International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 3, May-June-2026, ISSN (Online): 2395-566X

Synthetic Aperture Radar into Comprehensive Colorized Images Using Deep Learning Model

Sneha Kanawade, Dr. Suvarna Patil, Siddhi Kadu, Aniruddha Ojha,
Aryan Sahu, Indra Pratap Singh Rajawat
Artificial Intelligence and Data Science
Dr. D.Y. Patil Institute of Engineering, Management and Research, Akurdi, Pune, D. Y. Patil International University

Abstract — Synthetic Aperture Radar is vital remote sensing technology, offering all-weather, day-and- night imaging capabilities. However, its inherent grayscale nature, along with speckle noise, presents significant challenges for interpretation by non-specialists. This review addresses recent advancements in applying deep learning to SAR colorization, a technique aimed at enhancing visual interpretability of these images while preserving unique radiometric properties. The primary motivation is to bridge the gap between complex radar data, intuitive visual analysis, thereby broadening its application in fields like disaster management, environmental monitoring. Major themes covered include critical distinction between grayscale colorization, SAR-tooptical translation, evolution of methodologies from tradi-tional regression to advanced deep learning models, lack of standardized evaluation protocols that has hindered progress. Existing technologies often involve convolutional neural networks, Generative Adversarial Networks (GANs). This review high-lights a proposed methodology centered on conditional GAN within a complete benchmarking protocol utilizing synthetically generated ground truth via intensity-high saturation (IHS) fusion. Key features of this approach include an end-to-end supervised learning framework, use of domain-specific evaluation metrics (Q4, NRMSE, SAM). This advancement holds significant impli-cations for real-time disaster response, contributes to Sustainable Development Goals (SDGs) such as "Sustainable Cities and Com-munities", "Climate Action" by making critical environmental data more accessible, actionable.

Keywords— Synthetic Aperture Radar, SAR Colorization, Conditional Generative Adversarial Network, Image-to-Image Translation, Deep Learning, Remote Sensing.

I. INTRODUCTION

1. Background of SAR Imaging

Synthetic Aperture Radar (SAR) is an advanced remote sensing technology [9] that enables high-resolution imaging of the Earth's surface regardless of environmental conditions. Unlike optical imaging systems that rely on reflected sunlight, SAR operates using microwave signals [9], allowing it to capture images during both day and night, and under adverse weather conditions such as cloud cover, rain, or fog.

SAR systems function by transmitting electromagnetic waves [23] toward the earth surface and measuring the backscattered signals. The intensity and phase of these re-turned signals provide valuable information about surface roughness, moisture content, and structural characteristics. As a result, SAR imagery plays a crucial role in applications such as disaster monitoring, environmental analysis, military surveillance, and infrastructure assessment.

However, SAR images are inherently different from optical images. They are typically represented in grayscale, where pixel intensity corresponds to radar backscatter rather than visible light reflection. Additionally, SAR data is affected by speckle noise [10], [24], a granular interference pattern caused by the coherent nature of radar signals. While this noise contains useful statistical information, it makes visual interpretation challenging, especially for non-expert users.

With the increasing availability of satellite data from plat-forms such as Sentinel-1, SAR imagery has become more accessible. However, its limited visual interpretability remains a significant barrier to widespread adoption outside specialized domains. This gap has motivated the exploration of advanced computational techniques, particularly deep learning [11], [12], to enhance SAR image usability by transforming it into more intuitive visual representations.

2. Problem Statement

Despite the significant advantages of Synthetic Aperture Radar (SAR) in providing reliable, all-weather and day–night imaging, its practical usability remains limited due to inher-ent visualization constraints. SAR images are fundamentally grayscale [10] and represent radar backscatter rather than natural color information, making them difficult to interpret, especially for users without domain expertise.

The absence of intuitive visual cues such as color differentiation restricts the ability to quickly identify land-cover types, structural patterns, and environmental changes. Additionally, the presence of speckle noise further complicates perception by introducing granular distortions that obscure fine details. As a result, tasks that require rapid decision-making—such as disaster response, urban planning, and environmental monitoring—become more challenging when relying solely on raw SAR data.

Traditional approaches to improving SAR interpretability often involve manual analysis or handcrafted enhancement techniques. However, these methods either require expert knowledge or fail to produce visually meaningful outputs. Furthermore, SAR-to-optical translation techniques [3], [16], while capable of generating realistic images, may alter or suppress important radar-specific characteristics, leading to a loss of scientific reliability.

Therefore, the core problem addressed in this research is the development of an automated, data-driven approach that can transform grayscale SAR images into visually in-terpretable color representations while preserving essential spatial structures and radar-specific properties. The challenge lies in achieving a balance between visual realism and physical fidelity, ensuring that the generated outputs are both meaningful for human interpretation and consistent with the underlying SAR data.

3. Motivation

The growing reliance on remote sensing technologies for critical applications has highlighted the importance of making complex data more accessible and interpretable. Synthetic Aperture Radar (SAR), despite its robustness and versatility, remains underutilized outside expert communities due to its unintuitive grayscale representation. This creates a significant gap between the availability of valuable SAR data and its effective use in real-world decision-making scenarios.

In many applications such as disaster management, environmental monitoring, and urban development, rapid interpretation of satellite imagery is essential. Non-specialists, includ-ing policymakers, emergency responders, and field operators, often depend on visually intuitive information to make timely decisions. However, SAR imagery lacks the natural color cues that are commonly associated with optical images, making it difficult for such users to extract meaningful insights quickly. Advancements in deep learning, particularly in image-to-image translation [2], provide a promising opportunity to address this challenge. Conditional Generative Adversarial Networks (cGANs) [15] have demonstrated the ability to learn complex mappings between different image domains, enabling the transformation of grayscale inputs into realistic color outputs. By leveraging these techniques, it becomes possible to enhance the perceptual quality of SAR images while retaining their underlying structural and radiometric information.

The motivation behind this work is to bridge the gap between complex radar data and intuitive visual understand-ing. By developing a deep learning-based SAR colorization framework, this research aims to improve accessibility, reduce interpretation time, and expand the applicability of SAR imagery across various domains. Ultimately, the goal is to enable a broader range of users to effectively utilize SAR data without requiring specialized expertise.

4. Challenges in SAR Colorization

Colorizing Synthetic Aperture Radar (SAR) images is a fundamentally complex task due to the intrinsic differences between radar and optical imaging modalities. Unlike natural images, SAR data does not capture color information; instead, it represents physical properties such as surface roughness, dielectric constant, and geometric structure through backscatter intensity. This makes the process of assigning meaningful colors to SAR images inherently ambiguous.

One of the primary challenges lies in the ill-posed nature of the problem. A single SAR image can correspond to multiple plausible color interpretations, as there is no direct one-to-one mapping between radar signals and visible color space. This uncertainty makes it difficult for models to generate consistent and accurate color representations.

Another significant challenge is the presence of speckle noise [10], which introduces granular distortions [24] across the image. While speckle contains useful statistical informa-tion, it degrades visual quality and complicates the learning process for deep models. Models must learn to distinguish between meaningful structural patterns and noise without losing important radar-specific details.

A critical issue in existing approaches is the trade-off between visual realism and physical fidelity. Techniques that focus heavily on generating visually appealing outputs often resemble optical images but fail to preserve essential SAR characteristics such as texture and scattering behavior. Conversely, methods that prioritize radar fidelity may produce outputs that remain difficult to interpret visually.

The dependence on paired datasets also presents a major limitation. Supervised learning approaches require accurately aligned SAR and optical image pairs, which are difficult to obtain at large scales due to differences in sensor characteristics, acquisition times, and environmental conditions. Misalignment between pairs can significantly affect model performance.

Additionally, there is a lack of standardized evaluation metrics specifically designed for SAR colorization. Common image quality metrics such as PSNR and SSIM [8], [18] do not fully capture the preservation of radar-specific features, making it challenging to assess model performance compre-hensively.
Deep learning models used for this task often operate as black boxes, reducing interpretability and trust in the generated outputs. This becomes particularly important in high-stakes applications such as disaster response, where reliability and transparency are critical.

Addressing these challenges requires a carefully designed framework that balances perceptual quality with scientific accuracy, while also ensuring robustness, scalability, and interpretability.

5. Contributions of This Work
This research presents a comprehensive approach to enhanc-ing the interpretability of Synthetic Aperture Radar (SAR) imagery through deep learning-based colorization. The work is designed not only as a model development effort but as a complete system that integrates data processing, model training, evaluation, and real-world deployment.

The primary contribution of this work is the develop-ment of an end-to-end SAR colorization framework based on Conditional Generative Adversarial Networks (cGANs). The proposed approach leverages a Pix2Pix-based architecture [2], combining a U-Net generator [4] and a PatchGAN discriminator [2], to learn complex mappings between grayscale SAR inputs and corresponding color representations. This enables the generation of visually meaningful outputs while preserving essential spatial structures.

A significant contribution lies in the integration of a real-world paired dataset using Sentinel-1 and Sentinel-2 image pairs. The dataset is systematically divided into training, validation, and testing splits to ensure balanced learning and reliable evaluation. This structured dataset design supports reproducibility and allows consistent benchmarking of model performance.

Another key contribution is the implementation of a com-plete system pipeline (SAR-RANG platform) that bridges research and practical application. The system integrates fron-tend user interaction, backend processing, and cloud-based model deployment using modern technologies. This enables users to upload SAR images and receive colorized outputs in real time, demonstrating the practical viability of the proposed approach. The work also includes a comprehensive evaluation strategy, combining both qualitative and quantitative analysis. While visual comparisons highlight improvements in interpretability across different land-cover types, quantitative metrics such as FID, PSNR, and SSIM [8] are used to assess model performance. Additionally, training behavior is analyzed through loss curves to understand convergence and stability.

Furthermore, this research emphasizes the distinction be-tween SAR colorization and SAR-to-optical translation, en-suring that the generated outputs maintain radar-specific characteristics rather than simply mimicking optical imagery. This preserves the scientific integrity of the data while enhancing its usability.

Finally, the study identifies key limitations and challenges in current approaches and provides insights into future improvements, including the adoption of advanced architectures and improved evaluation methodologies.

II. RELATED WORK

1. Traditional SAR Interpretation Methods
Before the emergence of deep learning techniques, the interpretation of Synthetic Aperture Radar (SAR) imagery primarily relied on domain-specific knowledge and classical signal processing approaches. These traditional methods focused on understanding the physical principles of radar backscatter and translating them into meaningful visual or analytical insights. One of the most widely used approaches is manual visual interpretation, where experts analyze SAR images based on intensity variations, texture patterns, and contextual information. Different land-cover types exhibit distinct backscatter characteristics for example, urban areas typically produce strong reflections due to double-bounce scattering, while water bodies appear dark due to specular reflection. Although effective, this method requires significant expertise and is not scalable for large datasets or real-time applications.

Another common technique involves polarimetric analysis, where multiple polarization channels (such as HH, HV, VV) are used to derive additional information about surface prop-erties. By analyzing scattering mechanisms, polarimetric SAR can provide insights into vegetation structure, soil moisture, and surface roughness. However, this approach often requires specialized sensors and complex mathematical modeling, limiting its accessibility.

International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 3, May-June-2026, ISSN (Online): 2395-566X

Speckle reduction and filtering methods have also been extensively used to improve SAR image quality. Techniques such as Lee filtering, Frost filtering, and wavelet-based denoising aim to reduce speckle noise while preserving important features. While these methods enhance visual clarity, they do not fundamentally improve interpretability in terms of semantic understanding or color representation.

In addition, feature-based classification methods have been applied to SAR imagery. These include texture analysis, edge detection, and statistical modeling, followed by traditional machine learning classifiers such as Support Vector Machines (SVMs) or Random Forests. Although these approaches can automate certain tasks, they rely heavily on handcrafted features and often struggle to generalize across diverse datasets. Some early attempts at enhancing visualization involved rule-based color mapping, where different intensity ranges or polarimetric features were assigned specific colors. While this provided a basic level of visual enhancement, the resulting images lacked realism and often did not correspond to nat-ural color distributions, limiting their usefulness for intuitive interpretation.

Traditional SAR interpretation methods have played a cru-cial role in advancing remote sensing applications. However, they are limited by their dependence on expert knowledge, lack of scalability, and inability to produce visually intuitive out-puts. These limitations have driven the transition toward data-driven approaches, particularly deep learning-based models, which aim to automate and enhance SAR image interpretation.

2. Deep Learning in SAR Colorization

The rapid advancement of deep learning has significantly transformed the field of remote sensing [25], particularly in tasks involving image interpretation and enhancement. In the context of Synthetic Aperture Radar (SAR), deep learning methods have emerged as powerful tools for overcoming the limitations of traditional approaches by enabling automated, data-driven learning of complex patterns and mappings.

Early applications of deep learning in SAR focused on feature extraction and classification, where Convolutional Neu-ral Networks (CNNs) were used [5]–[7] to learn hierarchical representations of radar data. These models demonstrated improved performance over traditional feature-engineering methods by capturing spatial patterns and textures directly from the data. However, CNN-based regression models for SAR colorization often produced blurred and desaturated outputs, as they optimized pixel-wise loss functions that fail to capture perceptual quality.

To address these limitations, researchers began exploring image-to-image translation frameworks, where the task is formulated as learning a mapping from SAR images to corresponding color or optical images. Encoder–decoder architectures [4], [27], particularly those based on U-Net, became popular due to their ability to preserve spatial information through skip connections. These models improved structural consistency but still struggled with generating realistic color distributions.

The introduction of Generative Adversarial Networks (GANs) marked a significant breakthrough [1] in SAR col-orization. GAN-based models consist of a generator that produces synthetic images and a discriminator that evaluates their realism. This adversarial training process encourages the generator to produce sharper and more visually convincing outputs compared to traditional regression-based methods.

Among these, Conditional GANs (cGANs) [15] have proven especially effective for SAR colorization. By conditioning both the generator and discriminator on the input SAR image, cGANs learn a structured mapping between input and output domains. Architectures such as Pix2Pix [2] utilize a U-Net generator combined with a PatchGAN discriminator, enabling the model to capture both global structure and local texture details. This approach has demonstrated superior performance in producing realistic and structurally consistent colorized images.

Recent advancements [21] have further extended these mod-els by incorporating attention mechanisms, multi-scale feature extraction, and hybrid architectures. These improvements aim to enhance the preservation of fine-grained details, reduce noise artifacts, and improve generalization across different datasets and sensor modalities. Additionally, emerging ap-proaches such as diffusion models are being explored [22] for their potential to generate high-quality and stable outputs. Despite these advancements, challenges remain in terms of dataset availability, evaluation standardization, and balancing visual realism with radar-specific fidelity. Nevertheless, deep learning, particularly GAN-based frameworks, has established itself as the most promising direction for advancing SAR colorization and making radar imagery more accessible and interpretable.

3. GAN-based Image Translation (Pix2Pix, CycleGAN)

Generative Adversarial Networks (GANs) have become a foundational approach for image-to-image translation tasks, enabling the transformation of images from one domain to another while preserving structural consistency. In the context of SAR colorization, GAN-based models provide a powerful

International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 3, May-June-2026, ISSN (Online): 2395-566X

framework for learning complex mappings between grayscale radar images and their corresponding color representations.
One of the most influential models in this domain is Pix2Pix, a supervised image-to-image translation framework based on Conditional GANs (cGANs). Pix2Pix operates on paired datasets, where each input image has a corresponding ground truth output. The model consists of two main compo-nents: a U-Net-based generator and a PatchGAN discriminator [2], [20]. The generator learns to produce colorized outputs from SAR inputs, while the discriminator evaluates whether the generated images are realistic when compared to true optical images.

A key strength of Pix2Pix lies in its combined loss function, which integrates adversarial loss with a reconstruction loss (typically L1). The adversarial component encourages realism, while the reconstruction loss ensures structural similarity with the target image. This balance allows Pix2Pix to generate outputs that are both visually sharp and spatially consistent, making it well-suited for SAR colorization tasks where align-ment between input and output is available.

In contrast, CycleGAN is designed for unpaired image-to-image translation [3], [26], where corresponding input-output pairs are not required. Instead of relying on direct supervision, CycleGAN introduces a cycle consistency mechanism, ensur-ing that an image translated from one domain to another can be mapped back to its original form. This approach is particularly useful when paired SAR-optical datasets are unavailable or difficult to obtain.

However, CycleGAN has certain limitations in the context of SAR colorization. Since it does not enforce pixel-level alignment, the generated outputs may lack precise structural correspondence with the input SAR image. This can lead to inconsistencies in spatial features and reduced reliability for applications that require accurate representation of underlying radar characteristics.

Comparatively, Pix2Pix offers better performance in sce-narios where paired datasets are available, as it directly learns the mapping between SAR and optical domains with strong supervision. It produces more stable and structurally aligned outputs, making it more suitable for applications that demand both visual realism and physical consistency.

Both approaches have contributed significantly to the ad-vancement of image translation tasks. While CycleGAN pro-vides flexibility in handling unpaired data, Pix2Pix remains the preferred choice for SAR colorization when high-quality paired datasets are accessible. This research builds upon the Pix2Pix framework, leveraging its strengths to develop a robust and reliable SAR colorization system.

4. SAR-to-Optical vs SAR Colorization (Important Distinc-tion)
A critical aspect in the field of SAR image enhancement is the distinction between SAR-to-optical translation [16], [17] and SAR colorization, as both approaches aim to improve visual interpretability but differ fundamentally in objectives, methodology, and outcomes.

SAR-to-optical translation focuses on generating images that closely resemble true optical imagery. The goal is to reconstruct how a scene would appear in the visible spectrum by learning mappings between SAR data and corresponding optical images. This approach often prioritizes visual realism and semantic similarity to optical data, making the outputs appear natural and intuitive. However, in doing so, it may suppress or alter important radar-specific characteristics, such as speckle patterns, scattering behavior, and structural nuances unique to SAR imaging. As a result, while the generated images may look realistic, they may not accurately reflect the underlying physical properties captured by radar sensors.

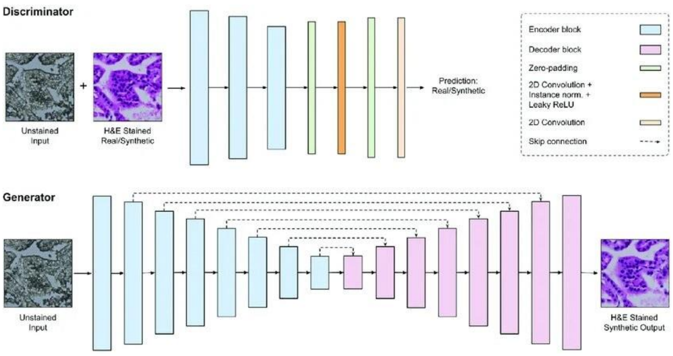

Fig. 1. Detailed architecture of the proposed GAN-based SAR colorization model. The generator utilizes an encoder–decoder structure with skip connections, while the discriminator evaluates real and synthetic samples to guide the learning process.

In contrast, SAR colorization aims to enhance the inter-pretability of SAR images without fundamentally transforming them into optical equivalents. Instead of recreating optical imagery, the objective is to assign meaningful color represen-tations to SAR data while preserving its intrinsic structural and radiometric properties. This approach emphasizes maintaining the authenticity of radar information while improving visual perception through color cues.

The distinction becomes particularly important in applica-tions where data fidelity is critical, such as disaster assessment, environmental monitoring, and scientific analysis. In such scenarios, preserving radar-specific features ensures that the information remains reliable and physically consistent. Overly realistic optical-like outputs may lead to misinterpretation or loss of critical insights derived from SAR data.

From a methodological perspective, SAR-to-optical translation often relies on models that prioritize perceptual similarity to optical images, whereas SAR colorization frameworks are designed to balance visual enhancement with structural preservation. This balance is typically achieved through carefully designed loss functions and architectural choices that retain spatial alignment and radar characteristics.

In this research, the focus is explicitly on SAR colorization rather than SAR-to-optical translation. The proposed approach aims to generate visually interpretable outputs while ensuring that the essential information contained in SAR imagery is not compromised. This distinction forms a key foundation of the work and guides the design of the model, dataset usage, and evaluation strategy.

5. Identified Research Gaps

Despite significant advancements in SAR image processing and deep learning-based colorization, several critical research gaps remain that limit the effectiveness, scalability, and reliability of existing approaches.

One of the primary gaps is the ambiguity in problem definition, particularly the confusion between SAR coloriza-tion and SAR-to-optical translation. Many existing studies do not clearly distinguish between enhancing interpretability and reconstructing optical imagery.

Another major limitation is the lack of standardized datasets and benchmarking protocols. Differences in data sources and preprocessing make comparison difficult.

The inadequacy of evaluation metrics is also a concern. Metrics like PSNR and SSIM do not fully capture SAR-specific features.

Dependence on paired datasets limits scalability. Lack of interpretability reduces trust in outputs.

Another gap is the limited focus on real-world deployment.

Finally, balancing realism and fidelity remains an open challenge.

III. DATASET AND PREPROCESSING

1. Dataset Description (Sentinel-1 & Sentinel-2)

The effectiveness of any deep learning-based image translation model is highly dependent on the quality, diversity, and structure of the dataset used for training. In this work, a paired dataset consisting of Synthetic Aperture Radar (SAR) and corresponding optical images is utilized to enable supervised learning for SAR colorization. The dataset is derived from Sentinel-1 and Sentinel-2 [9] satellite missions, which are part of the Copernicus Earth observation program. Sentinel-1 provides SAR imagery captured using microwave sensors, while Sentinel-2 offers high-resolution optical images in multiple spectral bands. The complementary nature of these two data sources makes them ideal for learning mappings between radar and visual domains.

The dataset used in this study is the Sentinel-1 & Sentinel-2 Image Pairs dataset, curated by Michael Schmitt from the Technical University of Munich (TUM) and made publicly available through Kaggle by Paritosh Tiwari. It consists of co-registered SAR and optical image pairs, ensuring spatial alignment between the input (SAR) and target (RGB) images. This alignment is crucial for supervised training, as it allows the model to learn pixel-level correspondences between the two modalities.

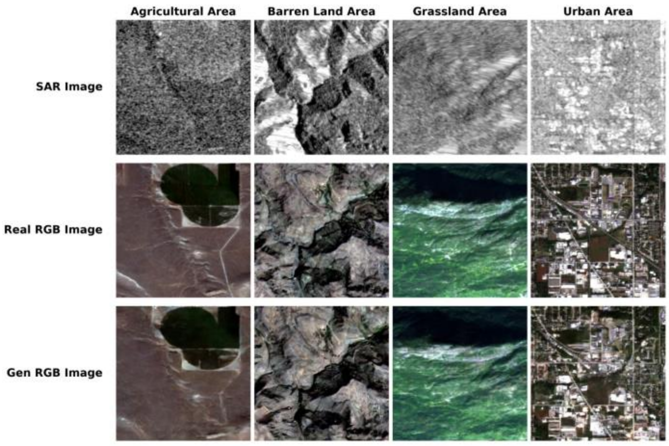

Fig. 2. Qualitative comparison across different land-cover types: SAR input (top row), real RGB images (middle row), and generated RGB outputs (bottom row), including agricultural, barren land, grassland, and urban areas.

Each image pair represents diverse geographical regions and land-cover types, including urban areas, agricultural fields, water bodies, and barren land. This diversity enhances the

International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 3, May-June-2026, ISSN (Online): 2395-566X

model's ability to generalize across different environmental conditions and surface characteristics.

The dataset is organized into three distinct subsets: training, validation, and testing. A total of 16,000 image pairs are utilized, with 12,800 pairs allocated for training, and 1,600 pairs each for validation and testing. The splits are designed to maintain similar category distributions, ensuring consistency in model evaluation and reducing bias.

All images are preprocessed to ensure uniformity in size, format, and alignment. The SAR images are provided as grayscale inputs, while the optical images serve as the ground truth RGB targets. This paired structure enables the application of supervised learning techniques, particularly the Pix2Pix framework, for effective SAR colorization.

The use of a well-structured and diverse paired dataset forms the foundation for training a robust model capable

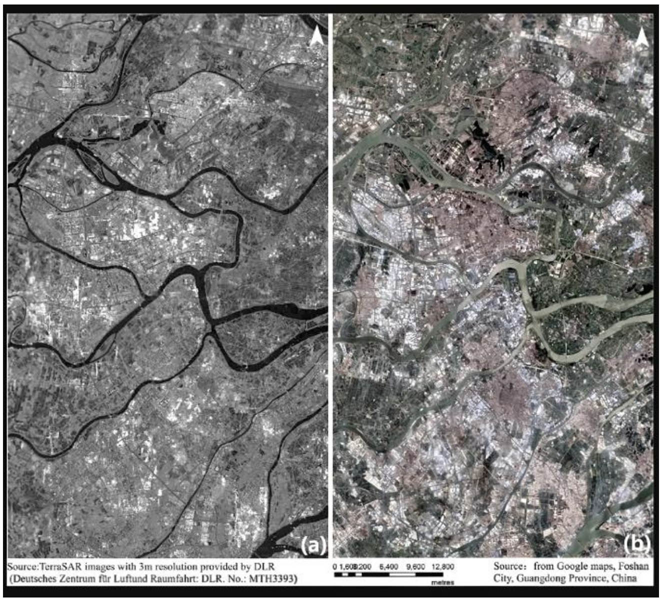

Fig. 3. Comparison between (a) Synthetic Aperture Radar (SAR) image and

corresponding optical image. of generating meaningful and interpretable colorized SAR images.

2. Data Splitting Strategy
A well-defined data splitting strategy is essential to ensure reliable model training, unbiased evaluation, and generalization to unseen data. In this work, the dataset is systematically divided into three subsets: training, validation, and testing, following a structured and balanced approach.

Out of the total 16,000 paired SAR–optical images, the dataset is split as follows:
- Training Set: 12,800 image pairs
- Validation Set: 1,600 image pairs
- Test Set: 1,600 image pairs

The training set constitutes the majority of the data and is used to learn the mapping between SAR inputs and corre-sponding RGB outputs. This large proportion allows the model to capture diverse spatial patterns, textures, and land-cover variations present in the dataset.

The validation set is used during training to monitor model performance and prevent overfitting. By evaluating the model on unseen data after each training phase, hyperparameters such as learning rate, number of epochs, and loss balancing can be adjusted effectively. This ensures that the model maintains a balance between learning and generalization.

The test set is reserved exclusively for final evaluation and is not used during training or validation. This ensures an unbiased assessment of the model's performance on completely unseen data, providing a realistic measure of its generalization capability.

To maintain consistency and fairness across all subsets, the dataset is split such that each subset contains a similar distribution of land-cover categories, including urban regions, vegetation, water bodies, and barren land. This balanced distribution minimizes bias and ensures that the model is evaluated across diverse scenarios.

The splitting is performed using predefined identifiers stored in a structured configuration (e.g., split.json), ensuring reproducibility of experiments. This standardized approach allows consistent comparison of results and supports future extensions of the work.

3. Data Preprocessing
Data preprocessing is a crucial step in ensuring that the input data is consistent, noise-resilient, and suitable for effective model training. Given the inherent differences between SAR and optical imagery, careful preprocessing is required to align both modalities and enhance the learning capability of the model.

The first step involves image resizing and normalization. All SAR and optical images are resized to a fixed resolution compatible with the model architecture. This ensures uniformity across the dataset and allows batch processing during training. Pixel values are normalized to a standard range,

International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 3, May-June-2026, ISSN (Online): 2395-566X

typically between 0 and 1, to stabilize training and improve convergence.

Next, spatial alignment (co-registration) is ensured between SAR and optical image pairs. Since the dataset is derived from Sentinel-1 and Sentinel-2 sources, slight spatial misalignments may exist due to differences in acquisition time, sensor characteristics, and viewing geometry. Proper alignment is critical, as the model relies on pixel-level correspondence to learn accurate mappings between input and output images.

The SAR images, which are originally in grayscale for-mat, are processed to enhance their compatibility with deep training dataset. Augmentation artificially increases the diver-sity of the data without requiring additional data collection, allowing the model to learn invariant features across different transformations and environmental variations.

One of the primary techniques used is geometric transformation, including horizontal flipping, vertical flipping, and random rotations. These transformations simulate variations in orientation and viewpoint, enabling the model to learn spatial patterns independent of image alignment.

Another important technique is random cropping and resiz-ing, which introduces variability in spatial regions. By training on different cropped sections of the same image, the model learns to focus on both local and global features.

Scaling and translation are also applied to simulate slight shifts in image positioning.

In addition, intensity-based augmentations can be applied selectively to SAR images.

Importantly, all augmentation operations are applied consistently to both the SAR input and its corresponding optical target to maintain alignment.

By incorporating these augmentation techniques, the model becomes more resilient to variations in real-world data.

IV. METHODOLOGY

1. Problem Formulation (Image-to-Image Translation)

The task of SAR image colorization can be formally framed as an image-to-image translation [2] problem, where the objective is to learn a mapping between two different visual domains: the grayscale SAR domain and the RGB optical domain.

Let X denote the domain of input SAR images and Y denote the domain of corresponding optical (RGB) images Given a paired dataset of aligned samples {(xi, yi)}Ngoal is to learn a mapping function:

$$G : X \to Y \quad (1)$$

such that the generated output G(x) is both visually realistic and structurally consistent with the ground truth y.

2. Conditional GAN Overview

A standard GAN consists of two competing neural net-works: a generator G and a discriminator D [1].
The generator learns a mapping:

$$G(x, z) \to y \quad (2)$$

The discriminator evaluates:

$$D(x, y) \quad (3)$$

The objective function is:

$$\mathcal{L}_{cGAN}(G, D) = \mathbb{E}_{x,y}[\log D(x, y)] + \mathbb{E}_{x,z}[\log(1 - D(x, G(x, z)))] \quad (4)$$

$$\mathcal{L}_{L1}(G) = \mathbb{E}_{x,y,z}[\|y - G(x, z)\|_1] \quad (5)$$

$$G^* = \arg\min_G \max_D \ \mathcal{L}_{cGAN}(G, D) + \lambda \mathcal{L}_{L1}(G) \quad (6)$$

3. Objective Function

The model employs a composite objective function consist-ing of adversarial loss and reconstruction loss (L1 loss).

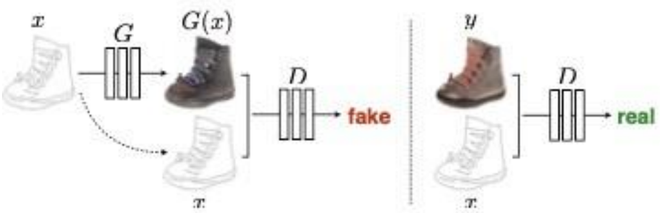

Fig. 4. Training process of a Conditional GAN (cGAN) [15] , where the generator produces synthetic outputs and the discriminator distinguishes between real and generated image pairs.

The adversarial loss encourages realism, while the L1 loss ensures structural similarity between generated outputs and ground truth images.

International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 3, May-June-2026, ISSN (Online): 2395-566X

This combination balances perceptual quality and fidelity.

4. Generator Architecture (U-Net)
The generator plays a central role in the SAR colorization framework, as it is responsible for transforming grayscale SAR images into meaningful RGB representations. In this work, a U-Net-based architecture is adopted for the generator due to its proven effectiveness in image-to-image translation tasks.
The U-Net architecture [4] follows an encoder–decoder structure with symmetric layers. The encoder progressively downsamples the input image, extracting high-level feature representations, while the decoder upsamples these features to reconstruct the output image at the original resolution.

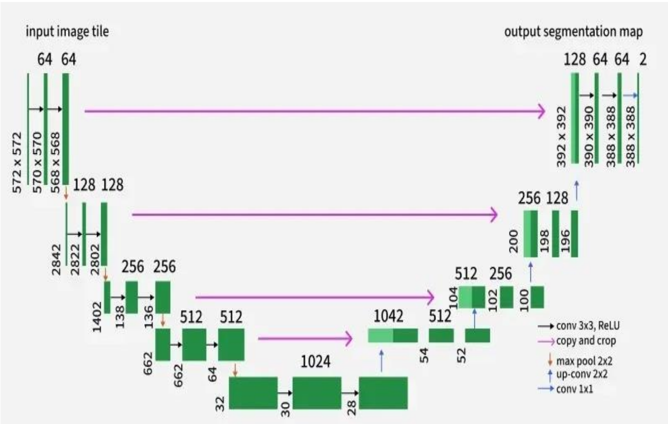

Fig. 5. U-Net architecture used as the generator in the proposed model. The network consists of an encoder–decoder structure with skip connections that preserve spatial information during reconstruction.

A key feature of the U-Net architecture is the use of skip connections between corresponding layers in the encoder and decoder. These connections allow low-level spatial information, such as edges and textures, to bypass the bottleneck and be directly transferred to the decoding stages.

5. Discriminator Architecture (PatchGAN)
The discriminator is designed using a PatchGAN architec-ture. Instead of classifying the entire image, PatchGAN evaluates smaller patches and determines whether each patch is real or fake.

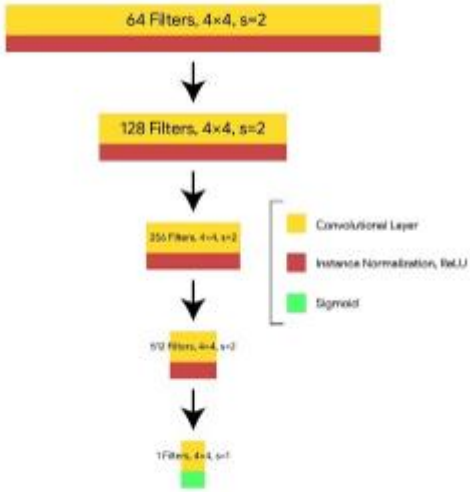

Fig. 6. Patch GAN discriminator architecture used in the proposed model. The network classifies image patches as real or fake, enabling fine-grained texture learning.

Advantages
- Captures fine-grained textures
- Improves sharpness
- Efficient computation

The discriminator takes SAR input and generated/real out-put as pairs and evaluates them.

6. Training Strategy
The training process follows adversarial learning.
- Train discriminator on real vs fake images
- Train generator to fool discriminator
- Optimizer: Adam

Training is done over multiple epochs with validation mon-itoring.

Stability techniques include:
- Batch normalization
- Learning rate tuning
- Balanced updates

V. SYSTEM ARCHITECTURE AND IMPLEMENTATION

1. Overall System Architecture
The proposed SAR colorization system is designed as an end-to-end integrated pipeline that seamlessly connects user interaction, data processing, model inference, and result visu-

alization. The architecture combines modern web technologies with deep learning-based image translation to deliver a practical and scalable solution.

At a high level, the system follows a client–server archi-tecture, where the frontend handles user interaction and the backend manages processing and communication with the deployed deep learning model. The overall workflow begins with user input and progresses through multiple stages until the final colorized output is generated and displayed.

The system consists of the following key components:

- User Interface Layer (Frontend) The frontend serves as the interaction point between the user and the system. It allows users to upload grayscale SAR images, view results, and download the generated outputs. The interface is designed to be intuitive and responsive, ensuring accessibility across different devices.
- Image Processing Layer Once an image is uploaded, it undergoes validation and preprocessing. This includes check-ing file format, resizing the image to the required input dimensions, normalizing pixel values, and converting it into a format suitable for model inference. This layer ensures that all inputs are standardized before being passed to the model.
- Backend Processing Layer The backend acts as an intermediary between the frontend and the model. It receives the processed image, encodes it into an appropriate format (such as Base64 or tensor representation), and forwards it to the model inference service. It also handles API communication, error management, and response formatting.
- Model Inference Layer This layer hosts the trained Pix2Pix Conditional GAN model. The model processes the input SAR image and generates a corresponding colorized output. The inference is performed on a cloud-based platform, enabling scalability and efficient computation.
- Output Visualization Layer After inference, the generated image is returned to the frontend, where it is displayed alongside the original SAR input. This allows users to visually compare the transformation. The system also provides an option to download the output image for further use.

2. Frontend–Backend Pipeline

The SAR colorization system follows a well-structured frontend–backend pipeline that ensures smooth communication between user interaction and model inference. This pipeline is designed to handle data flow efficiently while maintaining low latency and high reliability.

The process begins at the frontend, where the user uploads a grayscale SAR image through a web-based interface. Once the image is selected, the frontend performs initial validation checks, including file format and size verification. After validation, the image is prepared for transmission by converting it into a suitable format, such as Base64 encoding or binary payload.

The processed input is then sent to the backend server through secure HTTP requests. The backend, implemented using cloud-based services, acts as the central processing unit of the system. Upon receiving the request, it performs additional validation and prepares the data for model inference. This includes formatting the input to match the requirements of the deployed deep learning model.

The backend then forwards the request to the model in-ference service, which is hosted on a cloud platform. This communication is handled through API calls, ensuring that the frontend remains lightweight and does not directly interact with the model. The inference service processes the input im-age using the trained Pix2Pix model and generates a colorized output.

Once the model completes inference, the generated image is returned to the backend. The backend then converts the output into a format suitable for display, ensuring compatibility with the frontend interface. This may include decoding the image and attaching necessary metadata.

Finally, the processed output is sent back to the frontend, where it is displayed alongside the original input image. The user can visually compare both images and optionally download the result.

This pipeline ensures a clear separation of concerns:

- The frontend focuses on user interaction and visualization
- The backend manages data processing and communication
- The inference service handles computational tasks

Such a design enhances system scalability, security, and maintainability. It also allows independent upgrades of each component without affecting the overall system functionality.

3. Model Deployment (Hugging Face)

To enable real-time inference and scalable access, the trained SAR colorization model is deployed using a cloud-based inference service. In this work, the Pix2Pix Conditional GAN model is hosted on the Hugging Face platform, which provides

International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 3, May-June-2026, ISSN (Online): 2395-566X

a reliable and efficient environment for serving ma-chine learning models.

The trained model weights, including both the generator and discriminator components, are exported and uploaded to the Hugging Face Model Hub. This allows the model to be accessed through a dedicated inference endpoint, which acts as an API for processing incoming requests.

The deployment process involves configuring the model to accept input images in a predefined format and return the corresponding colorized outputs. The inference endpoint is designed to handle requests asynchronously, ensuring that multiple users can interact with the system simultaneously without significant performance degradation.

One of the key advantages of using Hugging Face is its scalability and ease of integration. The platform manages infrastructure requirements such as GPU allocation, request handling, and load balancing, allowing the system to focus

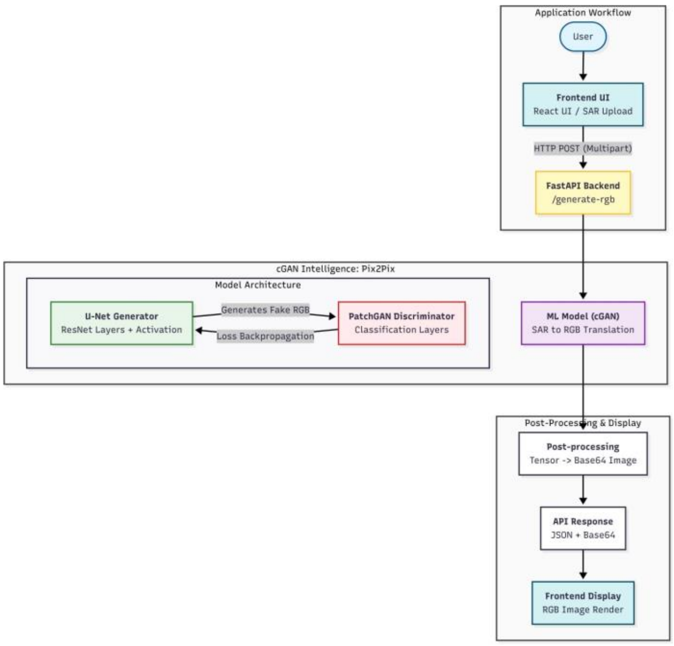

Fig. 7. Overall system architecture of the proposed SAR colorization platform, illustrating the interaction between frontend, backend, workflow orchestration, model inference, and data storage components.

on functionality rather than resource management. This is particularly important for deep learning models, which require substantial computational power for inference.

The backend communicates with the Hugging Face endpoint through secure API calls. When a request is received, the backend forwards the processed SAR image to the endpoint, which performs model inference and returns the generated RGB image. The response is then relayed back to the frontend for visualization.

Additionally, version control of model files is maintained on the platform, enabling easy updates and experimentation with improved models. This allows the system to evolve over time by incorporating enhancements without disrupting the existing pipeline.

4. API Communication Flow

The communication between different components of the SAR colorization system is facilitated through a structured and secure API communication flow. This flow ensures that data is transmitted efficiently between the frontend, backend, and model inference service while maintaining reliability and security.

The process begins when the frontend sends an HTTP re-quest containing the preprocessed SAR image to the backend. This request typically includes the image data encoded in a suitable format, along with any necessary metadata. The communication is carried out over secure protocols to ensure data integrity and prevent unauthorized access.

Upon receiving the request, the backend processes the input and prepares it for inference. It constructs a new API request directed toward the model inference endpoint hosted on the cloud platform. This request includes the encoded image and any required configuration parameters for the model.

The backend then sends this request to the inference API, which triggers the execution of the deployed Pix2Pix model. The inference service processes the input and generates the corresponding colorized output. Once the computation is com-plete, the result is returned to the backend as a response.

After receiving the output, the backend performs necessary post-processing steps, such as decoding the image and en-suring compatibility with frontend display requirements. The processed result is then packaged into a response and sent back to the frontend.

The frontend receives the response and renders the output image alongside the original SAR input. This completes the communication cycle, providing the user with the final result.
To ensure robustness, the API communication flow includes mechanisms for:

- Error handling, to manage invalid inputs or failed requests
- Timeout management, to handle delays in model inference
- Data validation, to ensure correct formatting of requests and responses

This structured API-based approach provides flexibility and modularity, allowing each component of the system to operate independently while maintaining seamless interaction. It also enables easy integration with external services and supports scalability for handling multiple concurrent users.

The API communication flow is a critical component of the system, ensuring efficient and reliable interaction between user inputs and deep learning-based processing.

5. SAR-RANG Platform Overview
The SAR-RANG platform represents the practical implementation of the proposed SAR colorization framework, combining deep learning capabilities with a user-friendly web-based interface. It is designed to bridge the gap between advanced research models and real-world usability, enabling users to interact with SAR data in an intuitive and efficient manner.

At its core, SAR-RANG is a web-based application that allows users to upload grayscale SAR images and obtain colorized outputs in real time. The platform integrates multiple technologies, including a modern frontend interface, cloud-based backend services, and a deployed Pix2Pix Conditional GAN model.

The user interface is designed with a focus on simplicity and accessibility, allowing users with minimal technical ex-pertise to interact with the system easily. Users can upload images, view side-by-side comparisons of input and output, and download the generated results. The platform also includes additional sections such as example galleries, informational content, and contact features to enhance user experience.

From a system perspective, SAR-RANG follows a modular architecture, where each component—frontend, backend, and model inference—operates independently while maintaining seamless integration. This modular design ensures flexibility, allowing the system to be easily extended with new features or improved models.

The platform leverages cloud-based deployment for scala-bility and performance. By hosting the model on an external inference service, SAR-RANG ensures that computationally intensive tasks are handled efficiently without overloading the frontend or backend systems. This enables the platform to support multiple users simultaneously while maintaining consistent response times.

Another key aspect of SAR-RANG is its role as a demonstration of real-world applicability. Unlike purely theoretical models, this platform showcases how deep learning techniques can be integrated into a complete system that delivers tangi-ble value. It highlights the potential of SAR colorization in practical scenarios such as disaster monitoring, environmental analysis, and infrastructure assessment.

VI. EXPERIMENTAL SETUP

1. Hardware and Software Configuration
The performance and efficiency of the proposed SAR colorization model are influenced by the computational environment used during training and deployment. This section outlines the hardware and software setup utilized for model development, experimentation, and system integration.

Hardware Configuration The model training and experimentation were conducted on a system equipped with standard computational resources suitable for deep learning tasks. The configuration includes:

- Processor: Intel i5 / AMD Ryzen 5 or equivalent
- RAM: Minimum 8 GB (recommended for stable training)
- GPU: NVIDIA GPU (for accelerated training and inference)
- Storage: Sufficient disk space for dataset storage and model checkpoints

The use of GPU acceleration is particularly important for training GAN-based models, as it significantly reduces training time and enables efficient handling of large datasets and complex architectures.

Software Configuration The software environment is de-signed to support both deep learning model development and full-stack system implementation. The key components include:

- Programming Language: Python
- Deep Learning Framework: PyTorch (for model development and training)
- Frontend Technologies: React.js with TypeScript
- Styling Framework: Tailwind CSS with UI components
- Backend Services: Supabase Edge Functions
- Model Deployment Platform: Hugging Face Inference Endpoint

- Development Tools: Visual Studio Code, Git, GitHub The model is trained using PyTorch due to its flexibility and strong support for custom architectures such as Pix2Pix. The frontend is built using modern web technologies to ensure a responsive and interactive user experience, while the backend manages API communication and secure data handling.

System Integration The overall system integrates these hardware and software components into a cohesive pipeline. The training phase utilizes local computational resources, while the deployment phase leverages cloud-based infrastruc-ture for scalability and real-time inference.

This configuration ensures a balance between development efficiency, system performance, and scalability, enabling both effective experimentation and practical deployment of the SAR colorization model.

2. Training Configuration

The training configuration defines the key parameters and settings used to optimize the Pix2Pix Conditional GAN model for SAR colorization. Proper configuration is essential to ensure stable adversarial training, effective convergence, and high-quality output generation.

The model is trained using a supervised learning setup on paired SAR–RGB image data. Training is performed over multiple epochs, allowing the generator and discriminator to iteratively improve through adversarial interaction.

Training Parameters

- Number of Epochs: Multiple training checkpoints were saved at different stages (e.g., 180, 265, and 295 epochs) to analyze performance variations over time.
- Batch Size: A moderate batch size is used to balance memory efficiency and training stability.
- Learning Rate: A small learning rate is selected to ensure gradual convergence and prevent instability in adversarial training.
- Optimizer: Adam optimizer is used for both generator and discriminator, due to its adaptive learning capability and faster convergence.

Loss Function Configuration The training process is guided by a composite loss function, combining:

- Adversarial loss (to ensure realism)
- L1 reconstruction loss (to maintain structural similarity) A weighting factor is applied to balance these two compo-nents, ensuring that the model does not overemphasize one objective at the expense of the other.

Training Strategy The generator and discriminator are trained in an alternating fashion:

- The discriminator is updated to distinguish real SAR–RGB pairs from generated ones.
- The generator is updated to produce outputs that can fool the discriminator while remaining close to ground truth images.

This iterative process continues until both networks reach a stable equilibrium.

Checkpointing and Model Saving To monitor training progress and evaluate performance at different stages, model checkpoints are saved at specific epochs. These checkpoints al-low comparison of results and selection of the best-performing model based on evaluation metrics such as FID score.

Stability Considerations GAN training is inherently un-stable; therefore, several techniques are employed to improve stability:

- Use of batch normalization
- Controlled learning rates
- Balanced updates between generator and discriminator Training progress is also monitored using loss curves, which provide insights into convergence behavior and potential issues such as overfitting or imbalance between networks.

3. Evaluation Metrics

Evaluating the performance of a SAR colorization model requires a combination of quantitative metrics and qualitative analysis, as the task involves both structural accuracy and perceptual realism. In this work, multiple evaluation metrics are employed to comprehensively assess the effectiveness of the proposed model.

Fre´chet Inception Distance (FID) The primary metric used in this study is the Fre´chet Inception Distance (FID) [19], which measures the similarity between the distribution of generated images and real images in a feature space. Lower FID values indicate that the generated images are closer to real images in terms of visual quality and statistical distribution.
The FID score is computed by comparing feature represen-tations extracted from a pretrained network for both real and generated images. It captures differences in both mean and covariance of feature distributions, making it a robust metric for evaluating generative models.

International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 3, May-June-2026, ISSN (Online): 2395-566X

In this work:
- A baseline FID score (30–40) is established by comparing two sets of real images
- The trained model achieves higher FID scores (around 185–190), indicating a gap between generated and real distributions

This highlights the challenge of achieving high-quality SAR colorization and provides a benchmark for future improvements.

Peak Signal-to-Noise Ratio (PSNR) PSNR is used to measure the pixel-level similarity between the generated image and the ground truth. Higher PSNR values indicate better reconstruction quality.
While PSNR is useful for evaluating structural similarity, it does not fully capture perceptual quality.

Structural Similarity Index (SSIM) SSIM evaluates similarity based on luminance, contrast, and structural in-formation. It provides a more perceptually aligned measure compared to PSNR.

Qualitative Evaluation Visual inspection is used to compare generated and ground truth images across land-cover types.

Training Loss Analysis Loss curves are analyzed to understand convergence and stability.

4. Baseline Definition
Establishing a baseline is essential for evaluating the effectiveness of the proposed SAR colorization model.

In this work, the baseline is defined using the Fre´chet Inception Distance (FID) computed between two independent sets of real optical images.

The baseline FID score is observed to lie in the range of 30 to 40.

The trained Pix2Pix models produce FID scores in the range of approximately 185 to 190.

This highlights a significant gap between generated and real distributions.

The baseline definition helps in:
- Contextualizing model performance
- Identifying performance gaps
- Guiding future improvements

VII. RESULTS AND ANALYSIS

1. Training Behavior Analysis
Understanding the training dynamics of the Pix2Pix Conditional GAN is essential to evaluate model stability, convergence, and overall learning behavior. Due to the adversarial nature of GANs, training involves a continuous competition between the generator and discriminator, making it important to analyze how both networks evolve over time.

The training process is monitored using loss curves for both the generator and the discriminator across epochs. These curves provide valuable insights into how effectively the model is learning and whether a balance is maintained between the two networks.

Generator Loss Behavior The generator loss reflects how well the model is able to produce outputs that can fool the discriminator while also maintaining similarity to the ground truth. During the initial stages of training, the generator loss is typically high, as the model produces coarse and unrealistic outputs.

As training progresses:
- The generator begins to capture structural patterns and texture information
- The loss gradually decreases, indicating improved performance
- Stabilization of the loss suggests convergence toward a consistent mapping

However, fluctuations in generator loss are expected due to the adversarial nature of training.

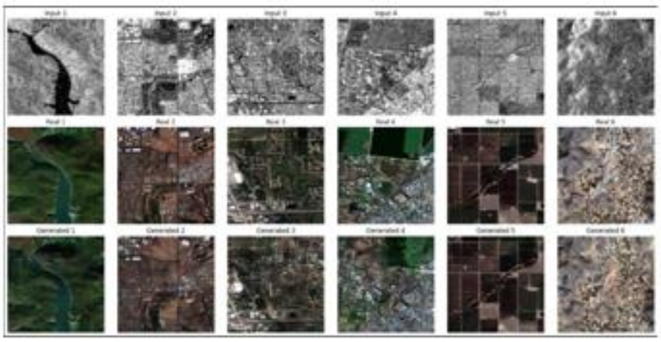

Fig. 8. Extended qualitative comparison of SAR inputs, real RGB images, and generated outputs across multiple samples, demonstrating the model’s ability to preserve structural details while enhancing visual interpretability.

Discriminator Loss Behavior The discriminator loss rep-resents the model's ability to distinguish between real and generated image pairs. In the early stages:

- The discriminator quickly learns to identify fake images, resulting in low loss
- As the generator improves, the discriminator faces more difficulty, leading to increased loss

A well-balanced training process is indicated by:

- Neither network dominating the other
- Oscillatory but stable loss patterns
- Gradual convergence over epochs

Training Stability and Convergence The loss curves ob-served in this work indicate that:

- The model achieves stable adversarial training without significant divergence
- Both generator and discriminator losses show controlled fluctuations
- Convergence is observed after sufficient epochs

Insights from Loss Graphs The training graphs demon-strate:

- Progressive improvement in generator performance
- Adaptive behavior of the discriminator
- Overall convergence toward a stable solution

2. Quantitative Results

Table 1: Fid Score Analysis

Model Name	Epochs	FID Score
Pix2Pix Generator	180	185.04
Pix2Pix Generator	265	189.81
Pix2Pix Generator	295	187.73

From the results, it can be observed that:

- The FID scores remain within a relatively close range across different epochs
- The model shows consistent performance but does not significantly improve beyond a certain point
- The lowest FID is achieved at epoch 180

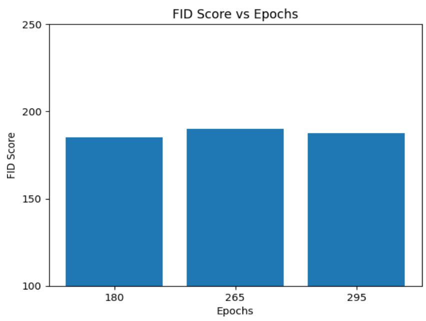

Fig. 9. FID score comparison across different training epochs. Lower values indicate better alignment with real image distributions.

When compared to the baseline FID score (30–40), there is a clear gap between real and generated distributions.

PSNR and SSIM Analysis In addition to FID, PSNR and SSIM are used to evaluate reconstruction quality:

- PSNR reflects pixel-level similarity
- SSIM measures structural consistency

The model achieves reasonable PSNR and SSIM values, indicating that structural features are preserved.

3. Qualitative Results

The qualitative evaluation compares input SAR images, ground truth optical images, and generated outputs across different land-cover categories.

Urban Areas

- Structural layouts such as buildings and roads are pre-served
- Color representation may not always match perfectly

Agricultural Regions

- Texture differentiation is visible
- Reasonable color variations are observed

Water Bodies

- Consistent identification
- Uniform color distribution

Barren Land

- Structural patterns captured
- Less consistent color assignment

General Observations

- Improved visual interpretability
- Preservation of spatial structures
- Some inconsistencies in color accuracy

4. Observations

Strengths of the Model

- Strong preservation of edges and spatial layouts
- Stable adversarial training
- Improved interpretability compared to raw SAR

Limitations of the Model

- High FID score

International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 3, May-June-2026, ISSN (Online): 2395-566X

- Color inconsistencies in complex regions
- Limited fine-detail capture

Training and Convergence Insights
- Increasing epochs does not significantly improve performance
- The model reaches a plateau after a certain point
- Best performance observed at intermediate epochs Overall Interpretation The model successfully enhances

SAR interpretability while maintaining structural consistency, but there remains a gap in perceptual realism compared to real optical images.

VIII. DISCUSSION

1. Performance Insights
The performance of the proposed SAR colorization model reflects both the strengths of conditional GAN-based architectures and the inherent complexity of mapping radar data to visually interpretable representations. A detailed analysis of the results reveals several key insights into how the model behaves and where improvements can be targeted.

One of the most notable observations is that the model demonstrates strong capability in preserving structural information. The generated outputs maintain spatial alignment with the input SAR images, successfully capturing features such as edges, boundaries, and geometric layouts. This indicates that the U-Net generator effectively retains low-level details while incorporating higher-level contextual information.

Another important insight is the model's ability to produce visually coherent outputs across different land-cover types. Despite variations in terrain, the model consistently generates distinguishable patterns for urban areas, vegetation, and water bodies. This suggests that the model has learned meaningful feature representations from the training data and can generalize across diverse scenarios.

However, the performance analysis also highlights a clear gap in perceptual realism, as reflected by the high FID scores. While the model produces structurally accurate outputs, it struggles to fully capture the statistical distribution of real optical images.

Another key insight is related to training saturation. The results show that increasing the number of training epochs does not significantly improve performance beyond a certain point. This suggests that the model reaches a learning plateau. The model also exhibits a balance between consistency and variability, although diversity in color representation remains limited.

From a system perspective, the integration of the model into a real-time pipeline demonstrates practical applicability.

2. Model Limitations
Despite the effectiveness of the proposed SAR colorization framework, several limitations are observed.
- High FID score indicating distribution gap
- Color inconsistency in complex regions
- Difficulty in capturing fine-grained textures

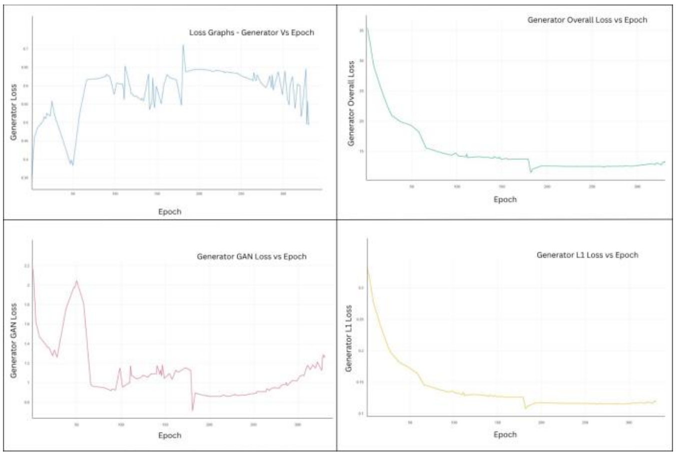

Fig. 10. Training loss curves of the generator across epochs, including overall loss, GAN loss, and L1 loss. The graphs illustrate convergence behavior and stability of the model during training.

- Dependency on paired datasets
- Limited interpretability (black-box nature)
- Training saturation
- Computational requirements

3. Comparison with Existing Methods
Traditional Methods
- Require expert knowledge
- Limited scalability
- Strong physical interpretability

CNN-based Approaches
- Produce blurred outputs
- Lack perceptual realism

CycleGAN-based Methods
- Work with unpaired data

- Lack strict structural alignment

Proposed Approach
- Better structural preservation
- Improved interpretability
- Real-world deployment capability

4. Research Gap Fulfillment
The proposed work addresses several key research gaps identified in prior studies.
- Clear distinction between SAR colorization and SAR-to-optical translation
- Use of structured paired dataset
- Multi-metric evaluation framework (FID, PSNR, SSIM)
- Integration into a complete system (SAR-RANG platform)

This bridges the gap between theoretical models and real-world applications.

Applications
The proposed SAR colorization framework has significant real-world applications across multiple domains.
- Disaster Management: Enables faster interpretation of affected regions during floods, earthquakes, and other natural disasters.
- Environmental Monitoring: Helps in tracking deforestation, water bodies, and climate-related changes.
- Urban Planning: Assists in analyzing infrastructure, land-use patterns, and urban expansion.
- Agriculture: Supports crop monitoring and soil condition analysis.
- Defense and Surveillance: Enhances situational awareness using SAR data in adverse conditions.

By improving interpretability, the model makes SAR data more accessible to non-experts and decision-makers.

Future Work
While the proposed framework demonstrates promising results, several directions can be explored for further improve-ment.
- Incorporation of advanced architectures such as attention-based GANs [21] and diffusion models [22]
- Development of improved evaluation metrics tailored for SAR colorization
- Expansion of dataset diversity to improve generalization
- Reduction of dependency on paired datasets using semi-supervised or unsupervised approaches
- Enhancement of model interpretability and explainability
- Optimization for faster inference and lower computational cost

These improvements can help bridge the gap between generated and real image distributions.

IX. CONCLUSION

This work presents a comprehensive framework for SAR image colorization using Conditional Generative Adversar-ial Networks (cGANs). The proposed approach successfully transforms grayscale SAR images into visually interpreable color representations while preserving structural and radiometric characteristics.

The integration of a Pix2Pix-based architecture, structured dataset, and complete system pipeline demonstrates both theoretical and practical contributions. The SAR-RANG platform further validates the real-world applicability of the approach by enabling interactive and scalable deployment.

Experimental results show that the model effectively pre-serves spatial structures and enhances interpretability, although challenges remain in achieving high perceptual realism, as indicated by FID scores.

Overall, this research contributes to advancing SAR image understanding and makes radar data more accessible for a wide range of applications, including disaster management and environmental monitoring.

REFERENCES

1. I. Goodfellow et al., ”Generative Adversarial Nets,” NeurIPS, 2014.
2. P. Isola et al., ”Image-to-Image Translation with Conditional Adversarial Networks,” CVPR, 2017.
3. J.-Y. Zhu et al., ”Unpaired Image-to-Image Translation using Cycle-GAN,” ICCV, 2017.
4. O. Ronneberger et al., ”U-Net: Convolutional Networks for Biomedical Image Segmentation,” MICCAI, 2015.
5. A. Krizhevsky et al., ”ImageNet Classification with Deep CNNs,” NeurIPS, 2012.
6. K. He et al., ”Deep Residual Learning for Image Recognition,” CVPR, 2016.
7. Y. LeCun et al., ”Gradient-Based Learning Applied to Document Recog-nition,” Proceedings of the IEEE, 1998.
8. Z. Wang et al., ”Image Quality Assessment: From Error Visibility to Structural Similarity,” IEEE TIP, 2004.

9. M. Schmitt and X. Zhu, ”Data Fusion and Remote Sensing: An Ever-Growing Relationship,” IEEE GRSM, 2016.
10. F. Argenti et al., ”Speckle Removal in SAR Images: A Review,” IEEE GRSM, 2013.
11. G. Camps-Valls et al., ”Deep Learning for Remote Sensing Data,” IEEE GRSM, 2021.
12. L. Ma et al., ”Deep Learning in Remote Sensing Applications: A Meta-Analysis,” IEEE GRSM, 2019.
13. R. Zhang et al., ”Colorful Image Colorization,” ECCV, 2016.
14. C. Ledig et al., ”Photo-Realistic Single Image Super-Resolution Using GANs,” CVPR, 2017.
15. M. Mirza and S. Osindero, ”Conditional Generative Adversarial Nets,” arXiv, 2014.
16. X. Yang et al., ”SAR-to-Optical Image Translation using GANs,” IEEE Access, 2020.
17. S. Wang et al., ”Deep Learning-Based SAR Image Colorization,” Re-mote Sensing, 2021.
18. R. C. Gonzalez and R. E. Woods, ”Digital Image Processing,” Pearson, 2008.
19. M. Heusel et al., ”GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium,” NeurIPS, 2017.
20. C. Li and M. Wand, ”Precomputed Real-Time Texture Synthesis with Markovian GANs,” ECCV, 2016.
21. H. Zhang et al., ”Self-Attention Generative Adversarial Networks,” ICML, 2019.
22. J. Ho et al., ”Denoising Diffusion Probabilistic Models,” NeurIPS, 2020.
23. G. Franceschetti and R. Lanari, ”Synthetic Aperture Radar Processing,” CRC Press, 1999.
24. J. S. Lee, ”Digital Image Enhancement and Noise Filtering by Use of Local Statistics,” IEEE TPAMI, 1980.
25. Y. Zhu et al., ”Deep Learning in Remote Sensing: A Comprehensive Review,” IEEE GRSM, 2017.
26. M. Liu et al., ”Unsupervised Image-to-Image Translation Networks,” NeurIPS, 2017.
27. J. Johnson et al., ”Perceptual Losses for Real-Time Style Transfer and Super-Resolution,” ECCV, 2016.