International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0224

e ISSN: 2584-2854
Volume: 04 Issue: 05
May 2026
Page No: 1492 - 1497

SAR-to-Optical Image Translation Using Pix2Pix GAN

Anjana C[1], Abhijith B[2], Amrith M Nair[3], Mohamed Harshil[4], Alen George[5], Abhimanyu A K[6]
[1,2,3,4,5,6]Department of Computer Science, Yenepoya University, Bangalore, India.
Email ID: *anjanachevara@gmail.com[1], abhijith87b@gmail.com[2], amrithmnair2004@gmail.com[3], mohamedharshil@gmail.com[4], alengeorge807@gmail.com[5], abhimanyueak@gmail.com[6]*

Abstract
Synthetic aperture radar (SAR) is very useful in remote sensing because it can capture images all year round, day and night. However, its complex electromagnetic scattering properties can be hard for humans to understand and inter- pret. Generative models based on deep learning have shown promise in translating between SAR and optical images. Still, their effectiveness depends on having large training datasets, which can be costly in terms of logistics and finances. To address these issues, we propose an improved image-to-image translation framework using a Pix2Pix conditional generative adversarial network (CGAN) to facilitate high-quality SAR-to- optical image transformation. The primary goal of this work is to increase the variety and structural accuracy of the generated images, especially when there are few training samples; this scenario is often called few-shot image generation (FSIG). The methodology employs a U-Net generator structure combined with Spectral Normalization (SN) in the discriminator to provide stability during adversarial training. A key innovation of this framework is the use of a new loss function called Pairwise Distance (PD). This approach normalizes the synthesized results using simulated SAR data, allowing the model to reflect the inherent variability found in real scattering phenomena. We expect improved performance in terms of structural similarity (SSIM) and higher classification accuracy, particularly when the generated images serve as data augmentation. This study has successfully converted the complex geometric and radiometric features of SAR into optical representations that are easier to interpret. This development enhances the usefulness of SAR data in resource-limited environments. Overall, the proposed model will create a scalable path for advancing automated target recognition and environmental surveillance by synthesizing cross- modal features effectively.

Keywords: *Synthetic Aperture Radar, Pix2Pix GAN, Image Translation, Attention U-Net, Remote Sensing.*

1. Introduction

Synthetic Aperture Radar (SAR) and optical visual imaging represent There are two valuable technologies in remote sensing. SAR works through passing radar signals and reflecting back on them at the ground level which allows imaging in all weather and all day conditions Nonetheless deciphering of SAR images is hazardous since they are radar backscatter instead so as compared to visual light they are abstract and hard to human interpretation Optical sensors on the other hand can capture light in the visible and near-infrared spectra and that is why it produces the kind of natural intuitive imagery that our eyes are already used to understanding. infrared light and generates visual information that is intuitive is more comprehensible to human beings though it is restricted by daylight and cloud cover conditions availability Achieving a correlation between SAR and optical imagery is a significant line of research in remote sensing to enhance the interpretability and usability of SAR data. Due to its substantial benefits SAR imagery provides notable benefits to its ability to take pictures under all-weather condition and lighting But the want of visual clearness renders SAR images that can hardly be interpreted limiting their use. in fields like agriculture, environmental surveillance and urban planning. The study will solve these problems by with a deep learning model to transform SAR images to optical images which are more visually rich and easier to read Image-to-image translation is done through the Pix2Pix system where U-Net is viewed as a baseline architecture and Attention U-Net enhancing maintenance of structural

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0224
e ISSN: 2584-2854
Volume: 04 Issue: 05
May 2026
Page No: 1492 - 1497

information throughout the process of translation. The main objective of this study is to design and improve a Pix2Pix GAN model Precisely created to render SAR to optical images. [1]- [3] The analysis measures the performance of the approach that was proposed based on SSIM and PSNR measures and examining its possibilities This technology in our opinion is the most essential to emergency team members in the event of natural disasters In such high stress scenarios a heavy cloud cover may usually make traditional satellite images useless but our translated SAR images can give the team the clear visual data required to execute rescue efforts at a quicker pace environmental control, city development, and agriculture analysis.

2. Literature Review

Synthetic Aperture Radar (SAR) imagery finds wide application in remote sensing because it can be used to capture images regardless of weather and lighting. Nevertheless, the interpretation of SAR images is difficult due to speckle noise and the different imaging processes relative to optical images. As a result, there has been a lot of research work focused on converting SAR images to optical images to ensure better visualization. Previous studies had used conventional image processing and machine learning, such as filtering, feature ex- traction, and statistical models, to analyse SAR images. Although they did positively impact image quality they were limited in their capability to accurately translate SAR images to optical images During the advancement of deep learning Convolutional Neural Networks (CNNs) and Generative Adversarial Networks (GANs) have been widely used in image- to-image translation problems. Pix2Pix and CycleGAN are models that have shown potential outcomes in creating optical images using SAR data through the learning of the correlation between the two image domains. With these developments there are still a number of challenges. [4]- [6] The main problems are variations in SAR and optical imaging characteristics, speckle noise of SAR measurements and the lack of matched datasets to develop models. Such constraints affect the image quality and precision of produced images. Consequently, there is a knowledge gap of creating more effective models to generate realistic optical images whilst maintaining structural details of SAR images. The proposed research will fill this gap by using GAN-based methods to enhance SAR-to-optical image translation.

3. Methodology

3.1.System Overview

The dataset applied in this research is QXSLABSAROPT which contains Synthetic Aperture Radar (SAR) and optical images with geometrically aligned pairs. In this first set, there were approximately 20,000 pairs of images. This dataset was then down-sampled to 99 paired samples to serve the goals of the current study, therefore, maximizing computational savings in the process of building and empirically testing a model. Using encoder-decoder-generated generator structure, the generalized paradigm of the Pix2Pix Generative Adversarial Network (GAN) has been used in the architectural design. This system in essence is a generator and a discriminator. The generator is carried out by using the U-Net architecture with skip connections and the discriminator is carried out by the concept of Patch GAN architecture, which measures image verisimilitude in a localized and patch-based granularity. [7] Two generator configurations were explored in this work. The baseline model uses a standard U-Net generator, whereas the enhanced model incorporates Attention Gates within the U- Net architecture, forming an Attention U-Net. These attention mechanisms allow the network to focus on important spatial features while suppressing irrelevant information. The models were trained for 20 epochs using a batch size of 1 and a learning rate of 0.0002. The implementation was carried out in Python using the Tensor Flow and Keras deep learning frameworks. The experiments were executed on the Google Colab platform with the support of NumPy and Matplotlib libraries. Prior to training, SAR and optical images were resized to a resolution of 256×256 pixels. The pixel values were normalized within the range of $[-1, 1]$ and stored as com- pressed NumPy arrays. During training, the encoder component of the generator extracts hierarchical features from the SAR input images. In the Attention U-Net model, Attention Gates generate

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0224
e ISSN: 2584-2854
Volume: 04 Issue: 05
May 2026
Page No: 1492 - 1497

feature importance weights using a sigmoid activation function. These weighted features are then combined during the decoder phase through up sampling operations to reconstruct the final optical image. This cross-modal image translation framework aims to re- duce the visual complexity of SAR imagery and convert it into optical-like representations that are easier to interpret. Such translation techniques can support applications including dis- aster management, maritime monitoring, agriculture analysis, environmental observation, and urban planning. Generative Adversarial Networks (GANs) consist of two neural networks that are trained simultaneously: a generator and a discriminator. The generator produces synthetic images that resemble the training data, while the discriminator evaluates whether an image is real or generated. [8]- [9] Through this adversarial process, all networks are optimized. the generator to generate more and more realistic outputs. The model used in this study is Pix2Pix GAN model on SAR- to-optical image translation. Pix2Pix is a conditional GAN structure that is specific to image-to-image correspondence. translation tasks. The traditional models of GAN also produce unlike the new models. Pix2Pix a direct translation of an, random noise images. Mapping of an image in one space to an image in another space. and without any change of figure. Pix2Pix model generator takes the form of U-Nature contained encoder- decoder pathways and skip connections. The hierarchical nature of the encoder is ascertained out of. input image and the contracting of spatial dimensions with con- volutational functions and pooling functions. The decoder reconstructs the reconstruction of the image that is carried out by up sampling and the layers are omitted. connections transfer high-resolution data of the encoder. to the two equal decoder layers. This structure preserves Local and global information is applied in the creation of image. The discriminator is based on the PatchGAN design. operates on small scale blocks and not on the bigger picture. This will enhance network capture of fineness of the network. description and texture details. The Pix2Pix framework also and pixel-wise L1 loss and adversarial loss. to derive form similarity and generation of natural images. As Shown in Figure 1.

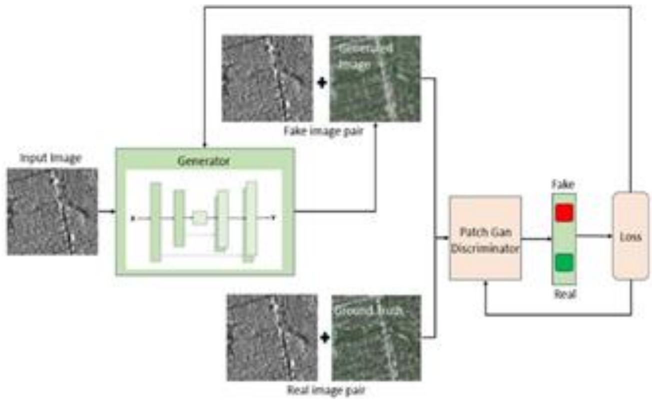

Figure 1 Architecture of Pix2Pix GAN

In the Pix2Pix framework, the U-Net architecture plays a crucial role as the generator network responsible for translating SAR images into optical images while preserving structural information. Originally proposed by Ronneberger et al. in 2015, U-Net is widely recognized for its effectiveness in image segmentation tasks due to its encoder–decoder design and the presence of skip connections. [10] The encoder component learns the hierarchical features by the use of convolutional layers and Although the decoder recreates the final, pooling operations. As Shown in Figure 2.

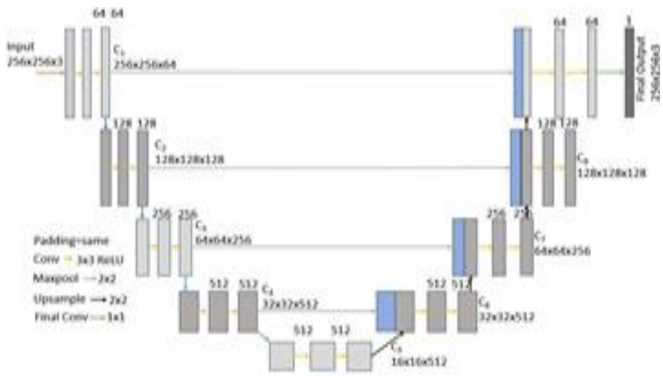

Figure 2 Architecture of U-Net

optical image with up sampling layers and skip connection. To improve the feature extraction ability, an Attention U-Net as an enhanced generator model, architecture is included. The Attention U-Net is a continuation of the U-Net architecture. through the addition of Attention Gates (AG) in each skip connect- between the decoder and encoder tracks.

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0224
e ISSN: 2584-2854
Volume: 04 Issue: 05
May 2026
Page No: 1492 - 1497

These attention mechanisms facilitate the network to perform attention towards pertinent spatial. regions and allowing irrelevant background information to be suppressed. The Attention Gates are biased in highlighting valuable features. as the network undergoes the decoding process, thus enhancing the network. capacity to represent a vast range of spatial detail found in SAR.

3.2.Proposed System

The proposed system is based on the Pix2Pix conditional Generative Adversarial Network framework for direct image- to-image translation. In this framework, both the generator and discriminator are conditioned on the input SAR image. Within this system, the generator is responsible for converting SAR images into corresponding optical images while maintaining high structural fidelity. The encoder component extracts hierarchical features through convolutional layers and pooling operations, while the decoder reconstructs the final optical image using up sampling layers and skip connections. To enhance feature extraction capability, an Attention U-Net architecture is incorporated as an improved generator model. The Attention U-Net extends the traditional U-Net architecture by introducing Attention Gates (AG) into each skip connection between the encoder and decoder paths. These attention mechanisms enable the network to focus on relevant spatial regions while suppressing irrelevant background information. The Attention Gates selectively emphasize important features during the decoding process, thereby improving the network's ability to capture complex spatial patterns present in SAR images. This mechanism results in improved reconstruction of optical images and better preservation of structural details. As Shown in Figure 3 & 4.

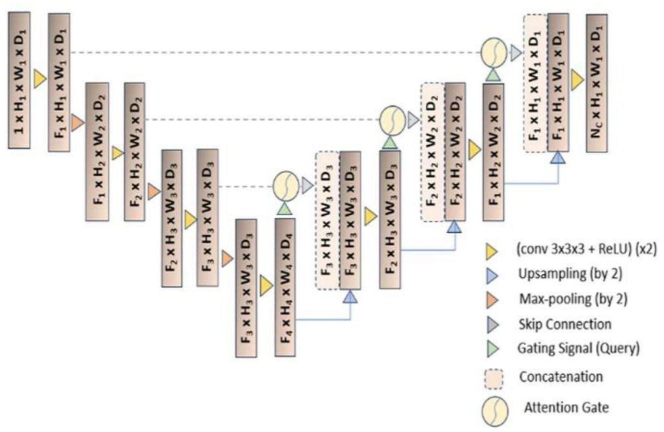

Figure 3 Architecture of Attention U-Net

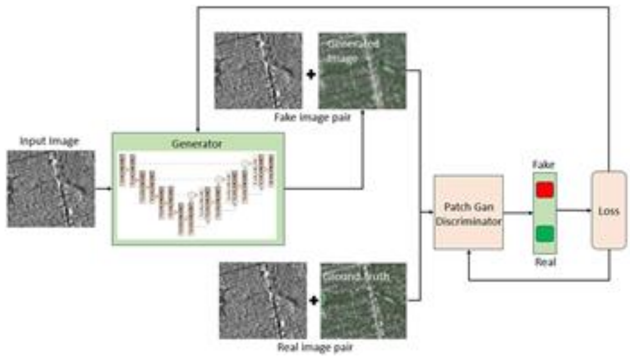

Figure 4 Pix2Pix GAN with Attention U-Net Generator

4. Results and Analysis

The performance of the proposed system was evaluated using both qualitative and quantitative analysis. During testing, SAR images were provided as input, and the generated optical images were compared with the ground truth optical images. The training process was analyzed using generator and discriminator loss curves to evaluate convergence and stability. [11]

For quantitative evaluation, two metrics were used:

- Peak Signal-to-Noise Ratio (PSNR)
- Structural Similarity Index (SSIM). As Shown in Table 1.

Table 1 Performance Comparison

Model	PSNR	SSIM
Pix2Pix + U-Net	7.03	0.0145
Pix2Pix + Attention U-Net	11.79	0.6482

The results show that the Attention U-Net model significantly improves image reconstruction quality compared to the baseline model.

4.1.Training Result

These screenshots illustrate that the model effectively generates optical images that closely approximate the ground truth optical images.

4.2.Training Loss

To evaluate the effectiveness of the proposed method, a comparison was made between the traditional Pix2Pix GAN with a standard U-Net generator and

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2026.0224

e ISSN: 2584-2854
Volume: 04 Issue: 05
May 2026
Page No: 1492 - 1497

the Pix2Pix GAN utilizing an Attention U-Net generator. During the testing phase, an unseen SAR image from the QXSLAB SAROPT dataset was employed as input, and the outputs from both models were visually compared to analyze improvements in reconstruction quality. As Shown in Figure 5,6 & 7.

Figure 5 Training result of the Pix2Pix GAN – 1

Figure 6 Training result of the Pix2Pix GAN – 2

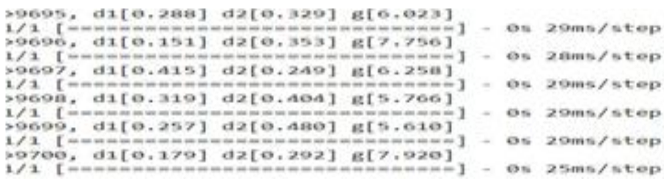

Figure 7 Training Loss

4.3.Testing Result

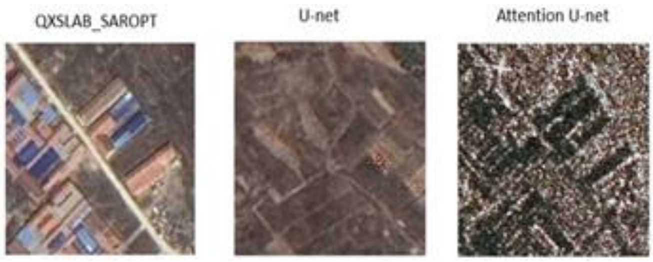

Figure 8 Test Result

For quantitative evaluation, the Structural Similarity Index (SSIM) and Peak Signal-to-Noise Ratio (PSNR) were employed. The standard U-Net achieved a PSNR of 7.0300 and SSIM of 0.0145, whereas the Attention U-Net achieved a PSNR of 11.7926 and SSIM of 0.6482. The findings indicate that the attention mechanism boosts the preservation of structural features and enhances the overall quality of the produced optical images.

Conclusion and Future Work

The paper was able to create and deploy a deep learning structure that improves the interpretability of Synthetic Aperture Radar (SAR) images through conversion into optical images that could be understood visually. The main contribution of the study is the comparative analysis of two Generative adversarial network (GAN) architectures: the basic Pix2Pix GAN with a typical U-Net generator and the improved Pix2Pix GAN with an Attention U-Net generator which are tested on the dataset of QXSLAB SAROPT. The experimental results show that the Attention U-Net model is much superior to the usual U-Net model in cross- modal image translation. The model can be effectively able to concentrate on important spatial details as well as reduce undesirable noise by means of attention gates which leads to a higher reconstruction quality and restoration of more structural information. It is shown in the performance indicators according to which the Attention U-Net has a better PSNR and SSIM, indicating a greater fidelity of image and similarity in structure than the baseline model. In the future this research can be expanded by training the model on more diverse and greater quantities of data to improve the model's strength and generalization. Other developments in architecture like multi-thread attention or advanced head attention mechanisms might enhance the quality of feature extraction and image reconstruction. Also a more efficient computational model would help to process the model faster and even achieve real-time applications in areas like remote sensing, environmental monitoring, and disaster management. Other cross-modal translation tasks such as medical imaging or other satellite imaging applications could also be adapted to the proposed approach.

References

[1]. Z. Zuo and Y. Li, "A SAR-to-optical image translation method based on PIX2PIX," in Proc. IEEE Int. Geoscience and Remote Sensing Symp. (IGARSS), 2021.
[2]. J. Bao et al., "Improved Few-Shot SAR Image Generation by Enhancing Diversity," IEEE J.

Sel. Topics Appl. Earth Observ. Remote Sens., 2024.

[3]. X. Bai, X. Pu, and F. Xu, “Conditional Diffusion for SAR to Optical Image Translation,” IEEE Geosci. Remote Sens. Lett., 2023.

[4]. P. Isola et al., “Image-to-image translation with conditional adversarial networks,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2017.

[5]. X. Yang et al., “SAR-to-optical image translation based on improved CGAN,” Pattern Recognition, vol. 121, 2022.

[6]. L. Wang et al., “SAR-to-optical image translation using supervised cycle-consistent adversarial networks,” IEEE Access, vol. 7, 2019.

[7]. H. Wang et al., “SAR-to-Optical Image Translation with Hierarchical Latent Features,” IEEE Trans. Geosci. Remote Sens., vol. 60, 2022.

[8]. J. Zhang, J. Zhou, and X. Lu, “Feature-guided SAR-to-optical image translation,” IEEE Access, vol. 8, 2020.

[9]. Y. Zhao et al., “A Comparative Analysis of GAN-based Methods for SAR-to-optical Image Translation,” IEEE Geosci. Remote Sens. Lett., vol. 19, 2022.

[10]. QXS-SAROPT Dataset. [Online]. Available: https://github.com/yaoxu008/QXSSAROPT/tree/main.

[11]. SAR2Opt Heterogeneous Dataset. [Online]. Available:https://github.com/MarsZhaoYT/SAR2Opt-Heterogeneous-Dataset.