

# SSIM 기반 cGAN 을 이용한 SAR 영상으로부터 광학 영상으로의 변환 기법

유초시, 황지언, 신요안\*

승실대학교 전자정보공학부

(\*교신저자)

cysu@soongsil.ac.kr, hwangjieon@soongsil.ac.kr, yashin@ssu.ac.kr

## 요 약

합성 개구 레이더 (Synthetic Aperture Radar; SAR) 영상과 광학 (Optical) 영상은 서로 다른 원격 탐지 방식이다. SAR 는 전파를 사용하므로 광학 센서와 비교했을 때 작업 시간과 기후 조건의 영향으로부터 자유롭게 관측이 가능한 장점이 있다. 본 논문에서는 cGAN (Conditional Generative Adversarial Network) 딥러닝 알고리즘을 이용하여 SAR 영상을 광학 영상으로 변환하는 기법을 제안한다. 제안하는 cGAN 에서는 생성기 (Generator) 네트워크에서 가상의 광학 영상을 생성하고 판별기 (Discriminator) 네트워크는 샘플 영상이 실재인지 가상의 광학 영상인지의 여부를 판별한다. 본 논문에서는 풍부한 구조적 정보를 유지하기 위해 cGAN 의 기본 손실 함수에 구조적 유사 지수를 추가하고, 학습은 가상 광학 영상과 실제 광학 영상을 구별하기 어려울 때까지 계속된다. 실험 결과, cGAN 을 이용한 효과적인 SAR 영상과 광학 영상 변환과 융합 가능성을 확인하였다.

## I. 서 론

원격 탐지 (Remote Sensing)가 발달함에 따라 점점 더 많은 위성 영상 자원이 다른 방식으로 획득되고 있으며, 지도 제작, 에너지 개발, 항법, 물체 검출 등 다양한 분야에서 중요하게 사용되고 있다. 그중 광학 영상 (Optical Image)을 통한 원격 탐지는 기후와 시간의 영향을 받는 경우가 많다. 반면, 합성 개구 레이더 (Synthetic Aperture Radar; SAR) 영상 원격 탐지는 전파를 이용하므로 작업 시간과 기후 조건의 영향을 받지 않으며 더 많은 정보를 얻을 수 있으나, 광학 영상에 비해 직관적으로 이해하기는 어렵다. 따라서 SAR 영상과 광학 영상을 융합하여 분석하는 연구는 최근 상당히 중요한 과제가 되었다.

최근 많은 딥러닝을 기반으로 SAR 영상과 광학 영상 전환에 활용될 수 있는 방법들이 제시되고 있다. 본 논문에서는 구조적 정보를 유지하면서 영상 간의 연결을 강화하기 위해, 향상된 조건부 생성 적대적 네트워크 (Conditional Generative Adversarial Network; cGAN)를 사용하여 SAR 영상을 광학 영상으로 변환 (SAR-to-Optical; S2O) 하는 방안을 제안한다.

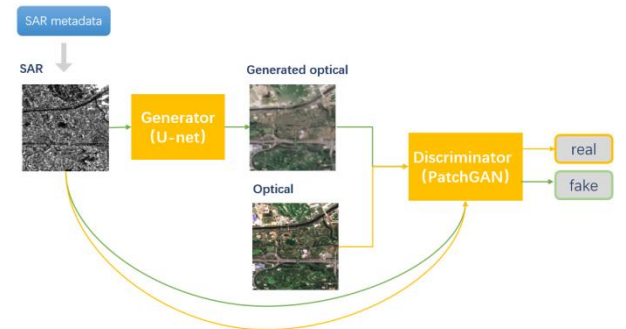
## II. SSIM 기반 cGAN을 이용한 S2O 영상 변환

cGAN 은 기본 GAN[1]에 조건을 설정하여 변형된 구조로서 영상과 영상 사이의 전환에 적용될 수 있다. 본 논문에서는 참고문헌 [2]에서 제안된 cGAN 모델을 사용하며, 생성기 (Generator) 네트워크인  $G$ 와 판별기 (Discriminator) 네트워크인  $D$ 가 포함된 구조로서, 다음의 목적 함수에 대한 최소-최대 문제를 해결하기 위해 번갈아 가며 최적화된다.

$$\begin{aligned} G^* &= \underset{G}{\operatorname{argmin}} \max_D \mathcal{L}_{cGAN}(G, D) + \lambda \mathcal{L}_{L1}(G) \\ &= \mathbb{E}_{x, y} [\log(D(x, y))] \\ &\quad + \mathbb{E}_{x, z} [\log(1 - D(x, G(x, z)))] \\ &\quad + \lambda \mathbb{E}_{x, y, z} [\|y - G(x, z)\|_1] \end{aligned} \quad (1)$$

$x$ : input image  $y$ : output image  $z$ : random noise vector

cGAN 은 관찰된 영상  $x$  및 랜덤 노이즈 벡터  $z$  에 대해  $G: \{x, z\} \rightarrow y$  의 매핑을 학습한다. 그림 1 은 cGAN 기반의 S2O 영상 변환에 대한 초기 네트워크를 보여준다. 여기서 생성기와 판별기가 번갈아 훈련되는데, 생성기  $G$  는 가상의 광학 영상을 생성하고 판별기  $D$  는 생성된 광학 영상을 실제 광학 영상과 비교하여 진위를 구별한다. 본 논문에서 제안된 향상된 cGAN 모델의 생성기와 판별기는 pix2pix 의 네트워크 구조 (U-net 및 PatchGAN)와 동일하다[2]. 그림 2 은 본 논문에서 고려된 cGAN 의 상세 구조이다.



$$\mathcal{L}_G = \mathbb{E}[\log D(sar, G(sar))] \quad \mathcal{L}_D = \mathbb{E}[\log D(sar, op)] + \mathbb{E}[\log(1 - D(sar, G(sar)))]$$

Figure 1. cGAN model for S2O image conversion

본 논문에서는 보다 풍부한 구조적 정보를 유지하면서 상세 정보를 추출하고 광학 영상을 생성하기 위해,  $\mathcal{L}_{L1}$  손실을 대신에 구조적 유사 지수 (Structural Similarity Index Measure; SSIM)[3]의 사용을 제안한다. 이 경우 제안 방식의 최종 목표 함수는 식 (2)와 같이 표현된다. 일반적으로 SSIM 값이 클수록 두 영상 사이의 유사성이 높아지며, 두 영상 사이의 SSIM 은 식 (3)와 같이 정의된다.

$$G^* = \underset{G}{\operatorname{argmin}} \max_D \mathcal{L}_{cGAN}(G, D) + \lambda \mathcal{L}_{SSIM}(G) \quad (2)$$

( $\lambda$ : SSIM 손실의 가중치)

$$SSIM(x, y) = L(x, y) * C(x, y) * S(x, y) \quad (3)$$

( $L$ : Luminance,  $C$ : Contrast,  $S$ : Structure)

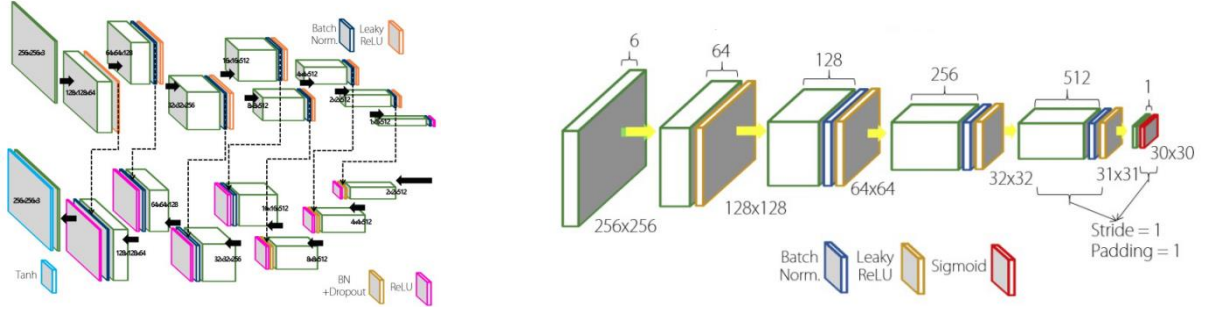


Figure 2. Structure of the generator and discriminator networks in the cGAN

본 논문에서는 참고문헌 [4]에서 사용된 중국 9 개 도시의 SAR-Optical 매칭 영상 데이터 (WHU-SEN-City)를 이용하여 제안 기법의 성능을 평가하였다. 해당 데이터셋을 전처리 하고 SAR 영상과 광학 영상 쌍을 이루었으며, 그 중 4000 개 SAR-Optical 패치 쌍을 훈련 과정에 사용했고, 200 개 패치 쌍을 테스트에 사용하였다. Python3.8 과 TensorFlow 를 통해 제안된 cGAN 네트워크를 구현하였고, 생성기 입력과 출력 영상 크기는  $256 \times 256 \times 3$ , 판별기 입력 크기는  $256 \times 256 \times 6$  이다. 훈련 과정에서 SSIM 가중치  $\lambda$  는 100 이었고, 학습율은 0.0002 로 설정하였다.

그림 3 은 50 번째 Epoch 에서의 실험 결과로, 개선된 cGAN 을 바탕으로 S2O 변환된 영상이 세부적인 성능, 질감, 색상 정보 측면에서 실제 광학 영상과 매우 가까운 것을 확인할 수 있다.

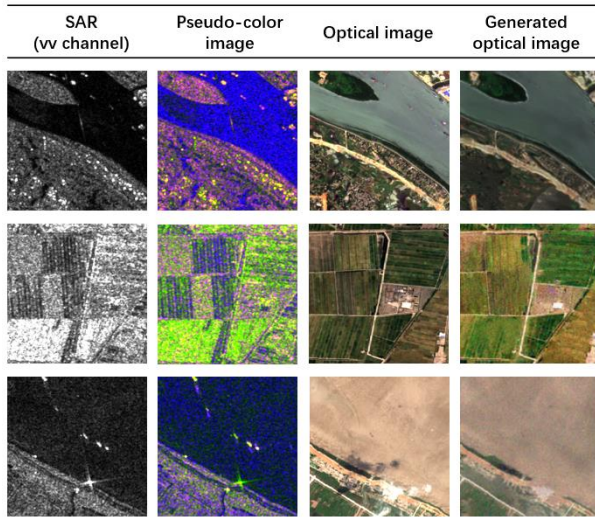


Figure 3. Examples of results (patch images from Changsha/Yinchuan/Jiujiang)



Figure 4. Result of Jiujiang's patch image in detail (red box: shadowing, yellow box: cloud)

그림 4 에서 제시하고 있는 바와 같이, 생성된 영상과 실제 광학 영상을 자세하게 대조 시 제안 기법을 통해 음영과 구름의 영향도 제거가 가능하여 분할 검사와 같은 영상의 후처리에도 도움이 될 것으로 기대된다[5].

### III. 결론

본 논문에서는 SSIM 을 손실 함수에 추가한 향상된 cGAN 알고리즘을 제안하였으며, 이를 통해 S2O 영상 변환에 유용한 구조적 정보를 보존하여 인간의 시각 시스템에 좀 더 적합한 S2O 영상 변환 결과를 얻을 수 있었다. 제안 기법을 통해 실제 광학 영상에 필요한 정보가 부족하거나 획득할 수 없는 경우 SAR 영상에 대한 가상 광학 영상을 생성할 수 있고, 이들 정보의 종합 분석을 통해 더욱 정밀한 원격 탐사에 도움을 줄 수 있을 것으로 기대된다.

### 감사의 글

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터지원사업의 연구결과로 수행되었음 (IITP-2020-2018-0-01424)

### References

- [1] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," *Proc. NIPS 2014*, pp. 2672-2680, Montreal, Canada, Dec. 2014.
- [2] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-Image-to-Image translation with conditional adversarial networks," *Proc. CVPR 2017*, pp.1125-1134, Honolulu, USA, July 2017.
- [3] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600-612, Apr. 2004.
- [4] L. Wang, X. Xu, Y. Yu, R. Yang, R. Gui, Z. Xu, and F. Pu, "SAR-to-optical image translation using supervised cycle-consistent adversarial networks," *IEEE Access*, vol. 7, pp. 129136-129149, 2019.
- [5] L. Liu and B. Lei, "Can SAR images and optical images transfer with each other?" *Proc. IGARSS 2018*, pp. 7019-7022, Valencia, Spain, July 2018.