GAN-based EO-SAR Domain Transfer

Yuanyuan Qing, Jiang Zhu, Weixian Liu and Bihan Wen
Nanyang Technological University

ABSTRACT

EO images suffer from poor visibility when they are captured at night. Besides, quality of EO images is also varied during day time, e.g., they are usually corrupted by various types of occlusions, such as cloud or haze. SAR imagery provides reliable measurements for all weather and time, but has relatively low resolution as well as speckle noise corruption, which limit the interpretation. Besides, SAR data are rare and usually obtained from different platforms (airborne or spaceborne) with varied imaging parameters (imaging angles, bands, etc.). We aim to explore the use of deep generative adversarial networks (GANs) for cross-modality correlation learning in remote sensing applications.

1. Challenges in Current Algorithms

For existing relevant works that are focusing on domain transfer, most of them can be summarized into three categories according to the availability of training data: self-supervised, semi-supervised, and supervised learning methods in Table 1.

	Training Data Requirement
Self-supervised (TuiGAN [1])	**One reference image** from the target domain, i.e., single SAR image
Semi-supervised (CycleGAN [2])	**A training corpus** of both EO and SAR images
Supervised (Pix2Pix [3])	A training corpus of **paired** EO and SAR images

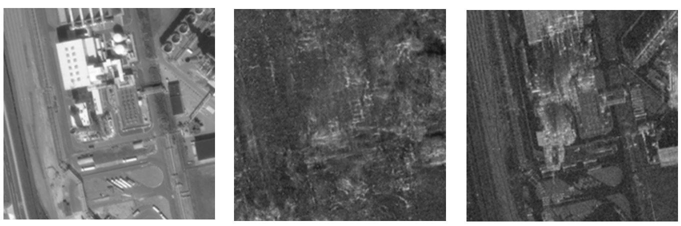
Table 1: Different requirements of training data for most domain transfer methods.

The performance of self-supervised learning method largely depends on the similarity between the used training reference image and the input EO image. Our observation is that when input EO image differs from the reference image, the model cannot transfer the corresponding SAR-specific structures effectively.

Yet a dilemma lies in semi-supervised and supervised learning methods for EO/SAR domain transfer. For semi-supervised methods, a training corpus of EO and SAR images is required, which does not need to be paired or aligned, e.g., if you would like to transfer an EO image to SAR, you just need one EO image set, and another independent SAR image set. While for supervised methods, a training corpus of perfectly aligned in the pixel level EO and SAR images is required, e.g., there must be no pixel shift at all for each pair of EO and SAR images that are used for training. However, for EO/SAR datasets, paired but slightly misaligned images are the most common type. Figure 1 and Figure 2 give samples of SAR images generated by CycleGAN [2] (semi-supervised learning algorithm) and Pix2Pix [3] (supervised learning algorithm). Our observation is that Pix2Pix completely fails and CycleGAN gives suboptimal results on EO/SAR domain transfer task. Therefore, we plan to build our EO-SAR domain transfer framework with CycleGAN as the backbone model.

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

Fig. 1: The generated SAR image using CycleGAN from the input EO image.



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

Fig. 2: The generated SAR image using Pix2Pix from the input EO image.

2. Incorporating Pair Loss into Semi-supervised Learning

Pairwise correspondence information in EO/SAR datasets should be further exploited. We propose shift-insensitive pair loss to reinforce low-level reconstruction quality given that perfect alignment is unachievable in EO/SAR pairs.

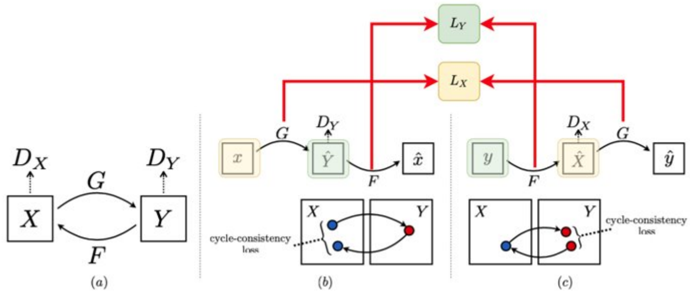


Fig. 3: CycleGAN network structure with shift-insensitive pair loss L_X and L_Y. The two mapping functions, i.e., G that maps X to Y, and F that maps Y back to X, are meanwhile used as the feature extractors to project inputs to feature space.

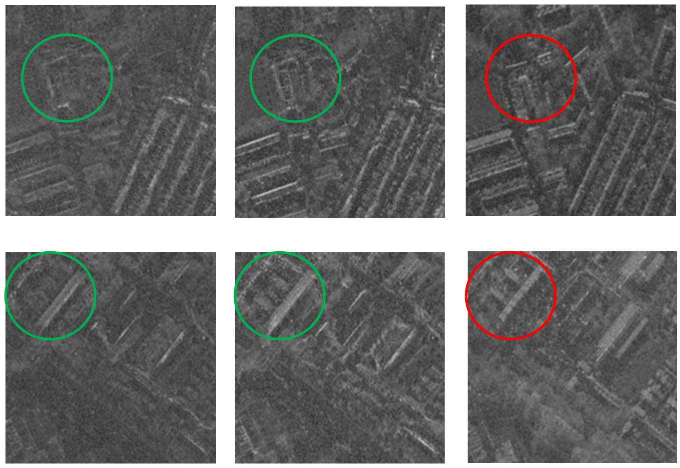
As shown in Figure 3, the two shift-insensitive pair losses L_X and L_Y are calculated by l_1-distance between real images (x, y) and fake images $(\hat{X}, \hat{Y})$ in feature space:

$$L_X = \left\| h_G(x) - h_G(\hat{X}) \right\|_1$$

$$L_Y = \left\| h_F(y) - h_F(\hat{Y}) \right\|_1$$

where $h_G(\cdot)$ and $h_F(\cdot)$ represent the first N-layers of generator G and F respectively. Traditional way of imposing pairwise loss is to make comparisons in image space. However, misalignment between EO/SAR pairs will introduce errors. We notice that the misalignment between current EO/SAR pairs are mild, which means that the pixel value of SAR images at location (i, j) corresponds to that of EO images at location (i', j'), where (i', j') is within the neighbourhood of (i, j). Thus after the two feature extractors $h_G(\cdot)$ and $h_F(\cdot)$, we expect such mismatch can be largely alleviated as convolutional layers are used.

We conducted experiments to verify the effectiveness of shift-insensitive pair loss. We set N=3, which corresponds to feature maps right after the down-sampling layers. As shown in Figure 4, it can be observed that more fine-scale spatial structures are captured with shift-insensitive pair loss.



(a) Without Pair Loss (b) With Pair Loss (c) Real SAR image

Fig. 4: Examples of the transferred SAR images using CycleGAN without (left) and with (middle) shift-insensitive pair loss.

3. Conclusion

We presented an EO-SAR domain transfer framework, which incorporates shift-insensitive pair loss into CycleGAN. For future work, we plan to further exploit geometry information in EO-SAR images as man-made structures, e.g., buildings, are of more interest to our task.

References

[1] 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.

[2] 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).

[3] 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).