Denoising Synthetic Aperture Radar / Aerial Images Using HOTV Deep Learning Models with Bayesian MAP Approach

Ashok Shrimant Hake[1], Krishnendu Remesh[2][a] and Vishal Subhash Chavan[1]

[1]*Dept. of Computer Science and Business Systems, KIT's College of Engineering Kolhapur, Maharashtra, India*

[2]*Department of Mathematics, Christ University, Bengaluru, Karnataka, India*

Keywords: Image Denoising, Total Variation Model, HOTV, CNN Model.

Abstract: Denoising plays an essential role in Synthetic Aperture Radar (SAR) and aerial image restoration. These images are distorted with various noises due to atmospheric changes. Therefore, the images should be analyzed using proper restoration and enhancement techniques. Many authors proposed traditional and deep learning models to perform this task. This paper employed the Bayesian Maximum A Posteriori (MAP) approach to the Higher Order Total Variation (HOTV) deep learning model. We assumed that the Poisson noise distorts the images. We also used the model to restore the images degraded by noises such as Gamma, Gaussian, and Rayleigh. Quantitative and qualitative analyses are provided.

1 INTRODUCTION

A wide variety of noise and distortions, including blur, decreased contrast, intensity, and inhomogeneity, frequently deteriorate sensor data Rasti et al. (2021). One of the primary causes of the deterioration of satellite, remote-sensed, and aerial images is the presence of several types of noise. Poisson noise is one of the most common noises in Synthetic Aperture Radar (SAR) images, and this type of noise ultimately increases the difficulty of interpreting images Febin et al. (2020).

A noisy satellite or aerial image can be formulated by

$$x_0 = x * n, \tag{1}$$

where x represents the original image, while n denotes the multiplicative noise, which is observed to follow a Poisson distribution. Mathematically, the denoising problem of the images is ill-posed.

Many studies have been carried out in denoising the images. The models employed local filters that could not retain essential details, such as edges, due to the assumption that neighboring pixels shared identical statistical characteristics Rasti et al. (2021). Non-local models estimate the weighted non-local similarity between small image patches to preserve resolution while removing the noise. Many other authors

---

[a] https://orcid.org/0000-0001-6713-8483

have used non-local denoising models for SAR images Deledalle et al. (2014); Parrilli et al. (2011).

Total Variational (TV) Model (Rudin Osher Fatemi (ROF) model) are well known for image denoising Rudin et al. (1992). TV-based models are efficient because of their ability to preserve edges. However, the staircase effect is the main flaw in the TV-based method Strong and Chan (2003). The optimization function in the TV model is a trade-off between data fidelity and regularization. It is formulated by:

$$\min_x \frac{\lambda}{2} \|x_0 - x\|_2^2 + \|\nabla x\|_1, \tag{2}$$

where $\lambda > 0$ and x is the original image or desired image. Even though the large values of λ retain the features in the images, the small values of λ provide better denoising. Therefore, an optimal choice of λ is mandatory.

Many modified approaches to TV model Li and Li (2021) have been proposed to rectify the staircase effect. By considering the higher-order gradient, the Higher-Order Total Variation model is introduced Lysaker et al. (2003). It is formulated by:

$$\min_x \frac{\lambda}{2} \|x_0 - x\|_2^2 + \|\nabla^2 x\|_1. \tag{3}$$

The higher order gradient with the L_1 norm reduces the staircase effect of the regular TV model in the image restoration.

Aubert and Aujol introduced a Bayesian Maximum A Posteriori (MAP) model using MAP estima-

tor with the total variation regularization in Aubert and Aujol (2008). The model facilitates handling multiplicative gamma noise. While effective to some extent, traditional models may struggle to handle complex noise patterns and may introduce undesirable artifacts in the denoised images.

Deep learning models CNNs to adapt the relation between clean and noisy images. By leveraging large datasets and learning complex patterns directly from data, CNNs have demonstrated remarkable performance in various image denoising Jebur et al. (2024). These models are widely studied in SAR image denoising also. Chierchia et al. proposed a residual-based learning model in Chierchia et al. (2017), which has a faster convergence. However, Training involves using a large multitemporal SAR image to approximate a clean image. A Bayesian despeckling method inspired by blind-spot denoising networks and incorporating a TV regularizer is employed by Molini et al. in Molini et al. (2021). We consider a CNN model based on Bayesian MAP approach.

2 DATA FIDELITY TERMS USING BAYESIAN MAP

According to Bayesian rule,

$$P(U|V) = \frac{P(V|U)P(U)}{P(V)}, \qquad (4)$$

where $P(U|V)$ is the conditional probability of the random variable U given V. Here, we use the above Bayesian rule and try to restore the image by maximizing the posterior probability $P(x|x_0)$ given by

$$P(x|x_0) = \frac{P(x_0|x)P(x)}{P(x_0)}. \qquad (5)$$

That is,

$$\max_x P(x|x_0) = \max_x P(x_0|x)P(x), \qquad (6)$$

The term $P(x_0)$, the prior probability on x_0, is a constant w.r.t. x that can be neglected.

Assume that the speckles in SAR images follow the Poisson noise. Therefore, the posterior probability function $P(x_0|x)$ is given as

$$P(x_0|x) = \frac{\exp(-x)x_0^x}{x_0!} \qquad (7)$$

One can consider the image (x and x_0) as a set of independent pixels of the image, say x_i, (The joint

probability equals the product of the marginal probabilities of each random variable $x(x_i)$), therefore, (6) can be written as

$$\max_x P(x(x_i)|x_0(x_i)) = \max_x \prod_{i=1}^{N} P(x_0(x_i)|x(x_i))P(x(x_i)), \qquad (8)$$

where N is the total number of image samples.

Since the function log is a monotone function, maximizing $P(x|x_0)$ is equivalent to minimizing the negative log-likelihood, and hence from (7) and (8), we can obtain the following;

$$\min_x \left\{ \sum_{i=1}^{N} x(x_i) - x_0(x_i)log(x(x_i)) - \sum_{i=1}^{N} log(P(x(x_i))) \right\} \qquad (9)$$

where the prior of x, say $P(x)$, follows a regularization prior. For the sake of simplicity, we eliminate x_i, thus we get,

$$\min_x \{-logP(x|x_0)\} = \min_x \left\{ x - x_0logx + \lambda\phi(x) \right\}, \qquad (10)$$

where $\phi(x)$ be the prior probability function. Many authors considered $\phi(x)$ is the the total variation of x.

3 MAP MODEL WITH HOTV REGULARIZATION

We implemented a deep learning model using a Convolutional Neural Network (CNN) architecture designed for Higher Order Total Variation (HOTV). Generally, we use the loss function of the HOTV model as in (3). In this paper, we designed the model for the loss function Poisson + HOTV which works well to restore the SAR/Ariel images distorted with the poisson noise. We consider the objective function as in (10) with the assumption that the prior probability $\phi(x)$ follows HOTV, provided the noise follows the Poisson distribution. That is, the fidelity term is $x - x_0logx$ and the prior regularization of x is $\phi(x) = \|\nabla^2 x\|_1$.

Also, we consider the model feature a custom loss function that integrates the HOTV loss function with other loss functions' Bayesian approach to address different noise types. So the model can be easily adapted to handle other noise distributions, such as Gamma, Gaussian, and Rayleigh, by modifying the data fidelity term. We experimented with the model with three variations of loss functions: Gamma + HOTV, Gaussian + HOTV, and Rayleigh + HOTV. By employing the same architecture and dataset of original images, we evaluated the performance of these

combined loss function models to determine their effectiveness in denoising. Note that, we used the specific loss function to reconstruct the image distorted by the corresponding noise.

We consider the data fidelity terms according to the nature of noise (see Table I) along with the fixed prior regularization term. Note that the Gaussian noise is additive.

Table 1: Data Fidelity Term for Various Noises.

Noise Distribution	Data Fidelity Term
Gamma	$logx + \frac{x_0}{x}$
Gaussian	$(x - x_0)^2$
Rayleigh	$2log(x) + \frac{x_0^2}{2x}$

3.1 Model Architecture

The deep learning model employed for HOTV utilized a CNN architecture. The details of the model layers, output shapes, and parameters are given in Table II. We use a total of 121,355 parameters such that 40,451 trainable parameters and 80,904 Optimizer parameters.

Table 2: Model architecture with layer details, output shapes, and parameter counts.

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 256, 256, 64)	1,792
conv2d_1 (Conv2D)	(None, 256, 256, 64)	36,928
conv2d_2 (Conv2D)	(None, 256, 256, 3)	1,731

3.2 Hyperparameters and Training Configuration

We consider the Adam optimizer with a learning rate of 0.001 for all the models except the Rayleigh + HOTV model. For the Rayleigh model, the learning rate considered 0.0001 for better performance and improved quality metrics. We use the ReLU activation function for the first two convolutional layers and Sigmoid for the final convolutional layer. The λ value for balancing the regularization term with the data fidelity term in the HOTV loss function is set to 0.0001. The models were trained using a dataset consisting of 2000 aerial images.

Algorithm 1 Training and Evaluation of Image Denoising Models

1: **Input:** Noisy images X, Clean images Y, Number of epochs E, Batch size. B
2: **Output:** Trained model, Evaluation metrics.
3: Initialize model parameters.
4: Define the combined loss function as in (10).
5: **for** each model type **do**
6:     Compile the model with Adam optimizer and loss function.
7:     **for** epoch = 1 to E **do**
8:         Shuffle the training data.
9:         **for** batch = 1 to N/B **do**
10:             Select a batch of noisy images X_{batch} and clean images Y_{batch}.
11:             Perform forward pass to compute predictions $\hat{Y}_{batch}$.
12:             Compute loss using the combined loss function.
13:             Perform backward pass to update model parameters.
14:         **end for**
15:     **end for**
16:     Save the trained model.
17:     **Evaluation:**
18:     **for** each test image **do**
19:         Load noisy image X_{test} and corresponding clean image Y_{test}.
20:         Denoise the image using the trained model to get $\hat{Y}_{test}$.
21:         Compute evaluation metrics (MSE, PSNR, SSIM).
22:         Store the computed metrics.
23:     **end for**
24:     Save evaluation metrics to an Excel file.
25: **end for**

4 QUANTITATIVE AND VISUAL ANALYSIS

To evaluate the performance of our deep learning-based image denoising models, we assessed them using three standard quality metrics: Mean Squared Error (MSE), Peak Signal-to-Noise Ratio (PSNR), and Structural Similarity Index (SSIM). These metrics provide a comprehensive evaluation of image quality and denoising effectiveness.

MSE measures the average squared difference between the noisy and denoised images, with lower values indicating better performance. The PSNR quantifies the ratio of the maximum possible signal power to the noise power, with higher values signifying better

quality. It is formulated as

$$PSNR = 10.log\left(\frac{x_{max}^2}{MSE}\right),$$

where x_{max} denotes the maximum pixel intensity. A higher PSNR value (in dB) indicates better image quality. SSIM is defined as

$$SSIM(x,x_0) = \frac{(2\mu_x\mu_{x_0} + c_1)(2\sigma_{x_0} + c_2)}{(\mu_x^2\mu_{x_0}^2 + c_1)(\sigma_x^2 + \sigma_{x_0}^2 + c_2)},$$

where μ_x, μ_{x_0} are the means and σ_x, σ_{x_0} are the variances of x, x_0, respectively. The variables c_1 and c_2 stabilize the division with a weak denominator. SSIM assesses the similarity between the original and denoised images, with values closer to 1 indicating higher similarity. The following tables present the detailed quality metrics for various noise models applied in our study.

Table 3: Quality Metrics for Poisson + HOTV Model

Image	MSE	PSNR (dB)	SSIM
img1	8.2905	38.846	0.9932
img2	10.4559	36.068	0.9828
img3	3.6985	42.432	0.9933
img4	4.8233	41.280	0.9921
img5	5.5569	40.646	0.9830
img6	3.6464	42.494	0.9939
img7	4.9850	41.147	0.9933
img8	13.4772	35.860	0.9883
img9	3.2891	42.945	0.9916
img10	4.7166	41.313	0.9939
Average	**6.0652**	**40.469**	**0.9890**

Table 4: Quality Metrics for Standard HOTV Model

Image	MSE	PSNR (dB)	SSIM
img1	33.93	32.40	0.847
img2	39.90	31.20	0.972
img3	24.11	33.87	0.865
img4	50.55	29.60	0.963
img5	18.27	35.42	0.968
img6	22.56	34.01	0.933
img7	26.78	33.12	0.912
img8	20.34	34.58	0.941
img9	29.45	32.78	0.918
img10	31.56	32.54	0.906
Average	**29.50**	**32.74**	**0.927**

Table 5: Quality Metrics for Gamma + HOTV Model

Image	MSE	PSNR (dB)	SSIM
img1	11.45	37.50	0.958
img2	24.00	33.77	0.916
img3	23.89	33.86	0.910
img4	34.17	32.30	0.912
img5	22.97	33.93	0.925
img6	12.34	36.92	0.946
img7	28.56	32.98	0.904
img8	19.43	34.71	0.929
img9	15.67	35.83	0.939
img10	21.12	34.24	0.921
Average	**22.23**	**34.29**	**0.922**

Table 6: Quality Metrics for Gaussian + HOTV Model

Image	MSE	PSNR (dB)	SSIM
img1	52.12	29.05	0.891
img2	32.24	31.71	0.801
img3	54.43	29.13	0.795
img4	34.56	32.32	0.823
img5	39.51	31.48	0.835
img6	45.67	30.28	0.812
img7	38.29	31.52	0.804
img8	47.12	30.17	0.829
img9	43.67	30.76	0.845
img10	51.34	29.29	0.820
Average	**43.39**	**30.70**	**0.818**

Table 7: Quality Metrics for Rayleigh + HOTV Model

Image	MSE	PSNR (dB)	SSIM
img1	47.6392	30.178	0.8298
img2	68.1799	28.234	0.8519
img3	117.3053	17.346	0.8055
img4	31.1744	31.814	0.8541
img5	61.6765	28.565	0.9192
img6	57.8432	29.310	0.8999
img7	80.7006	26.937	0.8610
img8	66.2137	27.874	0.9063
img9	41.7139	31.200	0.8537
img10	71.9193	28.108	0.8612
Average	**65.5326**	**27.786**	**0.8740**

5 CONCLUSION

In conclusion, the HOTV + Poisson model is the most effective in preserving image quality and reducing noise, while the other models, demonstrate varying degrees of effectiveness and quality trade-offs. Pois-

son + HOTV Model exhibits the highest overall performance with an average MSE of 6.0652, indicating superior noise reduction. It achieves the highest average PSNR of 40.469 and SSIM of 0.9890, demonstrating excellent preservation of image quality and structural similarity. These results highlight its effectiveness in producing high-fidelity denoised images. Also, Gamma + HOTV model demonstrates moderate performance when the images are distorted with Gamma noise. It provides a balanced approach to noise reduction and image fidelity but does not achieve the superior quality seen with the Poisson model. The standard HOTV model, also strikes a balance between effective noise reduction and maintaining image quality, though it does not reach the level of performance achieved by the Poisson model. The other two models indicate a loss in image quality and noticeable distortions.

Figure 1: Denoising results for Poisson + HOTV Model with Poisson Noise. In each figure, the first image is clean, the second is noisy and the third is restored.

Figure 2: Denoising results for Standard HOTV Model with Speckle Noise. In each figure, the first image is clean, the second is noisy, and the third is restored.
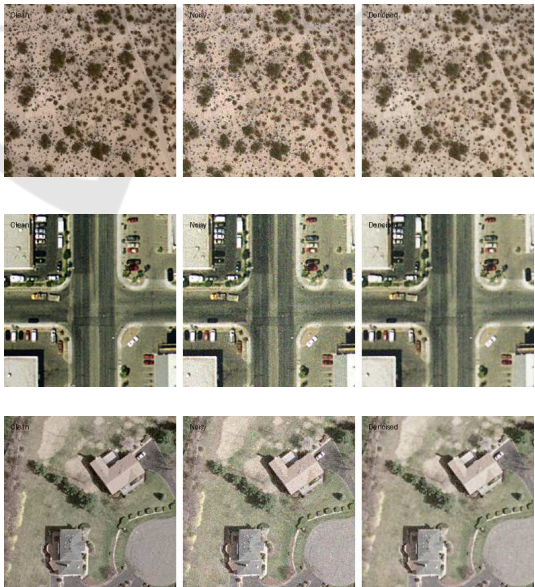
Figure 3: Denoising results for Gamma + HOTV Model with Gamma Noise. In each figure, the first image is clean, the second is noisy, and the third is restored.
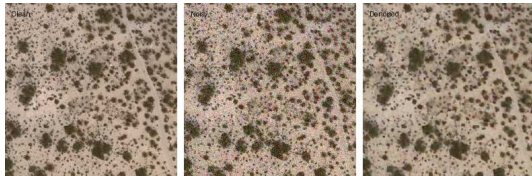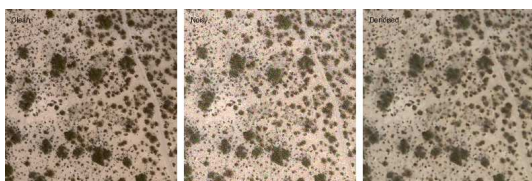

Figure 4: Denoising results for Gaussian + HOTV Model with Gaussian Noise. In each figure, the first image is clean, the second is noisy, and the third is restored.

Figure 5: Denoising results for Rayleigh + HOTV Model with Gaussian Noise. In each figure, the first image is clean, the second is noisy, and the third is restored.

REFERENCES

Aubert, G. and Aujol, J.-F. (2008). A variational approach to removing multiplicative noise. *SIAM journal on applied mathematics*, 68(4):925–946.

Chierchia, G., Cozzolino, D., Poggi, G., and Verdoliva, L. (2017). Sar image despeckling through convolutional neural networks. In *2017 IEEE international geoscience and remote sensing symposium (IGARSS)*, pages 5438–5441. IEEE.

Deledalle, C.-A., Denis, L., Tupin, F., Reigber, A., and Jäger, M. (2014). Nl-sar: A unified nonlocal framework for resolution-preserving (pol)(in) sar denoising. *IEEE Transactions on Geoscience and Remote Sensing*, 53(4):2021–2038.

Febin, I., Jidesh, P., and Bini, A. (2020). A retinex-based variational model for enhancement and restoration of low-contrast remote-sensed images corrupted by shot noise. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13:941–949.

Jebur, R. S., Zabil, M. H. B. M., Hammood, D. A., and Cheng, L. K. (2024). A comprehensive review of image denoising in deep learning. *Multimedia Tools and Applications*, 83(20):58181–58199.

Li, M.-M. and Li, B.-Z. (2021). A novel weighted total variation model for image denoising. *IET image processing*, 15(12):2749–2760.

Lysaker, M., Lundervold, A., and Tai, X.-C. (2003). Noise removal using fourth-order partial differential equation with applications to medical magnetic resonance images in space and time. *IEEE Transactions on image processing*, 12(12):1579–1590.

Molini, A. B., Valsesia, D., Fracastoro, G., and Magli, E. (2021). Speckle2void: Deep self-supervised sar despeckling with blind-spot convolutional neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–17.

Parrilli, S., Poderico, M., Angelino, C. V., and Verdo-

liva, L. (2011). A nonlocal sar image denoising algorithm based on llmmse wavelet shrinkage. *IEEE Transactions on Geoscience and Remote Sensing*, 50(2):606–616.

Rasti, B., Chang, Y., Dalsasso, E., Denis, L., and Ghamisi, P. (2021). Image restoration for remote sensing: Overview and toolbox. *IEEE Geoscience and Remote Sensing Magazine*, 10(2):201–230.

Rudin, L. I., Osher, S., and Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. *Physica D: nonlinear phenomena*, 60(1-4):259–268.

Strong, D. and Chan, T. (2003). Edge-preserving and scale-dependent properties of total variation regularization. *Inverse problems*, 19(6):S165.