

Mask Generation of Inpainting Model for Moire Pattern based 3D Reconstruction

Tae-Jung Kim¹, Min-Ho Ha², Saba Arshad³ and Tae-Hyoung Park⁴

Abstract—In this paper, we aim to enhance the performance of 3D height reconstruction from 2D printed circuit board (PCB) moiré images by removing and reconstructing the noise caused by light reflection and shadows, which interfere with the reconstruction process. To achieve this, we conducted comparative experiments using various generative adversarial network (GAN) models applied to an inpainting model, discovering that mask images play a crucial role in this context. By applying white mask areas over the noise regions in the 2D PCB moiré images using the inpainting model, and reconstructing only those areas, we significantly improve the performance in reconstructing the heights of components within the PCB by removing and reconstructing only the specified noise regions. For setting the mask areas, we employ the fast unsupervised anomaly detection with generative adversarial networks (f-AnoGAN model) as an outlier detection model using GAN, aiming to identify anomalies between images with and without noise to set the mask areas accordingly. Through this approach, we accurately identify and reconstruct only the noise regions, proposing a method for reconstructing the heights of components within PCB using 2D PCB moiré images.

I. INTRODUCTION

Moiré pattern refer to structured light with a periodic pattern resembling a cosine function, characterized by shifts at 0, 90, 180, and 270 degrees. By projecting these patterns[1] onto a 2D PCB, we can acquire four moiré pattern projection images at different degrees. Utilizing these, it is possible to calculate a phase map and, through 2D phase unwrapping, obtain phase height values for 3D height reconstruction.

However, most components mounted on PCB substrates are made of metal, and the light reflected from these components can interfere with the calculation of accurate phase maps. Furthermore, shadows cast by components due to their height can obscure other components, also hindering accurate phase map calculation[2][3].

Previous research[4][5] addressing this issue have required adjustments to the brightness of the lighting for each component, the addition of hardware for measuring heights, or the addition of projected moiré grids. This paper proposes

a method for accurate height reconstruction in areas affected by light reflection and shadows, not through hardware improvements but by employing artificial intelligence (AI) algorithms. To apply AI algorithms, the designated areas of light reflection and shadows are identified using 2D PCB images shifted at 0, 90, 180, and 270 degrees. Areas in all four images with less than 30 pixels are classified as shadow regions, while any area in at least one image exceeding 230 pixels is classified as a light reflection area. However, this approach lacks a clear criterion for determining light reflection and shadow areas. Moreover, when the model inputs are phase height values rather than conventional 2D PCB images, the classification of light reflection and shadow areas changes, necessitating a flexible and clear criterion for setting these areas for accurate height reconstruction.

Therefore, among several other [6-10] anomaly detection models, this paper applies the f-AnoGAN[11] model, an anomaly detection model, to establish precise criteria for light reflection and shadow areas and proposes a method for accurate height reconstruction. Furthermore, it aims to compare this method with conventional techniques using an Inpainting model.

For experimentation, phase height value obtained by pre-processing the 2D PCB moiré image is used as Z . GT is the actual component height value of the PCB component. Z_{GT} is the value of GT fitted to the value of Z using a linear least squares formula. For mask images, the proposed mask image, Z_{mask} , is generated using the f-AnoGAN model with Z and Z_{GT} .

Among several other [12-17] inpainting models, the models for comparative experimentation include Generative image inpainting with contextual attention (CA inpainting) [18], Coherent semantic attention for image inpainting (CSA inpainting) [19], Generative image inpainting with segmentation confusion adversarial training and contrastive learning (SCAT inpainting) [20]. These models are trained and compared before and after applying Z_{mask} to determine the reconstruction rate and the extent of height reconstruction among PCB components. The experimental results revealed that the CSA inpainting model showed the highest reconstruction rate when using Z_{mask} , compared to previous methods.

Section II summarizes the existing 3D reconstruction models. Section III explains the moire pattern, preprocessing of images, and generation of moire images for training while Section IV presents the proposed mask generation method. The experimental setup and evaluation results are detailed in Section V. Finally, Section VI concludes this research.

¹T. J. Kim is with the Department of Intelligent Robotics Engineering, Chungbuk National University, Cheongju-si, Chungcheongbuk-do, Republic of Korea (28644). (e-mail: ktajung@chungbuk.ac.kr)

²M. H. Ha is with the Department of Control and Robot Engineering, Chungbuk National University, Cheongju-si, Chungcheongbuk-do, Republic of Korea (28644). (e-mail: minho6988@chungbuk.ac.kr)

³Saba Arshad is with the Industrial Artificial Intelligence Research Center, Chungbuk National University, Cheongju-si, Chungcheongbuk-do, Republic of Korea (28644). (e-mail: sabarshad@chungbuk.ac.kr)

⁴T. H. Park is with the Department of Intelligent Systems and Robotics, Chungbuk National University, Cheongju-si, Chungcheongbuk-do, Republic of Korea (28644). (e-mail: taehpark@cbnu.ac.kr)

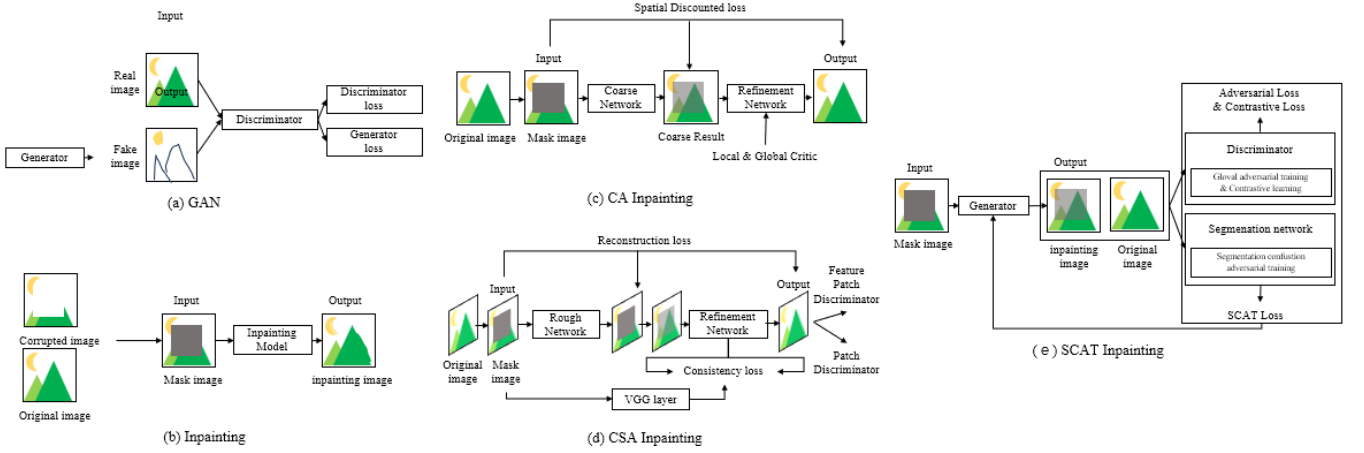


Fig. 1: Working pipeline; (a) GAN, (b) Inpainting, (c) CA Inpainting model architecture, (d) CSA Inpainting model architecture, (e) SCAT inpainting model architecture

II. 3D RECONSTRUCTION MODEL

With the prepared dataset, we conduct comparative experiments on the CA-Inpainting, CSA Inpainting, and SCAT Inpainting models. The GAN[21] model, as shown in Figure 1(a), learning occurs through the competition between two artificial neural networks, resulting in the generation of images that are increasingly realistic. The Generator (G) neural network is trained to create images that more closely resemble real images. Meanwhile, the Discriminator (D) neural network is trained to better distinguish between real images and those created by G. This can be analogously compared to the interaction between a police officer and a counterfeiter. The counterfeiter (G) aims to create currency so realistic that it deceives the police, while the police (D) must perfectly differentiate between genuine and counterfeit currency to catch the counterfeiter.

The Inpainting model[13-18], as shown in Figure 1(b), overlays random holes or noise on the input images as corrupted image, designating the damaged areas as mask regions. The goal is then to naturally reconstruct these masked areas by inputting the damaged image into the model.

The CA Inpainting model[19], while integrating the GAN framework, specifically designed for image reconstruction. This model operates in two stages of image reconstruction, as illustrated in Figure 1(c), and is categorized into a Coarse network and a refinement network. In [19], the model is trained to reconstruct original images from those damaged by rectangular or random mask areas, focusing on removing the obscured sections to recover the image fully.

This model enabling image reconstruction specifically for damaged areas rather than the entire image. However, to inform the model of the damaged areas, an additional mask image must be provided. Unlike previous research that supplied rectangular or random mask images, this paper utilizes Z_{mask} , a mask image generated through the earlier f-AnoGAN model.

For the model's input, a mask image overlaid on Z_{GT} ,

which lacks shadows and light reflections, is used. During training, the model learns the specified mask area by considering the surrounding regions. When Z , containing shadows and light reflections, given as input to the trained model, the overlaying mask leverages prior learning to remove shadows and light reflections within the mask area, resulting in the output.

The CSA inpainting model[20], as illustrated in Figure 1(d), follows the basic structure of the CA inpainting model but introduces the addition of a CSA layer, suggesting enhanced performance in image reconstruction. In [20], the model from the existing literature has been modified from VGG16 to VGG19 architecture, and the Fréchet Inception Distance (FID) has been utilized as a performance metric for GANs during the model saving process. The FID measures the similarity between two images, serving as an indicator of GAN model performance. Calculating the FID between Z_{GT} and the reconstructed Z , a lower score indicates higher similarity between the two images. The input and result of the model proceed in the same manner as the previous CA inpainting model.

The SCAT inpainting model[21], as shown in Figure 1(e), employs segmentation confusion adversarial training to perform labeling in the segmentation network (S) for areas where the mask region is 0 and areas where it is not. The Generator (G) is enhanced with the addition of a SCAT loss function, which aims to improve the reconstruction of the mask area to prevent S from labeling it. Conversely, in contrastive learning, the training progresses in a manner that makes the reconstructed image more distantly related to the damaged image and closer to the correct image.

III. MOIRE PATTERN

In this paper, we first acquire 2D PCB moiré images $I_n(x, y)$ (where $n = 1, 2, 3, 4$), which are displaced at four angles of 0, 90, 180, and 270 degrees with the presence of shadows and light reflections, from a moiré-based optical system as shown in figure 2. figure 2(a) corresponds to I_1 , figure 2(b) to I_2 , figure 2(c) to I_3 , and figure 2(d) to I_4 . The

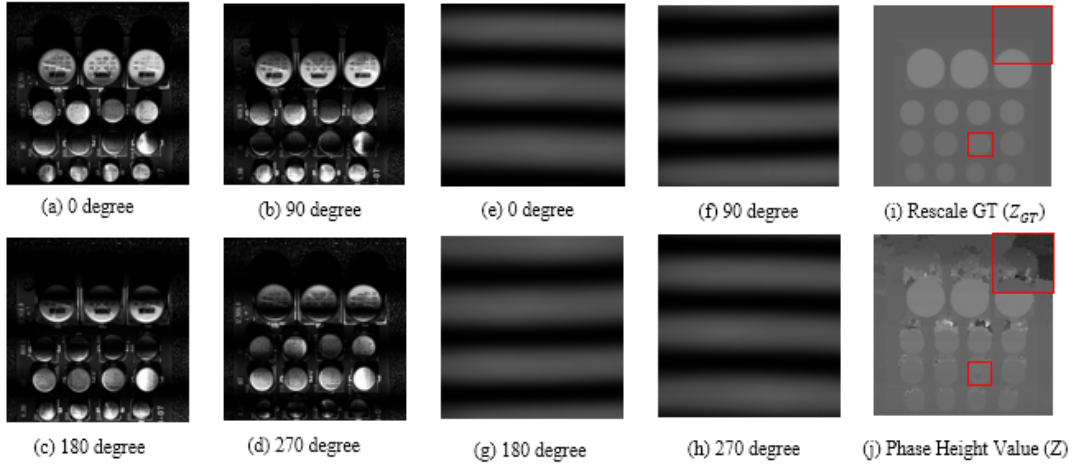


Fig. 2: Image with moiré pattern projected (a) - (d) projection to the ground, Image with moiré pattern projected (e) - (h) projection to the PCB, (i) Z_{GT} that needs to be reconstructed and (j) Z that is the correct answer

moiré images are transformed into a phase map $\varphi_p(x, y)$ as described by Equation (1). The transformed phase map undergoes a phase map unwrapping process as described by Equation (2). Subsequently, as shown in figure 2(j), the phase height value Z of the PCB is derived by Equation (3) through the difference between the 2D unwrapped phase map of the background image without the PCB $\varphi_{u-ref}(x, y)$ as shown in figure 2(e) (h), and the 2D unwrapped phase map of the PCB image, $\varphi_{u-p}(x, y)$. Here, x and y denote the x and y axes in two dimensions, respectively, and Z represents the total period of the moiré lattice pattern.

$$\varphi_p(x, y) = \arctan \frac{I_4(x, y) - I_2(x, y)}{I_1(x, y) - I_3(x, y)} \quad (1)$$

$$\varphi_{u-p}(x, y) = \varphi_p(x, y) + 2N\pi \quad (2)$$

$$Z(x, y) = (\varphi_{u-ref}(x, y) - \varphi_{u-p}(x, y)) \quad (3)$$

Additionally, for training purposes, a process is required to adjust the phase height values Z to match the actual heights of PCB components as contained in the Ground Truth (GT). To facilitate this, the linear least squares formula described in Equation (4) is applied, and the formula for the preprocessed Z_{GT} is given in Equation (5). Here, w represents the weight, and b signifies the intercept. GT_{1D} represents 2D GT as a 1D array, and Z_{1D} represents 2D Z in Equation (3) as a 1D array. Z_{GT} created through this process are depicted in figure 2(i).

$$\begin{pmatrix} w \\ b \end{pmatrix} = (GT_{1D}^T GT_{1D} + \lambda I)^{-1} GT_{1D}^T Z_{1D} \quad (4)$$

$$Z_{GT} = GT * w + b \quad (5)$$

Upon examining the differences between the Z image and the Z_{GT} image represented in figure 2(j),5(i), it is observable

that areas damaged in height reconstruction are due to shadows cast behind the components and light reflections within the component areas, which are considered noise. These areas, affected by noise, lead to issues in reconstructing the PCB component heights in Z . It is intended to apply a mask over these areas to facilitate the reconstruction of PCB component heights.

IV. MASK GENERATE METHOD

A. Motivation

The original mask image is created as shown in Figure 4(a). However, in this case, when the 2D PCB image is converted into a phase map and then undergoes 2D phase unwrapping, the values in the image change. This change leads to a problem where the areas damaged by shadows and light reflections differ from those initially identified. As a result, using the original mask may inadvertently reconstruct areas of the components that should not be reconstructed, leading to improper reconstruction.

B. Proposed method

Therefore, in cases where only Z and Z_{GT} data are available and there is no accurate criterion (ground truth) for the mask, an anomaly detection model utilizing GANs can be employed. In this paper, to define the mask image among several anomaly detection model [10-13], we utilize the f-AnoGAN model along with the Z_{GT} and Z images transformed as shown in figure 2(i),5(j).

The training process of the f-AnoGAN model is as shown in Figure 3. First, the GAN model is trained by generating realistic images to generate Z_{GT} , which is the correct image without outliers. Then, the encoder is trained, and the image is converted into a latent vector and the latent vector into an image to quickly obtain the latent vector, which is a feature of the image. When the training is completed and Z , an image with outliers, is input to the model, the image is converted to a latent vector through the encoder, and the converted latent vector is input to the previously trained GAN model

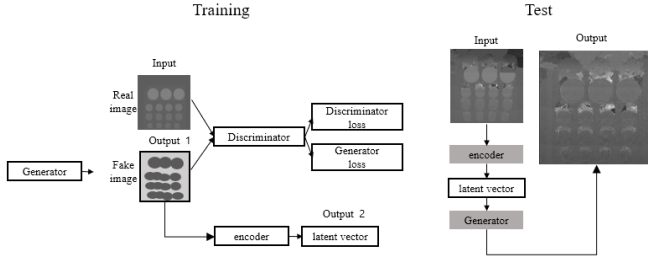


Fig. 3: f-AnoGAN model architecture

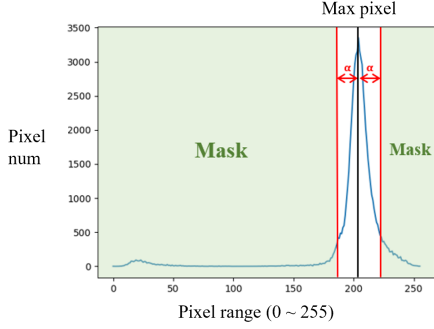


Fig. 4: 2D abnormal image histogram

to generate an image. We get the anomaly image A , which is the image of the absolute value of the difference between the generated image and the correct answer Z_{GT} .

Then examine the pixel distribution of A . By representing the pixel values of the 2D image ranging from 0 to 255 in a histogram as A_{hist} , we can observe, as depicted in Figure 4, that certain pixel values are concentrated. Based on this observation, Equation (6) sets a constant value α with respect to the pixel *Threshold* where the most pixel values are concentrated.

$$A(x, y) = \begin{cases} 255, & A(x, y) > Threshold + \alpha \\ 255, & A(x, y) < Threshold - \alpha \end{cases} \quad (6)$$

Subsequently, pixels that meet the condition of the equation are designated as mask areas and assigned a pixel value of 255 else 0 as shown in algorithm 1. Here x and y represent the x and y axes in two dimensions, respectively, and A represents the anomaly image obtained from the output of the f-AnoGAN model. The mask image regions are divided into 30 categories ($\alpha = 0, 1, \dots, 29, 30$) because different values of α result in different mask image regions. The reason why the α value is limited to 30 is because any value above that creates unnecessary mask regions that are visually too large. To set the α value, a full search from 0 to 30 is performed on Z and Z_{GT} to extract only the mask regions according to the α value and compare the pixel differences. As the α value gets larger, the pixel differences get larger, and at some point the α value starts to get smaller. This means that the mask region has become unnecessarily large and no longer needs to be reconstructed, and this point corresponds to the first maximum point in the 1D histogram of pixel differences

Algorithm 1: Generate Abnormal Mask Image

Input: Abnormal image A from f-anogan model result, Interval α
Output: Abnormal mask image (Z_{mask})

```

1  $A_{hist} [256]=0$ 
2 for each  $x \leftarrow A_{width}$  do
3   for each  $y \leftarrow A_{height}$  do
4      $A_{hist} [A(x,y)]++$ 
5   end
6  $A_{max} = \max(A_{hist})$ 
7 for each  $x \leftarrow 256$  do
8   if  $A_{hist}[x] == A_{max}$  then
9      $Threshold = x$ 
10  end
11 for each  $x \leftarrow A_{width}$  do
12   for each  $y \leftarrow A_{height}$  do
13     if  $A(x,y) > Threshold + \alpha$  and  $A(x,y) < Threshold - \alpha$  then
14        $Z_{mask}(x,y) = 255$ 
15     end
16   else
17      $Z_{mask}(x,y) = 0$ 
18   end
19 end

```

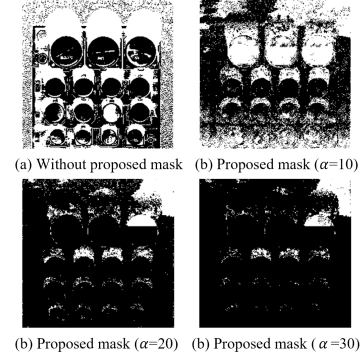


Fig. 5: 4 types of mask images (a) - (d)

obtained by full search. The α value at the first maximum point is the optimal α . The optimal α value obtained by this method is different every time, even for the same type of PCB, and will be called α_{opt} in this paper.

The mask images created by these methods are shown in Figure 5(b)-(c). (where $\alpha = 10, 20, 30$) The mask regions thus defined are used to overlay on Z_{GT} for training.

V. EXPERIMENTS

A. Experimental setup

All, the experiments are conducted using two NVIDIA A100 80GB PCIe GPUs and one GTX GeForce 3080 GPU.





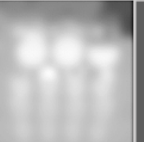


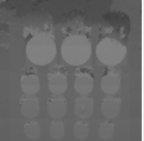








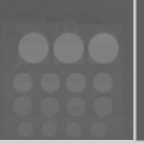

Model /Mask	Before mask	Proposed mask ($\alpha = 30$)	Proposed mask ($\alpha = 20$)	Proposed mask ($\alpha = 10$)	Proposed mask ($\alpha = \alpha_{opt}$)	Z_{GT}
CA inpainting						
SCAT inpainting						
CSA inpainting						

Fig. 6: Inpainting model result image (5 types of mask)

The data used in the experiments are obtained by manufacturing PCB substrates with reflective materials and directly photographing them using a moiré pattern generator and camera assembly. This process enables the acquisition of 2D PCB moiré images with shadows and light reflections, along with the ground truth (GT) information on the actual component heights. Due to the mismatch between the positions of components in the moiré images and the GT , a matching process is implemented using C++. Consequently, the experiments are performed with a dataset of 420 matched images, comprising 10 images from each of the 42 different types of PCBs.

In the experiments, preprocessing of the given dataset is conducted to generate Z and Z_{GT} , with the number of training epochs fixed at 500, and the size of the images standardized to 256×256 . Comparative experiments are carried out on a total of three models, and the applied mask images are compared in five scenarios: existing mask images and proposed mask images ($\alpha = \alpha_{opt}$, 10, 20, 30). The performance comparison of the model training results is presented through the reconstruction rate.

B. Results

Initially, training is conducted for the CA inpainting model. Mask image is required, and the training input sequentially included Z_{GT} along with the mask image Z_{mask} . The training proceeded with Z_{mask} overlaid on Z_{GT} . The model is configured to save in the direction of minimizing the loss function during the training process. The GPU used is the NVIDIA A100 80GB PCIe, with the model size around 16152MiB, and an average of approximately 2 seconds is required for each training session. The reconstructed Z from this training appears as in the second row of Figure 6.

Subsequent training is conducted on the CSA inpainting model, following a similar procedure to the CA inpainting model, with the model being saved in the direction where the Fréchet Inception Distance (FID), a performance metric

for GANs, is minimized. The GTX GeForce 3080 GPU is utilized, with the model size being approximately 3359MiB and an average training time of about 7 seconds. The reconstructed Z results are displayed in the third row of Figure 6.

Lastly, training is conducted on the SCAT inpainting model. The main point of interest is the performance comparison with the preceding CSA inpainting model, following a similar training methodology. The NVIDIA A100 80GB PCIe GPU is utilized, with the model size being approximately 25046MiB and an average training time of about 90 seconds. The reconstructed Z results are displayed in the fourth row of Figure 6.

After the experiment, to compare the performance of the reconstructed Z through the model, it is necessary to define the reconstruction rate. The criterion for reconstruction is focused on how naturally the model reconstructed the mask area and, as a result, how well it reconstructed the heights of components in the PCB image. To this end, Equation (7) specifies the reconstruction area by combining the area region, which extracts only the component area excluding the PCB board from the Z_{GT} , with the mask area. The pixel difference between $Area$ and the reconstructed Z within this reconstruction area is compared, and the error rate is calculated as shown in Equation (8) to obtain the Error in percentage. The final reconstruction rate is compared by subtracting Error/(sum of Error counts) from 100percent, and the performance of the model is evaluated by the average reconstruction rate of a total of 420 images. Z refers to the reconstructed Z from which shadows and light reflections have been removed and reconstructed by the previous models, and $Area$ refers to the area used to compare Z_{GT} and Z for calculating the reconstruction rate. Threshold refers to the boundary pixel value for extracting only the component area excluding the PCB board from Z_{GT} . Error represents the error rate between Z_{GT} and Z where $Area \neq 0$.

TABLE I: reconstruction rate of height [%]

Model / Mask	Before mask	Proposed mask ($\alpha = 30$)	Proposed mask ($\alpha = 20$)	Proposed mask ($\alpha = 10$)	Proposed mask ($\alpha = \alpha_{opt}$)
CA inpainting	65.05	65.46	62.59	63.20	62.55
CSA inpainting	95.79	94.98	95.57	96.72	98.02
SCAT inpainting	96.64	95.93	96.29	97.72	97.93

$$Area = \begin{pmatrix} 255, & Z_{GT} > Threshold \ \& \ Z_{mask} \neq 0 \\ 0, & ; otherwise \end{pmatrix} \quad (7)$$

$$Error = \begin{pmatrix} 1 - Z/Z_{GT} * 100, & Area \neq 0 \\ 0, & otherwise \end{pmatrix} \quad (8)$$

We conducted a performance comparison using the reconstruction rate formula provided in Equation (7)(8) for Z reconstructed through the three models discussed. For the CA Inpainting, CSA Inpainting, and SCAT Inpainting models, the reconstruction rate is calculated for five scenarios: using the existing mask and proposed masks ($\alpha = 30, 20, 10, \alpha_{opt}$). The results are presented in Table 1, showing, contrary to expectations, that the highest reconstruction rate is achieved using the proposed mask ($\alpha = \alpha_{opt}$) with the CSA Inpainting model. This represents the average reconstruction rate of total heights across all data, which consists of 10 images each for 42 types of PCBs, with the number of training epochs fixed at 500.

VI. CONCLUSION

The results demonstrate the impact of different masks on the reconstruction outcomes of PCB component heights in 2D PCB images. It is found that neither the excessively broad areas of the existing masks nor the overly narrow areas of the proposed masks ($\alpha = 30$) performed optimally. Instead, the proposed mask utilizing f-AnoGAN with a moderate α value ($\alpha = \alpha_{opt}$) yielded the best performance. This finding suggests that future research should focus on the 3D GAN model that generates 3D images from 2D images.

VII. ACKNOWLEDGMENT

This work is supported by Innovative Human Resource Development for Local Intellectualization program through the Institute of Information & Communications Technology Planning & Evaluation(IITP) grant funded by the Korea government(MSIT)(IITP-2024-2020-0-01462)

REFERENCES

- [1] J. E. Greivenkamp, J. H. Bruning, D. Malacara, "Phase shifting in terferometry", In Optical Shop Testing, pp.547-666, 1992.
- [2] S. Zhang, "High-speed 3D shape measurement with structured light methods:a review," In Optics and Lasers in Engineering, pp. 119-131, July. 2018.
- [3] Z. Wang, Z. Zhang, N. Gao, Y. Xiao, F. Gao, and X.Jiang, "Single-shot 3D shape measurement of discontinuous objects based on a coaxial fringe projection system," In Applied Optics, pp. A169-A178, 2019.
- [4] J.-S. Lee, T.-H. Park, "Defect classification of components for SMT inspection machines," In Journal of Institute of Control, Robotics and Systems (in Korean), pp. 982-987, 2019.
- [5] J.-D. Song, Y.-G. Kim, T.-H. Park, "Defect classification method of PCB solder joint by color features and region segmentation," In Journal of Institute of Control, Robotics and Systems (in Korean), pp. 1086-1091, 2017.
- [6] T. Schlegl, P. Seeböck, S. M. Waldstein, U. Schmidt-Erfurth, G. Langs, "Unsupervised anomaly detection with generative adversarial networks to guide marker discovery," In Int. Conf. on information processing in medical imaging, pp. 146-157, May 2017
- [7] K. Batzner, L. Heckler, and R. König, "Efficientad: Accurate visual anomaly detection at millisecond-level latencies," In Proc. of the IEEE/CVF Winter Conference on Applications of Computer Vision, pp. 128-138, 2024
- [8] A. Mousakhan, T. Brox, and J. Tayyub, "Anomaly Detection with Conditioned Denoising Diffusion Models," In arXiv preprint arXiv, 2305.15956, 2023
- [9] A. Berg, J. Ahlberg, and M. Felsberg, "Unsupervised learning of anomaly detection from contaminated image data using simultaneous encoder training," In arXiv preprint arXiv, 1905.11034, 2019
- [10] P. Liznerski, L. Ruff, R. A. Vandermeulen, B. J. Franks, K. R. Müller, M. Kloft, "Exposing outlier exposure: What can be learned from few, one, and zero outlier images," In arXiv preprint arXiv, pp. 2205-11474, 2022
- [11] T. Schlegl, P. Seeböck, S. M. Waldstein, G. Langs, and U. Schmidt-Erfurth, "f-AnoGAN: Fast unsupervised anomaly detection with generative adversarial networks," In Medical image analysis, pp.54, 30-44, 2019
- [12] H. Zheng, Z. Lin, J. Lu, S. Cohen, E. Shechtman, C. Barnes, and J. Luo, "Image inpainting with cascaded modulation GAN and object-aware training," In European Conf. on Computer Vision, pp. 277-296, October 2022
- [13] R. Suvorov, E. Logacheva, A. Mashikhin, A. Remizova, A. Ashukha, A. Silvestrov, and V. Lempitsky, "Resolution-robust large mask inpainting with fourier convolutions," In Proc. of the IEEE/CVF winter conference on applications of computer vision, pp. 2149-2159, 2022
- [14] W. Li, Z. Lin, K. Zhou, L. Qi, Y. Wang, and J. Jia, "Mask-aware transformer for large hole image inpainting," In Proc. of the IEEE/CVF conference on computer vision and pattern recognition, pp. 10758-10768, 2022
- [15] D. P. Kingma, and M. Welling, "Auto-encoding variational bayes," arXiv preprint arXiv, 1312.6114, 2013
- [16] S. Zhao, J. Cui, Y. Sheng, Y. Dong, X. Liang, E. I. Chang, Y. Xu, "Large scale image completion via co-modulated generative adversarial networks," In arXiv preprint arXiv, pp. 2103-10428, 2021
- [17] P. Jeevan, D. S. Kumar, A. Sethi, "WavePaint: Resource-efficient Token-mixer for Self-supervised Inpainting," In arXiv preprint arXiv, pp. 2307-00407, 2023
- [18] J. Yu, Z. Lin, J. Yang, X. Shen, X. Lu, and T. S. Huang, "Generative image inpainting with contextual attention," In Proc. of the IEEE conference on computer vision and pattern recognition, pp. 5505-5514, 2018.
- [19] H. Liu, B. Jiang, Y. Xiao, and C. Yang, "Coherent semantic attention for image inpainting," In Proc. of the IEEE/CVF International Conference on Computer Vision, pp. 4170-4179, 2019
- [20] Z. Zuo, L. Zhao, A. Li, Z. Wang, Z. Zhang, J. Chen, and D. Lu, "Generative Image Inpainting with Segmentation Confusion Adversarial Training and Contrastive Learning," In arXiv preprint arXiv, pp. 2303-13133, 2023
- [21] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, and Y. Bengio, "Generative adversarial networks," In IEEE signal processing magazine, pp. 35(1), 53-65, 2018