

일반논문 (Regular Paper)

방송공학회논문지 제30권 제3호, 2025년 5월 (JBE Vol.30, No.3, May 2025)

<https://doi.org/10.5909/JBE.2025.30.3.480>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

확산 모델을 활용한 SAR-광학 영상 변환을 위한 효과적인 데이터 전처리와 증대 기법

서 형 원^{a)}, 이 지 상^{a)}, 박 우 진^{a)}, 최 해 철^{b)}, 김 규 현^{c)}, 장 한 얼^{a)†}

Effective Data Preprocessing and Augmentation Methods of SAR Images for Optical Image Translation

Hyoungwon Seo^{a)}, Jisang Lee^{a)}, Woojin Park^{a)}, Haechul Choi^{b)}, Kyuheon Kim^{c)}, and Haneol Jang^{a)†}

요 약

합성 개구 레이더(SAR) 영상은 야간 및 기상 조건에 제약 없이 고해상도 영상을 획득할 수 있다는 장점이 있지만, 스펙클 노이즈로 인해 영상 품질 저하와 해석의 어려움이 있다. 이러한 한계를 극복하기 위해 SAR 영상을 이용하여 광학 영상과 유사한 보다 유용한 영상으로 변환하는 연구가 필요하다. 본 논문에서는 조건부 확산 모델을 활용하여 SAR 영상을 광학 영상으로 변환할 때 수평 반전 및 90도 회전과 같은 데이터 증강 기법과 함께 제안하는 전처리 방법인 스트레칭 및 $\pm 0.2\%$ 스케일링 기법을 적용하여 실험을 진행하였다. 실험 결과, 조건부 확산 모델에서의 데이터 증대 및 전처리 방법이 FID 평가 지표에서 최대 6.83% 감소하여 광학 영상과 유사한 영상을 생성하도록 유도함을 입증하였다. 제안된 방법은 SAR 영상의 활용 범위를 확대하고 재난 관리, 환경 모니터링, 도시 계획 등 다양한 분야에서 위성 영상 분석 발전에 기여할 수 있을 것으로 기대된다.

Abstract

Synthetic Aperture Radar (SAR) imagery offers the advantage of acquiring high-resolution images regardless of weather or lighting conditions, but its utility is limited by speckle noise, which degrades image quality and complicates interpretation. To overcome these limitations, research is needed to convert SAR images into more useful, optical-like imagery. In this paper, we apply a conditional diffusion model to transform SAR images into optical images, incorporating data augmentation techniques (horizontal flip and 90° rotation) alongside our proposed preprocessing methods-stretching and $\pm 0.2\%$ scaling. Experimental results demonstrate that these augmentation and preprocessing techniques reduce the Frechet Inception Distance (FID) by up to 6.83%, guiding the model to generate images closely resembling real optical imagery. The proposed approach expands the applicability of SAR data and is expected to advance satellite image analysis across disaster management, environmental monitoring, urban planning, and other domains.

Keyword : Synthetic Aperture Radar, Image Translation, Conditional Diffusion Model, Data Preprocessing, Data Augmentation

Copyright © 2025 Korean Institute of Broadcast and Media Engineers. All rights reserved.

“This is an Open-Access article distributed under the terms of the Creative Commons BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited and not altered.”

I. 서론

오늘날 인공위성 기술의 급속한 발전에 힘입어, 위성 영상을 활용한 다양한 응용 분야가 각광을 받고 있다. 재난 관리, 군사 정찰, 도시 계획, 환경 모니터링 등에서 신뢰도 높은 영상 데이터에 대한 수요가 증가하고 있으며, 이 가운데 합성개구레이더(Synthetic Aperture Radar, SAR)는 마이크로파를 이용하여 영상 정보를 획득하기 때문에 악천후나 야간에도 이미지를 취득할 수 있는 장점을 지닌다. 이를 통해 광학 센서가 가진 기상 및 조도 제약을 효과적으로 보완할 수 있다. 그러나 SAR 영상에는 스펙클 노이즈(Speckle Noise)가 존재하여, 영상 내부의 객체나 지형 정보를 식별하기가 까다롭다는 문제가 발생한다. 이는 SAR 영상의 품질을 저하시킬 뿐 아니라, 여러 후속 분석 과정에서 정확한 해석을 어렵게 만드는 요인이 되기도 한다. 이러한 문제를 극복하기 위해서는 SAR 영상을 보다 해석이 용이한 광학 영상 형태로 변환하는 기술이 필요하다.

SAR 영상에서 광학 영상으로 변환하는 연구는 전통적으로 Generative Adversarial Network(GAN) 기반 기법을 사용해 SAR-광학 간의 매핑을 학습하고자 하는 연구가 주를 이루었으나, 최근에는 확산 모델(Diffusion Model)이 제안되어 조건부 확산 모델(Conditional Diffusion Model)을 이용한 새로운 접근 방식이 각광받고 있다. 하지만 기존 연구들은 위성 영상 데이터가 가진 특성을 고려하지 않고 있다. 대부분의 위성 영상 데이터는 한 픽셀이 가지고 있는 값은 0부터 255까지가 아닌 0부터 65,535까지의 범위를 가지고 있다. 이러한 경우 스케일 차이로 인해 모델이 수렴하기 어

렵거나 gradient가 불안정해질 수 있다.

본 논문에서는 조건부 확산 모델을 이용해 SAR 영상을 광학 영상으로 변환하는 과정에서, 더욱 실제 광학 영상에 근접한 결과를 얻기 위해 효과적인 데이터 전처리 방법을 제시하고 증대 방안을 탐색한다. 우선적으로 SAR 영상에서 발생할 수 있는 극단값이나 범위 편차를 줄이는 전처리 방법을 적용하여 확산 모델의 가우시안 노이즈 분포와 입력 데이터의 분포를 맞추고 스펙클 노이즈의 영향을 최소화하였다. 나아가 다양한 데이터 증대 기법을 도입함으로써, 모델이 SAR 영상의 복잡한 장면 특성과 회전·반사 변환 등에 대해 강인성을 갖추도록 하였다. 이를 통해 생성된 영상은 스펙클 노이즈가 완화될 뿐 아니라, 광학 영상에 가깝게 표현되어 도시 구조물 인식이나 지형 분석과 같은 실제 활용 분야에서 큰 기여를 할 것으로 기대된다.

본 논문의 구성은 다음과 같다. 2장에서는 확산 모델 및 이미지 변환 분야에 관한 관련 연구 동향을 살펴보고, 특히 SAR-광학 변환과 관련된 선행 연구를 다룬다. 3장에서는 본 논문에서 제안하는 조건부 확산 모델 기반 데이터 전처리 및 증대 기법을 상세히 설명한다. 4장에서는 주요 실험 설정과 결과 분석을 통해 제안 기법의 우수성을 검증하며, 마지막으로 5장에서 본 논문의 결론과 향후 연구 방향을 논의한다. 본 연구의 주요 의의는 다음과 같다. 첫째, 조건부 확산 모델 기반의 효과적인 데이터 전처리 방법으로 SAR-광학 변환 성능을 향상하였다. 모델이 스펙클 노이즈와 극단값에 더욱 강건하도록 설계하여 학습 안정성과 전반적 화질 개선 효과를 동시에 달성한다. 둘째, 조건부 확산 모델 기반의 효과적인 데이터 증대 기법을 탐색하였다. SAR 이미지 특성을 고려한 적절한 증대 방안을 탐색하여 화질 개선 효과를 얻는다.

II. 관련 연구

1. 확산 모델

최근 몇 년 동안 GAN 모델은 이미지 생성, 이미지 변환 등 다양한 분야에서 주목받아 왔다. 그러나 GAN 모델은 학습 과정에서의 불안정성과 모드 붕괴(mode collapse) 문제로 인해 제한적인 경우에만 효과적이었다. 이러한 문제를

a) 국립한밭대학교 컴퓨터공학과(Department of Computer Engineering, Hanbat National University)

b) 국립한밭대학교 지능미디어공학과(Department of Intelligence Media Engineering, Hanbat National University)

c) 경희대학교 전자공학과(Department of Electronic Engineering, Kyung Hee University)

✉ Corresponding Author : 장한얼(Haneol Jang)

E-mail: hejang@hanbat.ac.kr

Tel: +82-42-821-1142

ORCID: <https://orcid.org/0000-0002-7253-7646>

※ This work was supported by Korea Research Institute for defense Technology planning and advancement(KRIT) grant funded by the Korea government(DAPA(Defense Acquisition Program Administration)) (KRIT-CT-22-047, Space-Layer Intelligent Communication Network Laboratory, 2022).

· Manuscript May 2, 2025; Revised May 7, 2025; Accepted May 7, 2025.

해결하기 위해 확산 모델(Diffusion Models)이 제안되었다^[1]. 확산 모델은 데이터의 분포를 학습하여 점진적으로 노이즈를 추가하고 제거하는 과정을 통해 새로운 데이터를 생성한다. 이 모델은 GAN에 비해 학습이 안정적이며, 고품질의 이미지를 생성할 수 있는 장점이 있다. 조건부 확산 모델은 확산 모델의 한 종류로, 입력 데이터와 추가적인 조건을 기반으로 원하는 특성을 가진 데이터를 생성할 수 있다.

2. 이미지에서 이미지로의 변환

이미지에서 이미지로의 변환 연구는 한 도메인의 이미지를 다른 도메인의 이미지로 변환하는 작업으로, 컴퓨터 비전 분야에서 중요한 연구 주제이다. Pix2Pix^[2]는 짝지어진 데이터셋을 활용하여 조건부 GAN을 적용한 모델로, 입력 이미지와 목표 이미지 사이의 맵핑을 학습시켰다. 그러나 Pix2Pix는 페어링된 데이터셋을 필요로 하기 때문에 제약이 존재한다. 이를 해결하기 위해 CycleGAN^[3]이 제안되었으며, 짝지어지지 않은 데이터셋에서도 이미지 간의 변환을 가능하게 하였다. CycleGAN은 Cycle-Consistency Loss를 도입하여 입력 이미지와 다시 복원된 이미지 간의 차이를 최소화함으로써 두 도메인 간의 변환을 안정적으로 학습할 수 있게 한다. 최근에는 StyleGAN^[4]과 같은 고품질의 이미지 생성 모델이 등장하여 이미지 변환의 품질을 크게 향상시켰으며, GAN의 한계를 극복하기 위해 확산 모델을 사용하여 성능을 높이는 연구가 진행되었다.

3. 합성개구레이더 영상에서 광학 영상으로의 변환

SAR 영상을 광학 영상으로 변환하는 연구는 스펙클 노

이즈같은 SAR 영상의 단점을 해결하고 더 유용한 이미지로 변환할 수 있게 하여 이미지에서 이미지로의 변환 분야에서 중요한 위치를 차지하고 있다. 기존 연구에서는 주로 GAN 모델을 사용하여 SAR 영상에서 광학 영상으로의 변환을 시도했으나^[5], 최근 조건부 확산 모델의 등장으로 SAR 영상 변환 분야에서도 새로운 접근 방식이 가능하게 되어 이와 관련한 연구도 진행되었다^[6]. GAN 모델과 조건부 확산 모델에 대한 블록 다이어그램을 [그림 1]에서 확인할 수 있다. Z. Guo 등^[7]은 쌓이 맞지 않은 SAR 영상과 EO 영상으로 학습할 수 있도록 이미지 사이의 변환 맵핑을 구축하였다. SAR 영상의 복잡한 장면 정보를 효과적으로 학습하기 위해 멀티스케일 표현 생성기와 멀티리셉티브 필드 판별기를 제안하여 SAR 영상의 다양한 스케일의 특징을 추출하고, 노이즈의 영향을 최소화하고 고품질의 광학 영상을 생성하는 MS-GAN 프레임워크를 제안하였다. J. Lee 등^[8]은 SAR 영상을 광학 영상으로 변환하는 과정에서 노이즈를 효과적으로 제거하고 맥락 정보를 잘 보존시키며 저해상도에서 고해상도로의 단계적 학습과 NIR 밴드를 이용한 보조 학습을 통해 SAR의 특성을 잘 반영한 고품질의 광학 영상을 생성하는 CFCA-SET 프레임워크를 제안하였다. 또한 촬영 시점, 각도, 센서 등의 차이로 인해 SAR 영상과 광학 영상이 정렬되어 있지 않을 수 있어 극복하기 위해 Misalignment Resistant Loss를 제안하였다. 하지만 기존 연구들은 위성 영상 데이터가 가진 특성을 고려하지 않아 전처리에 대한 연구가 필요하다. 이에 본 논문에서는 확산 모델에서 효과적으로 가우시안 노이즈와 입력 데이터의 분포를 맞추는 전처리를 제안하여 성능을 개선하였다.



그림 1. GAN 모델과 조건부 확산 모델에 대한 블록 다이어그램

Fig. 1. Block diagrams of the GAN model and the conditional diffusion model

III. 제안 기법

본 연구에서는 조건부 확산 모델^[6]을 기반으로 SAR 영상을 광학 영상으로 변환할 때 더 효과적인 데이터 증대법을 탐색하고 전처리 방법을 제안한다. 전체적인 블록 다이어

그램은 [그림 2]에 추가하였고, Preprocessing 블록에서 그림과 같은 수식을 적용하여 전처리를 진행하였고, 전처리가 적용된 이미지로 데이터 증대를 적용하여 확산 모델을 학습하였다. 보통 SAR 이미지와 광학 이미지가 짝지어진 데이터는 부족하여 데이터 증대가 필요하다. 또한 SAR 영

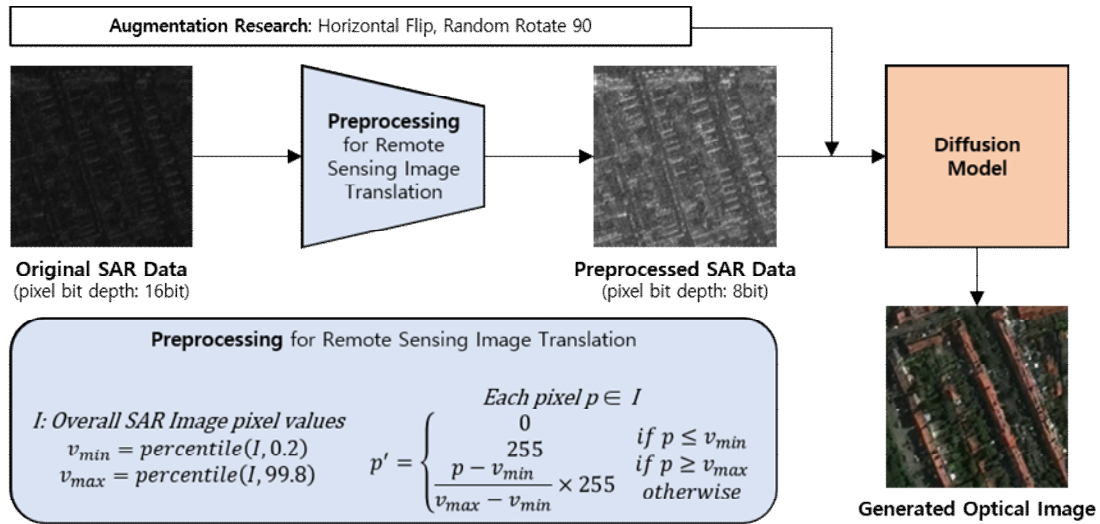


그림 2. 제안하는 방법에 대한 블록 다이어그램
Fig. 2. Block diagram of the proposed method

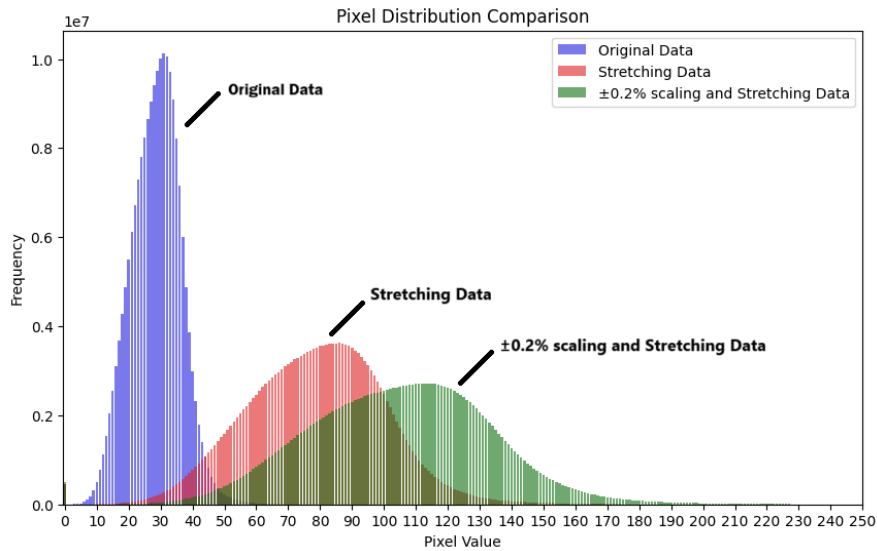


그림 3. 데이터 전처리 방법에 따른 학습 데이터 중 무작위로 선택된 1,000개의 예시 이미지의 픽셀값 분포
Fig. 3. Pixel value distributions according to data preprocessing methods for 1,000 randomly selected training data samples

상은 픽셀값의 변화에 민감하여 정보 손실이 발생할 수 있다. 따라서 본 연구에서는 영상의 기하학적 구조를 유지하면서 데이터를 증대시킬 수 있는 Horizontal Flip과 Rotate 90 기법을 적용하였다. Horizontal Flip은 50%의 확률로 이미지를 수평으로 뒤집어 모델이 다양한 방향의 객체를 학습할 수 있도록 도움을 준다. Rotate 90은 각 25% 확률로 90도, 180도, 270도로 이미지를 회전시켜 모델이 회전 불변성을 학습할 수 있게 한다.

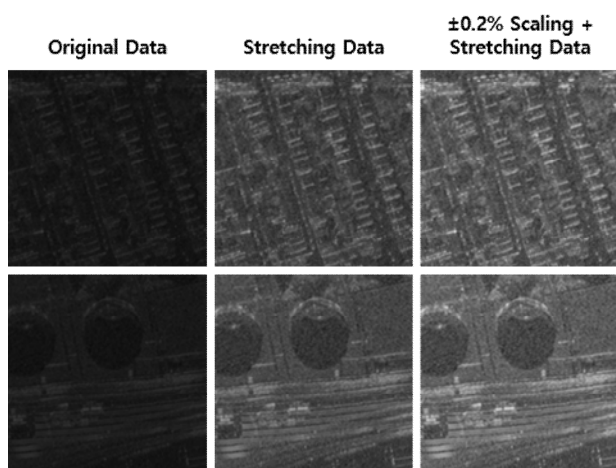


그림 4. 전처리 방법에 따른 SAR 이미지 예시

Fig. 4. Sample SAR images according to the preprocessing methods

일반적인 확산 모델은 uint8의 데이터 형식을 가진 이미지를 정규화하여 학습한다. 하지만 SAR 이미지와 같은 위성 영상 데이터는 uint16의 데이터 형식을 가지고 있다. 이를 동일하게 정규화를 하더라도, 일반적인 데이터값을 가지고 있지 않아 노이즈 주입과 제거 과정에서 모델이 수렴하기 어렵거나 gradient가 불안정해질 수 있다. 이를 해결하기 위해 본 연구에서는 SAR 이미지를 모델의 입력으로 사용하기 위한 전처리 방법으로 Stretching과 $\pm 0.2\%$ Scaling을 적용하여 정보 손실을 최소화하고자 하였다. [그림 3]은 전처리 방법에 따른 픽셀값 분포를 보여준다. 이를 통해 정보 손실이 발생되지 않아 표현되는 픽셀 수가 다양해져 광학 영상으로 변환을 잘 수행할 수 있게 된다. Stretching은 이미지의 픽셀값 범위를 0에서 255로 확장하여 명암 대비를 향상시키고, $\pm 0.2\%$ Scaling은 0.2%~99.8% 범위 바깥에서 극단적인 픽셀 값을 제거하여 노이즈를 줄인다. 이러한

데이터 증대 및 전처리 기법을 통해 변환된 영상의 품질을 향상시키고, 모델의 일반화 능력을 증가시킬 수 있다. [그림 8]을 통해 SAR 이미지에 각 전처리 방법에 따른 이미지의 모습을 확인할 수 있다.

IV. 실험

1. 데이터셋

본 연구에서는 SAR 영상을 광학(EO) 영상으로 변환하기 위해 SpaceNet6 데이터셋^[9]을 활용하였다. SpaceNet6은 Capella Space 위성으로부터 획득한 SAR 영상과 Maxar WorldView-2 위성으로부터 획득한 광학 영상이 한 쌍을 이루는 형태로 총 3,401개의 이미지 쌍을 제공한다. SAR 영상은 완전 편파(full polarization)로 수집되어 지표 구조에 대한 정밀한 분석이 가능하며, 광학 영상은 원본상에서 RGB 및 근적외선(NIR) 밴드를 모두 포함하지만, 본 연구에서는 RGB 채널만을 주로 이용한다. 각 패치는 900×900 크기를 가지며, 공간 해상도(GSD)는 약 0.5m 수준이어서 도시 지역의 세밀한 특징과 건물 구조를 식별하는 데 적합

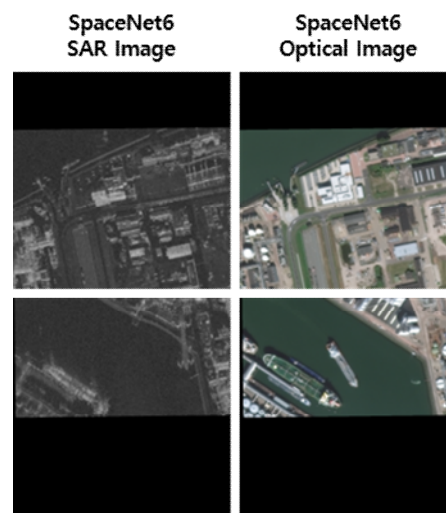


그림 5. SpaceNet6 데이터셋의 SAR과 광학 쌍의 예시

Fig. 5. Sample matched SAR and optical images from the SpaceNet6 dataset

하다. 다만 위성 이미지의 특성상 패치 가장자리에 검은색으로 빈 영역이 다수 포함되어 있는 경우가 많아, 본 실험에서는 전체 픽셀 중 검은색 배경이 5% 이상을 차지하는 이미지는 학습 시 불필요한 정보로 간주하여 제거하였다. 또한, 실험에 사용하기 용이하도록 이미지를 256×256 크기로 잘라 모델의 입력으로 활용하였다. 이 같은 전처리 과정을 거쳐 최종적으로 총 13,627장의 이미지를 학습용으로 확보하였으며, 변환 성능 검증을 위해 별도의 1,024장을 추가적으로 분리하여 사용하였다. [그림 5]에서 SAR 영상과 광학 영상이 짝이 맞는 SpaceNet6 데이터를 확인할 수 있다.

2. 평가 지표

SAR 영상에서 광학 영상으로의 변환 작업은 복잡한 도메인 차이를 다루어야 하므로, 결과물의 품질을 여러 지표를 통해 다양하게 평가하는 것이 중요하다. 본 연구에서는 변환된 이미지를 정량적으로 평가하기 위해 흔히 사용되는 픽셀 신호 대 잡음비(Peak Signal-to-Noise Ratio, PSNR)와 프레셋 인셉션 거리(Frechet Inception Distance, FID)를 활용하였다.

2.1 픽셀 신호 대 잡음비 (PSNR)

$$PSNR = 10 \log_{10} \left(\frac{R^2}{MSE} \right) \quad (1)$$

PSNR은 MSE에 로그 스케일을 취하여, 사람이 인지하는 화질 특성을 보정한 지표로 널리 사용된다. 식 (1)에서 R은 영상의 최대 픽셀값으로, 일반적으로 8비트 영상의 경우 R은 255를 사용한다. PSNR 값이 높을수록 변환된 영상이 원본 영상과 가깝다는 뜻이며, 주로 화질 손실의 정도를 모니터링할 때 유의미하게 활용된다. MSE와 달리 PSNR은 결과값이 데시벨(dB) 단위이므로, 다양한 범위의 영상 품질을 한눈에 비교하기에 편리하다.

2.2 프레셋 인셉션 거리 (FID)

$$FID(x, g) = \|\mu_x - \mu_g\|_2^2 + \text{Tr}(\Sigma_x + \Sigma_g - 2(\Sigma_x \Sigma_g)^{\frac{1}{2}}) \quad (2)$$

FID는 GAN 계열의 연구에서 탄생한 지표로, 변환된 이미지들이 원본 이미지와 얼마나 유사한 분포를 형성하는지를 측정한다. 식 (2)에서 x와 g는 각각 원본과 변환된 이미지를 사전 학습된 분류 모델 Inception-v3을 통해 추출한 특징 벡터 세트로, μ_x , μ_g 는 해당 벡터들의 평균, Σ_x , Σ_g 는 공분산 행렬을 나타낸다. 두 특징 분포가 비슷할수록 FID 값이 작아지며, 이는 변환된 이미지가 원본 이미지의 특징을 더 정확히 학습하고 있음을 의미한다. 단순 픽셀 단위 비교에서 벗어나, 분포적·특징적 유사도를 고려한다는 점에서 실험적 활용 가치가 높다.

PSNR은 기본적으로 픽셀 레벨에서 오차를 측정하지만, FID는 딥러닝 모델을 통해 추출한 고차원 특징 공간에서 분포 간 거리를 측정한다는 점에서 상호 보완적인 지표로 간주된다. 따라서 실제 SAR-광학 변환 작업의 성능 평가 시에는, 다각적인 관점에서 이미지 품질을 파악하기 위해 두 지표 모두를 종합적으로 분석하는 것이 권장된다.

3. 실험 세부 사항

실험은 629GiB의 메모리와 AMD EPYC 7642 CPU를 사용하고 있는 서버에서 NVIDIA A6000 GPU 하나를 사용하여 진행되었다. 실험 코드는 [6]의 확산 모델의 데이터 입력단계에서 전처리 및 증대부분을 수정하였다. 조건부 확산 모델에서 사용하는 파라미터로 timestep resampling, diffusion steps, noise schedule을 설정하였다. Timestep resampling은 검증 시 확산 과정에서의 시간 단계 수를 조정하는 기법으로, 250으로 설정하여 원래의 단계 수를 250개로 줄였다. Diffusion steps는 노이즈를 점진적으로 추가하고 제거하는 단계의 총 수를 나타내며, 2,000단계로 설정하여 높은 품질의 이미지 생성을 위해 세밀한 노이즈 제어가 가능하게 하였다. noise schedule은 확산 과정 중 노이즈를 추가하는 비율을 시간에 따라 변화시키는 일정으로, 'linear' 스케줄을 사용하여 노이즈를 일정한 비율로 추가하고 제거하도록 하였다. 학습은 총 batch size 12로 100,000 iterations 동안 진행되었으며, 학습률은 처음 1,000 iterations 동안 0.0001까지 선형적으로 증가한 후, 100,000 iterations까지 Cosine으로 감소하는 스케줄러를 적용하였고 weight

decay는 0.001로 설정하였다.

4. 실험 결과

본 연구에서는 총 1,024장의 검증 이미지를 활용하여 여러 전처리와 증대 기법이 SAR-광학 변환 성능에 미치는 영향을 정량적, 정성적으로 비교하였다. 구체적으로, [표 1]과 [표 2]에서는 전처리·증대 방법을 Baseline과 비교하는 결과를, [표 3]에서는 CycleGAN, 조건부 확산 모델(Baseline), 그리고 전처리·증대를 적용한 제안 기법 간의 성능 차이를 각각 살펴본다.

표 1. 데이터 전처리에 따른 실험 결과 비교

Table 1. Comparative experimental results according to data pre-processing methods

Evaluation Metrics	Baseline	Stretching	$\pm 0.2\%$ scaling + stretching
PSNR (\uparrow)	11.9302	11.9612 (+0.26%)	12.1849 (+2.13%)
FID (\downarrow)	110.2133	105.9517 (-3.87%)	102.6817 (-6.83%)

[표 1]에서는 Baseline과 비교하여 데이터 전처리 방법에 따른 비교 실험 결과를 나타내고 있다. Stretching 기법 적용 시, SAR 영상의 원시 픽셀값 범위를 0~255로 확장함으로써 명암 대비가 향상되고, 스펙클 노이즈가 상대적으로 완화되는 경향이 있다. 이를 통해 FID가 -3.87% 개선되는 모습을 보였다. 이는 단순 Baseline보다 더 균일한 픽셀 분포를 모델이 학습하게 하여, 이미지 재현 과정에서 픽셀 간 오차가 감소했음을 시사한다. $\pm 0.2\%$ Scaling + Stretching 기법은 극단값을 제거한 뒤 명암 범위를 확장하는 방식으로, FID는 -6.83% 더 크게 감소하였다. 즉, 분포 상 극단값이 줄어들고 동시에 전체적인 대비 향상이 함께 이루어져 조건부 확산 모델의 학습 안정성을 높인 것으로 볼 수 있다.

[표 2]에서는 Baseline과 비교하여 데이터 증대 방법에 따른 비교 실험 결과를 나타내고 있다. Horizontal Flip은 이미지의 좌·우를 반전하는 방법으로, 50% 확률로 적용되어 다양한 시야각을 모델이 학습하도록 돕는다. 결과적으로 PSNR이 소폭 감소하였지만 FID는 Baseline 대비 소폭 낮아지고 SAR 영상 특유의 구조적 특성을 좀 더 폭넓게

표 2. 데이터 증대에 따른 실험 결과 비교

Table 2. Comparative experimental results according to data augmentation methods

Evaluation Metrics	Baseline	Horizontal Flip	Rotate 90
PSNR (\uparrow)	11.9302	11.8685 (-0.52%)	11.9595 (+0.25%)
FID (\downarrow)	110.2133	109.2459 (-0.88%)	109.1778 (-0.94%)

학습할 수 있게 된 것으로 평가된다. Rotate 90(90도·180도·270도 회전) 또한 SAR 영상에서의 시점 변화를 가상으로 만들어냄으로써 회전에 대한 불변성(rotation invariance)을 모델에 부여한다. 이로 인해 FID가 -0.94%로 개선되었으며, 이는 다양한 회전 각도에서도 SAR-광학 변환 맵핑을 보다 정확하게 학습했음을 의미한다.

표 3. 제안하는 방법과 조건부 확산 모델, CycleGAN과의 실험 결과 비교
Table 3. Comparison of experimental results among CycleGAN, conditional diffusion model, and the proposed method

Evaluation Metrics	CycleGAN ^[3]	Conditional Diffusion ^[6]	Purposed Method
PSNR (\uparrow)	14.5706	11.9302	12.1849
FID (\downarrow)	188.5958	110.2133	102.6817

[표 3]에서 CycleGAN과 데이터 전처리 및 증대를 가하지 않은 기본 모델, 제안 기법과의 비교를 확인할 수 있다. 여기서 제안 기법은 가장 높은 성능을 기록한 $\pm 0.2\%$ scaling + stretching 전처리와 Rotate 90 데이터 증대를 적용한 모델이다. CycleGAN은 PSNR이 14.5706으로 상대적으로 높고 낮은 수치를 보이지만, FID가 188.5958로 다소 큰 편이다. 이는 픽셀 레벨 복원도(PSNR)에서는 제법 우수한 편일 수 있으나, 생성 이미지의 전체 분포적 유사성은 떨어진다는 해석이 가능하다. 조건부 확산 모델(Baseline)은 CycleGAN에 비해 PSNR(11.9302) 측면에서 소폭 열세를 보이지만, FID(110.2133)가 크게 향상되었다. 즉, 객체나 장면의 전반적인 분포·특징 정보를 보다 사람이 느끼기 자연스러운 형태로 생성하는 데 강점을 보인다. 제안 기법($\pm 0.2\%$ scaling + stretching 전처리를 적용한 조건부 확산 모델)은 FID가 102.6817로 가장 낮게 측정되어, 전체 분

포적 유사도가 크게 개선되었다. 이는 전처리로 인해 노이즈나 극단값이 줄어든 상태에서 모델이 안정적으로 학습함으로써, CycleGAN보다 낮은 FID를 달성하는 결과로 이어졌다. PSNR 역시 12.1849로 Baseline보다 소폭 상승했으며, 이는 전처리·증대가 정밀한 화질 표현에도 기여했음을 의미한다.

[그림 6]을 보면 주어진 SAR 이미지를 조건으로 CycleGAN, Baseline 및 제안 기법의 모델로 변환된 광학 이미지를 확인할 수 있다. CycleGAN의 경우 명암이나 색 표현은 선명할 수 있으나, 건물이나 지형 경계선이 흐릿하게 나타나는 경우가 있다. 조건부 확산 모델(Baseline)은 전반적으로 객체들이 자연스럽게 연결되어 있지만, 일부 지역에서 윤곽이 덜 표현되는 한계가 나타난다. 반면, 제안 기법은 전처리·증대가 결합되면서 건물의 형태적 디테일을 좀 더 선명하게 복원하고, SAR 영상 특유의 스펙클 노이즈가 효과적으로 제거된 양상을 보인다. 결과적으로, 사람이 바라보았을 때 실제 광학 영상에 가까운 선명도와 경계 표현을 확보했다는 점이 특징적이다.

종합적으로, 전처리 기법(특히 $\pm 0.2\%$ scaling + stretching)과 증대 기법(Rotate 90)을 적용한 조건부 확산 모델은 SAR 영상에서 광학 영상으로의 실감도 높은 변환을 달성하는 데 매우 효과적인 것으로 판단된다. FID가 크게 개선되었다는 점은 곧 학습된 이미지 분포의 질적 향상을 의미하므로, 실제 응용 분야에서 광학 영상 대체재로 활용될 가능성을 높인다. PSNR 수치적 상대적 차이는 SAR 영상의

화질 특성과 전처리 여부에 따라 변동이 있지만, 전처리를 적절히 수행함으로써 오차 최소화와 분포 유사성 향상을 동시에 달성할 수 있다는 점이 본 연구의 핵심 성과라고 할 수 있다.

V. 결 론

본 논문에서는 조건부 확산 모델을 이용하여 SAR 영상을 광학 영상으로 변환할 때, 효과적인 데이터 증대를 탐색 및 전처리 방법을 제안하였다. 제안 기법을 통해 영상의 품질을 더욱 향상된 것을 실험 결과를 통해 입증하였다. 이러한 연구를 통해 SAR 영상의 단점을 극복하여 활용도 높은 광학 영상으로 변환할 수 있게 한다. 이는 SAR 영상의 활용 범위를 확대하여 재난 관리, 환경 모니터링, 도시 계획 등 다양한 분야에서 응용 가능성을 높일 수 있을 것으로 기대된다. 본 연구를 통해 조건부 확산 모델 기반의 SAR 영상 변환 기술이 인공위성 영상 분야의 발전에 기여할 수 있을 것으로 기대된다.

참 고 문 헌 (References)

- [1] Ho, Jonathan, Ajay Jain, and Pieter Abbeel, "Denoising diffusion probabilistic models," *Advances in neural information processing systems*, vol. 33, pp. 6840-6851, 2020.

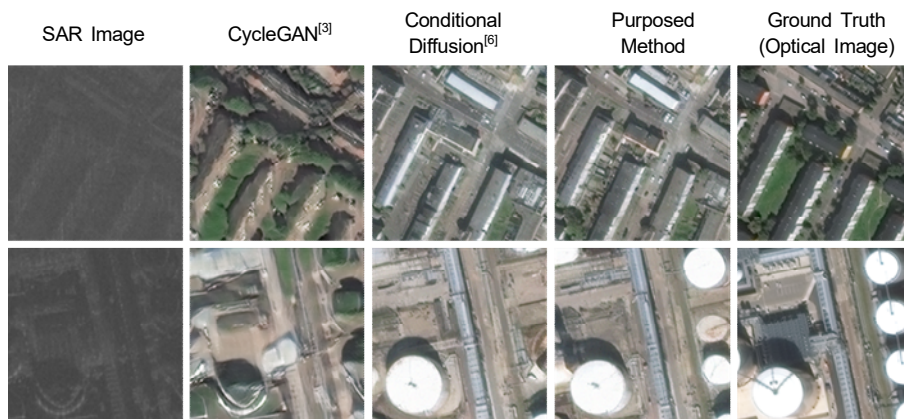


그림 6. SAR 이미지 변환의 정성적 결과

Fig. 6. Qualitative results of SAR image translation

- doi: <https://doi.org/10.48550/arXiv.2006.11239>
- [2] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, Alexei A. Efros. "Image-To-Image Translation With Conditional Adversarial Networks," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1125-1134, 2017.
doi: <https://doi.org/10.1109/cvpr.2017.632>
- [3] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A. Efros. "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4401-4410, 2019.
doi: <https://doi.org/10.1109/iccv.2017.244>
- [4] Tero Karras, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, Timo Aila, "Analyzing and Improving the Image Quality of StyleGAN," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 8110-8119, 2020.
doi: <https://doi.org/10.1109/cvpr42600.2020.00813>
- [5] Y. Li, R. Fu, X. Meng, W. Jin and F. Shao, "A SAR-to-Optical Image Translation Method Based on Conditional Generation Adversarial Network (cGAN)," *IEEE Access*, vol. 8, pp. 60338-60343, 2020.
doi: <https://doi.org/10.1109/access.2020.2977103>
- [6] X. Bai, X. Pu and F. Xu, "Conditional Diffusion for SAR to Optical Image Translation," *IEEE Geoscience and Remote Sensing Letters*, vol. 21, pp. 1-5, 2024.
doi: <https://doi.org/10.1109/lgrs.2023.3337143>
- [7] Z. Guo, Z. Zhang, Q. Cai, J. Liu, Y. Fan and S. Mei, "MS-GAN: Learn to Memorize Scene for Unpaired SAR-to-Optical Image Translation," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 17, pp. 11467-11484, 2024.
doi: <https://doi.org/10.1109/jstars.2024.3411691>
- [8] J. Lee, H. Cho, D. Seo, H. -H. Kim, J. Jeong and M. Kim, "CFCA-SET: Coarse-to-Fine Context-Aware SAR-to-EO Translation With Auxiliary Learning of SAR-to-NIR Translation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1-18, 2023.
doi: <https://doi.org/10.1109/tgrs.2023.3318980>
- [9] Shermeyer, Jacob, et al. "SpaceNet 6: Multi-sensor all weather mapping dataset," *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*. 2020.
doi: <https://doi.org/10.1109/cvprw50498.2020.00106>

저 자 소 개

서 형 원



- 2025년 2월 : 국립한밭대학교 컴퓨터공학과 공학사
- 2025년 3월 ~ 현재 : 국립한밭대학교 컴퓨터공학과 석사과정
- ORCID : <https://orcid.org/0009-0008-1685-5915>
- 주관심분야 : 컴퓨터비전, 항공/위성 영상처리

이 지 상



- 2025년 2월 : 국립한밭대학교 컴퓨터공학과 공학사
- 2025년 3월 ~ 현재 : 국립한밭대학교 컴퓨터공학과 석사과정
- ORCID : <https://orcid.org/0009-0001-3249-1997>
- 주관심분야 : 컴퓨터비전, 객체 탐지

저 자 소 개



박 우 진

- 2021년 3월 ~ 현재 : 국립한밭대학교 컴퓨터공학과 공학사과정
- ORCID : <https://orcid.org/0009-0006-6361-1476>
- 주관심분야 : 컴퓨터비전, 영상 생성, 영상 변환 등



최 해 철

- 1997년 : 경북대학교 전자공학과 공학사
- 1999년 : 한국과학기술원 전기공학과 공학석사
- 2004년 : 한국과학기술원 전기공학과 공학박사
- 2010년 ~ 현재 : 한밭대학교 지능미디어공학과 교수
- ORCID : <https://orcid.org/0000-0002-7594-0828>
- 주관심분야 : 영상처리, 비디오 코딩, 컴퓨터 비전



김 규 현

- 1989년 2월 : 한양대학교 전자공학과 공학사
- 1992년 9월 : 영국 University of Newcastle upon Tyne 전기전자공학과 공학석사
- 1996년 7월 : 영국 University of Newcastle upon Tyne 전기전자공학과 공학박사
- 1996년 ~ 1997년 : 영국 University of Sheffield, Research Fellow
- 1997년 ~ 2006년 : 한국전자통신연구원 대화형미디어연구팀장
- 2006년 ~ 현재 : 경희대학교 전자정보대학 교수
- ORCID : <https://orcid.org/0000-0003-1553-936X>
- 주관심분야 : 디지털 방송, 영상처리, 멀티미디어 통신, 디지털 대화형 방송



장 한 열

- 2012년 2월 : 아주대학교 정보컴퓨터공학부 공학사
- 2014년 2월 : 한국과학기술원 전산학과 공학석사
- 2018년 2월 : 한국과학기술원 전산학부 공학박사
- 2018년 4월 ~ 2018년 10월 : 네이버 클로바 인공지능 연구원
- 2018년 10월 ~ 2020년 8월 : 국가보안기술연구소 선임연구원
- 2020년 9월 ~ 현재 : 국립한밭대학교 컴퓨터공학과 교수
- ORCID : <https://orcid.org/0000-0002-7253-7646>
- 주관심분야 : 항공/위성 영상처리, 멀티미디어 보안, 컴퓨터비전, 멀티모달