공학박사학위논문

딥러닝을 이용한 반도체 리소그래피 공정에서의 핫스팟 검출 및 예측

Detection and prediction of semiconductor lithography hotspots using deep learning models

2022 년 8 월

서울대학교 대학원

기계항공공학부

김 재 훈

딥러닝을 이용한 반도체 리소그래피 공정에서의 핫스팟 검출 및 예측

Detection and prediction of semiconductor lithography hotspots using deep learning models

지도교수 김 도 년

이 논문을 공학박사 학위논문으로 제출함

2022 년 4 월

서울대학교 대학원

기계항공공학부

김 재 훈

김재훈의 공학박사 학위논문을 인준함

2022 년 6 월

위 원 장 김 윤 영 (인)

부위원장 김 도 년 (인)

위 원 박 종 우 (인)

위 원 윤 병 동 (인)

위 원 김 기 현 (인)

Abstract

Since the invention of transistors and integrated circuits, advances in semiconductor processes have grown rapidly over the decades. Current microchips contain hundreds of millions of transistors. Due to the increased complexity of semiconductor process technology and device structure, many defects called hotspots occurs in the semiconductor manufacturing process. In order to improve yields and reduce costs, it is essential not only to detect hotspots quickly and efficiently, but also to predict whether hotspots occur in advance before semiconductor lithography process. In particular, predicting hotspots during the design stage is essential for efficient productivity in the semiconductor industry. Therefore, research fields focusing on these problems have gained much attention. This paper presents a deep learning-based hotspot detecting and prediction model that outperforms standard methods for more efficient semiconductor process development and manufacture.

Our proposed hotspot detection model utilized the architecture of the conditional Generative Adversarial Network (cGAN) to generate a more precise detection heat map, allowing small-sized hotspots to be detected more accurately and faster than the conventional hotspot detection model. In this research, newly developed hotspot prediction model is proposed. Proposed hotspot prediction model combines segmentation models and style transfer models in parallel based on cGAN. Our developed hotspot prediction model helps predict hotspots through the Scanning Electron Microscopy (SEM) image generated by translation from layout design. In addition, hotspot prediction model proposed in this research uses the constrained partial cross entropy loss function to guess which part of the layout patterns are involved in hotspots. It is expected that more efficient semiconductor process development will be possible because this will greatly help layout pattern design. The performance of the hotspot detection and prediction model presented in this study was tested with real industrial dataset and showed better performance than the sssconventional models.

Keyword : semiconductor lithography process, hotspot detection model, hotspot prediction model, deep learning, hotspot, defect

Student Number : 2016-3779

Contents

List of Figures

predictions and the correlation between G.T and normal pattern. This shows that predictions of our model are highly correlated with to G.T, real hotspots.

Figure 4.9. Visualization in two-dimensional feature map utilizing T-distributed Stochastic Neighbor Embedding (T-SNE) with extra model features. This figure shows that the model's predictions are closer to G.T., while at the same time roughly showing that hotspots (prediction and real), have different latency values than normal patterns.

Figure 4.10. Illustration of layout designs, SEM image with ground-truth bounding boxes and ground-truth segmentation images

Figure 4.11. Experiment results of hotspot prediction model on the industrial test dataset. The third and fourth columns show segmentation prediction map and generated SEM images, respectively. Segmentation prediction maps predicted by Constrained partial cross entropy loss allow us to guess which patterns are associated with hotspots in layout designs. From the segmentation map of the third column, we can guess that two long horizontal patterns in the layout design were associated with hotspots.

List of Tables

Chapter 1

Introduction

1.1 Background and objectives

The one of most important process in the semiconductor process is lithography. To produce a circuit, a process known as lithography is used to imprint the desired pattern on the surface of the wafer. When light passing through the mask is received after photoresist coating, the polymer organic solution causes a chemical reaction, and desired pattern is drawn on the wafer. At this time, the mask is manufactured in four times larger size than the wafer to draw a finer pattern, refracts light through the lens or reflect the mirror, and irradiates light according to the chip size. Since the semiconductor manufacturing process is very expensive, as many chips as possible must be manufactured in a single wafer.

Due to the complicated lithography process, patterning defects called hotspots often occurs during semiconductor process. All these hotspots must be fully detected during the lithography development and manufacturing process. In addition, in order to create an effective lithography process, it is also essential to predict hotspots during the pattern design phase. It is very difficult to predict the occurrence of hotspots in advance without a lithography process. Many studies are currently underway because this will greatly help develop a more efficient lithography process.

Scanning Electron Microscope (SEM) images can be automatically analyzed and reviewed with a tool called Review Scanning Electron Microscope (Review SEM), which can also determine the position and type of hotspots in the image. However, The disadvantage of this method is that it requires a reference SEM image or the extraction of a contour from the SEM image. A new approach to addressing

the problems with the detection of hotspots has been made possible by recent developments in deep learning algorithms. Deep learning model can learn important features of hotspots as they are trained, as opposed to using predetermined features. Although there have been several performance enhancements thanks to deep learning, many models now have the drawback of utilizing simply SEM images or layout design and not classifying the sorts of hotspots.

In general, hotspot prediction is largely classified into rule-based, model-based and machine-learning-based methods. In the rule-based method, hotspots are predicted by matching the pattern with the patterns that are already known to contain hotspots. Since the rule-based method is considerably affected by the user's decision and the model-based method takes a long time and has the disadvantage that the performance depends heavily on the model calibration, Methods based on machine learning have been studied to overcome this issue. However, there is a limit to performance because it still relies on handcrafted features. Recent advances in convolutional natural networks have significantly enhanced the overall performance of deep learning-based vison models including object classification, object detection, super-resolution, style transfer, segmentation, and anomaly detection. This tendency has led to numerous studies on the deep learning-based hotspot prediction model. However, most hotspot prediction conventional models only predict hotspots in clip layouts and do not predict the exact location of hotspots. Accordingly, in the large-scale layout, the window sliding method is applied to predict hotspots for each clip and then aggregate them, which leads to a disadvantage that it takes a long time to predict.

In this study, we focus on developing hotspot detection and prediction models based on deep learning. The hotspot detection model we aim to develop should be able to detect very small-sized hotspots with high accuracy and classify whether the detected hotspot is hard defect or soft defect. In case of the hotspot prediction model, it should be possible to predict where hotspots occur at the design stage. In addition, beyond simply predicting the location of the hotspot, this model should generate SEM image from layout design so that you can see in advance what form the hotspot

will occur. It is also expected that more efficient lithography process development will be possible by suggesting which layout patterns caused hotspots.

1.2. Research outline

In Chapter **1**, we present background objectives and research outline of this study.

In Chapter **2**, the idea of the photolithography process and the defects known as hotspots, which are critical errors in the semiconductor manufacturing process, is presented. Hotspots can result in significant financial and time losses. In order to build an effective semiconductor process, it is crucial to identify and predict these hotspots. In this chapter, we describe the limits of conventional hotspot detection and prediction methods.

In Chapter **3**, we suggest a deep learning-based method for detecting hotspots and is effective at locating even the tiniest of hotspots. It employ a conditional generative adversarial network as the basis for its design of the adversarial network. When an adversarial network architecture is used, the discriminator lends support to the detection model so that it can accurately locate hotspots. In order to attain a greater level of performance than the baseline, CenterNet, a discriminator employing the PixelGAN technique and multi-scale level features is applied. The effectiveness of the model in locating hotspots was validated by using data taken from real-world industrial settings. We have high hopes that the hotspot detection methodology will result in an improvement to the semiconductor manufacturing as well as the development processes. Obtaining the datasets required to develop a model for predicting hotspots will also be made easier by hotspot detection model.

In Chapter **4**, We present a hotspot prediction model that is based on deep learning and can make accurate predictions of hotspots in layout patterns during the design stage of the semiconductor process. Our model integrated the segmentation model with the style transfer model using a conditional generative adversarial network. Through our model, we can directly check the location and appearance of the hotspots through the generated SEM image translated from the layout design. In addition, with the help of constrained partial cross entropy loss, it is also possible to roughly guess which part of layout patterns causes hotspots. The performance of our

model was verified by testing with real industrial data that it shows better performance than baseline models. The development of semiconductor lithography process will be more efficient by predicting hotspots during the layout design stage using hotspot prediction model.

Chapter 2

Photolithography process and hotspots

2.1 Introduction

Integrated circuits are used in the production of many modern electronic products, including smartphones, smartwatches, action cameras, gaming consoles, vehicles, and laptop computers. These integrated circuits are incredibly small yet very sophisticated. An integrated circuit made of semiconductors is a type of electronic component that combines a number of different devices onto a single chip in order to process and store a variety of different operations. Making a wafer, which is a large disk in the shape of a disk, out of silicon, which is the primary component of semiconductors, and then engraving a circuit onto the wafer before cutting it to the appropriate size for the product is the most straightforward method for producing semiconductor integrated circuits. It is achieved through the completion of a number of stages, including wafer processing, oxidation, photolithography process, and etching. Figure 2.1 describes semiconductor process briefly.

The term "wafer" refers to a disk that is produced when a single crystal pillar composed of Silicon (Si), Gallium Arsenide (GaAs), or another material is sliced to the desired thickness. In order to produce a semiconductor integrated circuit, the wafer is covered with a number of parallel circuits that are all the same. The process of obtaining wafers is referred to as "wafer processing" in the semiconductor industry. The production of ingots is the initial stage of the wafer manufacturing process. A purification process that results in an increase in purity is required before silicon that has been recovered from sand can be used as a material for semiconductors. Silicon

needs to be dissolved in order to make a solution that has a high purity level. And then crystallize it and make it as solid as possible. Ingots are the name given to the columns of silicon that are produced in this manner. In order to transform an ingot into a wafer in the shape of a disk, it is required to cut the ingot into even and thin slices using diamonds. The thinner the wafer, the lower the production cost, and the wider the diameter, the greater the number of semiconductor chips that can be made simultaneously; therefore, as technology improves, the wafer becomes thinner and larger. The surface of the wafer is then slickly polished using a polishing machine.

This thin, spherical wafer is in a state of nonconductivity. Because of this, it is essential to give the wafer a conducting as well as a nonconducting state. The oxidation process is the stage that is considered to be the most essential of this activity. The formation of SiO2, which performs the function of an insulating film on the wafer, is what makes the oxidation process necessary in order to stop current from leaking out of the circuits. The oxide coating also serves the purpose of an anti-etching film, which prevents the essential components from being etched wrongly. In addition, the oxidation process plays a part in the formation of a protective film on the surface of the wafer. It is able to protect the wafer from a variety of possible dangers, including chemical contaminants, diffusion during ion implantation, and slipping off during the etching process.

The photolithography process is one of the most vital steps in the fabrication of semiconductors. It is called photolithography because light is used to draw a circuit on a wafer using a mask which is also called a reticle. In this process, the formation of a pattern is analogous to the printing of a black-and-white photograph from a film onto printed paper. The greater the number of transistors in a dense integrated circuit, the greater the probability of defects in the semiconductor process; therefore, the photolithography process must also be executed with care and advanced technology. The specifics are described in Chapter 2.2.

The etching procedure is the next phase in the process. This process, which is analogous to the etching method of engraving and makes use of chemical reactions like corrosion, is used to generate semiconductor circuit layouts by selectively

deleting superfluous sections from a wafer. Etching is a form of engraving that was developed in the 19th century. The remaining photoresist is etched away using etchants, leaving only the circuit that was generated during the photographic process. The etching process can be broken down into wet and dry categories according to the etching reaction. Wet etching is using a chemical reaction and a solution, whereas dry etching is a method of eliminating particular regions of a surface by employing reactive gases, ions, and other such substances. Dry has the drawbacks of being more demanding and expensive than wet, although in recent years, highly integrated semiconductor technology has led to a reduction in the width of the circuit. This has resulted to a narrower footprint for the components. As a result, dry etching is gradually replacing wet etching as the method of choice for increasing yield.

Other stages, including packaging, are also essential in the semiconductor manufacturing process, but they are omitted because they are beyond the scope of this study.

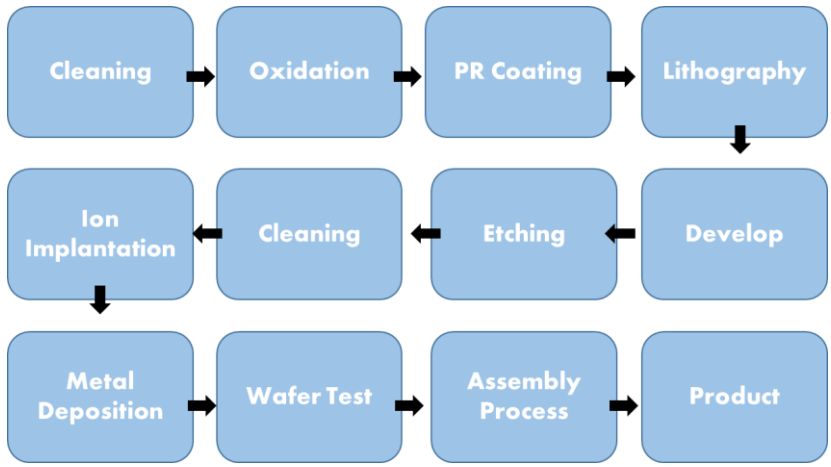

Figure 2.1. Illustration of overall semiconductor process flow

2.2 Photolithography and hotspot

The manufacturing of semiconductors relies heavily on the lithography technique, which involves designing circuits directly onto silicon wafers. In the process of lithography, a light source is shone through a mask that is holding the patterns that are going to be printed on a wafer that has been covered with a photoresist-sensitive substance. After the designs have been developed, the layer behind them will be etched in order to remove the section that has not been masked. The overall process of lithography is depicted in more detail in Figure 2.2.

Semiconductor lithography process has become more complex because of the evolution of technology. This change led to the advance of the semiconductor process from 65 nm to 5 nm. Extreme ultraviolet (EUV) lithography employs a wavelength of 13.5 nm, which, when measured in relation to the dimensions of the node, is regarded as being extremely large. Whenever there is a mismatch between the wavelength and node size in semiconductor process, the result is typically the development of wafer patterns that are undesired.

Patterning defects, called hotspots, often occur during semiconductor lithography process. Hotspot is a defect that occurs during the semiconductor process, where certain patterns in the layout design are not well printed and fatal errors occur. The main cause of hotspots is that the wavelength of light, the source of the lithography process, is much larger than the size of the pattern to be printed, which is called sub-wavelength lithography. Therefore, various Resolution Enhancement Techniques (RET) such as Off-Axis Illumination [1], Sub-Resolution Assist Features (SRAF) [2], Phase Shift Mask (PSM) [3], and Optical Proximity Correction (OPC) [4] have been developed so that the intended design can be printed on the wafer. Phase Shift Mask is a mask device that can improve image resolution by using phase shifters that produce phase differences. Sub-Resolution Assist Feature (SRAF) improves pattern fidelity with features placed next to features that need to be printed on the wafer. Sub-Resolution Assist Features are usually isolated and are not printed

on the wafer. The edges of the patterns in the layout design generate diffraction patterns, and interference from adjacent patterns degrades pattern fidelity, reducing printability. To cope with this unwanted behavior, add additional features to the layout pattern so that the result is closer to the layout design. A conceptual diagram of the Optical Proximity Correction (OPC) is shown in Figure 2.3. Rule-based Optical Proximity Correction (OPC) is a method of using a set of rules empirically learned from a lithography process. On the other hand, model-based Optical Proximity Correction (OPC) uses an optical physics-based simulation model.

Various physical/chemical factors and errors, such as the thickness of the thin film, photosensitivity of the material, and chemical materials used for development, resolution and depth of focus directly or indirectly affect lithography process and can cause hotspots. Appropriate values of theses parameters should be determined through trial and error, which can lead to hotspots. Additionally, hotspots may be caused by random variables, such as non-uniformity of the material and variations in the process.

Another cause of hotspots entails problems concerning the layout design. Hotspots are known to occur from incorrect layout designs; however, knowing exactly which layout patterns are susceptible to hotspots is difficult. Hotspots lead to poor performance or serious problems at the product operation level. If hotspots are found during the semiconductor manufacturing process, huge losses will occur. Therefore, it is imperative to predict hotspots in the early layout design stage and design the layout accordingly to prevent hotspots from occurring, as inspecting the wafer patterns after lithography is an expensive and a time-consuming task. Figure 2.4 describes why hotspot detection and prediction models are important in the lithography process development and manufacturing process. Hotspot detection model will help the development process more efficiently as well as manufacturing process. It will also help to secure the datasets needed to develop hotspot prediction model. Also, the development process will be more efficient by predicting hotspots in layout design stage using hotspot prediction model.

In accordance with the degree of damage they cause, hotspots are further

subdivided into either hard defect or soft defect categories. Hard defects have high possibility of considerable effect on the degree to which the performance of the semiconductor is compromised, whereas soft defects may or may not have any effect. Figure 2.5 illustrates some examples of defects such as these.

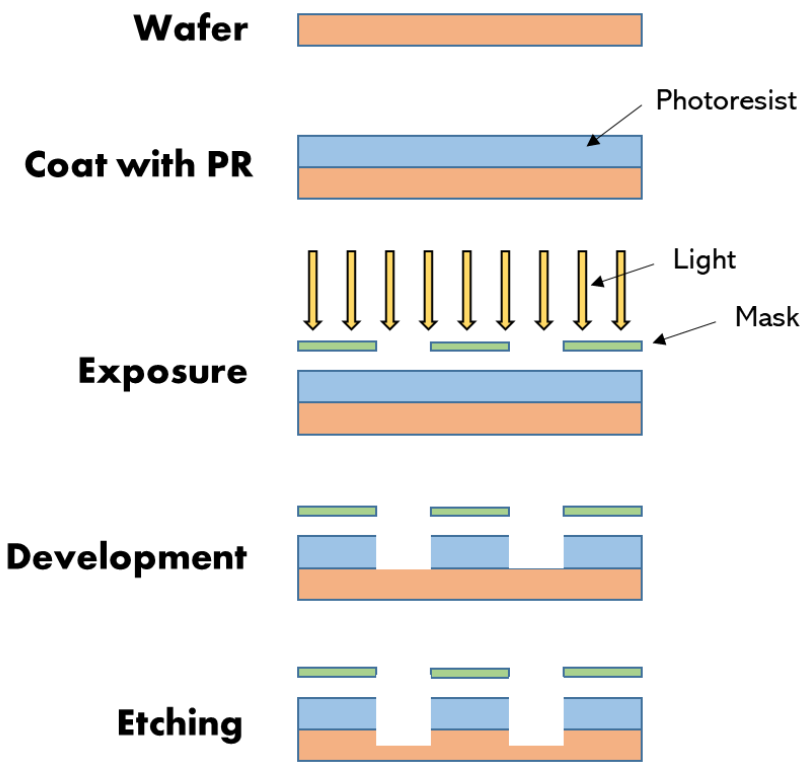

Figure 2.2. Illustration of photolithography process flow

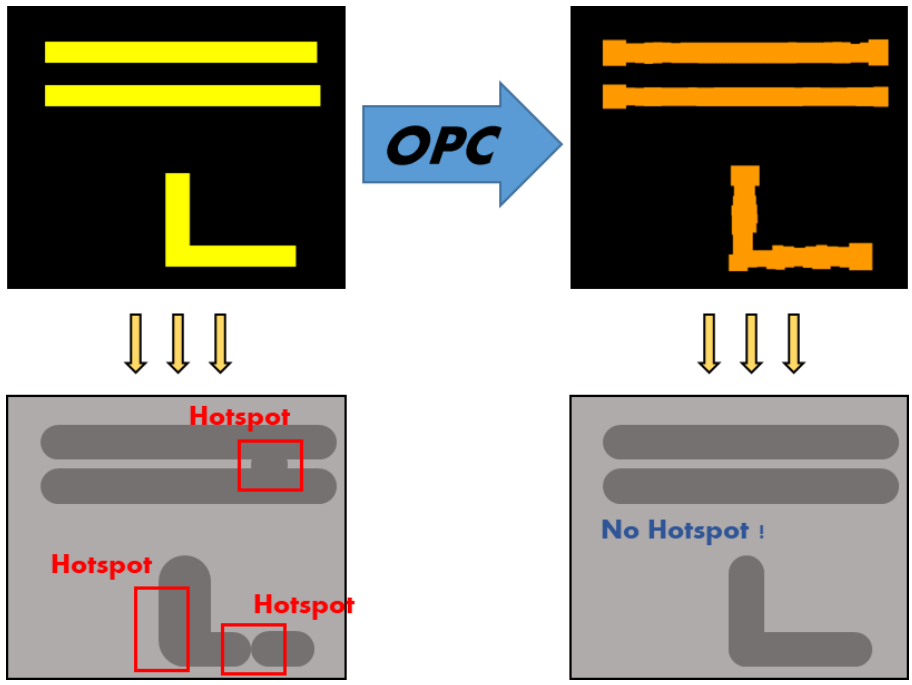

Figure 2.3. Illustration of Optical Proximity Correction (OPC)

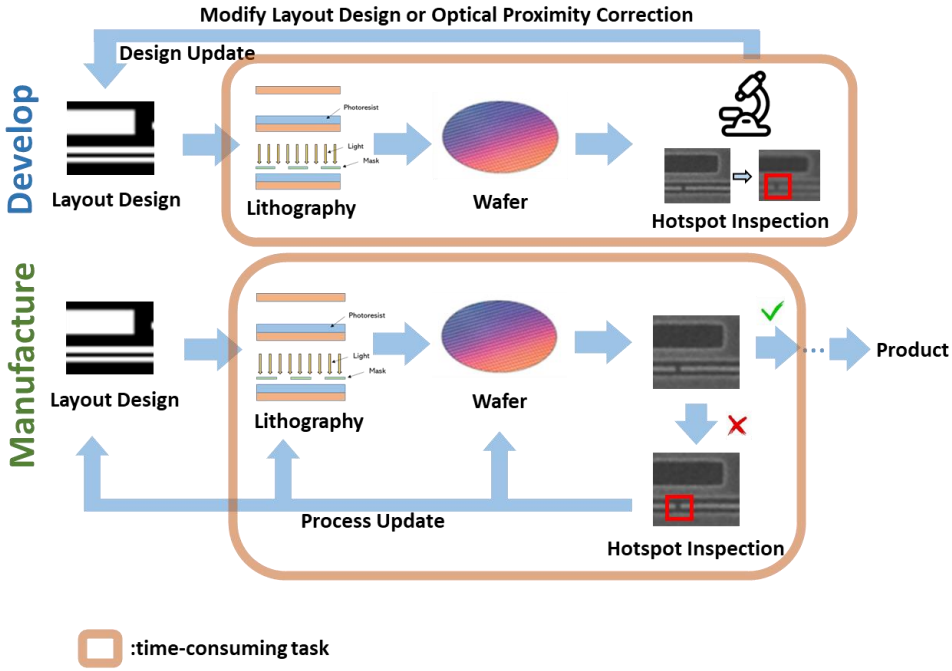

Figure 2.4. Brief Illustration of semiconductor photolithography process for development and manufacturing. In the figure, the tasks of the orange boxes, lithography process and inspection of hotspots, are time-consuming tasks in the process and supplementing them using hotspot detection and prediction models would save a lot of time and money.

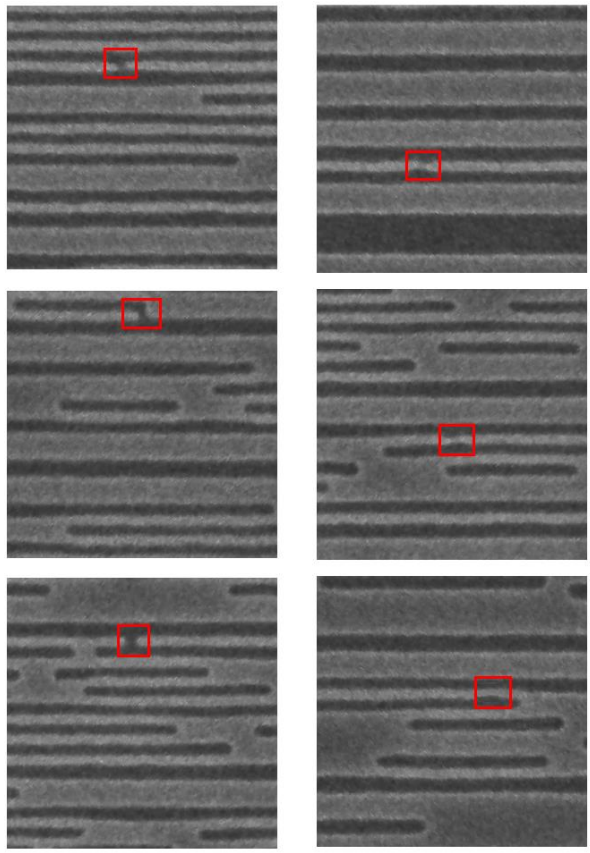

Figure 2.5. Illustration of examples of hotspots, hard defects(left) and soft defects(right) in semiconductor photolithography process

2.3 Conventional method to detect and predict hotspots

2.3.1 Conventional hotspot detection method

As technology advances, modern electronic devices require smaller, more sophisticated integrated circuits, resulting in many defects, or hotspots, in the manufacturing process. Since these hotspots must not occur to manufacture flawless digital products, all these hotspots must be detected from semiconductor process development to manufacturing process. Therefore, hotspot detection methods are being developed for more efficient semiconductor process development and manufacturing process.

The majority of the time, the Review Scanning Electron Microscope (Review SEM) is utilized to perform an analysis of the SEM images of the wafer in order to ascertain the location of hotspots as well as the type of hotspots that are in existence. Die-to-Die inspections (also known as D2D inspections) and Die-to-Database inspections (also known as D2DB inspections) are the two primary procedure types that are utilized by the Review SEM. The conventional approach to D2D inspection involves making comparisons between SEM images and a reference SEM image before classifying the results according to criteria that are saved on the server [5]. This enables the discovery of hotspots. The most significant disadvantage of D2D inspection is that it is not possible to automate the process of picking a reference SEM image. This is the primary reason why D2D inspection is not widely used. Even if a reference SEM image can be easily acquired, it is difficult to ascertain whether the image has been corrupted in any way and whether or not it contains any mistakes such as defects.

When compared to the reference layout design, SEM images are used in D2DB assessment to locate potential hotspots [6]. The contour lines from the SEM image are generally extracted using this method, which is very similar to the D2D approach. The contour lines are then compared to the associated layout plan using

predetermined criteria. However, due to imaging noise, it is challenging to accurately extract contour lines from SEM images. As a result, the method of review calls for an exhaustive and critical examination conducted by knowledgeable professionals for the purpose of additional verification. This examination is labor-intensive, requires a lot of time, and is prone to error.

To address these obstacles, several research have been done. Nakagaki et al. [7] suggested a method for fault identification based on several SEM images. Harada et al. [8] suggested a hotspot categorization model using distribution of typical patches in SEM images. There have also been numerous studies to find patterns with a high likelihood of hotspots using only layout design [9–11]. Recent advancements in deep learning have presented a novel method for addressing the challenges associated with hotspot detection. Rather of employing predetermined features, deep learning model learn critical features from the training dataset and utilize them to predict hotspots. Patel et al. [12] presented convolutional neural model for the categorization and identification of hotspots using SEM images. Imoto has proposed a technique for hotspot categorization that employs transfer learning [13]. Several approaches [14–16] have been suggested the categorization of layout pattern clips with a risk of hotspots. However, some models merely utilize SEM images or layout designs. The second technique is limited in its ability to detect random hotspots, whereas the first method has trouble evaluating pattern-related systematic hotspots. Ouch [17] has suggested a D2DB approach based on deep learning that leverages paired SEM images and layout design to learn luminosity distribution. This model is able to identify both random and systematic hotspots due to its ability to identify outliers as errors in the learnt distribution.

2.3.2 Conventional hotspot prediction method

Predicting hotspots is essential for the development of more efficient semiconductor process. Designing layout patterns without hotspots involves detecting hotspots after photolithography and modifying the layout pattern to prevent these hotspots from occurring, which are heavy, time-consuming and inefficient

processes. Therefore, from the stage of designing layout patterns, a model that can predict the occurrence of hot spots in advance is needed. Therefore, many researchers are struggling to find a solution because predicting the occurrence of hotspots is very difficult.

In general, hotspot prediction is largely classified into rule-based, model-based and machine-learning-based methods. In the rule-based method, the shape of the pattern is varied, and hotspots are predicted by matching the pattern with the pattern in which a hotspot exists. Yao et al. developed a range-pattern matching algorithm and analyzed the shape of the layout on a dual-scale basis to predict hotspots [18]. Kahng et al. created a layout graph that represented the critical dimension deformation and predicted hotspots based on pattern-shape variables [9]. Another way of predicting hotspots is a model-based approach, which utilizes lithography simulation. Rule-based methods are often used for simple patterns, and model-based methods are used for complex patterns. There are studies on layout pattern analysis and yield prediction for hotspot prediction through physical-based modeling [19], [20]. The rule-based method is considerably affected by the user's decision and the model-based method takes a long time and has the disadvantage that the performance depends heavily on the model calibration.

For hotspot prediction, machine learning-based methods have been studied. Ding et al. established a neural network or support vector machine (SVM) as a hotspot identifier to detect a hotspot in the layout [21]. Lin et al. and Wen et al. identified hotspots using fuzzy models after extracting features using layout encoding [22]–[24]. Matsunawa et al. significantly reduced false alarms by extracting layout features and distinguishing hotspot presence based on a probability distribution function [24]. Yu et al. detected hotspot areas based on CAD design and variable information using an SVM [25].

Convolutional natural networks (CNNs) have greatly enhanced deep learning models tacking vision problems such as image classification, object detection, style translation super resolution and semantic segmentation. Numerous studies on hotspot prediction models employing deep learning have resulted from this trend. By

eliminating the pooling layer, Yang et al. [26] established a strategy for finding hotspots in large-area patterns without information loss. Zhang et al. [27] proposed a model for extracting a hotspot area from a layout utilizing a deep learning-based model and updating the model in real time for fresh data via an online learning framework. Yang et al. [14] classified hotspots using CNN-based deep learning models and augmented hotspot in order to overcome the imbalance problem. Yang et al. presented a method for extracting features by evaluating the pattern's spatial relaxation [28]. Ye et al. created a loss function based on the region below the receiver operating characteristic curve and attained high accuracy, even with unbalanced data [29]. Shin and Lee improved the speed of the hotspot prediction process by incorporating CNN-based models with the following features: data augmentation, inspection region reduction, modified batch normalization, DBSCAN clustering, and fast image scanning [16].

Convolutional natural networks have improved the performance of predicting hotspots in layouts, but most models only predict hotspots in clip layouts and do not indicate the exact location of hotspots. Accordingly, in the large-scale layout, the window sliding method is applied to predict hotspots and then aggregate them, which leads to a disadvantage of taking a lot of time to predict. To handle this problem, recently, Chen et al. predicted hotspots in the layout of the large-scale chips through deep learning model [30]. This prediction model finds hotspots in layout at once, based on the object detection model,

Chapter 3

Hotspot detection model[①]

3.1 Introduction

We suggest the deep learning-based hotspot detection model in this chapter. To accurately detect small-sized hotspots, we developed a new hotspot detection model that was inspired by conditional Generative Adversarial Network (cGAN) [31][32].

The imbalance problem is the most generally known difficulty in object detection jobs. This indicates that the proportion of positive to negative data is extremely low, resulting in an imbalance that makes the object detection model poorly trained. Eventually, the object detection model's performance degrades. Unlike conventional object detection tasks, detecting hotspots is challenging due to the often-tiny size of hotspots in semiconductor processes. If the size of the item to be identified is very tiny, the imbalance problem in object detection gets more severe

---

[①] © 2021 IEEE. Reprinted, with permission, from: Kim, Jaehoon, et al. "Adversarial defect detection in semiconductor manufacturing process." IEEE Transactions on Semiconductor Manufacturing.

"In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of Seoul National University's products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink. If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation."

and training the model becomes more challenging. Hotspots in the semiconductor process have not only hard defects but also soft defects that are difficult to distinguish even humans. Soft defects may not have a substantial effect on pattern fidelity, which is less essential than finding hard defects, however discovering soft defect is of considerable use to engineers. To my knowledge, however, the vast majority of existing hotspot detection methods do not offer soft defect detection. Here, we proposed newly develop hotspot detection model combining the keypoint estimation-based object detection model with the conditional Generative Adversarial Network (cGAN) to solve these issues.

We will briefly describe object detection model and generative model for readers unfamiliar with them. Since the advent of deep learning, object detection has emerged as one of the most successful technologies in computer vision. It serves as the basis for a variety of applications, including pose estimation, autonomous driving, and anomaly detection in medical images. There are a great number of models for object detection that have been demonstrated to be effective [33–38]. The vast majority of object detection models utilize anchor boxes, which are pre-defined by bounding boxes with particular size. It is common knowledge that utilizing anchor boxes in object detection models improves their performance; however, as the number of anchor boxes used in the model increases, Increasing the number of anchor boxes increases the model's complexity and diminishes its performance enhancement effect. In order to find a solution, a fresh model for the detection of objects has been developed. It detects objects by utilizing keypoint estimation techniques such as heatmaps, in place of anchor boxes. The ConerNet [39], the ExtremeNet [40], and the CenterNet [41] models are all examples of this type of model. The performance of our proposed hotspot detection model will be demonstrated using CenterNet, which was chosen as the baseline for the study.

Generative Adversarial Networks (GAN) learn from the image of the target domain in a way designed to generate images from the latent space, generating images in the same domain. Mirza and Osindero developed by conditional Generative Adversarial Network (cGAN) and confirmed that when class information

is added to the latent space, the model generates an image corresponding to a specific class [32]. Isola et al. proposed a supervision style transfer model called Pix2Pix, utilizing a set of images with similar structural features but different styles [42]. Due to the excellent performance of the Pix2Pix model, many studies have been conducted based on this strategy. In particular, as the need for style transmission for high-quality and high-resolution images increases, Wang et al. proposed a Pix2PixHD model that converts segmentation information into sophisticated high-resolution images [43].

The keypoint-estimation-based detection model serves as a generator of the hotspot detection model that we developed by generating hotspot detection heatmaps and regressing the height and breadth of bounding boxes. And PixelGAN [42] is utilized to differentiate between a false and authentic heatmap. This discriminator, unlike a standard GAN discriminator, is more specialized, allowing it to detect small hotspots more effectively. Instead of employing the mean absolute loss function and mean squared loss function in typical GAN loss, we employ the Focal Loss [37], which is well-known to aid in resolving the issue of data imbalance. The evaluation of our suggested model utilizing a real-world industrial dataset demonstrates its improved performance.

3.2 Hotspot detection model

The whole of our hotspot detection model's design is depicted in Figures 3.1 and 3.2. It modifies the fundamental architecture of the conditional Generative Adversarial Network (cGAN) to improve hotspot detection performance. Primarily, it is comprised of the generator network and the discriminator network.

3.2.1 Generator network

Two primary networks are included in the one-stage detection model. The encoder network was the first network. Convolutional feature maps are generated by this network's backbone network over an input image. To reduce the amount of time spent on inference while maintaining a high level of accuracy, we use Resnet-18 [44] as the backbone of the hotspot detection model. The decoder network is the second network, and it is responsible for producing hotspot detection heat maps based on the features that have been generated by the encoder network.

Two sub-networks, each responsible for a certain function, are located at the tail of the decoder network. The first subnetwork, which is known as the detection subnetwork, is made up of two convolutional layers and one batch normalization layer. It is responsible for convolutional heatmap $\hat{T} \in [0,1]^{\frac{W}{OS} \times \frac{H}{OS} \times C}$ regression to the feature maps that are produced by the backbone network. Bounding box regression is the responsibility of the second sub-network, which is called the bounding box sub-network. The structure of this subnetwork is identical to that of the detection subnetwork; however, in addition to predicting the width and height of bounding boxes $\hat{B} \in R^{\frac{W}{OS} \times \frac{H}{OS} \times 2}$, it also predicts the offset $\hat{O} \in R^{\frac{W}{OS} \times \frac{H}{OS} \times 2}$. In inference procedure, the topset 20 heat peaks of the hotspot detection heat map are identified, and the width, height, and offset of the bounding boxes are estimated in the same way as CenterNet [41].

Early object detection models, in which the encoder and decoder networks are only connected to one another, make use of the backbone network's final features.

Although these features have robust semantic information, they are not suitable for detecting small objects, such as the hotspots that occur during the manufacturing process of semiconductors [38]. Encoder network's multi-scale features can be utilized to great effect in order to detect objects of a wide range of sizes, including the small hotspots. The problem, however, is that low-level features have very little information in terms of their semantic meaning. There have been a lot of different attempts made to overcome this problem. Pyramidal features are utilized in a top-down architecture with skip connections by the networks FPN [45] and RetinaNet [37]. These two networks are examples of networks that use pyramidal features. MDSSD [46] makes use of a fusion module that brings together features on multiple scales. In this investigation, we also made use of the same methodology. In our model, the skip connection was utilized to merge low-level encoder network features with features from high-level decoder network (Figure 3.3 and 3.4). As will be shown in the following section, the performance of a model that included multi-scale features was superior to that of a model that did not using these features.

3.2.2 Discriminator network

There are three kinds of discriminators that can be used for adversarial network architecture. These are VanillaGAN [31], PatchGAN [42], and PixelGAN [42]. VanillaGAN styled discriminator is a simple discriminator that does nothing more than compare a single input image to a single output value in order to decide whether or not an image is authentic. On the other hand, PatchGAN styled discriminator seeks to classify images by first splitting them into patches before continuing with the classification process. This discriminator operates under the presumption that the pixels contained within a patch are dependent upon and interdependent upon one another, but that they are independent of the pixels contained within other patches. To put it another way, the discriminator, in comparison to VanillaGAN one, is able to learn more high-frequency information from the input images by making use of patches.

In order to build the discriminator network for our model, the tool that is utilized

is called PixelGAN styled discriminator. Three layers of leaky ReLU, two layers of 1×1 convolution, and one layer of batch normalization that does not contain a sigmoid layer make up this structure. Because this method is better able to focus on high-frequency feature information related to small hotspots than PatchGAN styled discriminator. The performance of our discriminator network based on PixelGAN was superior to that of networks based on VanillaGAN and PatchGAN, as will be demonstrated in Chapter IV.

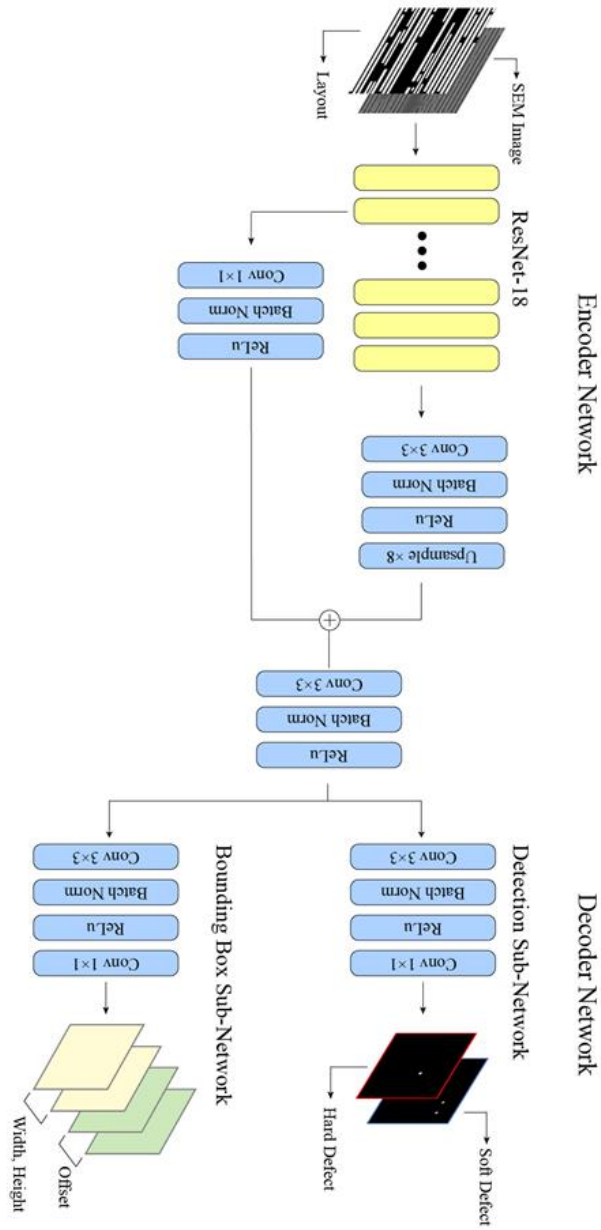

Figure 3.1. Architecture of generator network of hotspot detection model[②]

[②]© 2021 IEEE. Reprinted, with permission, from: Kim, Jaehoon, et al. "Adversarial defect detection in semiconductor manufacturing process." IEEE Transactions on Semiconductor Manufacturing.

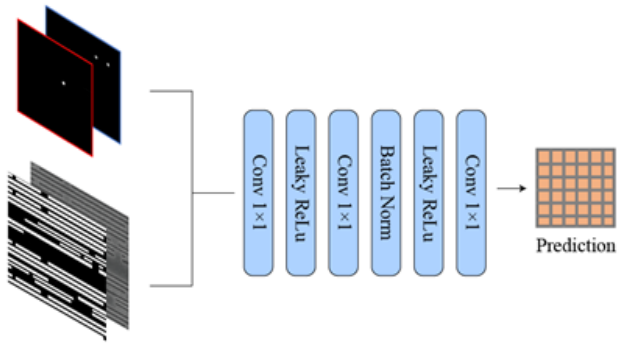

Figure 3.2. Architecture of discriminator network of hotspot detection model. In contrast to the discriminator of VanillaGAN, it is capable of focusing more on high-frequency features[③]

[③]© 2021 IEEE. Reprinted, with permission, from: Kim, Jaehoon, et al. "Adversarial defect detection in semiconductor manufacturing process." IEEE Transactions on Semiconductor Manufacturing.

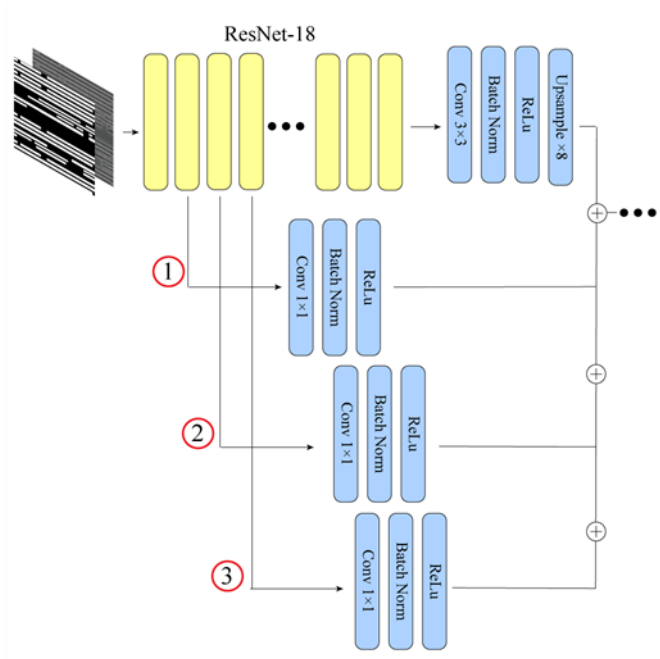

Figure 3.3. Utilization of multi-scale features in the model for hotspot detection. Utilizing multi-scale features acquired from the backbone network is an excellent method for locating objects of various sizes, even those of a relatively small scale.④

④© 2021 IEEE. Reprinted, with permission, from: Kim, Jaehoon, et al. "Adversarial defect detection in semiconductor manufacturing process." IEEE Transactions on Semiconductor Manufacturing.

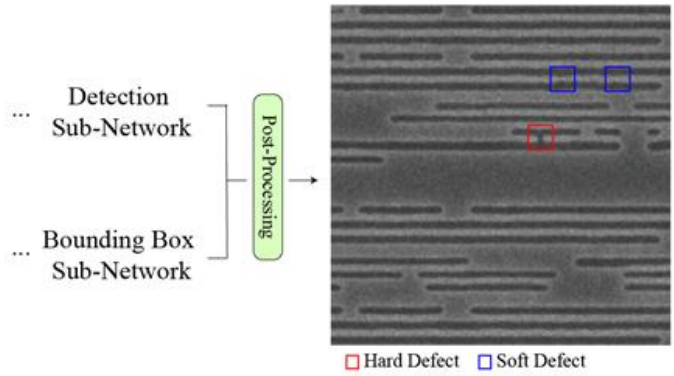

Figure 3.4. In the generator network, the predictions of the detection sub-network and bounding box sub-network are combined to derive the hotspot detection result as shown in the figure. The red bounding boxes represent hard defect, and the blue bounding boxes represent soft defect.[5]

[5]© 2021 IEEE. Reprinted, with permission, from: Kim, Jaehoon, et al. "Adversarial defect detection in semiconductor manufacturing process." IEEE Transactions on Semiconductor Manufacturing.

3.2.3 Loss function

The total loss function of our detection model is composed of two main loss functions. The GAN loss function is the first function, while the generator loss function is the second function.

3.2.3.1 GAN Loss

We made use of the loss function that was included in the Least Squares GAN [47].

$$L_{GAN}(G,D) = \mathrm{E}_{C,S,H}[\log D(C,S,T)] + \mathrm{E}_{C,S}[1-\log D(C,S,G(C,S))] \tag{3.1}$$

The discriminator network, the generator network, the layout image, the SEM image, and the ground truth heatmap are each denoted by the letters D, G, C, S, and T, respectively. $G(C,S)$ represents hotspot detection heat maps which are output results of the generator network's detection subnetwork. In a GAN, the discriminator D and generator G compete to maximize and minimize the target function, which is known as an adversarial loss function. The ultimate objective is to figure out which parameters of generator G produce the best hotspot detection heat maps that accurately represents the ground truth or comes as close as possible to accurately representing it.

3.2.3.2 Detection Sub-Network Loss

For the purpose of the detection sub-network, the loss function is implemented as an altered version of the Focal Loss [37]. The class imbalance problem manifests itself in the context of a general detection issue whenever there is a scenario in which the typical number of negative samples is higher than positive ones. When there is

an imbalance problem, the performance of the detection model suffers as a direct result. Because the detection objects we are trying to detect (hotspots in the semiconductor process) are quite smaller, class imbalance becomes more serious in our application than it does in other problems. The following is an explanation of how Focal Loss came to be developed as a solution to this problem: [37]:

$$L_H = -\frac{1}{N}\sum_{c}^{C}\sum_{y}^{Y}\sum_{x}^{X}\begin{cases}(1-\hat{T}_{xyc})^{\alpha}\log(\hat{T}_{xyc}) & \text{if } T_{xyc}=1\\ (1-T_{xyc})^{\beta}(\hat{T}_{xyc})^{\alpha}\log(1-\hat{T}_{xyc}) & \text{otherwise}\end{cases} \tag{3.2}$$

Where N denotes the total number of image defects. Both α and β are considered to be hyper-parameters of the loss function. The proportion of losses that occur in samples that are misclassified increases in direct proportion to the values of α and β. This enables the model to concentrate more on samples that are not correctly classified than it does on samples that are correctly classified. In order to comply with CornerNet and CenterNet, we set α to the value 2 and β to the value 4.

3.2.3.3 Bounding Box Sub-Network Loss

The bounding box sub-network makes use of L1 loss as its choice for the loss function.

$$L_{BO} = \frac{1}{N}\sum_{p}\left(\left|\hat{B}_p - B_p\right| + \left|\hat{O}_p - O_p\right|\right) \tag{3.3}$$

Where p indicates the center points of the bounding boxes of the various objects in the image $I \in R^{W\times H\times 3}$.

3.2.3.4 Overall Loss

The goal of our proposed hotspot detection model is to optimize the adversarial loss function. This function is made up of the GAN loss function, the detection sub-network loss function, and the bounding box sub-network loss function.

$$\min_G(\max_D(L_{GAN}(G,D)) + \lambda(L_H + L_{BO})) \tag{3.4}$$

The loss scale constant for the detection and bounding box sub-networks is denoted by the letter λ. In this study, alpha is going to be at 0.1. When it comes to recognizing small objects, the use of the Focal Loss function in conjunction with the GAN loss function demonstrated significantly higher performance than mixing the standard L1 or L2 loss functions [42]. After conducting research, we came to the realization that this was the case. The conversation will pick back up later on in Chapter 3.3.

3.3 Experiments

A single Titan RTX 24GB GPU served as our training platform, and we used Adam [48] as an optimizer and a learning rate of 5×10^{-4} to train our hotspot detection model and other baseline object detection methods over a period of 10-40 epochs.

3.3.1 Dataset and preprocessing

In this research, we make use of an industrial dataset that contains a total of 17,125 SEM images, layout designs, and ground truth images. These images include labeling information regarding the types of hotspots that hard defect and soft defect as well as the coordinate information of bounding boxes. We divided the total into two sets of data: 1,717 for the testing portion, and 15,407 for the training portion. A total of 512×512 images were taken from the dataset and used for the training process. We used standard data augmentation techniques. These techniques included rotating the data at random, flipping it randomly, and jiggling the colors at random. The random color jittering effect was only applied to the SEM images. In order for our model to make use of the benefits offered by D2DB inspection, as shown in Figure 3.5, it requires as input a SEM image as well as a layout design that is combined in the channel direction.

Our model's generator network employs a keypoint estimation-based detection model. Therefore, the ground truth should be converted into hotspot detection heatmap $T \in [0,1]^{\frac{W}{OS}\times\frac{H}{OS}\times C}$ depicting hotspots as a heat distribution using a Gaussian kernel. Here, W and H represent the width and height of the input image, while OS represents the output stride. In this study, the output stride OS is set to 4, which decreases the output size by an OS factor. C equals 2 because hard and soft defects are defined.

As a result of the inspection equipment limitations, even though we trained with

high-quality industrial datasets, many of the bounding boxes were incorrect (Figure 3.6). To counteract this, we added the following Gaussian random noise to the ground truth bounding boxes:

$$T_{xyc} = \sum_i \exp\{-(\frac{\left(x - \frac{p_{icx}}{OS}\right)^2}{\sigma_{icx}} + \frac{\left(y - \frac{p_{icy}}{OS}\right)^2}{\sigma_{icy}})\} \tag{3.5}$$

In this equtation, p_{icx} and p_{icy} represent the coordinates of the central point of ground truth bounding box i in channel c. σ_{icx} and σ_{icy} represents $w_{ig}^* / 6$ and $h_{ig}^* / 6$, where $w_{ig}^* = w_{ig} + \varepsilon$ and $h_{ig}^* = h_{ig} + \varepsilon$. Here, w_{ig} and h_{ig} stands for width and the height of ground truth bounding box i. And ε is a random noise sampled from the gaussian distribution, $N(0,1)$.

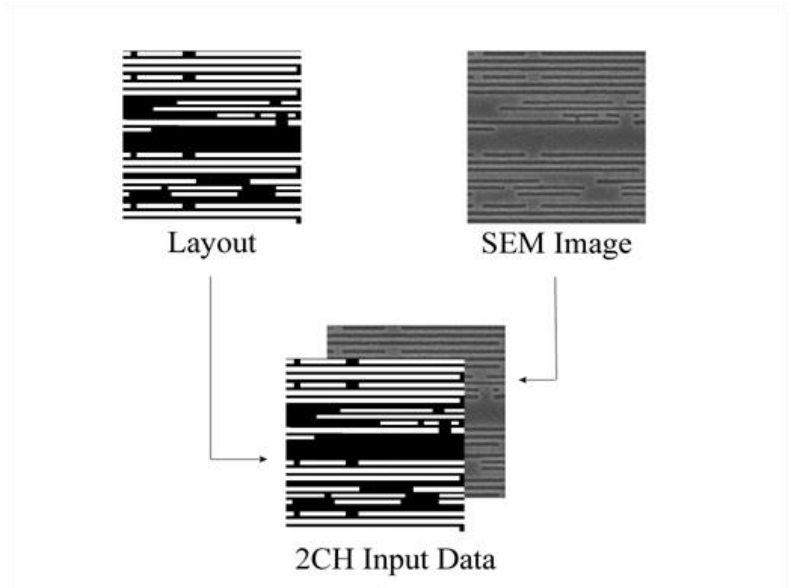

Figure 3.5. In hotspot detection model, channel-wise merged layout designs and SEM images are used.[6]

[6]© 2021 IEEE. Reprinted, with permission, from: Kim, Jaehoon, et al. "Adversarial defect detection in semiconductor manufacturing process." IEEE Transactions on Semiconductor Manufacturing.

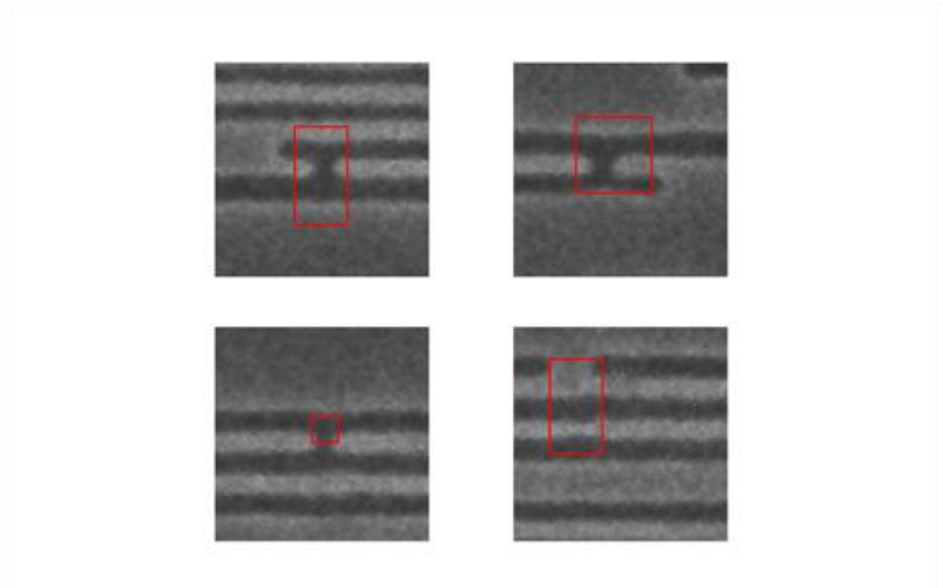

Figure 3.6. Examples of bounding boxes of hotspot in training dataset. It is not easy to train the hotspot detection model because the height and width of the bounding boxes are not constant and there are cases where it is incorrectly labeled.[7]

[7] © 2021 IEEE. Reprinted, with permission, from: Kim, Jaehoon, et al. "Adversarial defect detection in semiconductor manufacturing process." IEEE Transactions on Semiconductor Manufacturing.

3.3.2 Performance

When evaluating the effectiveness of the model, the mean average precision (mAP) at the intersection of union (IOU) 50 % is the metric that is utilized. When the predicted bounding box overlaps the ground truth by more than 50 %, this metric is considered to have been met. In addition, we measured AP_{Hard} and AP_{Soft} to assess the model's accuracy for every hotspot defect type. Table 3.1 compares performance to that of other well-known object detection models, including CenterNet, Faster-RCNN, RetinaNet, YOLOv3, and DTER. Because our model's generator network is based on CenterNet, we chose it as a representative baseline model to compare it to for the purpose of this study. The mAP evaluated by our model is approximately 5 % points higher than the score generated by CenterNet. In the case of hard defects, our model performed satisfactorily with a score of AP_{Hard}, 98.6 %, and it was demonstrated that its capability to detect soft defects had been significantly improved (Figure 3.7 and Figure 3.8). And Figures 3.9 and 3.10 depict, respectively, the normalized confusion matrix and the log-average miss rate for hard defect and soft defect. In addition, Table 3.1 demonstrates that our hotspot detection model outperforms Centernet and other well-known object detection models. These results demonstrate the significance of our model's adversarial network architecture and multi-scale features.

3.3.3 Ablation Study

Ablation studies were carried out so that we could obtain a better understanding of how the performance of our model was affected by its fundamental components. This was accomplished by comparing the results of the ablation studies to the results of the other studies. The use of generative adversarial networks and multi-scale features, the type of loss function, the loss scale constant λ and the discriminator were the primary aspects of our model that needed to be evaluated.

3.3.3.1 GAN and Multi-scale Features

We investigated the ways in which adversarial network architecture and multi-scale features improve the detection performance of the model. The findings are presented in Table 3.2 below. As was to be expected, the GAN and multi-scale features are necessary components for the performance of detection models. Both mAP and AP_{Hard} were enhanced by 5.1 % and 2.4 %, respectively, due to the use of both components. As evidenced by the 7.8 % improvement in AP_{Soft}'s performance for locating tiny hotspots, they were particularly important. Additionally, we examined the impact of the use of the number of multiscale features. According to Table 3.3, a model with three multi-scale features outperformed those with less features. However, in order to maximize the efficiency of the model, we decided to only use a single low-level feature. This decision was made due to the fact that using multiple low-level features did not result in a statistically significant increase in performance across models, and the addition of multiple features results in an increase in the number of parameters that need to be trained.

3.3.3.2 Loss Scale Constant

When compared to the loss scale constant, the performances were rated and analyzed. According to Table 3.4, the best possible outcome could be accomplished by setting the value to 0.1.

3.3.3.3 Detection Sub-Network Loss

In addition, we validated the manner in which the Focal Loss that was utilized as the Detection Sub-Network Loss improved the model's detection performance. In traditional conditional GAN models, the L1 or L2 Loss is used in conjunction with the GAN Loss. Instead, Focal Loss was included in this model as an additional

variable. According to Table 3.5, Focal Loss is required to achieve a high level of detection performance. When L1 or L2 Loss was utilized, the ability to locate hotspots was significantly hindered.

3.3.3.4 Discriminator

Table 3.6 demonstrates that the PixelGAN-based discriminator network is superior. The mAP of the model with PixelGAN was 3.8 % and 2.9 %, respectively, higher than that of the model with VanillaGAN and PatchGAN. This would imply that it is absolutely necessary to place a larger emphasis on high-frequency feature informations in order to detect hotspots with a greater degree of precision.

3.3.3.5 Input data

The performance of a model that only used SEM images as input was compared to the performance of a model that used SEM images with layout designs as input. When SEM images and layout designs were utilized, the model demonstrated a marginal improvement in accuracy, as demonstrated in Table 3.7. The improved performance of mAP, AP_{Hard}, and AP_{Soft}, respectively, of 1.6 %, 0.9 %, and 2.3 %, respectively, suggests that using of layout designs with SEM images is more advantageous for detecting soft defects.

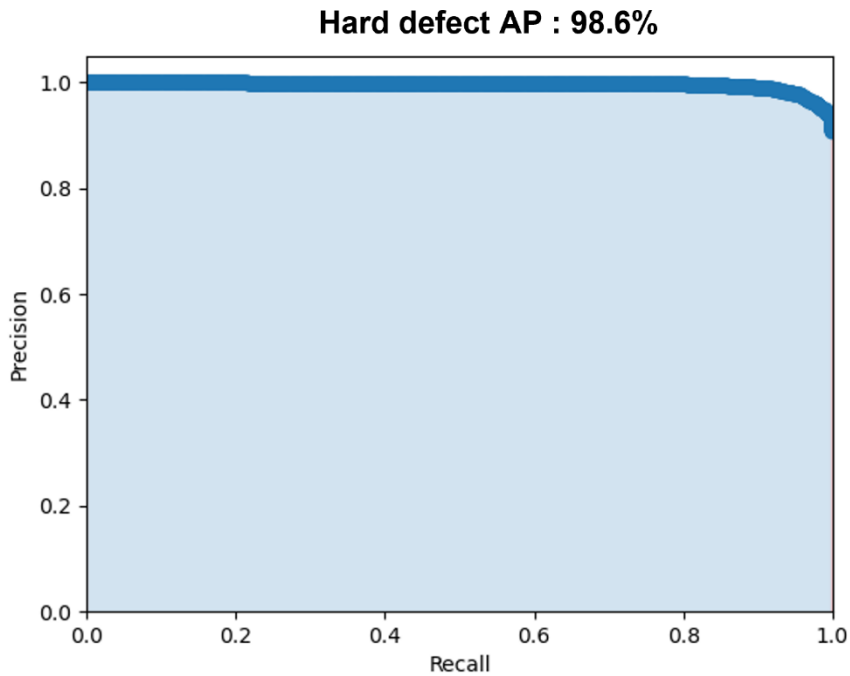

Figure 3.7. Average Precision for hard defect on hotspot detection model.

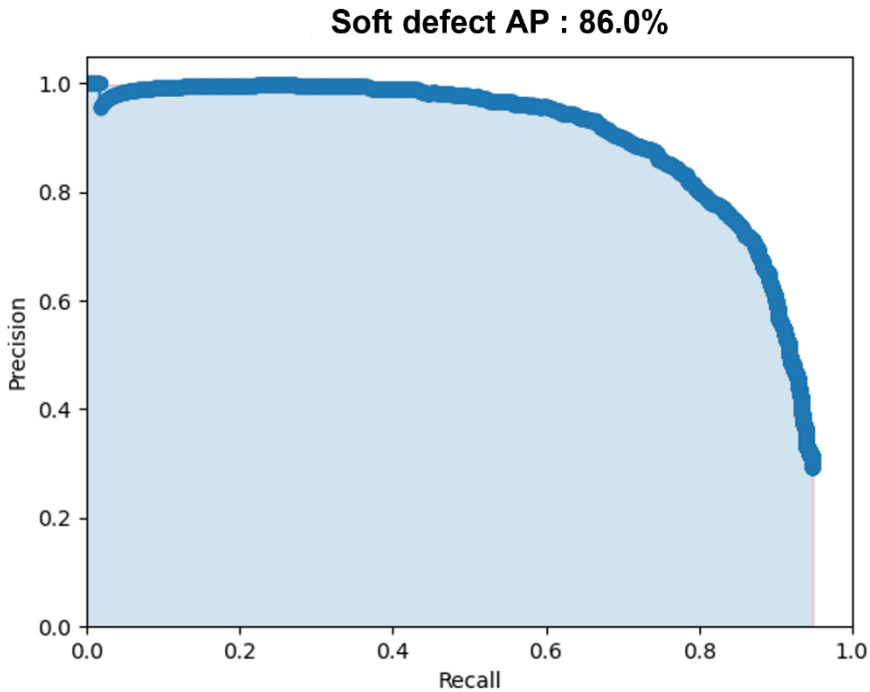

Figure 3.8. Average Precision for soft defect on hotspot detection model

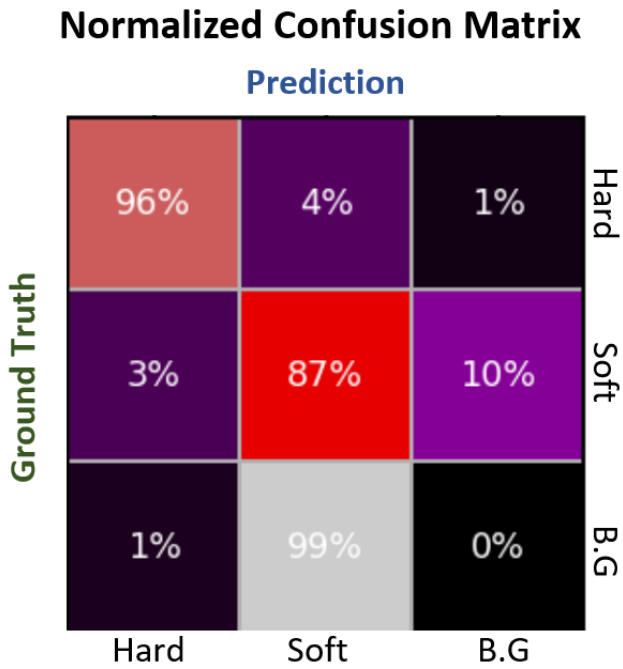

Figure 3.9. Normalized confusion matrix for hotspot detection model. Normalized confusion matrix was made at IoU threshold of 0.5, score threshold of 0.3.

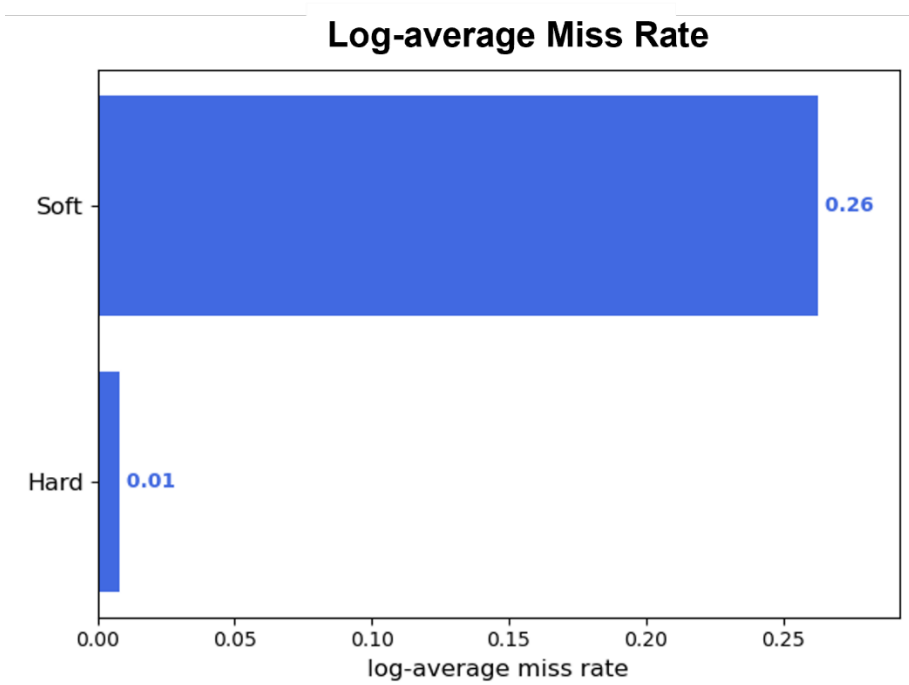

Figure 3.10. Log-average miss rate for hotspot detection model

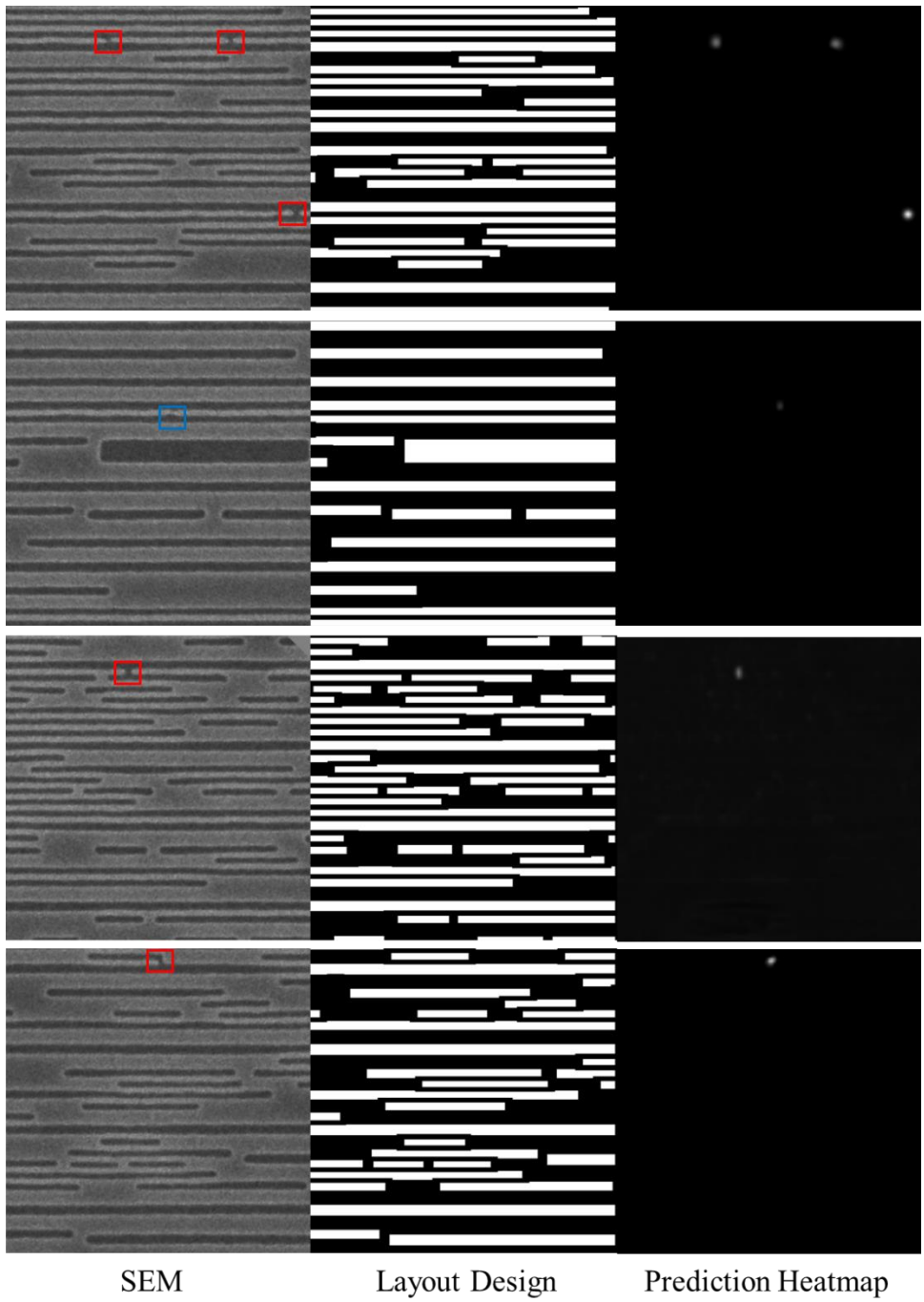

Figure 3.11. Examples of prediction result of hotspot detection model. The red box and blue box represents hard defect and soft defect.

Model	backbone	mAP	AP_{Hard}	AP_{Soft}	Param(M)
Faster R-CNN w/ FPN	ResNet-50	83.0	98.0	68.0	41.53
RetinaNet	ResNet-50	84.9	98.6	71.2	36.13
EfficientDet	EfficientNet B3	85.8	98.2	73.5	18.36
YOLOv3	DarkNet-53	89.1	98.3	79.8	61.53
Deformable Deter	ResNet-50	86.7	97.8	75.6	40.80
CenterNet	ResNet-18	86.1	97.7	74.5	14.21
CenterNet	ResNet-50	87.4	98.4	76.3	30.67
Our Model	**ResNet-18**	**92.3**	**98.6**	**86.0**	**12.64**

Table 3.1. Comparison of the proposed our model and baselines

GAN	Multi-scale features	mAP	AP_{Hard}	AP_{Soft}
NO	NO	87.2	96.2	78.2
NO	YES	90.9	97.5	84.3
YES	**YES**	**92.3**	**98.6**	**86.0**

Table 3.2. Ablation studies of GAN architecture and multi-scale features [8]

[8] © 2021 IEEE. Reprinted, with permission, from: Kim, Jaehoon, et al. "Adversarial defect detection in semiconductor manufacturing process." IEEE Transactions on Semiconductor Manufacturing.

Number of features	mAP	GAN
None	87.2	NO
1	**90.9**	NO
2	91.2	NO
3	91.3	NO

Table 3.3. Ablation studies for number of multi-scale features[⑨]

[⑨]© 2021 IEEE. Reprinted, with permission, from: Kim, Jaehoon, et al. "Adversarial defect detection in semiconductor manufacturing process." IEEE Transactions on Semiconductor Manufacturing.

λ	mAP	AP_{Hard}	AP_{Soft}
0.1	**92.3**	**98.6**	**86.0**
1	91.7	98.6	84.7
10	90.8	97.5	84.1

Table 3.4. Ablation studies for parameter λ [10]

[10] © 2021 IEEE. Reprinted, with permission, from: Kim, Jaehoon, et al. "Adversarial defect detection in semiconductor manufacturing process." IEEE Transactions on Semiconductor Manufacturing.

Loss Type	mAP
L1	50.4
L2	49.6
Focal Loss	**92.3**

Table 3.5. Ablation studies for loss type[11]

[11]© 2021 IEEE. Reprinted, with permission, from: Kim, Jaehoon, et al. "Adversarial defect detection in semiconductor manufacturing process." IEEE Transactions on Semiconductor Manufacturing.

Discriminator Type	mAP
VanilaGAN	88.5
PatchGAN	89.4
PixelGAN	**92.3**

Table 3.6. Ablation studies for architecture of discriminator [12]

[12] © 2021 IEEE. Reprinted, with permission, from: Kim, Jaehoon, et al. "Adversarial defect detection in semiconductor manufacturing process." IEEE Transactions on Semiconductor Manufacturing.

Input Data Type	mAP	AP_{Hard}	AP_{Soft}
SEM and Layout	**92.3**	**98.6**	**86.0**
Only SEM	90.7	97.7	83.7

Table 3.7. Ablation studies for input data of hotspot detection model [13]

[13] © 2021 IEEE. Reprinted, with permission, from: Kim, Jaehoon, et al. "Adversarial defect detection in semiconductor manufacturing process." IEEE Transactions on Semiconductor Manufacturing.

3.4 Conclusions

Within the scope of this research, we propose a hotspot detection model for the semiconductor process that is based on deep learning. The keypoint-estimation-based object detection model and the discriminator are combined in our proposed hotspot detection approach, which employs an adversarial network architecture. This allows us to locate hotspots, soft defects and hard defects, in the semiconductor process accurately and rapidly. To get a higher hotspot detection performance, we make use of several additional features, including Focal Loss, PixelGAN styled discriminator, and multi-scale. The proposed model's performance is confirmed to be of an excellent standard after being tested on a dataset derived from industry.

Chapter 4

Hotspot prediction model

4.1 Introduction

We propose a deep learning-based approach that predicts hotspots by translating the layout design into defective SEM images containing hotspots in this chapter. In order not only to generate a good quality pattern image from the layout but also to localize the hotspot more accurately, our model combines the segmentation model and style-transfer model based on conditional Generative Adversarial Network (cGAN) model in parallel.

Segmentation is a task of classifying classes in pixel units within an image, and Long et al. suggested an initial segmentation model based on a fully convolutional network [49]. Chen et al. improved localization capabilities by utilizing CNN and fully connected CRFs in combination [50]. Noh et e al. proposed a model of the encoder-decoder structure with deconvolution, showing good performance in the PASCAL VOC 2012 dataset [51]. Ronneberger et al. performed segmentation on biological images with a U-Net model which have skip connections with the encoder-decoder structure [52]. Lin et al. invented Feature Pyramid Network and proposed a method of dealing with multi-scale [45]. He et al. developed an RCNN-based Mask RCNN model and proposed a model that performs both detection and segmentation. this model introduced the concept of instance segmentation in which all objects belonging to the same class are segmented individually [34]. Chen et al. proposed a Deeplabv3+ model by adding atrous convolution to the encoder-decoder structure. Improve segmentation performance by effectively expanding receptive

fields through Atrous convolution [53]. Research on weakly supervision segmentation is currently attracting considerable attention in order to cut down on the time and money needed to create the segmentation model's training data. In terms of convenience, the advantage of using partially labeled images or image tags is clear, but there are limitations compared to supervised learning in terms of accuracy. To solve this, Kervadec et al. devised a model for learning domain knowledge and successfully applied it to medical images, achieving outstanding results [54].

Segmentation model in our proposed hotspot prediction model plays the role of the first step in predicting hotspots in the layout, such as the region proposal network of Faster R-CNN [33] and the generator is responsible for translating layout design to defective SEM image containing predicted hotspots. Therefore, unlike in other hotspot prediction models, hotspots can be visually identified. Although the proposed model predicts and generates fake hotspots, various interpretations are expected to be possible based on the size, appearance, and location of the predicted hotspots. In addition, we propose a way to guess approximately which part of the layout patterns are susceptible to hotspots, using the optical diameter. The proposed hotspot prediction model helps intuitively predict hotspots in semiconductor patterns in production lines and provides designers with insight into identifying the major causes of hotspots. Finally, it was confirmed that the performance of the model was improved by applying objective functions such as a multi-scale discriminator, feature discriminator loss, and constrained partial cross-entropy loss.

The main contributions of our proposed hotspot prediction model are two areas. First, we propose a model to predict hotspots by translating the layout design into an SEM image. This model performed better in predicting hotspots by training hotspot segmentation and SEM image generation simultaneously in parallel. Second, as far as we are aware, our model is the only one that includes previous knowledge for hotspot prediction. An optical diameter, the region known to be related to printed patterns on wafer, was used to approximately predict the layout patterns susceptible to hotspots.

4.2 Hotspot prediction model

Figures 4.1 and 4.4 illustrate the overall structure of our hotspot prediction model. Our model comprises a backbone network, a branch for hotspot segmentation, a branch for SEM generation, and two discriminators. Resnet-101[44] was utilized as the network's backbone. The low-level features of the backbone network are useful for detecting small objects, such as hotspots in semiconductor process; nevertheless, they provide limited semantic information. Using skip connection, features of the backbone network and each branch are merged to address this issue [42], [43], [49].

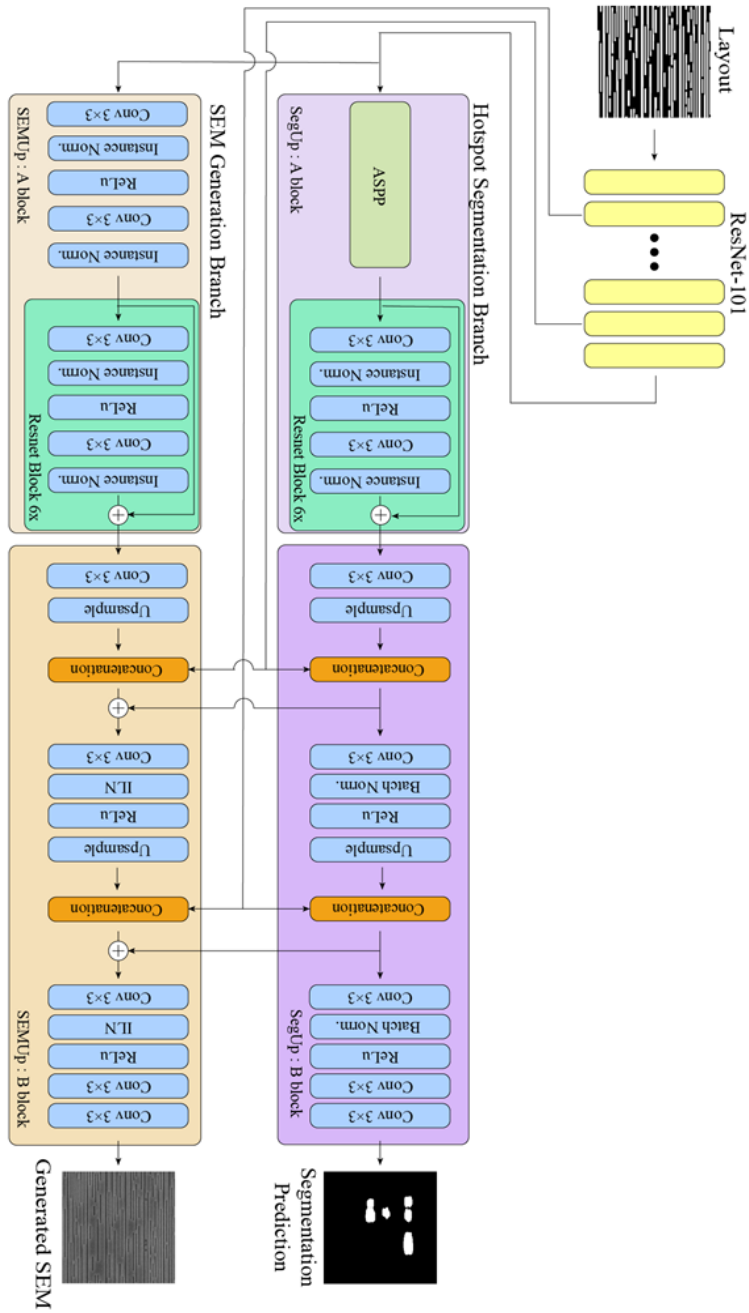

Figure 4.1. Architecture of generator network of hotspot prediction model. The hotspot prediction model consists of a parallel structure of the hotspot segmentation branch and the hotspot generation branch.

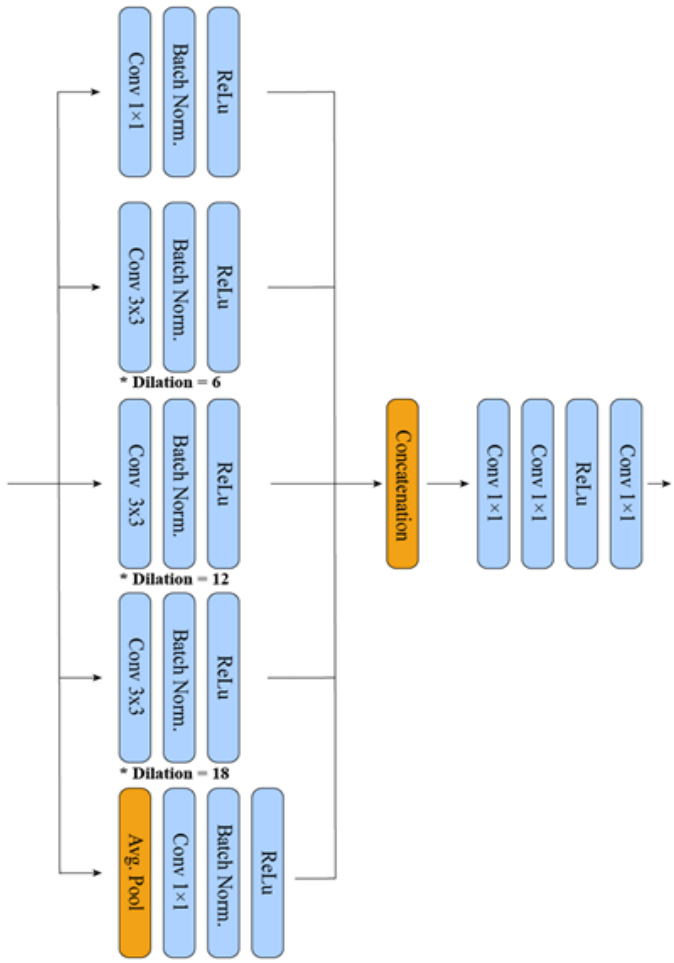

Figure 4.2. Architecture of Atrous Spatial Pyramid Pooling module (ASPP) of hotspot segmentation branch. this module utilizes atrous convolution to enable the model to have various receptive fields. This allows hotspot segmentation branch to handle multi-scale features well.

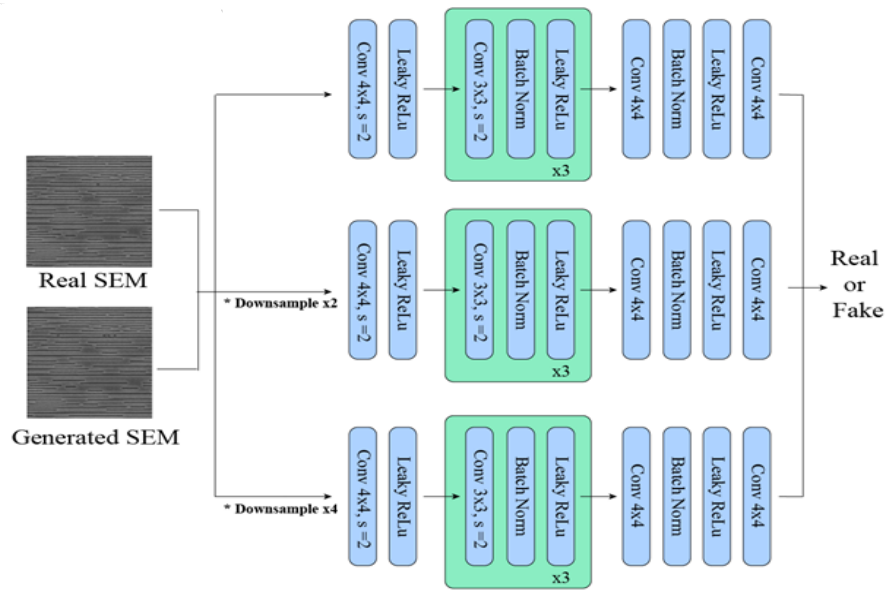

Figure 4.3. Architecture of SEM discriminator of hotspot prediction model. This multi-scale discriminator utilizes multiple features to enable the SEM generation branch to generate better SEM image at multiple scales

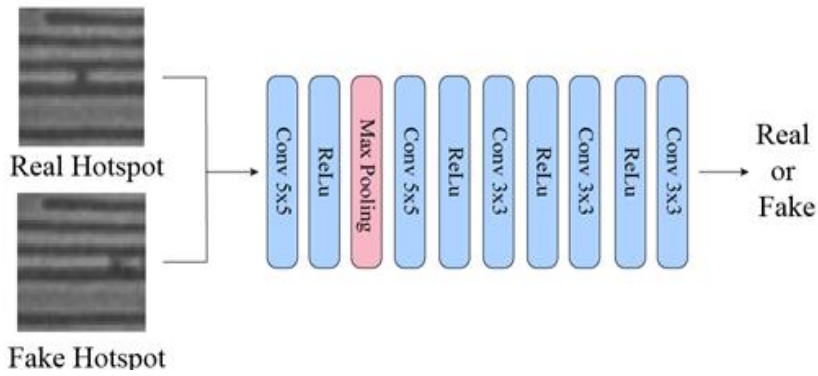

Figure 4.4. Architecture of Hotspot discriminator of hotspot prediction model

4.2.1 Hotspot segmentation branch

Similar to the region proposal network in two-stage object detection models, Faster R-CNN [33] and Mask R-CNN [34], hotspot segmentation branches help the SEM generation branch generate defective SEM images.

This branch consists of a SegUp block and atrous spatial pyramid pooling (ASPP) module. Introduced in Deeplabv2 [50] and Deeplabv3+ [55], the ASPP module is often used in segmentation models and utilizes atrous convolution to enable the model to have various receptive fields. This allows the model to handle multi-scale features well. The SegUp block consists of SegUp A and B blocks, the details of which are depicted in Figure 4.1.

The hotspot segmentation branch performs image segmentation tasks using features from the backbone network that received the layout as input and returns the hotspot prediction map $\hat{T}_p \in [0, 1]^{W \times H}$. p, W and H denote pixels, width and height of the prediction map respectively. This hotspot prediction map was used as a hotspot prediction region proposal to generate defective SEM images that contain the predicted hotspots.

Labeling is always considered to be a major challenge in semantic segmentation tasks. This is because fully supervised segmentation requires annotations at the pixel level, which is a laborious task. To lessen this burden, vigorous research is being performed on weakly supervised segmentation models that utilize weak annotations. These weak annotations include partial labels, bounding boxes, points, and scribbles. [56, 57] In particular, the weakly supervised segmentation model is utilized quite frequently for datasets that are either costly or difficult. As a result, it was also utilized in this research study. We had planned to label exactly which portion of the layout pattern is susceptible to hotspots in the dataset; however, we could only acquire the location and appearance of hotspots using ground-truth SEM images. Our original intention was to label exactly which portion of the layout pattern is susceptible to hotspots in the dataset.

However, the optical diameter, calculated by the point spread function, was part

of the prior knowledge that was applicable in this regard [58]. Optical diameter refers to the diameter of a circle including patterns involved in the generation of a specific layout pattern (Figure 4.5). The closer the layout patterns are to the hotspot, the more associated they are with the hotspot. If a hotspot occurs during the lithography process, it may be necessary not only to modify the layout pattern with the hotspot but also to modify all layout patterns within the optical diameter of the circle centered on the hotspot. The optical diameter may vary according to the process; in this study, the optical diameter was calculated to be 1 μm. We used constrained CNN loss for the hotspot segmentation branches to maximally utilize the optical diameter. Constrained CNN loss was designed for a weakly supervised segmentation task. This loss directly enforces an inequality constraint with prior knowledge of the partial cross-entropy loss [54], which is commonly used in weakly supervised segmentation models. Owing to the constrained CNN loss imposed by prior knowledge, the hotspot segmentation branch can approximately predict where hotspots occur in the layout pattern. These predictions can be utilized to correct layout patterns so that hotspots do not occur. Figure 4.6 shows the process of predicting layout design patterns related to hotspots using constrained partial cross-entropy loss. In this figure, the upper bound is defined by the optical diameter, which is prior knowledge.

However, Figure 4.7 also shows that the output of the hotspot segmentation branch had multiple false positives, which may suggest that the hotspot segmentation branch predicted incorrectly. However, the characteristics of the dataset covered in this study must be considered. As discussed in Section II, the dataset used in this study was obtained from an actual semiconductor process industry. The hotspot may not appear in the SEM image, even if the pattern is susceptible to hotspots because of temperature, humidity, and dust. The opposite is also possible.

We assume that the hotspot segmentation branch predicts layout patterns susceptible to hotspots, even if they are not in the ground truth. To examine this assumption, we derived the correlation between the predictions from the hotspot segmentation branch and the ground truths. Figure 4.8 shows that predictions of the hotspot segmentation branch are more highly correlated with ground truths than

normal patterns. Fig 4.9 shows a two-dimensional visualization using T-distributed stochastic neighbor embedding (T-SNE). This figure also indicates that the layout patterns predicted by the hotspot segmentation branch are closely related to the layout pattern susceptible to hotspots. If you look closely at this T-SNE two-dimensional visualization map, you will notice that the distribution of predictions, ground-truths, and normal layout pattens are not exactly distinguished. This is because the dataset used when the hotspot detection model is trained not includes masks which modified to prevent hotspots by the OPC procedure, but layout design that are the previous steps of mask. Due to this limitation, a clear distinction cannot be obtained from the T-SNE two-dimensional visualization map. If dataset from the mask is also used as training data, it is anticipated that results will be better.

4.2.2.1 Partial cross entropy constrained loss

The hotspot segmentation branch's training process proceeds as follows.

$$\begin{aligned} &\min_{\theta} L_T \quad \text{s.t.} \quad a \le \sum_{p \in \Omega_i} \hat{T}_p \le b, \quad i = 1,...,N \\ &L_T = -\sum_{N} \sum_{p \in \Omega_i} \log(\hat{T}_p) + (1 - T_p) \log(1 - \hat{T}_p) \end{aligned} \tag{4.1}$$

Here, L_T is a partial cross-entropy loss function, and T_p and $\hat{T}_p$ are the ground truth and prediction of the hotspot segmentation branch, respectively. N is the number of predictions in the hotspot segmentation and Ω_i is discreate image containing i-th hotspot. In this study, the ground-truth segmentation map generated using the bounding boxes was used as the partial ground truth. Further, a and b denote the lower and upper bound sizes, respectively, which are determined by prior knowledge and optical diameter. In this study, a was set close to 0, and b was set to 5,026. The upper bound size b is calculated from the relation $b = \pi(d/2)^2$, where d is 80 pixels, half of optical diameter of 1 μm, because the closer the layout patterns are to the

detector, the more they are associated with the hotspot. The detailed partial cross-entropy loss of the hotspot segmentation branch is as follows:

$$
\begin{aligned}
L_T + \lambda \sum_{N} C(V_i), \quad & \mathrm{i} = 1,\ldots,\mathrm{N} \\
V_i = \sum_{p \in \Omega_i} \hat{T}_p & \qquad (4.2) \\
C(V_i) = \begin{Bmatrix} (V_i - a)^2, & \text{if } V_i < a \\ (V_i - b)^2, & \text{if } V_i > b \end{Bmatrix} &
\end{aligned}
$$

Here, V_i denotes sum of pixels of i-th hotspot prediction and $C(V_i)$ is constraints of loss function. λ is a hyperparameter that determines the weight of constraints. In this study, we set $\lambda = 10$.

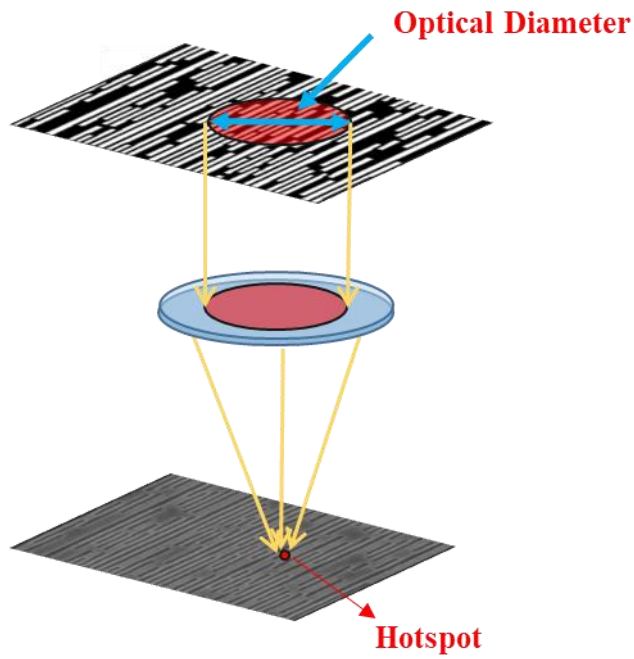

Figure 4.5. Illustration of concept of the optical diameter. The upper bound of constrained partial cross entropy loss is determined by the optical diameter.

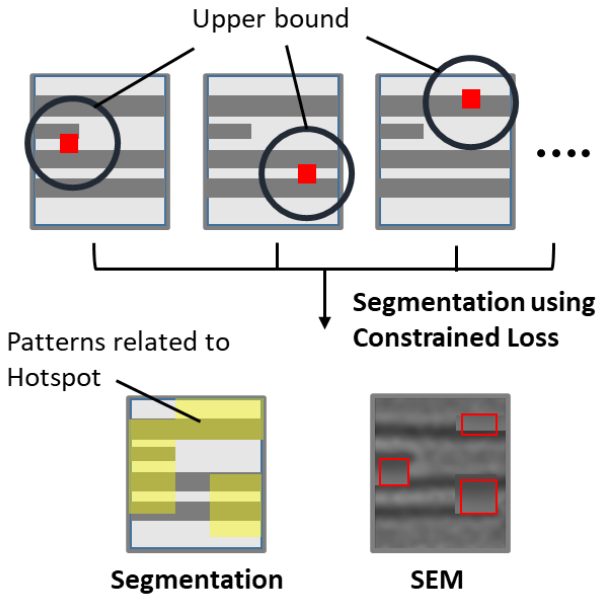

Figure 4.6. A brief conceptual diagram of the process of predicting the layout design patterns associated with hotspots using constrained partial cross entropy loss. When training the hotspot prediction model, it is possible to predict the layout patterns involved in hotspots while being constrained by the upper bound defined by the optical diameter which is prior knowledge.

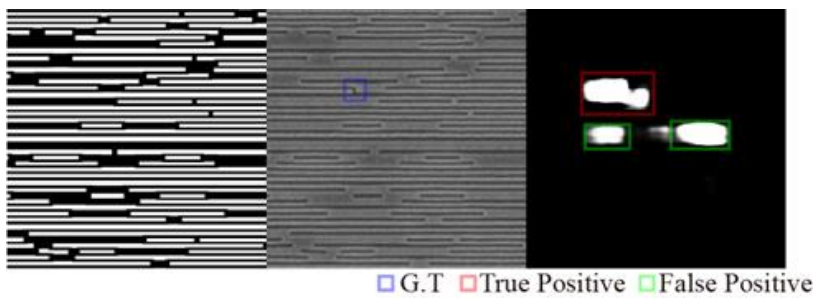

Figure 4.7. The results of the hotspot segmentation branch include not only accurate predictions (red box) but also several false positive (green box).

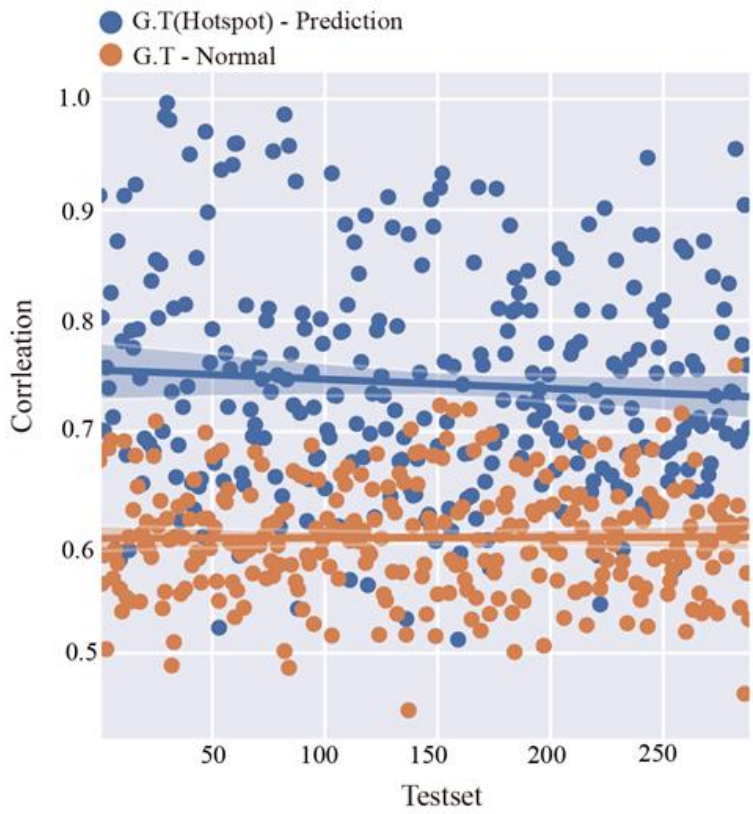

Figure 4.8. Illustration comparing the correlation between G.T(real hotspots) and predictions and the correlation between G.T and normal pattern. This shows that predictions of our model are highly correlated with to G.T, real hotspots

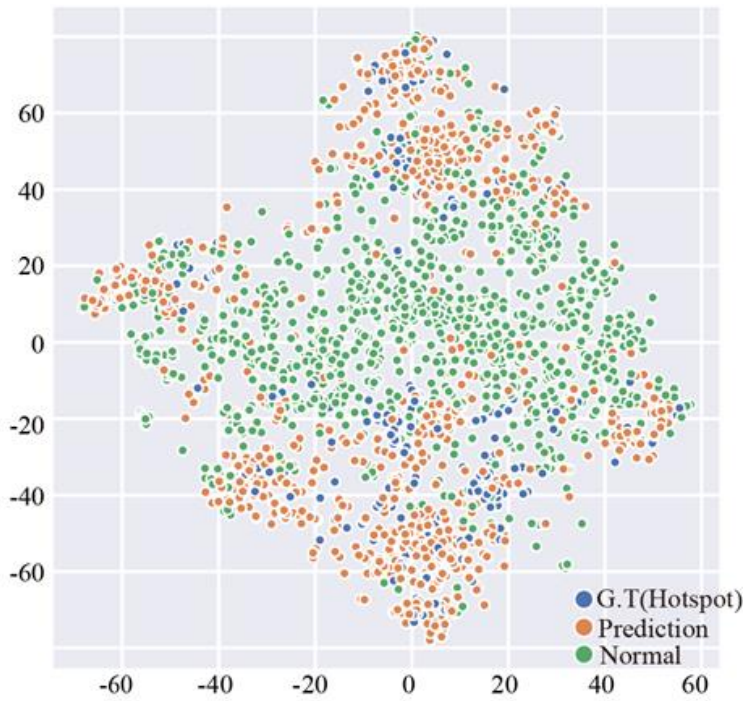

Figure 4.9. Visualization in two-dimensional feature map utilizing T-distributed Stochastic Neighbor Embedding (T-SNE) with extra model features. This figure shows that the model's predictions are closer to G.T., while at the same time roughly showing that hotspots (prediction and real), have different latency values than normal patterns.

4.2.2 SEM generation branch

The SEM generation branch translated the layout design into an defective SEM image containing predicted hotspots. Through the generated SEM image, one can directly check the appearance and location of potential hotspots in the layout pattern. The SEM generation branch consists of the bottleneck, SEMUp blocks, and SEM discriminator D_{SEM} for adversarial training.

In our model, error gradients from the hotspot segmentation branch are always propagated only with the backbone and do not provide direct information to the SEM generation branch. To solve this problem, multi-scale features were extracted from the hotspot segmentation branch and added to the SEM generation ranches, as shown in Figure 4.1. The SEM generation branch utilizes the learned features from the hotspot segmentation branch to generate a more accurate defective SEM image. Using these features, the SEM generation branch showed a better performance than the others did. We discuss this aspect in Chapter 4.3.

4.2.2.1 GAN loss

The objective function for the SEM generation branch as follows.

$$\begin{aligned} L_{G1}(G_{SEM}, D_{SEM}) = \mathrm{E}_{C,S}[\log D_{SEM}(C,S)] \\ +\mathrm{E}_{C}[\log(1-D_{SEM}(C, G_{SEM}(C)))] \end{aligned} \tag{4.3}$$

$$L_{G2}(G_{SEM}) = \left\| G_{SEM}(LD) - S \right\|_1 \tag{4.4}$$

$$\min_{G_{SEM}} \left(\max_{D_{SEM}} \left(L_{G1}(G_{SEM}, D_{SEM}) \right) + L_{G2}(G_{SEM}) \right) \tag{4.5}$$

Here, D_{SEM}, G_{SEM} are the SEM image discriminator, SEM generation branch and LD

and S are layout design ground-truth SEM images, respectively.

G_{SEM}(C) indicates the generated SEM image. Because L_{G1} is an adversarial loss, D_{SEM} try to maximize the target function and G_{SEM} attempt to minimize it. For stable training, we trained the SEM generation branch using the least-squares GAN [47] with L1 distance loss L_{G2} between the fake, generated defective SEM images and real SEM images.

4.2.2.2 SEM discriminator

We used a multi-scale discriminator and the feature matching loss used in Pix2PixHD. As the images are downsampled, the high-frequency features in the images can be removed, retaining the low-frequency features; therefore, the multi-scale discriminator enables individual learning of the high/low-frequency features in the images. The feature matching loss also plays a similar role. This loss utilizes multiple features extracted from the D_{SEM} to enable the SEM generation branch to generate better SEM images at multiple scales. Owing to these benefits, our model can predict hotspots more effectively than a typical discriminator. The details of the feature matching loss L_{FM} is as follows:

$$L_{FM}(G_{SEM}, D_{SEM}) = \mathrm{E}_{C,S} \sum_{l} \frac{1}{N_i} [\left\| D_{SEM}^{(i)}(LD, S) - D_{SEM}^{(i)}(LD, G_{SEM}(LD)) \right\|_1] \tag{4.6}$$

Here, l is the number of layers in the discriminator D_{SEM} and N_i denotes the number of elements in the i-th layer. In Chapter 4.3, we go over how the model's performance is enhanced by this loss function.

4.2.2.3 Hotspot discriminator

The hotspot discriminator is additional discriminator of the SEM generation branch, which connects the hotspot segmentation branch and SEM generation branch. Hotspots in the SEM image are aligned and extracted through region-of-interest alignment [34] according to the results of the hotspot segmentation branch and are entered into the hotspot discriminator to determine whether hotspots are real or fake. To deceive this discriminator, the SEM generation branch was trained to generate hotspots that are more plausible in the generated SEM image. The details are as follows:

$$L_{HD}(G_{SEM}, D_{HD}) = \mathrm{E}_S[\log D_{HD}(S^*)] + \mathrm{E}_C[\log(1 - D_{HD}(G^*_{SEM}(C)))] \tag{4.7}$$

$$\min_{G_{SEM}}\left(\max_{D_{HD}}\left(L_{HD}(G_{SEM}, D_{HD})\right)\right) \tag{4.8}$$

Here, D_{HD}, S^* and $G^*_{SEM}(C)$ are hotspot discriminator, real hotspots extracted from the ground-truth SEM image and fake hotspots extracted from the generated SEM image, respectively.

4.2.2.4 Loss function

Here is the overall loss function of our proposed hotspot prediction model:

$$\min_{G_{SEM}}\left(\begin{array}{l}\max\limits_{D_{SEM}}\left(L_{G1}(G_{SEM}, D_{SEM})\right) + \max\limits_{D_{HD}}\left(L_{HD}(G_{SEM}, D_{HD})\right) \\ +L_{G2}(G_{SEM}) + L_T + \lambda\sum\limits_{N} C(V_i)\end{array}\right) \tag{4.9}$$

4.3 Experiments

We trained the hotspot prediction model using a single Titan RTX 24GB GPU with a batch size of 5 and used Adam [48] as an optimizer with learning rate of 1×10^{-4} for 200 epochs.

4.3.1 Dataset and preprocessing

The dataset to train our proposed hotspot prediction model is from an industrial semiconductor process and consists of SEM images, layouts, and coordinates information of ground-truth bounding boxes containing real hotspots. The size of each image was 1024×1024 pixels consisting of 5,641 pieces each. When hotspots were located at the edge of the image, owing to the lack of layout pattern information for the hotspot from the optical diameter, areas of size 800 × 800 pixels were cut and extracted from the images to prevent the hotspots from being included in the dataset.

To train the hotspot segmentation branch, a ground-truth segmentation map $T_p \in \{0, 1\}^{W\times H}$, where p, W and H denote pixels, width and height of the prediction map respectively, was generated using information of the ground-truth bounding boxes of hotspots, as shown in Figure 4.10. 5,352 training datasets and 289 test datasets are randomly selected from the dataset. We extracted 800 × 800 images from the dataset and resized them to 256 × 256 pixels. We applied standard data augmentation, random color jittering, to the SEM image.

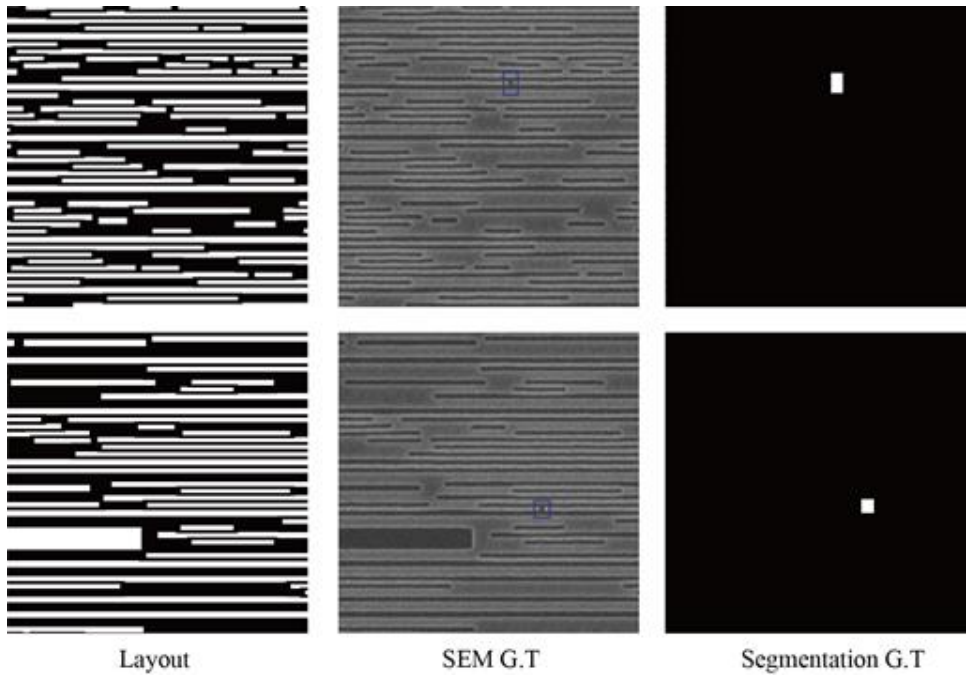

Figure 4.10. Illustration of layout designs, SEM image with ground-truth bounding boxes and ground-truth segmentation images

4.3.1 Performance

The performance was measured on the basis of how close the generated hotspot was to the real hotspot. This metric is similar to the percentage of correct parts used in the pose estimation model [59]; The generated hotspot is considered accurate if it is located within a certain distance of the ground-truth hotspot. We set the distances to 8, 16, 32, and 64 pixels. This metric is used because predicting as many hotspots as possible is very helpful in layout design for efficient development semiconductor lithography process. To evaluate our hotspot prediction model quantitatively, Pix2PixHD and DeepLabv3+, which exhibit excellent performance in supervised image translation and image segmentation tasks, were selected as the baseline.

The baseline model showed an accuracy of only 19.7% in 8-pixels as shown in Table 4.1. Expanding the range to 64 pixels improved the accuracy by 72.7 %; however, the prediction was still considered inaccurate.

In contrast, the combination of Deeplabv3+ and Pix2PixHD significantly contributed to the hotspot prediction performance. The generation accuracy within 8-pixels was 31.3 %, which is an improvement by 11.6 % from the baseline accuracy. Hence, Deeplabv3+ and Pix2PixHD played an crucial role in predicting hotspots in the segmentation model.

The performance of our proposed model is 42.9% in 8-pixels. This model outperformed the baseline model, combination model of Deeplabv3+ and Pix2PixHD, by 11.6 %. An accuracy of 42.9% may appear insufficient. However, the model exhibits an accuracy of 57.9% in16-pixels, 73.3% in 32-pixels, and 89% in 64-pixels. This indicates that the predicted result showed an 89% probability of actual hotspots occurring within 64 pixels. Therefore, this model is considered useful when designing a layout pattern and predicting hotspots.

4.3.2 Ablation study

4.3.2.1 Loss function type

In comparison to the loss function, the performance of the model was evaluated for its efficacy. The impacts of cross-entropy loss, dice loss, and constrained partial cross-entropy loss on performance were examined, as shown in Table 4.2 and Table 4.3.

Cross entropy loss is generally used mainly in classification tasks and is also used in image segmentation models. As shown in Table 4.2, an accuracy of 25.1% was recorded for in 8-pixels after only cross-entropy loss was applied. However, increasing the size of the criteria area does not result in a significant improvement in accuracy. However, our proposed hotspot prediction model showed higher performance at 32.9 % in 8-pixels. The larger the criteria area, the higher the accuracy was evaluated as shown in Table 4.3.

Dice loss is a loss function that employs the intersection over the union and is commonly utilized in segmentation models, especially when data are imbalanced; however, it recorded an accuracy of only 16% in 8-pixels. for hotspot prediction. Conversely, in our proposed hotspot prediction model, as specified in Table 4.3, it displayed better accuracy 37.6 % in 8-pixels than using cross-entropy loss and has a performance suitable for predicting hotspots. In case of constrained partial cross entropy Loss, it was confirmed that the model could learn more intensively in the hotspots, showing 31.3% accuracy in 8-pixels in baseline model. In the case of our prosed hotspot prediction model, it shows better performance 41.0 % in 8-pixels, 55.7 % in 16-pixels, 69.5 % in 32-pixels and 87.1% in 64-pixels. Constrained partial cross entropy loss shows superior performance compared to other loss functions as shown in Table 4.3.

4.3.2.2 Hotspot discriminator

As shown in Table 4.3, the prediction accuracy was improved by 1.9% for in 8-pixels using the hotspot discriminator, and an increase in overall accuracy was observed with respect to other criteria such as in 16-pixels, in 32-pixels and in 64-pixels.

4.3.2.3 Multi-scale features

We also tested how much the use of multi-scale features can help improve performance of the hotspot prediction model. As shown in Table 4,4, it has been shown that using features of various scales has a great effect. The prediction accuracy was improved by 9.7 % for in 8-pixels using multi-scale features from backbone network, Resnet-101 and improvement was observed with respect to other criteria such as in 16-pixels, in 32-pixels and in 64-pixels.

4.3.2.4 Feature matching loss

Table 4.5 demonstrates that utilizing multiple features in feature matching loss increases the prediction performance of our proposed hotspot prediction model; the accuracy was improved by 22.7% for in 8-pixels, and performance improvement was achieved in other criteria such as in 16-pixels, in 32-pixels and in 64-pixels.

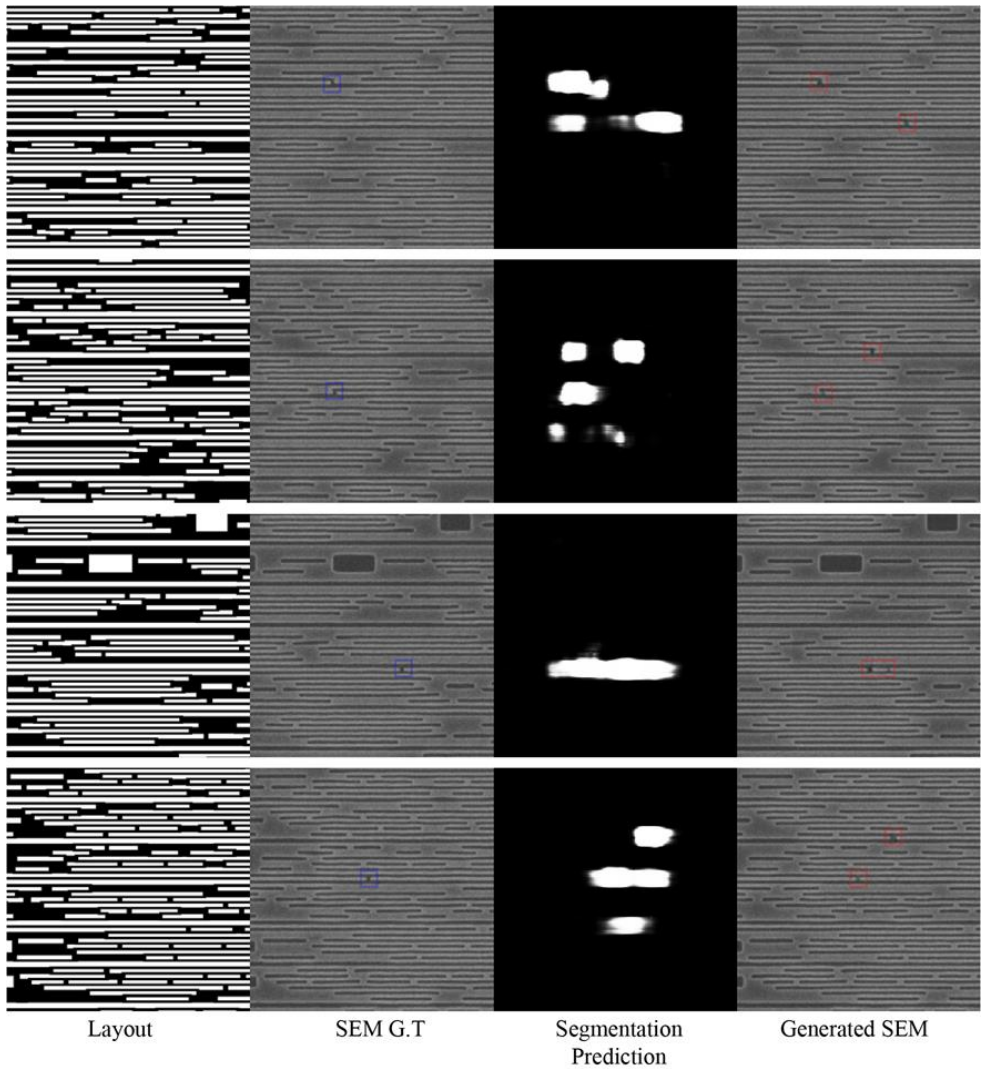

Figure 4.11. Experiment results of hotspot prediction model on the industrial test dataset. The third and fourth columns show segmentation prediction map and generated SEM images, respectively. Segmentation prediction maps predicted by Constrained partial cross entropy loss allow us to guess which patterns are associated with hotspots in layout designs. From the segmentation map of the third column, we can guess that two long horizontal patterns in the layout design were associated with hotspots.

Model	In 8 px	In 16 px	In 32 px	In 64 px
Pix2PixHD	0.197	0.298	0.470	0.727
Deeplabv3+ + Pix2PixHD	0.313	0.438	0.570	0.765
Our Model	0.429	0.579	0.733	0.890

Table 4.1. Comparison of the proposed our model, hotspot prediction model and baselines

Segmentation Loss Type	In 8 px	In 16 px	In 32 px	In 64 px
CE Loss	0.251	0.298	0.345	0.432
Dice Loss	0.160	0.191	0.269	0.398
Constrained Partial CE Loss	0.313	0.438	0.570	0.765

Table 4.2. Ablation experiments for segmentation loss type for baseline, Deeplabv3+ with Pix2PixHD

Segmentation Loss Type			Hotspot Discriminator	In 8 px	In 16 px	In 32 px	In 64 px
CE	Dice	Constrained Partial CE					
O				0.329	0.473	0.652	0.828
	O			0.376	0.517	0.655	0.833
		O		0.410	0.557	0.695	0.871
		O	O	**0.429**	**0.579**	**0.733**	**0.890**

Table 4.3. Ablation experiments for segmentation loss type and hotspot discriminator for our proposed model, hotspot prediction model

Multi-scale Feature	In 8 px	In 16 px	In 32 px	In 64 px
Single-scale	0.332	0.470	0.602	0.749
Multi-scale	**0.429**	**0.579**	**0.733**	**0.890**

Table 4.4. Ablation experiments for the use of multi-scale features in hotspot prediction model

Number of features	In 8 px	In 16 px	In 32 px	In 64 px
0	0.202	0.292	0.420	0.679
1	0.272	0.319	0.469	0.688
2	**0.429**	**0.579**	**0.733**	**0.890**

Table 4.5. Ablation experiments for number of features in feature matching loss

4.4 Conclusion

In this study, we propose a deep learning-based hotspot prediction model that can efficiently help layout design in the semiconductor lithography process. Our proposed hotspot prediction model combines the segmentation model and style-transfer model based on conditional generative adversarial network; unlike other hotspot prediction models, the location and appearance of the hotspots can be directly predicted through the generated SEM image which is output result of our proposed model.

It was founded that the model's overall performance was superior to that of the baseline models. Additionally, we created a technique to roughly identify and recommend layout design's patterns that are hotspot-prone. By allowing for the prediction and modification of layout patterns that are susceptible to hotspots in the layout design stage without the use of photolithography, this method is anticipated to aid in the prevention of the occurrence of hotspots.

Chapter 5

Concluding remark

With the rapid advance of technology, the need for more precise semiconductor process is required. The entire process of semiconductor lithography has become complicated enough to reach thousands of steps, resulting in many unwanted defects called hotspot in this field. Detecting and predicting these hotspots is essential not only for producing flawless electronic products but also for efficient semiconductor photolithography process with high yield and low cost.

In this study, we introduced a hotspot detection and prediction model based on deep learning to improve the efficiency of the semiconductor process. Unlike traditional hotspot detection methods, ours uses an adversarial network architecture that combines a discriminator and a keypoint estimation-based object detection model, thereby producing a more precise and accurate hotspot detection heat map for detecting small-sized hotspots. The keypoint-estimation-based detection model serves as a generator in the Generative Adversarial Network (GAN), producing hotspot detection heatmaps and regressing the width and height of hotspot bounding boxes using multiscale-level features of Resnet-18's backbone. And the model we propose achieves better hotspot detection performance by employing a PixelGAN-styled discriminator that can distinguish up to the details of the hotspot detection heat map, as opposed to a typical discriminator in Generative Adversarial Network discriminator (GAN).

Moreover, we showed a new hotspot prediction model based on deep learning method. The conventional hotspot prediction model can predict the occurrence of hotspots only by layout designs in units of clips of a predetermined size. However,

our proposed hotspot prediction model is a parallel combination of segmentation model and style transfer model utilizing the conditional Generative Adversarial Network (cGAN) architecture. In this model, segmentation model that can helps predict where hotspots occur, as well as translate layout designs into good-quality SEM images to predict the appearance of hotspots. In addition, our hotspot prediction model is trained on constrained partial cross entropy loss function using the prior knowledge, an optical diameter which is the region known to be related in wafer pattern generation, so our model can provide an approximate prediction of which layout pattern caused the hotspot. This is expected to be helpful in layout designing for semiconductor photolithography process. Thanks to the architecture and loss function of the model, our proposed hotspot prediction model shows better performance than conventional hotspot prediction models. We hope that our proposed hotspot detection model and prediction model will help construct a more efficient semiconductor photolithography process.

For further studies, we are considering two things. The first is an improvement in the hotspot detection model. The limitation of the hotspot detection model is that a large amount of datasets are required. Labeling this large amount of data sets is quite inefficient because it takes a huge amount of time. In addition, securing many datasets for each semiconductor process is also expensive. Therefore, it is necessary to develop the hotspot detection model that can produce high performance even with a small amount of datasets. One of the possible options is to apply the zero shot/few shot learning technique, which is currently widely studied. If the performance of the currently developed hotspot detection model and hotspot prediction model can be reached with only a small number of datasets, it will be of great help to the efficient lithography process. The second is a method of using the hotspot prediction model proposed in this study. Predicting hotspots and automatically modifying layout patterns to suppress their occurrence is expected to be of great help in the development of more efficient semiconductor process.

Bibliography

[1] F. Schellenberg, “A Little Light Magic,” *IEEE Spectrum*, vol. 40, no. 9, pp. 34–39, 2003, doi: 10.1109/MSPEC.2003.1228007.

[2] J. Li and L. S. Melvin, “Sub-resolution assist features in photolithography process simulation,” *Digest of Papers - Microprocesses and Nanotechnology 2007; 20th International Microprocesses and Nanotechnology Conference, MNC*, pp. 394–395, 2007, doi: 10.1109/IMNC.2007.4456270.

[3] M. D. Levenson, N. S. Viswanathan, and R. A. Simpson, “Improving Resolution in Photolithography with a Phase-Shifting Mask,” *IEEE Transactions on Electron Devices*, vol. 29, no. 12, pp. 1828–1836, 1982, doi: 10.1109/T-ED.1982.21037.

[4] X. Ma, “Resolution enhancement optimization methods in optical lithography with improved manufacturability,” *Journal of Micro/Nanolithography, MEMS, and MOEMS*, vol. 10, no. 2, p. 023009, Apr. 2011, doi: 10.1117/1.3590252.

[5] T. Hiroi and H. Okuda, “Robust Defect Detection System Using Double Reference Image Averaging for High Throughput SEM Inspection Tool,” in *The 17th Annual SEMI/IEEE ASMC 2006 Conference*, pp. 347–352. doi: 10.1109/ASMC.2006.1638781.

[6] T. Kitamura *et al.*, “Introduction of a die-to-database verification tool for the entire printed geometry of a die: geometry verification system NGR2100 for DFM,” in *Proc. SPIE 5853, Photomask and Next-Generation Lithography Mask Technology XII*, Jun. 2005, vol. 5853, p. 988. doi: 10.1117/12.620389.

[7] R. Nakagaki, T. Honda, and K. Nakamae, “Automatic recognition of defect areas on a semiconductor wafer using multiple scanning electron microscope images,” *Measurement Science and Technology*, vol. 20, no. 7, 2009, doi: 10.1088/0957-0233/20/7/075503.

[8] M. Harada, Y. Minekawa, and K. Nakamae, “Defect detection techniques

robust to process variation in semiconductor inspection," *Measurement Science and Technology*, vol. 30, no. 3, Feb. 2019, doi: 10.1088/1361-6501/aafd77.

[9] A. B. Kahng, C.-H. Park, and X. Xu, "Fast dual graph-based hotspot detection," in *n", Proc. SPIE 6349, Photomask Technology 2006*, Oct. 2006, vol. 6349. doi: 10.1117/12.692949.

[10] H. Nosato *et al.*, "Hotspot prevention and detection method using an image-recognition technique based on higher-order local autocorrelation," *Journal of Micro/Nanolithography, MEMS, and MOEMS*, vol. 13, no. 1, p. 011007, Feb. 2014, doi: 10.1117/1.JMM.13.1.011007.

[11] S. Shim and Y. Shin, "Topology-oriented pattern extraction and classification for synthesizing lithography test patterns," *Journal of Micro/Nanolithography, MEMS, and MOEMS*, vol. 14, no. 1, p. 013503, Jan. 2015, doi: 10.1117/1.jmm.14.1.013503.

[12] D. Patel, R. K. Bonam, and A. Oberai, "Engineering neural networks for improved defect detection and classification," in *Metrology, Inspection, and Process Control for Microlithography XXXIII*, Mar. 2019, p. 34. doi: 10.1117/12.2515065.

[13] K. Imoto, T. Nakai, T. Ike, K. Haruki, and Y. Sato, "A CNN-Based transfer learning method for defect classification in semiconductor manufacturing," *IEEE Transactions on Semiconductor Manufacturing*, vol. 32, no. 4, pp. 455–459, Nov. 2019, doi: 10.1109/TSM.2019.2941752.

[14] H. Yang, L. Luo, J. Su, C. Lin, and B. Yu, "Imbalance aware lithography hotspot detection: a deep learning approach," *Journal of Micro/Nanolithography, MEMS, and MOEMS*, vol. 16, no. 03, p. 1, Aug. 2017, doi: 10.1117/1.jmm.16.3.033504.

[15] H. Yang, J. Su, Y. Zou, Y. Ma, B. Yu, and E. F. Y. Young, "Layout Hotspot Detection with Feature Tensor Generation and Deep Biased Learning," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Jun. 2019, vol. 38, no. 6, pp. 1175–1187. doi:

10.1109/TCAD.2018.2837078.

[16] M. Shin and J.-H. Lee, “Accurate lithography hotspot detection using deep convolutional neural networks,” *Journal of Micro/Nanolithography, MEMS, and MOEMS*, vol. 15, no. 4, p. 043507, Nov. 2016, doi: 10.1117/1.jmm.15.4.043507.

[17] M. Ouchi *et al.*, “A trainable Die-To-Database for fast e-Beam inspection: learning normal images to detect defects,” Mar. 2020.

[18] H. Yao, S. Sinha, C. Chiang, X. Hong, and Y. Cai, “Efficient process-hotspot detection using range pattern matching,” *IEEE/ACM International Conference on Computer-Aided Design, Digest of Technical Papers, ICCAD*, pp. 625–632, 2006, doi: 10.1109/ICCAD.2006.320026.

[19] R. v. Joshi *et al.*, “Super fast physics-based methodology for accurate memory yield prediction,” *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 23, no. 3, pp. 534–543, Mar. 2015, doi: 10.1109/TVLSI.2014.2313815.

[20] P. Kumar, B. Srinivasan, and N. R. Mohapatra, “Fast and accurate lithography simulation using cluster analysis in resist model building,” *Journal of Micro/Nanolithography, MEMS, and MOEMS*, vol. 14, no. 2, p. 023506, May 2015, doi: 10.1117/1.JMM.14.2.023506.

[21] D. Ding, J. A. Torres, and D. Z. Pan, “High performance lithography hotspot detection with successively refined pattern identifications and machine learning,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 30, no. 11, pp. 1621–1634, Nov. 2011, doi: 10.1109/TCAD.2011.2164537.

[22] S.-Y. Lin, J.-Y. Chen, J.-C. Li, W. Wen, and S.-C. Chang, “A novel fuzzy matching model for lithography hotspot detection,” in *2013 50th ACM/EDAC/IEEE Design Automation Conference (DAC)*, 2013, pp. 1–6.

[23] 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,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*,

vol. 33, no. 11, pp. 1671–1680, Nov. 2014, doi: 10.1109/TCAD.2014.2351273.

[24] T. Matsunawa, J.-R. Gao, B. Yu, and D. Z. Pan, "A new lithography hotspot detection framework based on AdaBoost classifier and simplified feature extraction," *Proc. SPIE 9427, Design-Process-Technology Co-optimization for Manufacturability IX*, Mar. 2015, doi: 10.1117/12.2085790.

[25] 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," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 34, no. 3, pp. 460–470, Mar. 2015, doi: 10.1109/TCAD.2014.2387858.

[26] H. Yang, Y. Lin, B. Yu, and E. F. Y. Young, "Lithography hotspot detection: From shallow to deep learning," *International System on Chip Conference*, vol. 2017-September, pp. 233–238, Dec. 2017, doi: 10.1109/SOCC.2017.8226047.

[27] H. Zhang, B. Yu, and E. F. Y. Young, "Enabling online learning in lithography hotspot detection with information-theoretic feature optimization," *IEEE/ACM International Conference on Computer-Aided Design, Digest of Technical Papers, ICCAD*, vol. 07-10-November-2016, Nov. 2016, doi: 10.1145/2966986.2967032.

[28] H. Yang, J. Su, Y. Zou, Y. Ma, B. Yu, and E. F. Y. Young, "Layout Hotspot Detection with Feature Tensor Generation and Deep Biased Learning," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 38, no. 6, pp. 1175–1187, Jun. 2019, doi: 10.1109/TCAD.2018.2837078.

[29] W. Ye, Y. Lin, M. Li, Q. Liu, and D. Z. Pan, "Lithoroc: Lithography hotspot detection with explicit ROC optimization," *Proceedings of the Asia and South Pacific Design Automation Conference, ASP-DAC*, pp. 292–298, Jan. 2019, doi: 10.1145/3287624.3288746.

[30] R. Chen *et al.*, "Faster Region-Based Hotspot Detection," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 41, no.

3, pp. 669–680, Mar. 2022, doi: 10.1109/TCAD.2020.3021663.

[31] I. Goodfellow *et al.*, “Generative adversarial networks,” *Commun ACM*, vol. 63, no. 11, pp. 139–144, Oct. 2020, doi: 10.1145/3422622.

[32] M. Mirza and S. Osindero, “Conditional Generative Adversarial Nets,” Nov. 2014, [Online]. Available: http://arxiv.org/abs/1411.1784

[33] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017, doi: 10.1109/TPAMI.2016.2577031.

[34] K. He, G. Gkioxari, P. Dollar, and R. Girshick, “Mask R-CNN,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 2, pp. 386–397, Feb. 2020, doi: 10.1109/TPAMI.2018.2844175.

[35] B. Singh, M. Najibi, and L. S. Davis, “SNIPER: Efficient Multi-Scale Training,” *Advances in Neural Information Processing Systems*, vol. 2018-December, pp. 9310–9320, May 2018, doi: 10.48550/arxiv.1805.09300.

[36] J. Redmon and A. Farhadi, “YOLO9000: Better, Faster, Stronger,” Dec. 2016, [Online]. Available: http://arxiv.org/abs/1612.08242

[37] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar, “Focal Loss for Dense Object Detection,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 2, pp. 318–327, Feb. 2020, doi: 10.1109/TPAMI.2018.2858826.

[38] D. and E. D. and S. C. and R. S. and F. C.-Y. and B. A. C. Liu Wei and Anguelov, “SSD: Single Shot MultiBox Detector,” in *Computer Vision – ECCV 2016*, 2016, pp. 21–37.

[39] H. Law and J. Deng, “CornerNet: Detecting Objects as Paired Keypoints,” *Proceedings of the European conference on computer vision (ECCV)*, pp. 734–750, Aug. 2018, [Online]. Available: http://arxiv.org/abs/1808.01244

[40] X. Zhou, J. Zhuo, and P. Krahenbuhl, “Bottom-up Object Detection by Grouping Extreme and Center Points,” *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2019-

June, pp. 850–859, Jan. 2019, doi: 10.48550/arxiv.1901.08043.

[41] X. Zhou, D. Wang, and P. Krähenbühl, “Objects as Points,” Apr. 2019, doi: 10.48550/arxiv.1904.07850.

[42] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, “Image-to-Image Translation with Conditional Adversarial Networks,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jul. 2017, pp. 5967–5976. doi: 10.1109/CVPR.2017.632.

[43] T. C. Wang, M. Y. Liu, J. Y. Zhu, A. Tao, J. Kautz, and B. Catanzaro, “High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs,” *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 8798–8807, Dec. 2018, doi: 10.1109/CVPR.2018.00917.

[44] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 2016-December, pp. 770–778, Jun. 2016, doi: 10.1109/CVPR.2016.90.

[45] T.-Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, and S. Belongie, “Feature Pyramid Networks for Object Detection,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jul. 2017, pp. 936–944. doi: 10.1109/CVPR.2017.106.

[46] L. Cui *et al.*, “MDSSD: multi-scale deconvolutional single shot detector for small objects,” *Science China Information Sciences*, vol. 63, no. 2, Feb. 2020, doi: 10.1007/s11432-019-2723-1.

[47] X. Mao, Q. Li, H. Xie, R. Y. K. Lau, Z. Wang, and S. P. Smolley, “Least Squares Generative Adversarial Networks,” in *2017 IEEE International Conference on Computer Vision (ICCV)*, Oct. 2017, pp. 2813–2821. doi: 10.1109/ICCV.2017.304.

[48] D. P. Kingma and J. L. Ba, “Adam: A method for stochastic optimization,” *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 2015.

[49] E. Shelhamer, J. Long, and T. Darrell, “Fully Convolutional Networks for Semantic Segmentation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 4, pp. 640–651, Apr. 2017, doi: 10.1109/TPAMI.2016.2572683.

[50] L. C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, “DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 4, pp. 834–848, Apr. 2018, doi: 10.1109/TPAMI.2017.2699184.

[51] H. Noh, S. Hong, and B. Han, “Learning Deconvolution Network for Semantic Segmentation,” in *2015 IEEE International Conference on Computer Vision (ICCV)*, Dec. 2015, pp. 1520–1528. doi: 10.1109/ICCV.2015.178.

[52] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” Springer Verlag, 2015, pp. 234–241. doi: 10.1007/978-3-319-24574-4_28.

[53] L. C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, “Encoder-decoder with atrous separable convolution for semantic image segmentation,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11211 LNCS, pp. 833–851, 2018, doi: 10.1007/978-3-030-01234-2_49/COVER/.

[54] H. Kervadec, J. Dolz, M. Tang, E. Granger, Y. Boykov, and I. ben Ayed, “Constrained-CNN losses for weakly supervised segmentation,” *Medical Image Analysis*, vol. 54, pp. 88–99, May 2019, doi: 10.1016/J.MEDIA.2019.02.009.

[55] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, “Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation,” Sep. 2018.

[56] A. Khoreva, R. Benenson, J. Hosang, M. Hein, and B. Schiele, “Simple Does It: Weakly Supervised Instance and Semantic Segmentation,” *2017 IEEE*

Conference on Computer Vision and Pattern Recognition (CVPR), vol. 2017-January, pp. 1665–1674, Jul. 2017, doi: 10.1109/CVPR.2017.181.

[57] J. Lee, E. Kim, S. Lee, J. Lee, and S. Yoon, "Ficklenet: Weakly and semi-supervised semantic image segmentation using stochastic inference," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2019-June, pp. 5262–5271, Jun. 2019, doi: 10.1109/CVPR.2019.00541.

[58] B. M. la Fontaine *et al.*, "Flare and its impact on low-k 1 KrF and ArF lithography," in *Proc. SPIE 4691, Optical Microlithography XV*, Jul. 2002, vol. 4691. doi: 10.1117/12.474598.

[59] A. Toshev and C. Szegedy, "DeepPose: Human pose estimation via deep neural networks," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1653–1660, Sep. 2014, doi: 10.1109/CVPR.2014.214.

국 문 초 록

트랜지스터와 집적회로의 발명 이후, 반도체 공정의 발전은 수십 년 동안 급속하게 성장했다. 현재의 마이크로칩은 수억 개의 트랜지스터를 포함하고 있을 정도이다. 반도체 공정 기술과 소자구조의 복잡성 증가로 인해 제조 공정에서 핫스팟이라 불리는 많은 결함들이 발생하게 되었다. 반도체 수율 개선 및 비용 절감을 위해선 신속하고 효율적으로 핫스팟들을 검출하는 것뿐만 아니라 반도체 공정을 수행하기 전에 핫스팟의 발생 여부를 미리 예측하는 것이 필수적이다. 그렇기 때문에 핫스팟을 검출하거나 예측하기위한 많은 시도가 이루어지고 있다. 특히 핫스팟의 발생을 미리 예측하는 것은 효율적인 반도체 공정 개발을 위해 꼭 필요하나 매우 어렵기 때문에 이와 관련된 많은 연구가 현재 진행되고 있다. 본 연구에서는 보다 효율적인 반도체 공정 개발 및 제조 공정을 위해 기존의 핫스팟 검출 및 예측 모델보다 더 나은 성능을 보여주는 딥 러닝 기반 핫스팟 검출 및 예측 모델을 제안하고자 한다.

본 연구에서 제시하는 핫스팟 검출 모델은 Conditional Generative Adversarial Network의 구조를 도입하여 좀 더 정밀한 검출 히트맵을 생성함으로써 기존의 핫스팟 검출 모델보다 더 정확하고 빠르게 작은 사이즈의 핫스팟들을 검출할 수 있다. 또한 본 연구에서 제안하는 핫스팟 예측 모델의 경우 Segmentation 모델과 Style Transfer 모델을 Conditional Generative Adversarial Network 기반으로 하여 병렬로 합친 구조로 레이아웃 디자인에서 변환되어 생성된 SEM 이미지를 통해 핫스팟이 어디서 생성이 되는지 그리고 또 핫스팟이 어떤 모습일지 예측할 수 있다. 또한 본 연구에서 제안하는 핫스팟 예측 모델은 Constrained Partial Cross Entropy 손실 함수를 사용하여 레이아웃 디자인 패턴들 중 어느 패턴들이 핫스팟 발생에 관여했는지 추측할 수 있다. 이를 이용하면 레이아웃 패턴 디자인에 큰 도움이 될 것이기 때문에 더 효율적인 반도체 공정 개발이 가능할 것으로 기대된다. 본 연구에서 제시하는 핫스팟 검출 및 예측 모델의 성능은 실제 산업 데이터로 테스트하였고 기존 모델보다 더 나은 성능을 보여주었다.

주요어 : 반도체 공정, 포토리소그래피 공정, 핫스팟, 핫스팟 검출, 핫스팟 예측, 딥러닝

학 번 : 2016-37799