Satellite Image Denoising

A PROJECT REPORT

Submitted by

Pradoom Varma (234101036)

Ayush Agarwal (234101060)

Shantanu Shrivastav (234101063)

Shubham Mourya (234101064)

in partial fulfilment for the course

DA 526: Image Processing with Machine Learning

Year – 2024

Project Supervisor: Prof. Debanga Raj Neog

Problem Statement

Developing a satellite image denoising algorithm targeting speckle noise removal, ensuring clarity enhancement for applications like environmental monitoring. It should be robust against speckle noise, preserving crucial details.

Introduction

In satellite imaging, it's crucial to have clear and accurate pictures for tasks like keeping an eye on the environment, planning cities, and managing disasters. But sometimes, these pictures get distorted by unwanted stuff like speckle noise. To fix this, we've made a smart tool called a denoising algorithm. It cleans up the images, making them clearer and better for understanding. This tool is really good at dealing with speckle noise, the most common issue in satellite images. It knows how to get rid of the noise while keeping the important details safe. It's perfect for quickly working through lots of satellite pictures. And it's set up to work smoothly on cloud systems, meaning it can process images in real-time without any hassle. This makes it super easy for people to get the cleaned-up pictures they need, exactly when they need them, for making smart decisions.

Related Work

Deep Learning Based SAR Image Denoising Using Convolutional Autoencoder:

This work utilizes convolutional autoencoders to denoise SAR images. It employs deep learning techniques to learn the complex mapping between noisy and clean SAR images. Convolutional layers are utilized to capture spatial information effectively.

SAR Image Despeckling Using a Stacked Sparse Denoising Autoencoder:
This approach uses stacked sparse denoising autoencoders for despeckling SAR images. It leverages the ability of autoencoders to learn hierarchical representations of input data. Sparse coding is employed to learn features that are robust to speckle noise.

SAR Image Despeckling Using Deep Learning with Convolutional Denoising Autoencoder:
In this study, a convolutional denoising autoencoder is proposed for SAR image despeckling. The autoencoder is trained on a dataset of noisy and clean SAR image patches to learn a mapping function between them. The learned model is then used to denoise full SAR images.

Unsupervised SAR Image Denoising Using Variational Autoencoder:
Variational autoencoders (VAEs) are employed for unsupervised SAR image denoising in this research. VAEs are capable of learning a probabilistic distribution of clean SAR images given noisy inputs. The model is trained in an unsupervised manner, requiring only noisy SAR images during training.

Dataset

We have created our own dataset by downloading images from the Sentinel-1 satellite developed by the European Space Agency (ESA). Sentinel-1 is a Synthetic Aperture Radar (SAR) satellite developed by the European Space Agency (ESA) as part of the Copernicus program. It provides high resolution radar imaging of the Earth's surface, regardless of weather conditions and time of day. The data products include Level-1, Level-2, and Level-3 products, which are used for a wide range of applications such as land and ocean monitoring, disaster management, climate change, and scientific research. The Copernicus Open Access Hub provides free access to Sentinel-1 data products, which are invaluable tools for decision-making and policy-making. Sentinel-1 data products enable scientists, researchers, and decision-makers to gain a better understanding of our planet and the challenges we face.

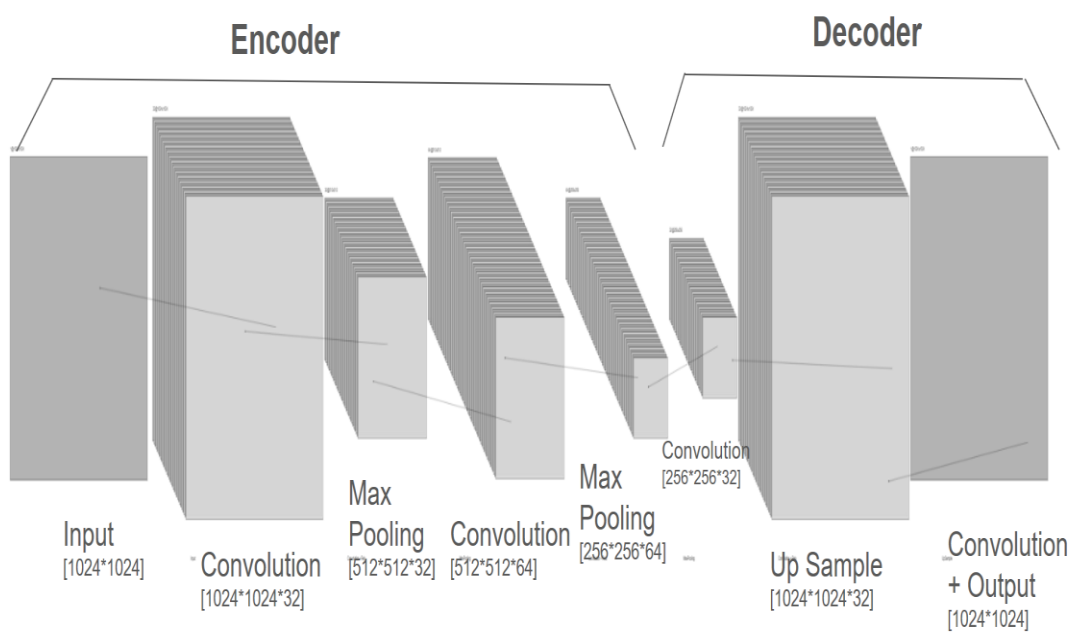
Speckle Noise

Speckle noise, often encountered in radar and synthetic aperture radar (SAR) images, appears as grainy interference that distorts the clarity of the image. It arises due to the coherent nature of the imaging system, where random constructive and destructive interference occurs among the coherent waves reflected from different surfaces within each resolution cell. As a result, speckle noise degrades the quality of satellite imagery, making it challenging to extract accurate information and discern subtle features.

After obtaining the image from the ESA website, it needs to be converted into JPEG format since the original image was in TIFF format. To create the dataset, we must divide the large image (approximately 25400*16800 pixels) into smaller images of size 1024*1024 pixels. Given that speckle noise is the most common noise in satellite imagery, we need to add this multiplicative noise to our dataset for training. Through this process, we generated images for both training and testing purposes. To evaluate the results and accuracy, we utilized the same original images.

Source:- https://dataspace.copernicus.eu/explore-data/data-collections

Method - Denoising Autoencoder



Denoising Auto-encoder

Denoising autoencoders are an extension of the basic autoencoders architecture. An autoencoder neural network tries to reconstruct images from hidden code space. In denoising autoencoders, we are introducing speckle noise to the images. The denoising autoencoder network will try to reconstruct the images. But before that, it will have to cancel out the noise from the input image data. In doing so, the autoencoder network will learn to capture all the important features of the data.

Encoding

Input Data: The encoder takes the input data, which could be an image, text, or any other type of data.

Feature Extraction: The input data is passed through several layers of the encoder neural network. Each layer typically applies linear transformations (e.g., matrix multiplications) followed by non-linear activation functions (e.g., ReLU, sigmoid) to transform the input data into a new representation.

Dimensionality Reduction: As the input data passes through the layers of the encoder, its dimensionality gradually decreases. This reduction in dimensionality forces the encoder to learn a compressed representation of the input data, capturing its essential features.

Latent Space Representation: The final output of the encoder is a compressed, lower-dimensional representation of the input data.

Decoding

Latent Space Representation: The compressed representation from the encoder serves as the input to the decoder.

Feature Reconstruction: The decoder consists of several layers that reverse the transformations applied by the encoder. These layers perform operations that gradually reconstruct the original input data from its compressed representation.

Dimensionality Expansion: Like the encoding process, the dimensionality of the data gradually increases as it passes through the layers of the decoder. This expansion process aims to recover the details lost during encoding and reconstruct the input data as faithfully as possible.

Output Reconstruction: The final output of the decoder is the reconstructed data, which ideally closely resembles the original input data. In the case of denoising autoencoders, the reconstruction process is aimed at removing noise and recovering the clean input data.

Experiments

Satellite images are often high-resolution and cover large geographical areas, resulting in large file sizes. Processing such large images can quickly exhaust the available memory resources, especially for computers with limited RAM. Patch-based processing involves dividing the large image into smaller, manageable patches that can fit into memory, thus alleviating memory constraints. Patch based processing was deployed to process SAR images.

Patch Based Processing

Working in patches also reduced computational complexity. Performing complex image processing operations, such as feature extraction or deep learning-based analysis, on full-resolution satellite images can be computationally intensive and time-consuming. By processing smaller patches individually, computational complexity is reduced, allowing for more efficient parallel processing and faster analysis.

Deciding for optimal patch size is the next challenge. There is a trade-off between Patch size and Context. Choosing a patch size that provides sufficient contextual information while maintaining computational efficiency is crucial. Larger patches capture more contextual information but may increase computational complexity and memory requirements. Conversely, smaller patches reduce computational demands but may lack sufficient context for accurate analysis, especially in areas with diverse features.

Due to above reason, experiments further were carried on low patch size.

Local features

Local features in images refer to patterns or structures that are discernible within a relatively small neighbourhood of pixels. Initially we started experimenting with 640x640 image sizes but were getting desirable results from it, the reason was lack of local features in the small sized image. While the fundamental principles of local feature extraction apply regardless of image size, the scale, context, computational complexity, robustness, and application requirements can influence how local features are characterized and utilized in large images compared to small images.

In-order to solve this, experiments were carried out on varying image sizes and finally 1024x1024 turns out to be optimal size.

Architecture

Deciding which deep learning architecture to use for denoising satellite images, particularly employing denoising autoencoders, involves various challenges due to the complexity of the data and the specific requirements of the denoising task,

1. Satellite images often have high dimensionality and contain intricate spatial patterns. Choosing a deep learning architecture with sufficient model capacity to capture these complex patterns while avoiding overfitting is challenging. Balancing model complexity with computational resources and training data availability is crucial.

2.Autoencoders are capable of learning rich hierarchical representations of input data. By training on noisy satellite images, autoencoders can learn to extract relevant features while filtering out noise. This feature learning capability enables autoencoders to capture complex patterns and structures present in satellite imagery, enhancing denoising performance.

Next, we decide to follow default Auto encoder architecture and started parameter tuning which includes, including the number of layers, the number of filters per layer, and the size of filters. Experiments were carried out with various parameter values and based on loss; optimal parameters were selected.

Overfitting in the model

Increasing the number of layers and filters in the autoencoder can enhance its capacity to learn complex representations from noisy input data. However, overly complex models may lead to overfitting, where the autoencoder memorizes noise patterns in the training data rather than learning meaningful features. Finding the right balance between model complexity and generalization performance is challenging and often requires empirical experimentation and validation.

Spatial and Temporal Context

Satellite images often contain spatial and temporal dependencies that span multiple scales and resolutions. Designing an autoencoder architecture that effectively captures these contextual dependencies while mitigating noise is challenging. Determining the appropriate receptive field size (size of filters) and spatial pooling operations (e.g., max pooling or average pooling) to preserve spatial and temporal context is crucial for effective denoising.

Feature Representation

The number of filters and their sizes determine the richness and granularity of the features learned by the autoencoder. Choosing an architecture that can extract informative features from noisy input data while suppressing noise is challenging. Experimenting with different filter sizes, activation functions, and network depths can help identify architectures that yield the most discriminative and noise-resistant feature representations.

Loss Function and Optimizer

Deciding on the appropriate loss function and optimizer for a denoising autoencoder model involves several challenges due to the complex interplay between model architecture, training dynamics, and the characteristics of the denoising task,

Choosing a loss function that balances reconstruction accuracy with noise suppression is challenging. Loss functions such as mean squared error (MSE) emphasize pixel-wise fidelity between the denoised output and the clean target. MSE is used as the loss function for our experiments.

Selecting an optimizer that balances convergence speed and stability during training is challenging. Optimizers such as Adam offer fast convergence rates and adaptive learning rates, making them suitable choices for denoising autoencoder training. Adam optimizer is selected as the optimizer for our experiments.

Results

Comparison of PSNR values

Peak Signal-to-Noise Ratio (PSNR) is commonly used as a metric to evaluate the quality of denoising algorithms. It measures the quality of a denoised image by comparing it to the original noisy image. The higher the PSNR value, the better the denoising performance.

PSNR is calculated using the mean squared error (MSE) between the original and denoised images. The formula for PSNR is:

$$MSE = \frac{\sum_{M,N} [I_1(m,n) - I_2(m,n)]^2}{M * N}$$

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

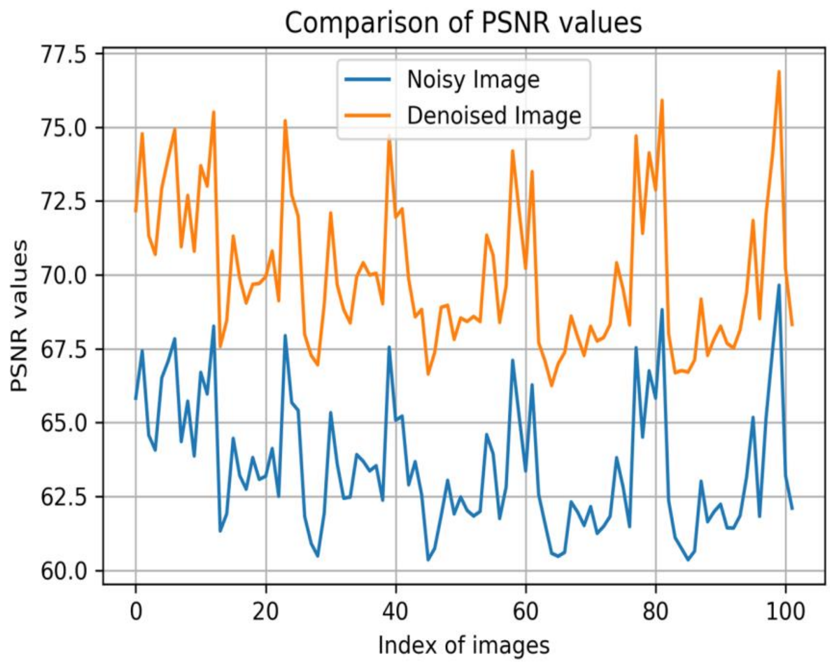
Where:

MSE is the mean squared error between the original and denoised images, calculated as the average of the squared pixel-wise differences between corresponding pixels in the two images.

Higher PSNR values indicate less distortion between the original and denoised images, implying better denoising performance.

Calculating PSNR values for all 102 test images (noisy) with clean images (denoised images) and plotting them.

Then calculating PSNR values for 102 test images denoised by Auto-encoder.



Clearly, Auto-encoder is successful in denoising the images.

Comparison with filters

Various filters and techniques have been developed to denoise SAR images while preserving important image features. Two widely used filters for SAR image denoising include:

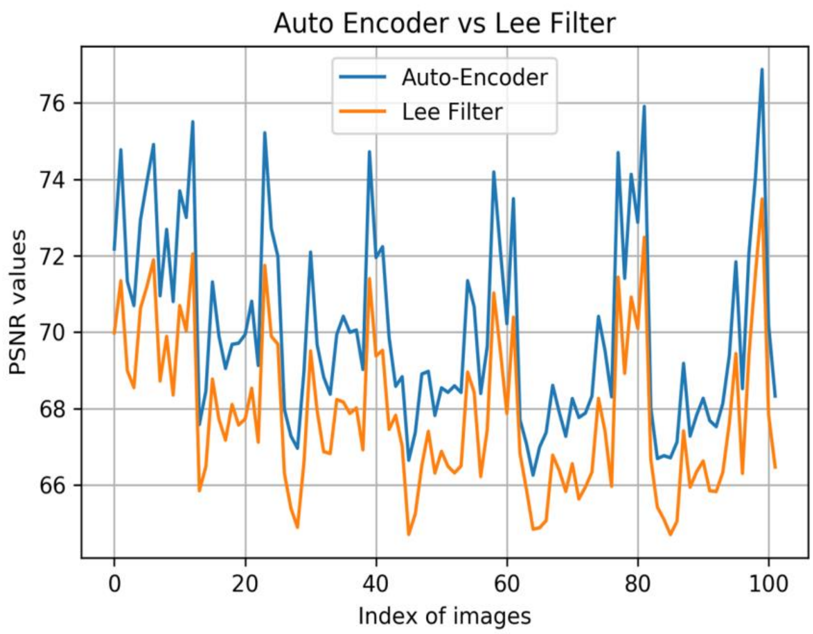
1. **Lee Filter**: The Lee filter is a simple and effective speckle reduction filter for SAR images. It operates by replacing each pixel with a weighted average of neighbouring pixels, with weights determined by the local statistics of the image.

2. **Frost Filter**: The Frost filter is another popular choice for SAR image denoising. It is based on a statistical model of speckle noise and employs an adaptive filter that estimates the noise variance from local image statistics.

Comparison with Lee Filter

For this we picked 102 test images which were noisy then

1. Denoised test images using our Auto-encoder and calculated PSNR values using 102 test images which were clean.
2. Denoised test images using Lee Filter and calculated PSNR values using 102 test images which were clean.



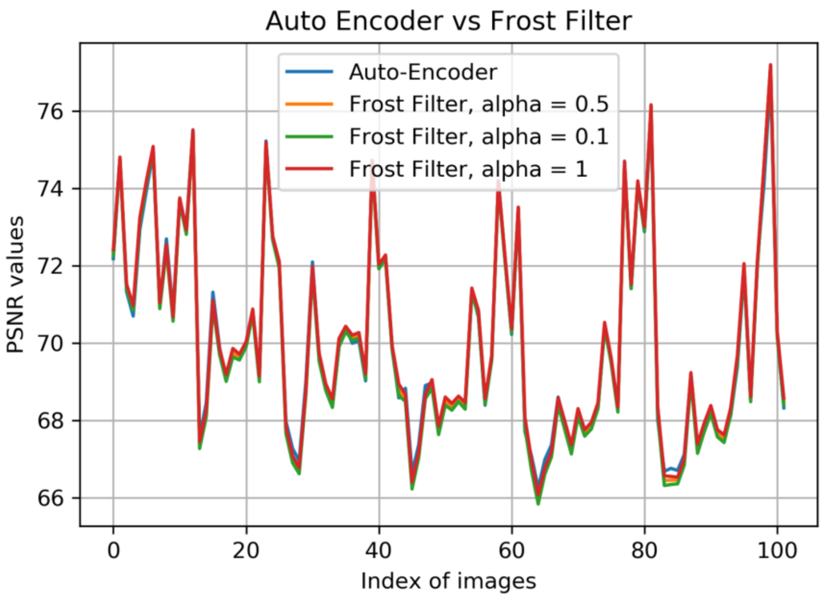
From the above plot, Auto-encoder is performing better than Lee filter for denoising of SAR images.

Comparison with Frost Filter

For this we picked 102 test images which were noisy then

1. Denoised test images using our Auto-encoder and calculated PSNR values using 102 test images which were clean.
2. Denoised test images using Frost Filter and calculated PSNR values using 102 test images which were clean with varying alpha values (0.1, 0.5 and 1.0).

Here alpha is denoising strength, higher alpha means higher denoising.



Model/Filter	**Avg. PSNR**
Auto Encoder	70.0352
Frost Filter Alpha = 0.5	70.0203
Frost Filter Alpha = 0.1	69.9348
Frost Filter Alpha = 1	70.0894

From given Avg. PSNR values Auto-encoder is performing better than Frost filter for alpha = 0.1 but close by when values of alpha are 0.5 and 1.0.

Alpha here indicates denoising strength, with higher alpha values there will be intense smoothing/blurring. This might lead to loss in features which will not be possible to recover. On doubling alpha values from 0.5 to 1.0 there is only slight improvement in denoising (approx. 0.098% increase in PSNR value) but this doubling of alpha would lead to undesirable amount of blurring.

Whereas Auto-encoder will give similar results without these issues.

Auto-encoder

Frost Filter (alpha = 1.0)

Conclusion

In this project an Auto-encoder based denoising model has been developed and then compared with widely used denoising filters (Lee filter and Frost filter). The proposed Auto encoder performs better in terms of denoising speckle noise from SAR images compare to Lee filter and Frost filter and in turns does not include excessive blurring/smoothing of the images.

Dataset was created from scratch and speckle noise was added. Peak SNR ratio taken as metric confirms above mentioned claims regarding the performance of the model.

Purpose of this model is to replace standard denoising filters used in data preprocessing step while working on SAR images.

References

[1] K. Langley and S. J. Anderson, “The Riesz transform and simultaneous representations of phase, energy and orientation in spatial vision,” Vis. Res., vol. 50, no. 17, pp. 1748–1765, Aug. 2010

[2] Y. Pu et al., “Variational autoencoder for deep learning of images, labels and captions,” in Proc. Adv. Neural Inf. Process. Syst., vol. 29, 2016, pp. 1–15

[3] S. Li, W. Dai, Z. Zheng, C. Li, J. Zou and H. Xiong, "Reversible Autoencoder: A CNN-Based Nonlinear Lifting Scheme for Image Reconstruction," in *IEEE Transactions on Signal Processing*, vol. 69, pp. 3117-3131, 2021

[4] Kumar, S., Aetesam, H., Saha, A., Maji, S.K. (2022). Attention-Based Deep Autoencoder for Hyperspectral Image Denoising. In: Raman, B., Murala, S., Chowdhury, A., Dhall, A., Goyal, P. (eds) Computer Vision and Image Processing. CVIP 2021.

[5] Ghose, S., Singh, N., Singh, P.: Image denoising using deep learning: convolutional neural network. In: 2020 10th International Conference on Cloud Computing, Data Science Engineering (Confluence), pp. 511–517 (2020).

Github

https://github.com/shantanu-shriv/Satellite-Image-Denoising