Generating Scanning Electron Microscopic Images from Optical Designs using image to image translation with Conditional GANs.

Jacob George
KLA-Tencor, India
Jacob.george@kla-tencor.com
jacobgeo@stanford.edu

Abstract

Detection of defects during semiconductor manufacturing is critical in enhancing the production yield of electronic chips. Scanning Electron Microscope (SEM) tools are used to identify these defects by scanning the silicon wafers in each manufacturing step. In this work, a new methodology is implemented to enhance the defect detection by artificially generating SEM images from optical designs of the electronic chip using conditional GANs[1]. These images could potentially serve as reference images to the original images obtained from SEM tools to identify the defects. Moreover these could also serve as a chip fabrication simulation from its optical design.

1. Introduction

Simulating the fabrication of silicon chips from its optical designs is crucial for the success of semi-conductor industry for creating optimum designs and identifying defects in the early steps of manufacturing process. Because of the complexity involved in the optical physics of the photolithography process, it is a hard problem to simulate the chip fabrication. Often what is intended in the design may not get fabricated in the chip and therefore, lot of optimizing steps of designing the silicon chips involve multiple trial and error fabrication of various designs. With the recent development in the Deep learning especially with regards to GAN networks [2], artificially synthesizing real world images has come into a reality. This work explores the possibility of simulating the chip fabrication from its optical design and translating it to images as seen under SEM tool with the help of conditional GAN networks [1], and thereby could potentially help in reducing the effective cost of the chip manufacturing. Simulating chip fabrication from optical design to SEM images serves two main purpose:

1) Model the physical relationship between optical design to the actual chip fabrication and thereby provide faster and optimized way of creating right designs.
2) It also helps in identifying any defects in the chip fabrication by comparing the simulated Image to the actual grabbed image from SEM microscope. These defect detections are carried out in factory automation without any manual intervention in the SEM tools.

Generating SEM images from the optical design images can be treated as an Image-to-Image translation problem in which an input image from one visual space is mapped on to a different visual space without altering the context. In this work, the Image-to-Image translation with conditional GAN network as described by Isola et. al. **[1]** is used to demonstrate that Image translation from Optical design to SEM images could be used to simulate chip fabrication from optical images. The input for the trained network is a binary image rendered from optical design and the network generates SEM image which is a single channel grayscale image (See Fig 1).

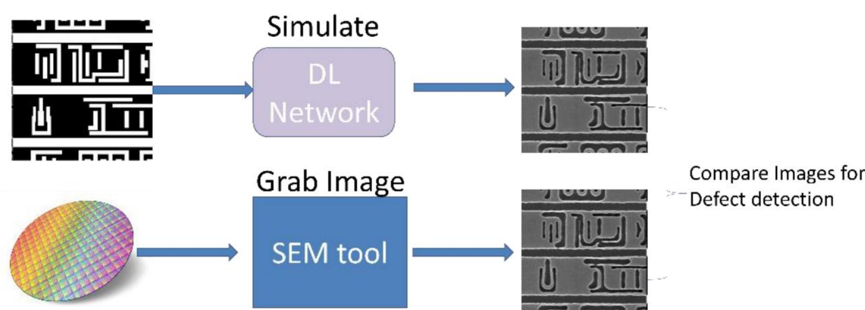

Figure 1. The proposed application

2. Dataset

To train GAN networks, the data was prepared by obtaining optical designs, rendering them to binary images and grabbing its corresponding SEM images from the Semiconductor wafers using KLA-Tencor Ebeam SEM review tool. In order to train the network, these images need to be aligned and therefore image alignment is done using image correlation. The SEM images were downscaled to image size of 256x256 (from its grabbed size of 720x720), while the design images were upscaled to the same size (from 139x139). These aligned image pairs form the data set for training the GAN network. 500 such aligned image pair samples were obtained and partitioned into two sets: 400 for training and 100 for validation.

3. Method

The conditional GAN network consists of two competing neural networks, the Generator and the Discriminator. The goal of the Generator here is to generate the SEM image from the binary design image, while the goal of the Discriminator is to identify whether the given SEM image is generated from Generator network or is it the original grabbed image from the SEM tool for a given optical design. Fig. 2 illustrates this.

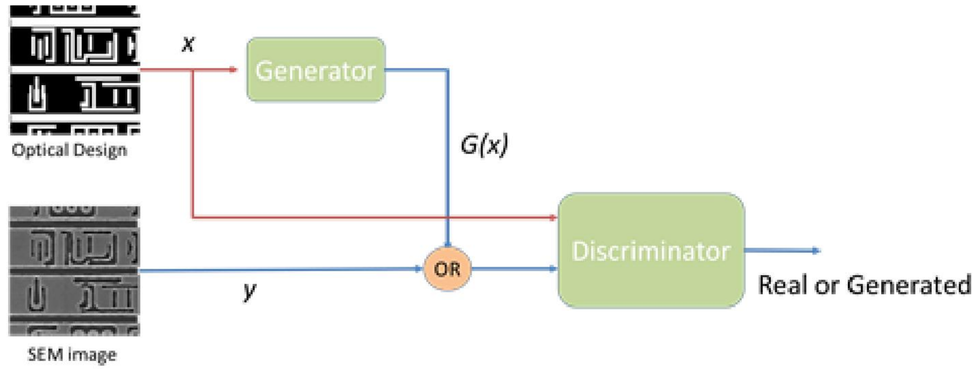

Figure 2. The GAN network architecture

On optimally training the GAN network, the generator network can be used to simulate the chip fabrication from the optical design. Since the goal of the generator is not only to fool the discriminator network, but also to generate images as close to the real images, an L1 loss with respect to the real image needs to be minimized besides adversarial loss. As per Isola et. al. [1], the choice of L1 loss over L2 is because L2 loss causes undesirable blurring effect on the generated images. The L1 loss of the generator is given by

$$L_{L1}(G) = \mathbb{E}_{x,y,z}[\|y - G(x)\|_1]$$

Whereas the adversarial loss can be expressed as

$$L_{cGAN}(G,D) = \mathbb{E}_{x,y}[\log(D(x,y))] + \mathbb{E}_x[\log(1 - D(x,G(x))]$$

Therefor the total objective of the GAN network can be formulated as

$$G^* = \arg max_G min_D \; L_{cGAN}(G,D) + \lambda \, L_{L1}(G)$$

4. The Network Architecture

4.1 Generator:

The UNet network architecture [3] is used as the generator in this work. The Unet architecture consists of two parts, the encoder and the decoder network. The encoder network encodes the given 256x256 binary image to a 512 vector. While the decoder translates this vector to a SEM image. The encoder network consists of 8 convolution layers. The first layer has 64 filters, the second 128, the third 256 and the rest of the layers has got 512 filters. Each convolution operation is applied with a stride of 2 and hence the 2D space gets truncated from 256x256 at the input layer to 1x1 in the final layer. Fig. 3 illustrates the overall UNet architecture.

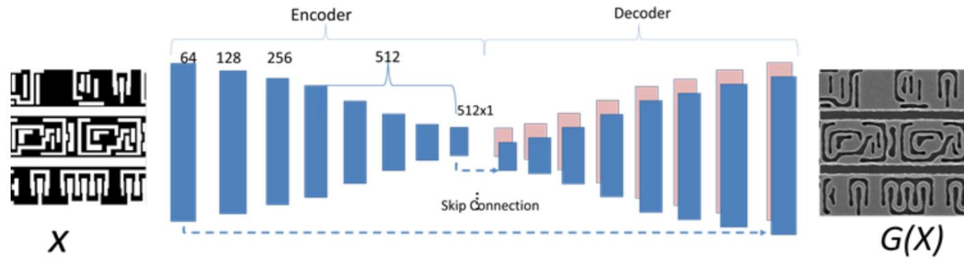

Figure 3. The UNet architecture used in the Generator

The decoder network mirrors the encoder network with 8 2D deconvolution operation. The input to the each of the deconvolution layer is the output of the previous layer and the skip connection from the corresponding encoder layer (as illustrated in the figure). The output of the decoder layer is the synthesized SEM image.

4.2 Discriminator:

The discriminator network is a simple DL network consisting of four convolution layers, followed by a sigmoid layer. There are 64 filters used in the first conv layer, 128 in the second, 256 in the third and 512 in the fourth. The discriminator takes in a 2-channel image as input. The optical design is fed as the first channel and the SEM image (either the original or generated) is fed as the second channel. Fig 4 illustrates the discriminator network.

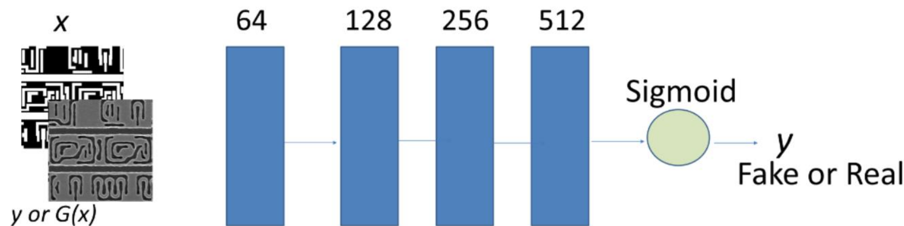

Figure 4. The Discriminator network

4.3 Training the network

The network is trained using the parameters suggested by [1]. Adam optimizer ($\beta1=0.5$, $\beta2=0.999$) is used for training both discriminator and the generator. The initial learning rate is set to 0.0002. The value of λ for L1 loss is set to 100. The network was trained for 200 epochs. Since the discriminator learns faster than the generator, for each iteration, the gradient descent optimizer is applied twice for the generator. During the training of the network data augmentation techniques of Image flipping and random cropping were applied.
For each epoch, a batch of Optical Design-SEM Image pairs are sampled from the dataset. The optical design is fed into the generator and 'fake' images are generated. The discriminator is run twice, once on the set of optical Design-Real SEM images pair and the second on the Optical Design-Fake SEM Images pair, in order to compute discriminator loss.

5. Results and Discussion

Fig 5. Shows the results of the generator after training the GAN for 200 epochs. The generated Image from the generator closely resembles the SEM image grabbed from the silicon wafer. Thus, generators from the conditional GAN can be used to simulate the chip fabrication from the optical design. One of the objective of this work is to find any defects any defects in the chip fabrication process, given the optical design. This could be achieved by comparing the grabbed image from SEM tool to the generated image from the generator network. In *Fig 5. Ex1.b*, the original SEM image shows a line breakage in the fabrication step, but the generated

image clearly depicts the separation between the lines as intended in the design. Hence generated images can be used as a reference for defect detection.

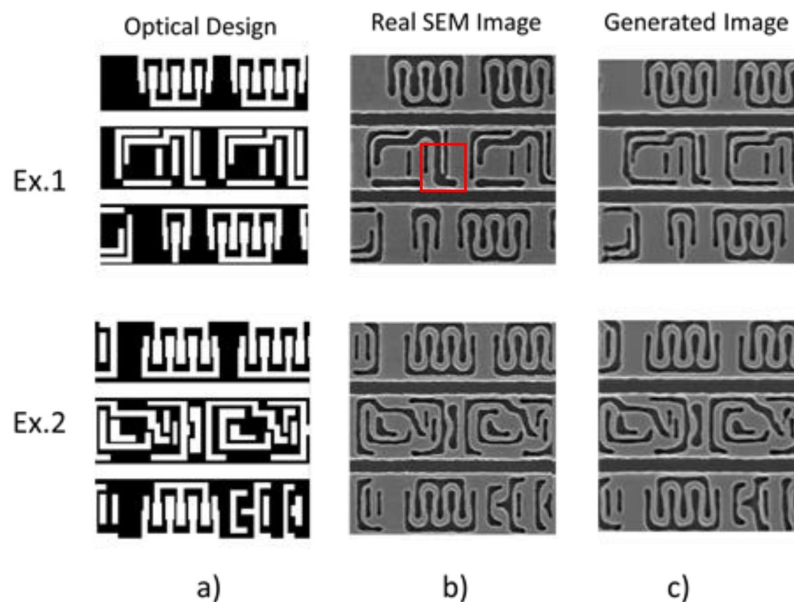

Figure 5. Results from the generator. a) The optical Design b) The original SEM Image grabbed from the Silicon wafers c) The Generated image from the UNet generator.

Though the generator can synthesize real-like SEM images from the optical design, the current model fails to create sharp images especially on the edges. The edges seem to be highly blurred especially when two lines are very close to each other in the design image. See appendix for more results.

6. Conclusion and Future Works

The Image-to-image translation using conditional GAN is very effective in synthesizing SEM images from its optical design and therefore helps in simulating the chip fabrication. This work proves that cGAN architectures could be vitally used for enhancing the yield in the semiconductor industry.

Due to the blur effect of the generator on edges and on smaller complex designs, the current model is not suitable for real field application. This needs to be studied and more robust architectures such as various other Resnet architectures needs to be explored for the case of generators.

7. Acknowledgement

I would like to thank KLA-Tencor for providing the infrastructure and data for this study. This work heavily realizes on the tensorflow implementation of pix2pix code published at https://github.com/yenchenlin/pix2pix-tensorflow. I would like to thank the contributors of the code.

8. References:

1. Isola, Phillip, et al. "Image-to-image translation with conditional adversarial networks." *arXiv preprint* (2017).
2. Goodfellow, Ian, et al. "Generative adversarial nets." Advances in neural information processing systems. 2014
3. Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." *International Conference on Medical image computing and computer-assisted intervention*. Springer, Cham, 2015.

Appendix