SAR Image Colorization using cGAN and Gradient Difference in LAB Color Space

Md Nafis Islam*, Ashutosh Anand†, Mukulgopal Mandal‡, Lipika Mohanty§, Adyasha Dash¶
School of Computer Science and Engineering, Kalinga Institute of Industrial Technology, Bhubaneswar, India
{ *nafis71041.work, †anandashutosh555, ‡mukulgopalmandal }@gmail.com
{ §lipika.mohantyfcs, ¶adyasha.dashfcs }@kiit.ac.in

***Abstract*—SAR image colorization enhances the interpretability of radar-based imagery, making it more accessible for various remote sensing applications. However, SAR images are inherently grayscale and suffer from speckle noise, limiting the effectiveness of traditional colorization methods. The proposed framework employs a Conditional GAN that uses a novel gradient loss computed in the CIELAB color space to improve edge preservation. Our approach uses U-Net (generator) and PatchGAN (discriminator), with BCE loss, pixel-wise loss, and our proposed gradient loss, which captures structural differences in the L-channel of LAB space. Experimental results show that our model outperforms prior CNN and GAN-based methods, achieving PSNR of 21.99, SSIM of 0.63, and LPIPS (AlexNet) of 0.3059. These findings show the effectiveness of our approach in generating colorized SAR images, advancing remote sensing visualization. GitHub: https://github.com/nafis71041/sar_colorization_gradient**

***Index Terms*—SAR Image Colorization, Remote Sensing, Conditional GAN (cGAN), Deep Learning, Gradient Loss, CIELAB Color Space**

I. Introduction

A. Background

Synthetic Aperture Radar (SAR) is a sophisticated remote sensing method that uses microwave frequencies to generate high-resolution images. Unlike optical imaging, which depends on external illumination, SAR actively transmits microwaves and records their backscatter signals. The resulting grayscale images encode valuable information about surface roughness [1], moisture content [2], and dielectric properties [3]. SAR imagery is effective in all type of weather conditions, including at night and through cloud cover, fog, and light precipitation.

Despite its utility, SAR imagery presents significant challenges for interpretation. The interference of reflected waves from multiple scatterers at the same pixel location leads to speckle noise, characterized by a salt-and-pepper appearance. Unlike additive noise, which can be reduced with traditional denoising techniques, SAR noise is multiplicative [4], making its removal more complex. Furthermore, the absence of spectral color information limits the accessibility of SAR images to non-experts, as grayscale representations do not intuitively convey object types or material compositions. This has led to increased interest in developing automated SAR colorizing methods.

979-8-3315-2261-2/24/$31.00 ©2024 IEEE

B. Motivation

However, SAR image colorization is a complex problem due to the difficulty of preserving structural details while generating realistic colors. Current methods frequently fail to maintain realistic textures, resulting in visual distortions that reduce image quality. Hence, we propose a SAR image colorization framework incorporating a novel gradient loss computed in the CIELAB color space and Conditional GAN. The key motivation behind our work is to enhance structural preservation and suppress noise while achieving visually realistic colorization.

C. Contribution

Our key contributions include:

- Introduction of a novel gradient loss function in the CIELAB space to enhance edge retention and structure-aware colorization.
- A novel Conditional GAN architecture, specifically designed for high-fidelity SAR-to-optical image translation.
- Comprehensive evaluation demonstrating superior performance compared to existing GAN-based models, with improvements in PSNR, SSIM, and LPIPS.

The paper unfolds as this: Section II discusses existing approaches in SAR image colorization. Section III details our proposed method, featuring the Conditional GAN architecture with its U-Net generator, PatchGAN discriminator, and our innovative gradient loss function. Section IV evaluates experimental outcomes and provides comprehensive analysis. We conclude in Section V by summarizing our principal contributions and suggesting potential avenues for future investigation.

II. Literature Review

Various approaches have been proposed to improve SAR image colorization by leveraging deep learning and traditional signal processing techniques. Deng et al. [5] utilized scattering characteristics to assign colors to SAR images, helping to distinguish objects based on their dielectric properties. However, this rule-based method is limited in its generalization to complex terrains. Nowdays GANs are one of the most preferred techniques for SAR-to-color image transformation. Isola et al. [6] established a cGAN-based solution, for generic image-to-image transformation tasks, which later inspired SAR colorization models. In particular, Shen et al. [7] proposed IcGAN4ColSAR, a multispectral cGAN approach

tailored for SAR images, demonstrating improvements in spectral fidelity. Ji et al. [8] extended this idea with a multidomain cycle-consistency GAN to refine texture preservation during colorization.

In parallel, Schmitt et al. [9] introduced the Sentinel 1 & 2 dataset, which has become a benchmark for SAR-optical fusion models. They proposed a variational autoencoder-based approach to improve the realism of generated images while retaining SAR-specific properties. Shen et al. [10] further contributed by establishing a benchmarking protocol, highlighting the importance of evaluating SAR colorization models using multiple metrics, including PSNR, SSIM, and perceptual similarity scores. These studies provide the foundation for our work.

III. Methodology

In our work, we propose a cGAN-based SAR image colorization framework. Our key innovation is the introduction of a novel gradient loss function, specially designed to preserve structural details by computing forward differences of the generated and target images in the L channel of the CIELAB color space. This ensures that fine-grained textures and edges are better retained during the colorization process. By leveraging adversarial loss, pixel-wise loss, and our gradient loss, we achieve superior performance in generating visually coherent and structurally accurate colorized SAR images. In this section, we describe the dataset, preprocessing techniques, model architecture, and loss functions used for training our SAR image colorization framework.

A. Dataset & Preprocessing

We trained our model on a subset of 2,000 SAR-optical image pairs, comprising 1,000 urban and 1,000 agricultural images, extracted from the Sentinel 1 & 2 SAR-Optical dataset available on Kaggle [11]. This dataset is a curated subset of the larger Sentinel 1 & 2 dataset provided by Schmitt et al. [9]. To ensure consistency and compatibility with deep learning frameworks, all images were resized to 256×256 pixels. Additionally, pixel intensities were normalized to the range [-1, 1], a necessary preprocessing step for stabilizing GAN training and improving convergence.

B. Model Architecture

a) Generator: The generator follows the U-Net architecture [12], which uses multiple encoding and decoding blocks with skip connections to maintain fine-grained spatial details throughout the transformation process. The encoder has multiple convolutional layer, batch normalization and LeakyReLU activation aggregate block. Each encoder block progressively down-samples the input SAR image, capturing hierarchical features at different spatial resolutions. The decoder has multiple transposed convolutional layer and ReLU activation aggregate block, progressively up-sampling the latent representation to reconstruct the colorized image. Skip connections directly join two same level encoder and decoder layers to prevent the loss of small details and improve the structure of the generated images. Fig. 1 describes the detailed structure of the generator.

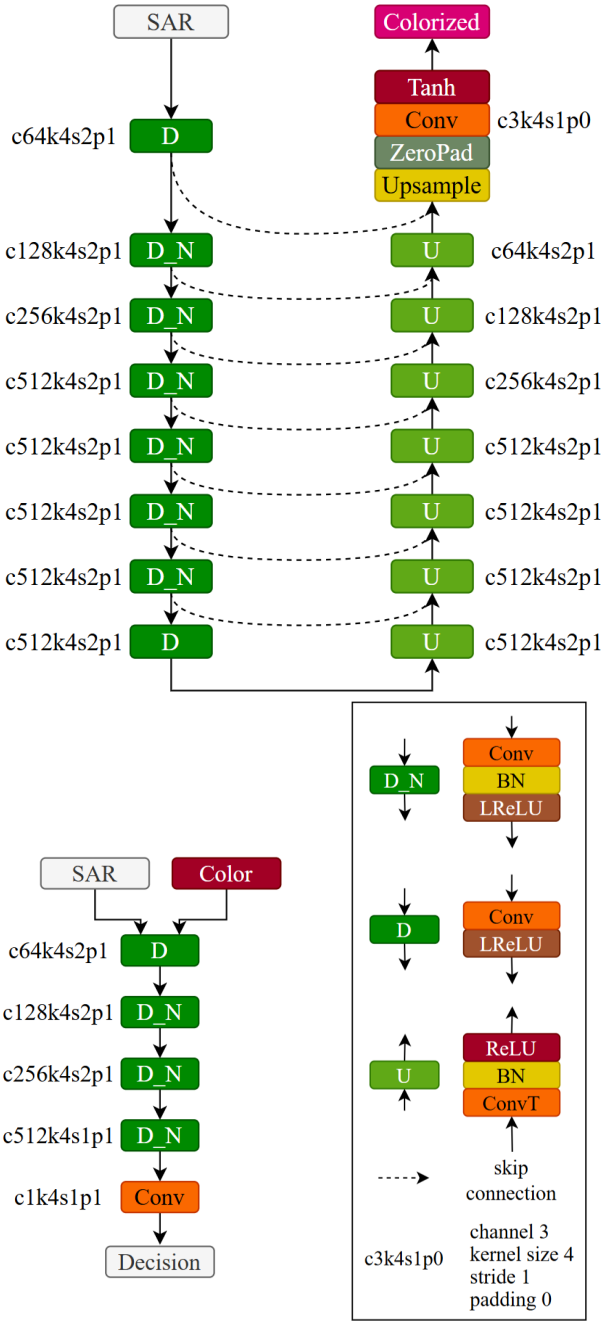

Fig. 1: Model Architecture: Generator (Top), Discriminator (Bottom-Left)

b) Discriminator: The discriminator is based on the PatchGAN architecture [6], which classifies whether small patches of the generated image are real or fake instead of evaluating the entire image at once. This approach increases the model's effectiveness to capture small details. The discriminator takes as input a concatenated one-channel SAR image and the three-channel optical/generated image. It has multiple convolutional layers with progressively increasing depth with batch normalization and LeakyReLU activation functions. The output is a 30×30 matrix, where each value

represents the probability of the corresponding patch being real or synthetic. This patch-based evaluation provides a finer level of discrimination and helps in generating sharper, more realistic images. Fig. 1 also describes the detailed structure of the discriminator.

C. Loss Functions

To train the GAN, we use an aggregation of adversarial, pixel-wise, and gradient-based loss functions.

a) Adversarial Loss: The adversarial loss drives the generator to create so realistic images that technically the discriminator becomes incapable of telling it from the real ones. This is achieved by using Cross Entropy Loss (Binary) as the adversarial loss, given by (1).

$$L_{\text{BCE}} = -\mathbb{E}_x \left[0.9 \log D(G(x)) + 0.1 \log(1 - D(G(x)))\right] \quad (1)$$

where:

- x is the SAR image,
- $G(x)$ is the generated image (colorized),
- $D(G(x))$ is the confidence level of the discriminator that the generated image is real or not.

To stabilize training, label smoothing is applied, replacing 1's with 0.9's and 0's with 0.1's in original BCE loss function.

b) Pixel Loss: Pixel-wise loss ensures that the generated image maintains similarity to the target image in terms of absolute intensity values. We use Mean Absolute Error (shown in (2)) for pixel loss.

$$L_{\text{MAE}} = \frac{1}{N} \sum |G(x) - y| \quad (2)$$

where N describes all pixel points.

c) Gradient Loss: To preserve structural details, we propose to use a gradient-based loss. The CIELAB color space is used since its L channel represents image structure. Gradients in vertical and horizontal directions are calculated using (3):

$$\nabla_x L = L_{(i,j)} - L_{(i-1,j)}, \nabla_y L = L_{(i,j)} - L_{(i,j-1)} \quad (3)$$

Then the gradient loss is computed as follows:

$$L_{\text{GRAD}} = \frac{1}{N} \sum \left(|\nabla_x L_G - \nabla_x L_y| + |\nabla_y L_G - \nabla_y L_y|\right) \quad (4)$$

where L_G and L_y represent the L channels of the generated and target images in CIELAB color-space, respectively.

d) Total Losses: The total loss for the generator is given by:

$$L_{\text{gen}} = L_{\text{BCE}} + \lambda_1 \times L_{\text{MAE}} + \lambda_2 \times L_{\text{GRAD}} \quad (5)$$

where $\lambda_1 = \lambda_2 = 100$.

The discriminator loss is given by:

$$L_{\text{disc}} = \frac{1}{2} \left[L_{\text{BCE}}(D(y), 0.9) + L_{\text{BCE}}(D(G(x)), 0.1)\right] \quad (6)$$

D. Training Strategy

Training was conducted for 100 epochs in an NVIDIA T4 GPU via Google Colab. We have taken 16 as mini-batch size. We employed the Adam optimization algorithm with the following parameters:

- Generator learning rate of 2×10^{-4}.
- Discriminator learning rate 10^{-5}.
- $\beta_1 = 0.5$, $\beta_2 = 0.999$ to stabilize training.

Label smoothing and lower learning rate were used to prevent discriminator overpowering.

E. Assessment Metrics

We have used three metrics to evaluate the quality of generated images by our model:

a) PSNR: This measures generation quality/effectiveness in decibels (dB). PSNR is calculated as:

$$PSNR = 10 \log_{10} \left(\frac{MAX_I^2}{MSE}\right) \quad (7)$$

where MAX_I is the maximum intensity.

b) SSIM: This shows how much two images are similar by considering their brigthness, contrast and overall structure [13]. This is computed by:

$$SSIM(x, y) = \frac{(2\sigma_{xy} + C_2)(2\mu_x\mu_y + C_1)}{(\sigma_x^2 + \sigma_y^2 + C_2)(\mu_x^2 + \mu_y^2 + C_1)} \quad (8)$$

c) LPIPS: This calculates the perceptual distance [14] with the features extracted from a already trained neural network (AlexNet). This is calculated as by:

$$LPIPS(x, y) = \sum_l w_l \left|\phi_l(x) - \phi_l(y)\right|^2 \quad (9)$$

where ϕ_l represents deep features at layer l.

IV. Results and Discussion

By integrating gradient loss in the CIELAB color space, our model achieves better preservation of fine details and reduces noise, leading to more visually realistic outputs. Some exemplary images are shown in Fig. 2. The left most image is the SAR image, right most image is the corresponding optical pair and middle image is the generated image by our model.

The discriminator plays a key role in this improvement. The PatchGAN-based discriminator evaluates the generated images at a patch level rather than globally, making it more sensitive to texture and structural inconsistencies. Meanwhile, the generator benefits from multiple loss functions. Gradient Loss reinforces structural preservation by minimizing edge distortions.

Table I compares of results of a baseline cGAN (without Gradient Loss) and that of the designed model. All the reported metrics were evaluated on 500 test images, which exclusively consist of urban and agricultural regions. The PSNR improved from 16.35 to 21.99, indicating lower reconstruction error. The SSIM increased significantly from 0.187 to 0.630, showing better preservation of local structures and textures. Moreover, the LPIPS metric suggests a reduction in perceptual differences, implying a higher visual quality.

TABLE I: Model Comparison

Metrics /Model	cGAN (without Gradient Loss	cGAN (with Gradient Loss
PSNR	16.35*	21.99
SSIM	0.187*	0.630
LPIPS	-	0.306

*values are taken from Ji et al. [8].

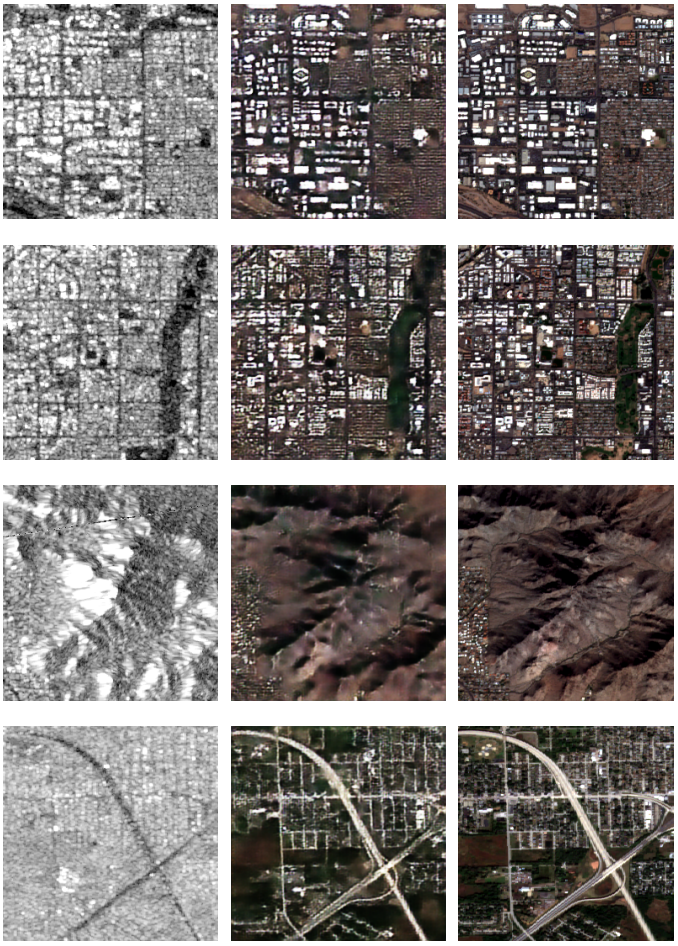

Fig. 2: SAR Images (left), Generated Images (middle) and Optical Images (right)

V. Conclusion

Our work proposes a new methodology for SAR image colorization leveraging a Conditional GAN (cGAN) with an enhanced loss function incorporating gradient differences in the CIELAB color space. Our approach effectively mitigates the challenges of SAR image translation, particularly preserving structural details despite the presence of speckle noise. By integrating adversarial loss, pixel-wise loss, and our novel gradient difference loss, the model maintains both global consistency and local texture fidelity. The results demonstrate significant improvements in PSNR, SSIM, and LPIPS scores over traditional cGAN-based methods. Additionally, label smoothing and adaptive learning rate scheduling contributed to stable and efficient training, preventing mode collapse and ensuring high-quality colorization.

Future research directions can include exploring attention mechanisms, multi-scale feature extraction, and transformer-based architectures to further enhance the colorization quality. Moreover, expanding the dataset with diverse terrain types can improve the applicability of the model for real-world SAR applications.

References

[1] J. Shi, J. Wang, A. Y. Hsu, P. E. O'Neill, and E. T. Engman, "Estimation of bare surface soil moisture and surface roughness parameter using l-band sar image data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 35, no. 5, pp. 1254–1266, 1997.

[2] C. Yin, Q. Liu, and Y. Zhang, "Using the aiem and radarsat-2 sar to retrieve bare surface soil moisture," *Water*, vol. 16, no. 11, p. 1617, 2024.

[3] A. Singh, M. Niranjannaik, S. Kumar, and K. Gaurav, "Comparison of different dielectric models to estimate penetration depth of l-and s-band sar signals into the ground surface," *Geographies*, vol. 2, no. 4, pp. 734–742, 2022.

[4] Y. Zhou, Y. Li, Z. Guo, B. Wu, and D. Zhang, "Multiplicative noise removal and contrast enhancement for sar images based on a total fractional-order variation model," *Fractal and Fractional*, vol. 7, no. 4, p. 329, 2023.

[5] Q. Deng, Y. Chen, W. Zhang, and J. Yang, "Colorization for polarimetric sar image based on scattering mechanisms," in *2008 Congress on Image and Signal Processing*, vol. 4. IEEE, 2008, pp. 697–701.

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

[7] K. Shen, G. Vivone, S. Lolli, M. Schmitt, X. Yang, and J. Chanussot, "Icgan4colsar: A novel multispectral conditional generative adversarial network approach for sar image colorization," *IEEE Geoscience and Remote Sensing Letters*, 2025.

[8] G. Ji, Z. Wang, L. Zhou, Y. Xia, S. Zhong, and S. Gong, "Sar image colorization using multidomain cycle-consistency generative adversarial network," *IEEE Geoscience and Remote Sensing Letters*, vol. 18, no. 2, pp. 296–300, 2020.

[9] M. Schmitt, L. H. Hughes, and X. X. Zhu, "The sen1-2 dataset for deep learning in sar-optical data fusion," *arXiv preprint arXiv:1807.01569*, 2018.

[10] K. Shen, G. Vivone, X. Yang, S. Lolli, and M. Schmitt, "A benchmarking protocol for sar colorization: from regression to deep learning approaches," *Neural Networks*, vol. 169, pp. 698–712, 2024.

[11] P. Tiwari, "Sentinel-1&2 image pairs (sar & optical)," March 2021, accessed on Feb 20, 2025. [Online]. Available: https://www.kaggle.com/datasets/requiemonk/sentinel12-image-pairs-segregated-by-terrain

[12] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*. Springer, 2015, pp. 234–241.

[13] Z. Wang, "Image quality assessment: Form error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 604–606, 2004.

[14] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The unreasonable effectiveness of deep features as a perceptual metric," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 586–595.