# AI-Empowered Remote Sensing Imagery Generation

Yuanyuan Qing, Jiang Zhu, Shuxin Cao, Weixian Liu and Bihan Wen

Nanyang Technological University

## ABSTRACT

The earth observation domain involves multiple important tasks, ranging from multi-sensor co-registration, multi-temporal and multi-modality fusion, task-driven image reconstruction, data mining, etc. Recent advance in AI and ML technologies can potentially enable highly intelligent remote sensing system that automatically extracts valuable information from the large-scale data, leading to superior performance and efficiency in these tasks. Electro-optical (EO) imagery has traditionally been applied by Earth Observation Satellites of different resolutions and spectral. Besides, Synthetic Aperture Radar (SAR), as a unique form of radar that can penetrate clouds, can collect image data under all-weather conditions, and during day and night. This project aims to exploit the advanced domain-aware generative adversarial network (GAN) models for SAR-EO domain transfer and modality adaptation.

## 1. Baseline Algorithms

Generative adversarial networks (GAN) [1] have created great impact and success in various computer vision tasks, e.g., image generation [2] and image-to-image translation [3,4,5]. We have surveyed existing GAN-based style transfer works, which are closely related to SAR-EO transfer task, as there are few studies on transferring SAR-EO modalities directly. GAN-based style transfer methods can be summarized into three categories, namely self-supervised, semi-supervised, and supervised learning methods. TuiGAN [3] (self-supervised learning), CycleGAN [4] (semi-supervised learning) and Pix2Pix [5] (supervised learning) are implemented and evaluated on our EO/SAR dataset. TuiGAN does not require any training corpus, but only one reference image from the target domain. We observe that TuiGAN model can only produce reasonable results when the reference image is similar, which makes it impractical to use.

CycleGAN requires a training corpus of both EO and SAR images, which does not need to be paired or aligned. Figure 1 illustrates the network architecture and the data flow of the CycleGAN design. The two mapping functions $G$ and $F$ are expected to generate fake SAR and EO images from real EO and SAR images, respectively. Moreover, cycle-consistency is used to regularize the learned mapping functions, i.e., $G$ and $F$ are required to map the generated fake SAR and EO images back to real EO and SAR. The final objective function is:

$$L(G,F,D_X,D_Y) = L_{GAN}(G,D_Y,X,Y) + L_{GAN}(F,D_X,X,Y) + \lambda L_{cyc}(G,F) \quad (1)$$

where $X$ and $Y$ represent EO and SAR domains, $D_X$ and $D_Y$ are two discriminators to distinguish real/fake EO and SAR images, $L_{GAN}(\cdot)$ and $L_{cyc}(\cdot)$ are adversarial loss and cycle-consistency loss, and $\lambda$ is the hyper-parameter determining the relative importance of the two loss terms.

Figure 2 shows examples of SAR images generated using CycleGAN. We observe that CycleGAN provides very promising results, even without pixel-level losses and supervision at all. Comparing the transferred/real SAR images, the consistent domain style is observed, and most of the key spatial features are preserved. However, there are some structures which involve large change of height, cannot be well reconstructed using CycleGAN. It is due to the limitation of the pure data-driven approach, as certain information is less visible in EO, while SAR can better present information like height. Also, the construction of fine details in the transferred SAR images are problematic as fake structures and artifacts are observed.

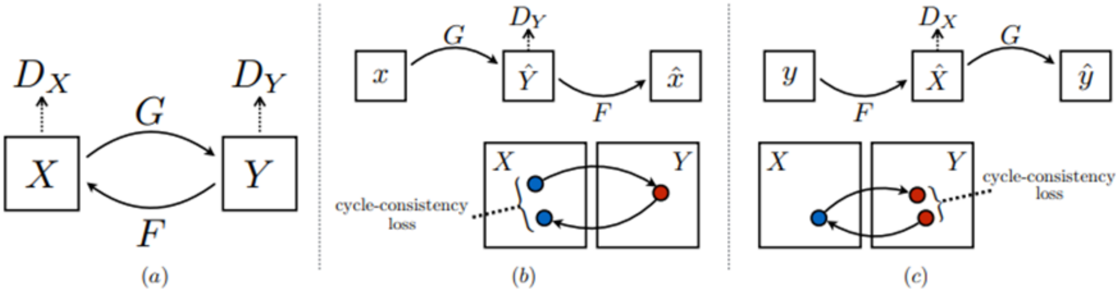
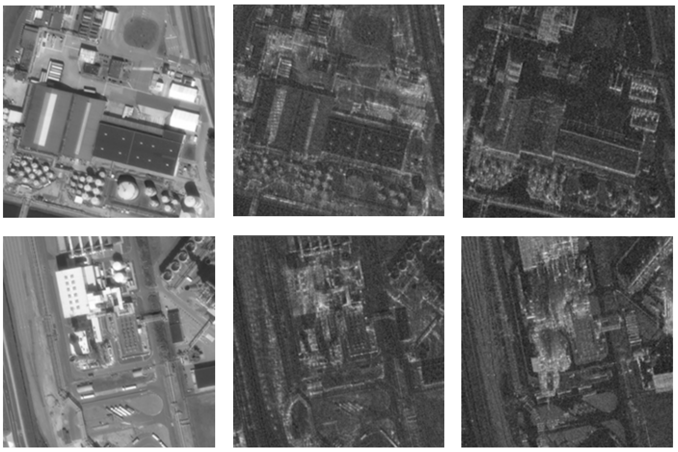


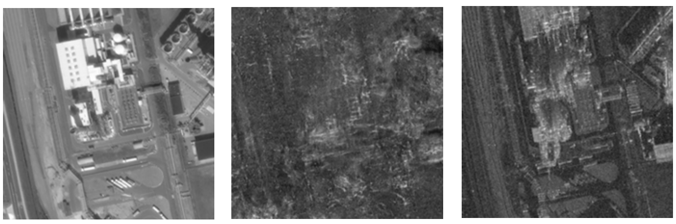

Fig 1: CycleGAN network structure, which contains two mapping functions, i.e., $G$ that maps $X$ to $Y$, and $F$ that map $Y$ back to $X$. The associated adversarial discriminators encourage $G$ to translate $X$ into indistinguishable domain of $Y$, and vice versa for $F$.



(a) Input EO image (b) Transferred SAR (c) Real SAR image

Fig. 2: The generated SAR images using CycleGAN from the input EO images.

Pix2Pix requires a training corpus of paired EO and SAR images, which must be perfectly aligned in the pixel level, e.g., there must be no pixel shift at all for each pair of EO and SAR images that are used for training. Unfortunately, due to the limitation of our EO/SAR dataset, small pixel shifts/misalignment is inevitable. We observed large area of the blurry artifacts in the generated SAR images using Pix2Pix, as shown in Figure 3.



(a) Input EO image (b) Transferred SAR (c) Real SAR image

Fig. 3: The generated SAR images using Pix2Pix from the input EO images.

## 2. Limitations of CycleGAN

According to Equation (1), pixelwise supervision is not exploited in CycleGAN. Adversarial loss and cycle-consistency loss are basically used to guarantee high-level visual similarity between real and fake images, i.e., the overall style. However, the construction of low-level structures are required in SAR-EO transfer task. Therefore, the supervision in vanilla CycleGAN is insufficient for our task.

## 3. Conclusion

We surveyed current GAN-based style transfer methods and conducted experiments on EO-SAR dataset. We have demonstrated that CycleGAN is capable of producing relatively good, transferred SAR images. However, the limitations of CycleGAN should be addressed in future work.

## References


[1] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. *Advances in neural information processing systems*, *27*.

[2]Denton, E. L., Chintala, S., & Fergus, R. (2015). Deep generative image models using a laplacian pyramid of adversarial networks. *Advances in neural information processing systems*, *28*.

[3] Lin, J., Pang, Y., Xia, Y., Chen, Z., & Luo, J. (2020, August). Tuigan: Learning versatile image-to-image translation with two unpaired images. In *European Conference on Computer Vision* (pp. 18-35). Springer, Cham.

[4] Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision* (pp. 2223-2232).

[5] Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1125-1134).