

저비용 고해상도 불량 이미지 생성을 위한 확산 모델 기반 인페인팅

김태형*, 이상민**

Diffusion Model-based Inpainting for Low-Cost, High-Resolution Defect Image Generation

Tae-Hyong Kim*, Sang-Min Lee**

이 연구는 국립금오공과대학교 대학 연구과제비로 지원되었음(2022-2024)

요 약

딥러닝 기반 확산 모델은 높은 품질의 이미지를 생성할 수 있어 다양한 산업 분야에서 주목받고 있으나, 학습과 생성 과정에 많은 시간과 비용이 소요된다는 한계가 있다. 본 연구는 생산 제품의 불량 검출에 활용할 수 있는 고해상도 불량 이미지를 저비용으로 생성하기 위한 확산 모델 기반 인페인팅 기법을 제안한다. 제안 기법은 불량 영역이 제품의 일부에 국한되는 점을 활용하여 전체 이미지에서 생성 대상 영역을 크롭한 후, 해당 영역의 마스크 부분에 인페인팅을 적용함으로써 제한된 비용으로 원하는 크기의 고해상도 불량 이미지를 생성한다. 특히, 제안하는 LWCoPaint 인페인팅 기법은 경량 저해상도 모델에서 생성된 이미지를 고해상도 모델로 개선하는 방식을 채택하여, 기존 CoPaint 모델 대비 생성 비용을 대폭 절감하면서도 보다 우수한 품질의 불량 이미지를 생성한다. 본 기법은 불량 검출을 위한 고품질 데이터셋 구축에 기여할 것으로 기대된다.

Abstract

Deep learning-based diffusion models have attracted attention in various industries due to their ability to generate high-quality images. However, these models require substantial time and cost during both training and generation processes. This study proposes a diffusion model-based inpainting technique to generate high-resolution defect images at low cost that can be used for defect detection of manufactured products. The proposed technique crops the target generation region from the entire image by utilizing the fact that the defects are typically confined to a portion of the product, and then applies inpainting to the masked area to produce a high-resolution defect images of the desired size at a limited cost. In particular, the proposed LWCoPaint inpainting technique refines images generated by a lightweight low-resolution model using a high-resolution model, significantly reducing the generation cost compared to the existing CoPaint model while generating defect images of better quality. We expect that this approach will contribute to the development of high-quality datasets for defect detection.

Keywords

diffusion model, inpainting, image generation, defect detection, smart factory

* 국립금오공과대학교 인공지능공학과 교수(교신저자)

- ORCID: <https://orcid.org/0000-0003-3806-2517>

** 국립금오공과대학교 컴퓨터공학과 학부과정

- ORCID: <https://orcid.org/0009-0009-9063-9708>

· Received: Jan. 20, 2025, Revised: Feb. 17, 2025, Accepted: Feb. 20, 2025

· Corresponding Author: Tae-Hyong Kim

Dept. of Artificial Intelligence Engineering, Kumoh National Institute of
Technology, 61 Daehak-ro, Gumi, Gyeongbuk, Korea

Tel.: +82-54-478-7528, Email: taehyong@kumoh.ac.kr

I. 서 론

최근 인공지능 기술의 발전은 스마트 제조, 공정 자동화, 자동화된 품질 관리, 비용 절감, 생산 효율성 향상, 작업 환경 개선 등 제조 산업 전반에 걸쳐 혁신을 가져왔다. 품질 관리 분야에서는 컴퓨터비전 기술과 시계열 분석 기술을 활용하여 제조 공정상의 이상 상황과 제품의 품질의 문제를 감지하는 기술 등이 활발히 연구되고 있다. 특히, 딥러닝 기반의 영상 분석 기법을 활용한 제품의 불량 검출 기술은 비약적인 성능 향상을 보이며 상용화되어 실제 제조업 현장에 도입되고 있다[1].

그러나 다양한 제조 환경에 필요한 딥러닝 기반의 불량 검사 시스템을 개발하기 위해서는 적지 않은 준비 과정이 요구된다. 딥러닝 학습을 위해 다수의 정상 제품 이미지와 유형별 불량 이미지를 확보해야 하며, 지도 학습을 위해서는 수집된 데이터에 대해 불량 여부, 종류, 위치 정보 등을 표시하는 라벨링 작업이 필요하다. 이러한 과정은 대부분 전문가에 의한 정교한 수작업이 요구되기 때문에 대부분의 중소기업은 이러한 비용을 부담하기 어려운 상황이다.

딥러닝을 위한 학습 데이터가 부족한 문제를 극복하기 위해 전이 학습(Transfer learning)[2], 퓨샷 학습(Few-shot learning)[3] 등을 활용할 수 있지만, 제조 현장에 도입할 수 있는 수준의 정확도를 얻기 위해서는 학습 데이터의 양이 중요하기 때문에, 부족한 데이터의 증강(Augmentation)[4]이 필수적이다. 이미지의 기하학적 변환 및 색상 기반의 변환 등 전통적인 증강 기법은 생성 이미지의 다양성을 확보하는데 한계가 있어 최근에는 GAN(Generative Adversarial Networks)[5], 확산(Diffusion) 모델[6] 등 생성 모델을 통한 데이터 증강 기법이 활발히 연구되고 있다.

제조 현장에서 생성 모델을 기반으로 불량 검출용 학습 이미지를 생성하기 위해서는 다음과 같은 문제를 해결해야 한다. 제품에 존재할 수 있는 미세한 결함을 인식하기 위해서 일반적으로 고해상도의 머신비전 카메라로 제품 이미지를 촬영한다. 만약 강판(Steel plate)과 같이 제품이 큰데 결함 부분은

작을 경우 촬영, 이미지의 해상도가 매우 높아야 하기 때문에 모델 학습과 생성에 많은 시간이 요구될 수 있다. 특히 최근 가장 우수한 품질의 이미지 생성 성능을 보여주는 확산 모델의 경우 각 시간 단계(Time step)를 반복하여 모델 학습과 추론이 수행되기 때문에, 고해상도의 이미지를 생성하기 위해서는 상당한 시간 비용이 발생한다.

본 연구는 이러한 문제를 해결하기 위해 고성능의 확산 모델을 사용하면서도 고해상도의 불량 이미지 생성 비용을 절감할 수 있는 방법을 제안한다. 일반적으로 결함은 제품의 일부 영역에서 작은 크기로 관찰되기 때문에, 제안하는 기법은 불량 이미지 생성 영역을 제품 이미지의 일부로 국한하고, 해당 영역의 마스크 부분에 인페인팅(Inpainting)[7] 방식으로 불량을 생성한다. 이를 통해 생성하는 이미지의 크기와 관계 없이 제한된 비용으로 이음매가 없는(Seamless) 고해상도 이미지를 생성할 수 있다. 특히, 최신 확산 모델 기반 인페인팅 기법을 분석하고 개선하여, 높은 품질의 불량 이미지를 생성하면서도 생성 시간은 최소화할 수 있는 LWCoPaint 인페인팅 기법을 제안한다.

본 논문의 구성은 다음과 같다. II장에서 불량 이미지 생성에 사용할 수 있는 확산 모델 기반의 최신 이미지 생성 기술을 분석한 후, III장에서 제안하는 확산 모델 인페인팅 기반 저비용 고해상도 불량 이미지 생성 방법의 세부 내용을 설명한다. IV장에서 학습 데이터셋 및 전처리 과정을 소개하고, 제안 기법의 성능 평가 방법과 평가 결과를 설명한다. 마지막으로 V장에서 결론 및 향후 연구 방향을 제시한다.

II. 관련 연구

본 절에서는 불량 이미지 생성에 활용될 수 있는 확산 모델 기반의 최신 이미지 생성 기술에 대해 소개한다. 우선 기본적인 DDPM(Denoising Diffusion Probabilistic Models)[8]과 DDIM(Denoising Diffusion Implicit Models)[9]에 대해 설명한 후, 특정 조건 기반 생성이 가능한 유도 확산 모델(Guided diffusion models)[10]을 소개한다. 다음으로 최신 확산 모델 기반 인페인팅 기법들에 대해 살펴보고 비교 분석한다.

2.1 DDPM과 DDIM

확산 모델 중 가장 널리 사용되고 있는 모델은 DDPM이다. DDPM은 데이터에 가우시안 노이즈 (Gaussian noise)를 조금씩 추가하여 정보가 소실되도록 한 후 반대의 과정을 통해 노이즈를 조금씩 제거하여 원래의 데이터를 얻는 생성 모델이다. 가우시안 노이즈를 단계적으로 추가하여 원래 데이터를 완전한 가우시안 노이즈로 만드는 과정을 통해 노이즈가 복잡한 데이터의 분포를 모두 포함하게 하여, 데이터 생성 시 모드 붕괴 발생 가능성을 줄이고 단계적 생성 과정을 통해 고정밀의 데이터를 생성할 수 있게 해 준다. DDPM의 순방향 및 역방향 확산 과정은 그림 1과 같이 나타낼 수 있다.

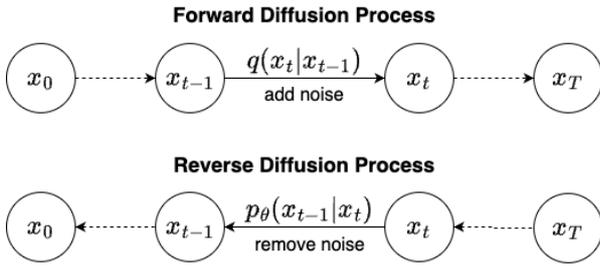


그림 1. DDPM의 순방향 및 역방향 확산과정

Fig. 1. Forward and reverse diffusion process of DDPM

DDPM의 순방향 확산 과정은 데이터에 서서히 노이즈를 추가해가는 과정으로 중간 단계의 잠재변수 x_1, \dots, x_T 에 대한 고정된 추론 과정이라 할 수 있다. 시간 단계 t 에서 추가되는 노이즈의 양을 β_t 라 할 때, 데이터 x_t 의 확률 분포는 식 (1)과 같이 정규분포로 나타낼 수 있다.

$$q(x_t|x_{t-1}) \sim N(\sqrt{1-\beta_t}x_{t-1}, \beta_t I) \quad (1)$$

순방향 확산과정에서는 수학적으로 확산 커널 분포를 도출할 수 있으므로, 노이즈를 반복적으로 추가하는 일련의 중간 단계를 거치지 않고, 식 (2)와 같이 학습에 필요한 시간 단계 t 에서의 노이즈가 있는 데이터를 원래 데이터로부터 직접 구할 수 있

다. 식 (2)에서 $\alpha_t = 1 - \beta_t$, $\bar{\alpha}_t = \prod_{i=1}^t \alpha_i$ 이다.

$$q(x_t|x_0) \sim N(\sqrt{\bar{\alpha}_t}x_0, (1-\alpha_t)I) \quad (2)$$

역방향 확산 과정은 순방향 확산 과정과 달리 노이즈를 제거하기 위한 조건부 확률 $q(x_{t-1}|x_t)$ 을 알 수 없으므로 이를 근사화할 수 있는 식 (3)과 같은 신경망 모델 $p_\theta(x_{t-1}|x_t)$ 를 학습해야 한다. 즉, 역방향 확산 과정에서는 신경망 모델을 학습하여 단계적으로 데이터의 노이즈를 제거한다. 이때, 모델은 식 (4)와 같이 시간 단계 t 에서 추가된 노이즈 ϵ_t 만 예측하면 되고, 샘플링 단계에서는 다양성을 위해 식 (5)의 분산을 가지는 노이즈를 섞어준다.

$$p_\theta(x_{t-1}|x_t) \sim N(\tilde{\mu}_t(x_t, x_0), \tilde{\beta}_t I) \quad (3)$$

$$\tilde{\mu}_t(x_t, x_0) = \frac{1}{\sqrt{\alpha_t}} \left(x_t - \frac{1-\alpha_t}{\sqrt{1-\bar{\alpha}_t}} \epsilon_t \right) \quad (4)$$

$$\tilde{\beta}_t = \frac{1-\bar{\alpha}_{t-1}}{1-\alpha_t} \beta_t \quad (5)$$

DDPM의 간소화된 목적함수 \mathcal{L} 는 식 (6)과 같이 노이즈의 예측 오차로 나타낼 수 있다. 식 (6)에서 $\epsilon_\theta(x_t, t) = \epsilon_\theta(\sqrt{\bar{\alpha}_t}x_0 + \sqrt{1-\bar{\alpha}_t}\epsilon_t, t)$ 이다.

$$\mathcal{L} = E_{t \sim [1, T], x_0, \epsilon_t} [\|\epsilon_t - \epsilon_\theta(x_t, t)\|^2] \quad (6)$$

DDIM은 DDPM의 데이터 생성을 위한 샘플링 속도가 GAN에 비해 현저히 느린 문제를 개선하고자 확률론적으로 진행되던 샘플링 과정을 결정론적으로 전환할 수 있게 함으로써 더 적은 시간 단계로 DDPM과 비견되는 성능을 낼 수 있게 만든 모델이다. DDPM의 순방향 확산과정이

$$q(x_{1:T}|x_0) = \prod_{t=1}^T q(x_t|x_{t-1}) \text{인 마르코프 체인(Markov chain)[11]인 반면 DDIM에서는 순방향 확산과정을}$$

식 (7)~(8)과 같은 추론 분포 q_σ 를 가지는 비마르코프 과정(non-Markovian process)으로 나타낼 수 있다.

$$q_\sigma(x_{t-1}|x_t, x_0) = N(\mu_t, \sigma_t^2 I) \quad (7)$$

$$\mu_t = \sqrt{\bar{\alpha}_{t-1}}x_0 + \sqrt{1-\bar{\alpha}_{t-1}-\sigma_t^2} \frac{x_t - \sqrt{\bar{\alpha}_t}x_0}{\sqrt{1-\bar{\alpha}_t}} \quad (8)$$

DDIM은 DDPM의 역방향 확산에 대한 확률론적 예측 $p_\theta(x_{t-1}|x_t)$ 을 현재 상태 x_t 와 초기 상태 예측 $\hat{x}_0^{(t)} = f_\theta^{(t)}(x_t)$ 을 이용한 결정론적 예측이 가능한 $q_\sigma(x_{t-1}|x_t, \hat{x}_0^{(t)})$ 으로 치환하였다. 학습을 통해 t 가 작아질 때 예측값 $\hat{x}_0^{(t)}$ 은 실제값 x_0 와 유사해진다. 여기서 σ_t 는 확산 과정의 확률적 특성을 제어하는 값으로 $\sigma_t \neq 0$ 이면 확산 과정이 DDPM과 같이 확률적 특성을 가진다. $\sigma_t = 0$ 일 때 역확산 과정은 각 시간 단계에서 노이즈 샘플링이 없는 결정론적 과정이 되므로 전체 단계 수를 줄여 빠르게 샘플링할 수 있다. DDIM은 DDPM의 학습된 매개변수를 생성 과정에 사용할 수 있어 DDPM으로 학습된 모델을 DDIM의 빠른 생성 방식으로 샘플링할 수 있다. DDIM의 결정론적 샘플링 과정은 식 (9)와 같다. 식 (9)에서 $\epsilon_\theta^{(t)} = \epsilon_\theta(x_t, t)$ 이다.

$$x_{t-1} = \sqrt{\alpha_{t-1}} \left(\frac{x_t - \sqrt{1 - \alpha_t} \epsilon_\theta^{(t)}}{\sqrt{\alpha_t}} \right) + \sqrt{1 - \alpha_{t-1}} \epsilon_\theta^{(t)} \quad (9)$$

2.2 유도 확산 모델

유도 확산 모델은 확산 모델에 클래스 정보를 제공하여 조건부 생성이 가능하도록 만든 모델로 분류기 유도(CG, Classifier Guidance) 모델[12]과 분류기 없는 유도(CFG, Classifier-Free Guidance) 모델[13]로 나눌 수 있다.

CG 모델은 노이즈가 있는 데이터 x_t 에 대해 분류기 $p_\phi(y|x_t, t)$ 를 학습한 다음 $\nabla_{x_t} \log p_\phi(y|x_t, t)$ 를 이용해 임의의 클래스 y 로 샘플링 과정을 유도한다. 샘플링에는 DDPM 및 DDIM 방식을 모두 사용할 수 있다. 식 (10)~(11)은 DDIM 방식 CG 모델의 샘플링 방법을 보여준다.

$$\hat{\epsilon} \leftarrow \epsilon_\theta(x_t) - \sqrt{1 - \alpha_t} \nabla_{x_t} \log p_\phi(y|x_t, t) \quad (10)$$

$$x_{t-1} \leftarrow \sqrt{\alpha_{t-1}} \left(\frac{x_t - \sqrt{1 - \alpha_t} \hat{\epsilon}}{\sqrt{\alpha_t}} \right) + \sqrt{1 - \alpha_{t-1}} \hat{\epsilon} \quad (11)$$

CFG 모델은 별도의 분류기를 학습시키는 대신 $\epsilon_\theta(x_t)$ 의 매개변수를 갖는 무조건부(Unconditional) 모델 $p_\phi(y)$ 와 $\epsilon_\theta(x_t, y)$ 의 매개변수를 갖는 조건부 모델 $p_\phi(x|y)$ 을 함께 학습시켜 노이즈를 예측한 다음, 식 (12)와 같이 두 노이즈의 보간(Interpolation)을 통해 최종 노이즈 $\tilde{\epsilon}_\theta(x_t, y)$ 를 예측한다. 이 식에서 ω 는 유도 강도(Guidance strength)로 이 값이 커질수록 특정 클래스의 데이터를 더 많이 생성하게 된다.

$$\tilde{\epsilon}_\theta(x_t, y) = (1 + \omega)\epsilon_\theta(x_t, y) - \omega\epsilon_\theta(x_t) \quad (12)$$

2.3 확산 모델 인페인팅

확산 모델의 학습 및 추론 비용을 줄이기 위해 지식 증류(Knowledge distillation), 양자화, 가지치기(Pruning) 및 알고리즘 최적화 등 다양한 경량화 연구가 시도되었다[13]. 본 연구에서는 생성하고자 하는 제품 이미지에서 불량 부분이 전체의 일부분이라는 점에 착안하여 이미지 크롭과 인페인팅 기법을 함께 사용하여 확산 모델의 비용을 줄이고자 한다. 본 연구를 위한 최적 인페인팅 기법을 도출하기 위해 최신 확산 모델 기반 인페인팅 기법을 살펴본다.

RePaint[14]는 이미 학습된 무조건부 DDPM을 이용해 노출된 영역은 순방향 확산 과정으로, 마스크로 감추어진 영역은 역방향 확산 과정으로 생성한다. 그림 2는 RePaint의 인페인팅 개념을 보여준다.

이미지 x_t 에서 x_{t-1} 로 가는 역방향 확산 과정에서 m 을 마스크, \odot 를 요소별 곱(Elementwise product)이라 하면 인페인팅을 통해 생성하는 이미지 x_{t-1} 는 식 (13)~(15)로 나타낼 수 있다.

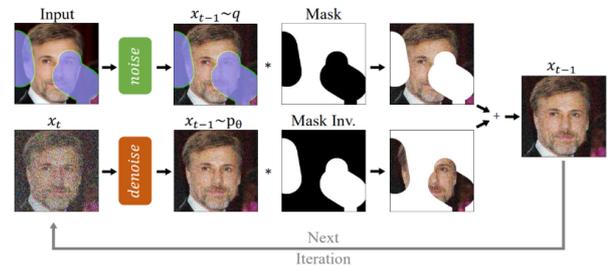


그림 2. RePaint의 인페인팅 과정[14]
Fig. 2. Inpainting process of the RePaint[14]

$$x_{t-1}^{\text{known}} \sim N(\sqrt{\bar{\alpha}_t} x_0, (1 - \bar{\alpha}_t) I) \quad (13)$$

$$x_{t-1}^{\text{unknown}} \sim N(\mu_\theta(x_t, t), \Sigma_\theta(x_t, t)) \quad (14)$$

$$x_{t-1} = (1 - m) \odot x_{t-1}^{\text{known}} + m \odot x_{t-1}^{\text{unknown}} \quad (15)$$

인페인팅에서 중요한 것은 주변의 노출된 영역과 매칭되면서도 의미적으로 정확한 영역 이미지를 생성하는 것이다. RePaint는 마스크된 영역 x_{t-1}^{unknown} 을 생성할 때 노출된 조건 영역 x_{t-1}^{known} 과 더욱 조화될 수 있도록 x_{t-1} 을 다시 x_t 로 확산하는 리샘플링(Resampling) 기법을 사용한다.

CoPaint[15]는 노출된 영역과 감주어진 영역 사이의 불연속성 문제를 해결하기 위해 베이지안 프레임워크를 기반으로 두 영역을 동시에 생성함으로써 보다 자연스럽게 일관된 전체 이미지를 얻을 수 있도록 복원을 수행한다. CoPaint는 생성 이미지에서 노출된 영역을 원래 이미지로 대체하는 대신 직접적인 인페인팅 제약 조건(constraint) $C: r(x_0) = s_0$ 를 사용하여 확산 과정을 수행한다. 여기서 $r(\cdot)$ 은 노출된 영역을 출력하는 연산자, s_0 는 주어진 참조 이미지의 노출 영역이다. 제약 조건 C 기반의 학습을 위해 식 (16)의 로그가능도가 최대가 되도록 x_t 를 최적화한다.

$$\begin{aligned} \log p_\theta(x_{t-1}|x_t, C) = & \quad (16) \\ & -\frac{1}{2\sigma_t^2} \|x_{t-1} - \tilde{\mu}_t\|_2^2 \\ & -\frac{1}{2\xi_{t-1}^2} \|s_0 - r(f_\theta^{(t-1)}(x_{t-1}))\|_2^2 + c \end{aligned}$$

CoPaint는 계산 효율성을 위해 DDPM 대신 DDIM을 사용하므로 식 (7)~(8)의 단계적인 과정을 통해 인페인팅된 이미지를 생성한다.

M2S(Minutes to Seconds)[16]는 RePaint와 같은 DDPM 기반의 이미지 인페인팅을 위해 제안된 가속화 기법이다. M2S는 모델의 매개변수 수를 줄이기 위한 사전학습 경량 확산 모델(LWDM, Light-Weight Diffusion Model)과 확산 단계 수를 줄이기 위한 DDIM 샘플링을 사용하고, 추론 속도 향상을 위해 저해상도 이미지 샘플링과 고해상도 이

미지 샘플링의 두 단계로 샘플링을 수행하는 CFS(Coarse-to-Fine Sampling) 기법을 사용한다. 저해상도 샘플링을 먼저 수행함으로써 상대적으로 시간이 많이 걸리는 고해상도 샘플링의 시간 단계를 줄여 시간 비용을 감소시킬 수 있다. LWDM은 이미지 x_t 의 신호대 잡음비 $SNR(t)$ 를 사용하여 노이즈가 큰 단계에서 손실함수의 가중치를 낮추고 노이즈가 작은 단계에서는 손실함수의 가중치를 높이는 방식으로 학습을 최적화하여 모델을 경량화하는 기법이다.

III. 제안 기법

본 절에서는 제안하는 확산 모델 인페인팅 기반 고해상도 불량 이미지 생성 기법을 설명한다. 먼저 전체 구조를 제시하고, 제안하는 인페인팅 기법을 소개한다.

3.1 전체 구조

그림 3은 본 논문에서 제안하는 불량 이미지 생성 절차를 보여준다. 대상 제품의 불량 이미지를 생성을 위해 우선 제품을 촬영하여 학습 데이터를 수집해야 한다. 본 연구에서는 학습에 필요한 수집 데이터 수를 줄이기 위해 비교적 단순한 표면 패턴을 가지는 강판 데이터를 사용하였다. 유사한 특성을 가지는 제품으로 조직(Texture) 구조를 가지는 옷감, 가죽 등이 있다. 특정 외형을 가지는 제품을 사용할 경우, 외형이 복잡하다면 학습 데이터를 충분히 확보할 필요가 있다.

먼저 머신 비전 카메라 등을 사용하여 촬영된 원본 이미지로부터 정상 및 불량 데이터셋을 구축한다. 이 데이터셋은 원본 강판 이미지에서 각각 불량이 포함되지 않은 영역과 포함된 영역을 정사각형 모양의 $s \times s$ 크기로 잘라내어 만든다. 일반적으로 GAN 또는 확산 모델 등의 생성 모델은 이러한 불량 데이터셋을 학습하여 동일한 크기의 불량 이미지를 생성하게 된다.

제안하는 기법은 생성 모델의 학습 및 추론 비용을 절감하기 위해 고해상도의 불량 데이터셋 이미

지를 다시 크롭(Crop)하여 저해상도의 확산 모델 학습용 이미지를 만든다. 데이터셋 이미지에서 불량 부분이 차지하는 영역은 대부분 크기가 작아서 전체 이미지를 학습할 필요가 없기 때문이다. 크롭 이미지의 크기는 불량 부분보다 크게 정하되, 학습 효율성을 고려하여 전체 이미지에 대해 동일한 정사각형 크기 $c \times c$ ($c < s$)가 되도록 한다.

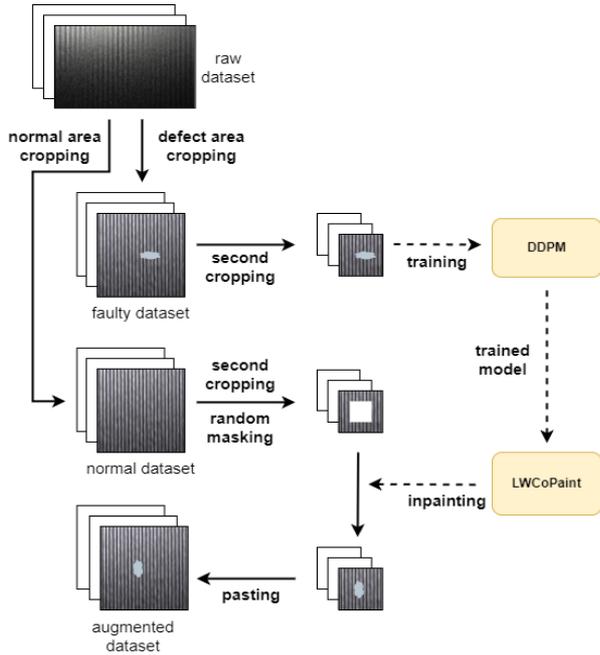


그림 3. 제안하는 고휘상도 불량 이미지 생성 과정
Fig. 3. Proposed generation procedure of high-resolution defect images

이제 불량 이미지를 생성할 수 있는 모델을 얻기 위해 DDPM을 이용하여 $c \times c$ 크기로 2차 크롭된 불량 이미지를 학습한다. 본 연구에서는 불량 이미지만을 학습하므로, 불량 생성 조건을 주기 위한 유도 확산 모델은 사용하지 않는다. 학습된 DDPM 모델을 사용해 샘플링을 수행하면 동일 크기의 불량 이미지를 생성할 수 있다. 그런데, 이 모델은 불량 부분뿐만 아니라 배경도 함께 새로 생성되므로, 2차 크롭 후 남은 외부 강판 이미지를 다시 덧붙이면 경계면이 일치하지 않아 확장된 이미지를 사용할 수 없다.

제안하는 기법은 이 문제를 해결하기 위해 불량 데이터셋과 동일하게 정상 데이터셋도 2차 크롭을 수행하고 크롭 이미지에서 테두리 영역은 그대로 유지하면서 $p \times p$ ($p < c$) 크기의 내부 마스크 영역

만 생성하는 인페인팅 방식으로 이미지를 샘플링한다. 이때 인페인팅 모델은 위에서 학습된 DDPM 모델을 그대로 사용할 수 있다. 인페인팅을 통해 생성한 불량 이미지에서 마스크 밖의 테두리 배경 부분은 원래의 2차 크롭 이미지와 동일하므로 2차 크롭 후 남은 원래의 외부 강판 이미지를 다시 이어 붙이면 이음매가 없는 $s \times s$ 크기의 고휘상도 불량 이미지를 만들 수 있다.

제안하는 기법의 이미지 생성 비용은 기존 $s \times s$ 크기에서의 생성 기법 대비 대략 $(c/s)^2$ 만큼으로 대폭 감소하게 된다. 또한, $c \times c$ 크기의 크롭 영역은 불량 부분의 위치에 따라 결정되며, $p \times p$ 크기의 내부 마스크 영역의 위치는 임의의 내부 위치로 자유롭게 정할 수 있다.

추가로, 기존의 확산 모델 기반 인페인팅 기법의 분석을 통해 적은 비용으로 우수한 품질의 불량 이미지를 생성하기 위한 LWCoPaint 인페인팅 기법을 제안한다.

3.2 LWCoPaint

2.3절에서 살펴본 바와 같이 확산 모델 기반의 인페인팅 기법은 마스크 영역에 대한 생성 이미지의 의미론적 정확성, 마스크 영역 경계면에서의 자연스럽게 매끄러운 일치성, 인페인팅을 위한 확산 모델 학습 및 추론 비용 등의 측면에서 우수한 성능을 얻는 것을 목표로 하고 있다.

RePaint와 M2S는 모두 DDPM으로 학습된 모델을 사용하는데, M2S는 다양한 경량화 기법을 통해 RePaint 대비 추론 시간을 최대 1/60까지 감소시키면서도 생성 이미지 품질 면에서 거의 동등한 성능을 보여주었다. CoPaint는 DDIM으로 추론을 수행하고 마스크 영역만이 아닌 전체 이미지를 생성하는 방식을 사용하여 RePaint 대비 생성 이미지의 품질을 높이면서도 추론 시간을 약 50%까지 감소시켰다.

본 연구에서는 우수한 품질의 불량 이미지를 저비용으로 생성하기 위해 가장 우수한 이미지 품질을 보여주는 CoPaint 기법에 M2S의 경량화 전략을 적용한 LWCoPaint(LightWeight CoPaint) 인페인팅 기법을 제안한다.

제안하는 LWCoPaint 기법은 기존의 CoPaint 기법에 두 가지 방식의 경량화를 적용하였다. 첫 번째 전략은 CFS 샘플링을 사용하는 것이다. CFS 샘플링에서는 원래 이미지를 가로, 세로 각각 1/4 크기로 축소시킨 후, 축소된 저해상도 단계에서 샘플링 수행하여 \tilde{x}_0 를 구하고 이를 원래의 고해상도로 업샘플링하여 다시 샘플링을 진행한다. 이때, 고해상도 샘플링의 시간 단계 T^f 를 저해상도 샘플링 시간 단계 T^c 의 약 1/3가량으로 줄일 수 있어, 전체적으로는 추론 시간을 약 2.5배 감소시킬 수 있다.

두 번째 전략은 학습 및 추론 속도를 높이기 위해 DDPM 학습 모델을 경량화하는 것이다. 일반적으로 가중치 수를 줄이거나 양자화하여 학습 모델을 경량화하면 생성 데이터의 품질이 저하될 수 있다. 그러나 CFS 사용 시 저해상도 모델에서 생성된 이미지를 고해상도 모델에서 더욱 세밀하게 개선하는 방식을 사용하기 때문에, 두 모델 모두 생성 비용이 줄어들고 모델을 경량화하더라도 생성 데이터의 품질 저하를 막을 수 있다. 경량화가 이루어지는 부분은 DDPM의 U-Net[17] 인코더에서 어텐션(Attention)이 존재하는 다운샘플링 단계 및 다운샘플링 단계에서의 잔차(Residual) 블록 수, 어텐션 헤드 수, 모델의 채널 수 등이다.

추가로, 경량화에 따른 품질 저하를 방지하기 위해 M2S에서 채택한 시간 단계 손실의 P2 가중치(Perception prioritized weighting)[18]의 효과를 시험할 예정이다. 이 기법은 디노이징으로 인한 이미지의 품질 차이가 작은 시간 단계에서는 손실을 감소시키고 품질 차이나 크게 나는 단계에서는 손실을 증가시켜 모델이 효율적으로 학습을 수행하도록 한다.

LWCoPaint 인페인팅 기법의 알고리즘은 표 1의 알고리즘 1과 같다. LWCoPaint 알고리즘은 이미지 x_0^f 와 마스크 밖 영역 s_0^f 을 입력받아 인페인팅 이미지 \tilde{x}_0 를 출력한다. 우선 x_0^f 과 s_0^f 을 다운샘플링하여 저해상도의 x_0^c 과 s_0^c 을 얻고, 이를 x_0 와 s_0 로 갱신하여 행 3~16까지를 실행함으로써 저해상도의 인페인팅 이미지 \tilde{x}_0 를 얻는다. 이 과정은 하이퍼파라미터 G, K, H를 각각 1로 설정하고 τ 를 5로 설정한 CoPaint와 동일하다. 다음으로 저해상도의 인페인팅 이미지 \tilde{x}_0 를 원래 해상도로 업샘플링하여 새로운

x_0 를 얻고, s_0^f 를 s_0 로 갱신하여 행 3~16까지를 반복한다. 이때 샘플링 시간 단계는 저해상도 시간 단계 T^c 보다 작은 T^f 로 주어진다. 알고리즘 실행을 마치면 최종 인페인팅 이미지 \tilde{x}_0 를 얻을 수 있다.

표 1. LWCoPaint 알고리즘
Table 1. LWCoPaint algorithm

Algorithm 1. LWCoPaint	
Input	<ul style="list-style-type: none"> • image x_0^f, unmasked area s_0^f • coarse stage timesteps T^c, • fine stage timesteps T^f • estimated image at time step t, $f_\theta^{(t)}(x_t)$ • learning rate at time step t, η_t
1:	$x_0^c, s_0^c \leftarrow \text{downsample}(x_0^f, s_0^f)$
2:	$x_0 \leftarrow x_0^c, s_0 \leftarrow s_0^c$
3:	$T \leftarrow T^c, t \leftarrow T, k \leftarrow 1, m \leftarrow \text{coarse}$
4:	initialize $\tilde{x}_T: \tilde{x}_T \sim N(0, I)$
5:	while $t \neq 0$ do
6:	optimize \tilde{x}_t to maximize Equation 17
7:	generate \tilde{x}_{t-1} with Equation 9
8:	$t \leftarrow t - 1$
9:	if $t \bmod 2 = 0$ and $t \leq T - 5$ then
10:	if $k > 0$ then
11:	generate $\tilde{x}_{t+2}: \tilde{x}_{t+2} \sim q(\tilde{x}_{t+2} \tilde{x}_t)$
12:	$t \leftarrow t + 4, k \leftarrow k - 1$
13:	else
14:	$k \leftarrow 1$
15:	endif
16:	endif
17:	if $t = 0$ and $m = \text{coarse}$ then
18:	$x_0 \leftarrow \text{upsample}(\tilde{x}_0), s_0 \leftarrow s_0^f$
19:	$T \leftarrow T^f, t \leftarrow T, k \leftarrow 1, m \leftarrow \text{fine}$
20:	continue
21:	endif
22:	endwhile
output	• \tilde{x}_0

IV. 성능 평가

4.1 데이터셋 및 전처리

전술한 바와 같이 본 연구는 직접 수집한 강판 데이터를 사용하여 성능 평가를 진행하였다. 원본 강판 이미지 데이터셋은 가로, 세로 최대 1.5m 크기의 강판을 초고해상도 머신비전 카메라를 이용해 촬영하여 수집한 것으로, 이미지의 크기는 카메라 설정에 따라 조금씩 다르다.

표 2는 수집된 원본 강판 이미지 데이터셋의 정보를 요약한 것으로, 강판 외부까지 촬영된 이미지도 일부 존재하였다.

표 2. 원본 강판 이미지 데이터셋 정보
Table 2. Raw steel plate image dataset information

Resolution	No. of images	Inclusion of the outside
4096×1300	2,304	O
4096×3000	70	X
3600×1300	6	X
2300×3000	40	X
2280×1600	53	X

수집된 원본 강판 이미지 데이터셋은 가로, 세로 길이의 차이가 크게 나는 비대칭 이미지로 구성되어 있다. 불량 이미지 학습과 인페인팅을 위해 이 데이터셋의 각 이미지를 가로, 세로 512×512 픽셀 크기의 격자 형태로 크롭하였다. 이때 누락 영역이 발생하지 않도록 크롭 이미지가 서로 약간 중첩되도록 하였다. 다음으로 강판 외부가 포함된 이미지는 폐기하고 불량이 포함된 이미지와 그렇지 않은 이미지를 모아 각각 불량 데이터셋과 정상 데이터셋으로 정리하였다. 강판의 불량 종류로는 강판 표면에 발생하는 스크래치, 얼룩, 오염, 부식 등이 있는데, 본 연구에서는 이를 구분하지 않고 사용하였다. 불량 이미지에 대한 모델 학습을 위해 불량 이미지는 다시 256×256 크기의 불량 부분이 포함된 이미지로 2차 크롭하였다. 전처리된 불량 및 정상 데이터셋의 이미지 수는 표 3과 같다.

표 3. 정상 및 불량 데이터셋의 이미지 개수
Table 3. Number of images in the normal and faulty datasets

Resolution	Normal images	Defect images	Total images
512×512	39,276	3,689	42,965
256×256	-	3,689	3,689

4.2 평가 지표 및 평가 방법

제안된 저비용 고해상도 불량 이미지 생성 기법의 성능은 최종 512×512 픽셀 크기로 생성된 불량 이미지의 품질과 학습 및 추론 시간 비용 측면으로 평가하였다. 생성 이미지의 품질은 생성 모델의 평

가에 가장 널리 사용되는 FID(Fréchet Inception Distance)[19]를 사용하였다. FID는 실제 이미지 데이터셋과 생성된 이미지 데이터셋 간의 분포 차이를 측정하여 품질을 평가하는 지표이다. 실제 이미지들과 생성된 이미지들을 Inception-v3 모델[20]에 입력하여 중간 레이어의 특징 벡터를 추출한 후, 각 특징 벡터의 평균과 공분산을 계산하여 두 데이터셋 분포 간의 거리를 계산한다. 실제 이미지 데이터셋 R 와 생성된 이미지 데이터셋 G 간의 FID는 식 (17)과 같이 계산한다.

$$FID(R, G) = \|\mu_R - \mu_G\|_2^2 + \text{Tr}(\Sigma_R + \Sigma_G - 2(\Sigma_R \Sigma_G)^{1/2}) \quad (17)$$

생성 모델 품질 평가 지표 중 IS(Inception Score)[21]는 데이터의 클래스 정보를 사용하는데, 본 연구에서는 정상 이미지 없이 불량 이미지만을 단일 클래스로 사용하므로 평가 지표에서 제외하였다. 또한 인페인팅 기법의 평가에 주로 사용되는 SSIM (Structural Similarity Index Measure)[22]나 LPIPS (Learned Perceptual Image Patch Similarity)[23]는 원본 이미지와 복원 이미지 쌍의 유사도를 평가하는데 본 연구에서는 인페인팅에 의한 이미지 복원이 목적이 아니라 다양한 불량 이미지 생성을 통한 데이터 증강이 목적이므로 역시 평가 지표에서 배제하였다.

생성 시간 비용을 측정하는 지표로는 1 에폭(epoch) 당 학습 시간(분)과 이미지 1장당 생성 시간(초)을 사용하였다. 확산 모델 학습과 추론은 Nvidia RTX6000 GPU와 Intel Xeon 5416S CPU가 장착된 시스템에서 실행하였다.

4.3 평가 결과 및 분석

제안된 LWCoPaint 기반의 고해상도 불량 이미지 생성 기법의 성능을 비교하기 위한 참조 모델로 DDPM, DDIM, RePaint, CoPaint 및 M2S를 사용하였다. DDPM과 DDIM로는 512×512 크기의 불량 데이터셋을 직접 학습하여 동일 크기의 불량 이미지를 생성하였고, 나머지 인페인팅 모델로는 256×256 크

기의 크롭 이미지를 학습하고, 같은 크기의 정상 이미지에 대해 마스크 영역 내에 불량 이미지를 생성하였다. M2S와 LWCoPaint는 CFS를 사용하므로 64×64 크기의 저해상도 불량 이미지를 학습한 모델을 추가로 사용하였다.

먼저 각 모델의 학습 시간과 추론 시간을 비교하여 표 4에 정리하였다. 학습 시간 측면에서는 크롭 이미지와 CFS를 사용하는 모델이 가장 적은 학습 시간을 보여주었다. 크기가 다른 두 개의 모델을 학습해야 하지만 상대적으로 학습 시간이 긴 256×256 크기 이미지 모델의 학습 단계가 짧기 때문이다. 추론 시간에서는 DDIM이 이미지 크롭을 사용하지 않아 생성 이미지의 크기가 큰데도 추론 시간이 가장 짧았다. 반면 RePaint와 CoPaint는 상대적으로 매우 긴 추론 시간을 나타내 인페인팅 샘플링 비용이 상당히 크다는 것을 알 수 있었다. 한편, M2S와 LWCoPaint는 경량화 전략을 통해 인페인팅 시간을 큰 폭으로 감소시켰다. 더욱이, 생성 이미지의 크기가 커질수록 DDPM과 DDIM은 학습 및 추론 시간이 지속적으로 증가하는 반면, 제안 기법처럼 이미지 크롭과 인페인팅을 함께 활용하면 생성 비용이 증가하지 않는다.

표 4. 모델 학습 및 추론 시간
Table 4. Model training and inference time

Configurations		Training time (min/epoch)	Inference time (sec/image)
Resolution	Model		
512×512	DDPM	40	50
	DDIM	40	4
256×256	RePaint	24	192
	CoPaint	24	201
256×256 + 64×64	M2S	14.4 (=8+6.4)	28
	LWCoPaint	14.4 (=8+6.4)	21

다음은 생성 이미지의 품질에 대한 평가 결과를 살펴본다. 여기서 주목해야 할 점은 DDPM 학습 모델의 불량 이미지 생성 비율이다. DDPM 학습은 비지도 학습이므로 학습 시 검증 데이터셋의 손실값을 이용해 과적합이 발생하지 않은 최적 학습 지점을 찾을 수 없다. 따라서 일정한 에폭마다 학습된 모델이 원하는 이미지를 잘 생성하는지 확인해야

한다. 본 연구는 불량 이미지를 생성하는 게 목적이므로 불량 데이터셋만 학습하는데, 학습 에폭이 늘어나면 불량 이미지 생성 비율이 감소하는 결과가 나타났다. 불량 데이터라도 이미지 전체 면적의 상당 영역은 정상 부분에 해당하므로, 일정한 모양을 가지는 정상 강판을 생성하는 것이 학습 손실을 확실히 줄이는 방법이므로 학습을 계속 진행하면 이러한 방식으로 일종의 과적합이 발생했다. 반면, 학습 초기 단계 모델에서는 상대적으로 불량 이미지가 많이 발생하지만, 학습이 충분히 되지 않았기 때문에 이미지 품질이 떨어지는 문제가 있다. 따라서 각 모델의 이미지 생성 시간은 표 4의 추론 시간과 같지만 실제로는 생성 이미지 중 불량 이미지를 선별해야 하므로 보다 많은 이미지 생성이 필요하였다. 그런데 CFS를 사용하는 M2S와 LWCoPaint는 저해상도 모델의 초기 학습 단계에서 생성되는 불량 이미지의 품질이 다소 떨어지더라도 고해상도 모델에서 확산을 통해 이를 극복하게 되므로 불량 이미지 생성 효율이 타 모델 대비 높았다. 따라서 실질적인 불량 이미지 추론 시간에서는 DDIM과 거의 대등한 성능을 보여주었다.

각 모델이 생성한 불량 이미지의 품질을 평가하기 위해 512×512 픽셀 크기의 최종 불량 이미지를 각각 1000개씩 생성하고, 동일 크기 불량 데이터셋과의 FID를 산출한 결과를 표 5에 정리하였다.

표 5. 생성 이미지 품질 평가 결과
Table 5. Quality evaluation result of generated images

model	FID ↓	
Whole image generation (512×512)	DDPM	0.6804
	DDIM	0.9454
Generation by inpainting (256×256)	RePaint	0.4014
	CoPaint	0.2856
	M2S	0.2304
	LWCoPaint	0.1975

DDIM은 추론 시간은 빠른 대신 가장 낮은 품질을 나타냈고, 인페인팅 모델로 생성한 이미지의 품질이 상대적으로 우수했다. M2S와 LWCoPaint는 각각 베이스 모델인 RePaint와 CoPaint 대비 생성 이미지의 품질이 개선됨을 보여주었다.

제안된 LWCoPaint는 M2S보다 약 14% 정도 낮은 FID 값을 나타내 비교군 중에서 가장 우수한 품질을 보여주었다.

LWCoPaint에 적용된 경량화 기법들의 영향을 확인하기 위해 제거 연구(Ablation study)를 수행한 결과가 표 6에 정리되어 있다. 표 6에서 경량 모델은 DDPM의 U-Net 인코더에서 어텐션을 다운 샘플링 비율 16에서만 사용하고, 잔차 블록 수, 어텐션 헤드 수, 모델의 채널 수를 각각 1, 1, 128로 설정한 모델이다. 추론 시간 측면에서는, CFS의 적용과 모델 경량화가 각각 58.7%와 69.7%의 비용 감소를 나타내었고, 두 기법을 같이 적용함으로써 추론 시간이 원래 보다 약 1/10로 줄어들었다. 이미지 품질 측면에서는 CFS만 적용하였을 때 이미지 품질이 약간 상승했지만, 모델 경량화만 적용하였을 때는 이미지 품질이 다소 저하되었다. 그러나 CFS와 경량 모델을 함께 사용하였을 때 이미지 품질이 뚜렷하게 개선되었다. 추가로 P2weight의 효과를 시험하였는데 추론 시간 측면에서 이득이 없고 생성 이미지 품질도 저하되어 최종 LWCoPaint 모델에서는 제외하였다.

표 6. LWCoPaint의 성능에 대한 제거 연구
Table 6. Ablation study of LWCoPaint performance

Lightweighting CoPaint ¹⁾			Inference time (sec/image)	FID ↓
CFS ²⁾	Light model	P2weight		
×	×	×	201	0.2856
○	×	×	83	0.2776
×	○	×	61	0.2919
×	×	○	201	0.4283
○	○	×	21	0.1975

주1) hyperparameter configuratin: $G=1, K=2, H=1, \tau=5$

주2) downsampling rate=1/4, $T^c = 250, T^f = 75$

그림 4는 불량 데이터셋의 실제 이미지와 각 모델을 통해 생성한 불량 이미지 샘플을 보여준다. 제한된 샘플이지만 DDPM과 DDIM 대비 인페인팅 기법으로 생성한 불량 이미지의 품질이 더 우수한 것을 확인할 수 있고, 특히 CoPaint, M2S, LWCoPaint이 생성한 불량 이미지는 실제 이미지와 매우 유사한 우수한 품질을 보여준다.

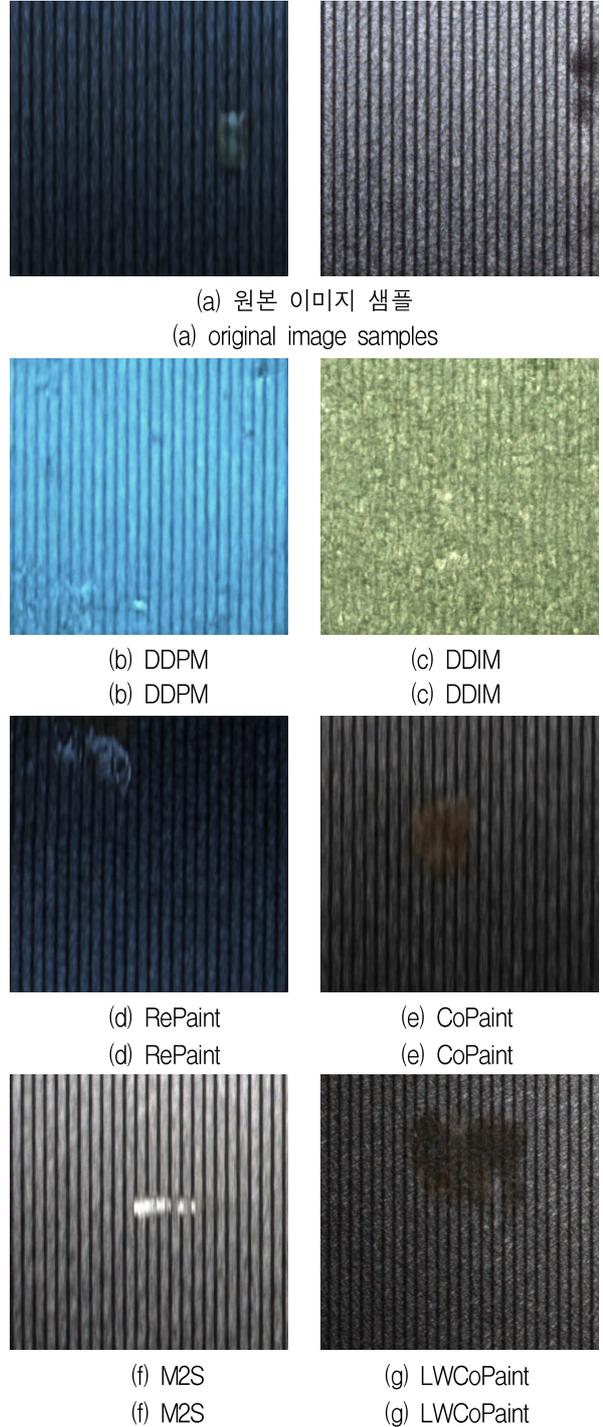


그림 4. 각 모델의 생성 이미지 샘플
Fig. 4. Sample generated images of each model

V. 결론

확산 모델의 등장으로 딥러닝 기반 이미지 생성 기법의 성능이 더욱 우수해져 다양한 산업 분야에서의 활용성이 높아지고 있다.

그런데 확산 모델은 학습과 생성에 모두 확산 과정을 사용하고 있어 시간 비용이 높다는 단점이 있다. 본 연구는 생산 제품의 불량 검출 등에 활용할 수 있는 고해상도 불량 이미지를 저비용 확산 모델 인페인팅 기법으로 생성하는 방법을 제안하였다. 일반적으로 불량 영역은 제품의 일부분에 국한되므로, 불량 영역을 특정 크기 이하로 제한할 경우, 제안한 이미지 크롭 기반 인페인팅은 생성 이미지의 크기와 관계없이 동일한 생성 비용을 가진다는 장점이 있다. 특히, LWCoPaint 기법은 저해상도 모델의 학습량을 줄여 불량 발생 비율을 높이고 고해상도 모델에서 이를 개선하는 방식으로 기존 CoPaint 모델 대비 학습 시간과 추론 시간을 각각 60.0%와 10.4%로 줄이면서도 생성된 불량 이미지의 품질은 FID 기준 30.7% 향상시켰다.

본 연구는 강판을 대상으로 한 불량 이미지 생성에 집중하였으나 강판과 유사한 특성을 가지는 옷감, 가죽 등의 재료에 대해서도 우수한 성능을 기대할 수 있다. 회로 기판과 같은 복잡한 패턴의 제품이나 복잡한 외형을 가지는 제품도 적용에 문제가 없으나 성능의 검증은 필요할 것이다. 한편, 유도 확산 모델을 사용하여 불량을 유형별로 생성하는 방법과 확산 모델 학습 정도와 불량 생성과의 관계에 관해서는 추가 연구가 필요하다. 향후 제안된 기법을 다양한 제품 및 다양한 유형의 불량에 적용하여 범용성을 확인하고 불량 생성을 위한 최적 모델 학습 방안에 관한 연구를 수행할 예정이다.

References

- [1] D. Leite, E. Andrade, D. Rativa, and A. M. A. Maciel, "Fault Detection and Diagnosis in Industry 4.0: A Review on Challenges and Opportunities", *Sensors*, Vol. 25, No. 1, pp. 60, Dec. 2024. <https://doi.org/10.3390/s25010060>.
- [2] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, and Q. He, "A Comprehensive Survey on Transfer Learning", *Proceedings of the IEEE*, Vol. 109, No. 1, pp. 43-76, Jan. 2021. <https://doi.org/10.1109/JPROC.2020.3004555>.
- [3] Y. Song, T. Wang, P. Cai, S. K. Mondal, and J. P. Sahoo, "A Comprehensive Survey of Few-shot Learning: Evolution, Applications, Challenges, and Opportunities", *ACM Computing Surveys*, Vol. 55, No. 13s, pp. 1-40, Jul. 2023. <https://doi.org/10.1145/3582688>.
- [4] A. Mumuni and F. Mumuni, "Data augmentation: A comprehensive survey of modern approaches", *Array*, Vol. 16, Dec. 2022. <https://doi.org/10.1016/j.array.2022.100258>.
- [5] I. J. Goodfellow, et al., "Generative Adversarial Networks", *arXiv preprint arXiv:1406.2661*, Jun. 2014. <https://doi.org/10.48550/arXiv.1406.2661>.
- [6] L. Yang, et al., "Diffusion Models: A Comprehensive Survey of Methods and Applications", *ACM Computing Surveys*, Vol. 56, No. 4, pp. 1-39, Nov. 2023. <https://doi.org/10.1145/3626235>.
- [7] Z. Qin, Q. Zeng, Y. Zong, and F. Xu, "Image inpainting based on deep learning: A review", *Displays*, Vol. 69, article. 102028, Sep. 2021. <https://doi.org/10.1016/j.displa.2021.102028>.
- [8] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models", *Proc. Advances in Neural Information Processing Systems (NeurIPS)*, Online, pp. 6840-6851, Dec. 2020.
- [9] J. Song, C. Meng, and S. Ermon, "Denoising Diffusion Implicit Models", *Proc. International Conference on Learning Representations (ICLR)*, Online, May 2021.
- [10] P. Dhariwal and A. Nichol, "Diffusion models beat GANs on image synthesis", *arXiv preprint arXiv:2105.05233*, May 2021. <https://doi.org/10.48550/arXiv.2105.05233>.
- [11] J. R. Norris, "Markov Chains", Cambridge University Press, 1997.
- [12] J. Ho and T. Salimans, "Classifier-Free Diffusion Guidance", *Advances in Neural Information Processing Systems (NeurIPS)*, Online, Dec. 2021.
- [13] W. Song, W. Ma, M. Zhang, Y. Zhang, and X. Zhao, "Lightweight diffusion models: a survey",

Artificial Intelligence Review, Vol. 57, article. 161, May 2024. <https://doi.org/10.1007/s10462-024-10800-8>.

[14] A. Lugmayr, et al., "RePaint: Inpainting using denoising diffusion probabilistic models", Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, pp. 11461-11471, Jun. 2022. <https://doi.org/10.1109/CVPR52688.2022.01117>.

[15] G. Zhang, et al., "Towards coherent image inpainting using denoising diffusion implicit models", Proc. International Conference on Machine Learning (ICML), Honolulu, Hawaii, USA, pp. 41164-41193, Jul. 2023.

[16] L. Zhang, X. Du, L. TomyEnrique, Y. Wang, Y. Zheng, and C. Jin, "Minutes to Seconds: Speeded-up DDPM-based Image Inpainting with Coarse-to-Fine Sampling", Proc. IEEE International Conference on Multimedia and Expo, Niagra Falls, Canada, pp. 1-6, Jul. 2024.

[17] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation", Proc. International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), Munich, Germany, Vol. 9351, pp. 234-241, Oct. 2015. https://doi.org/10.1007/978-3-319-24574-4_28.

[18] J. Choi, J. Lee, C. Shin, S. Kim, H. J. Kim, and S.-H. Yoon, "Perception Prioritized Training of Diffusion Models", Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, pp. 11462-11471, Jun. 2022. <https://doi.org/10.1109/CVPR52688.2022.01118>.

[19] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter. "GANs trained by a two time-scale update rule converge to a local nash equilibrium", Proc. International Conference on Neural Information Processing Systems (NIPS), California, USA, pp. 6629-6640, Dec. 2017.

[20] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision", Proc. IEEE

Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, pp. 2818-2826, Jun. 2016. <https://doi.org/10.1109/CVPR.2016.308>.

[21] S. Barratt and R. Sharma, "A Note on the Inception Score", arXiv:1801.01973, Jan. 2018. <https://doi.org/10.48550/arXiv.1801.01973>.

[22] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity", IEEE Transactions on Image Processing, Vol. 13, No. 4, pp. 600-612, Apr. 2004. <https://doi.org/10.1109/TIP.2003.819861>.

[23] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The Unreasonable Effectiveness of Deep Features as a Perceptual Metric", Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, pp. 586-595, Jun. 2018. <https://doi.org/10.1109/CVPR.2018.00068>.

저자소개

김 태 형 (Tae-Hyong Kim)



1995년 8월 : 연세대학교
전기전자공학과(공학석사)
2001년 2월 : 연세대학교
전기전자공학과(공학박사)
2002년 9월 ~ 현재 :
국립금오공과대학교
인공지능공학과, 컴퓨터공학과

교수

관심분야 : 딥러닝모델, 시계열분석, 스마트팩토리

이 상 민 (Sang-Min Lee)



2023년 3월 ~ 현재 :
국립금오공과대학교
컴퓨터공학과 학부과정
관심분야 : 생성모델