Deep Learning Driven: Approach for SAR Image Colorization

Prof. Shweta Jadhav[1], Ms. Disha Deshmukh[2], Ms. Vidya Dhage[3] , Ms. Prajakta Kapote[4], Ms. Manasi Sonawane[5] ,

[1]Department of Computer Engineering
[2,3,4,5]Undergraduate Students, Department of Artificial Intelligence and Data Science
K.K. Wagh Institute of Engineering Education and Research, Nashik

---***---

Abstract - To enhance the colorization of Synthetic Aperture Radar (SAR) images, which often lack fine details, a new approach is proposed. Traditional methods, such as CycleGAN, frequently result in blurry images with inaccurate colors. The proposed method utilizes a deep learning model to improve color accuracy and maintain fine details during colorization. It ensurs that the most important features are accurately represented. Furthermore, accuracy is assessed and refined after colorization to ensure optimal results. This innovative approach overcome the limitations of previous methods, making SAR images more interpretable and useful for applications like geological studies and environmental monitoring.

Key Words: SAR, Image Colorization, Deep Learning, Detail Preservation, Image Processing.

1.INTRODUCTION *(Size 11, Times New roman)*

Synthetic Aperture Radar (SAR) images are important in remote sensing, as they capture detailed views of the Earth's surface in all weather conditions, day or night. However SAR images are grayscale images, which makes it difficult to visually interpret features like land, water, and urban areas. Unlike optical images that show clear colors, SAR images need special processing to uncover patterns and textures. This makes it challenging to analyze important features like land, water, and urban areas, which are vital for environmental monitoring ,geological studies and urban planning.

Colorizing SAR images makes them easier to understand. Many current colorization techniques like CycleGAN face difficulty to create images that are both clear and realistic. They often result in blurry colors that don't look natural, this also leads to disruption of important details and ultimately reduces usefulness of the images.

Deep learning approach is used to bring realistic, vibrant color to SAR images while preserving critical details. The system uses a combination of Superpixel Segmentation, Principal Component Analysis (PCA), and a U-Net model to enhance colorization of images. Superpixel segmentation groups similar pixels to make color mapping simple ,PCA reduces the data complexity, focuses only on the essential features. The U-Net model helps the system to learn different regions in the image and apply appropriate colors to them.

Training the model on paired SAR and optical images enables it to learn realistic color mapping. This colorization approach finds its practical applications in many fields. In environmental monitoring, it keeps track of changes in forest regions, water bodies, and urban areas over time. In geological studies, it paves way for better understanding of land formations and rock structures. Even in space exploration, colorized SAR images can provide clear views of planetary surfaces. This system ultimately aims to transform grayscale SAR images into colorful, detailed versions of images, making SAR data more accessible and meaningful for researchers in various fields.

2. LITERATURE SURVEY

The colorization of Synthetic Aperture Radar (SAR) images is a challenging area within remote sensing due to the lack of color information in SAR data, which limits interpretability. Prior studies have proposed various deep learning methods to address this challenge, with varying degrees of success.

2.1 CycleGAN

CycleGAN is a generative model designed for unpaired image-to-image translation tasks. In SAR image colorization, it effectively maps grayscale SAR images to their colorized parts without requiring paired datasets. However, the model is challenging to train with noisy or small datasets, and often produces outputs like color inconsistencies and blurred regions. Jung-Hoon Lee et al. (2021) [1] introduced a deep learning-based colorization method using CycleGAN for SAR images. This model employed unpaired image-to-image translation to convert SAR to optical images, achieving colorized outputs. However, the CycleGAN model often produced blurry images with unnatural colors, resulting in a loss of critical structural details necessary for SAR image analysis.These limitations lead to need for different approaches that can maintain both color and detail clarity, especially for applications like environmental monitoring and disaster management.

2.2 Superpixel Segmentation

Superpixel segmentation is widely used to simplify image processing by dividing the image into homogeneous regions. These regions capture important structures and textures, making the processing more efficient. In SAR image colorization, superpixels help preserve edge details, which are important for accurate colorization.

Isabela Borlido et al. (2023) [2] conducted a comprehensive review of superpixel segmentation methods, which are important for enhancing object recognition and structural clarity in images. By categorizing segmentation techniques into classical and deep learning-based methods, the authors demonstrated how superpixel segmentation could improve object detection and detail preservation. This research supports the integration of superpixel segmentation in project to retain fine details during SAR image colorization.

2.3 U-Net for Image Segmentation

U-Net, a popular architecture for image segmentation tasks, has ability in pixel-level tasks like SAR image colorization. By

utilizing a convolutional network with a symmetric encoder-decoder structure, U-Net excels at preserving spatial information through skip connections,which are used for tasks that require high-resolution outputs.
N. Subrajaa and D. Venkatasekhar (2022) [3] explored the use of U-Net convolutional networks for satellite image segmentation. Their modified U-Net architecture, which included features like batch normalization and deconvolution layers, significantly improved segmentation accuracy for high-resolution satellite imagery. This study validates the suitability of U-Net for complex remote sensing tasks and informs our choice to apply U-Net with PCA and superpixel segmentation, aiming for better colorization with minimal detail loss in SARimagery.

2.4 Swin U-Net Model

Transformer models have recently made their way in image segmentation tasks due to their capacity to detect long-range dependencies and contextual information across images. These models, like the Swin Transformer, are being adapted for various segmentation tasks, offering state-of-the-art performance in capturing complex patterns. Ziyang Wang et al.(2023) [4] developed an advanced variant of U-Net (DCS-UNet), incorporating Swin Transformer blocks and dense skip connections to enhance feature aggregation across multiple scales. While designed for medical image segmentation, the
DCS-UNet model demonstrated the potential of
Transformer-based models to improve segmentation accuracy.This approach informs our work by illustrating how dense connections and multiscale processing can enhance model performance,
particularly in retaining structural details in colorized images.
These studies provide foundational understanding and identify limitations within existing SAR colorization techniques, such as color inaccuracies, loss of detail, and segmentation challenges. Our approach builds on these findings by integrating a U-Net model with PCA and superpixel segmentation, aiming to produce high-quality, colorized SAR images that retain features for environmental and geological applications.

3. IMPLEMENTATION

The proposed system uses a deep learning-based pipeline to transform grayscale Synthetic Aperture Radar (SAR) images into colorized versions while preserving important structural and textual features.

3.1 Feature Extraction

Before the colorization process, SAR images undergo feature extraction to enhance the model's ability to understand and preserve structural details during colorization. The following techniques are applied:

3.1.1 Superpixel Segmentation

Superpixel segmentation is used in SAR image colorization to break the image into smaller, more manageable regions, known as superpixels, where pixels with similar values are grouped together. This makes the colorization process easier by allowing the model to apply consistent colors to each region. At the same time, it helps preserve important details like edges and textures, ensuring that the colorization keeps the key features of the SAR images.

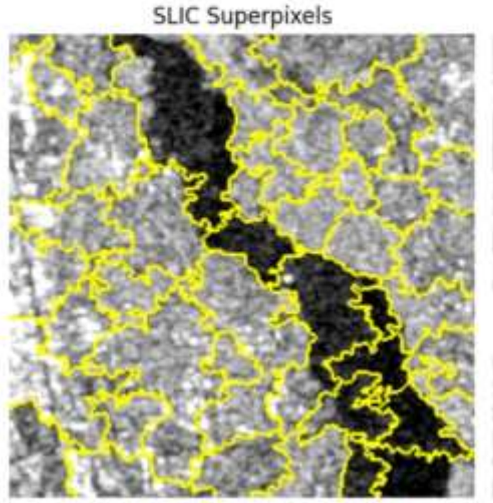

Figure 1: Application of Superpixel segmentation

3.1.2 Normalization

Normalization is an important step in preparing SAR images for colorization. It involves scaling the pixel values, usually from a range of [0, 255] down to [0, 1], so that the data is consistent across all images. Similarly, any features extracted from the images are also normalized, either by adjusting them to a set range or ensuring they have a mean of zero and a standard deviation of one. This process helps the model treat all inputs fairly, preventing any single feature from overwhelming the others. Ultimately, normalization makes the model training smoother and more stable, leading to more accurate and reliable colorized images.

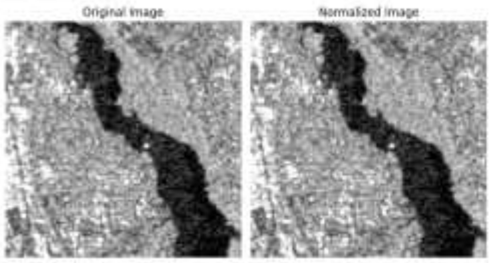

Figure 2: Normalization of image

3.1.3 Principal Component Analysis (PCA)

PCA is used to simplify the SAR images by focusing on the most important features. It reduces the image's complexity by removing unnecessary details, allowing the model to focus on key structures. This helps the model colorize the images more accurately while minimizing noise and irrelevant information.

3.2 Optimization

The optimization of the SAR image colorization system focused on enhancing efficiency, scalability, and accuracy. Several techniques were systematically implemented to achieve these objectives

- Feature Dimensionality Reduction : Principal Component Analysis (PCA) applied to high-dimensional SAR data into a compact feature space. By retaining only the most significant components, this step preserves essential patterns while significantly lowering computational overhead during training and inference.

- Loss Function Design : A combined loss function utilizing Mean Squared Error (MSE) and Structural Similarity Index Measure (SSIM) are applied to it.MSE ensures pixel-level accuracy by minimizing the difference between the predicted and actual pixel values. SSIM, on the other hand, preserves the structural and textural integrity of the image, preventing distortions in important features. This approach balances the need for high visual quality with strong quantitative performance.

- Model Optimization : The Adam optimizer is used to fine-tune model weights, offering adaptive learning rates and efficient gradient updates. A learning rate of 10^{-4} is selected to ensure stable convergence without overfitting.

- Hardware Acceleration : GPU acceleration in Google Colab is utilized to meet the high computational demands of SAR image colorization. Batch sizes are optimized to balance memory usage and processing speed, ensuring efficient training workflows.

- Evaluation : Performance metrics, including Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM),are monitored throughout training. Iterative adjustments in hyperparameters and architecture are made based on these evaluations, ensuring continuous performance improvements.

3.3 U-Net Model

The U-Net is an encoder-decoder architecture widely used for image-to-image tasks, including SAR image colorization. It is a non-generative model that effectively maps grayscale SAR images to corresponding colorized outputs. The U-Net consists of two main parts:

1. Encoder: Extracts features from the input SAR image using convolutional layers.
2. Decoder: Reconstructs the colorized image using upsampling layers, with skip connections preserving spatial details.

Input and Output Grayscale SAR image tensor of size H×W×1H×W×1.Colorized RGB image tensor of size H×W×3H×W×3.

Mathematical Representation:

- Encoder feature extraction:

$$f_l = \sigma(W_l * f_{l-1} + b_l)$$

- Decoder reconstruction:

$$fl' = \text{Upsample}(fl) + \text{Concat}(f\text{encoder}, l)$$

f_l: Feature map at layer l W_l: Weight matrix for layer l f_{l-1}: Feature map from the previous layer
b_l: Bias term for layer l σ: Activation function (e.g., ReLU) f_l': Reconstructed feature map at layer l in the decoder.

3.4 Mathematical Model

1. Input: SAR Image Grayscale SAR image:

$$ISAR \in RH\times W$$

where H is the image height and W is the image width.

2. Preprocessing
Superpixel Segmentation: Groups similar pixels to simplify color mapping:

$$S = \text{SuperpixelSegmentation}(I_{SAR})$$

where S is the set of superpixel groups.

Principal Component Analysis (PCA): Reduces dimensionality of superpixel features:

$$F = \text{PCA}(S)$$

where F represents the reduced feature set.

3. Colorization Using Deep Learning The U-Net model maps the grayscale input I_{SAR} to a colorized output

$$I_{RGB} = f_\theta(ISAR)$$

where f_θ is the U-Net model parameterized by weights θ.

- Encoder: Extracts latent features:

$$Z = \text{Encoder}(I_{SAR})$$

where Z represents the latent feature set.

- Decoder: Maps latent features to the colorized image:

$$I_{RGB} = \text{Decoder}(Z)$$

4. Loss Function The training objective minimizes the Mean Squared Error (MSE):

$$LMSE = N\ i\sum{=}1\ IRGB(i)$$

where I_{RGB}^{true} is the ground truth RGB image.

5. Optimization Parameters θ are optimized using the Adam optimizer:

$$\theta \leftarrow \theta - \eta \nabla\theta LMSE$$

where η is the learning rate

6. Evaluation Metrics :

- Structural Similarity Index (SSIM): Evaluates image quality:

$$\text{SSIM} = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu x2 + \mu y2 + C1)(\sigma x2 + \sigma y2 + C2)}$$

μ_x: Mean intensity of image x
μ_y: Mean intensity of image y
σ_y σ_{xy}: Covariance between images x and y
C_1: Stabilizing constant for mean terms
C_2: Stabilizing constant for variance terms

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 05 | May - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

- Peak Signal-to-Noise Ratio (PSNR): Measures reconstruction quality:

$$PSNR = 10 \times log_{10}\,(\,max(I^{\mathrm{RGB}})^2 \,/\, L_{\mathrm{mse}}\,)$$

$\max(I_{\mathrm{RGB}})^2$: Maximum possible pixel intensity in the RGB image
LMSE: Mean Squared Error between the original and reconstructed image

Figure 3: Training Stage

Figure 4: Testing Stage

4. PROCESS BLOCK DIAGRAM

The proposed system is divided into two key phases: the Training Phase and the Testing Phase. In the Training phase, the model is trained on a set of paired SAR and optical images. During this phase, the system learns how to map grayscale SAR images to their corresponding colorized optical images.In the Testing phase, a grayscale SAR image is provided as input, and the system generates the corresponding colorized output.

1. Input Image: The grayscale SAR image is provided as input to the system.
2. Feature Extraction: Features are extracted from the SAR image to represent the structural and textural elements.
3. Dimensionality Reduction: PCA (Principal Component Analysis) is applied to reduce the dimensionality of the extracted features, focusing on the most important components.
4. Superpixel Segmentation: The image is divided into superpixels, where pixels with similar values are grouped together for more efficient processing and colorization.
5. U-Net Model: The encoded features are passed through the U-Net model, which generates the colorized output.
6. Similarity Matching: After colorization, the generated image is compared against the ground truth to assess its accuracy using metrics such as MSE, SSIM, and PSNR.
7. Final Output: The colorized image is generated.

5. EXPERIMENTAL SETUP

5.1 Dataset

To evaluate the performance of the proposed system, a dataset of 3,000 training images and 250 testing images is used, sourced from the Kaggle Sentinel 1 and 2 Image Pairs Segregated by Terrain dataset. The images in the dataset are of size 256x256 pixels and are in PNG format. Each image consists of paired Synthetic Aperture Radar (SAR) and optical images, offering a diverse range of terrain types for training and testing purposes. Sentinel-1 provides Synthetic Aperture Radar (SAR) images, which are grayscale and offer high-resolution imagery of the Earth's surface, while Sentinel-2 provides optical images in the RGB (Red, Green, Blue) color spectrum.
The dataset is pre-processed to ensure that all images are aligned in terms of spatial resolution and format, making them suitable for deep learning models. The ground truth for the dataset consists of the paired optical images, which are used for comparison against the generated colorized SAR images. The performance of the proposed system is evaluated using standard metrics such as Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), and visual inspection for qualitative analysis.
The dataset provides a well-balanced mix of terrains, which will help ensure that our model generalizes well across different types of landscapes.

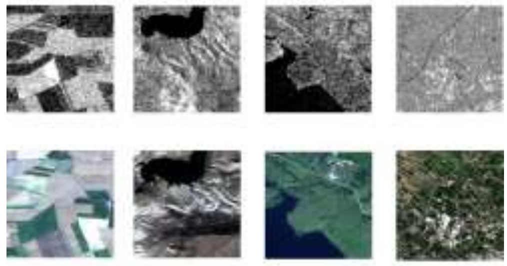

Figure 5: Dataset

5.2 Performance Metrics

Loss Function: The loss function measures how far the model's predictions are from actual values. For image colorization, Mean Squared Error (MSE) or Mean Absolute Error (MAE) are common choices, minimizing prediction errors to ensure generated images match the target[1].

Mean Squared Error (MSE):

$$MSE = (1\,/\,N) \times \Sigma\,(y_{\mathrm{i}} - \hat{y}_{\mathrm{i}})^2$$

Mean Absolute Error (MAE):

$$MAE = (1\,/\,N) \times \Sigma\,|y_{\mathrm{i}} - \hat{y}_{\mathrm{i}}|$$

where:
y_i : Ground truth value
$\hat{y}_i$: Predicted value
N: Total number of data points
Peak Signal-to-Noise Ratio (PSNR): PSNR is a common metric used to measure the quality of reconstructed images. It compares the maximum possible pixel value with the noise (errors) between the colorized and ground truth images. Higher PSNR values indicate better image quality.

$$PSNR = 10 \times log_{10}\,(\,max(I)^2 \,/\, L_{\mathrm{mse}}\,)$$

where $\max(I)$ is the maximum possible pixel value (typically 255 for 8-bit images), and L_{MSE} is the mean squared error.

6. RESULT

Parameter	U-Net	CycleGAN
Epochs	40	100
Batch Size	8	1
Train Dataset	12,800 images	2,000 images
Test Dataset	3,200 images	177 images

Figure 6: Parameters of U-Net Model

Metric	U-Net
SSIM	**0.8735** (Good)
MSE	**362.6798** (Lower Error)
PSNR	**16.7869 dB** (Moderate Quality)

Figure 7: Metrics Result of U-Net Model

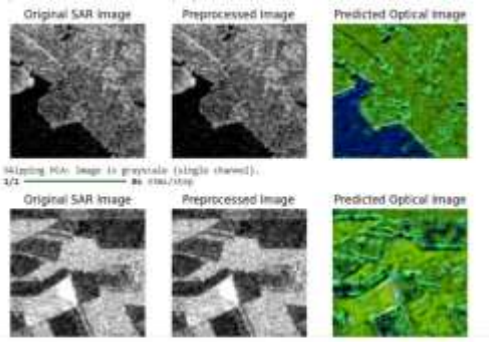

Figure 8: Image Result of U-Net Model

7. CONCLUSION

In conclusion, the system is still in development, with notable progress being made through continuous research and the acquisition of relevant datasets. It is currently experimenting with various preprocessing algorithms to identify the most effective methods for enhancing the images. This initial phase is essential as the system progresses, offering a clearer understanding of the challenges and opportunities in transforming SAR images. As development continues, the system is expected to provide valuable insights that will contribute to the final solution.

REFERENCES

[1] J.-H. Lee et al., "Design of CycleGAN Model for SAR Image Colorization," in *Proc. of APWCS*, IEEE, 2021, pp. 600–605.

[2] Borlido et al., "Superpixel Segmentation: From Theory to Applications," in *Proc. of SIBGRAPI*, IEEE, 2023, pp. 258–261.

[3] N. Subrajaa and D. Venkatasekhar, "Satellite Image Segmentation using Modified U-Net Convolutional Networks," in *Proc. of ICSCDS*, IEEE, 2022, pp. 1706–1712.

[4] Z. Wang et al., "Densely Connected Swin-UNet for Multiscale Information Aggregation in Medical Image Segmentation," in *Proceedings of IEEE International Conference on Image Processing (ICIP)*, 2023, pp. 940–943.

[5] Afsia T., Belhaj A., Tlijani, H.,Nouri K., "Implementing Canny Edge Detection Algorithm for Different Blurred and Noisy Images." in *Proceedings of the IEEE 21st International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA)*, Sousse, Tunisia, 2024, pp. 342–347

[6] Shen L., Lang B.,Song Z. (2023)." DS-YOLOv8-Based Object Detection Method for Remote Sensing Images." in *IEEE Access*, vol. 11, 2023, pp. 125122–125134.

[7] Addabbo P., Bernardi M. L., Biondi F., Cimitile M., Clemente C., Fiscante N., Giunta, G., Orlando D. Yan. ,L. (2023). "Super-Resolution of Synthetic Aperture Radar Complex Data by Deep-Learning." in *IEEE Access*, vol. 11, 2023, pp. 23647–23658.