LULC Guided SAR-to-RGB Image Translation with Generative Adversarial Network

K Ram Prabhakar* Veera Harikrishna Nukala* Jayavardhana Gubbi Arpan Pal Balamuralidhar P
Embedded Devices and Intelligent Systems
TCS Research, INDIA

Abstract

*Remote sensing data captured in optical modality suffers from cloud cover and invisible during poor illumination conditions. SAR images can handle these issues but are not easily interpretable. To address this problem, we propose a Generative Adversarial Network (GAN) based SAR to RGB image translation model. In our approach, we utilize Land Use and Land Cover (LULC) maps generated from SAR images, to provide semantic class information that aids in improving the generated RGB image quality. In addition, we also use DFC2020 dataset to pretrain our model. The presented approach is tested on the MultiEarth 2022 - Image to Image translation challenge under the team name "**SAR_RS_v6**" that achieves **0.0445** score in the leaderboard.*

1. Introduction

In Remote sensing, optical data or RGB images are very useful to many downstream tasks and most importantly they are easily interpretable by humans. However, they cannot be used to capture data in night or poor illumination conditions. Majorly, they are affected by frequent cloud covers. There are existing methods that can remove clouds from the RGB images, however they are cumbersome and may not always be reliable. In contrast, SAR offers visibility through clouds as well as during night. Hence, it an ideal candidate for many remote sensing applications. However, SAR is not easily interpretable.

We address this issue in this work by translating the SAR data into a RGB image. In our proposed method, we use Generative Adversarial Network (GAN) to synthesize three different RGB version for each SAR image. In addition, we use LULC maps generated from SAR image to generate semantically meaningful output. Our model achieves a score of **0.04451** in the MultiEarth 2022 - Image translation challenge leaderboard under the team name of "***SAR_RS_v6***".

*equal contribution

2. Proposed Method

The input to the model is SAR VV and VH bands after truncating them between (-25,0) and normalizing between (-1,1). With SAR data as input, the model is trained to predict three different RGB images. Training a model to directly generate RGB images from SAR can have following challenges: high frequency classes can dominate, and prediction can be semantically not meaningful. To address these issues, we propose to use LULC maps to guide the model and predict semantically plausible output RGB images.

The overall architecture of the proposed method is shown in Fig. 1. Our model consists of three major blocks: LULC generation, RGB generator, and PatchGAN discriminator.

LULC generation: In order provide semantic layout of the scene, we propose to use LULC maps in the GAN framework. The LULC maps are generated by a generator (G_1) with only SAR data as input. For G_1 network, we use U-net architecture [3]. G_1 is trained to predict 8-class LULC labels: Forest, Shrubland, Grassland, Wetlands, Croplands, Urban, Barren, and Water. We train G_1 using publicly available DFC2020 dataset [5]. The dataset consists of Sentinel-1 SAR image and the corresponding 10m high-resolution land cover map. We train G_1 for 200 epochs with batch size 8, learning rate of 10^{-4} and Adam optimizer [2]. We use categorical cross entropy loss function to compute loss between generated and ground truth LULC maps. G_1 is trained separately and the weights are frozen while using it to translate SAR to RGB image.

RGB image generation: The generated LULC map is concatenated with input SAR image and fed as input to a generator network, G_2. For G_2, we use U-net architecture with 59 layers. We replace the MaxPooling layers with strided convolutions for sharp results. All layers use ReLu activation except for Tanh activation in the final layer. The G_2 model is trained to generate three RGB images for each SAR input image. We train G_2 with learning rate of 2×10^{-4}, Adam optimizer with momentum of 0.5, and

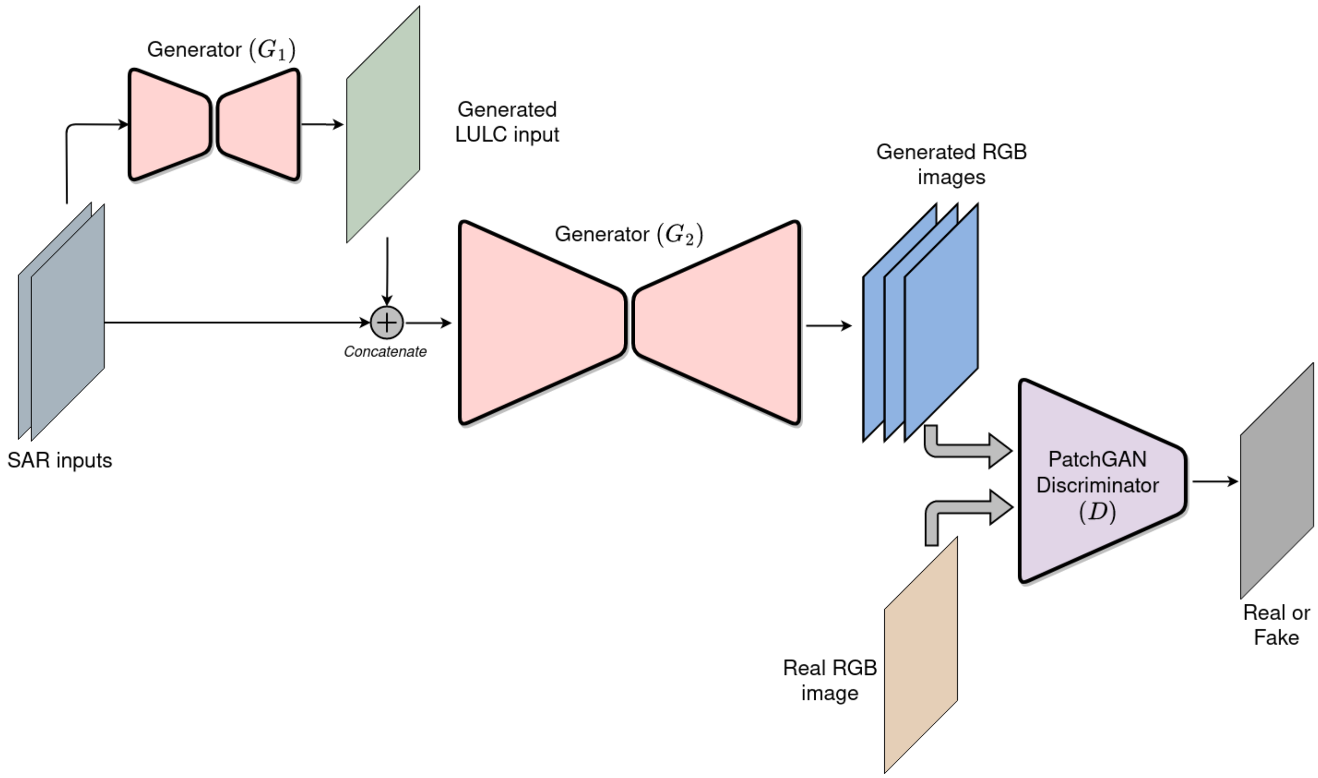

Figure 1. Architecture of our proposed model.

batch size of 8 for 400 epochs. We use mean square error, Structural Similarity Index Metric (SSIM) and Color SSIM loss functions to compute loss between generated RGB and ground truth RGB images.

Discriminator: The output of G_2 is passed onto a PatchGAN discriminator [1], D. D network aims to classify each pixel into either real or fake. We follow standard GAN training prodecure to pass batches of real and predicted RGB images to D and train it in turns. We train D with learning rate of 10^{-4} and Adam optimizer with 0.5 momentum. We use binary cross entropy loss function to train the discriminator network.

3. Experiments and Results

We pretrain our model with DFC2020 dataset [5] first before finetuning. In Fig. 2, we present qualitative results by our proposed method. In Table 1, we present ablation models with different loss functions used. Using only SSIM [4] or color SSIM loss functions for all three RGB generated images results in poor performance compared to using three different loss functions for three RGB images.

4. Conclusion

In this paper, we present a deep learning based model to synthesize RGB images from SAR data. We propose to use

Table 1. Quantitative comparison between proposed model submitted under team name "SAR_RS_v6" and other ablations.

Team name	Experiments	Test set score
	SSIM loss	0.0476
	Color SSIM loss	0.0705
SAR_RS_v6	Proposed model with SSIM, color SSIM and L2 loss	**0.0445**

LULC maps as auxillary information that helps the model to generate sharp and semantically meaningful outputs. The proposed method was tested on MultiEarth - 2022 Image translation challenge test set and achieves a score of **0.4451** under the team name of "**SAR_RS_v6**".

References

[1] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1125–1134, 2017. 2

[2] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014. 1

[3] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-

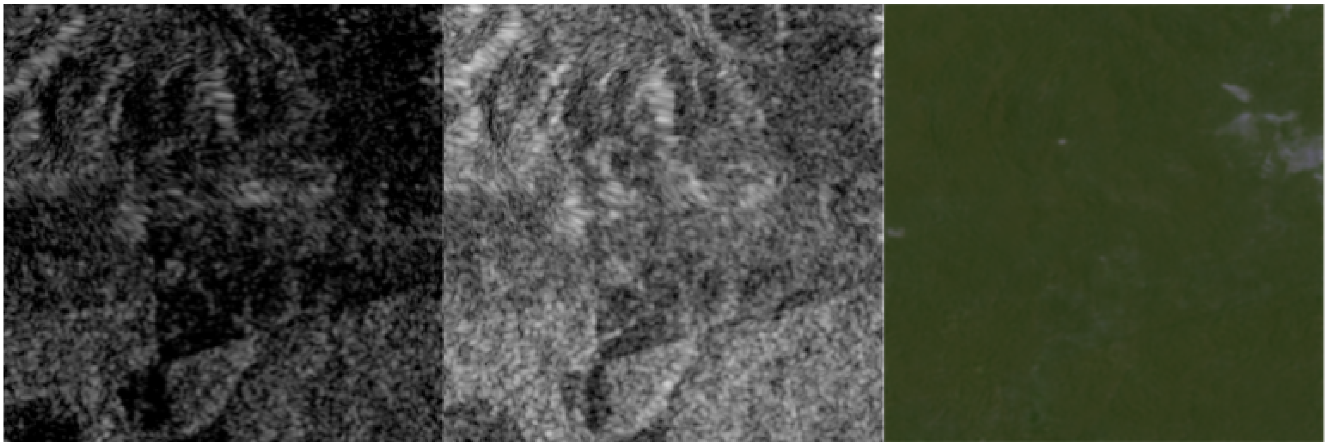

Figure 2. Qualitative results by our proposed model.

net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015. 1

[4] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4):600–612, 2004. 2

[5] Naoto Yokoya, Pedram Ghamisi, Ronny Hänsch, and Michael Schmitt. 2020 IEEE GRSS data fusion contest: Global land cover mapping with weak supervision. *IEEE Geoscience and Remote Sensing Magazine*, 8(1):154–157, 2020. 1, 2