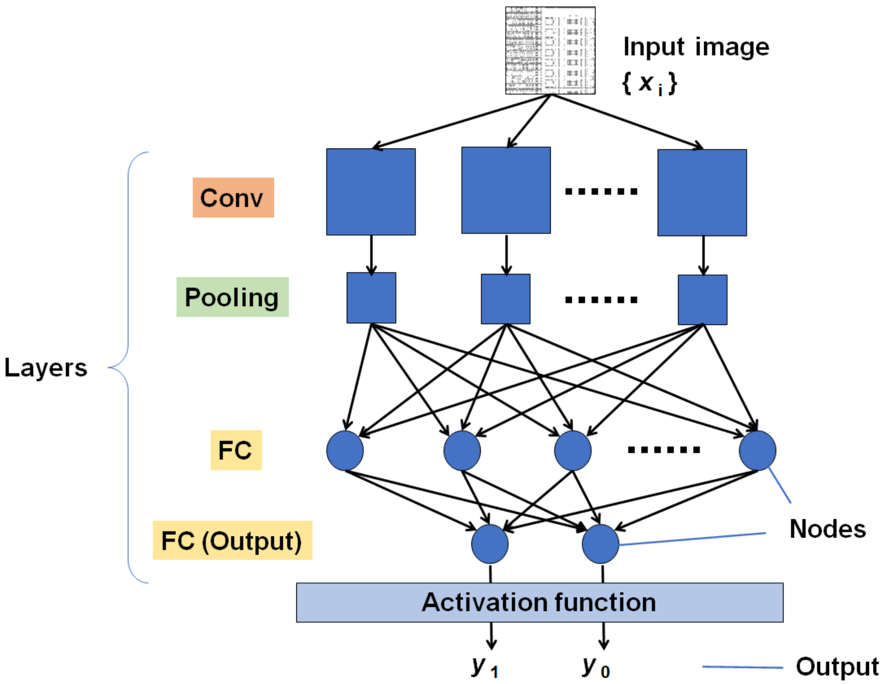


Figure 3: Conceptual diagram of CNN model for classifying the LSI layout segments (Conv: convolutional layer; FC: fully connected layer).

Then, the outputs of the layers are passed via a non-linear activation function in the output layer to produce the model outputs (i.e., the outputs of the entire network). The outputs y_1 and y_0 are numerical values that express the probability of classifying the input images into predefined classes 1 and 0. During the learning process, to obtain high classification accuracy, the models were trained multiple times with the training dataset. Generally, one cycle of training the model with the whole training dataset is called an epoch.

Furthermore, during the learning process, I adjusted the parameters of the models, filters of conv layers, and weights and biases of the sum-product arithmetic in the FC layers to minimize the difference between the model outputs and the target values using backpropagation. Before training, the parameters were initialized with random values to avoid the poor convergence of parameters during the training[118, 119]. Therefore, the parameters of the models were different after each learning process and produced different outputs, although the models were trained with the same training data. To observe variation in the model, an evaluation set comprising the learning process was repeated 100 times in the experiments. In this study, I used two types of CNN models: models that comprise 3 layers and 15 layers. Using the 3-layer model, the model ' s capability to classify the LSI layout segments under various training datasets were evaluated. Then, using the 15-layer model, the classification accuracy was confirmed. Figure 4 shows the configurations of the models. The 3-layer model comprises a conv layer, a pooling layer, and two FC layers. In the conv layer, a filter of 3×3 pixels and 32 channels was used with stride 1; in the pooling layer, a max-pooling of 2×2 pixels was also used. The output sizes of FC layers were as follows: 256 nodes for the first layer and 100 nodes for the second layer. For all the FC layers, sigmoid function was used as the activation function, whereas a softmax function was used for the output layer. For the 15-layer model, I used the transfer learning method of the VGG16 model, which is a part of the Keras library, i.e., the high-level API of TensorFlow[120, 121, 122]. In the VGG16 model ' s configuration, a 16-layer model trained with an image dataset was slightly modified. The original top FC layers were replaced using

a series of layers that included an FC layer with 256 channels, a dropout layer, and another FC layer that outputs two classes. The last three conv and two FC layers were fine-tuned during the learning process.

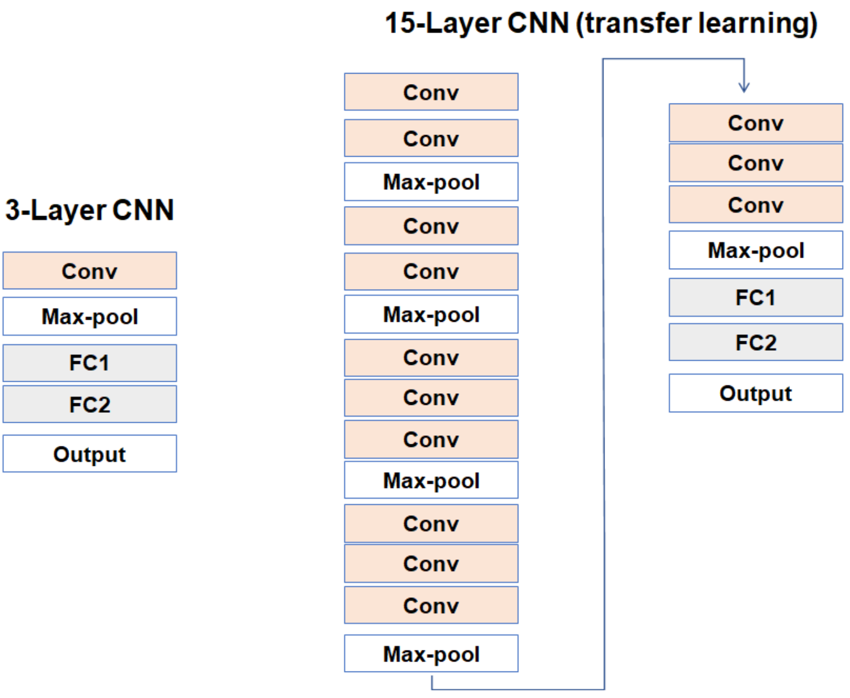


Figure 4: The 3- and 15-layer CNN models used in the experiments (transfer learning was applied with the VGG16 model to generate the 15-layer model).

3.3 Proposed scheme

3.3.1 Overview of layout segment classification

Figure 5 shows the proposed scheme for classifying the LSI layout segments using CNN to analyze the root cause of failures. The image data were prepared from the layout data (e.g., GDS-II) using a layout viewer. The training data comprised two types of image data labeled as “risk segments”

and “non-risk segments.” The risk segments in the training data are regions of the layout that include the coordinates of chip defects identified by the PFA. The non-risk segments in the training data are regions of the layout without defects. Note that the image data used for the evaluation were prepared from the LSI layout data, and all the data used for the evaluation were classified into either the “risk” or “non-risk” class based on the model outputs. Finally, I plotted a contour plot of the model outputs to visualize the layout regions using the outputs of the CNN classification.

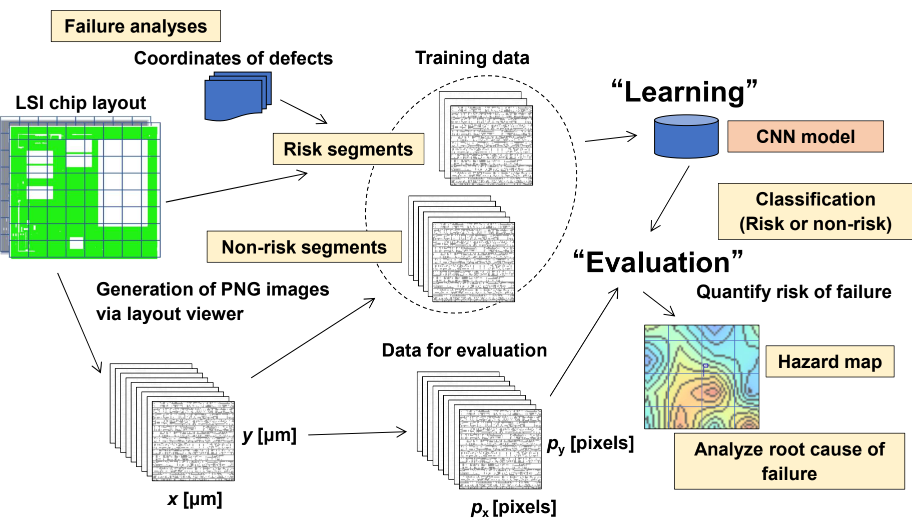


Figure 5: Proposed scheme of CNN-based layout segment classification for root cause analyses of LSI failures.

3.3.2 Input data for classification

To prepare the image data, the images of the LSI layout were cut into segments in a rectangular grid with a size of x [μm] $\times$ y [μm]. The size $x \times y$ should be determined based on the geometries in the LSI design. The images were quantized into a pixel block size of $p_x \times p_y$ [pixels] before the learning and evaluation of the models. Then, all the image data were used for the evaluation. For model training, the images of the logic circuit without defects were used as non-risk segments. In my experiments, the logic circuit of an LSI chip in another design was used. Moreover, the images with a size of $x \times y$ centered on the defect coordinates were separately cut and included in the training data as risk segments. Therefore, the image data used for the evaluation were not the same as those used for model training. The design layers used to generate the layout images should be selected by considering the relationship between the layouts and failures. In this study, the active area (AA), gate (Gate), and second metal (M2) layers were used to make datasets of images because the failures considered in this study were those in the Si substrates.

3.3.3 Learning process (model training)

The training dataset that I used to train the CNN models comprised images of risk segments and non-risk segments. The risk and non-risk segments were selected from the segments that included defect coordinates and those that excluded defect coordinates, respectively. The ratio of the numbers of risk to non-risk segments was ∼1:2 (Figure 6). To increase the number of

training images, images in the training dataset were modified using data augmentation techniques such as image flipping and magnification. Since information on defects is limited in the early stage of failure analyses and the number of images in the training dataset is small, overfitting in the model training must be considered. In my experiments, the training with an appropriate number of epochs showed acceptable classification ability of the 15-layer model (Figs. 10 and 11).

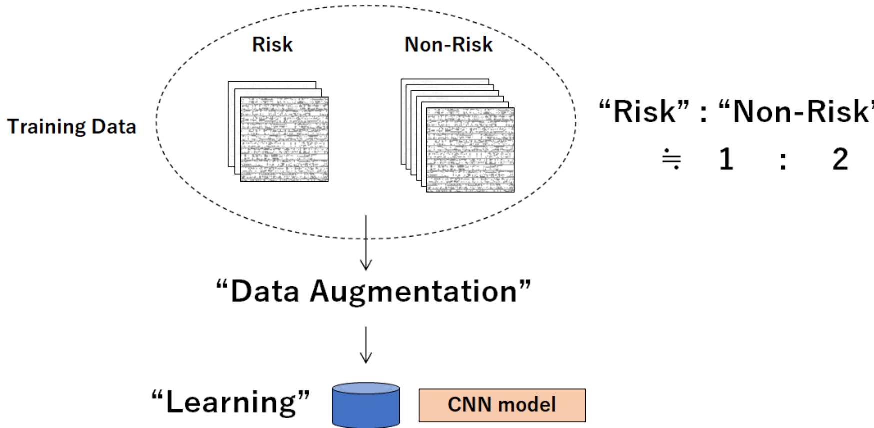


Figure 6: Training dataset preparation in the learning process of the CNN models. The ratio of risk to non-risk segments was 1:2 approx. All the segments in the dataset were processed with data augmentation to increase the size of datasets before the model training.

3.3.4 Outputs of CNN models

Generally, the outputs of CNN models range from 0 to 1, indicating the probability of each input image being classified as either "risk" or "non-risk." In this study, the outputs for the class of risk segments y_{Risk} are defined as the model outputs. When y_{Risk} is ≥ 0.5 (<0.5), the input segments are determined to be risk (non-risk) segments. For analyzing the impact of a

layout on failures, I examined images classified as risk segments by the CNN models.

3.4 Experiments and results

3.4.1 Datasets for model training

Table 1 shows the datasets used for the CNN model training. I used the layouts of the LSI chips based on a 130-nm design node. To prepare RGB images for the input, the chip layout was divided into segments with sizes of $x = 50$ μm and $y = 50$ μm. To prepare the datasets for the model training, the images of risk segments were arbitrarily selected from the 30 images centered on defects as identified by the PFA. The images of non-risk segments were then selected from the images without defects, and the ratio of the numbers of risk to non-risk segment images was $\sim$1:2. The datasets G100-#, which are sets of the images of Gate layer ($p_x = 100$, $p_y = 100$), comprise # = 3, 10, and 30 images of risk segments and 7, 20, and 60 non-risk segments. The datasets ##224-30 with ## = M, A, and G contain the images of M2, AA, and Gate layers ($p_x = 224$, $p_y = 224$).

3.4.2 Definition to judge " Risk " in defect positions

To review confusion matrices, I defined a rule for judging " Risk " in defect positions. As shown in Figure 7, the defect positions were determined as " Risk " when at least one segment among the segments attached to a segment ' s corner nearest the defect coordinates (UL, UR, LL, and LR) was determined as a risk segment by the CNN models. The number of the defect

Table 1: Datasets used for CNN Model Training

Dataset	Input Data		Number of Images In Training Datasets	
	Layer	Pixels p_x, p_y	Risk	Non-Risk
G100-3	Gate	100, 100	3	7
G100-10			10	20
G100-30			30	60
M224-30	M2	224, 224	30	60
A224-30	AA			
G224-30	Gate			

positions determined as "risk" was set in the row "Defect Position = Yes" and the column "CNN judgment = Risk" in the matrices. The total of the row "Defect Position = Yes" was 30, which is the number of defects detected by the PFA. Then, 1312 segments in the row "Defect Position = No," which excludes the four segments (UL, UR, LL, and LR) neighboring each defect position, were classified into risk and non-risk classes in the "CNN judgment" columns. The concept of the risk judgment of defect positions is discussed in Section 3.4.5.

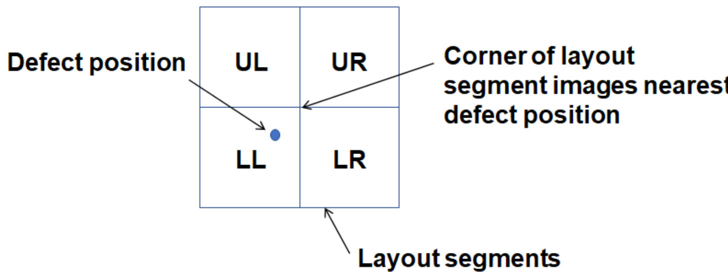


Figure 7: Layout segments UR to LR attached to a segment corner nearest the defect position (the defect position is defined to have been detected when at least one among these four segments is classified as a risk segment).

3.4.3 Comparison of 3- and 15-layer CNN classification

First, I used a CNN model with a simple three-layer configuration to investigate the influence of the number of images in the training datasets. The model was trained for 100 epochs in a learning process with the images of the Gate layer ($p_x = 100$, $p_y = 100$). I then developed an evaluation set that comprised the learning and evaluation of the CNN model 100 times. Figure 8(a) shows the confusion matrices of the 3-layer model classifications. The columns " CNN judgment = Risk " show the numbers of images and defect positions that the models judged as risk. Each segment was judged as a risk when the segment was classified as a risk >50 times in every 100 evaluation trials (threshold of risk judgment = 50%). The 3-layer model trained with a dataset of 10 segments, for 3 risk and 7 non-risk segments (G100-3), did not show enough ability to classify layout segments. The models trained with datasets of G100-10/30 (10 risk and 20 non-risk segments / 30 risk and 60 non-risk segments) determined more segments as "risk." As mentioned in Section 3.3.3, the images in the training datasets were augmented to increase the dataset volume for model training. Moreover, I applied a 15-layer model (a deeper and more practical neural network) for the image classification. The model was trained for 20 epochs with the same datasets used for the 3-layer model. The threshold of risk judgment was adjusted to 95% to compare the classification capability in a similar condition of a true positive detection. As shown by the confusion matrix for the 15-layer model in Figure 8(b), the model trained with 3 risk and 7 non-risk segments (G100-3) could not identify risk segments at all. However, when the model was trained with datasets

G100-10/30, the model showed better classification capability (i.e., less false positives) compared to the 3-layer model. The 15-layer model trained with a dataset of 30 risk and 60 non-risk segments (G100-30) demonstrated the best classification capability. The standard metrics of classification indicates that the accuracy was 92% ((1203 + 28)/(1312 + 30) = 0.92), the false positive rate (FPR) was 8% (109/(1203 + 109) = 0.08), and the false negative rate was 7% (2/(2 + 28) = 0.07).

Training dataset	(a) 3-Layer model				(b) 15-Layer model			
			CNN judgment				CNN judgment	
			Non-Risk	Risk			Non-Risk	Risk
3 risk + 7 non-risk	Defect position	No	1116	196	Defect position	No	1312	0
		Yes	12	18		Yes	30	0
10 risk + 20 non-risk	Defect position	No	1022	290	Defect position	No	1134	178
		Yes	1	29		Yes	1	29
30 risk + 60 non-risk	Defect position	No	1156	156	Defect position	No	1203	109
		Yes	2	28		Yes	2	28

Figure 8: Confusion matrices of (a) 3- and (b) 15-layer models trained with datasets of the Gate-layer images in 100 pixels (G100-3/10/30). The 3- and 15-layer models were trained for 100 and 20 epochs, respectively, in a classification trial, which was repeated 100 times. The thresholds of the CNN judgment were set to 50% for the 3-layer model and 95% for the 15-layer model (test samples were finally judged as "Risk" when they were classified as risk segments more than 50 times or 95 times in 100 trials.) "Defect position = Yes" is a class including four segments nearest each defect's coordinates and "No" is a class of segments excluding those.

Figures 9(a) and 9(b) show the locations of the images classified as risk segments by the 3- and 15-layer CNN models using the same training dataset in the LSI chip layout. Many images classified by the 3-layer model as risk segments are located far from the logic circuit area where STD cells that

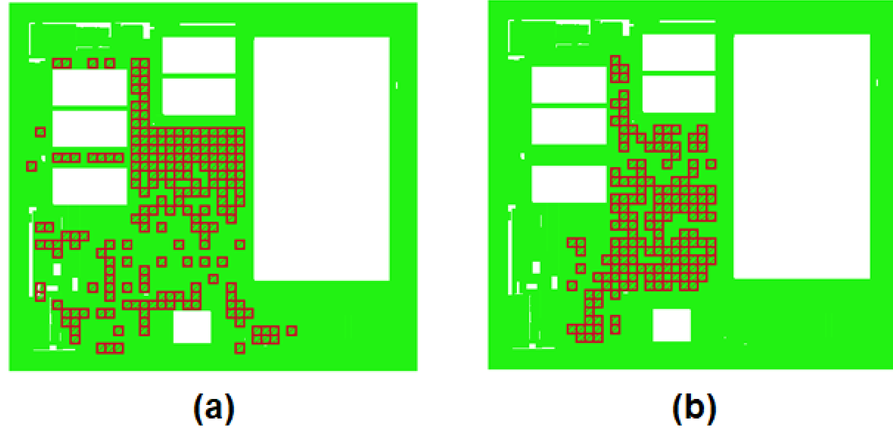


Figure 9: Locations of layout segments in LSI chip classified as risk segments by the (a) 3- and (b) 15-layer CNN models trained with the G100-30 dataset.

may cause failures to exist. However, the images classified by the 15-layer model as the risk segments are concentrated in the logic circuit area. Figures 10 and 11 show the histograms and confusion matrices of the results of classifications using the 15-layer model trained with various numbers of training epochs, respectively. The model was trained with 30 risk and 60 non-risk datasets of M2 layer segments. When the model was trained for 20 epochs (Figure 10(a)), most segments were judged as non-risk (i.e., the segments were classified as risk for <50 times in 100 trials). The segments of defect positions concentrated around the 50-times risk classification. They are the segments with the highest model outputs among the four segments around the defect coordinates. Then, when the model was trained for 100 and 200 epochs, more images were classified as risk segments with true positive rates (TPRs) of 83% (25/30 = 0.83) and 100% (30/30 = 1) and FPRs of 9.6% (126/(1186 + 126) = 0.096) and 9.5% (125/(1187 + 125) = 0.095), respectively. These results show that the 15-layer model trained for 100 and

200 epochs can be used to classify the layout segments. However, as shown in the histogram of the model trained for 200 epochs (Figure 10(c)), some segments were concentrated at 100 on the x-axis, which suggests weak overfitting. Thus, I utilized the 15-layer model trained for 100 epochs to confirm the model ' s performance using cross-validation.

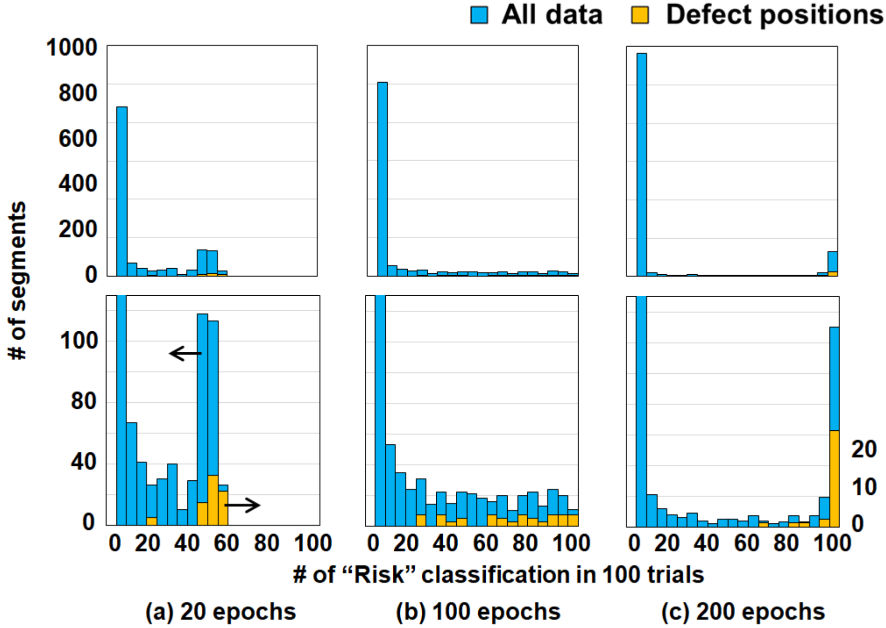


Figure 10: Histograms of the layout segment classification by the 15-layer model trained with training dataset M244-30 for (a) 20, (b) 100, and (c) 200 epochs. Top charts show whole data and bottom charts show enlarged views.

3.4.4 Cross-validation of 15-layer CNN classification

For image classification, cross-validation, a method generally used to evaluate the CNN model[123, 124] are applied and the performance of the 15-layer model is validated. From the dataset of M224-30 which comprises segments ($p_x = 224$, $p_y = 224$) of the M2 layer, 5 risk segments were randomly removed

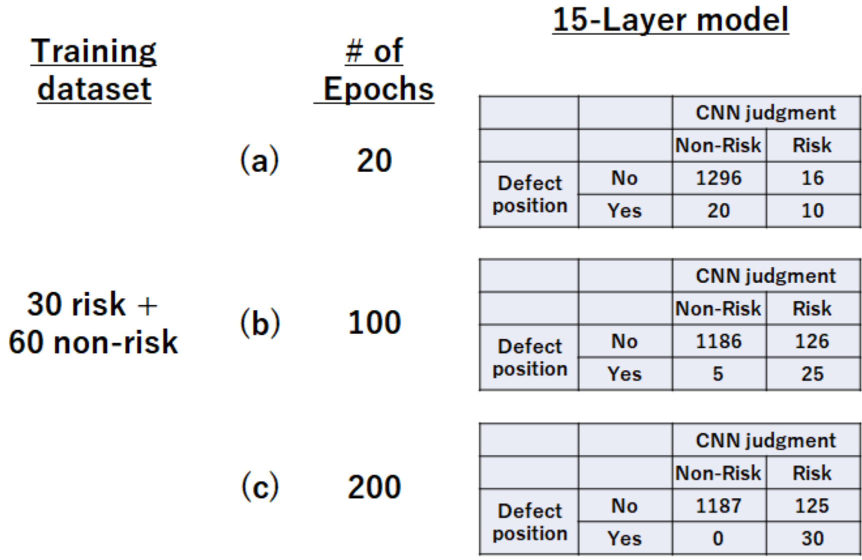


Figure 11: Confusion matrices of the 15-layer model trained with training dataset M244-30 for (a) 20, (b) 100, and (c) 200 epochs.

to prepare training datasets. Thus, the modified dataset consists of 25 risk and 60 non-risk segments. After the model was trained with each training dataset for 100 epochs, 30 defect positions and 1312 segment images without defect positions were classified, and the trial was conducted 300 times. Figure 12 shows the full prediction confusion matrix for the 300 trials. Among the 30 defect positions, 28 defects were determined as risk. The average TPR for the 300 trials was 93%. Furthermore, 101 segments without defect positions were determined as risk segments (FPR = 7.7%). In this work, false positives were not a problem; rather, they are the target that I hope to find. Because, unknown defects might exist in the area other than the area of defects which had been detected in PFA.

Table 2 summarizes the probabilities of judging defect positions by the CNN model trained with datasets that excluded each defect position. The

Table 2: Summary of Cross-validation of Layout Segment Classification using the 15-layer CNN model. [a]Probability of the risk classification of layout images in upper-right (UR), upper-left (UL), lower-left (LL), and lower-right (LR) regions around defect coordinates. The cells are shaded in green according to the probability. [b]R indicates that the defect position was judged by the CNN models as "Risk." [c]Number of evaluation sets, where the models were trained with datasets that excluded images with defects in each row. [d]Probability of judging the input images as risk segments in all the 30 defect positions.

Defect Position (Flag No.)	Probability [a]				Risk Judgment [b]	Number of Evaluation Sets [c]
	UR	UL	LL	LR		
1	0.203	0.814	0.119	0.000	R	59
2	0.936	0.638	0.745	0.745	R	47
3	0.162	0.000	0.000	0.595	R	37
4	0.321	0.375	0.054	0.357	-	56
5	0.500	0.286	0.482	0.179	R	56
6	0.000	0.623	0.472	0.000	R	53
7	0.228	0.614	0.719	0.825	R	57
8	0.275	0.225	0.175	0.375	-	40
9	0.565	0.543	0.957	0.065	R	46
10	1.000	0.060	0.100	0.400	R	50
11	0.857	0.571	0.179	0.768	R	56
12	0.776	0.837	0.735	0.184	R	49
13	0.682	0.568	0.659	0.136	R	44
14	0.581	0.721	0.465	0.837	R	43
15	0.571	0.000	0.265	0.163	R	49
16	0.184	0.921	0.842	0.368	R	38
17	0.200	0.560	0.280	0.540	R	50
18	0.305	0.508	0.017	0.356	R	59
19	0.426	0.459	0.197	0.361	-	61
20	0.686	0.490	0.667	0.373	R	51
21	0.026	0.205	0.744	0.154	R	39
22	0.041	0.061	0.306	0.592	R	49
23	0.784	0.686	0.157	0.667	R	51
24	0.714	0.381	0.024	0.667	R	42
25	0.714	0.000	0.000	0.551	R	49
26	0.286	0.161	0.054	0.268	-	56
27	0.000	0.000	0.288	0.212	-	52
28	0.489	0.404	0.638	0.957	R	47
29	0.491	0.263	0.930	0.175	R	57
30	0.158	0.439	0.123	0.544	R	57
Summary of Probability	-	-	-	-	0.83[d]	-

Training dataset	15-Layer model			
			CNN judgment	
			Non-Risk	Risk
25 risk + 60 non-risk	Defect position	No	1211	101
		Yes	2	28

Figure 12: Confusion matrix of the 15-layer model trained with the M244-30 dataset for 100 epochs using cross-validation. The threshold of the risk judgment is 50% (test samples were judged as risk when they were classified as risk segments in more than 150 times in 300 evaluation trials).

UR to LR are the segments around the corners of segments closest to the defect coordinates (Figure 7). The "R" in the column "Risk Judgment" shows that the layout at the defect position was determined as a risk segment, where at least one of the four images around the defect position was classified as a risk segment. The positions of 25 defects (83% of the 30 defect positions) were classified as risk segments by the model when trained with datasets excluding each defect position. This result indicates that the 15-layer model can predict risk segments with a TPR of >80%. Figure 13 shows locations of the top 100 images determined as risk segments in the cross-validation using the 15-layer CNN model and indicated by small squares. Most squares are in the logic circuit area, and the actual defect positions are indicated by small dots. The positions of most squares determined to be a risk are close to the actual defect positions, depicting that the CNN model did not directly recognize the positions of defects; rather, it selected the neighboring sites of defects.

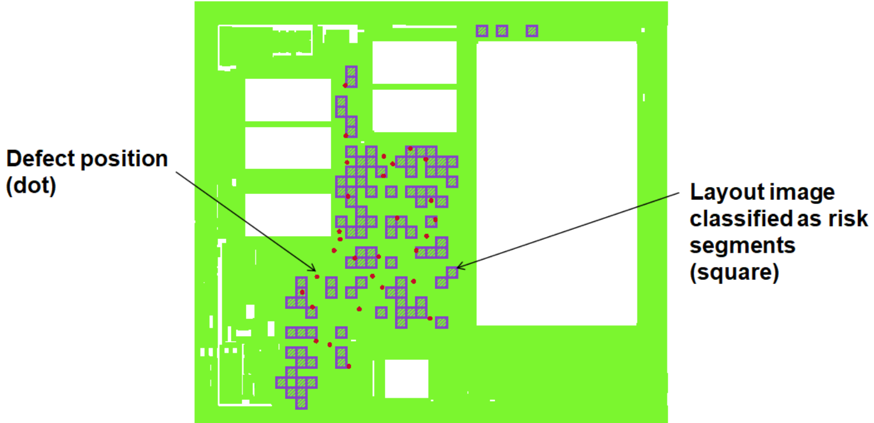


Figure 13: Locations of top 100 layout segments judged as risk segments in the LSI chip (small squares).

3.4.5 Visualization of impact on failures in the LSI layout

E. Visualization of impact on failures in the LSI layout I visualized the regional layout ' s impact on failures as a contour plot (hazard map) of the model outputs across the LSI chip. Because all the layout segments were classified by the model trained with layout segments that included defect positions, the contour plot indicates the regions of the layout with a risk of failure. the contour lines were determined from the slopes between the model outputs of the neighboring image segments. In Figure 14, the defect positions identified by the PFA are indicated by the flags in the contour plot. Most flags are between the peaks and valleys of the contours (i.e., not on the peaks where the risk is locally higher). This trend is consistent with the results in the previous sections, where most images classified by the CNN model as risk segments were neighboring to the defect positions.

Next, from the viewpoint of failure analysis, I present the results of the layout segment classification in this section and introduce an index for quan-

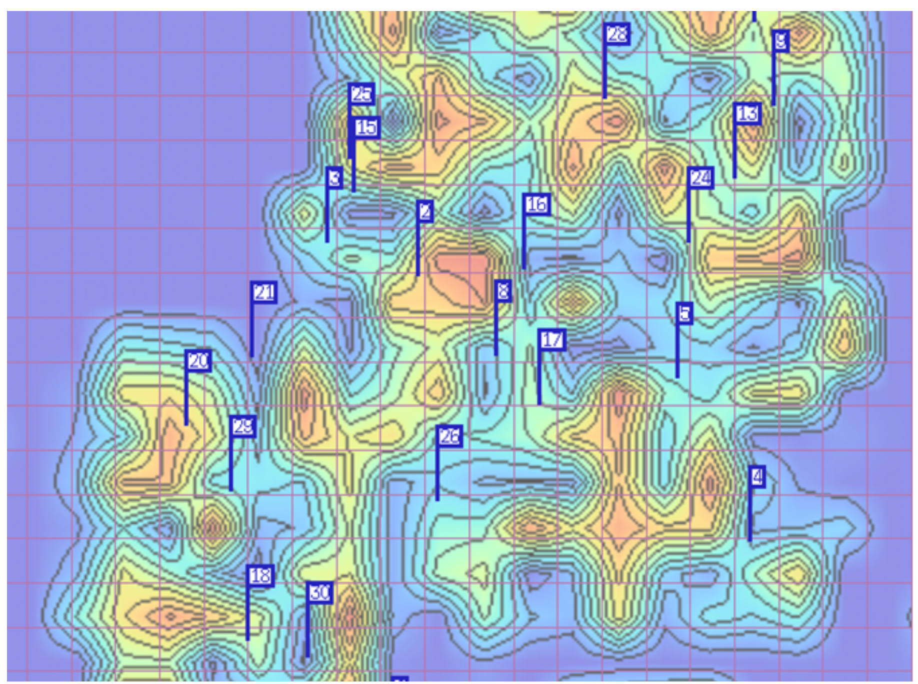
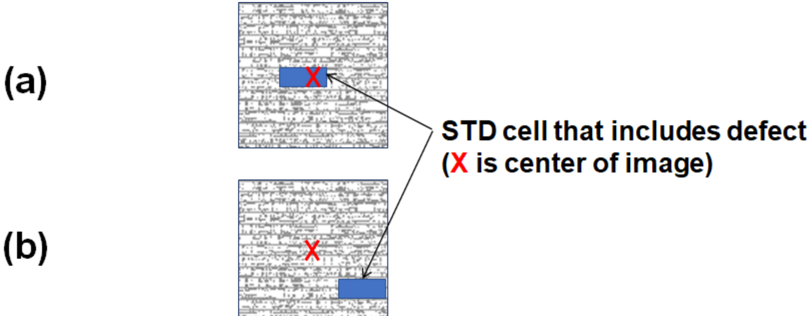


Figure 14: Contour plots of the 15-layer CNN model outputs in part of the LSI chip. The output of each segment was calculated from the average of the cross- validation results. A warmer color indicates higher output (i.e., higher risk of failure). Flags indicate coordinates of defects found by physical failure analysis.

tifying the classification capabilities of the CNN model. For the layout images used as training data labeled as risk segments, the defects are centered in the images, as shown in Figure 15(a). The images used for evaluation were cut in a rectangular grid without considering defect positions; therefore, the positions of defects where the failure occurred are not always in the center of the images. Instead, as shown in Figure 15(b), the defects are located close to the edges of the images. When a part of the layout image with a feature related to failure was far away from a defect, the region with these features might not exist in the images that include the defects; rather, it might exist in neighboring images. Therefore, most of the images that included defects

were determined as non-risk layouts by the 15-layer CNN model.



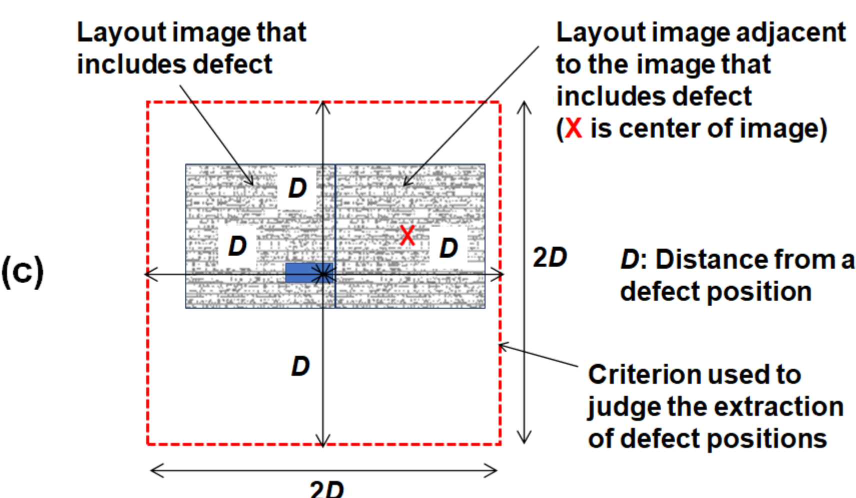


Figure 15: Examples of the layout segments used in the CNN classification and a criterion to judge the extraction of defect positions. (a) A layout segment centered in a defect position for the dataset used in model training. (b) A layout segment cut in a rectangular grid used for the dataset in model evaluation. (c) The layout segments for evaluation and a criterion for judging the extraction of defect position (dotted square). Four grayed segments, of which the centers are included in the dotted square of a criterion, from the center to bottom-right are the candidates to be judged as risk segments.

A general target of classification is a high TPR and a low FPR; however, in this work, the target of the FPR is not zero because the risk of defect generation is observed in the segments labeled as " Defect Position = No. "

To analyze defect candidates, the segments classified as risk segments with higher model outputs should be prioritized. Furthermore, the impact of the layouts on defect generation is related to the distance from the source of the defect generation. Considering these elements, I define a metric that focuses on segments with high model outputs and combines the distance between defect coordinates and layout segments to monitor the classification performances. Thus, to quantify the ability of the model to extract images of risk segments, I define the limited hit rate $R_a(D)$ as follows:

$$R_a(D) = \frac{n_a(D)}{N}, \tag{2}$$

where a is an arbitrary number used to limit the high-ranked images sequentially sorted by the outputs of the CNN model; $n_a(D)$ is the number of defects extracted from the limited number ($=a$) of high-ranked layout images classified as risk segments by the CNN model within distance D from the defect coordinates; and N is the number of all images that include defect positions used for the model training. If a square of size $2D \times 2D$, in which a defect is located at the center, includes at least one of the centers of risk segments which was extracted by CNN, then I say that the defect is successfully detected (Figure 15(c)). Figure 16 shows the limited hit rate $R_{50}(D)$ for the top 50 risk segments determined by the CNN model using the layout images of M2, Gate, and AA layers. The limited hit rate increases with increasing distance D, and the trials with the model trained with the layout images of the M2 layer showed the highest capability for extracting risk segments among trials with the model trained with the three design lay-

ers at the specified distances. At distance $D = 80$ μm, the limited hit rate, which depends on the conditions of model training, was 0.9 when the model was trained with the M2 layer. Furthermore, the limited hit rate is used to compare model capability and optimize the learning conditions for extracting risk segments in the layouts.

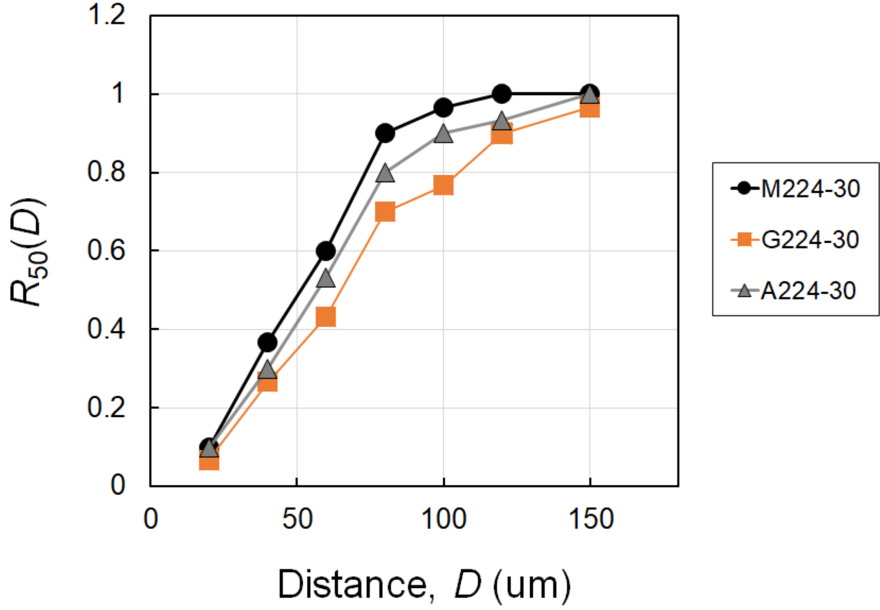


Figure 16: Limited hit rates of top 50 layout segments extracted by the 15-layer CNN model in cross-validation using images of three design layers (M2, Gate, and AA).

3.5 Conclusions

In this study, to determine the impact of a regional layout design on failure, I proposed a scheme for LSI layout segment classification based on CNNs. From the experimental results, the size of the training data influenced the classification performance. A 15-layer model with transfer learning using

90 images (i.e., 30 risk and 60 non-risk images) for 100 and 200 epochs showed the capability of classifying layout images that included real failures. The risk of failure was visualized in a hazard map based on the outputs of the CNN model. Thus, using limited hit rates, the conditions of model training can be compared and optimized. In the next chapter, based on the results of the present investigation, the layout features extracted by the CNN model are explored to clarify the root causes of failures using advanced neural network techniques[125, 126].

4 Layout feature extraction using CNN classification in root cause analysis of LSI defects

4.1 Introduction

ML and AI techniques are widely applied to various fields in semiconductor manufacturing. In particular, ML has been adopted to analyze wafer maps and root causes corresponding to LSI failures in the processes or tools [127, 128, 129]. Wafer maps indicate the positions of particles or defects in silicon wafers generated during production or the positions of failure chips in wafers, detected by LSI tests. Image classification techniques are used to analyze defect images captured during wafer inspection. Implementation of the automatic classification of defect images increases the operational efficiency. Moreover, providing high accuracy of defect identification can facilitate the solving of problems in the early production and save costs associated with losses in process optimization [104, 130]. In addition to the field of semiconductor manufacturing, ML has been applied to the field of DFM. LSI circuits are analyzed in DFM by conducting simulations aiming to reveal the weaknesses in layout designs that affect the quality of manufacturing [131, 132, 78]. “Hotspots,” which are defined as the layouts with a risk to induce defects during the lithography process, can be detected by lithography simulation using circuit layouts. However, conducting such simulations is time-consuming, and the cost of analyzing an entire LSI chip is unfeasible. Thus, ML has been applied to predict hotspots as a method substituting

simulations [111, 110, 133, 134]. As hotspots may be generated due to proximity effects, the positions and/or shapes of layout patterns, such as gates or metal-lines, need to be accurately recognized during image processing to perform defect prediction. Therefore, as model inputs, layout images used in ML must be split into small layout images (segments) including few geometries. In this chapter, I investigate the applicability of an ML scheme to analyze the layouts of LSI circuits to conduct the root cause analysis (RCA) of LSI failures. This study aims to search defects that may be generated in silicon substrates due to the stress of layout patterns in the isolation layer [85, 135]. I use the layout segments including relatively large areas that impact the defect generation instead of employing small segments, as suggested in hotspot detection. Utilizing a convolutional neural network (CNN), a deep learning algorithm that is commonly used in image analysis [136, 137, 138], ILayout segments are classified according to the prediction of the risk of defect generation and the defective layouts causing failures are specified. In general, ML-based methods require a large amount of data to train models. However, in my previous study [139], I demonstrated that a transfer-learning-based CNN model pretrained on large image datasets achieves acceptable classification performance concerning LSI layout images even though the training datasets included only 90 images. However, due to the low resolution of the input images including the large areas, I could not accurately predict the positions of defects using the classification results. I concluded that changing the size or resolution of image segments was one of the key elements affecting the performance of feature extraction in CNN models [140]. The partial layouts extracted using the models as the features of defective layouts could

provide valuable information for understanding the reasons underlying the occurrence of failures. In this chapter, I propose a scheme of image classification in LSI layouts using CNN models to analyze the root causes of failures. Moreover, the visualization of the layout features extracted by the models is realized. Then, layout images of different resolution are utilized to train the CNN models as well as to visualize the features extracted from the layout images.

The main contributions of this study can be summarized as follows.

- The impact of the resolution of image segments used to train CNN models on the resulting layout features outputted by these models is analyzed.

- A partial area within a layout segment can be accurately extracted as a common feature in the layout with the risk of defects using of small image segments of high resolution. Then, the extracted layout features can be used as references to infer the root causes of failures.

- It is proved that increasing the number of layout images in training datasets improves the classification performance of models even if image segments are small. The outputs of reliable models can be used to generate hazard maps with the purpose of supporting the search of unknown defects in physical failure analysis (PFA).

The remainder of this chapter is organized as follows. In Section 4.2, a scheme of image classification including the feature extraction step is proposed and the experimental conditions are outlined. In Section 4.3, the

results of feature extraction are presented and the effect of image resolution is discussed. In Section 4.4, the results of classifying defective layout images are demonstrated and the hazard maps derived based on the model outputs are discussed. Finally, I summarize this chapter in Section 4.5.

4.2 Overview of experiments

4.2.1 Proposed scheme of the layout analysis based on CNN classification

A flowchart of the proposed scheme is presented in Figure 17. Using CNN image classification, I extract the features of defective layouts to conduct RCA and to specify the locations with a risk of defects in an LSI chip to search unknown defects. The images of chip layouts (for example, GDS-II) are cut using a layout viewer into square grids to afford layout image segments (PNG files). Using the layouts of multiple LSIs having different designs is possible. The positions of particular defects that occur in LSIs corresponding to silicon wafers can be identified through PFA. The segments including defect positions and those excluding defect positions are denoted as " risk images " and " non-risk images, " respectively. Defect positions are defined as the coordinates of the defects detected in LSI chips by OBIRCH [68, 69]. The datasets including the risk and non-risk image segments are used to train CNN models. Other segments denoted as " images for tests " in Figure 17 are classified using CNN models into two classes: risk images and non-risk images. The features of risk images are visualized in the form of saliency maps using GradCAM++, a computational technique used to

explain the outcomes of model classification. Saliency maps are utilized to analyze the layouts of root causes of defects. The outputs of CNN models are real numbers from 0 to 1 representing the probability to classify test images for each target class. When the output is close to 1, the corresponding test image is similar to the images of a particular class, such as risk images. The contours of model outputs, referred to as hazard maps, are then utilized in PFA to identify the positions of unknown defects causing failures.

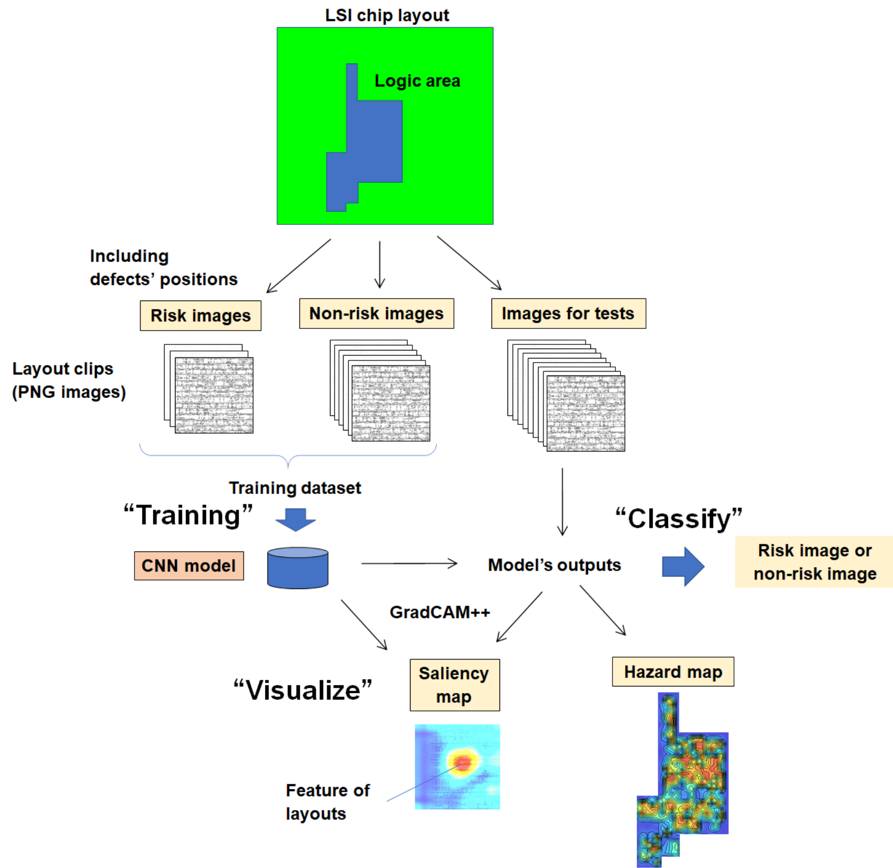


Figure 17: Proposed scheme of CNN classification for failure analysis.

In previous studies, I applied VGG16 transfer learning to develop a CNN model, which demonstrated excellent image classification capability [121,

122]. The top layers of the model were customized for fine-tuning using the specified " flat " layer to flatten multi-dimensional data. The top layer was customed into the sequence of dense layer composed of 256 nodes with a rectified linear unit (ReLu) [141] as the activation function, 50% dropout layer and dense layer of two nodes with softmax function. The parameters, weights, and biases of these layers were optimized by conducting additional trainings. To visualize the features extracted by CNN models, I applied GradCAM++ [142, 143]. This technique affords a saliency map that is a heat map of Lcij, denoting the importance of each spatial location $(i,\ j)$ for a particular class c. The model outputs for class c, Y_c, is calculated using a linear combination of k-th feature maps A_k corresponding to the last convolutional layer and weights for class c w_c^k as follows:

$$Y^c = \sum_k w_k^c \sum_i \sum_j A_{ij}^k. \tag{3}$$

Class-specific saliency map Lcij is calculated as follows:

$$L_{ij}^c = \sum_k w_k^c A_{ij}^k, \tag{4}$$

where weights wck are calculated as

$$w_k^c = \sum_i \sum_j \alpha_{ij}^{kc} \cdot relu(\frac{\partial Y^c}{\partial A_{ij}^k}). \tag{5}$$

Here, α_{ij}^{kc} denotes the coefficients introduced to level off the importance of each spatial location in all k-th feature maps in the last convolutional layer. Finally, L_{ij}^c expresses the visual explanation of the class prediction of CNN

models. Saliency maps indicate the locations of the features with high L_{ij}^{c} in an image, which highly impacts the results of image classification, concerning its correspondence to particular class (risk or non-risk images). When CNN models are trained on the layout images of defects, the map indicates the common features of the layouts exhibiting the risk of defects.

4.2.2 Experimental conditions

The layouts of an active area (AA) layer in a 130 nm design LSI product were used as the input images. The results of LSI tests before shipping detected leakage failures in several chips in a wafer, and PFA identified the defects in the silicon substrate in different locations within several chips. Most of the identified defects were detected within an identical STD cell, while the locations of cells differed from each other; thus, they were considered as layout-induced defects. The datasets used in the experiments are listed in Table 3. Dataset A was used to compare the performance of models with different image sizes. The subset of Dataset A used to train the models (the training set) comprises 30 risk images and 60 non-risk images. The CNN models trained on different datasets were named as A50, A30, and A20, based on the size of image segments: 50, 30, and 20 μm, respectively. In this study, " the image of s μm " denotes the image segment with the width and the height of s [μm]. The risk images in Dataset A were clipped centering at each position of the 30 defects, which were identified by PFA. Dataset B was used for the evaluation of the CNN models trained on the images of 20 μm. I compared the performance estimates of the models denoted as B60B600 based on the number of risk images used in training. The risk images in

Dataset B were clipped at random locations and included defect positions to increase the size of datasets.

Table 3: Numbers of layout images used in the training and testing datasets for CNN models.

Dataset	Image size *s* [um]	Numbers of image clips for training		Model name	Numbers of Test clips
		Risk images	Non-risk images		
A	50	30 Defects centered	60	A50	291
	30			A30	882
	20			A20	2150
B	20	60	2069	B60	2237
		100		B100	
		150		B150	
		300		B300	
		600		B600	

To avoid overfitting in CNN model training, the training datasets were extended by applying data augmentation procedures, such as horizontal and vertical flipping and magnifications. In the previous study [139], the models were not overfitted in this experimental condition based on the training curves and the histograms of the model outputs. Therefore, in all the experiments on each training dataset, the models were trained in 100 epochs same as the previous study.

4.3 Feature extraction using high-resolution images

4.3.1 Visualization of the circuit layout features extracted by the CNN models

I visualized the results of image classification outputted by the CNN models using GradCAM++. In Figure 18, the saliency maps of the two models, A50 and A20, are compared. Figure 18(a) represents a layout segment used to generate saliency maps. Within this segment, a position of a defect identified by PFA is located at the center (marked by x). The width and height of the segment are 50 μm, and the square in the middle of the segment indicates the boundary of the image of 20 μm. Figures 18(b) and 18(c) represent the saliency maps corresponding to the image segments of 50 μm and 20 μm, respectively. In these maps, the locations marked in yellow or orange (red) are the features explaining the prediction of the outcomes of risk image's classification, by which model outputs are strongly affected. On the contrary, dark blue (purple) area the impacts are relatively small. And sky-blue area is in the middle. Figure 18(b) illustrates that the features extracted by the model A50 are mainly located at the low pattern-density area of the image. The area in which several relatively large patterns are condensed is also highlighted. On the contrary, as shown in Figure 18(c), the feature of the layout is extracted by the model A20 at the center of the image in which an actual defect has been identified. These results indicated that the features in the layout images varied from the area of pattern density (macro) to the individual objects (micro) depending on the size of input images. Therefore, according to the aim of this study, selecting appropriate image sizes is cru-

cial. To accurately capture the features of layouts with a risk of defects, the images of 20 μm were deemed preferable in this study.

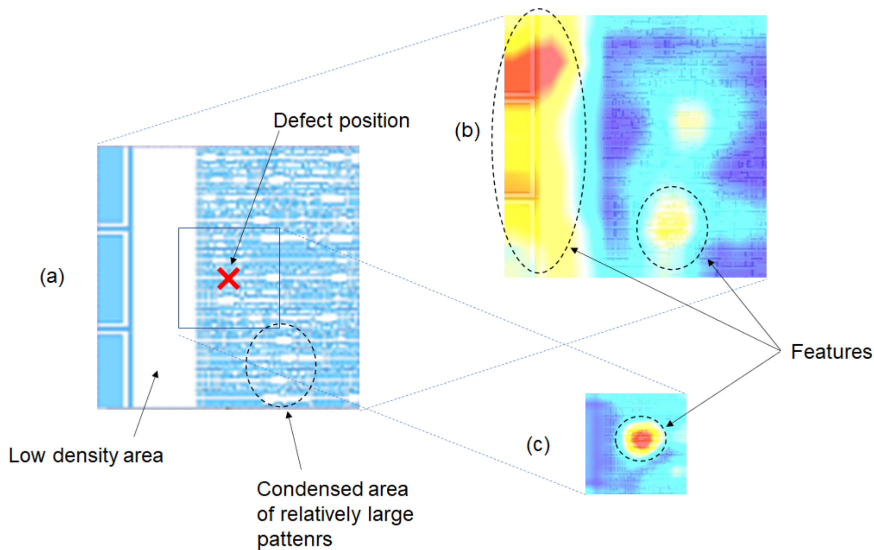


Figure 18: Layout segment and the saliency maps of the CNN models trained on the layout images of different size. (a) Input image segment centered on a defect position; (b), (c) The saliency maps of the images of 50 μm and 20 μm, respectively. In these maps, the locations marked in yellow or orange (red) are the features explaining the prediction of the outcomes of risk image classification. The CNN models trained on the images of 20 μm extracted the features in the center of the input images, which was close to the position at which a defect was detected by PFA.

Figure 19 represents the saliency map corresponding to an image of 20 μm. By carefully analyzing the area of the feature extracted near the center of this image, I noted that specific circuit patterns (polygons) were located at the feature location. These patterns (L- and T-shape, as marked in the figure) were common in most saliency maps obtained from the image segments including defect positions. I considered that the area including these patterns in the layout was related to failures. The size of the datasets used

in model training also impacted the results of image features extraction. The CNN models B600 and B150 trained on Dataset B, comprising 600 and 150 risk images, respectively, were applied to obtain saliency maps. The image features observed in these maps were investigated using three input images (#1-#3) centered on the defect positions (Figure 20). The defect positions at the center of images were almost completely extracted by each model. As shown in Figure 20(b), the sky-blue areas were widely spread in the segments. However, the sky-blue areas represented in Figure 20a are small. In addition, the location of the feature in segment #1 in Figure 20(b) was slightly shifted from the center of the segment. However, the features in all three segments represented in Figure 20(a) were constantly centered in the segments. These results indicate that with an increase in the number of risk images in the datasets, the image features were more accurately extracted, and the positions of defects were precisely captured.

4.3.2 Extracting layout features using a virtual layout image

I attempted to investigate the details of feature extraction in a layout segment. As mentioned in the previous section, the areas including the L- and T-shape patterns were extracted as the common feature of the layouts with a risk of failures. The various areas including such patterns in a risk image segment were cut and pasted together to generate a virtual layout segment, as shown in Figure 21(a). The L- and T- patterns cut with small surroundings were placed at the positions of red rectangles in an empty segment. Relatively large areas including both patterns and their surroundings were cut and placed in the positions of blue dot rectangles in the same segment. The

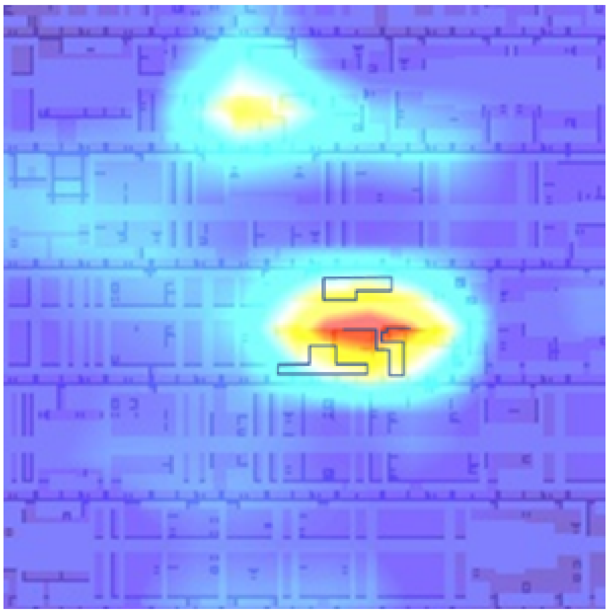


Figure 19: Saliency map obtained based on the input of a layout image of 20 μm.

spaces between the blue and red rectangles in the segment were filled with small-cut portions of dummy patterns. A saliency map was generated using the CNN model A20 (Figure 21(b)). The map demonstrated that the model accurately recognized the area in which the L- and T-shape patterns were close to each other in the image (the red dotted ovals). On the contrary, the intensities of the heat map were weak (marked in blue in the map) in the areas in which the L- and T- patterns solely existed and the areas in which the patterns did not exist. These results indicate that the CNN models recognized the feature in the layout images not according to a single pattern but based on a group of multiple patterns including their surroundings.

	(a) B600		(b) B150	
Segment #	Saliency map	Model's output	Saliency map	Model's output
#1		0.876		0.015
#2		0.999		0.066
#3		0.973		0.955

Figure 20: Saliency maps of the CNN models trained on 20 μm images. The CNN models of (a) B600 and (b) B150, which were trained on Dataset B, comprising 600 and 150 risk images, respectively, were utilized. The image segments that centered the defect positions were used to generate saliency maps. The location of the feature in segment #1 in (b) was slightly shifted from the center of the image.

4.3.3 Effect of image resolution on the results of feature extraction

As mentioned above, the CNN models trained on the 50 μm images classified according to the density of layout patterns, and not on the patterns themselves, when the layout of the AA layer in the 130 nm design was used. In turn, the models trained on the 20 μm images were capable of recognizing particular patterns (polygons). I considered that this variation in the results of feature recognition occurred due to the improvement of the resolution of

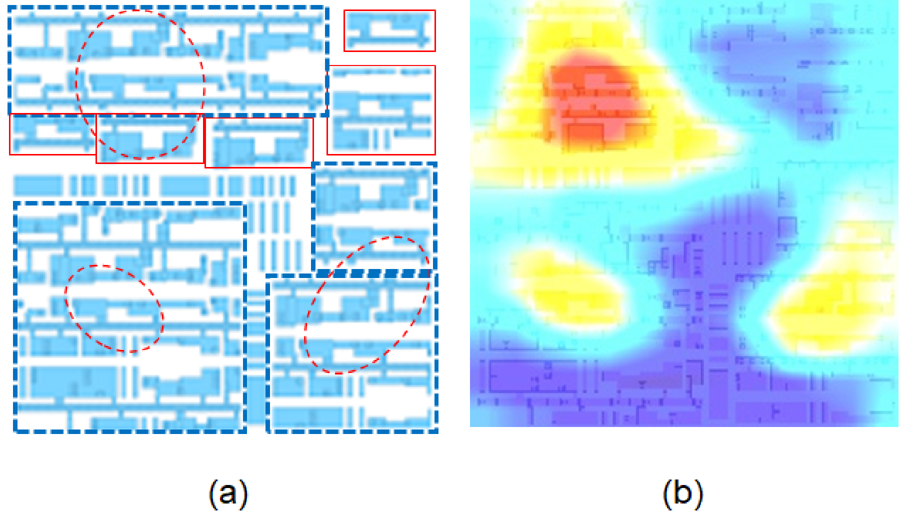


Figure 21: Visualization of the layout feature extracted by the CNN model in a virtual layout: (a) a virtual layout in which partial layouts of various sizes around the patterns extracted by a CNN model were assembled; (b) a saliency map derived from the input image (a).

layout patterns by more than approximately 10 pixels in both x and y directions. For comparison with the pixel sizes of the images, the dimensions of the L- and T-shape patterns and relatively large spaces between layout patterns (Figure 22), recognized by the CNN models as the features in the image segments, are listed in Table 4. Layout patterns corresponding to the shapes similar to the L- and T-shape patterns also existed in the non-risk images in the training datasets. However, those patterns were recognized as a common feature of the risk images, as they were distinguished by recognizing the differences in layout configurations including their surroundings. As the resolution of layout images increased, the results of image classification would be close to those of pattern matching, recognizing smaller patterns in images and extracting patterns of the same target shapes. However, the CNN models captured the features of the images and estimated the similarity

using not only particular patterns but also the groups of patterns in wider areas. This indicates that the use of CNN classification made it possible to distinguish the layouts with higher risk of defects and clarify the root causes of unknown defects.

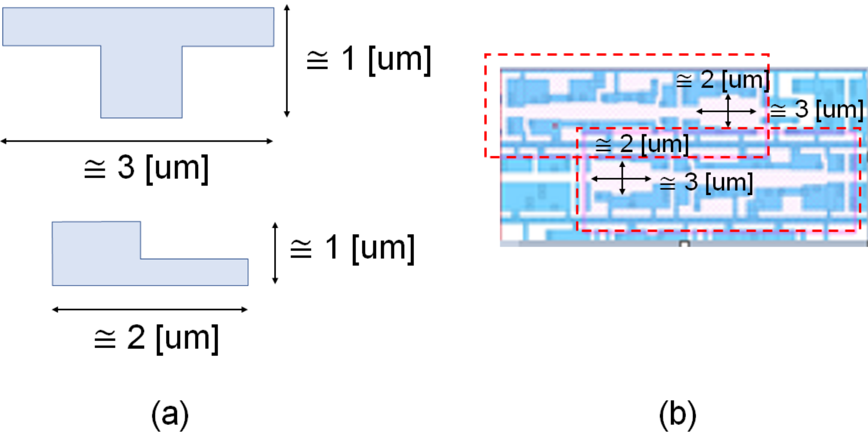


Figure 22: Dimensions of layout patterns of (a) L- and T- shape, and (b) spaces in the layout of STD cells.

4.4 Classification of layout images

4.4.1 Performance of CNN models in classifying layout images

The image segments for testing were cut into square grids from the entire logic circuit area in an LSI chip. Four segments that were the nearest to each position of all 30 defects were denoted as defect-neighboring segments and were used to evaluate the correctness of the detection of the defect positions through CNN classification. All other image segments, except defect-neighboring ones, were named test segments and were classified by CNN models. The test segments were classified into two categories: risk images

Table 4: Pixel sizes and pattern dimensions. (a) Sizes of pixels in the images of three sizes: 50, 30, and 20 μm; (b) approximate dimensions of the patterns extracted as features by the CNN models and the corresponding numbers of pixels.

(a)

	Image size [um]		
	50	30	20
Pixel size [um/pixel] A	0.223	0.134	0.089

(b)

patterns	x/y	Dimension [um] B	Number of pixels B/A [a]		
T-shape	x	≅ 3	13.5	22.4	33.7
	y	≅ 1	4.5	7.5	11.2
L-shape	x	≅ 2	9	14.9	22.5
	y	≅ 1	4.5	7.5	11.2
Space in a STD cell	x	≅ 3	13.5	22.4	33.7
	y	≅ 2	9	14.9	22.5

[a]Corresponds to image sizes in the upper table (a).

that have higher probability to generate defects and non-risk images exhibiting less probability of defects according to the CNN model outputs. The performance estimates of the models A50, A30, and A20 trained on Dataset A, comprising images of 50, 30, and 20 μm, respectively, are summarized in Table 5. I plotted the following indicators, the detection rate of defect positions (DTR), and the risk-image classification rate (RCR) in Figure 23. The number of test segments used to evaluate each model is listed in Table 3. DTR and RCR are defined as follows:

$$DTR = \frac{Number of defect positions which are judged as being detected}{Number of all defect positions},$$

and

$$RCR = \frac{Number of test segments classified as risk images}{Number of all test segments}.$$

The image segments used in mode testing were cut from an LSI chip layout without considering the defect positions; thus, most of the defect positions were not located in the center of the segments but were close to the edge of the segments. When the layout features extracted by the CNN models were far from the defect positions, the features could be located not in the segments including the defect position but in those next to such segments. I formulated a rule to judge the detection of defect positions in each classification as follows.

Rule to judge the detection of defect positions A defect position was judged as detected when at least one of the defect-neighboring segments (the four nearest image segments with respect to the defect position) was classified as risk images.

Table 5: Results of judging the defect position detection and test-segment classifications using the CNN models (a) A50, (b) A30, and (c) A20 trained on the layout images of different sizes.

Models	Defect-position judgement		Test-clip classification	
	Detected	Not detected	Risk image	Non-risk image
(a) A50	26	4	116	175
(b) A30	10	20	78	804
(c) A20	8	22	153	1997

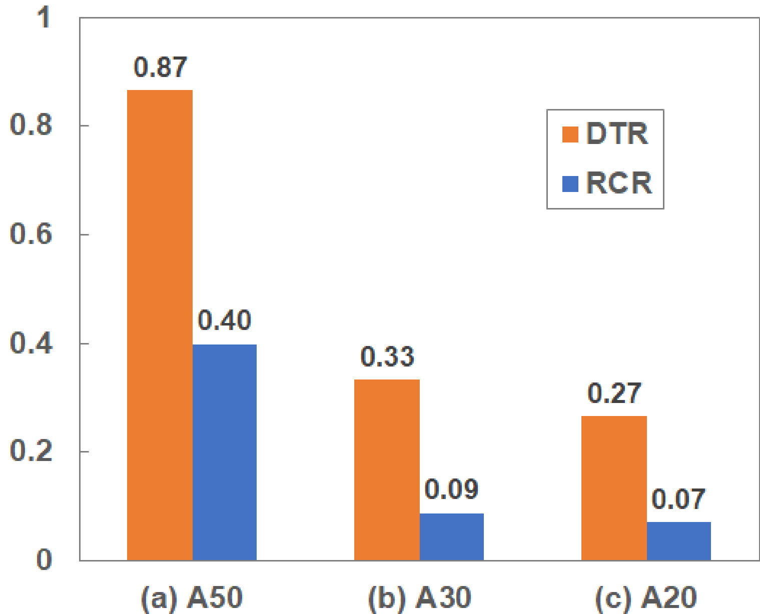


Figure 23: Classification performance of the CNN models: (a) A50, (B) A30, and (c) A20 trained on the layout images of different sizes. The CNN models A50, A30, and A20 were evaluated using the images of 50, 30, and 20 μm, respectively. The rule of defect position detection was applied. Defect positions were judged as detected when at least one of the defect-neighboring segments (the four nearest image segments with respect to the defect position) were classified as risk images.

DTR drastically declined when the image size was decreased. The model A50 trained on the images of 50 μm classified test segments with the DTR of 0.87. The indicator decreased to 0.33 and 0.27 when using the models A30 and A20 trained on the images of 30 and 20 μm, respectively. RCR also decreased when the A30 and A20 models were used. To recognize the features of LSI layouts with a risk of failures, the models trained on the 20 μm images were deemed preferable, as mentioned in the previous section. However, searching for the defects in silicon was difficult using the classification results

of the model A20 due to the low DTR. Then, I attempted to improve the classification performance of the CNN models trained on the images of 20 μm by increasing the size of training datasets.

4.4.2 Improvement of model classification results by increasing the size of training datasets

In the previous section, using a dataset of smaller image sizes resulted in decreasing DTR. When the image size was decreased, the number of layout patterns in an image decreased as well, implying that the volume of information from the images was reduced. In addition, smaller volume of datasets might not provide sufficient information to train the models. Therefore, I increased the size of the risk image datasets used for training by cutting the layout images in any location including defect positions, and not by cutting the images limitedly centering the defect positions. A maximum of 600 images (20 images for each of 30 defect positions) was generated to use as the training dataset. The impact of the dataset volume on the results of classification was investigated using Dataset B comprising the images of 20 μm (Table 6, Figure 24).

As shown in Table 6, the model (a) B600 trained on 600 risk images classified all 30 test segments including defect positions as risk images (DTR = 100%). However, the number of all images classified as risk images (similar to false positives) were equal to 673, and therefore, model performance was inadequate in terms of searching defects in silicon. Then, the model was trained on Dataset B with (b) 300, (c) 150, (d) 100, and (e) 60 risk images. Using the model (c), DTR was still 90%, and the number of test-segments

Table 6: Results of analyzing defect positions' detection and test segment classification using the CNN models: (a) B600(e) B60 trained on the layout images of 20 μm using the datasets of different volumes.

Models	Defect-position judgement		Test-clip classification	
	Detected	Not detected	Risk image	Non-risk image
(a) B600	30	0	673	1564
(b) B300	27	3	329	1908
(c) B150	27	3	223	2014
(d) B100	23	7	128	2109
(e) B60	20	10	21	2216

classified as risk images decreased to 223 ($RCR \fallingdotseq 10\%$). When the risk images in the datasets were decreased to 100 and below, DTR declined below 80% (model (d) B100 and (e) B60). Based on these results, the training dataset including 150 risk images was considered as preferable to obtain high DTR and low RCR.

4.4.3 Hazard maps of defect generation

All outputs of the models corresponding to the test segments were assigned the coordinates of the center of the segments within the LSI chip layout. Contour lines were determined based on the slope between the outputs of the neighboring segments. I referred to these contour lines as hazard maps. As the model outputs for the class of risk images expressed the similarities to the risk images in the training datasets, these plots were equivalent

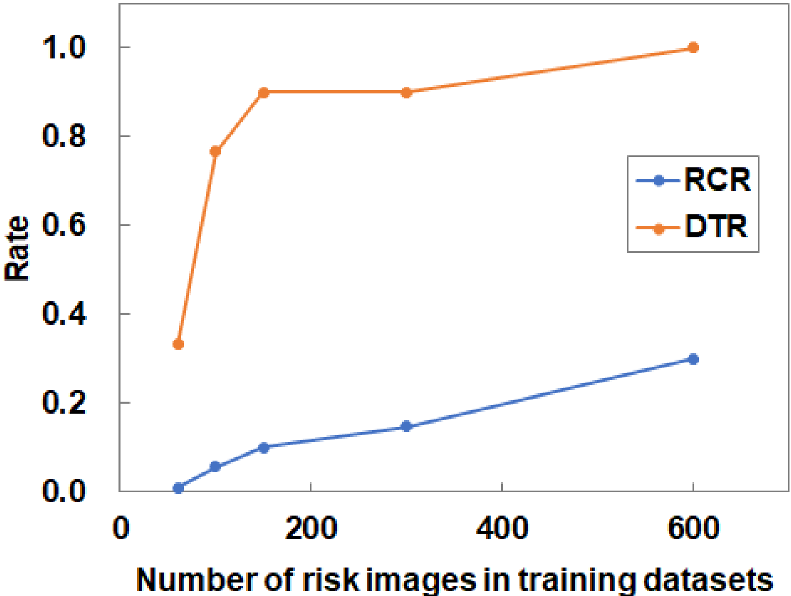


Figure 24: Classification of the CNN models (a) B600 to (e) B60 trained on different datasets of 20 μm images; the plots of RCR and DTR of the models trained on different training datasets. The special rule to judge on detecting defect positions was applied.

to the hazard maps representing the risk of defect generation. The hazard maps converted from the outputs of the CNN models A50, A30, and A20 were compared. The area of logic circuits in the LSI chip was drawn out as a map, as represented in Figure 25. The positions corresponding to higher risk of failures were highlighted in the map. As mentioned in Section 4.4.1, when the size of images in the dataset was reduced, DTR and RCR declined, meaning that the number of segments classified as risk images decreased. This result could be seen in the obtained hazard maps as well. Model A20 (Figure 25(a)) had the lower high-risk areas than model A50 (Figure 25(c)). The positions of the high-risk areas differed from each other in these maps. However, some parts of these areas were constant in all three models (dotted rectangles).

These areas, which were commonly highlighted in multiple maps, became the candidates to search unknown defects in silicon. Low high-risk areas are preferred in hazard mapsso that the positions of those areas could be utilized to search defects in silicon using PFA in conjunction with other information, such as the paths of critical nets determined by failure analysis [144].

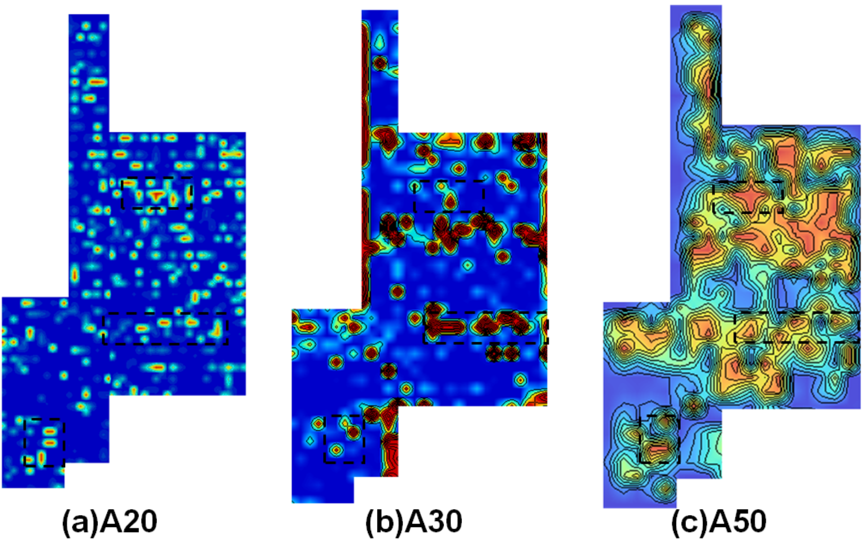


Figure 25: Contour plots of the outputs of models (a) A20, (b) A30 and (c) A50 which were trained on Dataset A of different image sizes.

Figure 26 represents the hazard maps obtained based on the outputs of the models B100 to B600. As mentioned in the previous section, when the models were trained on fewer risk images, fewer segments were classified as risk images. As indicated in the maps shown in Figures 26(b) to 26(d), most defect positions indicated as flags in the figures overlapped with the orange (or red) colored areas. This indicates that almost all defect positions could be detected by the models trained on the datasets including no less than 150 risk images. However, particular defect positions were not detected as a result

of the classification using the model B100 (Figure 26(a)). The hazard maps generated based on the outputs of the CNN models trained in an optimum condition to obtain high DTR and low RCR can serve as references to infer unknown defect positions and support the search of defects in PFA.

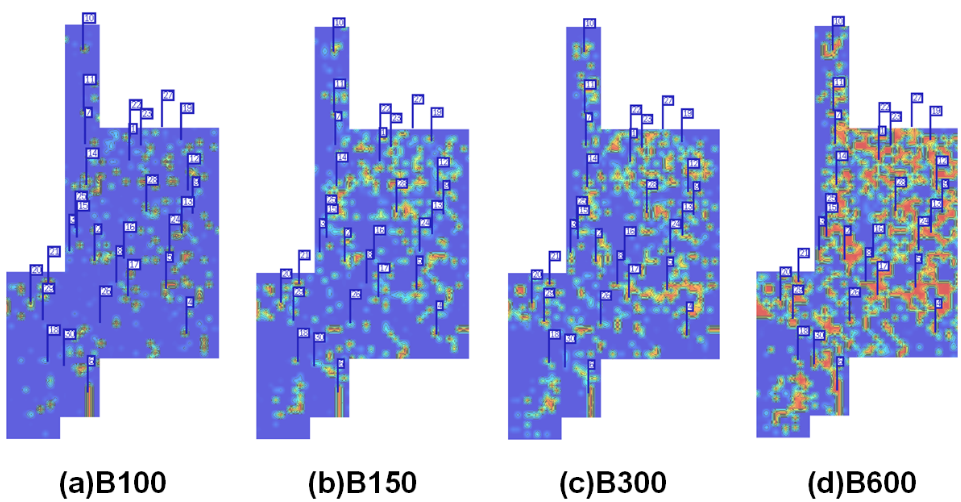


Figure 26: Contour plots of the outputs of the models (a) B100(d) B600. Layout images of 20 μm were used in the training and evaluation datasets. The number of risk images in training datasets was (a) 100, (b) 150, (c) 300, and (d) 600.

4.4.4 Using model outputs in root cause analysis

The hazard maps indicating the local areas with a risk of generating defects within the LSI layouts were determined based on the outputs of the CNN models. High model outputs indicated the high-risk areas in a map. However, even in the case of the images centered on the defect positions, the CNN models could output small values indicating that the images were classified as non-risk images. In other words, the input image exhibited little similarity with respect to the risk images in training datasets. For example,

the model trained on 150 risk images outputted <0.5 to the input images #1 and #2 in Figure 20(b) in which defect positions were located in the center of images. I consider the following two points as the reason of this inconsistency. First, I assume that the models cannot precisely capture the features located in the center of images when the models are trained on the risk images in which the defect positions are far from the center. Second, I assume that the model outputs may vary as the risk levels of generating defects are originally different among all 30 defects detected by PFA. To accurately capture the layout features, increasing the volume of training datasets is found to be effective. Although false positives can increase if the models trained on a large number of risk images in the datasets, specific layout patterns can be accurately captured by the models as the features of risk images (Figure 20(a)). The model performance depends on the resolution of images and the volume of training datasets. Thus, considering the ways of processing the information regarding defects and optimizing the model training procedure is crucial.

4.5 Conclusions

In this study, I proposed a scheme of layout image classification using CNN to infer the root causes of LSI defects. Si-substrate defects induced by the layouts of the isolation layer were the target of this study. CNN models based on VGG16 were trained on the image datasets comprising risk images including the positions of real defects and non-risk images excluding the positions of those defects. Then, these models were used to classify the test seg-

ments that were cut from the layout of a logic circuit area into square grids. The classification included two classes: risk images with a risk of defects and non-risk images exhibiting less risk of defects. The accuracy of the features extracted using the CNN models varied according to image resolution. In the case of low-resolution images, the models were capable of recognizing the local density of patterns in layouts. Using higher-resolution images, the models accurately captured the features of LSI layouts corresponding to defect positions and recognized the layout patterns as the features associated with particular shapes including their surrounding area. These layout features could be further used to infer the root causes of the layouts generating defects and could serve as inputs while developing future designs. However, note that even when images of higher resolution were employed, the use of images of smaller size deteriorated the classification performance of CNN models. This limitation was mitigated by cutting image segments in an arbitrary position including real defect positions to increase the size of the training datasets, which afforded a better performance (the detection rate of defect positions $DTR \cong 90\%$, and the risk-image classification rate $RCR \cong 10\%$) even in the case when the models were trained on small-sized segments. The risk of defects in a layout was visualized as a hazard map predicting the locations of defect candidates. The model outputs, which indicated the similarity of test images to the risk or non-risk images in the training datasets, were used to generate a hazard map. For models with low RCR, the number of the predicted high-risk image segments was limited. In other words, the number of predicted defective positions in the hazard maps was limited, which facilitated the search of unknown defects through

PFA. This study demonstrated the feasibility of CNN classification of layout images for analyzing LSI failures. For the future practical usage, to train the CNN models on fewer dataset of defective layouts and to predict unknown defects is a challenge.

5 Layout-based test coverage verification for high-reliability devices

5.1 Introduction

Reliability and quality are the prerequisites of LSI devices, which are now widely adopted in everyday appliances, especially automobiles, healthcare equipment, and security devices. LSI manufacturers have developed superior processes to reduce variation in production and highly sensitive tests to detect abnormalities. However, in reality, a small number of chips containing defects still manage to pass pre-shipment tests and are delivered to customers. To eliminate the distribution of defective chips, a great deal of effort, including failure analysis, is spent on defect reduction. Defective LSI chips with failures that are detected after shipping have been categorized into two failure mode types. The first is a defect that gradually degrades the chip performance during use, such as a dielectric breakdown in gate oxides. The second is a test escape, which is a defect that passes shipping tests because of a lack of test patterns. Test escapes are one of the critical problems for LSI quality assurance because they are difficult to identify through failure analysis. Additional patterns are generally implemented to eliminate test escapes, but their addition increases design and test costs. An effective solution is necessary to improve the test coverage and achieve zero-defect LSI quality. The fault coverage for a specified failure mode is defined as the ratio of the number of faults detected by tests to the number of all assumed faults in a fault simulation. For example, stuck-at fault coverage is determined from

the ratio of the number of detectable fault-assuming ports to the number of ports where the circuit nodes are tied to 0 or 1 independent of the switching sequence. Fault coverage has been adopted to measure the test coverage of quality assurance tests. The Automotive Electronics Council (AEC) defines test coverage as a test quality metric that is the ratio of the number of detected faults to the number of all possible faults excluding undetectable faults [145]. These values of test coverages do not accurately represent the failure detectability or the risk of test escapes, because the sensitivity of each net for defects can vary substantially with the large variation in the lengths of wires and the spaces between them on actual LSI devices. Test escapes have been estimated from the fault coverage [146], but this method is also based on the same assumption that faults in each net occur equally. In this paper, the conventional test coverage was calculated from the numbers of faults, as shown in Equation (6).

$$C_{logic} = \frac{n_{detected}}{n_{assumed} - n_{undetected}}. \tag{6}$$

Here, the number of detected faults divided by the substitution of the number of assumed faults minus the number of undetected faults is called the logical test coverage C_{logic} to clarify the difference between this quantity and the newly proposed physical test coverage.

5.2 Physical test coverage

To compensate for the shortcomings of logical test coverage in estimating test escapes, a new parameter called physical test coverage is proposed to analyze the test coverage based on the physical layout of LSI circuits. The physical test coverage is calculated by critical areas (CAs), which are determined from the physical layout of LSI circuits. In the following section, the concept of the CA is described.

5.2.1 Critical area (CA)

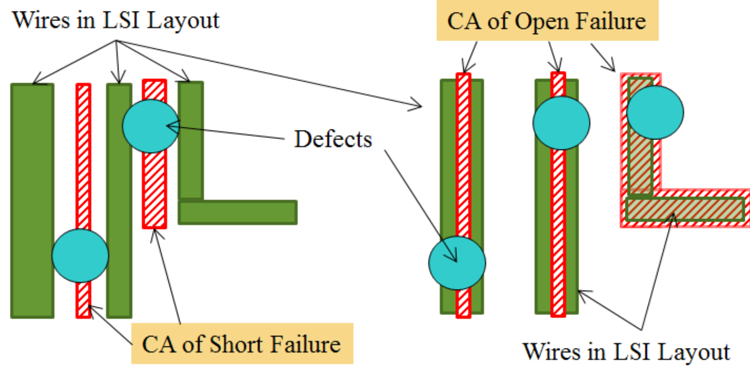


Figure 27: Critical areas of short and open defects in a wiring layer of a logic circuit.

The CA is the quantitative scale for area in an LSI circuit layout in which a particle of a specified size creates an open or a short defect. This defect breaks or connects wires on each mask layer when it falls onto the LSI wafer surface (Figure 27). Electronic design automation (EDA) tools can be used to compute the CA from the physical layout data of an LSI. The CA indicates the inherent potential of the layout to cause defects. Now, let us define I as the set of all layers in an LSI, and i as its variable. In the CA, $A_i(x)$ is

the quantity of a function of defect size x in each layer $i(i \in I)$. $A_i(x)$ is combined with a defect size distribution density function $h_i(x)$ to calculate the probability of the device failing or yielding [16]. The average CA of layer i, A_i, is given by the integral of the product of $A_i(x)$ and $h_i(x)$ as

$$A_i = \int_0^{\infty} A_i(x) \cdot h_i(x) dx. \tag{7}$$

The number of faults, λ_i in layer i, is obtained by multiplying the A_i and the average defect density D_i in layer i.

$$\lambda_i = D_i \cdot A_i. \tag{8}$$

Following Poisson ' s yield model, the yield Y_i of layer i is

$$Y_i = \exp(-\lambda_i). \tag{9}$$

The total device yield Y is the product of the yield Y_i of each mask layer, denoted as

$$\begin{aligned} Y &= \prod_{i \in I} Y_i \\ &= \exp(-\sum_{i \in I} \lambda_i) \\ &= \exp(-\sum_{i \in I} D_i \cdot A_i). \end{aligned} \tag{10}$$

This can be further simplified by using some approximation, that is,

$$
\begin{aligned}
Y &= \exp(-D_0 \cdot \sum_{i \in I} A_i) \\
&= \exp(-D_0 \cdot A_c),
\end{aligned}
\tag{11}
$$

where D_0 and A_c are the average of the defect densities and the sum of the A_i of all mask layers, respectively.

5.2.2 Calculation of physical test coverage to estimate the risk of test escapes

The physical test coverage C_{phys} is defined as the residual probability for the ratio of the CA of undetected wires, $A_{undetected}$, to the CA of all wires, A_{all}, in the LSI layout, as follows:

$$
C_{phys} = 1 - \frac{A_{undetected}}{A_{all}}. \tag{12}
$$

The $A_{undetected}$ is obtained by the CA calculation using the physical layout of undetected wires. The physical layout of undetected wires is extracted from the entire circuit layout by referring to the information of undetected nets. This information is described in a faults list. The fault list is a text file output from design for testability (DFT) tools or fault simulators, which lists the names of nets and the fault types. In contrast to the CA computation mentioned above, an extra layout operation is required to determine the $A_{undetected}$. When the defects are relatively large, the CA of undetected wires can be partly eliminated, because some faults caused by defects on the

undetected wires are detectable by the neighboring detected wires. A layout configuration was considered to calculate the CA of undetected wires in this study. An example of the localization of $A_{undetected}$ of a test layout is described in Section 5.4.2. The probability P_i of a test escape occurring in each layer i ($i \in I$) is calculated as Equation (13) [147]. The CA of undetected wires in layer i, $A_{i,undetected}$, is multiplied by the defect density of the same layer.

$$\begin{aligned} P_i &= 1 - \exp\left(-(1 - C_{phys}) \cdot \lambda_i\right) \\ &= 1 - \exp\left(-(1 - C_{phys}) \cdot D_i \cdot A_i\right) \\ &= 1 - \exp\left(-D_i \cdot A_{i,undetected}\right). \end{aligned} \tag{13}$$

Finally, the probability P of the test escape of a device is a product of P_i, as follows.

$$P = 1 - \prod_{i \in I} (1 - P_i). \tag{14}$$

5.2.3 Flow of physical test coverage calculation

The flowchart of the physical test coverage calculation is shown in Figure 28.

1. Extraction of physical layout of undetected wires: By referring to the net description in a fault list, the layout polygons of the undetected wires are extracted. In addition, the entire set of polygons of wires in an LSI device is extracted. New layout files including the extracted polygons are then generated. Basically, the nets described in the fault lists are the wires in random logic circuits. Therefore, the object of this

coverage evaluation is the wires of random logic circuits.

2. Critical area calculation: The CA of each mask layer is calculated by using the new layout files and an EDA tool (Calibre of Mentor Graphics Corp.). After generating the CA of undetected wires, the portion of the CA, in which the faults are detected by neighboring detected wires, is subtracted by the layout operation. The CA of undetected wires, the result from the layout operation and the CA of all wires are used to calculate the physical test coverage C_{phys}.

3. Physical test coverage calculation: Finally, the physical test coverage is estimated from the ratio of the CA of the undetected wires to that of all wires.

5.3 Comparison of logical and physical test coverages

The physical and logical test coverages of pilot devices of 40 nm design nodes were investigated. The logical test coverage C_{logic} was calculated by the number of faults in the faults list as shown in Equation (6). The physical test coverage C_{phys} was calculated by using the same fault lists used for calculating logical test coverages. The physical test coverage was lower than the logical test coverage for all investigated devices (A to M), as shown in Figure 29. The logical test coverage of all devices was higher than 97%, which is the minimum value satisfying the test coverage criterion specified in [145]. However, the physical test coverage of some devices did not reach the same level. In addition, the physical test coverage varied among the

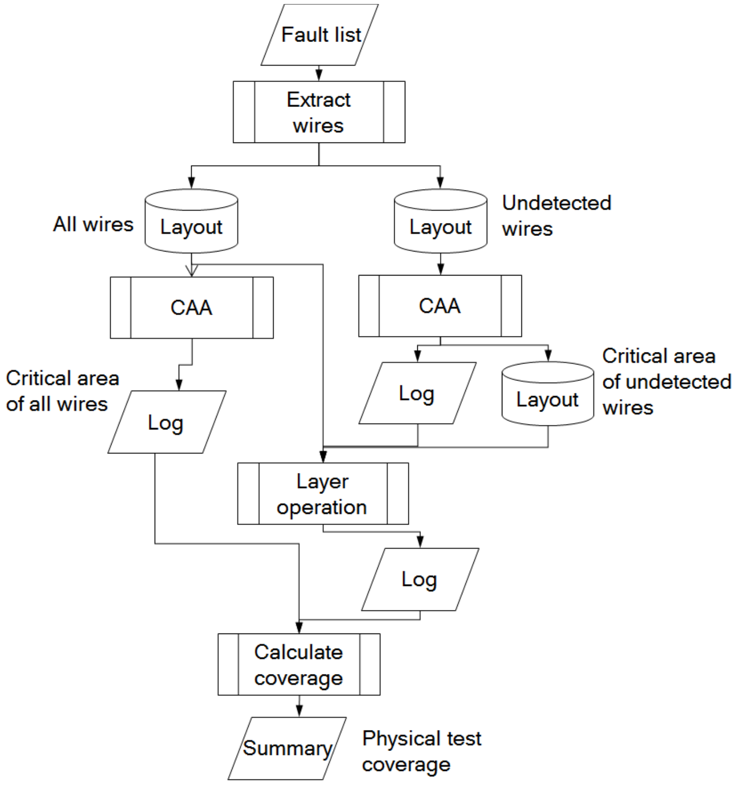


Figure 28: Flowchart of physical test coverage calculation using critical area analysis (CAA).

devices, whereas the logical test coverage remained nearly constant. No clear relation was seen between the two types of test coverage. This result indicated that even though the logical test coverage of a device was high enough to satisfy the test criterion, the physical test coverage might be low and have an unexpected risk of generating test escapes.

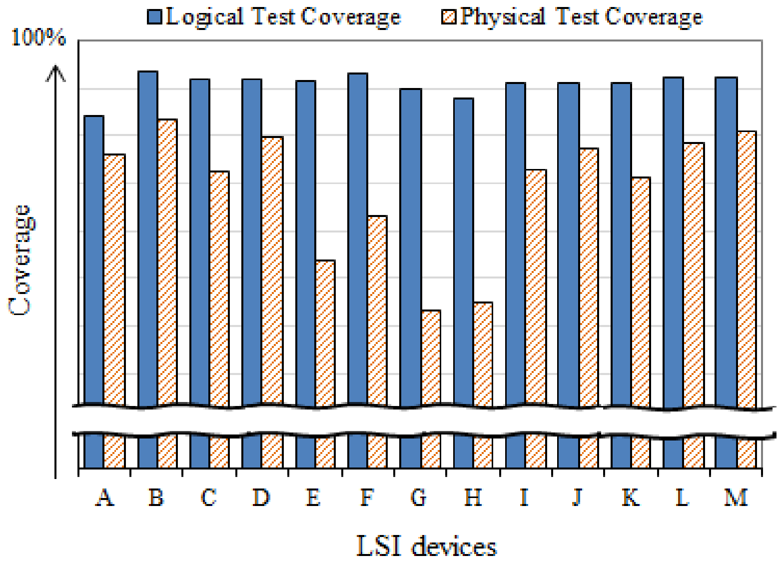


Figure 29: Physical and logical test coverages of LSI devices (A to M).

5.4 Discussion

5.4.1 CA and length of undetected wires

In this section, the physical layout of the wires extracted from LSI circuits is discussed. In investigations of the wire length distribution in the layouts of LSI devices, it was reported that approximately 80% of the total wire length was occupied by 20% of the total nets in a device [148], as shown in Figure 30. This indicates that long wires are used only for specified nets in an LSI layout. Therefore, development of test patterns, especially for physically long nets, is crucial to raise the physical test coverage. Even though the logical test coverage calculated from the number of ports in a circuit is high enough to assure test quality, the physical test coverage remains low when the test patterns are assigned only to physically short nets, and so long nets are undetected. The lower physical test coverage represents a higher risk of

test escapes.

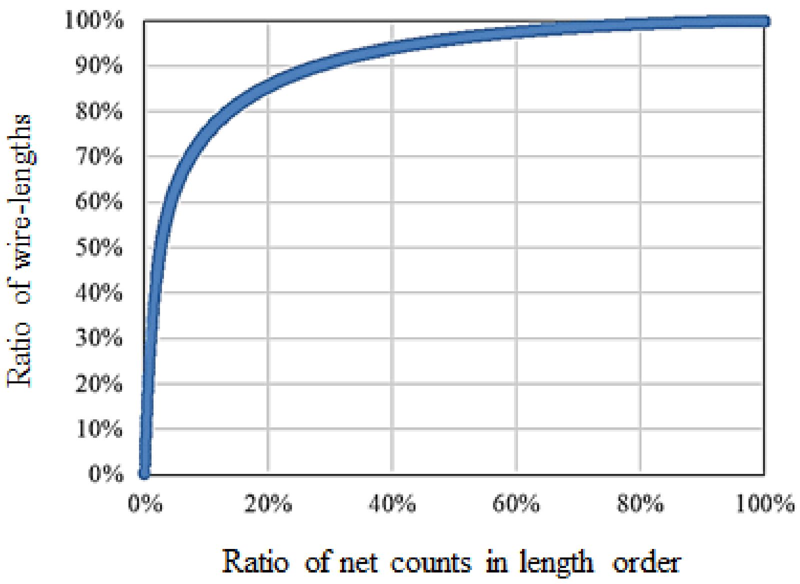


Figure 30: Wire-length distribution in an LSI logic circuit.

In most of the design processes of LSI logic circuits, DFT tools are used to improve the test coverage. When using these tools, the wires connecting functional blocks or hard macros tend to be long and remain undetected. LSI designers prioritize the specification of an LSI function and manually develop test patterns of the functional blocks that are most intensively used by their customers. As a result, long wires at the boundaries of circuit blocks are ignored when assigning test patterns and so are left undetected.

5.4.2 Relation between physical test coverage and lengths of undetected wires

In four LSI devices (N to Q) developed for the same design nodes, the relation between the physical test coverage and the average length of unde-

tected wires was investigated (Figure 31). The average length of undetected wires in each of the devices was longer than the average length of all wires in their logic circuits. Devices O and Q, which showed relatively low physical test coverage, had undetected wires longer than those in the other two devices. These results indicate that long wires tend to remain undetected and reduce the physical test coverage, and thereby increase the risk of test escapes.

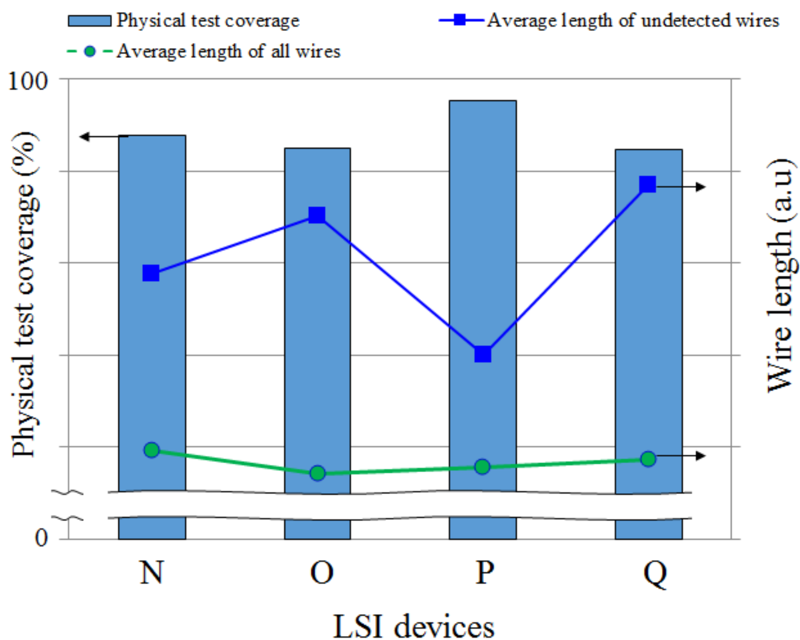


Figure 31: Average lengths of undetected wires and all wires, and the physical test coverages for four LSI devices (N to Q).

5.4.3 Relation between length and critical area of undetected wires

The relation between the total length of undetected wires and the critical area on the layout of various LSI devices was investigated (Figure 32).

Because the CA of open failure correlates with the wire length, the CA increases as the total length of the undetected wires increases. However, the variation in the CA is very wide (e.g. Device 1 and 2) so that the CA of the investigated devices could not be predicted from the wire length. Some devices had small CAs despite having long undetected wires.

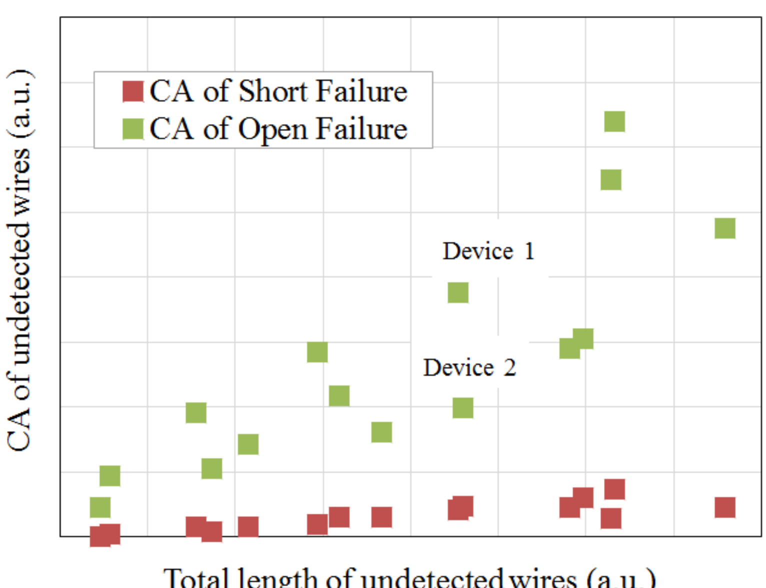


Figure 32: Relation between total length and critical area of undetected wires in logic circuits of LSI devices.

The probability of failure θ, which is the CA of all short or open failures, A_{all}, divided by logic circuit area S, was evaluated as an index of the wiring layout density as shown in Equation (15). This index shows the layout congestion of active wires excluding dummy and redundant figures. The 2nd to 5th metal layers were used to calculate the CAs.

$$\theta = \frac{A_{all}}{S}. \tag{15}$$

The θs of short and open failures of devices 1 and 2 in Figure 32 are compared in Figure 33. Although these devices had undetected wires in similar lengths, the θs of device 2 were higher than those of device 1. This result indicates that the wiring layout density of the logic circuit in device 2 is higher than that of device 1.

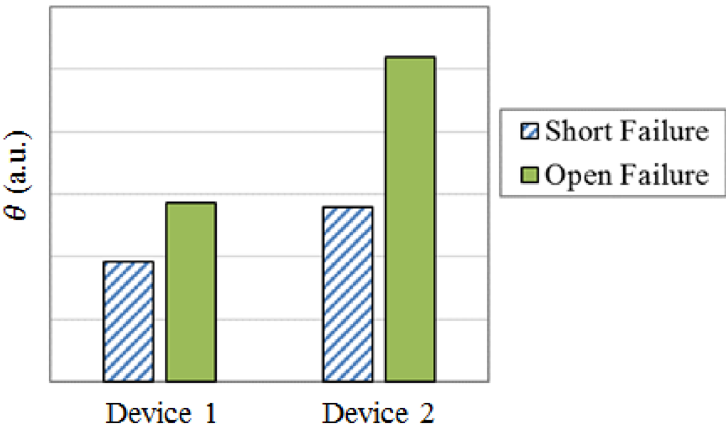


Figure 33: The short and open θs of metal layers in logic circuits of LSI devices that had undetected wires of similar lengths.

In high-layout density devices, the spaces between adjacent wires are so narrow that some CAs of undetected wires are eliminated, because the faults on undetected wires could be detected by the neighboring detected wires. This detection is only possible for some layout configurations. An example is the CAs of undetected wires in a small logic layout, generated by the $A_{undetected}$ layout operation in the workflow of the physical test coverage calculation (Figure 34). The test layouts consisted of (a) two neighboring undetected wires, and (b) both undetected and detected wires. The short and open CAs are simply merged and create the CA figures when both wires are undetected (Figure 34(a)). However, the CAs of a short failure and a

part of an open failure on the undetected wire were not generated, because of the fault detection by the neighboring detected wire (Figure 34(b)). As this example shows, the CA generation of undetected wires depends on the physical layout of the wires on the LSI device. To accurately predict the risk of test escapes, both the wire lengths and the physical configuration of the wires, such as their positions and shapes, must be considered.

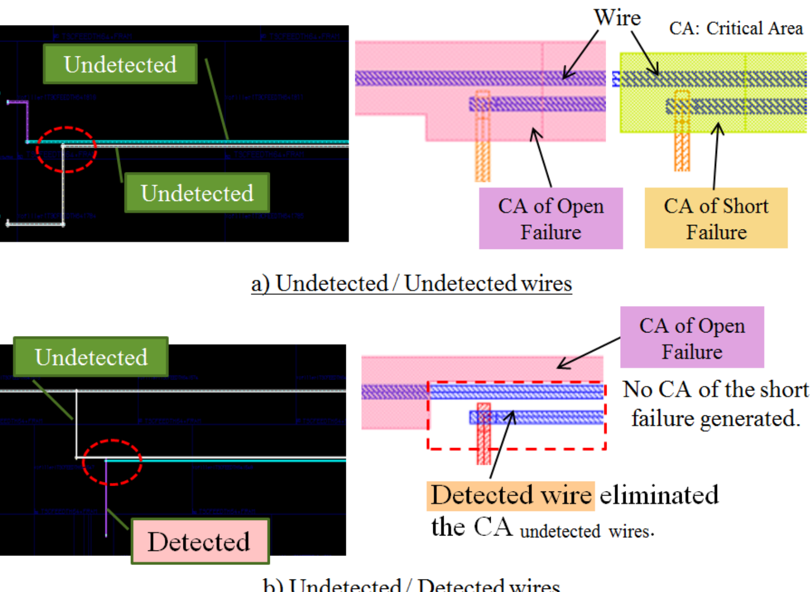


Figure 34: Examples of CA generation for undetected wires in a logic circuit.

5.4.4 Case studies of physical test coverage

The physical test coverage of three LSI devices (R, S, and T) in development was investigated. To provide some background, the relative values of the physical coverage and the probability of test escape of each device are shown in Figure 35. The logical test coverage (not shown) of the three devices was sufficiently high, but the physical test coverage of the devices was

low because of the long undetected wires in their layouts. When using the physical test coverage as an indicator, the test quality of these devices was improved by developing extra test patterns. In the best case, an estimated 90% of the test escapes were eliminated after the test pattern improvement. To effectively improve the test coverage by adding a minimum number of test patterns, the test patterns should be assigned to specified undetected wires that are prioritized by the CA [149].

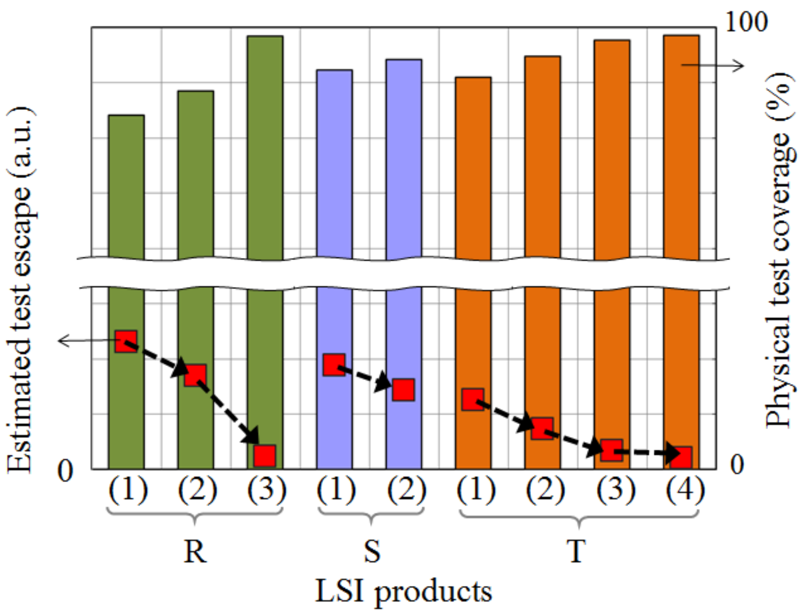


Figure 35: Test quality improvement using physical test coverage.

A workflow of the test development, including physical test coverage verification, is shown in Figure 36. The physical test coverage can be easily adopted into the LSI design workflow. Furthermore, this addition does not change the lead time of LSI design development, because the physical test coverage is calculated in parallel with an existing step in the test coverage verification process.

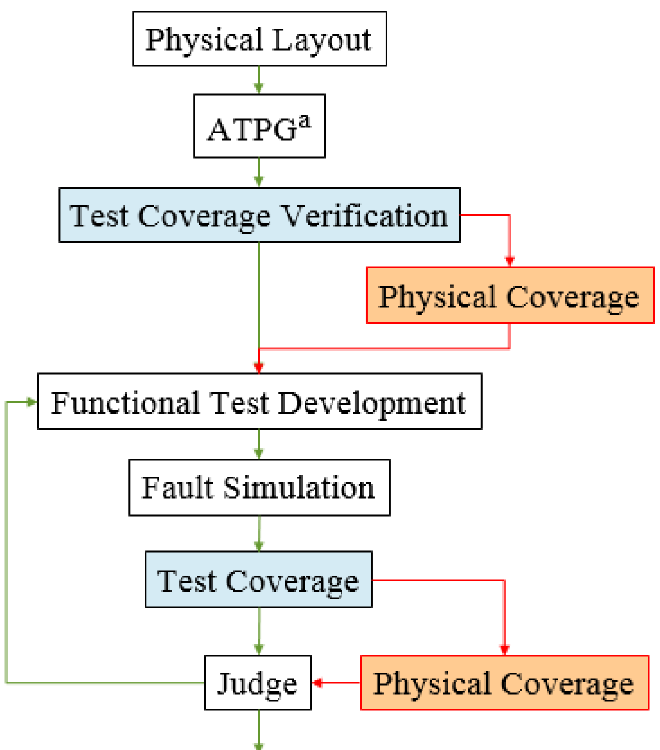


Figure 36: Test quality improvement using physical test coverage. [a]Automatic test pattern generation.

5.5 Conclusions

This paper proposed a new parameter called physical test coverage to estimate the test quality of LSI devices based on their physical layout in order to eradicate test escapes. The physical test coverage is calculated from the CAs of wires undetected by tests. The CAs of undetected wires, $A_{undetected}$ are determined by not only the wire lengths but also the physical configurations of the LSI circuit layout. The risk of test escapes is also quantified by using the $A_{undetected}$. Developing additional test patterns that are clearly prioritized by the $A_{undetected}$ effectively improves the test quality.

6 Conclusions

This dissertation has discussed the techniques of LSI layout analysis to search and evaluate the risk of failure in LSI devices. LSIs are used in many applications of modern society and thus require reliability. It is very vital to predict failures in early manufacturing processes and thus obtain reliable LSI products. The LSI layout is one of the factors that may become the root cause of failure. To address the layout-induced failure, new methods of LSI layout analysis are proposed and investigated. First, a scheme is proposed for classifying LSI circuit layout images using an AI technique and a CNN to identify the defective layout regions. Second, using CNN models, the image features of defective layout segments were extracted, which can be utilized to infer the root causes of failure. Third, to predict the risk of failure chips that could not be filtered in the shipping test (test escapes), a new concept of the indicator evaluating test performances is proposed using critical areas of the wirings undetected.

In Chapter 3, a scheme is proposed for LSI layout segment classification based on CNNs to determine the impact of a regional layout design on failure. CNN models were trained with layout segments, which are the images of LSI layouts cut into small square grids. CNNs showed the ability to classify the images of the LSI circuit layout. The size of the training dataset influenced the classification performance. Basically, it is difficult to train CNN models using small datasets. However, the experimental results showed that a 15-layer model based on VGG16 with transfer learning showed the capability of classifying layout images and identifying the segments that included real

failures using only 90 images (i.e., 30 risk and 60 non-risk images) in the training dataset. The risk of failure was visualized in a hazard map based on the outputs of the CNN model. Thus, using limited hit rates, the conditions of model training can be compared and optimized.

In Chapter 4, another scheme is proposed for layout image classification using CNN for the root cause analysis of LSI defects. Si-substrate defects induced by the layouts of the isolation layer were the targets of this study. CNN models trained on the datasets comprising risk and non-risk segments of the layout of isolation layer were used to classify the test segments that were cut from a logic circuit area. The accuracy of the features extracted using the CNN models varied according to the image resolution. Using higher-resolution images, the models accurately captured the features of LSI layouts corresponding to defect positions and recognized the layout patterns as the features associated with particular shapes, including their surrounding area. These layout features could be further used to infer the root causes of the layouts generating defects and could serve as inputs while developing future designs. However, owing to the limitation of the input size of the CNN model used in this study, a higher resolution requires smaller image segments, which deteriorate the classification performance of CNN models. We mitigated this limitation by cutting image clips in an arbitrary position, including real defect positions to increase the size of the training datasets, which provided a better performance even in the case when the models were trained on small-sized segments. Upon selecting the valance of training datasets, the number of predicted defective positions in the hazard maps decreased, which facilitated the search for unknown defects through PFA.

In Chapter 5, a new indicator called physical test coverage is proposed to estimate the test quality of LSI devices based on their layout and thus eradicate test escapes. To achieve zero defect, i.e., no test escape, LSI manufacturers must demonstrate their product reliability. Physical test coverage is calculated from the CAs of wirings undetected by tests. CA is a DFM tool used to estimate the impact of the layout on the occurrence of defects for predicting failure. Therefore, the risk of test escapes can be quantified using the CA of undetected wires, $A_{undetected}$. $A_{undetected}$ is determined by not only the wire lengths but also the configurations of the LSI circuit layout, which realizes the risk evaluation of LSI products in different designs. Developing additional test patterns that are clearly prioritized by $A_{undetected}$ effectively improve the test quality and show product reliability.

The research to identify the defective regions was the first to demonstrate the feasibility of CNN classification of layout images for analyzing LSI failures. For more practical applications, training CNN models on a small dataset of defective layouts and predicting unknown defects are the next challenges. The concept of physical coverage is widespread in the LSI industry, and techniques for test pattern generation considering CAs of undetected wiring are discussed [150, 151]. The collaboration of design that reduces the risk of failure based on the DFM rules and physical analysis, which determines the cause of previously unknown failures, is the key to ensuring future LSI reliability. Layout analysis plays an important role in determining LSI reliability.

References

[1] Semiconductor Reliability Handbook. Renesas Electronics Corp., 2017. [Online]. Available: https://www.renesas.com/us/en/doc/products/others/r51zz0001ej0250.pdf.' Accessed on: Oct. 20, 2020.

[2] Y. Huang, B. Benware, R. Klingenberg, H. Tang, J. Dsouza and W. T. Cheng, "Scan chain diagnosis based on unsupervised machine learning," in *IEEE 26th Asian Test Symposium (ATS)*, pp.220-225, 2017.

[3] K. Nikawa, "Status of current LSI failure analysis techniques (Japanese)," [Online]. Available: https://www.oeg.co.jp/Exhibition/pdf/13lsi.pdf. Accessed on: Oct. 17, 2020.

[4] A. Toriumi, M. Yoshimi, M. Iwase, Y. Akiyama and K. Taniguchi, "A study of photon emission from n-channel MOSFET's," in *IEEE Transactions on Electron Devices*, pp.1501-1508, Vol. 34, No. 7, 1987. doi: 10.1109/T-ED.1987.23112.

[5] T. C. Ong, K. W. Terrill, S. Tam and C. Hu, "Photon generation in forward-biased silicon p-n junctions," in *IEEE Electron Device Letters*, pp.460-462, Vol. EDL-4, No.12, 1983.

[6] Y. Y. Liu, J. M. Tao, D. S. H. Chan, J. C. H. Phang and W. K. Chim , "A new spectroscopic photon emission microscope system for semiconductor device analysis," in *5th IEEE International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA)*, pp.60-65, 1995.

[7] T. Koyama, Y. Mashiko, M. Sekine, H. Koyama and K. Horie, "New non-bias optical beam induced current (NB-OBIC) technique for evaluation of Al interconnects," in *IEEE International Reliability Physics Symposium (IRPS)*, pp.228-233, 1995.

[8] T. Koyama, K. Sonoda, J. Komori and Y. Mashiko, "Detection of defects in metal interconnects by nonbias-optical beam induced current technique," in *Journal of Applied Physics*, pp.5949-5956, Vol.86, No.11, 1999.

[9] Y. Yamada and H. Komoda, "A successful failure analysis using front and backside fault localization techniques on a deep sub-micron CMOS device," in *International Symposium for Testing and Failure Analysis (ISTFA)*, pp.305-311, 2001.

[10] EBAC Analysis, Renesas Engineering Service Co. Ltd., [Online]. Available: https://reg.renesas.com/business/analysis/physical-analysis-position-identification/ebac/, (Japanese). Accessed on: Jan. 19, 2021.

[11] K. Fujiyoshi, K. Sawai, K. Inoue, K. Saiki and K. Sakurai, "Voltage contrast for gate-leak failures detected by electron beam inspection," in *IEEE Transactions on Semiconductor Manufacturing*, pp.208-214, Vol. 20, No. 3, 2007.

[12] M. Sanada and H. Fujioka, "Yield enhancement for logic LSI by killer defect diagnosis technique using abnormal IDDQ phenomenon," in *IEEE International Symposium on Semiconductor Manufacturing (ISSM)*, pp.265-268, 1998.

[13] M. Sanada, "Evaluation and detection of CMOS LSI with abnormal IDDQ," in *Microelectronics and Reliability*, pp.619-629, Vol.35 No.3, 1995.

[14] H. Yamauchi, M. Yoshida, T. Ono, Y. Umeki and N. Takase, "Hardware and software application for IDDQ test," in *IEEE International Workshop on IDDQ Testing*, pp.26-29, 1995.

[15] J. A. Cunningham, "The use and evaluation of yield models in integrated circuit manufacturing," in *IEEE Transactions on Semiconductor Manufacturign*, pp.60-71, Vol. 3, No.2, 1990.

[16] C. H. Stapper, F. M. Armstrong and K. Saji, "Integrated circuit yield statistics," in *Proceedings of the IEEE*, pp.453-470, Vol. 71, No.4, 1983.

[17] R. C. Leachman and S. Ding, "Excursion yield loss and cycle time reduction in semiconductor manufacturing," in *IEEE Transactions on Automation Science and Engineering*, pp.112-117, Vol. 8, No. 1, 2011.

[18] M. Effron, "Integrated yield management: A systematic approach to yield management," in *IEEE/SEMI Advanced Semiconductor Manufacturing Conference and Workshop. Theme-Innovative Approaches to Growth in the Semiconductor Industry. (ASMC)*, pp.397-403, 1996.

[19] T. Winter, "Comprehensive critical area methodology for defect yield modeling," in *Proceedings of International Symposium on Semiconductor Manufacturing*, pp.483-486, 2002.

[20] C. Zhou, R. Ross, C. Vickery, B. Metteer, S. Gross, and D. Verret , "Yield prediction using critical area analysis with inline defect data," presented at *13th Annual IEEE/SEMI Advanced Semiconductor Manufacturing Conference. Advancing the Science and Technology of Semiconductor Manufacturing (ASMC) (Cat. No.02CH37259)*, pp.82-86, 2002.

[21] C. Hess and A. Strole, "Modeling of real defect outlines for defect size distribution and yield prediction," in *Proceedings of the 1993 International Conference on Microelectronic Test Structures*, pp.75-81, Vol.6, 1993.

[22] X. Jiang, Y. Hao and S. Horiguchi , "A new defect outline model used for critical area estimation in VLSI," in *IEEE International Symposium on Defect and Fault Tolerance in VLSI Systems*, pp.21-29, 2000.

[23] C. Hess, D. Stashower, B. E. Stine, Larg H, "Fast extraction of defect size distribution using a single layer short flow NEST structure," in *IEEE Transactions on Semiconductor Manufacturing*, pp.330-337, Vol. 14, No. 4, 2001. doi: 10.1109/66.964320.

[24] T. Kumazawa, Y. Hamamura, K. Tsunokuni, and H. Iwata, "Measurement method for short defect size distributions," in *Proceedings of IEEE International Symposium on Semiconductor Manufacturing (ISSM)*, pp.1-4, 2002.

[25] H. T. Heineken, J. Khare, and W. Maly, "Yield loss forecasting in the early phases of the VLSI design process,," in *Proceedings of IEEE Custom Integrated Circuits Conference*, pp.27-30, 1996.

[26] G. A. Allan and A. J. Walton, "Critical area extraction for soft fault estimation," in *IEEE Transactions on Semiconductor Manufacturing*, pp.146-154, Vol. 11, No. 1, 1998.

[27] T. Ouchi, K. Makabe, M. Ogasawara, E. Murakami, and N. Yoshioka, "Modeling of Cu IMD-TDDB caused by extrinsic defects," in *IEEE International Reliability Physics Symposium*, pp.120-124, 2010.

[28] Y. D. Chen, J. Huang, D. Zhao, D. Yim and J. Ervin, "Advanced 3D design technology co-optimization for manufacturability," in *2019 China Semiconductor Technology International Conference (CSTIC)*, pp.1-3, 2019. doi: 10.1109/CSTIC.2019.8755659.

[29] Y. Ham, Y. Choi, M. Green, M. Ramadan and C. Progler, "Effective mask patterning strategy as design complexity and CD stability," in *2019 China Semiconductor Technology International Conference (CSTIC)*, pp.1-3, 2019. doi: 10.1109/CSTIC.2019.8755776.

[30] R. Chen, W. Zhong, H. Yang, H. Geng, X. Zeng and B. Yu, "Faster region-based hotspot detection," presented at *2019 56th ACM/IEEE Design Automation Conference (DAC)*, pp.1-6, 2019.

[31] H. Yang, W. Zhong, Y. Ma, H. Geng, R. Chen, W. Chen and B. Yu, "VLSI mask optimization: From shallow to deep learning," in *2020 25th Asia and South Pacific Design Automation Conference (ASP-DAC)*, pp.434-439, 2020. doi: 10.1109/ASP-DAC47756.2020.9045241.

[32] Y. Shimada and K. Sakurai, "A new accurate yield prediction method for system-LSI embedded memories," in *IEEE Transactions on Semiconductor Manufacturing*, pp.436-445, Vol. 16, No. 3, 2003.

[33] S. Banerjee, Z. Li, and S. R. Nassif, "ICCAD-2013 CAD contest in mask optimization and benchmark suite," in *IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, pp.271-274, 2013.

[34] L. W. Liebmann and R. O. Topaloglu, "Design and technology co-optimization near single-digit nodes," in *IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, pp.582-585, 2014.

[35] M. Shafee, J.-W. Park, A. Aslyan, A. Torres, K. Madkour, and W. ElManhawy, "A novel algorithm for automatic arrays detection in a layout," in *Design for Manufacturability through Design-Process Integration VII*, Vol. 8684, p. 86840O, 2013. doi:10.1117/12.2011413.

[36] W. C. Chang and I. H. R. Jiang, "iClaire: A fast and general layout pattern classification algorithm with clip shifting and centroid recreation," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, pp.1662-1673, Vol. 39, No. 8, 2020.

[37] H. Yao, S. Sinha, C. Chiang, X. Hong, and Y. Cai, "Efficient process hotspot detection using range pattern matching," in *IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, pp.625-632, 2016.

[38] J. Y. Wuu, F. G. Pikus, A. Torres, and M. Marek-Sadowska, "Rapid layout pattern classification," in *16th Asia and South Pacific Design Automation Conference (ASP-DAC)*, pp.781-786, 2011.

[39] D. Ding, A. J. Torres, F. G. Pikus, and D. Z. Pan, "High performance lithographic hotspot detection using hierarchically refined machine learning," in *16th Asia and South Pacific Design Automation Conference (ASP-DAC)*, pp.775-780, 2011.

[40] W. Y. Wen, J. C. Li, S. Y. Lin, J. Y. Chen, and S. C. Chang, "A fuzzy-matching model with grid reduction for lithography hotspot detection," in *IEEE Transactions on Computer-Aided Design Integration Circuits System*, pp.1671-1680, Vol. 33, No. 11, 2014.

[41] Y. T. Yu, I. H. R. Jiang, Y. Zhang, and C. Chiang, "DRC-based hotspot detection considering edge tolerance and incomplete specification," in *IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, pp.101-107, 2014.

[42] W. C. J. Tam and R. D. S. Blanton, "" LASIC: Layout analysis for systematic IC-defect identification using clustering," in *IEEE Transactions on Computer-Aided Design Integration Circuits System*, pp.1278-1290, Vol. 34, No. 8, 2015.

[43] H. Y. Su, C. C. Chen, Y. L. Li, A. C. Tu, C. J. Wu, and C. M. Huang, "A novel fast layout encoding method for exact multilayer pattern matching with Prufer encoding," in *IEEE Transactions on Computer-Aided Design Integration Circuits System*, pp.95-108, Vol. 34, No. 1, 2015.

[44] Y. T. Yu, G. H. Lin, I. H. R. Jiang, and C. Chiang, "Machine-learning based hotspot detection using topological classification and critical feature extraction," in *IEEE Transactions on Computer-Aided Design Integration Circuits System*, pp.460-470, Vol. 34, No. 3, 2015.

[45] F. Yang, S. Sinha, C. C. Chiang, X. Zeng and D. Zhou, "Improved tangent space-based distance metric for lithographic hotspot classification," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, pp.1545-1556, Vol. 36, No. 9, 2017. doi: 10.1109/TCAD.2016.2638440.

[46] J. Yu and P. Guo, "Run-to-run control of chemical mechanical polishing process based on deep reinforcement Learning," in *IEEE Transactions on Semiconductor Manufacturing*, pp.454-465, Vol. 33, No. 3, 2020. doi: 10.1109/TSM.2020.3002896.

[47] J. Feng, X. Jia, F. Zhu, J. Moyne, J. Iskandar and J. Lee, "An online virtual metrology model with sample selection for the tracking of dynamic manufacturing processes with slow drift," in *IEEE Transactions on Semiconductor Manufacturing*, pp.574-582, Vol. 32, No. 4, 2019. doi: 10.1109/TSM.2019.2942768.

[48] K. P. Whiite, R. N. Athay and W. J. Trybula, "Applying DFM in the semiconductor industry," in *IEEE/CPMT International Electronics Manufacturing Technology Symposium*, pp.438-441, 1995.

[49] S. K. S. Fan, D. M. Tsai, F. He, J. Y. Huang, and C. H Jen, "Key parameter identification and defective wafer detection of semiconductor manufacturing processes using image processing techniques," in *IEEE Transactions on Semiconductor Manufacturing*, pp.544-552, Vol. 32, No. 4, 2019.

[50] S. J. Jang, J. S. Kim, T. W. Kim, H. J. Lee, and S. Ko, "A wafer map yield prediction based on machine learning for productivity enhancement," in *IEEE Transactions on Semiconductor Manufacturing*, pp.400-407, Vol. 32, No. 4, 2019.

[51] R. Wang and N. Chen, "Wafer map defect pattern recognition using rotation-invariant features," in *IEEE Transactions on Semiconductor Manufacturing*, pp.596-604, Vol. 32, No. 4, 2019.

[52] J. Yu, "Enhanced stacked denoising autoencoder-based feature learning for recognition of wafer map defects," in *IEEE Transactions on Semiconductor Manufacturing*, pp.613-624, Vol. 32, No. 4, 2019.

[53] J. Hwang and H. Kim, "Variational deep clustering of wafer map patterns," in *IEEE Transactions on Semiconductor Manufacturing*, pp.466-475, Vol. 33, No. 3, 2020.

[54] Y. Kong and D. Ni, "A semi-supervised and incremental modeling framework for wafer map classification," in *IEEE Transactions on Semiconductor Manufacturing*, pp.62-71, Vol. 33, No. 1, 2020.

[55] M. Saqlain, Q. Abbas, and J. Y. Lee, "A deep convolutional neural network for wafer defect identification on an imbalanced dataset in semiconductor manufacturing processes," in *IEEE Transactions on Semiconductor Manufacturing*, pp.436-444, Vol. 33, No. 3, 2020.

[56] J. Shim, S. Kang, and S. Cho, "Active learning of convolutional neural network for cost-effective wafer map pattern classification," in *IEEE Transactions on Semiconductor Manufacturing*, pp.258-266, Vol. 33, No. 2, 2020.

[57] N. Chaudhary and S. A. Savari, "Increasing the utilization of deep neural networks for SEM measurements through multiple task formulation and visualization," in *IEEE Transactions on Semiconductor Manufacturing*, pp.322-330, Vol. 33, No. 3, 2020.

[58] B. Guo, J. Zhuang, Y. Wu, W. Wu, F. Wu, and Q. Peng, "Color difference detection of polysilicon wafers using optimized support vector

machine by magnetic bacteria optimization algorithm with elitist strategy," in *IEEE Transactions on Semiconductor Manufacturing*, pp.267-282, Vol. 33, No. 2, 2020.

[59] B. M. Haddad, S. F. Dodge, L. J. Karam , N. S. Patel, and M. W. Braun, "Locally adaptive statistical background modeling with deep learning-based false positive rejection for defect detection in semiconductor units," in *IEEE Transactions on Semiconductor Manufacturing*, pp.357-372, Vol. 33, No. 3, 2020.

[60] J. O'Leary, K. Sawlani, and A. Mesbah, "Deep learning for classification of the chemical composition of particle defects on semiconductor wafers," in *IEEE Transactions on Semiconductor Manufacturing*, pp.72-85, Vol. 33, No. 1, 2020.

[61] M. Azamfar, X. Li , and J. Lee, "Deep learning-based domain adaptation method for fault diagnosis in semiconductor manufacturing," in *IEEE Transactions on Semiconductor Manufacturing*, pp.445-453, Vol. 33, No. 3, 2020.

[62] H. Kim, D. E. Lim, and S. Lee, "Deep learning-based dynamic scheduling for semiconductor manufacturing with high uncertainty of automated material handling system capability," in *IEEE Transactions on Semiconductor Manufacturing*, pp.13-22, Vol. 33, No. 1, 2020.

[63] A. Zhakov , H. Zhu, A. Siegel, S. Rank, T. Schmidt, L. Fienhold, and S. Hummel, "Application of ANN for fault detection in overhead transport systems for semiconductor fab," in *IEEE Transactions on Semiconductor Manufacturing*, pp.337-345, Vol. 33, No. 3, 2020.

[64] A. Chouichi, J. Blue, C. Yugma, and F. Pasqualini, "Chamber-to-chamber discrepancy detection in semiconductor manufacturing," in *IEEE Transactions on Semiconductor Manufacturing*, pp.86-95, Vol. 33, No. 1, 2020.

[65] G. R. Reddy, K. Madkour and Y. Makris, "Machine learning-based hotspot detection: fallacies, pitfalls and marching orders," in

IEEE/ACM International Conference on Computer-Aided Design (ICCAD), pp.1-8, 2019. doi: 10.1109/ICCAD45719.2019.8942128.

[66] M. B. Alawieh, Y. Lin, Z. Zhang, M. Li, Q. Huang and D. Z. Pan, "GAN-SRAF: Sub-resolution assist feature generation using generative adversarial networks," in *Transactions on Computer-Aided Design of Integrated Circuits and Systems*, pp.373-385, Vol. 40, No. 2, 2020. DOI 10.1109/TCAD.2020.2995338.

[67] P. Kareem, Y. Kwon, and Y. Shin, "Layout pattern synthesis for lithography optimizations," in *IEEE Transactions on Semiconductor Manufacturing*, pp.283-290, Vol. 33, No. 2, 2020.

[68] K. Nikawa, S. Inoue and K. Morimoto, "Killer defect detection using the IR-OBIRCH (infrared optical-beam-induced resistance-change) method," in *Proceedings of IEEE International Symposium on Semiconductor Manufacturing Conference*, pp.103-106, 1999.

[69] P. Jacob, "Defect- and structure-weakness-localization on power semiconductors using OBIRCH (optical beam induced resistivity change)," in *Proceedings of 9th International Symposium on the Physical and Failure Analysis of Integrated Circuits*, pp.152-156, 2002.

[70] W. S. Tam, C. K. Wong and C. W. Kok, "Study of microscopic defects in silicon oxynitride films prepared by plasma-enhanced chemical vapor deposition process," *IEEE International Conference on Electron Devices and Solid-State Circuits*, pp.1-4, 2018. [Online]. Available: https://ieeexplore.ieee.org/document/4760636.

[71] M. Komagata, "A new method of reducing the particle contamination in semiconductor manufacturing," in *Proceedings of Japan International Electronic Manufacturing Technology Symposium*, pp.146-149, 1995.

[72] T. Mogami, "Perspectives of CMOS technology from the viewpoint of variability," in *International Symposium on Semiconductor Manufacturing (ISSM)*, pp.1-27, 2010. [Online]. Available: https://ieeexplore.ieee.org/document/5750202.

[73] S. Banerjee, P. Elakkumanan, D. Chidambarrao, J. Culp and M. Orshansky, "Analysis of systematic variation and impact on circuit performance," in *Design for Manufacturability through Design-Process Integration II. International Society for Optics and Photonics*, Vol. 6925, p. 69250K, 2008. doi: 10.1117/12.772075.

[74] M. Cote and P. Hurat, "Standard cell printability grading and hot spot detection," in *IEEE 6th International Symposium on Quality Electronic Design (isqed' 05)*, pp.264-269, 2005. doi: 10.1109/ISQED.2005.113.

[75] F. Yang, Y. Cai, Q. Zhou and J. Hu, "SAT based multi-net rip-up-and-reroute for manufacturing hotspot removal," in *IEEE Design, Automation & Test in Europe Conference & Exhibition (DATE 2010)*, pp.1369-1372, 2010. doi: 10.1109/DATE.2010.5457024.

[76] S. Onishi, A. Ayukawa, K. Uda and K. Sakiyama, "Defect-free shallow P/N junction by point defect engineering," in *Proceedings of Reliability Physics*, pp.102-106, 1992.

[77] M. Orlowski, "Impurity and point defect redistribution in the presence of crystal defects," in *IEEE International Technical Digest Electron Device*, pp.729-732, 1990. doi: 10.1109/IEDM.1990.237097.

[78] J. R. Hu, J. Chen, B. Liew, Y. Wang, L. Shen and L. Cong, "Systematic co-optimization from chip design, process technology to systems for GPU AI chip," in *International Symposium of VLSI Technology, Systems and Application (VLSI-TSA)*, pp.1-2, 2018. [Online]. Available: https://ieeexplore.ieee.org/document/8403802.

[79] S. Banerjee, K. B. Agarwal and M. Orshansky, "SMATO: Simultaneous mask and target optimization for improving lithographic process window," in *IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, pp.100-106, 2010.

[80] Y. Nagamura, K. Hosono, L. Pang, K. K. Chan and Y. Tanaka, "Photomask quality assessment strategy at 90-nm technology node with

aerial image simulation," in *Photomask and Next-Generation Lithography Mask Technology X*, pp.476-483, Vol. 5130, 2003. [Online]. Available: https://doi.org/10.1117/12.504209.

[81] K. Agarwal and S. Banerjee, "Design driven patterning optimization for low k1 lithography," in *IEEE International Conference on IC Design & Technology*, pp.1-4, 2012. [Online]. Available: https://ieeexplore.ieee.org/document/6232834.

[82] K. W. Su, Y. M. Sheu, C. K. Lin, S. J. Yang, W. J. Liang, X. Xi, C. S. Chiang, J. K. Her, Y. T. Chia, C. H. Diaz and C. Hu, "A scalable model for STI mechanical stress effect on layout dependence of MOS electrical characteristics," in *Proceedings of IEEE Custom Integrated Circuits Conference*, pp.245-248, 2003.

[83] G. Pham, M. Ritter and M. Pfost, "Influence of metallization layout on aging detector lifetime under cyclic thermo-mechanical stress," in *IEEE International Reliability Physics Symposium (IRPS)*, pp.5B-5, 2016.

[84] R. B. R. van Silfhout, W. D. van Driel, Y. Li, M. A.J. van Gils, J. H. J. Janssen, G. Q. Zhang, G. Tao J. Bisschop and L. J. Ernst, "Effect of metal layout design on passivation crack occurrence using both Experimental and simulation techniques," in *Proceedings of 5th International Conference on Thermal and Mechanical Simulation and Experiments in Microelectronics and Microsystems, 2004. EuroSimE*, pp.69-74, 2004.

[85] F. Cacho, S. Orain, G. Cailletaud and H. Jaouen, "A constitutive single crystal model for the silicon mechanical behavior: applications to the stress induced by silicided lines and STI in MOS technologies," in *Proceedings of 6th International Conference on Thermal, Mechanical and Multi-Physics Simulation and Experiments in Micro-Electronics and Micro-Systems (EuroSimE)*, pp.635-640, 2005.

[86] D.A.B. Oliveira and M.P. Viana, "Fast CNN-based document layout analysis," in *Proceedings of the IEEE International Conference of Computer Vision Workshops*, pp.1173-1180, 2017.

[87] Y. Wang, M. Sonogashira, A. Hashimoto and M. IIYAMA, "Stroke recovery of handwritten chinese character using fully convolutional networks," in *IEICE Technical Report*, PRMU2018-62, Sep. 2018.

[88] J. Redmon and A. Farhadi, "YOLO9000: Better, faster, stronger," presented at *IEEE Conference Computer Vision and Pattern Recognition*, pp.6517-6525, 2017.

[89] W. Wang, D. Liang, L. Lin, H. Hu, Q. Zhang, Q. Chen, Y Iwamoto, X. Han and Y.W. Chen, "Computer-aided diagnosis of liver cancers using deep learning with fine-tuning," in *IEICE Technical Report*, PRMU2018-57, Sep. 2018.

[90] A. Alfarrarjeh, D. Trivedi, S.H. Kim and C. Shahabi, "A deep learning approach for road damage detection from smartphone images," in *IEEE International Conference on Big Data*, pp.5201-5204, 2018.

[91] B. Kloter, "Application of machine learning for production optimization," in *IEEE 7th World Conference on Photovoltaic Energy Conversion (WCPEC)*, pp.3489-3491, 2018.

[92] A. L. Huyet and J. L. Paris, "Optimizing simulated manufacturing systems using machine learning coupled to evolutionary algorithms," in *Proceedings of 8th International Conference on Emerging Technologies and Factory Automation (ETFA)*, pp.17-21, Vol.1, 2001. Cat. No.01TH8597.

[93] D. Y. Liao, C. Y. Chen, W. P. Tsai, H. T. Chen, Y. T. Wu and S. C Chang, "Anomaly detection for semiconductor tools using stacked autoencoder learning," in *International Symposium on Semiconductor Manufacturing (ISSM)*, pp.1-4, 2018. [Online]. Available: https://ieeexplore.ieee.org/document/8651179.

[94] R. Mih, "Trends in manufacturing productivity and yield enhancement for interconnected devices and industries," in *IEEE 2nd Electron Devices Technology and Manufacturing Conference (EDTM)*, pp.251-253, 2018.

[95] P. Tong, J. Lu and K. Yun, “Fault detection for semiconductor quality control based on Spark using data mining technology,” in *Chinese Control and Decision Conference (CCDC)*, pp.4372-4377, 2018.

[96] Y. Suzuki, S. Iwashita, T. Sato, H. Yonemichi, H. Moki and T. Moriya, “Machine learning approaches for process optimization,” in *International Symposium on Semiconductor Manufacturing (ISSM)*, pp.1-4, 2018. [Online]. Available: https://ieeexplore.ieee.org/document/8651142.

[97] Z. Huang, V .C. Angadi, M. Danishvar, A. Mousavi and M. Li, “Zero defect manufacturing of microsemiconductors - An application of machine learning and artificial intelligence,” in *5th International Conference on Systems and Informatics (ICSAI)*, pp.449-454, 2018.

[98] A. A. Popov, D. V. Bilevich, A. S. Salnikov, A. A. Kalentyev and D. V. Garays, “Plasma-chemical etching process behavioral models based on tree ensembles and neural network,” in *IEEE XIV International Scientific-Technical Conference on Actual Problems of Electronics Instrument Engineering (APEIE)*, pp.128-131, 2018. doi: 10.1109/APEIE.2018.8546139.

[99] B. Bringoltz, E. Rothstein, I. Rubinovich, Y. Kim, N. Tal, O. Cohen, S. Yogev, A. Broitman, E. Rabinovich and T. Zaharoni, “Machine learning and big data in optical CD metrology for process control,” in *e-Manufacturing & Design Collaboration Symposium (eMDC)*, pp.1-4, 2018. [Online]. Available: https://ieeexplore.ieee.org/document/8514236.

[100] T. Nishimi, Y. Sato, S. Kajihara and Y. Nakamura, “Good die prediction modelling from limited test items,” in *IEEE International Test Conference in Asia (ITC-Asia)*, pp.115-120, 2018. doi: 10.1109/ITC-Asia.2018.00030.

[101] F. Lin, A. Ahmadi, K. Sekar, Y. Pan and K. Huang, “IC layout weak point effectiveness evaluation based on statistical methods,” in *IEEE*

36th VLSI Test Symposium (VTS), pp.1-6, 2018. [Online]. Available: https://ieeexplore.ieee.org/document/8368655.

[102] W. T. Cheng, R. Klingenberg, B. Benware, W. Yang, M. Sharma, G. Eide, Y. Tian, S. M. Reddy, Y. Pan, S. Fernandes and A. Chittora., "Automatic Identification of Yield Limiting Layout Patterns Using Root Cause Deconvolution on Volume Scan Diagnosis Data," in *IEEE 26th Asian Test Symposium (ATS)*, pp.219-224, 2017.

[103] C. J. Shan, P. Babghian, L. C. Wang, "Systematic defect detection methodology for volume diagnosis: A data mining perspective," in *IEEE International Test Conference (ITC)*, pp.1-10, 2017 [Online]. Available: https://ieeexplore.ieee.org/document/8242050.

[104] S. Cheon, H. Lee, C. O. Kim and S. H. Lee, "Convolutional neural network for wafer surface defect classification and the detection of unknown defect class," in *IEEE Transactions on Semiconductor Manufacturing*, pp.163-170, Vol. 32, No.2, 2019 [Online]. Available: https://ieeexplore.ieee.org/document/8657760.

[105] S. Schrunner, O. Bluder, A. Zernig, A. Kaestner and R. Kern, "A comparison of supervised approaches for process pattern recognition in analog semiconductor wafer test data," in *IEEE 17th International Conference on Machine Learning and Applications (ICMLA)*, pp.820-823, 2018.

[106] P. Tulala, H. Mahyar, E. Hhalebi and R. Grosu, "Unsupervised wafermap patterns clustering via variational autoencoders," in *International Joint Conference on Neural Networks (IJCNN)*, pp.1-8, 2018. [Online]. Available: https://ieeexplore.ieee.org/document/8489422.

[107] K. Nakata, R. Orihara, Y. Mizuoka and K. Takagi, "A comprehensive big-data-based monitoring system for yield enhancement in semiconductor manufacturing," in *IEEE Transactions on Semiconductor Manufacturing*, pp.339-344, Vol. 30, 2017. doi: 10.1109/TSM.2017.2753251.

[108] Y. Lin, M. Li, Y. Watanabe, T. Kimura, T. Matsunawa, S. Nojima and D. Z. Pan, "Data efficient lithography modeling with

transfer learning and active data selection," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, pp.1900-1913, Vol. 38, No. 10, 2019 [Online]. Available: https://ieeexplore.ieee.org/document/8428441.

[109] B. Lantz, "Machine learning in practice in Machine Learning with R 2nd edition," in *Birmingham, UK*, pp.16-21, 2015.

[110] M. Shin and J.H Lee, "CNN based lithography hotspot detection," in *International Journal of Fuzzy Logic and Intelligent Systems (IJFIS)*, Vol. 16, No.3, pp.208-215, 2016. doi: 10.5391/IJFIS.2016.16.3.208.

[111] H. Yang, J. Su, Y. Zou, Y. Ma, B. Yu and E. F. Y. Yong, "Layout hotspots detection with feature tensor generation and deep biased learning," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, pp.1175-1187, 2018 [Online]. Available: https://ieeexplore.ieee.org/document/8360060.

[112] J. R. Gao, B. Yu, D. Ding and D. Z. Pan, "Lithography hotspot detection and mitigation in nanometer VLSI," in *IEEE 10th International Conference on ASIC*, pp.1-4, 2013 [Online]. Available: https://ieeexplore.ieee.org/document/6811917.

[113] Y. Tomioka, T. Matsunawa, C. Kodama and S. Nojima, "Lithography hotspot detection by two-stage cascade classifier using histogram of oriented light propagation," in *22nd Asia and South Pacific Design Automation Conference (ASP-DAC)*, pp.81-86, 2017.

[114] Y. Lin, X. Xu, J. Ou and D. Z. Pan, "Machine learning for mask/wafer hotspot detection and mask synthesis," in *Photomask Technology 2017. International Society for Optics and Photonics*, Vol. 10451, p. 104510A, 2017 [Online]. Available: https://doi.org/10.1117/12.2282943.

[115] R. Chen, W. Zhong, H. Yang, H. Geng, X. Zeng and B. Yu, "Faster region-based hotspot detection," in *56th ACM/IEEE Design Automation Conference (DAC)*, pp.1-6, 2019.

[116] B. Jiang, J. Yang, H. Zhang and E. F. Y. Young, "A fast machine learning-based mask printability predictor for OPC acceleration," in *Proceedings of 24th Asia and South Pacific Design Automation Conference*, pp.412-419, 2019. [Online]. Available: https://doi.org/10.1145/3287624.3287682.

[117] W. Ye, M. B. Alawieh, M. Li, Y. Lin and D. Z. Pan, "Litho-GPA: Gaussian process assurance for lithography hotspot detection," in *Design, Automation & Test Europe Conference & Exhibition (DATE 19)*, pp.54-59, 2019.

[118] I. Sutskever, J. Martens, G. Dahl and G. Hinton, "On the importance of initialization and momentum in deep learning," in *International Conference on Machine Learning*, pp.1139-1147, 2013.

[119] P. Krahenbuhl, C. Doersch, J. Donahue and T. Darrell, "Data-dependent initializations of convolutional neural networks," in *arXiv preprint arXiv:1511.06856*, 2015. [Online]. Available: https://arxiv.org/abs/1511.06856v3.

[120] T. R. Scott, K. Ridgeway and M. C. Mozer, "Adapted deep embeddings: A synthesis of methods for k-shot inductive transfer learning," in *arXiv preprint arXiv:1805.08402*, 2018 [Online]. Available: https://arxiv.org/abs/1805.08402.

[121] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *arXiv preprint arXiv:1409.1556*, 2014. [Online]. Available: https://arxiv.org/abs/1409.1556v6.

[122] An End-to-End Open Source Machine Learning Platform. Accessed: Jun. 16, 2018. [Online]. Available: https://www.tensorflow.org/.

[123] J. H. Kim, B. G. Kim, P. P. Roy and D. M. Jeong, "Efficient facial expression recognition algorithm based on hierarchical deep neural network structure," in *IEEE Access*, pp.41273-41285, Vol. 7, 2019. doi: 10.1109/ACCESS.2019.2907327.

[124] S. Darshan S. L and Jaidhar C.D, "Windows malware detector using convolutional neural network based on visualization images," in *IEEE Transactions on Emerging Topics in Computing*, 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8685181.

[125] K. Kitaguchi, Y. Nishizaki and M. Saito, "Analysis of the region of interest in visual inspection using CNN," in *IEICE Technical Report*, PRMU2018-59, 2018.

[126] M. Ancona, E. Ceolini, C. Oztireli and M. Gross, "Towards better understanding of gradient-based attribution method for deep neural networks," in *arXiv preprint arXiv:1711.06104*, [Online]. Available: https://arxiv.org/abs/1711.06104.

[127] B. Devika and N. George, "Convolutional neural network for semiconductor wafer defect detection," in *International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, pp.1-6, 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8944584.

[128] T. Nakazawa and D. V. Kulkarni, "Anomaly detection and segmentation for wafer defect patterns using deep convolutional encoder-decoder neural network architectures in semiconductor manufacturing," in *IEEE Transactions on Semiconductor Manufacturing*, pp.250-256, Vol. 32, No. 2, 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8634922.

[129] N. Yu, Q. Xu and H. Wang, "Wafer defect pattern recognition and analysis based on convolutional neural network," in *IEEE Transactions on Semiconductor Manufacturing*, pp.566-573, Vol. 32, No.4, 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8815875.

[130] K. Imoto, T. Nakai, T. Ike, K. Haruki and Y. Sato, "A CNN-based transfer learning method for defect classification in semiconductor manufacturing," in *IEEE Transactions on Semiconductor Manufacturing*, pp.455-459, Vol. 32, No.4, 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8839832.

[131] M. R. Nishat, U. Katakamsetty, V. Mehrotra, H. Landis, S. Nakagawa and G. Huda,, “ Topography hotspot-aware process control metrology cell insertion: DFM: Design for manufacturability,” in *2018 29th Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC)*, pp.234-237, 2018. doi: 10.1109/ASMC.2018.8373199.

[132] A. Sinha, S. Pandey, A. Singhal, A. Sanyal and A. Schmaltz, “DFM-aware fault model and ATPG for intra-cell and inter-cell defects,” in *2017 IEEE International Test Conference (ITC)*, pp.1-10, 2017. doi: 10.1109/TEST.2017.8242059.

[133] Y. Lin, Y. Watanabe, T. Kimura, T. Matsunawa, S. Nojima, M. Li and D. Z. Pan, “Data efficient lithography modeling with residual neural networks and transfer learning,” in *Proceedings of the 2018 International Symposium on Physical Design*, pp.82-89, 2018. [Online]. Available: https://doi.org/10.1145/3177540.3178242.

[134] H. Yang, Y. Lin, B. Yu and E. F. Y. Young, “Lithography hotspot detection: From shallow to deep learning,” in *2017 30th IEEE International System-on-Chip Conference (SOCC)*, pp.233-238, 2017. doi: 10.1109/SOCC.2017.8226047.

[135] N. Rai, H. Hashempour, Y. Xing, B. Kruseman and S. Hamdioui, “ A schematic-based extraction methodology by dislocation defects in analog/mixed-signal devises,” in *IEEE International Symposium on Defect and Fault Tolerance in VLSI and Nanotechnology Systems*, pp.139-145, 2011.

[136] P. Arena, L. Fortuna, G. Manganaro and S. Spina, “CNN image processing for the automatic classification of oranges,” in *Proceedings of the Third IEEE International Workshop on Cellular Neural Networks and their Applications (CNNA-94)*, pp.463-467, 1994. doi: 10.1109/CNNA.1994.381631.

[137] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, “You only look once: unified, real-time object detection,” in *2016 IEEE Conference on*

Computer Vision and Pattern Recognition (CVPR), pp.779-788, 2016. doi: 10.1109/CVPR.2016.91.

[138] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar and L. Fei-Fei,, "Large-scale video classification with convolutional neural networks," in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pp.1725-1732, 2014. doi: 10.1109/CVPR.2014.223.

[139] Y. Nagamura, T. Ide, M. Arai and S. Fukumoto, "CNN-based layout segment classification for analysis of layout-induced failures," in *IEEE Transactions on Semiconductor Manufacturing*, pp.597-605, Vol. 33, No.4, 2020. doi: 10.1109/TSM.2020.3029049.

[140] S. Hayashi and O. Hasegawa, "A detection technique for degraded face images," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.1-5, 2006. [Online]. Available: https://ieeexplore.ieee.org/document/1640935.

[141] G. E. Dahl, T. N. Sainath and G. E. Hinton, "Improving deep neural networks for LVCSR using rectified linear units and dropout," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp.8609-8613, 2013.

[142] A. Chattopadhay, A. Sarkar, P. Howlader and V. N. Balasubramanian, "Grad-CAM++: Generalized gradient-based visual explanations for deep convolutional networks," in *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp.839-847, 2018.

[143] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh and D. Batra,, "Grad-CAM: Visual explanations form deep networks via gradient-based localization," in *Proceedings of the IEEE International Conference on Computer Vision*, pp.618-626, 2017.

[144] S. Venkataraman and S. B. Drummonds, " POIROT: a logic fault diagnosis tool and its applications," in *Proceedings of International Test Conference 2000 (IEEE Cat. No.00CH37159)*, pp.253-262, 2000. doi: 10.1109/TEST.2000.894213.

[145] Automotive Electronics Council, "Fault simulation and fault grading," in *document AEC-Q100-007 Rev-B*, Sep. 2007.

[146] T. W. Williams and N. C. Brown, "Defect level as a function of fault coverage," in *IEEE Transactions on Computers*, pp.987-988, Vol. C-30, No. 12, 1981.

[147] J. T. Sousa et al., "Defect level evaluation in an IC design environment," in *IEEE Transactiosn on Computer-Aided Design Integration Circuits Systems*, pp.1286-1293, Vol. 15, No. 10, 1996.

[148] K. Shiozawa and H. Ito, "A study of effective detection rate based on net length distribution," in *Proceedings of 34th Annual NANO Testing Symposium*, pp.75-80, 2014.

[149] M. Arai, Y. Shimizu, and K. Iwasaki, "Note on layout-aware weighted probabilistic bridge fault coverage," in *Proceedings of IEEE 21st Asian Test Symposium*, pp.89-94, 2012.

[150] P. Maxwell, F. Hapke, M. Ryynnen and P. Weseloh, "Bridge over troubled waters: critical area based pattern generation," in *2017 22nd IEEE European Test Symposium (ETS)*, pp.1-6, 2017.

[151] Masayuki Arai, Shingo Inuyama and Kazuhiko Iwasaki, "Layout-aware 2-step window-based pattern reordering for fast bridge/open test generation," in *2017 IEEE International Test Conference (ITC)*, pp.1-8, 2017.

Acknowledgements

This dissertation was the result of the research and development on the engineering of LSI layout analysis for a period of three years, from 2018 to 2020, in the working doctor's programs (latter period) in Tokyo Metropolitan University (TMU) and many years prior to that in Renesas Electronics Corporation (RENESAS). I am grateful for the support and guidance of several people.

First and foremost, I would like to thank my professors at TMU. I would like to thank my PhD advisor, Prof. Satoshi Fukumoto for his guidance and support throughout my doctor's program in TMU. Because of his valuable support and advice, I have been able to successfully publish several papers in journals and in an academic meeting. In every seminar in the laboratory, he spent a good deal of time to discuss about my research and suggested a right direction. I would like to thank Prof. Kazuhiko Iwasaki, who retired this year, for his invaluable assistance in the development of the physical test coverages and for his advice on the path to my doctorate degree. I would like to thank Associate Prof. Kiyoshi Nishikawa, Associate Prof. Kan Okubo, and Prof. Yukiya Miura in the Faculty of System Design from TMU for their various discussions on my research in end-term evaluation meetings.

I would like to thank Prof. Hisashi Yamamoto in the Faculty of System Design from TMU and Prof. Toshinori Hosokawa in Department of Mathematical Information Engineering at College of Industrial Technology from Nihon University for their insightful suggestions and helpful comments in the review meetings and the public hearing of this dissertation.

I would like to thank Dr. Masayuki Arai (Associate Prof. from Nihon University) for his advice and support to proceed with my research. He supported me throughout my research period, from the physical test coverage to the layout segment classification. I would like to specially appreciate his help for revising the drafts for journal publication.

I would like to thank Dr. Kazuya Sakai (Associate Prof. from TMU) for his suggestions on my research at the laboratory seminar and for his advice on machine learning.

I would also like to thank the members of RENESAS for supporting my research. Dr. Takshi Ide (Former Director of Analysis and Evaluation Technology Department (AETD)), Mr. Koji Arima (Director of AETD), Mr. Fumihito Ota (Senior Manager of AETD), Dr. Tomohiro Yamashita (Vice President of Production and Technology Unit (PTU)), and Dr. Takio Ono (former Vice President of PTU) helped with my research. I would also like to thank Dr. Fumihiko Hayashi (retired former Vice President of PTU) and Dr. Shuichi Inoue (retired former Vice President of PTU) for their support and encouragement to develop new techniques.

I appreciate Mr. Tomomi Ukai (Principal Engineer from AETD) who conducted OBIRCH failure positioning to offer the defect data used in my research on layout analysis for failure analysis. He offered several suggestions with regard to the defects that occurred on silicon devices. I would like to thank Mr. Nobuo Fukuzawa (retired Engineer from AETD) who supported the consideration of machine learning and helped me use high-spec computers. I would like to thank Mr. Tsuyoshi Miyatake (Senior Manager from Equipment & Productivity Engineering Department), who gave me an

opportunity to have technical discussions with many experts on AI technologies.

I also thank Dr. Kenji Shiozawa (former member of AETD) and Dr. Toru Koyama (former director of AETD) for providing useful information relevant to layout analysis and helpful discussions for furthering my research.

For the research on the physical test coverage, I appreciate Mr. Jun Matsushima (Principal Engineer from Digital Design Technology Department), Mr. Kazuhiro Tomonaga (Senior Manager from Test Integration Department), Mr. Yutaka Hoshi (retired Engineer from Renesas Engineering Service (REG)), and Mr. Shuji Nomura (Engineer from REG) for their help and discussions on research and development of internal operation for reliability using physical test coverage. I appreciate the endeavors of all the retired members who developed DFM in RENESAS and the members of STARC for their support. They helped me understand the basics for the study of physical test coverage.

I thank the members of the TMU for their support and meaningful discussions at laboratory seminars. Mr. Yuki Takeuchi supported me in studying machine learning. Mr. Shingo Yokoyama set a special computer containing GPUs for use in image classification evaluation. Mr. Ryota Mizuhara helped me with programming. Mr. Yuya Fukuchi provided his deep insights and discussions at the seminar. I especially appreciate Mr. Kai Murakawa and Mr. Itsuki Fujita, members of the team of image classification, for their help in basic evaluation to classify LSI layout images and demonstrate the properties of convolutional neural networks (CNNs). The information was helpful and helped me consider the use of CNNs to classify LSI layout images

appropriately. Ms. Yumi Sano and Ms. Yuko Kikuchi kindly supported my administrative paper work at the university. I would like to thank all of them for their help and support.

Finally, I was supported by my family during the period of this research. I dedicate this thesis to my wife Yoko and my son Kan with my greatest appreciation. Thank you very much.