

합성개구레이다 영상에서 광학 영상으로의 이미지 변환을 위한 데이터 전처리 및 증대 기법

서형원, 이지상, 박천음, 장한얼†

한밭대학교 컴퓨터공학과

{20191735, 20191785}@edu.hanbat.ac.kr, {parkce, hejang}@hanbat.ac.kr

Data Preprocessing and Augmentation Methods for SAR to Optical Image Translation

Hyoungwon Seo, Jisang Lee, Cheoneum Park, Haneol Jang†

Dept. of Computer Engineering, Hanbat National University

요약

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 SAR 영상을 광학 영상으로 변환할 때 GAN 모델 대신 조건부 확산 모델을 사용할 때, 데이터 전처리 및 증대 방법을 제안한다. 수평 반전 및 90도 회전과 같은 데이터 증강 기법과 스트레칭 및 $\pm 0.2\%$ 스케일링과 같은 전처리 기법을 적용함으로써 변환된 영상의 품질을 더욱 향상하여 실제 광학 영상과 더 유사한 영상을 생성할 수 있게 된다. 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.

1. 서론

오늘날 인공위성 기술의 발달과 함께 위성영상에 대한 수요가 급격히 증가하고 있다. 특히, 합성개구레이다(Synthetic Aperture Radar, SAR)는 마이크로파를 이용하여 대상물에 반사된 신호를 수신하여 영상을 생성하기 때문에 기상 조건 및 야간에 대한 제약 없이 영상을 획득할 수 있어 광학 영상의 한계를 극복할 수 있는 기술로 주목받고 있다. 하지만, SAR 영상은 스펙클 노이즈(Speckle Noise)라는 마이크로파의 간섭에 의해 발생하는 잡음이 존재하며, 이는 영상의 품질을 저하하고 영상 해석을 어렵게 만드는 요인이 된다. 이러한 SAR 영상의 단점을 해결하고, 보다 활용도 높은 영상을 얻기 위해 SAR 영상을 광학 영상과 유사한 형태로 변환하는 기술에 관한 연구가 진행되고 있다. 기존의 연구에서는 주로 Generative Adversarial Network(GAN) 모델을 이용하여 SAR 영상을 광학 영상으로 변환하는 방법이 많이 사용되었지만, 확산 모델의 등장으로 조건부 확산 모델을 이용한 방법이 새롭게 사용되고있다. 본 논문에서는 조건부 확산 모델(Conditional Diffusion Model)에서 SAR 영상을 광학 영상으로 변환할 때, 더 실제 광학 영상처럼 변환하기 위한 데이터 전처리 및 증대에 대한 실험을 진행하였다. 이를 통해 SAR 영상의 스펙클 노이즈를 효과적으로 제거하고, 보다 해석하기 쉬운 광학 영상과 유사한 형태의 영상을 생성하여 다양한 분야에 활용될 수 있기를 기대한다.

2. 관련 연구

2.1 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 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.

2.2 이미지에서 이미지로의 변환

이미지에서 이미지로의 변환 연구는 SAR 영상을 광학 영상으로 변환하는 작업처럼 주어진 입력 데이터와 관련된 이미지로 출력하는 작업이다. 지금까지는 딥러닝 접근 방식에서 GAN을 사용하여 상당한 발전을 보였다. 특히, CycleGAN[2]은 페어링 되지 않은 데이터셋에서 이미지 간 변환을 수행할 수 있는 GAN 기반 모델로, 도메인 간 변환에 널리 사용되었다. 최근에는 확산 모델

* 본 연구는 2024년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음 (2022-0-01068)

의 등장으로 이미지 변환 분야에 새로운 접근 방식이 가능해져 확산 모델을 적용하는 연구가 활발히 진행되고 있다. SAR 영상을 광학 영상으로 변환하는 연구는 스펙클 노이즈같은 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[3], 최근 조건부 확산 모델의 등장으로 SAR 영상 변환 분야에서도 새로운 접근 방식이 가능하게 되어 이와 관련한 연구도 진행되었다[4].

3. 제안 기법

본 연구에서는 조건부 확산 모델[4]을 기반으로 SAR 영상을 광학 영상으로 변환할 때 더 효과적인 데이터 증대법 및 전처리 방법을 제안한다. SAR 영상은 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 90도, 180도, 270도로 이미지를 회전시켜 모델이 회전 불변성을 학습할 수 있게 한다.

SAR 이미지를 모델의 입력으로 사용하기 위한 전처리 방법으로 Stretching과 $\pm 0.2\%$ Scaling을 적용하였다. [그림 1]은 전처리 방법에 따른 픽셀값 분포를 보여준다. 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. 원본 SAR 데이터는 uint16 형식이지만, 딥러닝 모델은 uint8 형식을 요구하므로 변환 과정에서 정보 손실이 발생할 수 있다. 본 연구에서는 Stretching과 $\pm 0.2\%$ Scaling을 적용하여 정보 손실을 최소화하고자 하였다. 이러한 데이터 증대 및 전처리 기법을 통해 변환된 영상의 품질을 향상시키고, 모델의 일반화 능력을 증가시킬 수 있다.

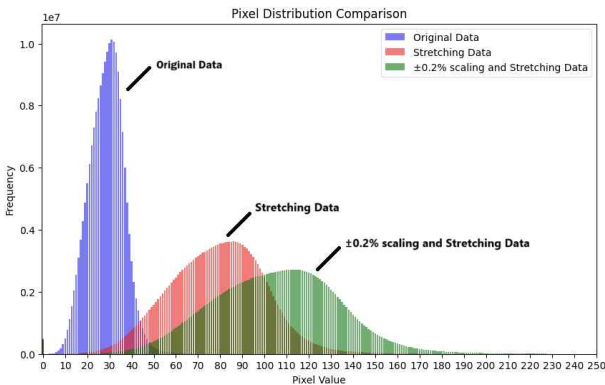


그림 1. 학습 데이터 중 무작위로 선택된 1,000장의 데이터 전처리 방법에 따른 픽셀값 분포

4. 실험

4.1 데이터셋

본 연구에서는 SAR 영상을 광학 영상으로 변환하기 위해 SpaceNet6 데이터셋[5]을 사용하였다. SpaceNet6은 광학 영상(Optical)과 SAR 영상이 쌍을 이루는 위성 영상 데이터셋으로, 다양한 지역에 대한 고해상도 이미지를 제공한다. 위성 영상의 특성상 이미지의 가장자리에는 검은색 배경이 많이 존재하는데, 검은색 배경이 전체 이미지의 5%가 넘는 경우 학습에 불필요한 정보로 판단되어 영상에서 제거하였다. 또한, 모델 학습을 위해 이미지를 256x256 크기로 잘라내어 사용하였다. 최종적으로 학습에 사용된 이미지는 총 13,627장이며, 검증을 위해 1,024장의 이미지를 별도로 사용하였다. 이렇게 전처리된 SpaceNet6 데이터셋을 통해 조건부 확산 모델을 학습하고, SAR 영상에서 광학 영상으로의 변환 성능을 평가하였다.

4.2 평가 지표

SAR에서 광학 영상으로의 이미지 변환 작업에서 사용하는 주요 평가 지표로는 PSNR, MSE, FID가 있다. PSNR은 원본과 변환된 이미지 간의 차이를 측정하며, 값이 높을수록 품질이 좋을 것을 의미한다. MSE는 평균 제곱 오차로, 낮을수록 원본과 변환된 이미지 간의 차이가 작음을 의미하고 FID는 생성된 이미지와 실제 이미지 간의 특징 분포 유사성을 측정하며, 낮을수록 더 좋은 품질을 나타낸다.

$$MSE = \frac{\sum_{m,n} [I_1(m,n) - I_2(m,n)]^2}{M * N} \quad PSNR = 10 \log_{10} \left(\frac{R^2}{MSE} \right)$$

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

수식 1. 평가 지표들에 대한 수식

4.3 실험 세팅

실험은 629GiB의 메모리와 AMD EPYC 7642 CPU를 사용하고 있는 서버에서 NVIDIA A6000 GPU 하나를 사용하여 진행되었다. 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.

4.4 실험 결과

검증을 위한 1,024장의 이미지를 평가 지표를 통해 데이터 전처리와 증대에 대한 실험 결과를 [표 1] 및 [표 2]에서 확인할

수 있으며, [표 3]에서 제안기법의 결과를 확인할 수 있다.

표 1. 데이터 전처리 방법에 따른 비교 실험 결과

평가 지표	Baseline	stretching	$\pm 0.2\%$ scaling + stretching
PSNR (↑)	11.9302	11.9612 (+0.26 %)	12.1849 (+2.13 %)
MSE (↓)	0.0826	0.0816 (-1.21 %)	0.0767 (-7.14 %)
FID (↓)	110.2133	105.9517 (-3.87 %)	102.6817 (-6.83 %)

[표 1]에서는 Baseline과 비교하여 데이터 전처리 방법에 따른 비교 실험 결과를 나타내고 있다. 가장 주목할 만한 성능 지표인 MSE와 FID를 살펴보면, stretching 및 $\pm 0.2\%$ scaling + stretching 방법을 적용하였을 때 각각 MSE가 1.21%, 7.14% 감소하고, FID가 3.87%, 6.83% 감소하여 더 나은 성능을 보여준다. 이는 전처리 방법을 통해 명암 대비를 향상시키고 노이즈를 제거하는 것이 조건부 확산 모델을 사용하여 SAR 영상을 광학 영상으로 변환할 때 더 효과적인임을 확인할 수 있다.

표 2. 데이터 증대법에 따른 비교 실험 결과

평가 지표	Baseline	Horizontal Flip	Rotate 90
PSNR (↑)	11.9302	11.8685 (-0.52 %)	11.9595 (+0.26 %)
MSE (↓)	0.0826	0.0825 (-0.10 %)	0.0792 (-4.12 %)
FID (↓)	110.2133	109.2459 (-0.88 %)	109.1778 (-0.94 %)

[표 2]에서는 Baseline과 비교하여 데이터 증대 방법에 따른 비교 실험 결과를 나타내고 있다. Horizon Flip 및 Rotate 90 증대 방법에서 MSE 및 FID에 대한 성능이 향상되는 것을 확인할 수 있다.

표 3. CycleGAN, 조건부 확산 모델, 제안기법 실험 결과 비교

평가 지표	CycleGAN [2]	Conditional Diffusion [4]	Purposed Method
PSNR (↑)	14.5706	11.9302	12.1849
MSE (↓)	0.0456	0.0826	0.0767
FID (↓)	188.5958	110.2133	102.6817

[표 3]에서 CycleGAN과 데이터 전처리 및 증대를 가하지 않은 기본 모델, 제안기법과의 비교를 확인할 수 있다. 여기서 제안기법은 가장 높은 성능을 기록한 $\pm 0.2\%$ scaling + stretching 전처리를 적용한 모델이다. 조건부 확산 모델이 CycleGAN에 비해 PSNR 및 MSE에 대한 성능은 하락하지만, FID에서 높은 점수 향상을 보였다. 또한 [그림3]에서 정성적 결과를 확인하였을 때, 객체의 윤곽을 잘 살리지 못하는 모습을 확인할 수 있으며 제안기법은 기본 모델과 비교하면 건물의 형태를 잘 표현하는 것을 확인할 수 있다.

5. 결론

본 논문에서는 조건부 확산 모델을 이용하여 SAR 영상을 광학 영상으로 변환할 때, 데이터 증대 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,

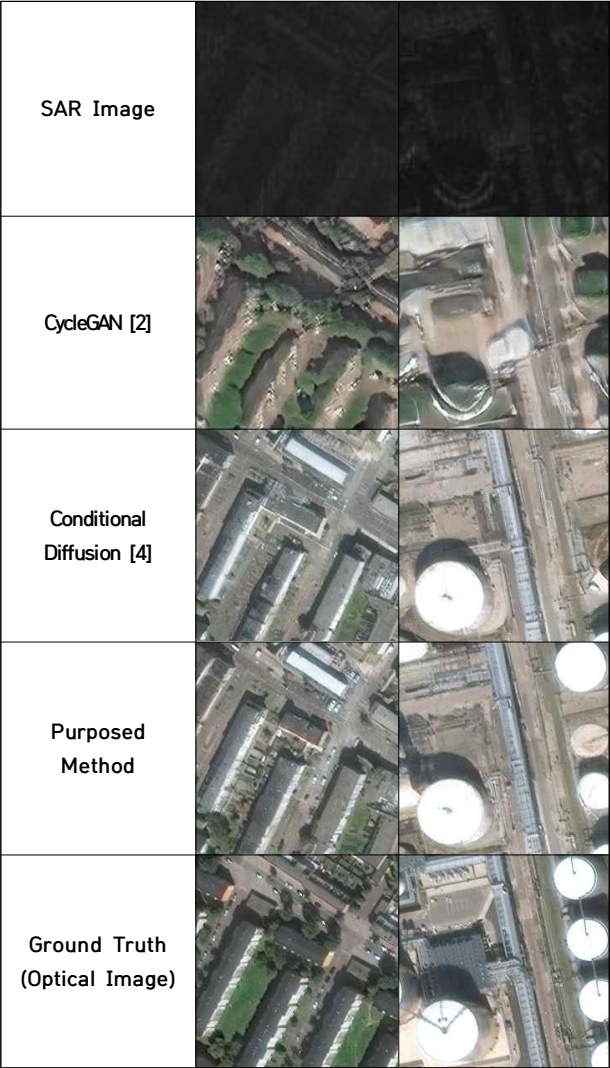


그림 3. SAR 이미지 변환 정성적 결과

도시 계획 등 다양한 분야에서 응용 가능성을 높일 수 있을 것으로 기대된다. 본 연구를 통해 조건부 확산 모델 기반의 SAR 영상 변환 기술이 인공지능 영상 분야의 발전에 기여할 수 있을 것으로 기대된다.

6. 참고문헌

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

[2] Jun-Yan Zhu*, Taesung Park*, Phillip Isola, and Alexei A. Efros. "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks", in IEEE International Conference on Computer Vision (ICCV), 2017.

[3] 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)," in IEEE Access, vol. 8, pp. 60338-60343, 2020.

[4] X. Bai, X. Pu and F. Xu, "Conditional Diffusion for SAR to Optical Image Translation," in IEEE Geoscience and Remote Sensing Letters, vol. 21, pp. 1-5, 2024.

[5] 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.