

# Self-Supervised Outpainting for Display Panel Defect Image Augmentation

Zhihong Pan<sup>1</sup>, Rahul Shenoy<sup>1</sup>, Kaushik Balakrishnan<sup>1</sup>, Yongmoon Jeon<sup>2</sup> and Jaewon Kim<sup>2</sup>

<sup>1</sup>Samsung Display America Lab, San Jose, CA, USA

<sup>2</sup>Samsung Display Co., Yongin-si, Gyeonggi-do, South Korea

## Abstract

Deep learning (DL) has gained increasing popularity over conventional computer vision methods in smart manufacturing for its potential in high performance. However, training good DL models needs a large set of training data with sufficient variations. In the case of defect detection and classification for display panels, it is desirable for the defect (NG) images to have variations in defect shape, size and location. In practice, NG images are often collected without these considerations, e.g., defects are often located at the center. To increase data variation of the dataset, a novel image outpainting method is proposed to support expanding an input NG image to larger size so that data augmentations like random scales are applicable. A series of special masks are designed to support self-supervised training of the outpainting model using NG images exclusively. More importantly, the same model can be applied to remove defect from an input NG image and output a matching normal (OK) image. These augmented OK-NG pairs can benefit other DL methods, e.g., a defect image generation AI model can be trained as an OK-conditional model to support generating defect on an input OK image while preserving the background. It is shown with experiments that a defect generation model, which is trained with augmented images using outpainting, is able to generate samples with higher defect quality and better background fidelity comparing to previous methods.

## Keywords

Deep Learning, Computer Vision, Defect Classification, Image Outpainting, Diffusion Model, Image Generation

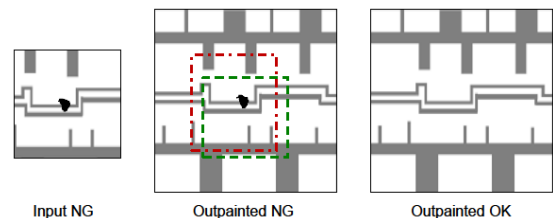
## 1. Introduction

Recently, various Deep Learning (DL) models are increasingly used in smart manufacturing. Among them, Computer vision (CV) models like object detection and classification [1] are often deployed for optical inspections where visual images of products are taken to identify defects for repairing. For example, defects in the display panel manufacturing process need to be identified with high accuracy for efficiency and robustness. Currently, a lot of the defect identification, classification and repair are conducted by human personnel, which is time consuming, expensive and prone to human error. To make a more robust process, an AI-based defect classification and repair system is warranted. However, to build an AI-based classifier, one needs data balance between the number of defect-free (OK) and defective (NG) sample images used to train the AI model. This is challenging in practice since the number of defective samples in manufacturing are typically a very small subset of the total (e.g., 1-2%).

This data imbalance issue could be mitigated by using an AI-based generative model to synthesize NG images. However, this generative model requires a large dataset too. Given the limited number and diversity of NG images for the product which a classifier is needed in production (the target product), a transfer

learning method can be applied to learn from other products where NG images are abundant (source products). To solve this problem, a Defect Transfer GAN (DTGAN) [2] was first proposed to train the GAN model with mask-conditional losses, learning from abundance of defect images from an existing Product B to generating defect on a normal Product A image.

But GAN has its limitation in unstable training and mode collapse which may result in lack of image quality and diversity. In recent works [3, 4], a new type of defect transfer generative method is proposed, which trains a diffusion model [5] with OK and NG images from both source and target products where the number of NG images from target products are very limited. To synthesize a target NG image, a target OK image is injected with partial noise and the diffusion model is applied with a NG-condition to convert it to target NG images. One challenge for this solution is to support random scale and crop augmentations. For random scaling, it is to bridge the scale difference between source and target products to make defect transfer more effective. For random cropping, it is to vary defect locations so that the diffusion model learns to generate defect in random areas so that it is not limited to generate defect on the OK images which is more prone to have defect occurred in the center. However, in practice, the NG images for source products could be collected some time back and they are often limited in size and resolution which makes data augmentation like random scaling and crop impossible or limited in extent. The other challenge is to maintain background fidelity as the injected noise may make the diffusion model confuse with another product. To solve this challenge, the diffusion model can also be trained as conditional to OK image in addition to class label, thus the injected noise is of no concern as the diffusion model refers to the OK image condition to maintain background fidelity. However, this is only applicable when paired OK-NG images are available, which is missing in most cases.

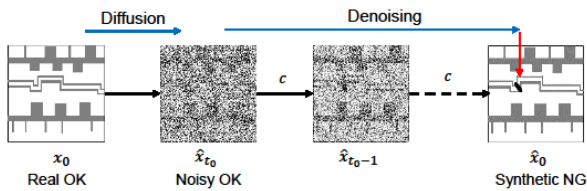


**Figure 1.** Dual purpose for NG image augmentations using self-supervised outpainting.

In this work, an innovative image outpainting method is proposed to solve both challenges above, lack of large NG and paired OK-NG, by supporting two types of data augmentation as shown in Fig. 1. In the first case, a larger size NG image is generated so random scaling and cropping augmentation is possible. For the latter, paired OK and NG images are generated so diffusion model conditional to OK images can be trained. While image outpainting [6] has been widely used in generative models like GAN [7] and diffusion models [8, 9], they often rely

on paired images, an input where the outside area is masked and an unmasked pair, to learn outpainting. In other words, they belong to the application of supervised learning. In the case of outpaint a NG image, it means large NG images where the defects reside in the inside area are needed. In practice, large NG images are often not available for outdated source products. To address this, this newly proposed outpainting model is designed to be trained using existing fixed-size NG images only, hence self-supervised learning.

First, a series of novel masks are designed to segment all images in three areas: the input area is used as input for outpaint model while the output area is repainted by the outpainting model, both of which are areas outside the defect area in general; and the excluded area where the defect is located are not included in output model training and inferencing. Once trained, these designed masks are applied sequentially to expand the input NG image to different surrounding areas step by step until a full outpainted image is generated. This large NG image can be used to support random crop and scaling for improve training performance of other DL models. Additionally, it also supports repainting the defect area of the initial NG image with OK contents so a matching OK-NG pair is generated. This self-supervised outpainting method was applied to augment a display panel NG image dataset. Enabling training a defect generation model using paired OK-NG images with random scale and crop, the outpainting augmentation is shown to improve defect generation quality significantly.



**Figure 2.** Illustration of defect transfer diffusion model sampling, where  $x_0$  is an OK image,  $t_0 < T$  is used to get noisy image  $\hat{x}_{t_0}$  and  $c = NG$  is used as the condition to generate defect in the final output. These setting could vary in other models introduced in this work.

## 2. Proposed Method

**Diffusion model background:** The Denoising Diffusion Probabilistic Model (DDPM) as proposed in [8] aims to model the bidirectional transformation between the distribution of real image  $x$  and a tractable random distribution like Gaussian as a bidirectional process. For the forward diffusion process, an image  $x_0$  is transformed progressively as  $x_1, x_2, \dots, x_T$  where  $x_T$  becomes a random noise. On the other hand, a diffusion model is designed to convert any random noise  $x_T$  to an image  $x_0$ . For this study, the latest latent diffusion model (LDM) [5] is adopted for its superior capability to generate high-resolution (HR) images. Following LDM, the diffusion model, denoted as  $\epsilon_\theta$ , is trained to estimate the noise of sampled  $x_t$ . And the sampling of next step  $x_{t-1}$ , using the DDIM [10] method, becomes

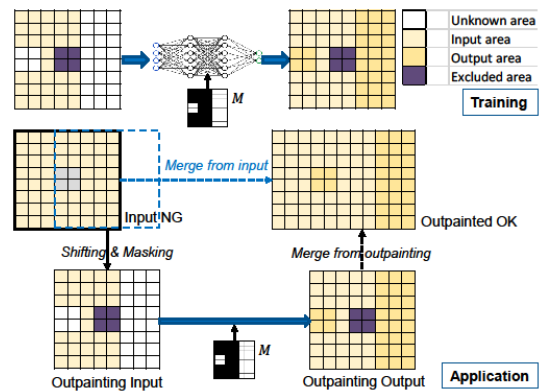
$$x_0^t = (x_t - \sqrt{1 - \bar{\alpha}_t} \epsilon_\theta(x_t, t, c)) / \sqrt{\bar{\alpha}_t} \quad (1)$$

$$x_{t-1} = \sqrt{\bar{\alpha}_{t-1}} x_0^t + \sqrt{1 - \bar{\alpha}_{t-1} - \sigma_t^2} \epsilon_\theta(x_t, t, c) + \sigma_t \epsilon_t \quad (2)$$

Here  $x_0^t$  is the predicted noise-free image,  $\epsilon_t$  is a random noise used for sampling for the next time step  $t - 1$ . For convenience, this sampling step is denoted simply as  $x_{t-1} = f(x_t, t, c)$ .

**Defect transfer diffusion model:** For this study, the primary

application is defect generation diffusion model and the proposed self-supervised outpainting model is used to augment training data for a better defect generation model. This defect generation model is trained from OK and NG images from both source and target products where target NG images are very limited. To help learning defect transfer from source to target, the product type is transparent so the condition is binary: NG and OK. As shown in Fig. 2, to generate a NG image of a target product, the diffusion model needs to know the product type. Instead of sampling from random noise, a noisy image of the given target product is used instead, denoted as  $\hat{x}_{t_0}$ . The setting of  $t_0$  is tricky in this case as it needs to be larger, closer to  $T$ , to make defect generation effective, but smaller to keep product fidelity, i.e., maintain the non-defective area of input image. As discussed in [3], a mask could be introduced to optimize both defect quality and product fidelity, but it adds significant computational overhead. Alternatively, the diffusion model can be implemented with a hybrid condition of  $c_1 = (OK, NG)$  and  $c_2 = x_{OK}$ . This is only applicable when an OK-NG pair dataset is available. For that, the self-supervised outpainting model is proposed to augment existing NG images as OK-NG pairs.



**Figure 3.** Illustration for self-supervised outpainting model training and application. For simplicity, only one mask which supports outpainting to the right is illustrated.

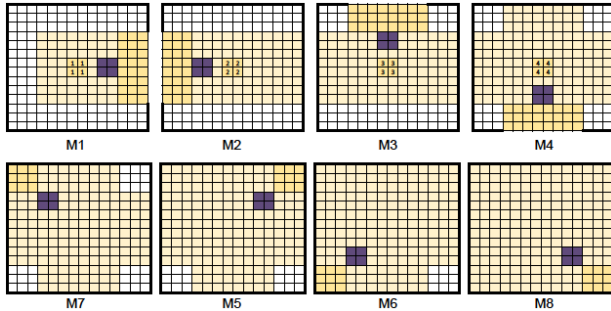
**Self-supervised outpainting:** As shown in Fig.3 (top), during training of the outpainting model, the white area on the right side as well as a small central left area are masked as unknown areas. The remained are not masked but only the input area, marked in light yellow, are used as inputs. Pixels in the excluded area are assigned with a constant value, like average of the input area. The outpainting model is optimized to learn filling the unknown areas using contexts from the input area, denoted as the output area in yellow. For such a mask as shown here, the outpainting model is learning to generate OK content exclusively as the defect is located inside the excluded area, neither used as input nor as output. And as both input and output images are of the same size as NG images used for training, the whole process is self-supervised. The conditional diffusion model illustrated earlier in Fig. 2 is also used for outpainting model training, except it is conditioned to both  $c_1 = M$  and  $c_2 = x_m$  as

$$x_{t-1} = \sqrt{\bar{\alpha}_{t-1}} x_0^t + \sqrt{1 - \bar{\alpha}_{t-1} - \sigma_t^2} \epsilon_\theta(x_t, t, c_1, c_2) + \sigma_t \epsilon_t \quad (3)$$

where  $M$  is the mask and  $x_m$  is the masked input image.

At application time, as shown in Fig. 3 (bottom), the outpainting input is shifted to the right to align the defect area (in gray) with the small masked unknown area and the exclude area is assigned to a constant. After passing through the outpainting process using

diffusion sampling, the output area is filled with OK content. By merging from two source, one from the outpainting output with excluded area restored from input NG, and the other from unused left side of NG, an outpainted larger OK image is generated. Alternatively, the small output area from outpainting could be restored from the defect area of input NG, an outpainted NG can also be generated to form an outpainted OK-NG pair.



**Figure 4.** Progressive outpainting using a novel set of 8 masks, which all supports self-supervised learning as shown in Fig.3.

**Progressive outpainting:** The mask shown in Fig. 3 is only for outpainting to the right. To expanding to all directions, a set of 8 masks are designed to apply outpainting progressively as shown in Fig. 4. The mask used in Fig. 1 is M1, while there 3 other similar masks, M2 to M4, are used to outpaint toward left, top and down respectively. After applying them sequentially, in addition to outpainting areas outside the original input, the defect areas are repainted four times. This area could be the fused version of these four duplicates to get outpainted OK, or restore for input NG to get outpainted NG. For M5 to M8, each has one output area toward one corner and applying them sequentially on the outpainted OK from M1 to M4, a full large OK with outpainted contents in all four directions is generated. Similar, restoring the small central area with defect in input NG results in a large NG which matches the large NG as a pair.

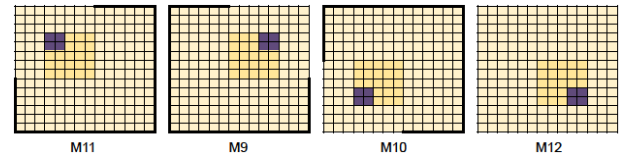
**OK-conditional defect generation model:** After applying progressive outpainting to the original NG images, the output set of OK-NG pairs enable training a dual conditional defect generation model as formulated in Equation (3), where the conditions are now  $c_1 = (OK, NG)$  and  $c_2 = x_{ok}$ . Here  $x_{ok}$  is the paired OK image when the input is NG, otherwise the OK image itself. With the OK-conditional model, the empirical value  $t_0$  as shown in Fig. 2 is not needed since the sampling process can start from  $T$  to use full power of diffusion model. Besides, the carefully selected mask  $M$  as in Equation (4) is also not necessary as the OK-conditional model learns to keep the background from the condition starting from  $T$ .

### 3. Experiments and Results

For the self-supervised outpaint model, there is no ground-truth reference to assess the results directly. Instead, we have conducted experiments using the augmented NG dataset after applying outpainting comparing to the original dataset. One is for the defect transfer diffusion model sampling, comparing the samples from the class-conditional (OK and NG) model and those from the OK-conditional model which is made possible by the outpainting application. The other is assessed in classification, comparing the robustness of classifiers trained without and with random scale and crop augmentation.

**Dataset:** For experiments, we use a proprietary dataset collected

from Samsung Display’s manufacturing line. It includes both normal and defect images from multiple NG and OK products. The original dataset is in  $256 \times 256$  resolution. For self-supervised outpainting, the excluded area is fixed as  $64 \times 64$  in the center, M1 to M4 have 2 output areas of  $96 \times 256$  and  $64 \times 64$ , and M5 to M8 have output areas of  $96 \times 96$ . To train the outpainting model, only NG images with small defects are used. To apply progressive outpainting to NG images with larger defects, an additional set of 4 masks are used to in both training and application. After sequential application of M9 to M12, as shown in Fig. 5, almost the full area of the original NG is repainted with OK content so the OK-NG pairs could be augmented from most of original NG images.



**Figure 5.** Additional progressive outpainting steps which repaint most of original input area to cover large defects at NG-to-OK conversion.

**Model training and sampling:** For LDM model training, the Stable Diffusion 1.5 code base [11] is adopted and it is trained on multiple A100 GPUs. DPM-Solver++ [12] is used for diffusion model sampling with a fixed sampling step of 50.

**Evaluation metrics:** There is no direct way to evaluate outpainted images quantitatively as there is no ground-truth available. Instead, we used the outpainted images to train an OK-conditional defect generation model and assess the quality of synthetic defect images. To evaluate defect fidelity and diversity, the FID [13] score is calculated between a set of synthetic images and a ground-truth real defect image set. The popular LPIPS [14] metric is also chosen to assess the background fidelity between the synthetic defect image and its input normal image. Although the generated defect causes degradation in LPIPS, the average LPIPS value is a good indicator of minimum change in the background area.

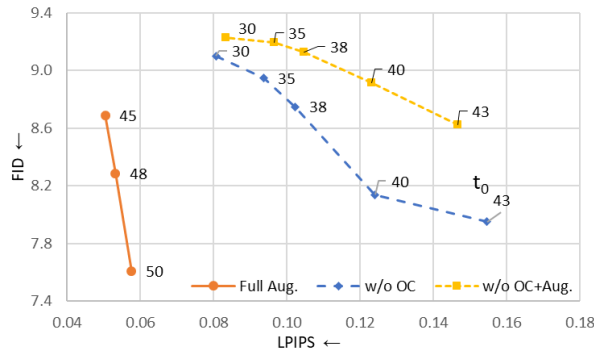
**Main Experiments:** For the defect generation model trained from outpainted images using our proposed method, it has two advantages comparing to the one trained from the original dataset: OK-NG pair augmentations which enables OK-conditional model, and random scale and crop augmentation which helps increasing training data variations. To fully demonstrate the benefit of our outpainting methods, an OK-conditional defect generation model was trained with random scale and crop augmentation. For comparison, two models of the same architecture were trained, one without OK-condition first, and one without data augmentation on top of that. For each model, 1,250 fake NG images were sampled for each combination of 4 defect types and 7 product types, 35,000 in total. For the model with full augmentation, including OK-condition, they were sampled from OK images starting from random noise as  $t_0 = 50$ . For the other two, they were sampled from the same set of OK images using  $t_0 = 40$  to keep background fidelity. As shown in Table 1, the OK-conditional samples with scale augmentation are superior in both background fidelity (lower LPIPS) and sample quality in general (lower FID). Without OK-condition, there is a significant drop in background fidelity and degradation in FID is obvious too. After further removing scale and crop augmentation, LPIPS is about the same but there is obvious loss in FID metric. This degradation is mostly caused by lost capabilities due to lack of

data augmentations. Without random crop, the training data has a bias in defect location (always centered) and the trained model may fail to generate defect when the random input OK image has a central area tends to be defect-free.

**Table 1.** Comparison of sample quality between diffusion models using dataset of different augmentations

	Full Aug.	No OK-Cond.	+ No Aug.
LPIPS ↓	0.058	0.124	0.123
FID ↓	7.61	8.14	8.91

**Ablation Study:** For results shown in Table 1, the optimal  $t_0$  is chosen for each method considering the trade-off between FID and LPIPS. Results from the ablation study to find the optimal values are shown in Fig. 6. For the one taking full advantages of data augmentation enabled by our proposed outpainting method, there is little impact on LPIPS with decreasing  $t_0$  while the degradation in FID metric is significant. It is optimal to start from  $t_0 = 50$  to get best defect quality. Not that lower LPIPS is not always preferred as one can get better LPIPS by generating smaller defects or even no defect at all. For the two methods without OK-condition (OC), there is an obvious trade-off between FID and LPIPS. For smaller  $t_0$ , it preserves the background better (smaller LPIPS) but it comes at the cost of defect quality (larger FID).



**Figure 6.** Comparison of sample qualities for different models using varying  $t_0$ .

#### 4. Impact

Generative AI has gained popularity in a broad range of applications, including display manufacturing applications. Built on top our previous works of defect transfer diffusion models [3, 4], we have developed a novel outpaint method to improve performance of these defect generations models, by enabling data augmentations which are otherwise limited in existing dataset. For these reasons, we have furthered the state-of-the-art in the use of AI/ML in the display panel manufacturing process.

#### 5. Conclusions

In this paper, we propose a novel outpainting method that can be used enable data augmentations for existing display panel defect image dataset. It converts an input NG image to a large one with outpainted normal patterns around the input area, and generates a matching large OK to form a matching OK-NG pair too. The large NG output enables random scale and crop augmentation which increases training data variations and make the trained model more robust to general applications. The generated OK-NG pair enables training a defect generation model as OK

conditional, which increase overall quality of the generated NG images while significantly improve background fidelity without need of a mask. Additionally, the proposed outpainting method is trained with a set of masks specifically designed to support learning from the original fixed-size NG images. This self-supervised learning ability is the first known solution to outpaint NG images without a large reference image.

#### 6. References

- Ming W, Cao C, Zhang G, Zhang H, Zhang F, Jiang Z, Yuan J. Application of convolutional neural network in defect detection of 3C products. IEEE Access. 2021 Sep 29;9:135657-74.
- Jeon Y, Kim H, Lee H, Jo S, Kim J. GAN-based Defect Image Generation for Imbalanced Defect Classification of OLED panels. In Proceedings of the Eurographics Symposium on Rendering 2022.
- Pan Z, Shenoy R, Balakrishnan K, Cheng Q, Lee J, Jeon Y, Jeong D, Kim J. 32-1: Improving QD Backplane Defect Image Generation Using Automatic Masking in Diffusion Models. In SID Symposium Digest of Technical Papers 2024 Jun (Vol. 55, No. 1, pp. 409-412).
- Shenoy R, Pan Z, Balakrishnan K, Cheng Q, Lee J, Jeon Y, Jeong D, Kim J. Diffusion Models for Generation of Synthetic QD Backplane Defect Images. In The 24th International Meeting on Information Display, 2024
- Rombach R, Blattmann A, Lorenz D, Esser P, Ommer B. High-resolution image synthesis with latent diffusion models. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition 2022 (pp. 10684-10695).
- Kim K, Kong K, Kang SJ. A Review on Deep Learning-based Image Outpainting. Journal of Broadcast Engineering. 2021;26(1):61-9.
- Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y. Generative adversarial nets. Advances in neural information processing systems. 2014;27.
- Ho J, Jain A, Abbeel P. Denoising diffusion probabilistic models. Advances in neural information processing systems. 2020;33:6840-51.
- Lugmayr A, Danelljan M, Romero A, Yu F, Timofte R, Van Gool L. RePaint: Inpainting using denoising diffusion probabilistic models. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition 2022 (pp. 11461-11471).
- Song J, Meng C, Ermon S. Denoising diffusion implicit models. arXiv preprint arXiv:2010.02502. 2020 Oct 6.
- <https://github.com/Kameronski/stable-diffusion-1.5>
- Lu C, Zhou Y, Bao F, Chen J, Li C, Zhu J. DPM-Solver++: Fast solver for guided sampling of diffusion probabilistic models. arXiv preprint arXiv:2211.01095. 2022 Nov 2.
- Heusel M, Ramsauer H, Unterthiner T, Nessler B, Hochreiter S. GANs trained by a two time-scale update rule converge to a local Nash equilibrium. Advances in neural information processing systems. 2017;30.
- Zhang R, Isola P, Efros AA, Shechtman E, Wang O. The unreasonable effectiveness of deep features as a perceptual metric. In Proceedings of the IEEE conference on computer vision and pattern recognition 2018 (pp. 586-595).