Wafer SEM Image Generation with Conditional Generative Adversarial Network

Huqiang Du\textsuperscript{1*}, Zheng Shi\textsuperscript{2}

\textsuperscript{1}College of Information Science & Electronic Engineering, Zhejiang University, Zhejiang, 310027, China
\textsuperscript{2}Corresponding author’s e-mail: 1499807918@qq.com

Abstract. Scanning Electron Microscope (SEM) images play an essential role in the analysis and evaluation of the defects of the circuit in advanced integrated circuit manufacturing. Currently, the image generation method is a useful means to solve the insufficiency of wafer SEM images due to the high cost of getting a large number of labeled SEM images. In this paper, an algorithm based on conditional generative adversarial network (cGANs) is proposed for SEM image generation. Firstly, a Sobel operator is used to calculate the gradient information of images to guide the discriminator. Then we apply two discriminators to discriminate images of different resolutions. Finally, Wasserstein distance and smooth L1 loss functions are applied to accelerate network convergence. Experimental results show that this method could learn to mimic the distribution of wafer SEM image data effectively. Compared with other image generating models, our model improves the quality of the generated image, and its 1-Nearest-Neighbour (1-NN) classification score is increased by 0.3. Therefore, this algorithm is more suitable for generating images to alleviate the shortage of wafer SEM images.

1. Introduction
An integrated circuit (IC) is an electronic component which is widely applied in industry and daily life. Its manufacture is complex and costly, which involves hundreds of steps, and many process parameters need to be monitored in the whole production process [1]. In the process of chip manufacturing, every process may produce structures that do not conform to expectations. And these structures often become wafer defects that make the circuits on the chip fail to work [2]. Today, the defects cannot be avoided, even precisely positioned equipment operated by well-trained process engineers in the near-dust-free-clean room. The existence of these defects often affects the reliability of the device and the yield of chip manufacturing [1]. However, reliability and yield will remain the concerns to the IC industry in the future. They are also two important factors affecting the profitability of semiconductor manufacturers [3]. Therefore, it is important for process engineers to effectively identify the causes of production defects and functional failures in the manufacturing process of IC.

Traditionally, the defects are identified and classified through visual inspection by quality engineers using a scanning electron microscope. Since a large amount of data are generated in the semiconductor manufacturing process, manual detection becomes time-consuming and laborious, and the error rate is high [1].

With the further study of machine learning, methods such as K nearest neighbor, support vector machine, hidden Markov, and neural network are increasingly used for wafer surface defect detection. Most of these pattern recognition methods require a large number of SEM images to train the classifier [4]. However, in the actual production process, the acquisition of these SEM images is costly and time-
consuming. And due to the noise of the SEM image itself and some unpredictable changes in the device manufacturing process, it is often unrealistic to obtain a large amount of labeled data.

To obtain more wafer SEM images, using algorithms is a feasible method to generate images manually. The existing algorithms used to generate SEM images include traditional simulator methods and the latest deep learning methods. Traditional simulator methods include the Monte Carlo simulator and the ASEM image analysis simulator proposed by Reference [5]. Since these simulators are not specifically used to generate SEM image samples, the synthesized SEM images are of poor quality. The methods based on deep learning mainly include generative adversarial networks (GANs) and its derivative models [6]. The existing GANs can effectively capture the data distribution of the image and predict the image based on this distribution. However, the typical generative adversarial network model has a significant mode collapse problem. Consequently, the generated SEM images suffer from coherent problems such as blurred circuit edges and uneven background noise distribution [7].

According to the characteristics of the wafer SEM image, this paper proposes an algorithm based on the pix2pix model to realize the transformation of the circuit design layout to the wafer SEM image [8]. Compared with other algorithms for generating SEM images, the main contribution of this paper is as follows:

1) According to the characteristics of the generated image and the consistency of the real image characteristics, we use the pix2pix model as the basic architecture and add a new discriminator structure to discriminate images of different resolutions. The image gradient information guides the generator to generate a distribution that more accurately matches the real distribution and improves the quality of the generated image.

2) Wasserstein distance and the smooth L1 loss function are combined to accelerate the model's convergence.

The algorithm in this paper not only meets the characteristics of high-resolution and high-definition for generated images but also avoids overfitting due to the limited number of training sets, which improves the robustness of the algorithm.

2. Related Work

2.1 Pix2Pix Model
Pix2Pix model is a supervised learning method proposed by P Isola and J Y Zhu, et al. (2016), which is used to transform a visual scene to another visual scene. The network uses a set of pictures with similar content but different scenes as input. In Pix2Pix model, the generator uses a U-Net architecture, which allows low-level information to shortcut across the network [8]. So low-frequency information between input and output is shared, and the loss of high-frequency information is reduced [9]. PatchGAN is used in the discriminator, which only penalized structure at the scale of patches, so it is suitable for modeling high-frequencies [8].

3. Method

3.1 Architecture
The model used in this paper is based on the Pix2Pix model; the framework is showed as infigure.1. In the generator, all convolutional and deconvolutional layers use a 5x5 spatial filter with a step size of 2. Leaky ReLU with a slope of 0.2 is used as the non-linear activation function in the convolutional layer, and Batch-normalization (BN) was used to accelerate convergence. In the deconvolution layer, ReLU is used as a non-linear activation function, and Dropout with a probability of 0.5 is used as a regularization method to prevent overfitting. Finally, the generated image is gotten after the Tanh activation function. In discriminator, we used Wasserstein distance to solve the unstable convergence of the GANs network [10].
Figure 1. The framework of algorithm for wafer SEM image generation

To match the distribution generated by the generator with training distribution better, we add a discriminator $D_2$ in our model. In this paper, the inputs of generator $G$ are the wafer layout image $x$ and random noise $z$. And the inputs of the discriminator $D$ are a layout image combined with its SEM image or generated image and a gradient image calculated by a Sobel operator. The discriminator $D_1$ and $D_2$ have the same structure, and the only difference between them is that the input of $D_1$ is input to $D_2$ after being down-sampled.

### 3.2 Loss Function

The loss functions in this paper are as follows:

$$l_1^G(x, v) = D_1(G(x, z)) - D_1(v)$$  \hspace{1cm} (1)

$$l_2^G(x, v) = D_2(Sobel(G(x, z))) - D_2(sobel(v))$$  \hspace{1cm} (2)

$$l^g(x, v) = \lambda_1 R(G(x, z)) - v + \lambda_2 R(G(x, z) - v) - D_1(G(x, z), Sobel(G(x, z), x)) - D_2(G(x, z), Sobel(G(x, z), x))$$  \hspace{1cm} (3)

Where $\lambda_1$ and $\lambda_2$ are constants, $R$ is a smooth L1 loss function [11]. Equation (1) and equation (2) force $D$ to distinguish between generated samples and real images, while equation (3) is used to train $G$ to fool $D$ with generated samples.

### 4. Datasets and Implementation Details

#### 4.1 Datasets

The datasets we used for generating SEM images are collected from a fab, and the details are as follows:

- **Circuit layout images**: The circuit layout images used in this paper are RGB pictures of the real circuit layout read from the GDS file. All the extracted circuit layout pictures have three layers, and the size is $256 \times 256$ pixels. Part of the circuit layout pictures is shown in figure 2a-c.

- **Wafer SEM images**: The wafer SEM images used in this paper are grayscale images from a fab, and we used three different datasets to test out model. The sizes are $256 \times 256$ pixels, and there are 8000 images for every dataset. Part of wafer SEM images is shown in figure 2d-f.

We randomly select 80% of the images of each dataset as training sets, and the rest is used as the testing sets.
Figure 2. Datasets used in our experiments. (a) to (c) are layout images, and (d) to (e) are wafer SEM images.

4.2 Experiment Procedure

The experiments in this paper are carried out on Tensorflow platform. During the experiment, all models use the Adam optimization algorithm. The initial value of the learning rate was 0.00015. After 10 iterations, the learning rate dropped to 0.0001. To increase the dataset, we reversed the training images randomly during the training process. At the same time, the image sizes were enlarged using bilinear interpolation, then we clipped these images to the size of the original picture. With a single NVIDIA GTX 1070ti GPU, the entire training process costs 3 hours.

5. Experimental Results and Analysis

5.1 Assessment Method

It is hard to compare our models with different typical methods generating the wafer SEM images because it is a new application in the wafer SEM image area. There is no paper to solve the same problem before as far as we have known. Hence, we compare our models with three popular networks used for generating images, including Pix2Pix, Pix2PixHD, and DualGAN [12-13]. Part of generated SEM images is shown in figure 3. We used a 1-NN classifier and Peak Signal to Noise Ratio (PSNR) algorithm to evaluate the models. The 1-NN classifier is the best indicator for evaluating GAN networks at present [7]. Currently, The PSNR algorithm is widely used as an index for evaluating the similarity of two pictures.

5.2 Experimental Results and Analysis

![Figure 3. Different models generate different quality of images. Each column shows results generated by a different model.](image)

| Method   | PSNR(dB) | 1-NN Score |
|----------|----------|------------|
| Pix2Pix  | 29.32    | 1          |
| Pix2PixHD| 29.45    | 1          |
| DualGAN  | 29.32    | 1          |
| Ours     | 29.84    | 0.623      |

Table 1. Comparison of experimental results in SEM image dataset 1.

| Method   | PSNR(dB) | 1-NN Score |
|----------|----------|------------|
| Pix2Pix  |          |            |
| Pix2PixHD|          |            |
| DualGAN  |          |            |
| Ours     |          |            |

Table 2. Comparison of experimental results in SEM image dataset 2.
Table 3. Comparison of experimental results in SEM image dataset 3.

| Method   | PSNR(dB) | 1-NN Score |
|----------|----------|------------|
| Pix2Pix  | 29.43    | 1          |
| Pix2PixHD| 29.37    | 1          |
| DualGAN  | 29.22    | 1          |
| Ours     | 29.53    | 0.668      |

Table 3. Comparison of experimental results in SEM image dataset 3.

To explain the rationality of the experimental results, we used the AlexNet network as a 1-NN classifier to classify an image from real SEM images or synthesized images and used the PSNR to compare the gray pixel values of the generated images with the real images. A 1-NN classifier score close to 0.5, indicates that the generated image has perfect similarity with the real image. And a higher PSNR indicates that a generated wafer SEM image is of higher quality.

Finally, we mixed the test sets with the corresponding generated SEM image sets, the average PSNR value and 1-NN classifier score of each wafer SEM image dataset are shown in Tables 1 to 3.

As can be seen from Table 1-3, the PSNR values of our method are higher than the other three algorithms, indicating that this algorithm is superior to other algorithms in terms of pixel values. When using the 1-NN classifier to evaluate the generated image and the real image, the score is significantly better than the other three algorithms. The 1-NN scores of our method are closer to 0.5, indicating that the classifier can no longer classify the images correctly. The experimental generated images by the algorithm show consistent behaviors with the real images in terms of low- and high-dimensional features. In wafer defect detection, the characteristics of defects often show significant differences in low dimensions. Therefore, the algorithm proposed in this paper achieves considerable improvements in solving the shortage of real SEM images to some extent.

6. Summary

In this paper, we present a new method used to generate wafer SEM images, that can successfully generate new SEM images that look realistic. Compared with other traditional deep learning algorithms, the images generated by this algorithm are more consistent with real SEM images in terms of high-dimensional features. The experimental results show that the algorithm can alleviate the difficulties with wafer surface defect detection by adding more realistic SEM images. In the future, we will continue to improve our model and add more functions to it, for instance, generating SEM images via the model with specified defects. Furthermore, the distribution of generated images should match the distribution of real images better, which makes the 1-NN score closer to the perfect score.

Acknowledgement

This study is supported by Institute of VLSI Design, Zhejiang University and DMO Corporation.

References

[1] Wang, C. H., Kuo, W., & Bensmail, H. (2006) Detection and classification of defect patterns on semiconductor wafers. IIE transactions, 38(12), 1059-1068.

[2] Fang, X. (2018) Study of Defect Detection and Classification of Wafer SEM Image based on CNN. Zhejiang University, Hangzhou.

[3] Kuo, W., & Kim, T. (1999) An overview of manufacturing yield and reliability modeling for semiconductor products. Proceedings of the IEEE, 87: 1329-1344.

[4] Lin, J., Wang, H. M., Yu, N. G., et al. (2018) Research on Online Detection of Wafer Surface Defects. Computer Measurement & Control, v.26, 236:20-22+26
[5] Babin, S., Borisov, S. S., Trifonenkov, V. P. (2015) Fast analytical modeling of SEM images at a high level of accuracy. In: SPIE Advanced Lithography. San Jose. Proc. of SPIE Vol. 9424 94240I-1

[6] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Bengio, Y., et al. (2014) Generative adversarial nets. In: Advances in neural information processing systems. Montreal, pp. 2672-2680.

[7] Xu, Q., Huang, G., Yuan, Y., Guo, C., Sun, Y., Wu, F., & Weinberger, K. (2018) An empirical study on evaluation metrics of generative adversarial networks. arXiv preprint arXiv:1806.07755.

[8] Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. (2017) Image-to-image translation with conditional adversarial networks. In: The IEEE conference on computer vision and pattern recognition. Hawaii. pp. 1125-1134.

[9] Li, C., Wand, M. (2016) Precomputed real-time texture synthesis with markovian generative adversarial networks, In European Conference on Computer Vision. Amsterdam. pp. 702-716.

[10] Arjovsky, M., Chintala, S., Bottou, L. (2017) Wasserstein gan. arXiv preprint arXiv:1701.07875.

[11] Girshick, R. (2015) Fast r-cnn. In: The IEEE international conference on computer vision. Santiago. pp. 1440-1448.

[12] Wang, T. C., Liu, M. Y., Zhu, J. Y., Tao, A., Kautz, J., Catanzaro, B. (2018) High-resolution image synthesis and semantic manipulation with conditional gans. In: The IEEE conference on computer vision and pattern recognition. Kauai. pp. 8798-8807.

[13] Yi, Z., Zhang, H., Tan, P., Gong, M. (2017) Dualgan: Unsupervised dual learning for image-to-image translation. In: The IEEE international conference on computer vision. Venice. pp. 2849-2857.